Section 1.1

Linear Systems and their solutions

Objective

- What is a linear equation and a linear system?
- What is a general solution of a LE/LS?
- What is the geometrical interpretation?
- How to find a general solution of a LE?

What is a linear equation?

Discussion 1.1.1

A line in the xy-plane

is represented algebraically by
a linear equation
in the variables x and/or y

e.g. $x + y = 1$
$x = 2$
$y = -3$

General form $ax + by = c$

a, b, c represent some real numbers

a and b are not both zero

If not it will be a point

What is a linear equation?

Definition 1.1.2

A linear equation in 3 variables $ax + by + cz = d$

geometrical meaning: plane

A linear equation in 4 variables $ax + by + cz + dw = e$

geometrical meaning: none

A linear equation in n variables

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

Indexed Notation

variables: $x_1, x_2, \ldots, x_n$ also called the unknowns

constants: $a_1, a_2, \ldots, a_n$ and $\boxed{b}$ constant term

Coeffiencts

What is a linear equation?

Example 1.1.3.1

The following are (specific) linear equations:

a) $x + 3y = 7$

b) $x_1 + 2x_2 + 2x_3 + x_4 = x_5$

c) $y = x - 0.5z + 4.5$

d) $x_1 + x_2 + \cdots + x_n = 1$

$$\boxed{a_1x_1 + a_2x_2 + \cdots + a_nx_n = b}$$

standard form

Example 1.1.3.2

The following are not linear equations:

cross term (multiplied together)
hyperbola

a) $\boxed{xy = 2}$

linear in x and y if converted

b) $\boxed{\sin(\theta) + \cos(\varphi) = 0.2}$ not linear in θ and φ

c) $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$ square terms

d) $x = \boxed{e^y}$ function of y

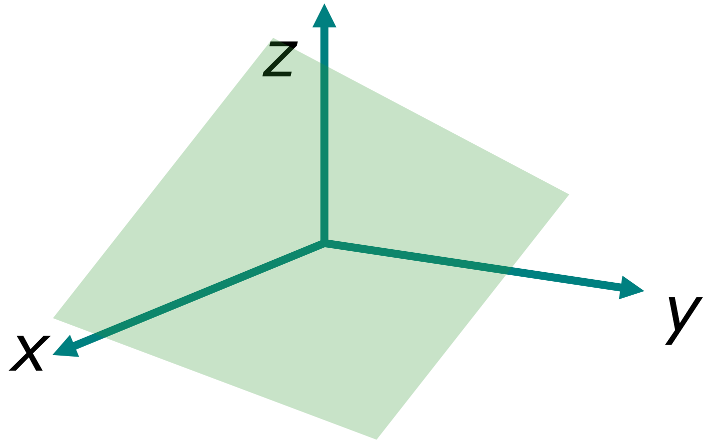
$$\boxed{a_1x_1 + a_2x_2 + \cdots + a_nx_n = b}$$

What is a linear equation?

Example 1.1.3.3

$ax + by + cz = d$ not all a, b, c are zero

represents a plane in the three dimensional space



If a, b, c all non-zero, the plane is "slanting" not parallel to any of the 3 coordinate plane

If some of a, b, c is zero, the plane is parallel to some axis

If d = 0, the plane passes through origin

What is a general solution of a LE?

Definition 1.1.4

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

variables: $x_1, x_2, \ldots, x_n$

constants: $a_1, a_2, \ldots, a_n, b$

real numbers $s_1, s_2, \ldots, s_n$

If the equation is satisfied,

$$\boxed{x_1 = s_1,\ x_2 = s_2,\ \ldots,\ x_n = s_n}$$

a solution of the linear equation

A linear equation has (infinitely) many solutions unless n = 1

The set of all solutions: solution set

An expression that represents all solutions: general solution

algebric expression to contain all the solutions

How to find a general solution of a LE?

Example 1.1.5.1

$$4x - 2y = 1$$

some solutions $\begin{cases} x = 1 \\ y = 1.5 \end{cases}$ $\begin{cases} x = 1.5 \\ y = 2.5 \end{cases}$ $\begin{cases} x = -1 \\ y = -2.5 \end{cases}$

infinitely many solutions

- pick a random value for x
- substitute this value into the equation
- solve the value of y

general solution

- set x = t (parameter)
- substitute t for x in the equation
- express y in terms of t

represents all the different kinds of answers

$$\boxed{\begin{cases} x = t \\ y = 2t - \frac{1}{2} \end{cases}}$$

How to find a general solution of a LE?

Example 1.1.5.1

$4x - 2y = 1$

some solutions
$\begin{cases} x = 1 \\ y = 1.5 \end{cases}$ $\begin{cases} x = 1.5 \\ y = 2.5 \end{cases}$ $\begin{cases} x = -1 \\ y = -2.5 \end{cases}$

general solution not a fixed answer - as long as can describe possible solutions

- set x = t (parameter)
- substitute t for x in the equation
- express y in terms of t

$$\begin{cases} x = t \\ y = 2t - \frac{1}{2} \end{cases}$$

general solution (alternative)

- set y = s (parameter)
- substitute s for y in the equation
- express x in terms of s

$$\begin{cases} x = \frac{1}{2}s + \frac{1}{4} \\ y = s \end{cases}$$

How to find a general solution of a LE?

Example 1.1.5.2

Finding the general solution for 3 variables

$$x_1 - 4x_2 + 7x_3 = 5$$

general solution

different names

- set $x_2 = s$ and $x_3 = t$ (parameters)
- substitute s for x_2 and t for x_3 in the equation
- express x_1 in terms of s and t

$$\begin{cases} x_1 = 5 + 4s - 7t \\ x_2 = s \\ x_3 = t \end{cases}$$

n variables means n-1 parameters

Geometrical interpretation

Example 1.1.5.3(a)

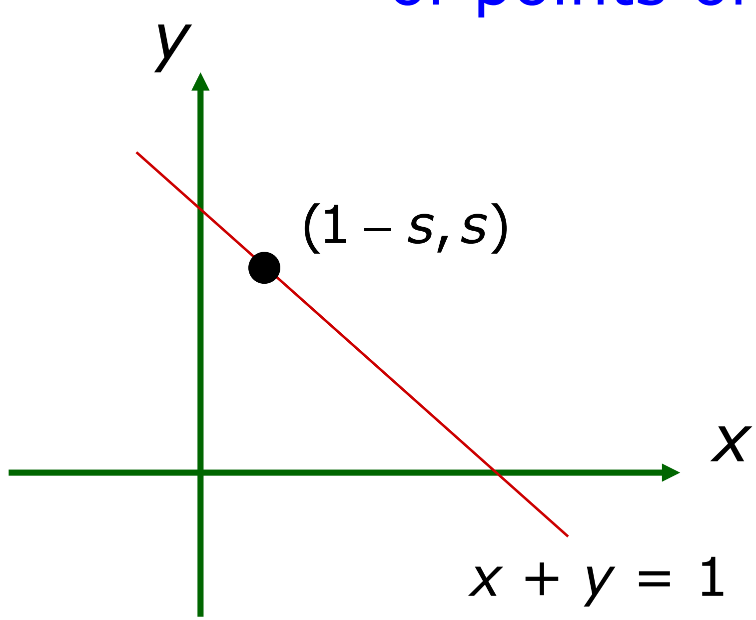
equation
$x + y = 1$

general solutions $\begin{cases} x = 1 - s \\ y = s \end{cases}$

Rewrite: $(x, y) = (1 - s, s)$

represents
a line in *xy*-plane

represents coordinates
of points on the line



Geometrical interpretation

Example 1.1.5.3(b)

equation

$x + y = 1$

regarded as

$x + y + 0z = 1$

represents

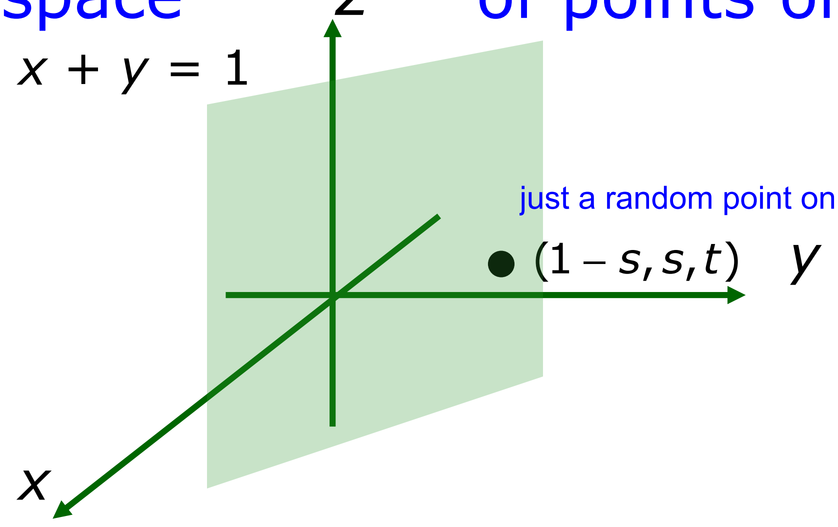
a plane in 3D space

general solutions $\begin{cases} x = 1 - s \\ y = s \\ z = t \end{cases}$

Rewrite: $(x, y, z) = (1 - s, s, t)$

represents coordinates

of points on the plane



What is a linear system?

Definition 1.1.6

Zero System
VS
Non-zero system

A system of linear equations (or a linear system)

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

Linear systems = Putting a few linear equations together

but must all have the same variables

m linear equations

variables

n $\boxed{\text{variables } x_1, x_2, \dots, x_n}$

$\boxed{a_{11}, a_{12}, \dots, a_{\boxed{mn}}}$ and $\boxed{b_1, b_2, \dots, b_m}$ are real constants

coefficients

double indices
m = tells us which equation
n = tells us which variable

constants

What is a general solution of a LS?

Definition 1.1.6

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

real numbers $s_1, s_2, \dots, s_n$

If all the equations are satisfied,

$$\boxed{x_1 = s_1,\ x_2 = s_2,\ \dots,\ x_n = s_n}$$

a solution of the linear system

solution set and general solution of the system are defined similarly as before.

Solution of a LS

Example 1.1.7

$$\begin{cases} 4x_1 - x_2 + 3x_3 = -1 \\ 3x_1 + x_2 + 9x_3 = -4 \end{cases}$$

$x_1 = 1,\ x_2 = 2,\ x_3 = -1$ is a solution

$x_1 = 1,\ x_2 = 8,\ x_3 = 1$ is not a solution

does not sastisfy BOTH equations

Remark 1.1.8

Not all systems of linear equations have solutions.

$$\begin{cases} x + y = 4 \\ 2x + 2y = 6 \end{cases}$$

This system has no solution both equtaions contridict with each other

every equations puts in a new contraints on the variable

more equations will reduce the number of solutions

What is a consistent/inconsistent LS?

Definition 1.1.9

A system of linear equations

no solution

at least one solution

inconsistent system

$$\begin{cases} x + y = 4 \\ 2x + 2y = 6 \end{cases}$$

consistent system

$$\begin{cases} 4x_1 - x_2 + 3x_3 = -1 \\ 3x_1 + x_2 + 9x_3 = -4 \end{cases}$$

Remark 1.1.10

IMPORTANT FEATURE OF LINEAR SYSTEM

Every system of linear equations has either

- no solution
- exactly one solution or
- infinitely many solutions

Chapter 2 will explain

Geometrical interpretation

Discussion 1.1.11.1

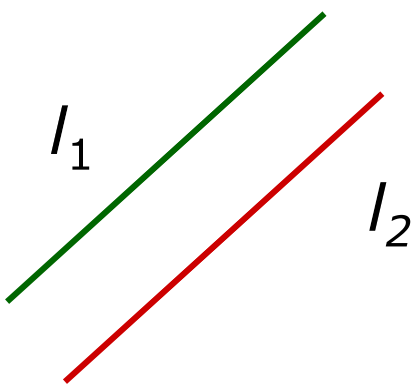
In the *xy*-plane, the system

$$\begin{cases} a_1x + b_1y = c_1 & (l_1) \\ a_2x + b_2y = c_2 & (l_2) \end{cases}$$

represent two straight lines.

a) l_1 and l_2 are parallel lines

The system has no solution



Discussion 1.1.11.1

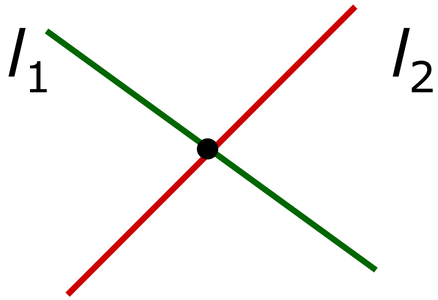
In the xy-plane, the system

$$\begin{cases} a_1x + b_1y = c_1 & (l_1) \\ a_2x + b_2y = c_2 & (l_2) \end{cases}$$

represent two straight lines.

b) l_1 and l_2 are not parallel lines.

The system has exactly one solution



Discussion 1.1.11.1

In the xy-plane, the system

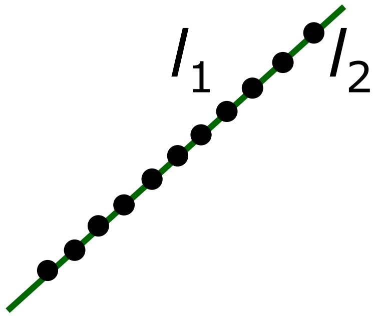
$$\begin{cases} a_1x + b_1y = c_1 & (l_1) \\ a_2x + b_2y = c_2 & (l_2) \end{cases}$$

represent two straight lines.

c) l_1 and l_2 are the same lines.

The system has infinitely many solutions

all the points are intersections



Geometrical interpretation

Discussion 1.1.11.2

In the xyz-space, the system

2 equations
3 variables

$$\begin{cases} a_1x + b_1y + c_1z = d_1 & (p_1) \\ a_2x + b_2y + c_2z = d_2 & (p_2) \end{cases}$$

represents two planes.

The system has either
no solution or infinitely many solutions.

parallel planes

the plane will intersect along a line

Section 1.2

Elementary Row Operations

Objective

- What are the three elementary row operations?
- How to perform ERO on an augmented matrix?
- What is meant by row equivalence between two augmented matrices?

What is an augmented matrix of a LS?

Definition 1.2.1

linear system

m equations
n variables

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

augmented matrix

dont show the variables

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

rectangular array

m rows
n+1 columns

n columns of variables
+1 for the constant

What is an augmented matrix of a LS?

Example 1.2.2

Consider the system of linear equations:

$$\begin{cases} x_1 + x_2 + 2x_3 = 9 \\ 2x_1 + 4x_2 - 3x_3 = 1 \\ 3x_1 + 6x_2 - 5x_3 = 0 \end{cases}$$

The augmented matrix of the system:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right)$$

take note of the sign

What are the three elementary row operations?

Definition 1.2.4

augmented matrix $\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$

Consider the following three operations on the augmented matrix:

1. Multiply a row by a nonzero constant.
2. Interchange two rows.
3. Add a multiple of one row to another row.

These are called elementary row operations.

How to perform elementary row operations?

Definition 1.2.4

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right) \xrightarrow{\text{Multiply first row by 3}} \left(\begin{array}{ccc|c} 3 & 3 & 6 & 27 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right)$$

Add 2 times of first row to second row

step 1: 1 1 2 | 9 => 2 2 4 | 18
step 2: add to the second row

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 4 & 6 & 1 & 19 \\ 3 & 6 & -5 & 0 \end{array}\right)$$

only second row is changed

Interchange second and third rows

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 3 & 6 & -5 & 0 \\ 2 & 4 & -3 & 1 \end{array}\right)$$

Why perform ERO ?

Discussion 1.2.3

Objective is to make equations simpler to solve

Elementary row operations

1. Multiply a row by a nonzero constant.
2. Interchange two rows.
3. Add a multiple of one row to another row.

These are the basic steps for solving linear system.

Correspond to the following action on the system

1. Multiply an equation by a nonzero constant.
2. Interchange two equations.
3. Add a multiple of one equation to another equation.

similar to normal steps in other maths

Why perform ERO ?

Example 1.2.5

example of: adding a multiple of 1 row to another row

x-2 - 2 - 2 - 6 = 0

$$\begin{cases} x + y + 3z = 0 & (1) \\ 2x - 2y + 2z = 4 & (2) \\ 3x + 9y = 3 & (3) \end{cases} \qquad \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 2 & -2 & 2 & 4 \\ 3 & 9 & 0 & 3 \end{array}\right)$$

Add -2 times of Equation (1) to Equation (2) to obtain Equation (4).

- 2 - 2 - 6 = 0

$$\begin{cases} x + y + 3z = 0 & (1) \\ -4y - 4z = 4 & (4) \\ 3x + 9y = 3 & (3) \end{cases} \qquad \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 3 & 9 & 0 & 3 \end{array}\right)$$

This is equivalent to adding -2 times of the first row of the matrix to the second row.

Why perform ERO ?

Example 1.2.5

x - 3 - 3 - 3 - 9 = 0

$$\begin{cases} x + y + 3z = 0 & (1) \\ \quad -4y - 4z = 4 & (4) \\ 3x + 9y = 3 & (3) \end{cases} \qquad \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 3 & 9 & 0 & 3 \end{array}\right)$$

Add -3 times of Equation (1) to Equation (3) to obtain Equation (5).

$$\begin{cases} x + y + 3z = 0 & (1) \\ \quad -4y - 4z = 4 & (4) \\ \quad 6y - 9z = 3 & (5) \end{cases} \qquad \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 0 & 6 & -9 & 3 \end{array}\right)$$

This is equivalent to adding -3 times of the first row of the matrix to the third row.

Why perform ERO ?

Example 1.2.5

$$\begin{cases} x + y + 3z = 0 & (1) \\ \quad - 4y - 4z = 4 & (4) \\ \quad 6y - 9z = 3 & (5) \end{cases} \qquad \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 0 & 6 & -9 & 3 \end{array}\right)$$

Add 6/4 times of Equation (4) to Equation (5) to obtain Equation (6).

$$\begin{cases} x + y + 3z = 0 & (1) \\ \quad - 4y - 4z = 4 & (4) \\ \quad - 15z = 9 & (6) \end{cases} \qquad \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 0 & 0 & -15 & 9 \end{array}\right)$$

This is equivalent to adding 6/4 times of the second row of the matrix to the third row.

Why perform ERO ?

Example 1.2.5

$$\begin{cases} x + y + 3z = 0 & (1) \\ \quad - 4y - 4z = 4 & (4) \\ \qquad\quad - 15z = 9 & (6) \end{cases}$$

allows for backwards substitution

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 0 & 0 & -15 & 9 \end{array}\right)$$

row echelon form

By Equation (6), $z = -3/5$.

Substituting $z = -3/5$ into Equation (4),

$-4y - 4(-3/5) = 4 \quad \Leftrightarrow \quad y = -2/5$.

Substituting $y = -2/5$ and $z = -3/5$ into Equation (1)

$x + (-2/5) + 3(-3/5) = 0 \quad \Leftrightarrow \quad x = 11/5$.

solution for the linear system

Section 1.2

Elementary Row Operations

Objective

- What is meant by row equivalence between two augmented matrices?

What is row equivalence ?

$\boxed{\text{Definition}}$ 1.2.6

Two augmented matrices are row $\boxed{\text{equivalent}}$ (to each other)

something is equal to something

if one can be obtained from the other by a $\boxed{\text{series}}$ of elementary row operations.

doesnt matter how many operations

In example 1.2.5,

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 2 & -2 & 2 & 4 \\ 3 & 9 & 0 & 3 \end{array}\right) \longleftrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 3 & 9 & 0 & 3 \end{array}\right) \longleftrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 0 & 6 & -9 & 3 \end{array}\right) \longleftrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 0 & 0 & -15 & 9 \end{array}\right)$$

Any 2 of the augmented matrices are row equivalent

1st and 4th are row equivalent

What can we say about 2 row equivalent LS ?

Theorem 1.2.7

If augmented matrices of two linear systems are row equivalent,
then the two systems have the same set of solutions.

Linear System ***X*** $\xleftrightarrow[\text{solutions}]{\text{same}}$ Linear System ***Y***

therefore

Augmented Matrix of System ***X*** $\xleftrightarrow[\text{equivalent}]{\text{row}}$ Augmented Matrix of System ***Y***

What can we say about 2 row equivalent LS ?

Example 1.2.8

All augmented matrices in Example 1.2.5 are row equivalent.

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 2 & -2 & 2 & 4 \\ 3 & 9 & 0 & 3 \end{array}\right) \quad \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 3 & 9 & 0 & 3 \end{array}\right) \quad \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 0 & 6 & -9 & 3 \end{array}\right) \quad \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & -4 & -4 & 4 \\ 0 & 0 & -15 & 9 \end{array}\right)$$

So all systems of linear equations in Example 1.2.5 have the same solution.

1st
$$\begin{cases} x + y + 3z = 0 & (1) \\ 2x - 2y + 2z = 4 & (2) \\ 3x + 9y = 3 & (3) \end{cases}$$

2nd
$$\begin{cases} x + y + 3z = 0 & (1) \\ -4y - 4z = 4 & (4) \\ 3x + 9y = 3 & (3) \end{cases}$$

3rd
$$\begin{cases} x + y + 3z = 0 & (1) \\ -4y - 4z = 4 & (4) \\ 6y - 9z = 3 & (5) \end{cases}$$

4th
$$\begin{cases} x + y + 3z = 0 & (1) \\ -4y - 4z = 4 & (4) \\ -15z = 9 & (6) \end{cases}$$

Remark 1.2.9

means there is a proof

To see why Theorem 1.2.7 is true,
we only need to check that
every elementary row operation applied to an augmented matrix
will not change the solution set of the corresponding linear system.

$x+y=1 > 2x+2y=2$
relationship of x&y are still similar

1. Multiply a row by a nonzero constant
2. Interchange two rows

   interchanging the position of 2 equation
   will not change solution

3. Add a multiple of one row to another

1: multiplying equation by constant = no change
2: adding together = no change

Section 1.3

Row-Echelon Forms

Objective

- How to identify a row-echelon form (REF) and a reduced row-echelon form (RREF)?
- How to use REF / RREF to get solutions of linear system?
- How to tell the number of solutions from REF?

How to identify REF?

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \times \qquad \begin{pmatrix} 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \checkmark$$

Definition 1.3.1

An augmented matrix is said to be in row-echelon form always be in the augmented matrix
if it has the following 2 properties:

1. If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.

just move all the 0 rows below

$$\left(\begin{array}{cccc|c} * & * & \cdots & \cdots & * \\ \vdots & \vdots & & & \vdots \\ * & * & \cdots & \cdots & * \\ \hline 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \end{array}\right) \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{nonzero rows} \\ \left.\begin{array}{c} \\ \\ \end{array}\right\} \text{zero rows (if any)} \end{array}$$

How to identify REF?

$$\begin{pmatrix} 0 & 0 & 0 & \underline{1} & 3 \\ 0 & \underline{1} & \underline{2} & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \times \qquad \begin{pmatrix} 0 & \underline{1} & \underline{2} & 0 & 1 \\ 0 & 0 & 0 & \underline{1} & \underline{3} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \checkmark$$

Definition 1.3.1

2. In any two successive $\boxed{\text{non-zero}}$ rows, the first nonzero number in the lower row occurs farther to the right than the first nonzero number in the higher row.

essentially is to form the staircase

$$\left.\left(\begin{array}{cccccc|c} \underline{0} & \underline{0} & \otimes & * & \cdots & \cdots & * \\ \underline{0} & \underline{\cdots} & \underline{\cdots} & \otimes & * & \cdots & \cdots \end{array}\right)\right\} \text{two successive rows}$$

$\boxed{\text{leading entries}}$ leading entries no need be same

numbers after leading entry are not essential

How to identify REF?

Definition 1.3.1

Combining properties 1 and 2:

width of stairs does not matter

BUT
height of stairs matter

$$\left(\begin{array}{ccccccc|c} 0 & \otimes & \cdots & \cdots & * & \cdots & * & * \\ 0 & \cdots & 0 & \otimes & * & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & \ddots & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & 0 & \otimes & * & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{array}\right)$$

nonzero rows

zero rows (if any)

everything below the staircase is 0 row

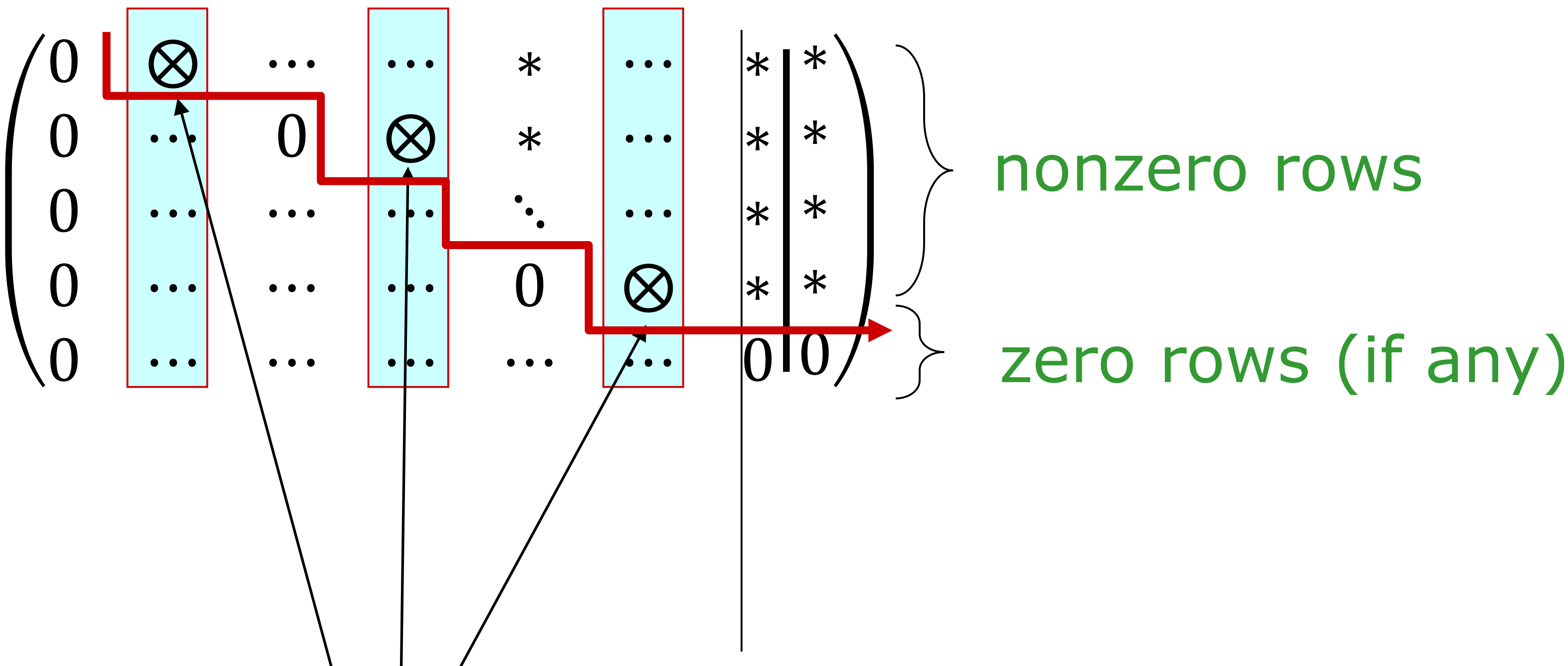
leading entries

This is a row-echelon form (REF)

How to identify REF?

Definition 1.3.1

columns that contain pivot points called pivot columns



leading entries also called pivot point

$$\begin{pmatrix} 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \times \qquad \begin{pmatrix} 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \checkmark$$

How to identify RREF?

Definition 1.3.1

An augmented matrix is said to be in **reduced row-echelon form** (RREF) special type of REF so must already have REF properties

if it is in row-echelon form and has the following properties:

3. The leading entry of every nonzero row is 1.

$$\left(\begin{array}{ccccccc|c} 0 & 1 & \cdots & \cdots & * & \cdots & * & * \\ 0 & \cdots & 0 & 1 & * & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & \ddots & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & * & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{array}\right)$$

How to identify RREF?

$$\begin{pmatrix} 0 & 1 & \boxed{2} & 0 & 0 \\ 0 & 0 & 1 & \boxed{3} & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \times \qquad \begin{pmatrix} 0 & 1 & \boxed{0} & \boxed{0} & 0 \\ 0 & 0 & 1 & \boxed{0} & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \checkmark$$

Definition 1.3.1

4. In each pivot column, except the pivot point, all other entries are zeros.

main focus is above the pivot point/leading entry must be 0

non pivot columns no restrictions

$$\left(\begin{array}{cccccccc|c} 0 & 1 & \cdots & 0 & * & 0 & \cdot & * & * \\ 0 & 0 & 0 & 1 & * & \vdots & \cdot & * & * \\ 0 & \vdots & \cdots & 0 & \ddots & 0 & \cdot & * & * \\ 0 & \vdots & \cdots & \vdots & 0 & 1 & & * & * \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdot & 0 & 0 \end{array}\right)$$

pivotal columns

How to identify (R)REF?

Remark 1.3.2

In this module

Properties 1 + 2: REF

Properties 1 + 2 + 3 + 4: RREF

In some textbooks

Properties 1 + 2 + 3: REF

Properties 1 + 2 + 3 + 4: RREF

Discussion 1.3.4

If the augmented matrix of a linear system is in REF or RREF,

we can get the solutions of the system easily.

$$\left(\begin{array}{ccccccc|c} 0 & \otimes & \cdots & \cdots & * & \cdots & * & * \\ 0 & \cdots & 0 & \otimes & * & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & \ddots & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & 0 & \otimes & * & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{ccccccc|c} 0 & 1 & \cdots & 0 & * & 0 & * & * \\ 0 & \cdots & 0 & 1 & * & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & \ddots & 0 & * & * \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & * & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{array}\right)$$

REF RREF

How to use REF / RREF to get solutions?

Example 1.3.5.1

$$\begin{array}{cccc} x1 & x2 & x3 & \\ \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right) \end{array}$$

= only 1 solution

RREF

$$\begin{cases} x_1 & & & = 1 \\ & x_2 & & = 2 \\ & & x_3 & = 3 \end{cases}$$

index notation

The system has only one solution:

$$x_1 = 1,\ x_2 = 2,\ x_3 = 3.$$

standard form

∴ RREF almost will providce immediate answers

How to use REF / RREF to get solutions?

Example 1.3.5.2

more variables than equations

means need to intro parameters

$$\left(\begin{array}{ccccc|c} 0 & 2 & 2 & 1 & -2 & 2 \\ 0 & 0 & 1 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 2 & 4 \end{array}\right)$$

REF

$$\begin{cases} 2x_2 + 2x_3 + x_4 - 2x_5 = 2 \\ x_3 + x_4 + x_5 = 3 \\ 2x_5 = 4 \end{cases}$$

introduce parameters
no. of variables - no. of equations

non-pivot columns

use variables that correspond to non-pivot column as parameters

x_1 = free parameter s

x_4 = free parameter t

$x_5 = 2$

$x_3 = 1 - t$

$x_2 = 2 + (1/2)t$

The general solution is

$$\begin{cases} x_1 = s \\ x_2 = 2 + \frac{1}{2}t \\ x_3 = 1 - t \\ x_4 = t \\ x_5 = 2 \end{cases}$$

back substitution

The system has infinitely many solutions.

Example 1.3.5.3

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 3 & -2 \\ 0 & 0 & 1 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \qquad \begin{cases} x_1 - x_2 + 3x_4 = -2 \\ x_3 + 2x_4 = 5 \\ \cancel{0 = 0} \end{cases}$$

RREF

non-pivot columns

need to intro 2 parameters

The general solution is

$$\begin{cases} x_1 = -2 + s - 3t \\ x_2 = s \\ x_3 = 5 - 2t \\ x_4 = t \end{cases}$$

The system has infinitely many solutions

Example 1.3.5.4

$$\boxed{\left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)}$$

zero system

no leading entry so everything is a non-pivot column

The general solution is

$$\begin{cases} x_1 = r \\ x_2 = s \\ x_3 = t \end{cases}$$

The system has infinitely many solutions

How to use REF / RREF to get solutions?

Example 1.3.5.5

$$\left(\begin{array}{cc|c} 3 & 1 & 4 \\ 0 & 2 & 1 \\ \boxed{0 \quad 0} & & \boxed{1} \end{array}\right)$$

$$\begin{cases} 3x_1 + x_2 = 4 \\ 2x_2 = 1 \\ \boxed{0 = 1} \end{cases}$$

usually is after performing ERO on the original equations

This system is $\boxed{\text{inconsistent}}$, i.e. no solution.

Recall:
Any linear system has
- no solution
- exactly one solution
- infinitely many solutions

Section 1.4

Gaussian Elimination

Objective

- What are Gaussian elimination and Gauss-Jordan elimination?
- How to use GE / GJE to reduce an augmented matrix to a REF / RREF ?

Row echelon form of augmented matrix

Definition 1.4.1

Carl Friedrich Gauss
(1777-1855)

Gaussian Elimination is an algorithm to reduce an augmented matrix to a row-echelon form by using elementary row operations.

$$\begin{pmatrix} \text{Augmented} \\ \text{matrix} \end{pmatrix} \xrightarrow{\text{e.r.o.}} \begin{pmatrix} \text{row - echelon} \\ \text{form} \end{pmatrix}$$

Systematic process

information for the computer to do

How to use GE to reduce a matrix to REF?

Algorithm 1.4.2

Step 1: Locate the leftmost column that does not consist entirely of zero.

Example A

$$\begin{pmatrix} \boxed{0} & 3 & \dots & \dots \\ \boxed{1} & -2 & \dots & \dots \\ \boxed{4} & 0 & \dots & \dots \end{pmatrix}$$

the first nonzero column

Example B

$$\begin{pmatrix} 0 & \boxed{3} & 1 & \dots \\ 0 & \boxed{2} & -3 & \dots \\ 0 & \boxed{-1} & 6 & \dots \end{pmatrix}$$

the first nonzero column

How to use GE to reduce a matrix to REF?

Algorithm 1.4.2

Step 2: Interchange the top row with another row, if necessary, to bring a nonzero entry to the top of the column found in Step 1.

Example A

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

Interchange the 1st row with the 2nd row.

Example B

$$\begin{pmatrix} 0 & 3 & 1 & \ldots \\ 0 & 2 & -3 & \ldots \\ 0 & -1 & 6 & \ldots \end{pmatrix}$$

No action is needed

How to use GE to reduce a matrix to REF?

Algorithm 1.4.2

Step 3: For each row below the top row, add a suitable multiple of the top row to it so that the entry below the leading entry of the top row becomes zero.

Example A

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

Example B

$$\begin{pmatrix} 0 & 3 & 1 & \dots \\ 0 & 2 & -3 & \dots \\ 0 & -1 & 6 & \dots \end{pmatrix}$$

Add – 4 times of the 1st row to the 3rd row so that the entry marked by ● becomes 0.

Add – 2/3 times of the 1st row to the 2nd row so that the entry marked by ● becomes 0.

Add 1/3 times of the 1st row to the 3rd row so that the entry marked by ● becomes 0.

How to use GE to reduce a matrix to REF?

Algorithm 1.4.2

Step 4: Now cover the top row in the matrix and begin again with Step 1 applied to the submatrix that remains.

Example A

$$\begin{pmatrix} 1 & -2 & \dots & \dots \\ 0 & 3 & \dots & \dots \\ 0 & 8 & \dots & \dots \end{pmatrix}$$

Cover the 1st row and work on the remaining rows.

Example B

$$\begin{pmatrix} 0 & 3 & 1 & \dots \\ 0 & 0 & -11/3 & \dots \\ 0 & 0 & 19/3 & \dots \end{pmatrix}$$

Cover the 1st row and work on the remaining rows.

Continue in this way until the entire matrix is in row-echelon form.

What is Gauss-Jordan elimination?

Algorithm 1.4.3

zam ERO until become RREF

Gauss-Jordan Elimination is an algorithm to reduce an augmented matrix to the reduce row-echelon form by using elementary operations.

$$\overbrace{\underbrace{\begin{pmatrix}\text{Augmented}\\ \text{matrix}\end{pmatrix} \xrightarrow{\text{e.r.o.}} \begin{pmatrix}\text{row - echelon}\\ \text{form}\end{pmatrix}}_{\text{Gaussian Elimination}} \xrightarrow{\text{e.r.o.}} \begin{pmatrix}\text{reduce row}\\ \text{echelon form}\end{pmatrix}}^{\text{Gauss-Jordan Elimination}}$$

How to use GJE to reduce a matrix to RREF?

Algorithm 1.4.3

Given an augmented matrix, use Algorithm 1.4.1 to reduce it to a row-echelon form.

Example A

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

$$\downarrow$$

$$\begin{pmatrix} 1 & -2 & \cdots & \cdots \\ 0 & 3 & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots \end{pmatrix}$$

Example B

$$\begin{pmatrix} 0 & 3 & 1 & \cdots \\ 0 & 2 & -3 & \cdots \\ 0 & -1 & 6 & \cdots \end{pmatrix}$$

$$\downarrow$$

$$\begin{pmatrix} 0 & 3 & 1 & \cdots \\ 0 & 0 & -11/3 & \cdots \\ 0 & 0 & 0 & \cdots \end{pmatrix}$$

Algorithm 1.4.3

Step 5: Multiply a suitable constant to each row so that all the leading entries becomes 1.

Example A

$$\begin{pmatrix} 1 & -2 & \cdots & \cdots \\ 0 & 3 & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots \end{pmatrix}$$

No action is needed.

Multiply the 2nd row by 1/3 so that the entry marked by ● becomes 1.

Example B

$$\begin{pmatrix} 0 & 3 & 1 & \cdots \\ 0 & 0 & -11/3 & \cdots \\ 0 & 0 & 0 & \cdots \end{pmatrix}$$

Multiply the 1st row by 1/3 so that the entry marked by ● becomes 1.

Multiply the 2nd row by – 3/11 so that the entry marked by ● becomes 1.

How to use GJE to reduce a matrix to RREF?

Algorithm 1.4.3

Step 6: Beginning with the last nonzero row and working upward, add a suitable multiples of each row to the rows above to introduce zeros above the leading entries.

Example C

$$\begin{pmatrix} 1 & -2 & \cdots & \cdots & -4 & \cdots \\ & 1 & \cdots & \cdots & 3 & \cdots \\ & & \ddots & \cdots & \cdots & \cdots \\ & & & 1 & 2 & \cdots \\ 0 & 0 & \cdots & 0 & 1 & \cdots \end{pmatrix}$$

do from bottom to top

Add 4 times of the last row to the 1st row so the entry marked by becomes 0.

Add – 3 times of the last row to the 2nd row so the entry marked by becomes 0.

Add – 2 times of the last row to the next row so the entry marked by becomes 0.

-2R4 + R3 > R3
since everything before the leading entry = 0 > only affects the value above

How to use GJE to reduce a matrix to RREF?

Algorithm 1.4.3

Step 6: Beginning with the last nonzero row and working upward, add a suitable multiples of each row to the rows above to introduce zeros above the leading entries.

Example C

$$\begin{pmatrix} 1 & -2 & \cdots & \cdots & 0 & \cdots \\ & 1 & \cdots & \cdots & 0 & \cdots \\ & & \ddots & \cdots & \cdots & \cdots \\ & & & 1 & 0 & \cdots \\ & & & & 1 & \cdots \end{pmatrix}$$

Apply the same process to the next pivot column on the left

How to use GJE to reduce a matrix to RREF?

Example 1.4.4

$$\left(\begin{array}{ccccc|c} 0 & 0 & 2 & 4 & 2 & 8 \\ 1 & 2 & 4 & 5 & 3 & -9 \\ -2 & -4 & -5 & -4 & 3 & 6 \end{array}\right)$$

Gaussian Elimination

Gauss-Jordan Elimination

$$\left(\begin{array}{ccccc|c} 1 & 2 & 4 & 5 & 3 & -9 \\ 0 & 0 & 2 & 4 & 2 & 8 \\ 0 & 0 & 0 & 0 & 6 & -24 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc|c} 1 & 2 & 0 & -3 & 0 & -29 \\ 0 & 0 & 1 & 2 & 0 & 8 \\ 0 & 0 & 0 & 0 & 1 & -4 \end{array}\right)$$

Do we have to strictly follow the steps ?

Remark 1.4.5.2

In the actual implementation of the algorithms, the steps mentioned in Algorithm 1.4.2 and Algorithm 1.4.3 are usually modified to avoid the round-off errors during the computation

many entries in a matrix
small change in entry can impact solution very big
from solution to no solution

$\boxed{\text{Ill-conditioned matrix}}$

very sensitive matrix

See Exercise 1 Q21

Do we have to strictly follow the steps ?

Additional remarks

Modification in GE

Example

Standard

$$\begin{pmatrix} 4 & 3 & \dots & \dots \\ 1 & -2 & \dots & \dots \\ 0 & 0 & \dots & \dots \end{pmatrix} \longrightarrow \begin{pmatrix} 4 & 3 & \dots & \dots \\ 0 & -11/4 & \dots & \dots \\ 0 & 0 & \dots & \dots \end{pmatrix}$$

Variation

$$\begin{pmatrix} 1 & -2 & \dots & \dots \\ 4 & 3 & \dots & \dots \\ 0 & 0 & \dots & \dots \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -2 & \dots & \dots \\ 0 & 11 & \dots & \dots \\ 0 & 0 & \dots & \dots \end{pmatrix}$$

swap first even though gaussian no need to follow strictly

make life easy

Is the REF/RREF of a matrix unique ?

Remark 1.4.5.1

why RREF over REF

Every matrix can have many different row-echelon forms.

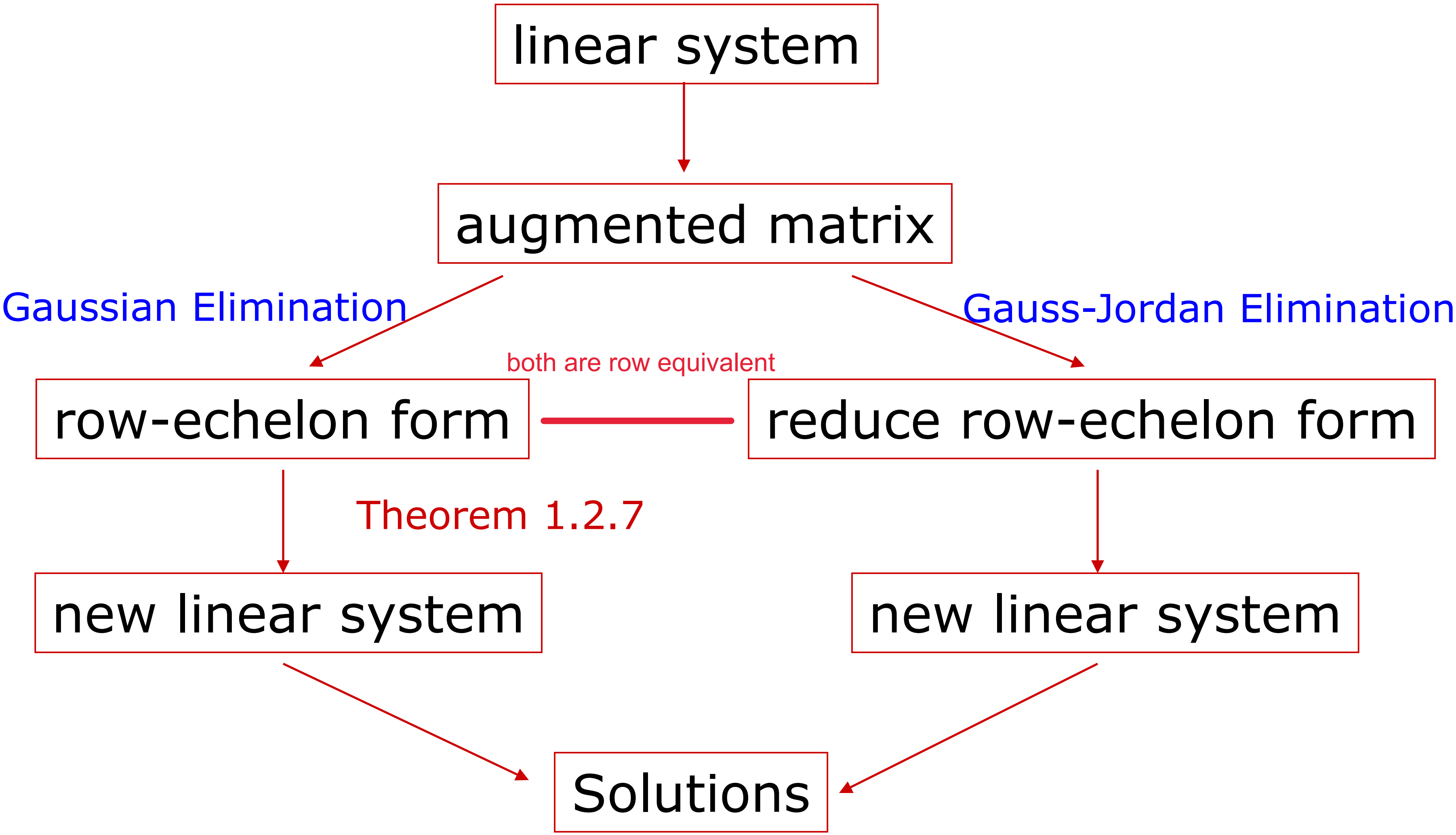
to check whether 2 augmented matrices are row equivalent, check RREF can help to check answers with other people

$$\boxed{\text{Every matrix has a unique reduced row-echelon form}}$$

$$\begin{tikzcd}
 & \left(\begin{array}{c}\text{row - echelon}\\ \text{form}\end{array}\right) \arrow[dr] & \\
\left(\begin{array}{c}\text{Augmented}\\ \text{matrix}\end{array}\right) \arrow[ur] \arrow[r, "\text{e.r.o.}"] \arrow[dr] & \left(\begin{array}{c}\text{row - echelon}\\ \text{form}\end{array}\right) \arrow[r, "\text{e.r.o.}"] & \left(\begin{array}{c}\text{reduce row}\\ \text{echelon form}\end{array}\right) \\
 & \left(\begin{array}{c}\text{row - echelon}\\ \text{form}\end{array}\right) \arrow[ur] &
\end{tikzcd}$$

WHY?

Discussion 1.4.6



How to tell the number of solutions from REF?

Remark 1.4.8.1

inconsistent system

A linear system has no solution if:

REF has a row with nonzero last entry but zero elsewhere.

The last column of REF is a pivot column.

If constant column is a pivot column,
then system will have no solution

$$\left(\begin{array}{ccccccc|c} 0 & \otimes & \cdots & \cdots & * & \cdots & * & * \\ 0 & \cdots & 0 & \otimes & * & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & \ddots & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & 0 & \otimes & * & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \otimes \end{array}\right) \text{ e.g. } \left(\begin{array}{ccc|c} 3 & 2 & 3 & 4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right)$$

How to tell the number of solutions from REF?

Remark 1.4.8.2

A linear system has exactly one solution if:

every column of REF is a pivot column, except the last column.

because no need to introduce parameter

no. of variables = no. of equations

$$\left(\begin{array}{cccc|c} \otimes & \cdots & * & \cdots & * \\ \cdots & \otimes & * & \cdots & * \\ \cdots & \cdots & \ddots & \cdots & * \\ \cdots & \cdots & 0 & \otimes & * \\ \cdots & \cdots & \cdots & \cdots & 0 \end{array}\right)$$

Remark 1.4.8.2

In other words, a consistent linear system has exactly one solution if:

# of variables in LS = # of nonzero rows in REF

e.g. $\left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & -1 & 2 \end{array}\right)$ $\left(\begin{array}{cccc|c} 1 & 1 & 2 & 3 & 4 \\ 0 & 2 & 0 & 1 & -1 \\ 0 & 0 & 4 & -1 & 2 \\ 0 & 0 & 0 & -1 & 2 \\ \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} \end{array}\right)$ 4 variables = 4 equations

How to tell the number of solutions from REF?

Remark 1.4.8.3

A linear system has infinitely many solutions if:

there is a non-pivot column in the REF, other than the last column.

at least 1 non-pivot column = intro parameters

$$\left(\begin{array}{ccccccc|c} 0 & \otimes & \cdots & \cdots & * & \cdots & * & * \\ 0 & \cdots & 0 & \otimes & * & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & \ddots & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & 0 & \otimes & * & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{array}\right)$$

How to tell the number of solutions from REF?

Remark 1.4.8.3

In other words, a consistent linear system has infinitely many solutions if:

# of variables in LS > # of nonzero rows in REF

e.g. $$\left(\begin{array}{cccc|c} 5 & 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right) \qquad \left(\begin{array}{cccc|c} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right)$$

4 variables > 3 equations

4 variables > 3 equations
with 1 hidden variable x1

Finding a REF/RREF first then find if inconsistent/consistent

Section 1.4

Gaussian Elimination

Objective

• How to use GE / GJE to solve indirect LS problems?

How to denote ERO?

Notation 1.4.9

When doing elementary row operations, we adopt the following notation:

1. cR_i

   "multiply the i^{th} row by the constant c".

2. $R_i \leftrightarrow R_j$

   "interchange the i^{th} and the j^{th} rows".

3. $R_i + cR_j$

   "add c times of the j^{th} row to the i^{th} row".

Linear system with "unknown" constant terms

Example 1.4.10.1

What is the condition that must be satisfied by a, b, c so that the system of linear equations

$$\begin{cases} x + 2y - 3z = a \\ 2x + 6y - 11z = b \\ x - 2y + 7z = c \end{cases}$$

has at least one solution?

Turn it into REF then see if can find infinitely many solutions

Example 1.4.10.1

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & a \\ 2 & 6 & -11 & b \\ 1 & -2 & 7 & c \end{array}\right) \xrightarrow{R_2 - 2R_1} \left(\begin{array}{ccc|c} 1 & 2 & -3 & a \\ 0 & 2 & -5 & b-2a \\ 1 & -2 & 7 & c \end{array}\right) \xrightarrow{R_3 - R_1}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & a \\ 0 & 2 & -5 & b-2a \\ 0 & -4 & 10 & c-a \end{array}\right) \xrightarrow{R_3 + 2R_2} \left(\begin{array}{ccc|c} 1 & 2 & -3 & a \\ 0 & 2 & -5 & b-2a \\ 0 & 0 & 0 & 2b+c-5a \end{array}\right)$$

If $2b + c - 5a \neq 0$, system has no solution

If $2b + c - 5a = 0$, system has infinitely many solns.

It has (infinitely many) solutions if and only if

$2b + c - 5a = 0$.

Linear system with "unknown" constant terms

Example 1.4.10.1

$$\begin{cases} x + 2y - 3z = a \\ 2x + 6y - 11z = b \\ x - 2y + 7z = c \end{cases}$$

How many solutions do these systems have?

$$\begin{cases} x + 2y - 3z = 1 \\ 2x + 6y - 11z = 1 \\ x - 2y + 7z = 1 \end{cases} \qquad \begin{cases} x + 2y - 3z = 1 \\ 2x + 6y - 11z = 2 \\ x - 2y + 7z = 1 \end{cases}$$

infinitely many solutions

$2b + c - 5a = -2$ $\qquad$ $2b + c - 5a = 0$

It has (infinitely many) solutions if and only if

$2b + c - 5a = 0$.

Linear system with "unknown" coefficients and constant terms

Example 1.4.10.2

$$\begin{cases} x + 2y + z = 1 \\ 2x + by + 2z = 2 \\ 4x + 8y + b^2 z = 2b \end{cases}$$

Determine the values of b so that the system of linear equations has

(a) no solution

(b) a unique solution

(c) infinitely many solutions

Linear system with "unknown" coefficients and constant terms

Example 1.4.10.2

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & b & 2 & 2 \\ 4 & 8 & b^2 & 2b \end{array}\right) \xrightarrow[R_3 - 4R_1]{R_2 - 2R_1} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & b-4 & 0 & 0 \\ 0 & 0 & b^2-4 & 2b-4 \end{array}\right)$$

Add -2 times of the first row to the second row.

Add -4 times of the first row to the third row.

Linear system with “unknown” coefficients and constant terms

Example 1.4.10.2

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & b-4 & 0 & 0 \\ 0 & 0 & b^2-4 & 2b-4 \end{array}\right)$$

(a) The system has no solution if

the last column is a pivot column

$b^2 - 4 = 0$ and $2b - 4 \neq 0$ $\rightarrow$ $b = -2$

Linear system with "unknown" coefficients and constant terms

Example 1.4.10.2

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & b-4 & 0 & 0 \\ 0 & 0 & b^2-4 & 2b-4 \end{array}\right)$$

(b) The system has a unique solution if

every column is a pivot column (except the last)

$b-4 \neq 0$ and $b^2-4 \neq 0$ $\Leftrightarrow$ $b \neq 4$, $b \neq 2$ and $b \neq -2$

Linear system with "unknown" coefficients and constant terms

Example 1.4.10.2

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & b-4 & 0 & 0 \\ 0 & 0 & b^2-4 & 2b-4 \end{array}\right)$$

(c) The system has infinitely many solutions if

some columns are non-pivot columns

(i) $b - 4 = 0 \quad \rightarrow \quad b = 4$

or

(ii) $b^2 - 4 = 0$ and $2b - 4 = 0 \rightarrow b = 2$

Linear system with "unknown" coefficients and constant terms

Example 1.4.10.2

$$\begin{cases} x + 2y + z = 1 \\ 2x + by + 2z = 2 \\ 4x + 8y + b^2 z = 2b \end{cases}$$

Determine the values of b so that the system of linear equations has

(a) no solution $b = -2$

(b) a unique solution $b \neq 4$, $b \neq 2$ and $b \neq -2$

(c) infinitely many solutions $b = 2$ or $b = 4$

Linear system with more than one "unknown" coefficients and constant terms

Example 1.4.10.3

Determine the values of a and b so that the system of linear equations

$$\begin{cases} ax + y = a \\ x + y + z = 1 \\ y + az = b \end{cases}$$

has

(a) no solution,

(b) a unique solution, and

(c) infinitely many solutions.

Linear system with more than one "unknown" coefficients and constant terms

Example 1.4.10.3

$$\left(\begin{array}{ccc|c} a & 1 & 0 & a \\ 1 & 1 & 1 & 1 \\ 0 & 1 & a & b \end{array}\right)$$

add $-1/a$ times of first row to second row

Cannot do this if $a = 0$

Need to consider two different situations:

Case 1: $a = 0$ and

Case 2: $a \neq 0$.

Case 1 $a = 0$ | Case 2 $a \neq 0$

Example 1.4.10.3

Solution Case 1: $a = 0$

Substitute $a = 0$ to the augmented matrix:

$$\left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & b \end{array}\right) \xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & b \end{array}\right) \xrightarrow{R_3 - R_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & b \end{array}\right)$$

Under the assumption $a = 0$,

- the system has no solution if $b \neq 0$;
- the system has infinitely many solutions if $b = 0$.

Case 1
$a = 0$

Case 2
$a \neq 0$

Example 1.4.10.3

Solution Case 2: $a \neq 0$

$$\left(\begin{array}{ccc|c} a & 1 & 0 & a \\ 1 & 1 & 1 & 1 \\ 0 & 1 & a & b \end{array}\right) \xrightarrow{R_2 - \frac{1}{a}R_1} \left(\begin{array}{ccc|c} a & 1 & 0 & a \\ 0 & \frac{a-1}{a} & 1 & 0 \\ 0 & 1 & a & b \end{array}\right)$$

add $-a/(a-1)$ times of second row to third row

Cannot do this if $a = 1$

Need to consider two cases again:

Case 2a: $a = 1$ and

Case 2b: $a \neq 1$.

Case 1 $a = 0$

Case 2 $a \neq 0$

Case 2a $a = 1$

Case 2b $a \neq 1$

Example 1.4.10.3

Solution Case 2a: $a = 1$

Substitute $a = 1$ to the last augmented matrix

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & b \end{array}\right) \xrightarrow{R_2 \leftrightarrow R_3} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & 0 \end{array}\right)$$

Under the assumption $a = 1$,

- the system has exactly one solution.

Case 1 $a = 0$ | Case 2 $a \neq 0$

Case 2a $a = 1$ | Case 2b $a \neq 1$

Example 1.4.10.3

Solution Case 2b: $a \neq 0$ and $a \neq 1$

$$\left(\begin{array}{ccc|c} a & 1 & 0 & a \\ 0 & \frac{a-1}{a} & 1 & 0 \\ 0 & 1 & a & b \end{array}\right) \xrightarrow{R_3 - \frac{a}{a-1}R_2} \left(\begin{array}{ccc|c} a & 1 & 0 & a \\ 0 & \frac{a-1}{a} & 1 & 0 \\ 0 & 0 & \frac{a^2-2a}{a-1} & b \end{array}\right)$$

- the system has no solution if

  $(a^2 - 2a)/(a - 1) = 0$ & $b \neq 0$ $\Leftrightarrow$ $a = 2$ & $b \neq 0$;

- the system has one solution if

  $(a^2 - 2a)/(a - 1) \neq 0$ $\Leftrightarrow$ $a \neq 2$;

- the system has infinitely many solutions if

  $(a^2 - 2a)/(a - 1) = 0$ & $b = 0$ $\Leftrightarrow$ $a = 2$ & $b = 0$.

Case 1 — $a = 0$

Case 2 — $a \neq 0$

Case 2a — $a = 1$

Case 2b — $a \neq 1$

Example 1.4.10.3

Answer (a)

The system has no solution:

by Case 1, $a = 0$ and $b \neq 0$

or

by Case 2b, $a \neq 0$ & $a \neq 1$ and $a = 2$ & $b \neq 0$

> The system has no solution if
>
> $b \neq 0$ and $a = 0$ or $a = 2$.

Case 1 $a = 0$

Case 2 $a \neq 0$

Case 2a $a = 1$

Case 2b $a \neq 1$

Example 1.4.10.3

Answer (b)

The system has a unique solution:

by Case 2a, $a = 1$;

or

by Case 2b, $a \neq 0$ & $a \neq 1$ and $a \neq 2$

The system has a unique solution if $a \neq 0$ and $a \neq 2$.

Case 1 $a = 0$ | Case 2 $a \neq 0$

Case 2a $a = 1$ | Case 2b $a \neq 1$

Example 1.4.10.3

Answer (c)

The system has infinitely many solutions:

by Case 1, $a = 0$ and $b = 0$

or

by Case 2b, $a \neq 0$ & $a \neq 1$ and $a = 2$ & $b = 0$

> The system has infinitely many solutions if
>
> $b = 0$ and $a = 0$ or 2.

Linear system with more than one "unknown" coefficients and constant terms

Remark on Example 1.4.10.3

$$\left(\begin{array}{ccc|c} a & 1 & 0 & a \\ 1 & 1 & 1 & 1 \\ 0 & 1 & a & b \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & a & b \\ a & 1 & 0 & a \end{array}\right)$$

If we rearrange the rows of the augmented matrix in the following way:

the 2nd row at the top,

the 3rd row in the middle and

the 1st row at the bottom,

the problem will be much easier to be solved by Gaussian Elimination.

Finding equation of a curve

Example 1.4.10.4

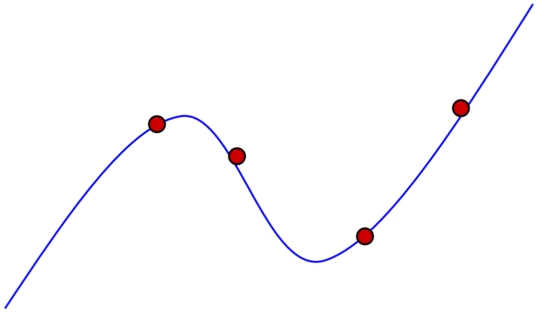
Given a cubic curve with equation

$$y = a + bx + cx^2 + dx^3,$$

where a, b, c, d are real constants, that passes through the points

$$(0, 10),\ (1, 7),\ (3, -11) \text{ and } (4, -14),$$

find the values of a, b, c, d.



4 points will determine the equation

Finding equation of a curve

Example 1.4.10.4

By substituting

$(x, y) =$ (0, 10), (1, 7), (3, –11) and (4, –14)

into the equation $y = a + bx + cx^2 + dx^3$,

we obtain a system of linear equations:

$$\begin{cases} a & & & & & & & = & 10 \\ a & + & b & + & c & + & d & = & 7 \\ a & + & 3b & + & 9c & + & 27d & = & -11 \\ a & + & 4b & + & 16c & + & 64d & = & -14 \end{cases}$$

where a, b, c, d are the variables

Note the role swap of notation

Finding equation of a curve

Example 1.4.10.4

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 10 \\ 1 & 1 & 1 & 1 & 7 \\ 1 & 3 & 9 & 27 & -11 \\ 1 & 4 & 16 & 64 & -14 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 10 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

So the solution is
$a = 10$, $b = 2$, $c = -6$ and $d = 1$.

The equation of the cubic curve is
$y = 10 + 2x - 6x^2 + x^3$.

Geometrical interpretation in 3D space

Discussion 1.4.11

LS of 3 variables
(with solutions)

REF	Solutions	Geometrical interpretation for 3 planes
3 non-zero rows	0 parameter	Intersect at 1 point
2 non-zero rows	1 parameter	Intersect at a line
1 non-zero row	2 parameters	Intersect at a plane
0 non-zero row	3 parameters	NA

Section 1.5

Homogeneous Linear Systems

Objective

- What is a homogeneous system?
- What is a trivial / non-trivial solution of a homogeneous system?

What is a homogeneous system?

Definition 1.5.1

A system of linear equations is said to be homogeneous if it has the form

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

all the constant terms are zero

If a linear system has some non-zero constant terms, we say it is non-homogeneous.

What is a trivial/non-trivial solution?

Definition 1.5.1

A system of linear equations is said to be homogeneous if it has the form

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

$x_1 = 0, \; x_2 = 0, \; \dots, \; x_n = 0$ is a solution

trivial solution

Any solution other than the trivial solution is called a non-trivial solution.

Example

Consider the following homogeneous system:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 - x_2 + x_3 - x_4 = 0 \end{cases}$$

$x_1 = 0,\ x_2 = 0, x_3 = 0,\ x_4 = 0$ trivial solution

$x_1 = 1,\ x_2 = 0, x_3 = -1,\ x_4 = 0$ non-trivial solution

Remark: Only in a homogeneous system do we talk about trivial / non-trivial solution.

Example 1.5.2

Given a quadric surface with equation

$$ax^2 + by^2 + cz^2 = d$$

where a, b, c, d are real constants,

that passes through the points

$(1, 1, -1)$, $(1, 3, 3)$ and $(-2, 0, 2)$,

find a formula for the quadric surface.

$$\begin{cases} a + b + c = d \\ a + 9b + 9c = d \\ 4a + 4c = d \end{cases} \qquad \left(\begin{array}{ccc|c} 1 & 1 & -1 & d \\ 1 & 9 & 9 & d \\ 4 & 0 & 4 & d \end{array}\right)$$

Example 1.5.2

Given a quadric surface with equation

$$ax^2 + by^2 + cz^2 = d$$

$$\begin{cases} a + b + c - d = 0 \\ a + 9b + 9c - d = 0 \\ 4a + 4c - d = 0 \end{cases}$$

homogeneous system

General solution

$$\begin{cases} a = t \\ b = \frac{3}{4}t \\ c = -\frac{3}{4}t \\ d = t \end{cases}$$

t = 0: a = 0, b = 0, c = 0, d = 0

trivial solution

t = 4: a = 4, b = 3, c = -3, d = 4

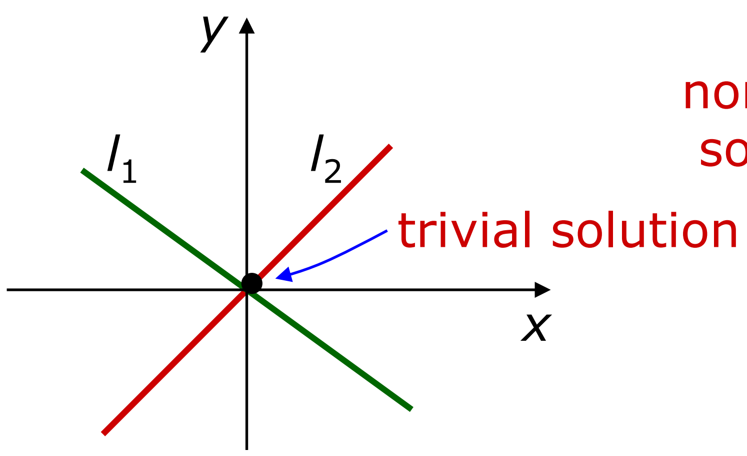
non-trivial solution

What is a trivial/non-trivial solution?

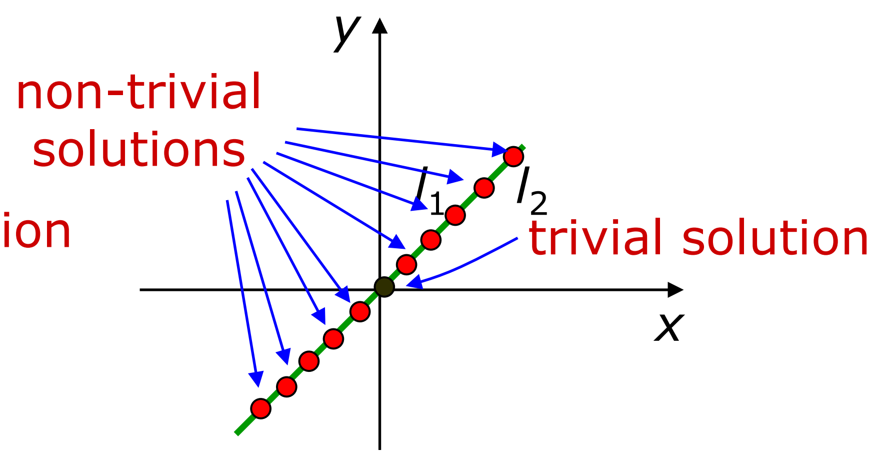
Discussion 1.5.3.1

$$\begin{cases} a_1 x + b_1 y = 0 & (l_1) \\ a_2 x + b_2 y = 0 & (l_2) \end{cases}$$

represent two straight lines through the origin.



exactly one solution



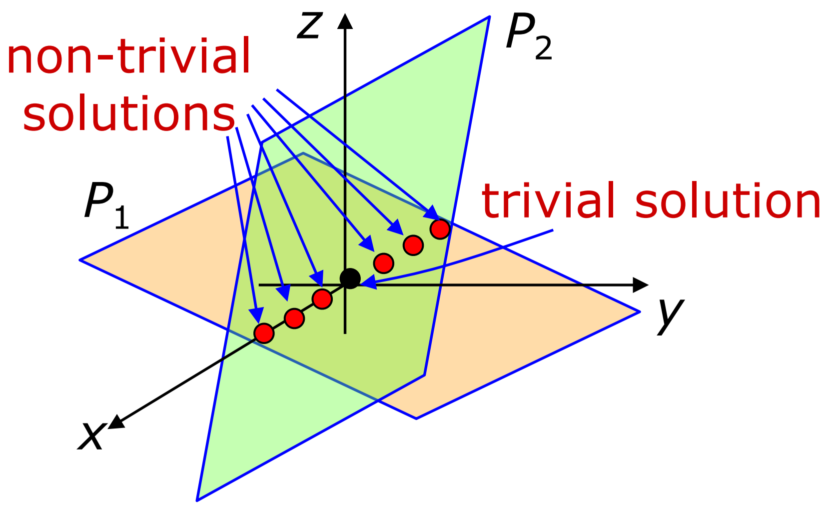
infinitely many solutions

What is a trivial/non-trivial solution?

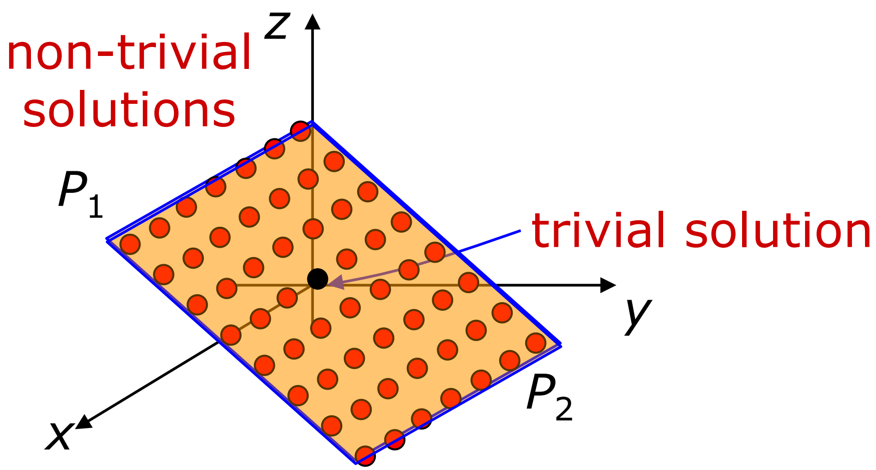
Discussion 1.5.3.2

$$\begin{cases} a_1x + b_1y + c_1z = 0 & (P_1) \\ a_2x + b_2y + c_2z = 0 & (P_2) \end{cases}$$

represent two planes through the origin.



infinitely many solutions



infinitely many solutions

How many solutions does a homogeneous solution have?

Remark 1.5.4

1. A homogeneous system of linear equations has either only the trivial solution
   or infinitely many solutions in addition to the trivial solution.
2. A homogeneous system of linear equations with more variables than equations has infinitely many solutions.

$$\begin{aligned} a_1x + b_1y + c_1z &= 0 \\ a_2x + b_2y + c_2z &= 0 \end{aligned}$$

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 &= 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 &= 0 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 &= 0 \end{aligned}$$

Section 2.1

Introduction to Matrices

Objective

- What are the size, entries, order of a matrix?
- What are diagonal, identity, symmetric, triangular matrices?
- How to express matrices using (i, j)-entries?

What are the size and entries of a matrix?

Summary 2.1.1-2.1.5

Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

row

column

can be simplified as $\boldsymbol{A} = (a_{ij})_{m \times n}$ or (a_{ij})

number of rows is m number of columns is n

We say: The size of the matrix $\boldsymbol{A}$ is m x n

$\boldsymbol{A}$ is an m x n matrix

a_{ij} denotes the number in the i^{th} row and j^{th} column.

We say: a_{ij} is the (i, j)-entry of the matrix $\boldsymbol{A}$

What are the size and entries of a matrix?

Example 2.1.6

1. $\boldsymbol{A} = (a_{ij})_{2\times3}$ where $a_{ij} = i + j$

$$\boldsymbol{A} = \begin{pmatrix} \, & \, & \, \\ \, & \, & \, \end{pmatrix}$$

2. $\boldsymbol{B} = (b_{ij})_{3\times2}$ where $b_{ij} = \begin{cases} 1 & \text{if } i+j \text{ is even} \\ -1 & \text{if } i+j \text{ is odd} \end{cases}$

$$\boldsymbol{B} = \begin{pmatrix} \, & \, \\ \, & \, \\ \, & \, \end{pmatrix}$$

Learn how to describe various types of matrices in terms of (i, j)-entries

What are the order and diagonal of a square matrix?

Summary 2.1.7-2.1.8

Square matrices

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

same number of rows and columns

$\boldsymbol{A}$ is an n x n matrix

$\boldsymbol{A} = (a_{ij})$ is a square matrix of order n

a_{11}, a_{22}, ..., a_{nn} are called the diagonal entries

a_{ij}, $i \neq j$, are called the non-diagonal entries

What are diagonal, scalar, identity matrices?

How to express them using (i, j)-entries?

Summary 2.1.7-2.1.8

Types of square matrices

Diagonal matrix	all non-diagonal entries are zero	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$	$a_{ij} = 0$ whenever $i \neq j$
Scalar matrix	diagonal matrix with all diagonal entries the same	$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$	$a_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ c & \text{if } i = j \end{cases}$
Identity matrix $\boldsymbol{I}_n$	diagonal matrix with all diagonal entries equal 1	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$a_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$

What are symmetric and triangular matrices?

How to express them using (i, j)-entries?

Summary 2.1.7-2.1.8

Types of square matrices

Zero matrix $\mathbf{0}_{m\times n}$	all entries equal to zero (can be non-square)	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$a_{ij} = 0$ for all i, j
Symmetric matrix	k^{th} row "equal" k^{th} column for all k	$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix}$	$a_{ij} = a_{ji}$ for all i, j
Upper triangular matrix	all entries below diagonals are zero	$\begin{pmatrix} 1 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 2 \end{pmatrix}$	$a_{ij} = 0$ for all $i > j$
Lower triangular matrix	all entries above diagonals are zero	$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 2 & 3 & 2 \end{pmatrix}$	$a_{ij} = 0$ for all $i < j$

Section 2.2

Matrix Operations

Objective

- How to perform matrix addition & multiplication, scalar multiplication and transpose?
- How to express these operations using (i, j)-entries?
- What are some properties of these operations?
- What are some different ways to express matrix multiplication?
- How to express LS in matrix equation form?

How to perform matrix addition, scalar multiplication?

Summary 2.2.1 - 2.2.5

Let $\boldsymbol{A} = (a_{ij})_{m \times n}$ $\boldsymbol{B} = (b_{ij})_{m \times n}$ and c a real constant.

Matrix Equality	$\boldsymbol{A} = \boldsymbol{B}$	$\boldsymbol{A}$ and $\boldsymbol{B}$ have same size and same corresponding entries	$a_{ij} = b_{ij}$ for all i, j
Matrix Addition	$\boldsymbol{A} + \boldsymbol{B}$	addition of corresponding entries of $\boldsymbol{A}$ and $\boldsymbol{B}$	$(a_{ij} + b_{ij})_{m \times n}$
Matrix subtraction	$\boldsymbol{A} - \boldsymbol{B}$	subtraction of corresponding entries of $\boldsymbol{A}$ and $\boldsymbol{B}$	$(a_{ij} - b_{ij})_{m \times n}$
Scalar multiplication	$c\boldsymbol{A}$	multiply every entry of $\boldsymbol{A}$ by scalar c	$(ca_{ij})_{m \times n}$
Negative of matrix	$-\boldsymbol{A}$	attach negative sign to every entry of $\boldsymbol{A}$	$(-a_{ij})_{m \times n}$

Summary 2.2.6 - 2.2.7

Properties

- On matrix addition and scalar multiplication
- Theorem 2.2.6
- Similar to ordinary numbers operations
- Commutative Law: $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$
- Associative Law: $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$
- Zero matrix behaves like number "0" in matrix addition

Definition 2.2.8 & Example 2.2.9.1

Matrix Multiplication

$$\underset{2 \times 3}{\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}} \underset{3 \times 2}{\begin{pmatrix} 1 & 1 \\ 2 & 3 \\ -1 & -2 \end{pmatrix}}$$

$$= \underset{2 \times 2}{\begin{pmatrix} & \\ & \end{pmatrix}}$$

$$= \begin{pmatrix} 2 & 1 \\ 8 & 7 \end{pmatrix}$$

Definition 2.2.8 (Matrix Multiplication)

Let $\boldsymbol{A} = (a_{ij})_{m\times p}$ and $\boldsymbol{B} = (b_{ij})_{p\times n}$ be two matrices.

The product $\boldsymbol{AB}$ is an $m \times n$ matrix

its (i, j)-entry is

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$$

summation notation

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{pmatrix} \qquad \boldsymbol{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{pmatrix}$$

How to perform matrix multiplication?

Remark 2.2.10.1

We can only multiply two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ (in the manner $\boldsymbol{AB}$) when the number of columns of $\boldsymbol{A}$ is equal to the number of rows of $\boldsymbol{B}$.

$\boldsymbol{A} = (a_{ij})_{m\times p}$ and $\boldsymbol{B} = (b_{ij})_{p\times n}$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 3 \\ -1 & -2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} 1 & 2 & -1 \\ 1 & 3 & -2 \end{pmatrix}$$

What are some properties of matrix multiplication?

Remark 2.2.10.2-4

Different from ordinary numbers multiplication

The matrix multiplication is not commutative.

i.e. $\boldsymbol{AB} \neq \boldsymbol{BA}$ in general, even if the product exist.

$\boldsymbol{AB}$: pre-multiplication of $\boldsymbol{A}$ to $\boldsymbol{B}$

$\boldsymbol{BA}$: post-multiplication of $\boldsymbol{A}$ to $\boldsymbol{B}$

$\boldsymbol{AB} = \mathbf{0}$ does not imply $\boldsymbol{A} = \mathbf{0}$ or $\boldsymbol{B} = \mathbf{0}$.

What are some properties of matrix multiplication?

Theorem 2.2.11.1-3

Similar to ordinary numbers multiplication

1. $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}$ Associative Law
2. $\boldsymbol{A}(\boldsymbol{B}_1 + \boldsymbol{B}_2) = \boldsymbol{AB}_1 + \boldsymbol{AB}_2$
   $(\boldsymbol{C}_1 + \boldsymbol{C}_2)\,\boldsymbol{A} = \boldsymbol{C}_1\boldsymbol{A} + \boldsymbol{C}_2\boldsymbol{A}$ Distributive Law
3. $c(\boldsymbol{AB}) = (c\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}(c\boldsymbol{B})$ c is a scalar

To prove these properties,
check LHS and RHS have same size
and same corresponding entries

What are some properties of matrix multiplication?

Theorem 2.2.11.4

Similar to ordinary numbers multiplication

Let $\boldsymbol{A}$ be a $m \times n$ matrix.

- $\boldsymbol{A}\mathbf{0}_{n \times q} = \mathbf{0}_{m \times q}$ and $\mathbf{0}_{p \times m}\boldsymbol{A} = \mathbf{0}_{p \times n}$
- $\boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A}$

Zero matrix behaves like number "0" in matrix multiplication

Identity matrix behaves like number "1" in matrix multiplication

What are the powers of a matrix?

Definition 2.2.12

Similar to ordinary numbers multiplication

$\boldsymbol{A}$: square matrix
n : nonnegative integer

We define $\boldsymbol{A}^n$ as follows:

$$\boldsymbol{A}^n = \underbrace{\boldsymbol{AA} \ldots \boldsymbol{A}}_{\text{n times}} \qquad n \geq 1$$

$$\mathbf{A}^0 = \mathbf{I}$$

Properties of matrix powers

Remark 2.2.14

1. $\boldsymbol{A}^r\boldsymbol{A}^s = \boldsymbol{A}^{r+s}$ Similar to ordinary number

2. $(\boldsymbol{AB})^n \neq \boldsymbol{A}^n\boldsymbol{B}^n$ Different from ordinary number

$$\underbrace{(\boldsymbol{AB})\,(\boldsymbol{AB})\,\ldots\,(\boldsymbol{AB})}_{\text{n times}}$$

$$\underbrace{\boldsymbol{AA}\,\ldots\,\boldsymbol{A}}_{\text{n times}}\ \underbrace{\boldsymbol{BB}\,\ldots\,\boldsymbol{B}}_{\text{n times}}$$

Other ways to "zip" a matrix

Notation 2.2.15

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{pmatrix}$$

$$\boldsymbol{A} = (a_{ij})_{m\times p} = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_m \end{pmatrix} \begin{array}{l} 1^{st} \text{ row of } \boldsymbol{A} \\ 2^{nd} \text{ row of } \boldsymbol{A} \\ \\ m^{th} \text{ row of } \boldsymbol{A} \end{array}$$

zipped along the rows

$\boldsymbol{a}_i$ is a $1 \times p$ row matrix

$$\boldsymbol{a}_1 = (a_{11} \quad a_{12} \quad \ldots \quad a_{1p})$$
$$\boldsymbol{a}_2 = (a_{21} \quad a_{22} \quad \ldots \quad a_{2p})$$
$$\vdots$$
$$\boldsymbol{a}_m = (a_{m1} \quad a_{m2} \quad \ldots \quad a_{mp})$$

Other ways to "zip" a matrix

$$\boldsymbol{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{pmatrix}$$

Notation 2.2.15

$$\boldsymbol{B} = (b_{ij})_{p\times n} = (\boldsymbol{b}_1 \quad \boldsymbol{b}_2 \quad \cdots \quad \boldsymbol{b}_n)$$

1st column of B

2nd column of B

nth column of B

zipped along the columns

$\boldsymbol{b}_i$ is a p x 1 column matrix

$$\boldsymbol{b}_1 = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{p1} \end{pmatrix} \qquad \boldsymbol{b}_2 = \begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{p2} \end{pmatrix} \qquad \cdots \qquad \boldsymbol{b}_n = \begin{pmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{pn} \end{pmatrix}$$

What are some different ways to express matrix multiplication?

Notation 2.2.15

Example 2.2.16

$$\begin{array}{c} \\ \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{array}\overset{A}{\begin{pmatrix} 1 & 1 \\ 2 & 3 \\ -1 & -2 \end{pmatrix}} \underset{\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3}{\overset{B}{\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}}} = \overset{AB}{\begin{pmatrix} \underset{\mathbf{a}_1\mathbf{b}_1}{5} & \underset{\mathbf{a}_1\mathbf{b}_2}{7} & \underset{\mathbf{a}_1\mathbf{b}_3}{9} \\ \underset{\mathbf{a}_2\mathbf{b}_1}{14} & \underset{\mathbf{a}_2\mathbf{b}_2}{19} & \underset{\mathbf{a}_2\mathbf{b}_3}{24} \\ \underset{\mathbf{a}_3\mathbf{b}_1}{-9} & \underset{\mathbf{a}_3\mathbf{b}_2}{-12} & \underset{\mathbf{a}_3\mathbf{b}_3}{-15} \end{pmatrix}}$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{pmatrix} \quad \mathbf{B} = (\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n) \Rightarrow \mathbf{AB} = \begin{pmatrix} \mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 & \cdots & \mathbf{a}_1\mathbf{b}_n \\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 & \cdots & \mathbf{a}_2\mathbf{b}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_m\mathbf{b}_1 & \mathbf{a}_m\mathbf{b}_2 & \cdots & \mathbf{a}_m\mathbf{b}_n \end{pmatrix}$$

What are some different ways to express matrix multiplication?

Notation 2.2.15

Example 2.2.16

$$\overset{\boldsymbol{A}}{\begin{pmatrix}1 & 1\\ 2 & 3\\ -1 & -2\end{pmatrix}}\underset{\begin{array}{ccc}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \boldsymbol{b}_3\end{array}}{\overset{\boldsymbol{B}}{\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix}}} = \underset{\begin{array}{ccc}\boldsymbol{Ab}_1 & \boldsymbol{Ab}_2 & \boldsymbol{Ab}_3\end{array}}{\overset{\boldsymbol{AB}}{\begin{pmatrix}5 & 7 & 9\\ 14 & 19 & 24\\ -9 & -12 & -15\end{pmatrix}}}$$

$\boldsymbol{A}$(j th column of $\boldsymbol{B}$) = j th column of $\boldsymbol{AB}$

$$\boxed{\mathbf{AB} = (\mathbf{Ab}_1 \quad \mathbf{Ab}_2 \quad \cdots \quad \mathbf{Ab}_n)}$$

What are some different ways to express matrix multiplication?

Notation 2.2.15

Example 2.2.16

$$\begin{array}{c} \\ \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \boldsymbol{a}_3 \end{array}\overset{\boldsymbol{A}}{\begin{pmatrix} 1 & 1 \\ 2 & 3 \\ -1 & -2 \end{pmatrix}} \overset{\boldsymbol{B}}{\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}} = \overset{\boldsymbol{AB}}{\begin{pmatrix} 5 & 7 & 9 \\ 14 & 19 & 24 \\ -9 & -12 & -15 \end{pmatrix}} \begin{array}{c} \boldsymbol{a}_1\boldsymbol{B} \\ \boldsymbol{a}_2\boldsymbol{B} \\ \boldsymbol{a}_3\boldsymbol{B} \end{array}$$

(*i* th row of ***A***) ***B*** = *i* th row of ***AB***

$$\boxed{\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1\mathbf{B} \\ \mathbf{a}_2\mathbf{B} \\ \vdots \\ \mathbf{a}_m\mathbf{B} \end{pmatrix}}$$

How to express LS in matrix equation form?

Example 2.2.18

$$\begin{cases} 4x + 5y + 6z = 1 \\ x - y = 2 \\ y - z = 3 \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} 4 & 5 & 6 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

matrix equation form

$$\Leftrightarrow \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix} x + \begin{pmatrix} 5 \\ -1 \\ 1 \end{pmatrix} y + \begin{pmatrix} 6 \\ 0 \\ -1 \end{pmatrix} z = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

vector equation form

Remark 2.2.17

Consider the system of linear equations:

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

rewrite the system using the matrix multiplication:

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

coefficient matrix

variable matrix

constant matrix

How to express LS in matrix equation form?

Example 2.2.18

$$\begin{cases} 4x + 5y + 6z = 1 \\ x - y = 2 \\ y - z = 3 \end{cases}$$

don't confuse matrix equation form
with augmented matrix

$$\begin{pmatrix} 4 & 5 & 6 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

matrix equation form

$$\left(\begin{array}{ccc|c} 4 & 5 & 6 & 1 \\ 1 & -1 & 0 & 2 \\ 0 & 1 & -1 & 3 \end{array}\right)$$

augmented matrix

A concise notation for linear system

Remark 2.2.17

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

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}}_{\boldsymbol{A}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{\boldsymbol{b}}$$

$$\boxed{x_1 = u_1 \quad x_2 = u_2 \quad \cdots \quad x_n = u_n}$$

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

We can represent the linear system as $\boldsymbol{Ax} = \boldsymbol{b}$

A solution of the linear system

is represented by an n x 1 column matrix.

$\boldsymbol{u}$ is a solution of $\boldsymbol{Ax} = \boldsymbol{b}$

if and only if $\boldsymbol{Au} = \boldsymbol{b}$

How to express LS in vector equation form?

Remark 2.2.17

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

Linear system can also be written in vector equation form:

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

use of this form in chapter 3

How to perform matrix transpose?

Summary 2.2.19 - 2.2.20

Let $\boldsymbol{A} = (a_{ij})_{m \times n}$

Matrix Transpose	$\boldsymbol{A}^T$ (or $\boldsymbol{A}^t$)	interchanging the rows and columns of $\boldsymbol{A}$	$\boldsymbol{A}^T = (a_{ji})_{n \times m}$

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix} \qquad \boldsymbol{A}^T = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}$$

The transpose operator interchanges i and j of the entries.

Relation between transpose and symmetric matrix

Remark 2.2.21

2. A square matrix is symmetric if and only if

$$\boldsymbol{A} = \boldsymbol{A}^T.$$

The transpose operator does not change a symmetric matrix.

We can determine whether an (implicit) matrix $\boldsymbol{A}$ is symmetric by checking whether $\boldsymbol{A} = \boldsymbol{A}^T$.

What are some properties of transpose?

Theorem 2.2.22

Let $\boldsymbol{A}$ be an $m \times n$ matrix.

1. $(\boldsymbol{A}^T)^T = \boldsymbol{A}$
2. If $\boldsymbol{B}$ is an $m \times n$ matrix, then $(\boldsymbol{A} + \boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$.
3. If a is a scalar, then $(a\boldsymbol{A})^T = a\boldsymbol{A}^T$.
4. If $\boldsymbol{B}$ is an $n \times p$ matrix, then $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$.

Section 2.3

Inverses of Square Matrices

Objectives

- What is an invertible matrix?
- What is the inverse of a matrix?
- What are some basic properties of invertible matrices?
- What are the powers of a matrix?

Two proving techniques:

- *Direct proof*
- *Proof by contradiction*

Motivation

Discussion 2.3.1

a, b real numbers such that $a \neq 0$

To solve the equation $ax = b$

inverse of a

$$x = b/a = (a^{-1})\cdot b$$

Let $\boldsymbol{A}$, $\boldsymbol{B}$ be two matrices.

To solve the matrix equation $\boldsymbol{AX} = \boldsymbol{B}$

Can we do this: $\boldsymbol{X} = \boldsymbol{B/A}$?

We do not have "division" for matrices.

Can we find "inverses" for matrices "$\boldsymbol{A}^{-1}$" which have the similar property as a^{-1}?

What is an invertible matrix?

For ordinary numbers:
$a(a^{-1}) = 1$ $(a^{-1})a = 1$

Definition 2.3.2

A : square matrix of order *n*.

Is $\mathbf{I}$ itself invertible?

A is invertible

if there exists a square matrix ***B*** of order *n* such that

OR

$$\boldsymbol{AB} = \boldsymbol{I} \quad \text{~~and~~} \quad \boldsymbol{BA} = \boldsymbol{I}$$

The matrix ***B*** here is called an inverse of ***A***.

Does every matrix have an inverse? No

A square matrix is called singular if it has no inverse.

non-singular = invertible

What is an invertible matrix?

Example 2.3.3.1

$$\boldsymbol{A} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \qquad \boldsymbol{B} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$$

$$\boldsymbol{AB} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{I}$$

$$\boldsymbol{BA} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{I}$$

So $\boldsymbol{A}$ is invertible and $\boldsymbol{B}$ is an inverse of $\boldsymbol{A}$

Also $\boldsymbol{B}$ is invertible and $\boldsymbol{A}$ is an inverse of $\boldsymbol{B}$

A simple application

Example 2.3.3.2

2x1 variable column matrix $\begin{pmatrix} x \\ y \end{pmatrix}$

$$\begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \boxed{\boldsymbol{X}} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$$

Linear system ***AX*** = ***b***

$$\Rightarrow \quad \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \boldsymbol{X} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ 0 \end{pmatrix}$$

$$\Rightarrow \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{X} = \begin{pmatrix} 12 \\ 4 \end{pmatrix}$$

$$\Rightarrow \quad \boldsymbol{X} = \begin{pmatrix} 12 \\ 4 \end{pmatrix}$$

Solution of the linear system

Given a matrix ***A***, how to find the inverse?

An example of a singular matrix

Example 2.3.3.3

No inverse

Show that $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ is singular.

Proof by Contradiction

proving technique

Suppose $\boldsymbol{A}$ has an inverse: Let $\boldsymbol{B} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be the inverse

assume the opposite of the claim

Represent the object

By definition of inverses, $\boldsymbol{BA} = \boldsymbol{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

using definition

On the other hand, $\boldsymbol{BA} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a+b & 0 \\ c+d & 0 \end{pmatrix}$

direct multiplication

The two results for $\boldsymbol{BA}$ contradict with each other.

arrive at a contradiction

Conclusion: $\boldsymbol{A}$ is singular.

Properties of invertible matrices

Remark 2.3.4.1 (Cancellation Law for Matrices)

Let $\boxed{\boldsymbol{A} \text{ be an invertible}}$ matrix.

given condition　　to prove

$$\boxed{\boldsymbol{AB_1} = \boldsymbol{AB_2} \;\Rightarrow\; \boldsymbol{B_1} = \boldsymbol{B_2}}$$

Is this true?

If $\boldsymbol{A}$ is not invertible, then the Cancellation Law may not hold.

$$\boldsymbol{C_1A} = \boldsymbol{C_2A} \;\Rightarrow\; \boldsymbol{C_1} = \boldsymbol{C_2}$$

Prove it yourself

Direct Proof

Start from $\boldsymbol{AB_1} = \boldsymbol{AB_2}$

$\Rightarrow \boldsymbol{A'AB_1} = \boldsymbol{A'AB_2}$

$\Rightarrow \boldsymbol{IB_1} = \boldsymbol{IB_2}$

$\Rightarrow \boldsymbol{B_1} = \boldsymbol{B_2}$

Since $\boldsymbol{A}$ is invertible,
let $\boldsymbol{A'}$ be an inverse of $\boldsymbol{A}$.

introduce the inverse

How many inverses can a matrix have?

Theorem 2.3.5 Uniqueness of Inverses

If $\boldsymbol{B}$ and $\boldsymbol{C}$ are inverses of a square matrix $\boldsymbol{A}$, then $\boldsymbol{B} = \boldsymbol{C}$.

i.e. every invertible matrix has exactly one inverse

Direct Proof

$\boldsymbol{B}$ is an inverse of $\boldsymbol{A}$ $\Rightarrow \boldsymbol{BA} = \boldsymbol{I}$ and $\boxed{\boldsymbol{AB} = \boldsymbol{I}}$

given condition

definition of inverse

$\boldsymbol{C}$ is an inverse of $\boldsymbol{A}$ $\Rightarrow \boxed{\boldsymbol{CA} = \boldsymbol{I}}$ and $\boldsymbol{AC} = \boldsymbol{I}$

$$\begin{aligned}
&\boldsymbol{AB} = \boldsymbol{I} \\
\Rightarrow\ &\boldsymbol{CAB} = \boldsymbol{CI} \\
\Rightarrow\ &\boldsymbol{IB} = \boldsymbol{C} \\
\Rightarrow\ &\boldsymbol{B} = \boldsymbol{C}
\end{aligned}$$

Notation of an inverse matrix

Notation 2.3.6

Let $\boldsymbol{A}$ be an invertible matrix.
By Theorem 2.3.5, we know that there is exactly one inverse of $\boldsymbol{A}$.

We use $\boldsymbol{A}^{-1}$ to denote this unique inverse of $\boldsymbol{A}$.

In example 2.3.3

$$\mathbf{A} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \qquad \mathbf{A}^{-1} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$$

How to show one matrix is the inverse of another?

Remark 2.3.7

If you are asked to show : $\boldsymbol{A}^{-1} = \boldsymbol{B}$
you just need to check

$$\boldsymbol{AB} = \boldsymbol{I} \quad \text{and} \quad \boldsymbol{BA} = \boldsymbol{I}$$

In fact, only need to check any one of these two conditions.
(See Theorem 2.4.12)

Example Given $\boldsymbol{A}^2 + \boldsymbol{A} = \boldsymbol{I}$ show : $\boldsymbol{A}^{-1} = \boldsymbol{A} + \boldsymbol{I}$

$$\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{I}) = \boldsymbol{A}^2 + \boldsymbol{A} = \boldsymbol{I}$$

algebraic manipulation use given condition

$$(\boldsymbol{A}+\boldsymbol{I})\boldsymbol{A} = \boldsymbol{A}^2 + \boldsymbol{A} = \boldsymbol{I}$$

Conclusion : $\boldsymbol{A}^{-1} = \boldsymbol{A} + \boldsymbol{I}$

Example 2.3.8

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$ Show that if $ad - bc \neq 0$, then

formula

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

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} = \begin{pmatrix} \frac{ad-bc}{ad-bc} & \frac{-ab+ba}{ad-bc} \\ \frac{cd-dc}{ad-bc} & \frac{-cb+da}{ad-bc} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \frac{da-bc}{ad-bc} & \frac{db-bd}{ad-bc} \\ \frac{-ca+ac}{ad-bc} & \frac{-cb+ad}{ad-bc} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Conclusion : $\boldsymbol{A}$ is invertible and $\mathbf{A}^{-1} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}$

Theorem 2.3.9

The inverses can be expressed in terms of inverses of **A** and **B**

$\boldsymbol{A}, \boldsymbol{B}$: two invertible matrices (same size)
a : non-zero scalar

Matrix	Invertible?	Inverse
$a\boldsymbol{A}$	yes	$(a\boldsymbol{A})^{-1} = (1/a)\boldsymbol{A}^{-1}$
$\boldsymbol{A}^T$	yes	$(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$
$\boldsymbol{A}^{-1}$	yes	$(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$
$\boldsymbol{AB}$	yes	$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$

Example

$$\mathbf{A} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \qquad \mathbf{A}^{-1\mathbf{T}} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}^{\mathbf{T}}$$

$$\mathbf{A}^T = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} \qquad (\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$$

Remark 2.3.10

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

Given $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_k$ are all invertible matrices of the same size.

1. The product $\boldsymbol{A}_1\boldsymbol{A}_2\ldots\boldsymbol{A}_k$ is an invertible matrix.

   This follows from Theorem 2.3.9.4

2. The inverse of $\boldsymbol{A}_1\boldsymbol{A}_2\ldots\boldsymbol{A}_n$ is

$$(\boldsymbol{A}_1\boldsymbol{A}_2\ldots\boldsymbol{A}_k)^{-1} = (\boldsymbol{A}_k)^{-1}\ldots(\boldsymbol{A}_2)^{-1}(\boldsymbol{A}_1)^{-1}$$

What are the powers of a matrix?

Definition 2.3.11

A : square matrix
n : nonnegative integer

Similar to ordinary number

We define $\boldsymbol{A}^n$ as follows:

$$\boldsymbol{A}^n = \underbrace{\boldsymbol{A}\boldsymbol{A} \ldots \boldsymbol{A}}_{\text{n times}} \qquad n \geq 1$$

$$\mathbf{A}^0 = \mathbf{I}$$

What about negative powers?

If **A** is invertible,

$$\mathbf{A}^{-n} = (\mathbf{A}^{-1})^n = \underbrace{\boldsymbol{A}^{-1}\boldsymbol{A}^{-1} \ldots \boldsymbol{A}^{-1}}_{\text{n times}}$$

Properties of matrix powers

Remark 2.3.13

1. $\boldsymbol{A}^r\boldsymbol{A}^s = \boldsymbol{A}^{r+s}$

   $\boldsymbol{A}^r\boldsymbol{A}^{-s} = \boldsymbol{A}^{r-s}$

Similar to ordinary number

2. $(\boldsymbol{A}^n)^{-1} = \boldsymbol{A}^{-n}$

inverse of nth power

nth power of inverse

Section 2.4

Elementary Matrices

Objectives

- What are elementary matrices?
- How are elementary matrices related to elementary row operations?
- How to find inverse of an elementary matrix?

Overview

- Perform e.r.o. R to a matrix $\boldsymbol{A}$ is the same as pre-multiply a certain square matrix $\boldsymbol{E}$ to $\boldsymbol{A}$

$$\boldsymbol{A} \xrightarrow{R} \boldsymbol{B} \qquad\qquad \boldsymbol{EA} = \boldsymbol{B}$$

- Every e.r.o. R has a "reverse" ("undo") operation R'

$$\boldsymbol{A} \xrightarrow{R} \boldsymbol{B} \xrightarrow{R'} \boldsymbol{A}$$

- R' is also an e.r.o.
- R' corresponds to a square matrix $\boldsymbol{E'}$
- $\boldsymbol{E'}$ is the inverse of $\boldsymbol{E}$

How to find the matrix E ?

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{2R_2} \mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_3 + 2R_1} \mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

Which matrices have the same effect as e.r.o.?

Discussion 2.4.2.1

elementary row operations of the first type:

Multiply a row by a constant

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix} \underset{\xleftarrow{\frac{1}{2}R_2}}{\xrightarrow{2R_2}} \boldsymbol{B} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 4 & -2 & 6 & 12 \\ 1 & 4 & 4 & 0 \end{pmatrix}$$

$$\boldsymbol{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{E}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{EA} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 4 & -2 & 6 & 12 \\ 1 & 4 & 4 & 0 \end{pmatrix} = \boldsymbol{B}$$

Elementary matrices of first type

Discussion 2.4.2.1

Let $\boldsymbol{A}$ be an m x n matrix.

Let $\boldsymbol{E}$ be a square matrix of order m:

$$\boldsymbol{E} = \begin{array}{c} \left(\begin{array}{ccccccc} 1 & & & & & & \\ & \ddots & & & & \mathbf{0} & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & \mathbf{0} & & & & \ddots & \\ & & & & & & 1 \end{array}\right) \end{array} \begin{array}{l} \\ \\ \\ \leftarrow i\text{th row} \\ \\ \\ \\ \end{array}$$

EA : multiplying the *i*th row of ***A*** by *c*. cR_i

How to find inverse of an elementary matrix?

Discussion 2.4.2.1

$$E = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i\text{th row} \\ \\ \\ \\ \end{array}$$

$$\begin{array}{c} \uparrow \\ i\text{th column} \end{array}$$

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

$$E^{-1} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & \frac{1}{c} & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i\text{th row} \\ \\ \\ \\ \end{array}$$

$$\begin{array}{c} \uparrow \\ i\text{th column} \end{array}$$

$E^{-1}A$: multiplying the i^{th} row of A by $1/c$. $(1/c)R_i$

Which matrices have the same effect as e.r.o.?

Discussion 2.4.2.2

elementary row operations of the second type:

Interchange two rows

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix} \underset{\xleftarrow{R_2 \leftrightarrow R_3}}{\xrightarrow{R_2 \leftrightarrow R_3}} \boldsymbol{B} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 1 & 4 & 4 & 0 \\ 2 & -1 & 3 & 6 \end{pmatrix}$$

$$\boldsymbol{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \boldsymbol{E}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\boldsymbol{EA} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 1 & 4 & 4 & 0 \\ 2 & -1 & 3 & 6 \end{pmatrix} = \boldsymbol{B}$$

Elementary matrices of second type

Discussion 2.4.2.2

Let $\boldsymbol{A}$ be an $m \times n$ matrix.

Let $\boldsymbol{E}$ be a square matrix of order m :

$$\boldsymbol{E} = \begin{array}{c@{}l} \left(\begin{array}{ccccccccc} 1 & & & & & & & & \\ & \ddots & & & & & & \text{O} & \\ & & 1 & & & & & & \\ & & & 0 & & & 1 & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & 1 & & & & 0 & \\ & & & & & & & & 1 \\ & \text{O} & & & & & & & \ddots \\ & & & & & & & & & 1 \end{array}\right) & \begin{array}{l} \\ \\ \\ \leftarrow i\text{th row} \\ \\ \\ \\ \leftarrow j\text{th row} \\ \\ \\ \\ \end{array} \end{array}$$

$\boldsymbol{EA}$: interchanging the ith and jth rows of $\boldsymbol{A}$. $R_i \leftrightarrow R_j$

$$\boldsymbol{E}^{-1} = \boldsymbol{E}$$

How to find inverse of an elementary matrix?

Which matrices have the same effect as e.r.o.?

Discussion 2.4.2.3

elementary row operations of the third type:

Add a multiple of a row to another row

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix} \underset{R_3 - 2R_1}{\overset{R_3 + 2R_1}{\rightleftarrows}} \boldsymbol{B} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 3 & 4 & 8 & 6 \end{pmatrix}$$

$$\boldsymbol{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \qquad \boldsymbol{E}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$$

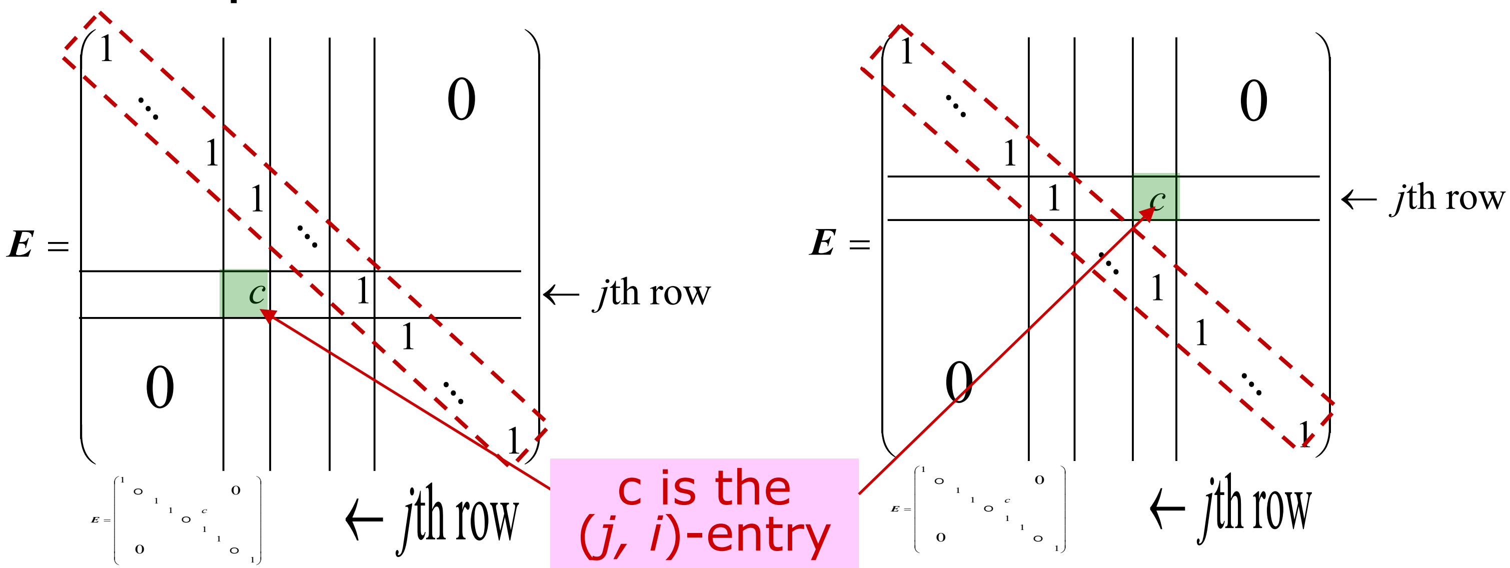
(3,1)-entry

$$\boldsymbol{EA} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 3 & 4 & 8 & 6 \end{pmatrix} = \boldsymbol{B}$$

Discussion 2.4.2.3

Let $\boldsymbol{A}$ be an m x n matrix.

$\boldsymbol{E}$ be a square matrix of order m as shown below



below diagonal if $i < j$ above diagonal if $i > j$

$\boldsymbol{EA}$: adding c times of ith row to jth row of $\boldsymbol{A}$ $R_j + cR_i$

Discussion 2.4.2.3

Let $\boldsymbol{A}$ be an $m \times n$ matrix.

$$E = \left(\begin{array}{ccccccccc} 1 & & & & & & & & \\ & \ddots & & & & & & 0 & \\ & & 1 & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & & c & & 1 & & & \\ & & & & & & 1 & & \\ & 0 & & & & & & \ddots & \\ & & & & & & & & 1 \end{array}\right) \begin{array}{c} \\ \\ \\ \\ \\ \leftarrow j\text{th row} \\ \\ \\ \\ \end{array}$$

$$\uparrow i\text{th column} \qquad \uparrow j\text{th column}$$

$$E^{-1} = \left(\begin{array}{ccccccccc} 1 & & & & & & & & \\ & \ddots & & & & & & 0 & \\ & & 1 & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & & -c & & 1 & & & \\ & & & & & & 1 & & \\ & 0 & & & & & & \ddots & \\ & & & & & & & & 1 \end{array}\right) \begin{array}{c} \\ \\ \\ \\ \\ \leftarrow j\text{th row} \\ \\ \\ \\ \end{array}$$

$$\uparrow i\text{th column} \qquad \uparrow j\text{th column}$$

if $i < j$

$$E^{-1} = \left(\begin{array}{ccccccccc} 1 & & & & & & & & \\ & \ddots & & & & & & 0 & \\ & & 1 & & & & & & \\ & & & 1 & & -c & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & 1 & & \\ & 0 & & & & & & \ddots & \\ & & & & & & & & 1 \end{array}\right) \begin{array}{c} \\ \\ \\ \leftarrow j\text{th row} \\ \\ \\ \\ \\ \\ \end{array}$$

$$\uparrow j\text{th column} \qquad \leftarrow j\text{th row}$$

if $i > j$

$R_j - cR_i$

$\boldsymbol{E}^{-1}\boldsymbol{A}$: adding $-c$ times of ith row to jth row of $\boldsymbol{A}$.

Definition 2.4.3 & Remark 2.4.4

A square matrix is called an elementary matrix if it can be obtained from an identity matrix by performing a single elementary row operation.

1. The matrices $\boldsymbol{E}$ in Discussion 2.4.2 are elementary matrices.
   Every elementary matrix is of one of the three types in Discussion 2.4.2.
2. All elementary matrices are invertible and their inverse are also elementary matrices.

Example 2.4.5

$$\mathbf{A} = \begin{pmatrix} 0 & 4 & 2 \\ -2 & 1 & -3 \\ 1 & 0 & 2 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

A and ***B*** are row equivalent

i.e. ***B*** can be obtained from ***A*** by performing a series of e.r.o. and vice versa

$$\mathbf{A} \longrightarrow \longrightarrow \longrightarrow \cdots \longrightarrow \mathbf{B}$$

i.e. ***B*** can be obtained from ***A*** by pre-multiplying ***A*** with a series of elementary matrices and vice versa

$$E_n \ldots E_2 E_1 A = B$$

Example 2.4.5

$$\underset{A}{\begin{pmatrix} 0 & 4 & 2 \\ -2 & 1 & -3 \\ 1 & 0 & 2 \end{pmatrix}} \xrightarrow[\mathbf{E}_1]{R_1 \leftrightarrow R_3} \underset{E_1A}{\begin{pmatrix} 1 & 0 & 2 \\ -2 & 1 & -3 \\ 0 & 4 & 2 \end{pmatrix}} \xrightarrow[\mathbf{E}_2]{R_2 + 2R_1} \underset{E_2E_1A}{\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 4 & 2 \end{pmatrix}}$$

$$\xrightarrow[\mathbf{E}_3]{R_3 - 4R_2} \underset{E_3E_2E_1A}{\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{pmatrix}} \xrightarrow[\mathbf{E}_4]{-\frac{1}{2}R_3} \overset{B}{\underset{E_4E_3E_2E_1A}{\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}}}$$

$$\mathbf{E}_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \mathbf{E}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{E}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix} \quad \mathbf{E}_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

$$E_4E_3E_2E_1A = B$$

Example 2.4.5

$$E_4 E_3 E_2 E_1 A = B$$

$$E_3 E_2 E_1 A = E_4^{-1} B$$

$$\Rightarrow E_2 E_1 A = E_3^{-1} E_4^{-1} B$$

$$\Rightarrow E_1 A = E_2^{-1} E_3^{-1} E_4^{-1} B$$

$$\Rightarrow \boxed{A = E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1} B}$$

A **and** ***B*** **are row equivalent**

B in terms of
A and elementary matrices

A in terms of
B and elementary matrices

$$\mathbf{E}_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \mathbf{E}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{E}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix} \quad \mathbf{E}_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

$$\mathbf{E}_1^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \mathbf{E}_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{E}_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix} \quad \mathbf{E}_4^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Section 2.4

Elementary Matrices

Objectives

- How to find the inverse of an invertible matrix?
- How to tell whether a matrix is invertible?
- What can we say about an invertible matrix?

Example 2.4.5

$\boldsymbol{A}$ and $\boldsymbol{B}$ are row equivalent

$$\boldsymbol{A} \rightarrow \ \rightarrow \ \rightarrow \ \ldots \ \rightarrow \boldsymbol{B}$$

$$\boldsymbol{E_n} \ldots \boldsymbol{E_2 E_1 A} = \boldsymbol{B}$$

$$\boldsymbol{A} \leftarrow \ \ldots \ \leftarrow \ \leftarrow \ \leftarrow \boldsymbol{B}$$

$$\boldsymbol{A} = \boldsymbol{E_1}^{-1} \boldsymbol{E_2}^{-1} \ldots \boldsymbol{E_n}^{-1} \boldsymbol{B}$$

Take note of the order

Remark 2.4.6

Proof of Theorem 1.2.7

If augmented matrices of two linear systems
are row equivalent,
then the two systems have the same set of solutions.

The idea is to use elementary matrices

Read up!

$\boldsymbol{Ax} = \boldsymbol{c}$ and $\boldsymbol{Bx} = \boldsymbol{d}$

$\boxed{\boldsymbol{E}_{n} \ldots \boldsymbol{E}_{2}\boldsymbol{E}_{1}\boldsymbol{A}}$ $\boxed{\boldsymbol{E}_{n} \ldots \boldsymbol{E}_{2}\boldsymbol{E}_{1}\boldsymbol{c}}$

How to find inverse matrix?

Example 2.4.9

Find the inverse of

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix} \qquad \boldsymbol{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

if it exists.

Form the 3x6 augmented matrix

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right)$$

Gauss-Jordan Elimination

How to find inverse matrix?

Example 2.4.9 Gauss-Jordan Elimination

$$\underset{A}{\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right)}\ \begin{matrix} \\ \\ \\ & I \end{matrix} \xrightarrow[R_3 - R_1]{R_2 - 2R_1} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_3 + 2R_2} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right) \xrightarrow{-R_3} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right)$$

$$\xrightarrow[R_2 + 3R_3]{R_1 - 3R_3} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right) \xrightarrow{R_1 - 2R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right)$$

$I \qquad A^{-1}$

RREF

$$E_k \cdots E_2 E_1 A = I \Rightarrow E_k \cdots E_2 E_1 I = A^{-1}$$

Why does it work?

Question:
What if the RREF is not $\boldsymbol{I}$?

Discussion 2.4.8

Suppose the RREF of $\boldsymbol{A}$ is the identity matrix

$\boldsymbol{A}$: invertible matrix of order n

$$\boldsymbol{A} \xrightarrow{R_1} \xrightarrow{R_2} \cdots \xrightarrow{R_k} \boldsymbol{I}$$

elementary matrices $\boldsymbol{E}_1$ $\boldsymbol{E}_2$ $\boldsymbol{E}_k$

$$\boldsymbol{E}_k \cdots \boldsymbol{E}_2 \boldsymbol{E}_1 \boldsymbol{A} = \boldsymbol{I} \quad \Rightarrow \boldsymbol{E}_k \cdots \boldsymbol{E}_2 \boldsymbol{E}_1 = \boldsymbol{A}^{-1}$$

Form an $n \times 2n$ "augmented matrix" $(\boldsymbol{A} \mid \boldsymbol{I})$

$$\boldsymbol{E}_k \cdots \boldsymbol{E}_2 \boldsymbol{E}_1 (\boldsymbol{A} \mid \boldsymbol{I})$$

$$(\boldsymbol{E}_k \cdots \boldsymbol{E}_2 \boldsymbol{E}_1 \boldsymbol{A} \mid \boldsymbol{E}_k \cdots \boldsymbol{E}_2 \boldsymbol{E}_1 \boldsymbol{I})$$

$$= (\boldsymbol{I} \mid \boldsymbol{A}^{-1})$$

applying e.r.o. to $(\boldsymbol{A} \mid \boldsymbol{I})$
same as
applying e.r.o. to both $\boldsymbol{A}$ and $\boldsymbol{I}$

$$\boxed{(\boldsymbol{A} \mid \boldsymbol{I}) \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} (\boldsymbol{I} \mid \boldsymbol{A}^{-1})}$$

A very^3 important theorem

Theorem 2.4.7

Any 1 of the 4 statements implies the other 3.

Let $\boldsymbol{A}$ be a square matrix.

The following statements are equivalent

1. $\boldsymbol{A}$ is invertible.
2. The linear system $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution.
3. The reduced row-echelon form of $\boldsymbol{A}$ is an identity matrix.
4. $\boldsymbol{A}$ can be expressed as a product of elementary matrices.

Equivalent Statements

Two equivalent statements

statement 1

Two implications

statement 2

Four equivalent statements

How many implications are there?

Do this for every pair of statements.

1 & 2	1 & 3	1 & 4
2 & 3	2 & 4	3 & 4

Twelve implications

1. $\boldsymbol{A}$ is invertible
2. $\boldsymbol{Ax} = \mathbf{0}$ has only trivial solution
3. RREF of $\boldsymbol{A}$ is $\boldsymbol{I}$
4. $\boldsymbol{A}$ a product of elementary matrices

What's it for?

Applications

1. $\boldsymbol{A}$ is invertible
2. $\boldsymbol{Ax} = \mathbf{0}$ has only trivial solution
3. RREF of $\boldsymbol{A}$ is $\boldsymbol{I}$
4. $\boldsymbol{A}$ a product of elementary matrices

(1) Given $\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$ is invertible.

How many solutions does the linear system have?

$$\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Ans: Only the trivial solution

Apply: Statement 1 ⇒ Statement 2

(2) Given $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \xrightarrow{\text{G.J.E.}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Is $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ invertible? Ans: Yes

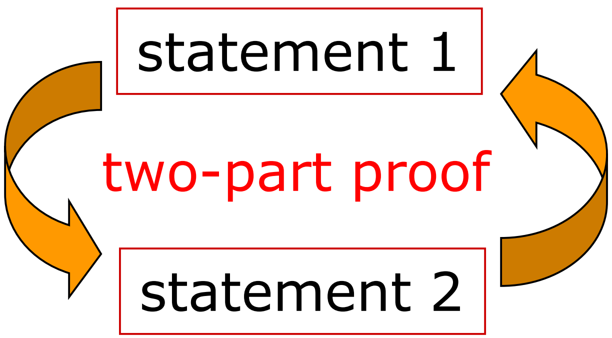
Apply: Statement 3 ⇒ Statement 1

How to prove it?

1. $\boldsymbol{A}$ is invertible
2. $\boldsymbol{Ax} = \boldsymbol{0}$ has only trivial solution
3. RREF of $\boldsymbol{A}$ is $\boldsymbol{I}$
4. $\boldsymbol{A}$ a product of elementary matrices

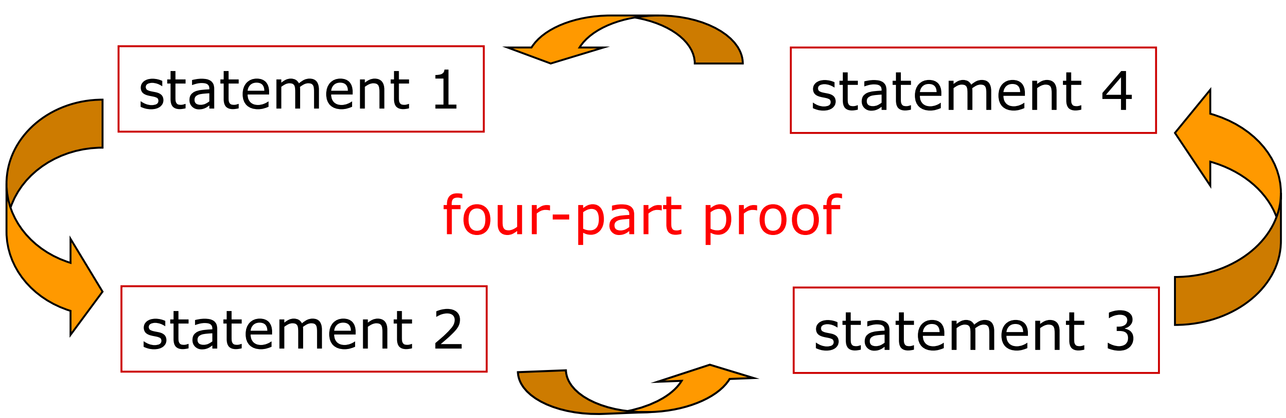
Proof

Two equivalent statements



Four equivalent statements

twelve-part proof ?



How to prove it?

1. $\boldsymbol{A}$ is invertible
2. $\boldsymbol{Ax} = \mathbf{0}$ has only trivial solution
3. RREF of $\boldsymbol{A}$ is $\boldsymbol{I}$
4. $\boldsymbol{A}$ a product of elementary matrices

Proof

$(1 \Rightarrow 2)$ Start with $\boldsymbol{Au} = \mathbf{0}$ and show $\boldsymbol{u} = \mathbf{0}$

$(2 \Rightarrow 3)$ Convert $\boldsymbol{Ax} = \mathbf{0}$ to augmented matrix $(\boldsymbol{A} \mid \mathbf{0})$ and consider the pivot columns of its RREF

$(3 \Rightarrow 4)$ Express the Gauss-Jordan Elimination from $\boldsymbol{A}$ to $\boldsymbol{I}$ in terms of elementary matrices

$(4 \Rightarrow 1)$ Product of invertible matrices is invertible

Remark 2.4.10

To check whether a square matrix is invertible:

- Look at the RREF
  - RREF = $\boldsymbol{I}$ implies invertible
  - RREF $\neq$ $\boldsymbol{I}$ implies not invertible
- Look at REF
  - REF has no zero row implies invertible
  - REF has zero rows implies not invertible

Example 2.4.11.1

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 4 \\ 3 & 6 & 3 \end{pmatrix} \xrightarrow{GE} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

$\boldsymbol{A}$ is not invertible.

Example 2.4.11.2

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Example 2.3.8
$\boldsymbol{A}$ is invertible if $ad - bc \neq 0$.

not quite the same!

$\boldsymbol{A}$ is invertible if and only if $ad - bc \neq 0$.

$\boxed{ad - bc \neq 0}$

Example 2.3.8

$\boxed{\boldsymbol{A} \text{ is invertible}}$

Theorem 2.4.7: 1 $\Rightarrow$ 3

1. $\boldsymbol{A}$ is invertible
3. RREF of $\boldsymbol{A}$ is $\boldsymbol{I}$

Read the solution in textbook

If we only know $\boldsymbol{AB} = \boldsymbol{I}$, can we say $\boldsymbol{A}$ and $\boldsymbol{B}$ are inverses of each other?

Theorem 2.4.12

Let $\boldsymbol{A}$, $\boldsymbol{B}$ be square matrices of the same size.

If $\boldsymbol{AB} = \boldsymbol{I}$,
then $\boldsymbol{BA} = \boldsymbol{I}$.
So $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible, $\boldsymbol{A}^{-1} = \boldsymbol{B}$, $\boldsymbol{B}^{-1} = \boldsymbol{A}$.

Outline of proof

$\boxed{\boldsymbol{AB} = \boldsymbol{I}} \rightarrow$ $\boxed{\text{1. } \boldsymbol{B} \text{ is invertible}} \rightarrow$ $\boxed{\text{2. } \boldsymbol{B}^{-1} = \boldsymbol{A}} \rightarrow$ $\boxed{\text{3. } \boldsymbol{BA} = \boldsymbol{I}}$

First prove $\boldsymbol{AB} = \boldsymbol{I} \Rightarrow \boldsymbol{B}$ is invertible

Theorem 2.4.12

Given $\boldsymbol{AB} = \boldsymbol{I}$ Show $\boldsymbol{B}$ is invertible

Consider the homogeneous system $\boldsymbol{Bx} = \mathbf{0}$.

$$
\begin{aligned}
\boldsymbol{Bu} &= \mathbf{0} && \text{Start with the system} \\
\Rightarrow \boldsymbol{ABu} &= \boldsymbol{A}\mathbf{0} && \text{algebraic manipulation} \\
\Rightarrow \boldsymbol{Iu} &= \mathbf{0} && \text{LHS: use given condition} \\
\boldsymbol{u} &= \mathbf{0}
\end{aligned}
$$

The system $\boldsymbol{Bx} = \mathbf{0}$ has only the trivial solution.
By Thm 2.4.7 ($2 \Rightarrow 1$), $\boldsymbol{B}$ is invertible.

Theorem 2.4.7
1. $\boldsymbol{A}$ is invertible.
2. $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution.

Next prove $\boldsymbol{B}$ is invertible $\Rightarrow \boldsymbol{B}^{-1} = \boldsymbol{A}$ and $\boldsymbol{BA} = \boldsymbol{I}$

Theorem 2.4.12

Given $\boldsymbol{AB} = \boldsymbol{I}$

We have shown $\boldsymbol{B}$ is invertible

To show $\boldsymbol{B}^{-1} = \boldsymbol{A}$

and $\boldsymbol{BA} = \boldsymbol{I}$

$$\begin{aligned} \boldsymbol{AB} &= \boldsymbol{I} \\ \Rightarrow \boldsymbol{AB}\boldsymbol{B}^{-1} &= \boldsymbol{I}\boldsymbol{B}^{-1} \\ \Rightarrow \boldsymbol{AI} &= \boldsymbol{B}^{-1} \\ \Rightarrow \boldsymbol{A} &= \boldsymbol{B}^{-1} \\ \Rightarrow \boldsymbol{BA} &= \boldsymbol{BB}^{-1} \\ \Rightarrow \boldsymbol{BA} &= \boldsymbol{I} \end{aligned}$$

use given condition

use **B** is invertible

Summary 2.4.15-16

elementary column operations of the first type:

Multiply a column by a constant

elementary column operations of the second type:

Interchange two columns

elementary column operations of the third type:

Add a multiple of a column to another column

Perform e.c.o. C to a matrix **A** is the same as post-multiply a certain square matrix **E** to **A**

$$\mathbf{A} \xrightarrow{C} \mathbf{B} \qquad \boldsymbol{AE} = \boldsymbol{B}$$

$$\mathbf{I} \xrightarrow{C} \mathbf{E}$$

Section 2.5

Determinants

Objectives

- What is the determinant of a matrix?
- What is cofactor expansion?
- How to find determinant?

Motivation

Discussion 2.5.1

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$\boldsymbol{A}$ is invertible if and only if $\underbrace{ad - bc}_{\text{determinant}} \neq 0$.

det(***A***) $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$

$\boldsymbol{A}$: nxn square matrix

$\boldsymbol{A}$ is invertible if and only if "*determinant of* ***A***" $\neq 0$.

What is a 3x3 determinant?

Example 2.5.4.2

$$\boldsymbol{B} = \begin{pmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{pmatrix}$$

Define determinant "inductively"

3x3 determinant defined in terms of 2x2 determinants

$$\det(\boldsymbol{B}) = (-3)\underbrace{\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix}}_{\boldsymbol{M}_{11}} - (-2)\underbrace{\begin{vmatrix} 4 & 1 \\ 0 & 4 \end{vmatrix}}_{\boldsymbol{M}_{12}} + 4\underbrace{\begin{vmatrix} 4 & 3 \\ 0 & 2 \end{vmatrix}}_{\boldsymbol{M}_{13}}$$

Submatrices of $\boldsymbol{B} \rightarrow$ $\boldsymbol{M}_{11}$ $\boldsymbol{M}_{12}$ $\boldsymbol{M}_{13}$

$$= -3(3 \times 4 - 1 \times 2) + 2(4 \times 4 - 1 \times 0) + 4(4 \times 2 - 3 \times 0)$$

$$= 34$$

What is a 4x4 determinant?

Example 2.5.4.3

4x4 determinant defined
in terms of 3x3 determinants

$$\mathbf{C} = \begin{pmatrix} 0 & -1 & 2 & 0 \\ 2 & -3 & 3 & -2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{pmatrix}$$

$\det(\boldsymbol{C})$

$$0\underbrace{\begin{vmatrix} -3 & 3 & -2 \\ 2 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix}}_{M_{11}} - (-1)\underbrace{\begin{vmatrix} 2 & 3 & -2 \\ 0 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix}}_{M_{12}} + 2\underbrace{\begin{vmatrix} 2 & -3 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{vmatrix}}_{M_{13}} - 0\underbrace{\begin{vmatrix} 2 & -3 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{vmatrix}}_{M_{14}}$$

in terms of 2x2 determinants

in terms of 2x2 determinants

in terms of 2x2 determinants

in terms of 2x2 determinants

What is an nxn determinant?

Definition 2.5.2

Let $\boldsymbol{A} = (a_{ij})$ be an n x n matrix.

If $\boldsymbol{A} = (a_{11})$ is a 1 x 1 matrix, then $\det(\boldsymbol{A}) = a_{11}$

For $n > 1$,
let $\boldsymbol{M_{1j}}$ be the $(n-1)$ x $(n-1)$ matrix obtained from $\boldsymbol{A}$ by deleting the 1*st* row and the *j*th column.

$A_{11} = \det(\boldsymbol{M_{11}})$ $A_{13} = \det(\boldsymbol{M_{13}})$

$A_{12} = -\det(\boldsymbol{M_{12}})$ $A_{14} = -\det(\boldsymbol{M_{14}})$

etc… cofactors of $\boldsymbol{A}$

The determinant of $\boldsymbol{A}$ is defined to be

$$\det(\mathbf{A}) = a_{11}A_{11} + a_{12}A_{12} + \dots + a_{1n}A_{1n}$$

not practical for large matrices

cofactor expansion along row 1

What is an (i, j)-cofactor ?

Definition 2.5.2

Let $\boldsymbol{A} = (a_{ij})$ be an n x n matrix.

$\boldsymbol{M_{ij}}$:deleting ith row and jth column from $\boldsymbol{A}$

$A_{ij} = (-1)^{i+j}\det(\boldsymbol{M_{ij}})$ (i, j)-cofactor of $\boldsymbol{A}$

$A_{11} = \det(\boldsymbol{M_{11}})$ $A_{13} = \det(\boldsymbol{M_{13}})$

$A_{12} = -\det(\boldsymbol{M_{12}})$ $A_{14} = -\det(\boldsymbol{M_{14}})$

etc… cofactors of $\boldsymbol{A}$

How to compute determinant?

Theorem 2.5.6 **(Cofactor Expansions)**

$$\boxed{\det(\mathbf{A}) = a_{11}A_{11} + a_{12}A_{12} + \ldots + a_{1n}A_{1n}}$$

det($\boldsymbol{A}$) can be expressed as a cofactor expansion using any row or column of $\boldsymbol{A}$.

for any $i = 1, 2, \ldots, n$

$$\det(\mathbf{A}) = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}$$

cofactor expansion along row i

for any $j = 1, 2, \ldots, n$

$$\det(\mathbf{A}) = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj}$$

cofactor expansion along column j

How to compute determinants

$$\mathbf{C} = \begin{pmatrix} 0 & -1 & 2 & 0 \\ 2 & -3 & 3 & -2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{pmatrix} = -2\begin{vmatrix} -1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix}$$

Example 2.5.7

$$\boldsymbol{B} = \begin{pmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{pmatrix}$$

(i, j)-cofactor of $\boldsymbol{B}$:

$$B_{ij} = (-1)^{i+j}\det(\boldsymbol{M_{ij}})$$

$$\det(\boldsymbol{B}) = \underbrace{-4\begin{vmatrix} -2 & 4 \\ 2 & 4 \end{vmatrix}}_{B_{21}} + \underbrace{3\begin{vmatrix} -3 & 4 \\ 0 & 4 \end{vmatrix}}_{B_{22}} \underbrace{-1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix}}_{B_{23}} = 34$$

cofactor expansion along row 2

$$\det(\boldsymbol{B}) = \underbrace{4\begin{vmatrix} 4 & 3 \\ 0 & 2 \end{vmatrix}}_{B_{13}} \underbrace{-1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix}}_{B_{23}} + \underbrace{4\begin{vmatrix} -3 & -2 \\ 4 & 3 \end{vmatrix}}_{B_{33}} = 34$$

cofactor expansion along column 3

Determinant of triangular matrix

Theorem 2.5.8 & Example 2.5.9

If $\boldsymbol{A}$ is a triangular matrix, diagonal matrix
then the determinant of $\boldsymbol{A}$ is equal to the product of the diagonal entries of $\boldsymbol{A}$.

$$\begin{vmatrix} -1 & 3.5 & 14 \\ 0 & 5 & -26 \\ 0 & 0 & 2 \end{vmatrix} = (-1)\times 5 \times 2 = -10$$

$$\begin{vmatrix} -2 & 0 & 0 \\ 99 & 0 & 0 \\ 10 & -4.5 & 10 \end{vmatrix} = (-2)\times 0 \times 10 = 0$$

Suppose we want to prove certain property holds for all (specific type of) square matrices

Mathematical Induction

Show: Property P holds for all square matrices.

We can try to prove the following:

1. P works for all 1×1 matrices   Base case
2. Show that, if P works for all $k \times k$ matrices, then P works for all $(k+1) \times (k+1)$ matrices

Inductive step

Mathematical Induction

works for 1×1 $\Rightarrow$ works for 2×2 $\Rightarrow$ works for 3×3 $\Rightarrow$ … $\Rightarrow$ works for $n \times n$ $\Rightarrow$ …

Repeatedly, we have shown that P works for all square matrices

Determinant and transpose

Theorem 2.5.10

If $\boldsymbol{A}$ is a square matrix, then $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^T)$.

Example

$$\det(\mathbf{C}) = \begin{vmatrix} 0 & -1 & 2 & 0 \\ 2 & -3 & 3 & -2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{vmatrix} = \det(\mathbf{C}^T) = \begin{vmatrix} 0 & 2 & 0 & 0 \\ -1 & -3 & 2 & 0 \\ 2 & 3 & 4 & 2 \\ 0 & -2 & 0 & -1 \end{vmatrix}$$

Prove by mathematical induction

Let P be the property: $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^T)$

Base case P works for 1 x 1 matrices

Inductive step

We assume $\det(\boldsymbol{B}) = \det(\boldsymbol{B}^T)$ for any k x k matrix $\boldsymbol{B}$

Show $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^T)$ for any (k + 1) x (k + 1) matrix $\boldsymbol{A}$

det($\boldsymbol{B}$) = det($\boldsymbol{B}^T$) for 3 x 3 matrix $\boldsymbol{B}$
□ det($\boldsymbol{A}$) = det($\boldsymbol{A}^T$) for 4 x 4 matrix $\boldsymbol{A}$

Example 2.5.11

$$\det(\boldsymbol{C}) = \begin{vmatrix} 0 & -1 & 2 & 0 \\ 2 & -3 & 3 & -2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{vmatrix} \qquad \det(\boldsymbol{C}^T) = \begin{vmatrix} 0 & 2 & 0 & 0 \\ -1 & -3 & 2 & 0 \\ 2 & 3 & 4 & 2 \\ 0 & -2 & 0 & -1 \end{vmatrix}$$

$\boldsymbol{M}_{12}$ $\qquad$ $\boldsymbol{M}_{12}{}^{\mathrm{T}}$

cofactor expansion along row 1 $\qquad$ cofactor expansion along column 1

Inductive step

4x4 $\qquad$ 3x3 $\qquad$ 3x3 $\qquad$ 3x3 $\qquad$ 3x3

$$\det(\boldsymbol{C}) = 0\det(\boldsymbol{M}_{11}) - (-1)\det(\boldsymbol{M}_{12}) + 2\det(\boldsymbol{M}_{13}) - (0)\det(\boldsymbol{M}_{14})$$

$$\det(\boldsymbol{C}^T) = 0\det(\boldsymbol{M}_{11}{}^T) - (-1)\det(\boldsymbol{M}_{12}{}^T) + 2\det(\boldsymbol{M}_{13}{}^T) - (0)\det(\boldsymbol{M}_{14}{}^T)$$

Section 2.5

Determinants

Objectives

- How do matrix operations affect determinants?
- What is the relation between invertibility and determinant?
- What is the adjoint of a matrix?
- What is Cramer's rule?

Two main results:

$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$

$\mathbf{A}$ is invertible if and only if $\det(\mathbf{A}) \neq 0$

Matrices with two identical rows (columns)

Theorem 2.5.12 & Example 2.5.13

1. The determinant of a square matrix with two identical rows is zero.

2. The determinant of a square matrix with two identical columns is zero.

$$\begin{pmatrix} 4 & -2 \\ 4 & -2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 4 \\ -1 & 10 & 4 \\ 1 & 2 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 1 \\ -1 & -3 & -3 & 9 \\ 2 & 4 & 4 & 0 \\ 0 & -2 & -2 & -1 \end{pmatrix}$$

det = 0 det = 0 det = 0

Theorem 2.5.12 (Exercise 2.58)

Prove by mathematical induction.

The determinant of a square matrix with two identical rows is zero

Base case

2×2: $\overset{\mathbf{0}}{\begin{vmatrix} a & b \\ a & b \end{vmatrix}} = ab - ab$

Inductive step $k \times k \Rightarrow (k+1) \times (k+1)$

3×3: $\overset{\mathbf{0}}{\begin{vmatrix} a & b & c \\ * & * & * \\ a & b & c \end{vmatrix}} = -\,*\overset{\mathbf{0}}{\begin{vmatrix} b & c \\ b & c \end{vmatrix}} + *\overset{\mathbf{0}}{\begin{vmatrix} a & c \\ a & c \end{vmatrix}} - *\overset{\mathbf{0}}{\begin{vmatrix} a & b \\ a & b \end{vmatrix}}$

cofactor expansion along row 2

Discussion 2.5.14 & Theorem 2.5.15

$$\boldsymbol{A} \xrightarrow{E.R.O.} \boldsymbol{B}$$

What is the relation between det($\boldsymbol{A}$) and det($\boldsymbol{B}$)?

E.R.O	Determinant
$\boldsymbol{A} \xrightarrow{kR_i} \boldsymbol{B}$	$\det(\boldsymbol{B}) = k\det(\boldsymbol{A})$
$\boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{B}$	$\det(\boldsymbol{B}) = -\det(\boldsymbol{A})$
$\boldsymbol{A} \xrightarrow{R_i + kR_j} \boldsymbol{B}$	$\det(\boldsymbol{B}) = \det(\boldsymbol{A})$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 0 & 3 \\ 0 & 1 & 1 \end{pmatrix}$$

$$\mathbf{A} = \begin{pmatrix} 3 & 3 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix}$$

Using e.r.o. to find determinants

Example 2.5.17.1

$$\begin{vmatrix} 3 & -1 & 1 & 1 \\ 3 & -1 & 2 & 1 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{vmatrix} \overset{R_2 - R_1}{=} \begin{vmatrix} 3 & -1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{vmatrix}$$

$$\overset{R_2 \leftrightarrow R_3}{=} - \begin{vmatrix} 3 & -1 & 1 & 1 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & -1 \end{vmatrix} \overset{R_4 - 2R_3}{=} - \begin{vmatrix} 3 & -1 & 1 & 1 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}$$

$$= -3 \times 2 \times 1 \times (-1) = 6$$

Gaussian Elimination

Using e.r.o. to find determinants

Example 2.5.17.2

$$\boldsymbol{A} \xrightarrow{R_1 + \frac{2}{9}R_2} \boldsymbol{C} \xrightarrow{R_2 \leftrightarrow R_3} \boldsymbol{D} \xrightarrow{4R_2} \boldsymbol{B} = \begin{pmatrix} 5 & 0 & 8 & -1 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix}$$

det($\boldsymbol{B}$) = 4 det($\boldsymbol{D}$)

Find det($\boldsymbol{A}$).

$$\det(\boldsymbol{B}) = 5 \times (-2) \times 1 \times \frac{1}{3} = -\frac{10}{3}$$

$$\det(\boldsymbol{A}) = \det(\boldsymbol{C}) = -\det(\boldsymbol{D}) = -\boxed{}\det(\mathbf{B}) = \frac{5}{6}$$

Proof of part 3

To prove: $\det(\boldsymbol{B}) = \det(\boldsymbol{A})$

Theorem 2.5.15

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \xrightarrow{R_2 + kR_1} \mathbf{B} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} + ka_{11} & a_{22} + ka_{12} & a_{23} + ka_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

the $(2, j)$-cofactor of $\boldsymbol{A}$ = the $(2, j)$-cofactor of $\boldsymbol{B}$

$$A_{21} = B_{21} \qquad A_{22} = B_{22} \qquad A_{23} = B_{23}$$

Cofactor expansion along row 2 of $\boldsymbol{B}$:

$$\begin{aligned} \det(\mathbf{B}) &= (a_{21} + ka_{11})B_{21} + (a_{22} + ka_{12})B_{22} + (a_{23} + ka_{13})B_{23} \\ &= (a_{21} + ka_{11})A_{21} + (a_{22} + ka_{12})A_{22} + (a_{23} + ka_{13})A_{23} \\ &= \underbrace{(a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23})}_{= \det(\mathbf{A})} + k\underbrace{(a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23})}_{\neq \det(\mathbf{A})} \end{aligned}$$

Proof of part 3

To prove: det(***B***) = det(***A***)

Theorem 2.5.15

shown = 0

entries on row 1

$$\boxed{\det(\mathbf{B}) = \det(\mathbf{A})} + k\boxed{(a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23})}$$

cofactors for row 2

replace row 2 of ***A*** by row 1

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

cofactor expansion (***A***′) along row 2

$$\det(\mathbf{A'}) = a_{11}A'_{21} + a_{12}A'_{22} + a_{13}A'_{23}$$

$$= a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23}$$

$$\det(\mathbf{A'}) = 0$$

$A_{21} = A'_{21}$

$A_{22} = A'_{22}$

$A_{23} = A'_{23}$

$$\mathbf{A'} = \left.\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\right\}$$

two identical rows

In terms of elementary matrices

Theorem 2.5.15 (part 4)

A : nxn square matrix

	det(***E***)	e.r.o.	Determinant
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}$	k	$\boldsymbol{A} \xrightarrow{kR_i} \boldsymbol{B}$	$\det(\boldsymbol{B}) = k\det(\boldsymbol{A})$ $= \det(\boldsymbol{E})\det(\boldsymbol{A})$
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$	-1	$\boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{B}$	$\det(\boldsymbol{B}) = -\det(\boldsymbol{A})$ $= \det(\boldsymbol{E})\det(\boldsymbol{A})$
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{pmatrix}$	1	$\boldsymbol{A} \xrightarrow{R_i + kR_j} \boldsymbol{B}$	$\det(\boldsymbol{B}) = \det(\boldsymbol{A})$ $= \det(\boldsymbol{E})\det(\boldsymbol{A})$

E : nxn elementary matrix

$$\boldsymbol{EA} = \boldsymbol{B} \Rightarrow \boxed{\det(\boldsymbol{EA})} = \det(\boldsymbol{B}) = \boxed{\det(\boldsymbol{E})\det(\boldsymbol{A})}$$

Remark

For any square matrix $\boldsymbol{A}$ and elementary matrix $\boldsymbol{E}$:

$$\det(\boldsymbol{EA}) = \det(\boldsymbol{E})\det(\boldsymbol{A})$$

$$\begin{aligned}\det(\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}) &= \det(\boldsymbol{E}_2)\det(\boldsymbol{E}_1\boldsymbol{A}) \\ &= \det(\boldsymbol{E}_2)\det(\boldsymbol{E}_1)\det(\boldsymbol{A})\end{aligned}$$

$$\det(\boldsymbol{E}_k\cdots\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}) = \det(\boldsymbol{E}_k)\cdots\det(\boldsymbol{E}_2)\det(\boldsymbol{E}_1)\det(\boldsymbol{A})$$

In particular

$$\det(\boldsymbol{E}_k\cdots\boldsymbol{E}_2\boldsymbol{E}_1) = \det(\boldsymbol{E}_k)\cdots\det(\boldsymbol{E}_2)\det(\boldsymbol{E}_1)$$

Note: We have not yet proved that

$$\boxed{\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})}$$

Column operations

Remark 2.5.18

Since $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^T)$ for any square matrix $\boldsymbol{A}$, Theorem 2.5.15 is still true if we change "rows" to "columns".

- We have 3 corresponding elementary column operations.
- Column operations have same effect as post-multiplying an elementary matrix.

$$\boldsymbol{A} \xrightarrow{\substack{C:\ column \\ operation}} \boldsymbol{B} \qquad\qquad \boldsymbol{B} = \boldsymbol{A}\boldsymbol{E}$$

Column operations

Remark 2.5.18 = (Theorem 2.5.15)^T

A : nxn square matrix ***E*** : nxn elementary matrix

Elementary column operation	Determinant
$\boldsymbol{A} \xrightarrow{kC_i} \boldsymbol{B}$	det(***B***) = *k* det(***A***)
$\boldsymbol{A} \xrightarrow{C_i \leftrightarrow C_j} \boldsymbol{B}$	det(***B***) = −det(***A***)
$\boldsymbol{A} \xrightarrow{C_i + kC_j} \boldsymbol{B}$	det(***B***) = det(***A***)

det(***AE***) = det(***A***)det(***E***)

Determinant and invertibility

Theorem 2.5.19

A square matrix $\boldsymbol{A}$ is invertible

if and only if

$\det(\boldsymbol{A}) \neq 0$.

contrapositive

converse

$\boldsymbol{A}$ is invertible $\Rightarrow \det(\boldsymbol{A}) \neq 0$

$\boldsymbol{A}$ is invertible $\Leftarrow \det(\boldsymbol{A}) \neq 0$

$\boldsymbol{A}$ is not invertible $\Rightarrow \det(\boldsymbol{A}) = 0$

$\boldsymbol{A}$ is not invertible $\Leftarrow \det(\boldsymbol{A}) = 0$

same meaning

different meaning

different meaning

The proof

Theorem 2.5.19

A square matrix $\boldsymbol{A}$ is invertible
if and only if
$\det(\boldsymbol{A}) \neq 0$.

$$\mathbf{A} \xrightarrow{R_1} \quad \xrightarrow{R_2} \quad \cdots \quad \xrightarrow{R_k} \mathbf{B} \text{ (RREF)}$$

$$\boldsymbol{E_k} \cdots \boldsymbol{E_2 E_1 A} = \boldsymbol{B}$$

$$\det(\boldsymbol{E_k} \cdots \boldsymbol{E_2 E_1 A}) = \det(\boldsymbol{B})$$

$$\underbrace{\det(\boldsymbol{E_k}) \cdots \det(\boldsymbol{E_2})\det(\boldsymbol{E_1})}_{\text{non-zero}}\det(\boldsymbol{A}) = \det(\boldsymbol{B})$$

$\boldsymbol{A}$ invertible $\Rightarrow$ RREF $\boldsymbol{B} = \boldsymbol{I} \Rightarrow \det(\boldsymbol{B}) = 1 \Rightarrow \det(\boldsymbol{A}) \neq 0$

$\boldsymbol{A}$ not invertible $\Rightarrow$ RREF $\boldsymbol{B}$ has zero row

$\Rightarrow \det(\boldsymbol{B}) = 0 \Rightarrow \det(\boldsymbol{A}) = 0$

Using determinant to check invertibility

Remark 2.5.21

check invertibility

Theorem 2.5.19

Remark 2.4.10

determinant

row echelon form

- When the determinant is easy to get
- Connecting concepts

In practice

What is the actual value of determinant for?

Theorem 2.5.22

A and ***B*** : square matrices of order n
c a scalar

$\det(c\boldsymbol{A}) \neq c \det(\boldsymbol{A}$

1. $\det(c\boldsymbol{A}) = c^n \det(\boldsymbol{A})$

2. $\det(\boldsymbol{A}\boldsymbol{B}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$ Multiplicative property

3. if ***A*** is invertible, then

$$\det(\boldsymbol{A}^{-1}) = \frac{1}{\det(\boldsymbol{A})}$$

4. $\det(\boldsymbol{A} + \boldsymbol{B}) \neq \det(\boldsymbol{A}) + \det(\boldsymbol{B})$

Proof of part 2

$\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$

$\det(\boldsymbol{EB}) = \det(\boldsymbol{E})\det(\boldsymbol{B})$

Theorem 2.5.22

Case 1: $\boldsymbol{A}$ is singular

By Theorem 2.4.14

$\boldsymbol{AB}$ is singular

$\Rightarrow \det(\boldsymbol{A}) = 0$

$\Rightarrow \det(\boldsymbol{A})\det(\boldsymbol{B}) = 0$

$\Rightarrow \det(\boldsymbol{AB}) = 0$

Case 2: $\boldsymbol{A}$ is invertible

Theorem 2.4.7: (1) implies (4)

$\Rightarrow \boldsymbol{A} = \boldsymbol{E}_1\boldsymbol{E}_2\cdots\boldsymbol{E}_k$ (product of elementary matrices)

$$\begin{aligned}\det(\boldsymbol{AB}) &= \det(\boldsymbol{E}_1\boldsymbol{E}_2\cdots\boldsymbol{E}_k\boldsymbol{B}) \\ &= \det(\boldsymbol{E}_1)\det(\boldsymbol{E}_2)\cdots\det(\boldsymbol{E}_k)\det(\boldsymbol{B}) \\ &= \det(\boldsymbol{E}_1\boldsymbol{E}_2\cdots\boldsymbol{E}_k)\det(\boldsymbol{B}) \\ &= \det(\boldsymbol{A})\det(\boldsymbol{B})\end{aligned}$$

What is adjoint?

Definition 2.5.24

Let $\boldsymbol{A}$ be a square matrix of order n.

The adjoint of $\boldsymbol{A}$ is the $n \times n$ matrix

$$\operatorname{adj}(\boldsymbol{A}) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}^{T} = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1\mathrm{n}} & A_{2\mathrm{n}} & \cdots & A_{nn} \end{pmatrix}$$

where A_{ij} is the (i, j)-cofactor of $\boldsymbol{A}$.

$(-1)^{i+j} \det(\boldsymbol{M_{ij}})$

What is adjoint?

Example 2.5.26.2

$$\boldsymbol{B} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

$$\operatorname{adj}(\boldsymbol{B}) = \begin{pmatrix} \begin{vmatrix} -1 & 0 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 0 & 0 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} -1 & 1 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \end{pmatrix}^T = \begin{pmatrix} -3 & 0 & 1 \\ 3 & 2 & -1 \\ 1 & 0 & -1 \end{pmatrix}^T = \begin{pmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{pmatrix}$$

What is adjoint for?

Theorem 2.5.25

Let $\boldsymbol{A}$ be a square matrix.

If $\boldsymbol{A}$ is invertible, then

$$\boxed{\boldsymbol{A}^{-1} = \frac{1}{\det(\boldsymbol{A})}\operatorname{adj}(\boldsymbol{A})}$$

Example 2.5.26.2

$$\boldsymbol{B} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{pmatrix} \qquad \det(\boldsymbol{B}) = -2 \qquad \operatorname{adj}(\mathbf{B}) = \begin{pmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{pmatrix}$$

$$\boldsymbol{B}^{-1} = \frac{1}{\det(\boldsymbol{B})}\operatorname{adj}(\boldsymbol{B}) = -\frac{1}{2}\begin{pmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{pmatrix}$$

The proof

Theorem 2.5.25

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \qquad \text{adj}(\boldsymbol{A}) = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

the (i, i)-entry of $\boldsymbol{A}[\text{adj}(\boldsymbol{A})]$ diagonal entries

$$= a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in} = \det(\boldsymbol{A})$$

cofactor expansion along row i

the (i, j)-entry of $\boldsymbol{A}[\text{adj}(\boldsymbol{A})]$ with $i \neq j$,

non-diagonal entries

$$= a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn}$$

$$= 0$$

see the proof of Theorem 2.5.15.3.

The proof

Theorem 2.5.25

$$\boldsymbol{A}\,\mathrm{adj}(\boldsymbol{A}) = \begin{pmatrix} \det(\boldsymbol{A}) & 0 & \ldots & 0 \\ 0 & \det(\boldsymbol{A}) & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \det(\boldsymbol{A}) \end{pmatrix} = \det(\boldsymbol{A}) \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

$$= \det(\boldsymbol{A})\,\boldsymbol{I}$$

$$\Rightarrow \quad \boldsymbol{A}\left[\boxed{\frac{1}{\det(\boldsymbol{A})}\mathrm{adj}(\boldsymbol{A})}\right] = \boldsymbol{I}$$

$$\Rightarrow \mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\mathrm{adj}(\mathbf{A})$$

Remark

find inverse

Theorem 2.5.25

Discussion 2.4.8

adjoint

reduced row echelon form

for smaller size

for larger size

give explicit formula for inverse

What is Cramer's rule?

Example 2.5.28 (Cramer's rule)

Use Cramer's rule to solve the system of linear equations

$$\begin{cases} x + y + 3z = 0 \\ 2x - 2y + 2z = 4 \\ 3x + 9y = 3 \end{cases}$$

Rewrite the linear system as $\begin{pmatrix} 1 & 1 & 3 \\ 2 & -2 & 2 \\ 3 & 9 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 3 \end{pmatrix}$

$$\boldsymbol{A}_1 = \begin{pmatrix} 0 & 1 & 3 \\ 4 & -2 & 2 \\ 3 & 9 & 0 \end{pmatrix} \quad \boldsymbol{A}_2 = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 4 & 2 \\ 3 & 3 & 0 \end{pmatrix} \quad \boldsymbol{A}_3 = \begin{pmatrix} 1 & 1 & 0 \\ 2 & -2 & 4 \\ 3 & 9 & 3 \end{pmatrix}$$

What is Cramer's rule?

Example 2.5.28

$$\begin{cases} x + y + 3z = 0 \\ 2x - 2y + 2z = 4 \\ 3x + 9y = 3 \end{cases}$$

Cramer's rule says:

$$x = \frac{\det(A_1)}{\det(A)} = \frac{\begin{vmatrix} 0 & 1 & 3 \\ 4 & -2 & 2 \\ 3 & 9 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 3 \\ 2 & -2 & 2 \\ 3 & 9 & 0 \end{vmatrix}} = \frac{132}{60} = 2.2$$

$$y = \frac{\det(A_2)}{\det(A)} = \frac{\begin{vmatrix} 1 & 0 & 3 \\ 2 & 4 & 2 \\ 3 & 3 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 3 \\ 2 & -2 & 2 \\ 3 & 9 & 0 \end{vmatrix}} = \frac{-24}{60} = -0.4$$

$$z = \frac{\det(A_3)}{\det(A)} = \frac{\begin{vmatrix} 1 & 1 & 0 \\ 2 & -2 & 4 \\ 3 & 9 & 3 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 3 \\ 2 & -2 & 2 \\ 3 & 9 & 0 \end{vmatrix}} = \frac{-36}{60} = -0.6$$

this gives the unique solution of the system

What is Cramer's rule?

Theorem 2.5.27 (Cramer's Rule)

Suppose $\boldsymbol{Ax} = \boldsymbol{b}$ is a linear system where $\boldsymbol{A}$ is an $n \times n$ invertible matrix.

terms and conditions

Let $\boldsymbol{A_i}$ be the matrix obtained from $\boldsymbol{A}$ by replacing the ith column of $\boldsymbol{A}$ by $\boldsymbol{b}$.

Then the system has a unique solution

$$\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})}\begin{pmatrix} \det(\boldsymbol{A}_1) \\ \det(\boldsymbol{A}_2) \\ \vdots \\ \det(\boldsymbol{A}_n) \end{pmatrix}$$

$$x_1 = \frac{\det(\boldsymbol{A_1})}{\det(\boldsymbol{A})}$$

$$x_2 = \frac{\det(\boldsymbol{A_2})}{\det(\boldsymbol{A})}$$

$$x_n = \frac{\det(\boldsymbol{A}_n)}{\det(\boldsymbol{A})}$$

The proof

Proof of Theorem 2.5.27

To prove:

$$x_i = \frac{\det(\boldsymbol{A}_i)}{\det(\boldsymbol{A})}$$

$$\boldsymbol{Ax} = \boldsymbol{b} \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} \quad = \frac{1}{\det(\boldsymbol{A})}\boxed{\operatorname{adj}(\boldsymbol{A})}\boldsymbol{b}$$

$$\Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det(\boldsymbol{A})} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

To show

$$x_1 = \frac{1}{\det(\mathbf{A})}(b_1A_{11} + b_2A_{21} + \dots + b_nA_{n1})$$

$$x_i = \frac{1}{\det(\mathbf{A})}(b_1A_{1i} + b_2A_{2i} + \dots + b_nA_{ni})$$

for $i = 1, 2, \dots, n$

The proof

To prove:

$$x_i = \frac{\det(\boldsymbol{A}_i)}{\det(\boldsymbol{A})}$$

Proof of Theorem 2.5.27

entries on column i

$$x_i = \frac{1}{\det(\mathbf{A})}\underbrace{\boxed{(b_1 A_{1i} + b_2 A_{2i} + \ldots + b_n A_{ni}}}_{\det(\boldsymbol{A}_i)}) = \frac{\det(\boldsymbol{A}_i)}{\det(\boldsymbol{A})}$$

cofactors for column i

cofactor expansion of $\boldsymbol{A}_i$ along column i

$$\det(\boldsymbol{A}_i) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,i-1} & b_2 & a_{2,i+1} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{vmatrix}$$

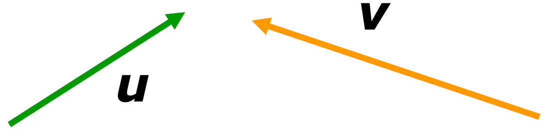
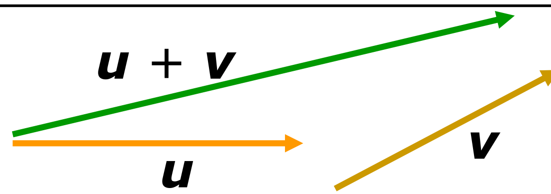
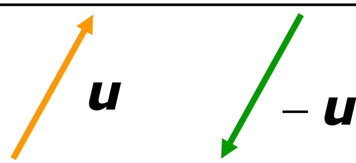
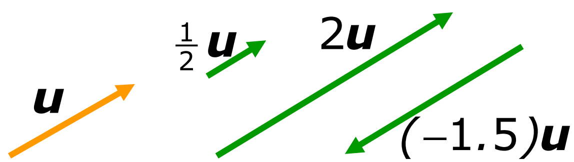
Section 3.1

Euclidean n-Spaces

Objectives

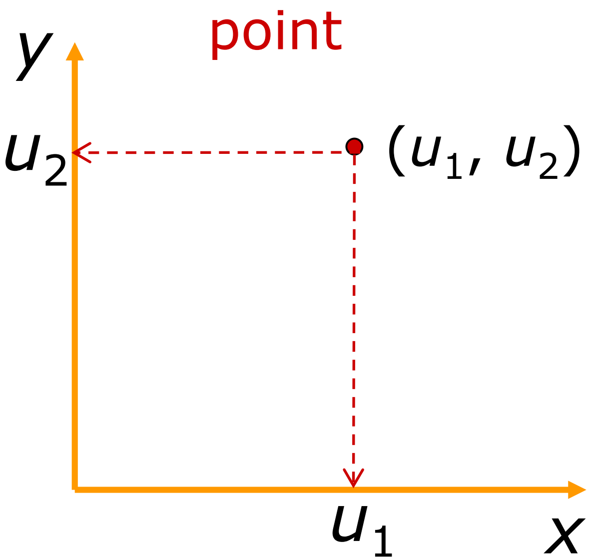
- What is an n-vector?
- What are some operations on n-vectors?
- What is a Euclidean n-space $\mathbf{R}^n$?
- How to express subsets of $\mathbf{R}^n$?

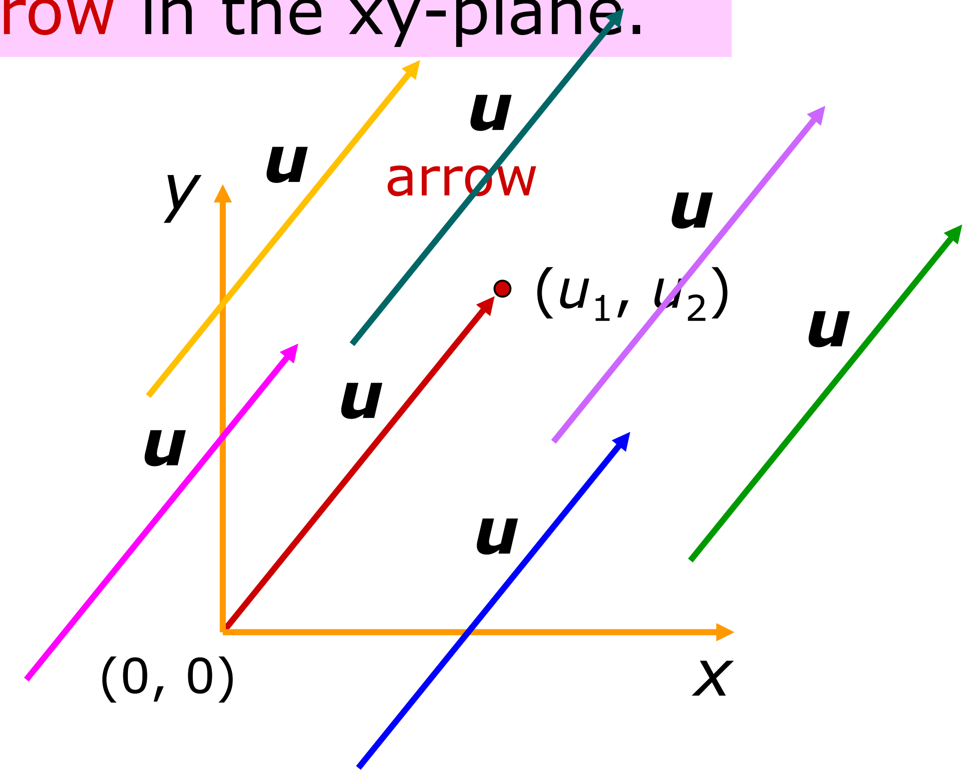
Discussion 3.1.1 - 3.1.2 (Vectors)

Notation	geometric	algebraic (2-dimension) (3-dimension)
Vector ***u*** Vector ***v***		(u_1, u_2) (v_1, v_2) (u_1, u_2, u_3) (v_1, v_2, v_3)
Addition ***u***+***v***		(u_1+v_1, u_2+v_2) $(u_1+v_1, u_2+v_2, u_3+v_3)$
Negative -***u***		$(-u_1, -u_2)$ $(-u_1, -u_2, -u_3)$
Scalar multiple a***u***		(au_1, au_2) (au_1, au_2, au_3)

Discussion 3.1.2.1

Geometrically, (u_1, u_2) can represent either a point and an arrow in the xy-plane.



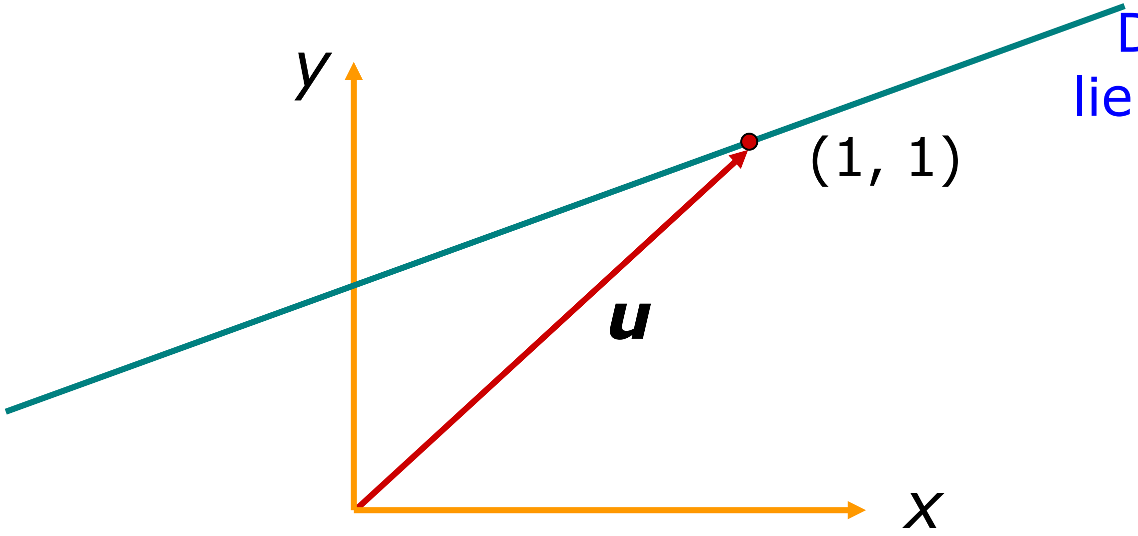


Similarly for (u_1, u_2, u_3) in the xyz-space.

Point or Arrow?

Geometrically, (u_1, u_2) can represent either a point and an arrow in the xy-plane.

Linear equation: $2y - x = 1$ A solution: $x = 1, y = 1$



The point (1, 1) lies on the line,
but the arrow (1, 1) does not lie on the line

What is an n-vector?

(u_1, u_2) 2-vector

(u_1, u_2, u_3) 3-vector

Definition 3.1.3

n-vector $(u_1, u_2, \ldots, u_i, \ldots, u_n) \neq \{u_1, u_2, \ldots, u_i, \ldots, u_n\}$

where $u_1, u_2, \ldots, u_n$ are real numbers

i^{th} component (or i^{th} coordinate) of the n-vector

Always think/view an n-vector
as a SINGLE object
and not n numbers

n-vectors as matrices

Notation 3.1.5

We can identify an n-vector $(u_1, u_2, \ldots, u_n)$ with a $1 \times n$ matrix $(u_1\ u_2\ \ldots\ u_n)$ (row vector)

or an $n \times 1$ matrix $\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ (column vector).

Which form to use depends on the context.

Definitions and properties of vector operations are similar to matrix operations (see 3.1.3 to 3.1.6)

What is a Euclidean n-space ?

Definition 3.1.7

The set of all n-vectors of real numbers is called the Euclidean n-space and is denoted by $\mathbf{R}^n$.

$\mathbf{R}^4$

(1,2,3,1)

(0,2,1,5)

(-1,2,3,19)

:

$\boldsymbol{u} \in \mathbf{R}^n \longleftrightarrow \boldsymbol{u}$ is an n-vector $\longleftrightarrow \boldsymbol{u} = (u_1, u_2, \ldots, u_n)$

Euclidean 2-space $\mathbf{R}^2$

all the 2-vectors (as points) in xy-plane

Euclidean 3-space $\mathbf{R}^3$

all the 3-vectors (as points) in xyz-space

How to express subsets of $\mathbf{R}^n$?

Example 3.1.8.1

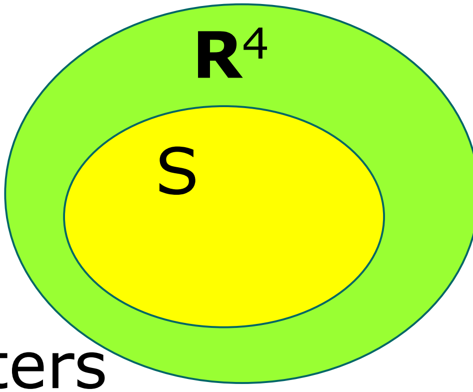
Set notation

$$S = \{ \underbrace{(u_1, u_2, u_3, u_4)}_{\text{type of elements is 4-vector}} \mid \underbrace{u_1 = 0 \ \text{ and } \ u_2 = u_4}_{\text{conditions on the components}} \}$$

implicit form

$(0, 0, 0, 0), (0, 1, 5, 1), (0, \pi, -3, \pi) \in S$

$(0, 2, 2, 3), (1, 1, 1, 1) \notin S$



general form $(0, a, b, a)$ $\quad a, b$: parameters

$$S = \{ (0, a, b, a) \mid a, b \in \mathbf{R} \}$$

explicit form

How to express subsets of $\mathbf{R}^n$?

Set notation for subsets of $\mathbf{R}^n$

Implicit form

$$\left\{ \text{general n-tuple } (u_1, u_2, \ldots, u_n) \;\middle|\; \text{conditions satisfied by } u_1, u_2, \ldots, u_n \right\}$$

$$S = \{ (u_1, u_2, u_3, u_4) \mid u_1 = 0 \text{ and } u_2 = u_4 \}$$

Explicit form Not always possible to express in explicit form

$$\left\{ \text{n-tuples with explicit form} \;\middle|\; \text{range of parameters appearing on the left} \right\}$$

$$S = \{ (0, a, b, a) \mid a, b \in \mathbf{R} \}$$

Don't write $\{a, b \in \mathbf{R} \mid (0, a, b, a) \}$

Solution set as a subset of $\mathbf{R}^n$

Example 3.1.8.2

$$\begin{cases} x + y + z = 0 \\ x - y + 2z = 1 \end{cases}$$

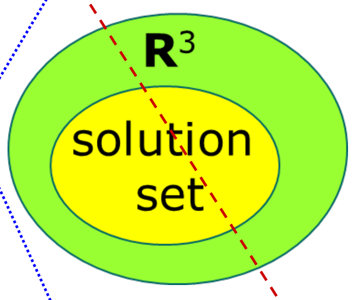
$x = 2, y = -1, z = -1$

$(2, -1, -1)$

a 3-vector

general solution: $\begin{cases} x = 0.5 - 1.5t \\ y = -0.5 + 0.5t \\ z = t \end{cases}$

solution set: subset of $\mathbf{R}^3$



Explicit form

$\{ (0.5 - 1.5t, -0.5 + 0.5t, t) \mid t \in \mathbf{R} \}$

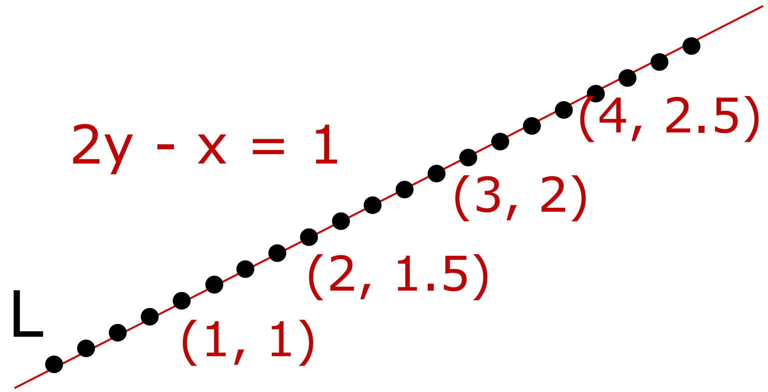
Implicit form

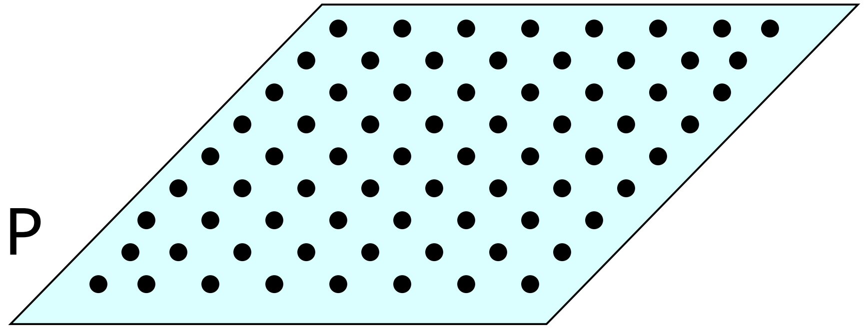
$\{ (x, y, z) \mid x + y + z = 0$ and $x - y + 2z = 1 \}$

Example 3.1.8.3

A line/plane in the xy-plane/ xyz-space can be regarded as a collection of points.

a collection of vectors





L is a subset of **R**^{2}

P is a subset of **R**^{3}

Set notations of lines and planes

Example 3.1.8.3

Lines in xy-plane

Implicit form: $\{ (x, y) \mid ax + by = c \}$

Explicit form: $\left\{ \left(\dfrac{c - bt}{a}, \ t \right) \mid t \in \mathbf{R} \right\}$

Planes in xyz-space

Implicit form: $\{ (x, y, z) \mid ax + by + cz = d \}$

Explicit form: $\left\{ \left(\dfrac{d - bs - ct}{a}, \ s, \ t \right) \mid s, t \in \mathbf{R} \right\}$

Lines in xyz-space

Implicit form: $\{ (x, y, z) \mid$ eqn of the line $\}$**?**

Explicit form: { (general solution) | 1 parameter }**?**

Line as a subset of $\mathbf{R}^3$

Example 3.1.8.2 (revisited)

$$\begin{cases} x + y + z = 0 \\ x - y + 2z = 1 \end{cases}$$

two planes

solution set (explicit form)

$$\{\ (0.5 - 1.5t,\ -0.5 + 0.5t,\ t)\ |\ t \in \mathbf{R}\ \}$$

This represents a line in the xyz-space

$$(0.5, -0.5, 0) + (-1.5t, 0.5t, t)$$

$$(0.5, -0.5, 0) + t(-1.5, 0.5, 1)$$

a point on the line

the direction of the line

Line as a subset of $\mathbf{R}^3$

Example 3.1.8.3(c)

Set notation (explicit)

$\{\ (a_0 + at,\ b_0 + bt,\ c_0 + ct)\ |\ t \text{ in } \mathbf{R}\ \}$ t: parameter

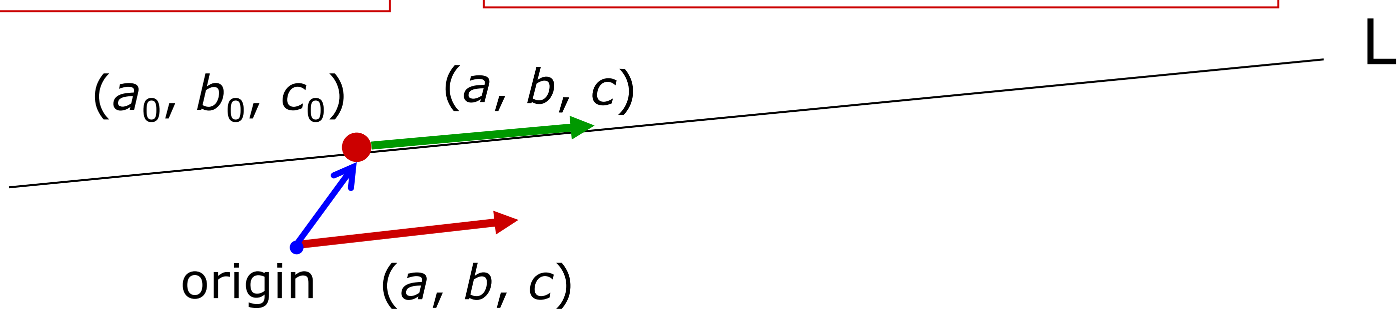
a_0, b_0, c_0, a, b, c are fixed real numbers

a, b, c not all zero

$\{\ (a_0, b_0, c_0) + t(a, b, c)\ |\ t \text{ in } \mathbf{R}\ \}$

a point on the line

the direction of the line



Line as a subset of $\mathbf{R}^3$

Example 3.1.8.3(c)

Set notation (Implicit) $\{ (x, y, z) \mid \text{eqn of the line} \}$**?**

A line in $\mathbf{R}^3$ cannot be represented by a single linear equation.

But it can be regarded as the intersection of two planes P_1 and P_2.

Suppose the equations of the two planes are given by

P_1: $a_1x + b_1y + c_1z = d_1$ $\quad$ P_2: $a_2x + b_2y + c_2z = d_2$

Set notation (implicit)

$\{ (x, y, z) \mid a_1x + b_1y + c_1z = d_1 \text{ and } a_2x + b_2y + c_2z = d_2 \}$

refer to 3.1.8.2

Number of elements in a set

Notation 3.1.9 & Example 3.1.10

For a finite set S, we denote the number of elements of S by $|S|$

$S_1 = \{ 1, 2, 3, 4 \}$ $\quad |S_1| = 4$

$S_2 = \{ (1, 2, 3, 4) \}$ $\quad |S_2| = 1$

$S_3 = \{ (1,2,3), (2,3,4) \}$ $\quad |S_3| = 2$

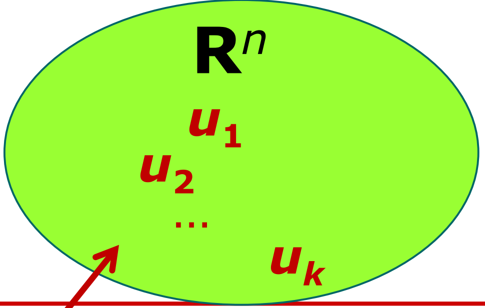
Section 3.2

Linear Combinations and Linear Spans

Objective

- What is a linear combination?
- How to express a vector as a linear combination?
- What is a linear span?

What is a linear combination?



Definition 3.2.1

$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$: a fixed set of vectors in $\mathbf{R}^n$

$c_1, c_2, \ldots, c_k$: real numbers

$$\boxed{c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k}$$

is called a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$.

Example $\boldsymbol{u}_1 = (2, 1, 0)$ $\boldsymbol{u}_2 = (-3, 0, 1)$

$c_1 = 1,\ c_2 = 1$

$1(2, 1, 0) + 1(-3, 0, 1) = (-1, 1, 1)$

a specific linear combination

$c_1 = s,\ c_2 = t$

$s(2, 1, 0) + t(-3, 0, 1)$ general linear combination with parameters s and t

Can every vector be expressed as a linear combination of a given set of vectors?

Example 3.2.2.1

$\boldsymbol{u_1} = (2, 1, 3)$, $\boldsymbol{u_2} = (1, -1, 2)$ and $\boldsymbol{u_3} = (3, 0, 5)$.

(a) $\boldsymbol{v} = (3, 3, 4)$

is a linear combination of $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$.

(3, 3, 4) can be expressed as

$a(2, 1, 3) + b(1, -1, 2) + c(3, 0, 5)$

(b) $\boldsymbol{w} = (1, 2, 4)$

is not a linear combination of $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$.

(1, 2, 4) cannot be expressed as

$a(2, 1, 3) + b(1, -1, 2) + c(3, 0, 5)$

How to express a vector as a specific linear combination of a given set of vectors?

Example 3.2.2.1(a)

$\boldsymbol{u}_1 = (2, 1, 3)$ $\boldsymbol{u}_2 = (1, -1, 2)$ $\boldsymbol{u}_3 = (3, 0, 5)$

Write $\boldsymbol{v} = a\boldsymbol{u}_1 + b\boldsymbol{u}_2 + c\boldsymbol{u}_3$

$$\boxed{(3, 3, 4) = a(2, 1, 3) + b(1, -1, 2) + c(3, 0, 5)}$$

solve for a, b, c

Equating components

$$\boxed{\begin{cases} 2a + b + 3c = 3 \\ a - b = 3 \\ 3a + 2b + 5c = 4 \end{cases}}$$

1st component

2nd component

3rd component

So we obtain a linear system in variables a, b, c

$$\boxed{\begin{pmatrix} 3 \\ 3 \\ 4 \end{pmatrix} = a\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + b\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + c\begin{pmatrix} 3 \\ 0 \\ 5 \end{pmatrix}}$$

vector equation form of the linear system (P.43)

How to express a vector as a specific linear combination of a given set of vectors?

Example 3.2.2.1(a)

$\boldsymbol{u_1} = (2, 1, 3)$ $\boldsymbol{u_2} = (1, -1, 2)$ $\boldsymbol{u_3} = (3, 0, 5)$

$$\boxed{(3, 3, 4) = a(2, 1, 3) + b(1, -1, 2) + c(3, 0, 5)}$$

can find specific values for a, b, c

$$\left(\begin{array}{ccc|c} 2 & 1 & 3 & 3 \\ 1 & -1 & 0 & 3 \\ 3 & 2 & 5 & 4 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gaussian}} \left(\begin{array}{ccc|c} 2 & 1 & 3 & 3 \\ 0 & -\frac{3}{2} & -\frac{3}{2} & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{array}\right)$$

system is consistent

So (3, 3, 4) is a linear combination of $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$.

To write (3, 3, 4) as a specific linear combination:

general solution of LS : $a = 2 - t,\ b = -1 - t,\ c = t$

Take t = 0: a = 2, b = -1, c = 0

$$(3, 3, 4) = 2\boldsymbol{u_1} - \boldsymbol{u_2} + 0\boldsymbol{u_3}$$

Take t = 1: a = 1, b = -2, c = 1

$$(3, 3, 4) = \boldsymbol{u_1} - 2\boldsymbol{u_2} + \boldsymbol{u_3}$$

How to show that a vector cannot be expressed as a linear combination of a given set of vectors?

Example 3.2.2.1(b) $\boldsymbol{u}_1 = (2, 1, 3)$ $\boldsymbol{u}_2 = (1, -1, 2)$ $\boldsymbol{u}_3 = (3, 0, 5)$

Write $\boldsymbol{w} = a\boldsymbol{u}_1 + b\boldsymbol{u}_2 + c\boldsymbol{u}_3$

$$(1, 2, 4) = a(2, 1, 3) + b(1, -1, 2) + c(3, 0, 5)$$

$$\begin{aligned} 2a + b + 3c &= 1 \\ a - b &= 2 \\ 3a + 2b + 5c &= 4 \end{aligned}$$

$$\left(\begin{array}{ccc|c} 2 & 1 & 3 & 1 \\ 1 & -1 & 0 & 2 \\ 3 & 2 & 5 & 4 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gaussian}} \left(\begin{array}{ccc|c} 2 & 1 & 3 & 1 \\ 0 & -\frac{3}{2} & -\frac{3}{2} & \frac{3}{2} \\ 0 & 0 & 0 & 3 \end{array}\right)$$

system is inconsistent

(1, 2, 4) is not a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$.

How to express a general vector as a linear combination of a given set of vectors?

Example 3.2.2.2

standard basis vectors

Directional vectors of the x-axis, y-axis, z-axis

Every vector in $\mathbf{R}^3$

is a linear combination of the following vectors

$$\boxed{\mathbf{e}_1 = (1, 0, 0),\ \mathbf{e}_2 = (0, 1, 0),\ \mathbf{e}_3 = (0, 0, 1)}$$

Take a general 3-vector (x, y, z)

$$(x, y, z) = (x, 0, 0) + (0, y, 0) + (0, 0, z)$$

$$= x\,(1, 0, 0) + y\,(0, 1, 0) + z\,(0, 0, 1)$$

$$= x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$$

e.g. $(1, 2, 5) = 1\mathbf{e}_1 + 2\mathbf{e}_2 + 5\mathbf{e}_3$

Span preview

How many linear combinations of (2,1,0) and (-3,0,1) are there?

Infinite

The set of all linear combinations of (2,1,0) and (-3,0,1)

$\{s(2, 1, 0) + t(-3, 0, 1) \mid s, t \in \mathbf{R}\}$

using set notation

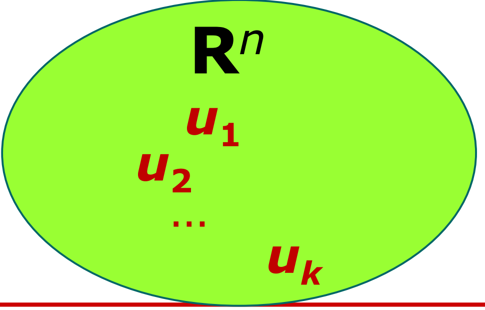
We call it: the linear span of (2,1,0) and (-3,0,1)

using words (in terms of linear span)

We write it: span{(2,1,0) , (-3,0,1)}

using linear span notation

What is a linear span?



Definition 3.2.3

$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$: k (finite) vectors in $\mathbf{R}^n$.

The set of all linear combinations of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$

$$\{c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k \mid c_1, c_2, \ldots, c_k \text{ in } \mathbf{R}\}$$

=

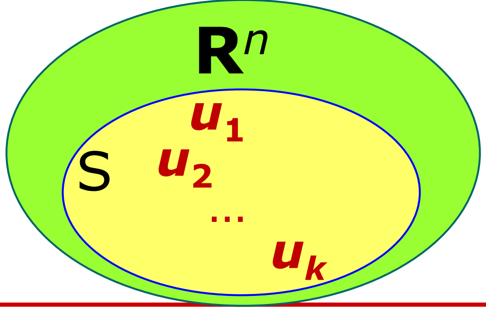
This set is called

the linear span of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$

"Linear span" is always used w.r.t. a set of vectors

This set is denoted by span$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$

What is a linear span?



Definition 3.2.3

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$: a (finite) subset of $\mathbf{R}^n$.

The $\boxed{\text{set of all linear combinations}}$ of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$

$$\{c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k \mid c_1, c_2, \ldots, c_k \text{ in } \mathbf{R}\}$$

$$= \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} = \operatorname{span}(S)$$

This set is called

the $\boxed{\text{linear span}}$ of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$

the linear span of S

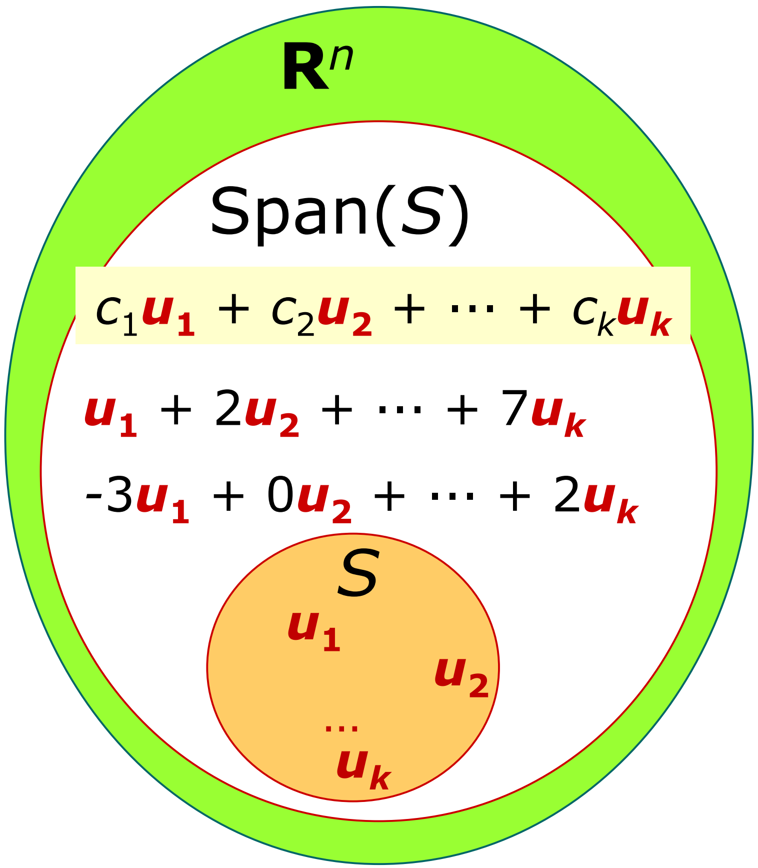
"Linear span" is always used w.r.t. a set of vectors

This set is denoted by

What is a linear span?

Definition 3.2.3

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$ a finite collection of vectors in $\mathbf{R}^n$



$\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k} \in \mathbf{R}^n$

$S \subseteq \mathbf{R}^n$

$\text{span}(S) \subseteq \mathbf{R}^n$

$S \subseteq \text{span}(S)$

span(S) can be equal to $\mathbf{R}^n$ but not always.

Example 3.2.4.1

In Example 3.2.2.1,

$\boldsymbol{u}_1 = (2, 1, 3)$, $\boldsymbol{u}_2 = (1, -1, 2)$ and $\boldsymbol{u}_3 = (3, 0, 5)$.

(a) $\boldsymbol{v} = (3, 3, 4)$ (b) $\boldsymbol{w} = (1, 2, 4)$

$\boldsymbol{v}$ is a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$.

$$\boldsymbol{v} \in \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$$

$\boldsymbol{w}$ is not a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$.

$$\boldsymbol{w} \notin \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$$

Express a linear span in explicit set notation form

Example 3.2.4.2

$S = \{(1, 0, 0, -1), (0, 1, 1, 0)\} \subseteq \mathbf{R}^4$ $\boxed{\text{span}(S) \subseteq \mathbf{R}^4}$

$$\text{span}(S) = \text{span}\{(1, 0, 0, -1), (0, 1, 1, 0)\}$$ linear span form

$$= \{a(1, 0, 0, -1) + b(0, 1, 1, 0) \mid a, b \in \mathbf{R}\}$$

$$= \{(a, b, b, -a) \mid a, b \in \mathbf{R}\}$$ explicit form

A general vector in span(S):

$$a(1, 0, 0, -1) + b(0, 1, 1, 0) = (a, b, b, -a).$$

Section 3.2

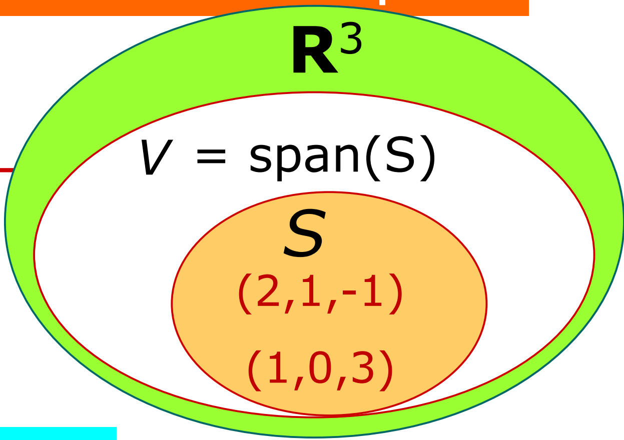
Linear Combinations and Linear Spans

Objective

- How to express a linear span in explicit set notation?
- How to express a set notation as a linear span?
- How to show a linear span is (is not) equal to $\mathbf{R}^n$?
- How to show a linear span is contained in another?

Express an explicit set notation form as linear span

Example 3.2.4.3



Let $V = \{\ (2a + b, a, 3b-a) \mid a, b \in \mathbf{R}\} \subseteq \mathbf{R}^3$.

explicit form

Rewrite the general form:

$(2a + b, a, 3b-a) = a\ (2, 1,-1) + b\ (1, 0, 3)$.

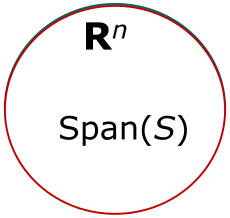
So $V = \text{span}\{(2, 1,-1), (1, 0, 3)\}$.

linear span form

The subset V is spanned by (2, 1,-1), (1, 0, 3)

(2, 1,-1), (1, 0, 3) spans the subset V.

How to show a linear span equal to $\mathbf{R}^n$?



Example 3.2.4.4

To show: span{(1, 0, 1), (1, 1, 0), (0, 1, 1)} = $\mathbf{R}^3$

Same as showing:

every vector (x, y, z) in $\mathbf{R}^3$ can be written as linear combination of the three vectors

Write $(x, y, z) = a\,(1, 0, 1) + b\,(1, 1, 0) + c\,(0, 1, 1)$

Convert into linear system

$$\begin{cases} a + b = x \\ b + c = y \\ a + c = z \end{cases}$$

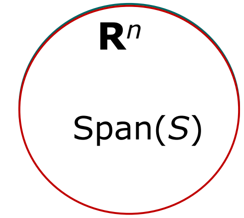
$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 0 & 1 & 1 & y \\ 1 & 0 & 1 & z \end{array}\right)$$

a, b, c are variables

x, y, z are treated as constants.

To show:
The system is consistent

How to show a linear span equal to $\mathbf{R}^n$?



Example 3.2.4.4

$(x, y, z) = a\,(1, 0, 1) + b\,(1, 1, 0) + c\,(0, 1, 1)$

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 0 & 1 & 1 & y \\ 1 & 0 & 1 & z \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gaussian}} \left(\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 0 & 1 & 1 & y \\ 0 & 0 & 2 & z - x + y \end{array}\right)$$

The system is consistent regardless of the values of x, y, z.

So we can always solve for a, b, c for any vector (x, y, z).

Every (x, y, z) in $\mathbf{R}^3$ is a linear combination of the three given vectors

So span{(1, 0, 1), (1, 1, 0), (0, 1, 1)} = $\mathbf{R}^3$

How to show a linear span equal to $\mathbf{R}^n$?

Example 3.2.4.4

Solve a, b, c in terms of x, y, z

$$(x, y, z) = a\,(1, 0, 1) + b\,(1, 1, 0) + c\,(0, 1, 1)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 0 & 1 & 1 & y \\ 0 & 0 & 2 & z-x+y \end{array}\right) \Rightarrow \begin{cases} a + b = x \\ b + c = y \\ 2c = z - x + y \end{cases}$$

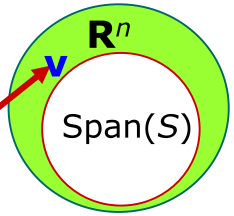
Solution: $c = \boxed{\frac{-x+y+z}{2}}$ $b = \boxed{\frac{x+y-z}{2}}$ $a = \boxed{\frac{x-y+z}{2}}$

$$(x,y,z) = \left(\frac{x-y+z}{2}\right)(1,0,1) + \left(\frac{x+y-z}{2}\right)(1,1,0) + \left(\frac{-x+y+z}{2}\right)(0,1,1)$$

e.g. $(1,2,5) = 2(1,0,1) + (-1)(1,1,0) + 3(0,1,1)$

Every (x, y, z) can be expressed as a linear combination of (1, 0, 1), (1, 1, 0) and (0, 1, 1) in exactly one way.

How to show a linear span not equal to $\mathbf{R}^n$?



Example 3.2.4.5

To show: span{(1,1,1), (1,2,0), (2,1,3), (2,3,1)} $\neq \mathbf{R}^3$

$$(x, y, z) = a\,(1, 1, 1) + b\,(1, 2, 0) + c\,(2, 1, 3) + d\,(2, 3, 1)$$

$$\left(\begin{array}{cccc|c} 1 & 1 & 2 & 2 & x \\ 1 & 2 & 1 & 3 & y \\ 1 & 0 & 3 & 1 & z \end{array}\right) \xrightarrow{\text{G.E.}} \left(\begin{array}{cccc|c} 1 & 1 & 2 & 2 & x \\ 0 & 1 & -1 & 1 & y-x \\ 0 & 0 & 0 & 0 & y+z-2x \end{array}\right)$$

The system is inconsistent when $y + z - 2x \neq 0$.

e.g. $x = 1$, $y = 0$, $z = 0$

So $(1, 0, 0) \notin$ span{(1,1,1), (1,2,0), (2,1,3), (2,3,1)}

How to determine whether a linear span is equal to $\mathbf{R}^n$ or not?

Discussion 3.2.5

$\text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} = \mathbf{R}^n$?

$\mathbf{u}_1 = (a_{11}, a_{12}, \ldots, a_{1n})$,
$\mathbf{u}_2 = (a_{21}, a_{22}, \ldots, a_{2n})$,
$\vdots$
$\mathbf{u}_k = (a_{k1}, a_{k2}, \ldots, a_{kn})$.

Consider the linear system

$$\underbrace{\left(\begin{array}{cccc|c} a_{11} & a_{21} & \cdots & a_{k1} & \\ a_{12} & a_{22} & \cdots & a_{k2} & \\ \vdots & \vdots & & \vdots & \\ a_{1n} & a_{2n} & \cdots & a_{kn} & \end{array}\right)}_{\mathbf{A}} \xrightarrow{\text{G.E.}} \underbrace{\left(\begin{array}{cccc|c} * & * & \cdots & * & \\ 0 & * & \cdots & * & \\ \vdots & \ddots & * & \vdots & \\ 0 & \cdots & 0 & * & \end{array}\right)}_{\mathbf{R}} \quad \text{REF}$$

R has no zero row
system is always consistent
$\text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} = \mathbf{R}^n$

R has a zero row
system may be inconsistent
$\text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} \neq \mathbf{R}^n$

Tutorial 3 Q43

A condition for a linear span to be not equal to $\mathbf{R}^n$

Theorem 3.2.7

Let S = $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a set of vectors in $\mathbf{R}^n$.

If $k < n$, then S cannot span $\mathbf{R}^n$. span(S) $\neq \mathbf{R}^n$

More rows than columns

$$\underset{\mathbf{A}}{\left(\begin{array}{cccc|c} a_{11} & a_{21} & \cdots & a_{k1} & x \\ a_{12} & a_{22} & \cdots & a_{k2} & y \\ \vdots & \vdots & & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{kn} & z \end{array}\right)} \xrightarrow{\text{.E.}} \underset{\mathbf{R}}{\left(\begin{array}{cccc|c} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \ddots & * & \vdots & \vdots \\ 0 & \ldots & & 0 & * \end{array}\right)} \quad \text{REF}$$

The REF **R** of **A** must have a zero row,
so the system may be inconsistent,
and span(S) $\neq \mathbf{R}^n$.

A condition for a linear span to be not equal to $\mathbf{R}^n$

Theorem 3.2.7

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a set of vectors in $\mathbf{R}^n$.

If $k < n$, then S cannot span $\mathbf{R}^n$. $\text{span}(S) \neq \mathbf{R}^n$

Example 3.2.8

$\text{span}\{\ \boldsymbol{u}\ \} \neq \mathbf{R}^2$ since $k = 1 < n = 2$

$\text{span}\{\ \boldsymbol{u}\ \} \neq \mathbf{R}^3$ since $k = 1 < n = 3$

$\text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\ \} \neq \mathbf{R}^3$ since $k = 2 < n = 3$

Every linear span contains the zero vector

Theorem 3.2.9.1

Let $S = \{ \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k \}$ ← any set

The zero vector $\mathbf{0} \in \text{span}(S)$.

Proof

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \ldots + c_k\boldsymbol{u}_k \in \text{span}(S)$$

for any $c_1, c_2, \ldots, c_k$ in $\mathbf{R}$

In particular

$$\underbrace{0\boldsymbol{u}_1 + 0\boldsymbol{u}_2 + \ldots + 0\boldsymbol{u}_k}_{\mathbf{0}} \in \text{span}(S)$$

$$\mathbf{0} \in \text{span}(S)$$

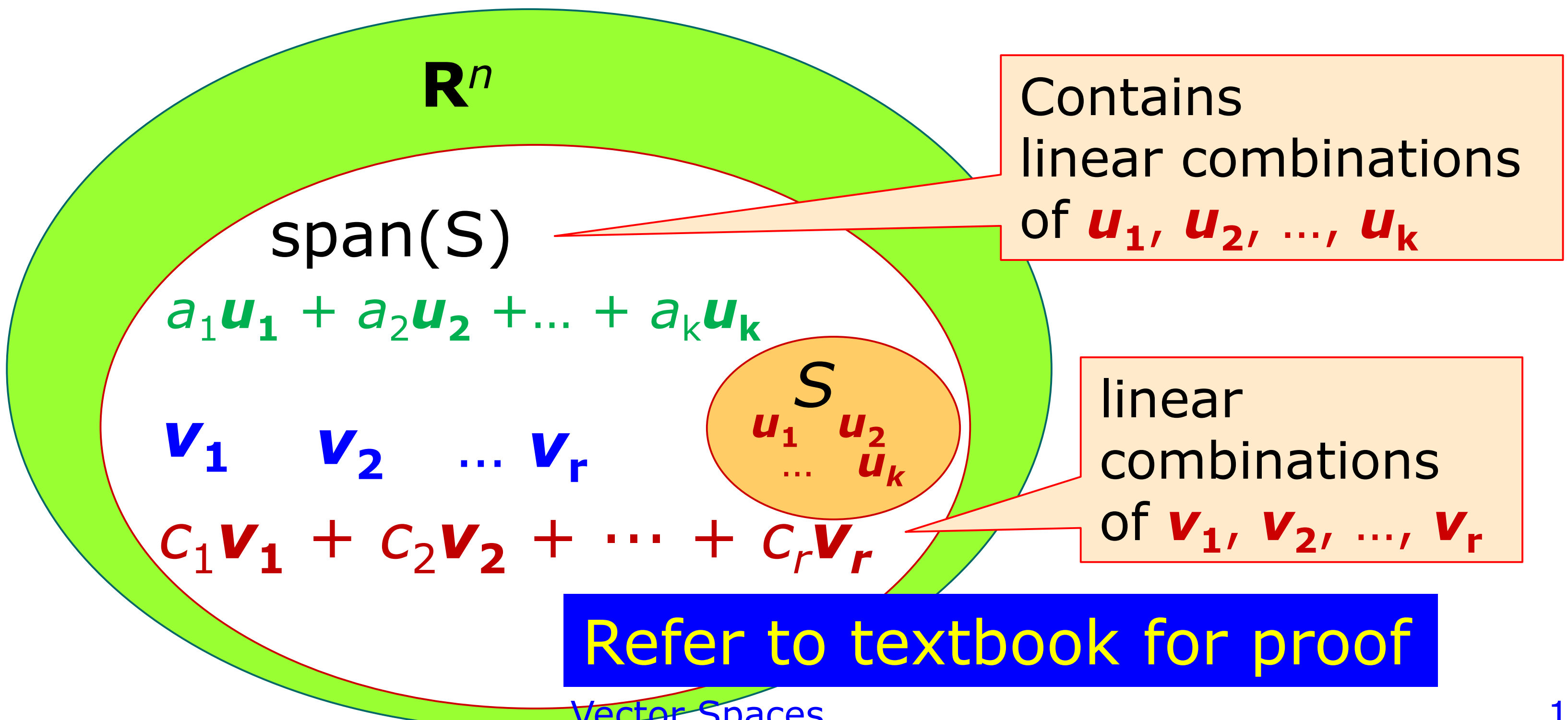
Any linear combination of vectors in a linear span is again a vector in the linear span.

Theorem 3.2.9.2

Let $S = \{ \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k \} \subseteq \mathbf{R}^n$

If $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r \in \text{span}(S)$ and $c_1, c_2, \ldots, c_r \in \mathbf{R}$

then $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_r\boldsymbol{v}_r \in \text{span}(S)$



Refer to textbook for proof

Any linear combination of vectors in a linear span is again a vector in the linear span.

Theorem 3.2.9.2

Let S = { $\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}$ } $\subseteq \mathbf{R}^n$

If $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_r} \in$ span(S) and $c_1, c_2, \ldots, c_r \in \mathbf{R}$

then $c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_r\boldsymbol{v_r} \in$ span(S)

Consequent of theorem

if $\boldsymbol{u}$ and $\boldsymbol{v} \in$ span(S), then $\boldsymbol{u} + \boldsymbol{v} \in$ span(S).

Closure property under vector addition

if $\boldsymbol{u} \in$ span(S) and $c \in \boldsymbol{R}$, then $c\boldsymbol{u} \in$ span(S).

Closure property under scalar multiplication

Motivation

Example 3.2.11.1

$$\text{span}\ \boxed{\begin{array}{c}\boldsymbol{u_1} = (1, 0, 1)\\ \boldsymbol{u_2} = (1, 1, 2)\\ \boldsymbol{u_3} = (-1, 2, 1)\end{array}} \overset{?}{=} \text{span}\ \boxed{\begin{array}{c}\boldsymbol{v_1} = (1, 2, 3)\\ \boldsymbol{v_2} = (2, -1, 1)\end{array}}$$

How are the two linear spans related?

Given two sets A and B.

To show $A = B$: We check $A \subseteq B$ and $B \subseteq A$.

How to show span(S_1) ⊆ span(S_2)?

$\boldsymbol{u_1} = (1, 0, 1)$ $\quad \boldsymbol{v_1} = (1, 2, 3)$
$\boldsymbol{u_2} = (1, 1, 2)$ $\quad \boldsymbol{v_2} = (2, -1, 1)$
$\boldsymbol{u_3} = (-1, 2, 1)$

Example 3.2.11.1

Show span$\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\} \subseteq$ span$\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$:

Need to show: each $\boldsymbol{u_i}$ can be written as $a\boldsymbol{v_1} + b\boldsymbol{v_2}$ for some real number a and b

$$a\boldsymbol{v}_1 + b\boldsymbol{v}_2 = \boldsymbol{u}_1 \quad\Downarrow\quad \begin{cases} a + 2b = 1 \\ 2a - b = 0 \\ 3a + b = 1 \end{cases}$$

$$a\boldsymbol{v}_1 + b\boldsymbol{v}_2 = \boldsymbol{u}_2 \quad\Downarrow\quad \begin{cases} a + 2b = 1 \\ 2a - b = 1 \\ 3a + b = 2 \end{cases}$$

$$a\boldsymbol{v}_1 + b\boldsymbol{v}_2 = \boldsymbol{u}_3 \quad\Downarrow\quad \begin{cases} a + 2b = -1 \\ 2a - b = 2 \\ 3a + b = 1 \end{cases}$$

Need to show all three linear systems are consistent

$$\begin{array}{c} \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 2 & -1 & 0 \\ 3 & 1 & 1 \end{array}\right) \\ \boldsymbol{v_1}\ \ \boldsymbol{v_2}\ \ \boldsymbol{u_1} \end{array} \qquad \begin{array}{c} \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 2 & -1 & 1 \\ 3 & 1 & 2 \end{array}\right) \\ \boldsymbol{v_1}\ \ \boldsymbol{v_2}\ \ \boldsymbol{u_2} \end{array} \qquad \begin{array}{c} \left(\begin{array}{cc|c} 1 & 2 & -1 \\ 2 & -1 & 2 \\ 3 & 1 & 1 \end{array}\right) \\ \boldsymbol{v_1}\ \ \boldsymbol{v_2}\ \ \boldsymbol{u_3} \end{array}$$

How to show span(S_1) ⊆ span(S_2)?

$\boldsymbol{u}_1 = (1, 0, 1)$ $\boldsymbol{v}_1 = (1, 2, 3)$
$\boldsymbol{u}_2 = (1, 1, 2)$ $\boldsymbol{v}_2 = (2, -1, 1)$
$\boldsymbol{u}_3 = (-1, 2, 1)$

Example 3.2.11.1

We can solve the three systems simultaneously:

$$\begin{array}{c} \left(\begin{array}{cc|c|c|c} 1 & 2 & 1 & 1 & -1 \\ 2 & -1 & 0 & 1 & 2 \\ 3 & 1 & 1 & 2 & 1 \end{array}\right) \\ \begin{array}{ccccc} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{array} \end{array} \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \left(\begin{array}{cc|c|c|c} 1 & 0 & \frac{1}{5} & \frac{3}{5} & \frac{3}{5} \\ 0 & 1 & \frac{2}{5} & \frac{1}{5} & -\frac{4}{5} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

All the three systems are consistent.

This shows each $\boldsymbol{u}_i$ can be written as $a\boldsymbol{v}_1 + b\boldsymbol{v}_2$ for some real number a and b,

So span{$\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$} ⊆ span{$\boldsymbol{v}_1, \boldsymbol{v}_2$}. Theorem 3.2.9.2

By solve the three systems, we get:

$$\boldsymbol{u}_1 = \tfrac{1}{5}\boldsymbol{v}_1 + \tfrac{2}{5}\boldsymbol{v}_2 \qquad \boldsymbol{u}_2 = \tfrac{3}{5}\boldsymbol{v}_1 + \tfrac{1}{5}\boldsymbol{v}_2 \qquad \boldsymbol{u}_3 = \tfrac{3}{5}\boldsymbol{v}_1 - \tfrac{4}{5}\boldsymbol{v}_2$$

How to show $\text{span}(S_1) \subseteq \text{span}(S_2)$?

Theorem 3.2.10

Let $S_1 = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $S_2 = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ be subsets of $\mathbf{R}^n$.

Every linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ belongs to $\text{span}(S_2)$

Then

$$\boxed{\text{span}(S_1) \subseteq \text{span}(S_2)}$$

if and only if

$\boxed{\text{each } \boldsymbol{u}_i \text{ is a linear combination of } \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m.}$

Every $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ belongs to $\text{span}(S_2)$

How to show span(S_1) = span(S_2)?

Example 3.2.11.1

$$\text{span}\left\{\begin{array}{l} \boldsymbol{u_1} = (1, 0, 1) \\ \boldsymbol{u_2} = (1, 1, 2) \\ \boldsymbol{u_3} = (-1, 2, 1) \end{array}\right\} \overset{\text{to show}}{=} \text{span}\left\{\begin{array}{l} \boldsymbol{v_1} = (1, 2, 3) \\ \boldsymbol{v_2} = (2, -1, 1) \end{array}\right\}$$

Need to show

span{$\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$} $\subseteq$ span{$\boldsymbol{v_1}, \boldsymbol{v_2}$}

$$\begin{array}{c} \left(\begin{array}{cc|c|c|c} 1 & 2 & 1 & 1 & -1 \\ 2 & -1 & 0 & 1 & 2 \\ 3 & 1 & 1 & 2 & 1 \end{array}\right) \\ \begin{array}{ccccc} \boldsymbol{v_1} & \boldsymbol{v_2} & \boldsymbol{u_1} & \boldsymbol{u_2} & \boldsymbol{u_3} \end{array} \end{array}$$

Check consistencies

span{$\boldsymbol{v_1}, \boldsymbol{v_2}$} $\subseteq$ span{$\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$}

$$\begin{array}{c} \left(\begin{array}{ccc|c|c} 1 & 1 & -1 & 1 & 2 \\ 0 & 1 & 2 & 2 & -1 \\ 1 & 2 & 1 & 3 & 1 \end{array}\right) \\ \begin{array}{ccccc} \boldsymbol{u_1} & \boldsymbol{u_2} & \boldsymbol{u_3} & \boldsymbol{v_1} & \boldsymbol{v_2} \end{array} \end{array}$$

How to show span(S_1) $\neq$ span(S_2)?

Example 3.2.11.2

span $\begin{array}{|l|}\hline \boldsymbol{u_1} = (1, 1, 0, 2) \\ \boldsymbol{u_2} = (1, 0, 0, 1) \\ \boldsymbol{u_3} = (0, 1, 0, 1) \\ \hline \end{array}$ to show $\subseteq$ $\neq$ span $\begin{array}{|c|}\hline \boldsymbol{v_1} = (1, 1, 1, 1) \\ \boldsymbol{v_2} = (-1, 1, -1, 1) \\ \boldsymbol{v_3} = (-1, 1, 1, -1) \\ \hline \end{array}$

span$\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\} \subseteq$ span$\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$

Show that the augmented matrix
$(\boldsymbol{v_1}\ \boldsymbol{v_2}\ \boldsymbol{v_3} \mid \boldsymbol{u_1} \mid \boldsymbol{u_2} \mid \boldsymbol{u_3})$ is consistent.

span$\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\} \neq$ span$\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$

Show that span$\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\} \not\subseteq$ span$\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$

Show that the augmented matrix
$(\boldsymbol{u_1}\ \boldsymbol{u_2}\ \boldsymbol{u_3} \mid \boldsymbol{v_1} \mid \boldsymbol{v_2} \mid \boldsymbol{v_3})$ is inconsistent.

What is a redundant vector in span(S)?

$\mathbf{R}^n$

$\boldsymbol{u}_1$

$\boldsymbol{u}_2$

$\cdots$

$\boldsymbol{u}_k$

Theorem 3.2.12

Suppose $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are vectors taken from $\mathbf{R}^n$.

If $\boldsymbol{u}_k$ is a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k-1}$,

then $\quad \boldsymbol{u}_k = d_1\boldsymbol{u}_1 + d_2\boldsymbol{u}_2 + \cdots + d_{k-1}\boldsymbol{u}_{k-1}$

$$\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k-1}\} = \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k-1}, \boldsymbol{u}_k\}$$

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_{k-1}\boldsymbol{u}_{k-1} \Longleftrightarrow c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_{k-1}\boldsymbol{u}_{k-1} + c_k\boldsymbol{u}_k$$

We say $\boldsymbol{u}_k$ is a "redundant" vector in span$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k-1}, \boldsymbol{u}_k\}$.

If $\boldsymbol{u} \in \operatorname{span}(S)$, then $\operatorname{span}(S) = \operatorname{span}(S \cup \boldsymbol{u})$

Geometrical meaning of linear span

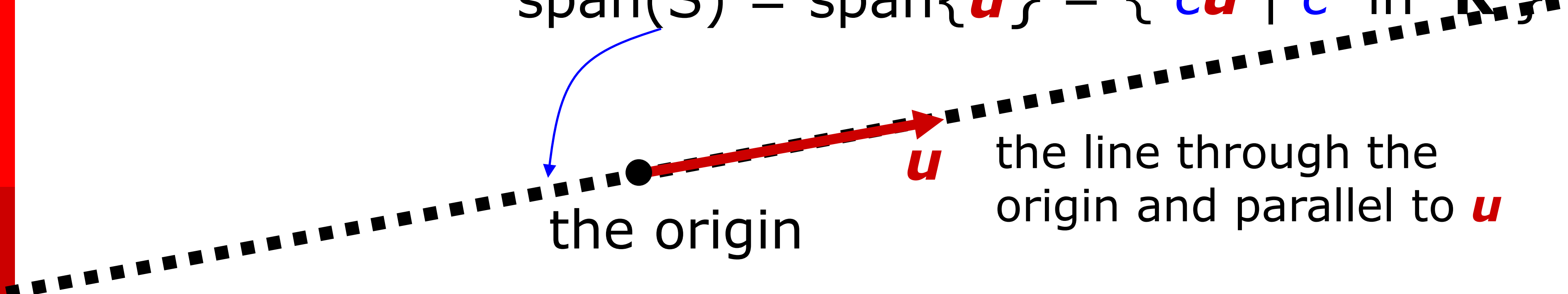
Discussion 3.2.14.1

span = extend across (Oxford Dictionary)

In $\mathbf{R}^2$ and $\mathbf{R}^3$

$S = \{\boldsymbol{u}\}$ ($\boldsymbol{u}$ is a non-zero vector)

$\text{span}(S) = \text{span}\{\boldsymbol{u}\} = \{\, c\boldsymbol{u} \mid c \text{ in } \mathbf{R} \,\}$



$\text{span}(S) = \text{span}\{\boldsymbol{u}\}$ represents a line through the origin

Discussion 3.2.14.2

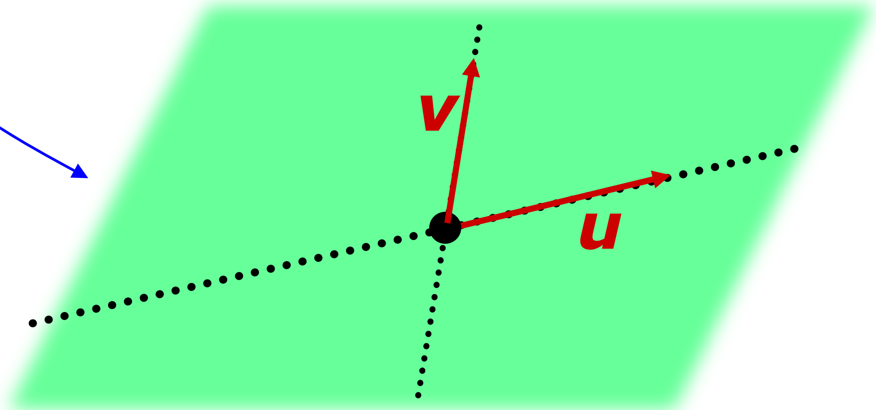
span = extend across (Oxford Dictionary)

In $\mathbf{R}^2$ and $\mathbf{R}^3$

$S = \{\boldsymbol{u}, \boldsymbol{v}\}$ ($\boldsymbol{u}$, $\boldsymbol{v}$ are two non-parallel vectors)

$$\begin{aligned} \text{span}(S) &= \text{span}\{\boldsymbol{u}, \boldsymbol{v}\} \\ &= \{\, s\boldsymbol{u} + t\boldsymbol{v} \mid s, t \in \mathbf{R} \,\} \end{aligned}$$

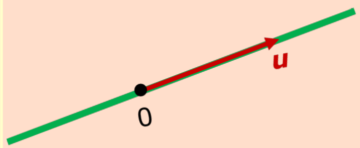
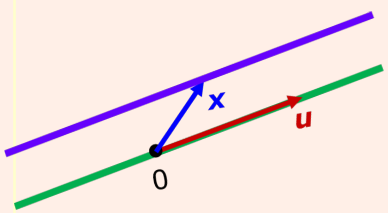
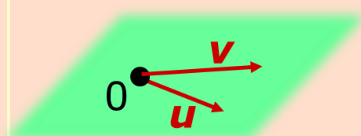
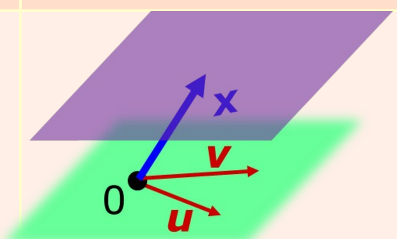
the plane containing the origin and parallel to $\boldsymbol{u}$ and $\boldsymbol{v}$



span(S) = span$\{\boldsymbol{u}, \boldsymbol{v}\}$ represents a plane through the origin

Lines and planes in terms of linear span

Discussion 3.2.15

Objects	Geometrical	Span	Set notation
Line through origin		$\mathrm{span}\{\mathbf{u}\}$	$\{t\mathbf{u} \mid t \in \mathbf{R}\}$
Line not through origin		$\mathbf{x} + \mathrm{span}\{\mathbf{u}\}$	$\{\mathbf{x} + t\mathbf{u} \mid t \in \mathbf{R}\}$ $\{\mathbf{x} + \mathbf{w} \mid \mathbf{w} \in \mathrm{span}\{\mathbf{u}\}\}$
Plane through origin		$\mathrm{span}\{\mathbf{u}, \mathbf{v}\}$	$\{t\mathbf{u} + s\mathbf{v} \mid t, s \in \mathbf{R}\}$
Plane not through origin		$\mathbf{x} + \mathrm{span}\{\mathbf{u}, \mathbf{v}\}$	$\{\mathbf{x} + t\mathbf{u} + s\mathbf{v} \mid t, s \in \mathbf{R}\}$ $\{\mathbf{x} + \mathbf{w} \mid \mathbf{w} \in \mathrm{span}\{\mathbf{u},\mathbf{v}\}\}$

Fill in the blanks

a vector in $\mathbb{R}^2$, a vector in $\mathbb{R}^3$, a line in $\mathbb{R}^3$, a plane in $\mathbb{R}^3$, the entire $\mathbb{R}^3$ space

1. A linear combination of two vectors in $\mathbb{R}^3$ is <u>a vector in $\mathbb{R}^3$</u>.

2. A linear combination of three vectors in $\mathbb{R}^3$ is <u>a vector in $\mathbb{R}^3$</u>.

non-zero

3. A linear span of one vector in $\mathbb{R}^3$ is <u>a line in $\mathbb{R}^3$</u>.

non-parallel

4. A linear span of two vectors in $\mathbb{R}^3$ is <u>a plane in $\mathbb{R}^3$</u>.

non-coplanar

5. A linear span of three vectors in $\mathbb{R}^3$ is <u>the entire $\mathbb{R}^3$ space</u>.

Section 3.3

Subspaces

Objective

- What is a subspace?
- How to show that a subset of $\mathbf{R}^n$ is a subspace?
- What are some subspaces of $\mathbf{R}^n$?
- What is a solution space of a LS?

What is a subspace of $\mathbf{R}^n$?

Definition 3.3.2

no condition

Let V be a subset of $\mathbf{R}^n$

condition applies

V is called a subspace of $\mathbf{R}^n$ provided …

there is a set $S = \{ \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k \}$ of $\mathbf{R}^n$

such that $V = \text{span}(S)$

condition of subspace

i.e. V can be expressed in linear span form.

Every subspace of $\mathbf{R}^n$ is a subset of $\mathbf{R}^n$.

Not every subset of $\mathbf{R}^n$ is a subspace of $\mathbf{R}^n$.

$\{\mathbf{0}\}$ and $\mathbf{R}^n$ are subspaces of $\mathbf{R}^n$

Remark 3.3.3

condition of subspace $V = \text{span}(S)$

1. $\{\mathbf{0}\}$ is a subspace of $\mathbf{R}^n$. zero space

   Take $S = \{\mathbf{0}\}$

   $\{\mathbf{0}\} = \text{span}\{\mathbf{0}\}$

2. $\mathbf{R}^3$ is a subspace of $\mathbf{R}^3$.

   Take S to be standard basis vectors for $\mathbf{R}^3$

   $\boldsymbol{v}_1 = (1, 0, 0)$, $\boldsymbol{v}_2 = (0, 1, 0)$, $\boldsymbol{v}_3 = (0, 0, 1)$

   $\mathbf{R}^3 = \text{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$

   Refer to Example 3.2.2

   $\mathbf{R}^n$ is a subspace of $\mathbf{R}^n$.

   Take S to be standard basis vectors for $\mathbf{R}^n$

   $\boldsymbol{e}_1 = (1, 0, \ldots, 0)$, $\boldsymbol{e}_2 = (0, 1, \ldots, 0)$, …, $\boldsymbol{e}_n = (0, \ldots, 0, 1)$

   $\mathbf{R}^n = \text{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$

How to show that a given subset is a subspace?

Example 3.3.4.1

$V_1 = \{ (a+4b, a) \mid a, b \in \mathbf{R} \}$ explicit form

$(a + 4b, a) = (a, a) + (4b, 0)$

$= a(1, 1) + b(4, 0)$ general linear combination

V_1 is the set of all linear combinations of $(1, 1)$ and $(4, 0)$

$\boxed{V_1 = \operatorname{span}\{(1, 1), (4, 0)\}}$ linear span form

V_1 is a subspace of $\mathbf{R}^2$

In fact $V_1 = \mathbf{R}^2$

How to show that a given subset is a subspace?

Example 3.3.4.2

$V_2 = \{ (x, y, z) \mid x + y - z = 0 \}$ implicit form

$V_2 = \{ (t - s, s, t) \mid s, t \in \mathbf{R} \}$ explicit form

$$(t - s, s, t) = (t, 0, t) + (-s, s, 0)$$

$$= t(1, 0, 1) + s(-1, 1, 0)$$

general linear combination

V_2 is the set of all linear combinations of $(1, 0, 1)$ and $(-1, 1, 0)$

$\boxed{V_2 = \operatorname{span}\{(1, 0, 1), (-1, 1, 0)\}}$ linear span form

V_2 is a subspace of $\mathbf{R}^3$

In fact V_2 is a plane in $\mathbf{R}^3$.

How to show a given subset is not a subspace?

Example 3.3.4.3

$V_3 = \{ (1, a) \mid a \text{ in } \mathbf{R} \}$ subset of $\mathbf{R}^2$

$(1, a) = (1, 0) + (0, a) = (1, 0) + a(0, 1)$

not a general linear combination

V_3 is not a linear span of "any" set of vectors

"So" V_3 is not a subspace of $\mathbf{R}^2$

There is an easier way: Use theorem 3.2.9.1

$(0, 0) \notin V_3 = \{ (1, a) \mid a \text{ in } \mathbf{R} \}$

$\Rightarrow$ not a subspace of $\mathbf{R}^2$

If a subset of $\mathbf{R}^n$ does not contain the zero vector $\mathbf{0}$, then it is not a linear span.

How to show a given subset is not a subspace?

Example 3.3.4.4

$V_4 = \{ (x, y, z) \mid x^2 \leq y^2 \leq z^2 \}$ subset of $\mathbf{R}^3$

e.g. (1, 1, 2), (1, 1, -2), (0, 0, 0) $\in V_4$

Note: Having zero vector in a set V does not guarantee V is a subspace

Use theorem 3.2.9.2

Take two vectors in V, show that the sum is not in V.

$(1, 1, 2) + (1, 1, -2) = (2, 2, 0) \notin V_4$ Not a linear span

Violate the closure property of linear span (theorem 3.2.9.2)

So V_4 is not a subspace of $\mathbf{R}^3$

Remark 3.3.5.1

The following are all the subspaces of $\mathbf{R}^2$:

a. $\{\mathbf{0}\}$ spanned by zero vector $\mathbf{0}$

b. any line that passes through the origin spanned by one non-zero vector $\boldsymbol{u}$

c. $\mathbf{R}^2$ spanned by two non-parallel vectors $\boldsymbol{u}, \boldsymbol{v}$

Why are there no other subspaces of $\mathbf{R}^2$?

$$V = \text{span}\{ \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \ldots, \boldsymbol{u}_k \}$$

all parallel → a line

at least two not parallel → XY-plane $\mathbf{R}^2$

Geometrical interpretation of subspaces of $\mathbf{R}^3$

Remark 3.3.5.2

The following are all the subspaces of $\mathbf{R}^3$:

a. $\{\mathbf{0}\}$ spanned by zero vector $\mathbf{0}$

b. any line through the origin — spanned by one non-zero vector $\boldsymbol{u}$

c. any plane containing the origin — spanned by two non-parallel vectors $\boldsymbol{u}, \boldsymbol{v}$

d. $\mathbf{R}^3$ — spanned by three vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ not lying on a plane

What is a solution space?

Closure properties under vector addition and scalar multiplication

Theorem 3.3.6

$\boldsymbol{Ax} = \mathbf{0}$

The solution set of a homogeneous linear system in n variables is a subspace of $\mathbf{R}^n$.

The solution set of every homogeneous LS can be written as a linear span

We call it the solution space of the system.

The solution set of non-homogeneous LS is not a subspace of $\mathbf{R}^n$.

Example 3.3.7

Homogeneous system

$$\begin{cases} x - 2y + 3z = 0 \\ 2x - 4y + 6z = 0 \\ 3x - 6y + 9z = 0 \end{cases}$$

general solution

$$\begin{cases} x = 2s - 3t \\ y = s \\ z = t \end{cases}$$

subspace of $\mathbf{R}^3$

linear span form

solution set

span{(2, 1, 0), (-3, 0, 1)} = {(2*s* – 3*t*, *s*, *t*) | *s*, *t* in **R**}

s(2, 1, 0) + t(-3, 0, 1)

general linear combination

Closure property of subspaces

Remark 3.3.8

Let V be a non-empty subset of $\mathbf{R}^n$.
Then

V is a subspace of $\mathbf{R}^n$

if and only if

for all $\boldsymbol{u}, \boldsymbol{v} \in V$ and $c, d \in \mathbf{R}$, $c\boldsymbol{u} + d\boldsymbol{v} \in V$.

closure properties under addition & scalar multiplication

This is the actual definition of subspaces in abstract linear algebra.

To show a subset V is a subspace,
(i) check that it contains the zero vector;
(ii) take two general vectors $\boldsymbol{u}$, $\boldsymbol{v}$ in V and $c, d \in \mathbf{R}$, show that $c\boldsymbol{u} + d\boldsymbol{v} \in V$.

To show subspace (or not)

To show a subset S of $\mathbf{R}^n$ is a subspace:

- Express S as a linear span
- Show that S is the solution set of a homogeneous system
- (For $\mathbf{R}^2$ and $\mathbf{R}^3$) show that S represents a line or plane through origin.

To show a subset S of $\mathbf{R}^n$ is not a subspace:

- Show that the zero vector is not in S
- Find $\mathbf{u}, \mathbf{v} \in S$ such that $\mathbf{u} + \mathbf{v} \notin S$
- Find $\mathbf{v} \in S$ and a scalar c such that $c\mathbf{v} \notin S$
- (For $\mathbf{R}^2$ and $\mathbf{R}^3$) show that S is not a line or plane through origin.

Section 3.4

Linear Independence

Objective

- What is a linearly independent/dependent set?
- How to show that a set is linearly (in)dependent?
- What are some conditions on linearly (in)dependent sets?

What is a redundant vector in span(S)?

Example

$S_1 = \{\ (1,1,1),\ (1,0,-2)\ \}$ $\quad$ $S_2 = \{(1,1,1),\ (1,0,-2),\ (2,3,5)\}$

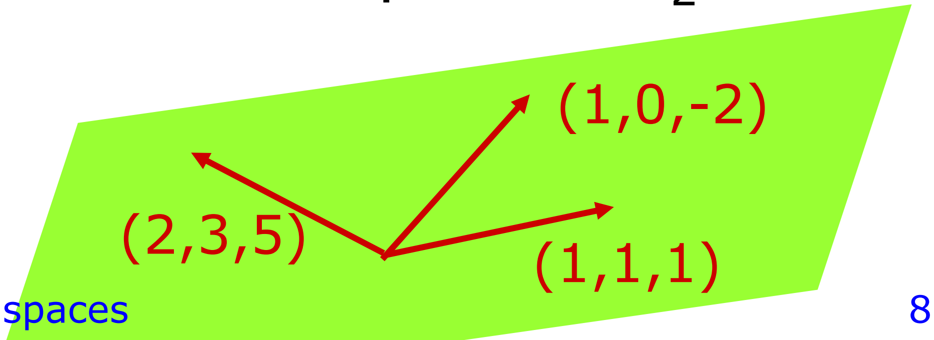
$$\text{span}(S_1) \xleftrightarrow{\text{equal}} \text{span}(S_2)$$

all linear combinations	all linear combinations
$a(1,1,1)+b(1,0,-2)$	$a(1,1,1)+b(1,0,-2)+c\boxed{(2,3,5)}$

$\boxed{3(1,1,1)+(-1)(1,0,-2)}$

Adding the vector $(2, 3, 5)$ to S_1 does not change the linear span of S_1

There is a "redundant" vector in the span of S_2



Homogeneous vector equation

$\boldsymbol{v}_1 = (1, -2, 3)$, $\boldsymbol{v}_2 = (5, 6, -1)$, $\boldsymbol{v}_3 = (3, 2, 1)$

0 0 0

$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$ vector equation

1 1 -2

c_1, c_2, c_3 variable scalars (in **R**)

Can we find scalars c_1, c_2, c_3 that satisfies this vector equation?
Is this the only solution?

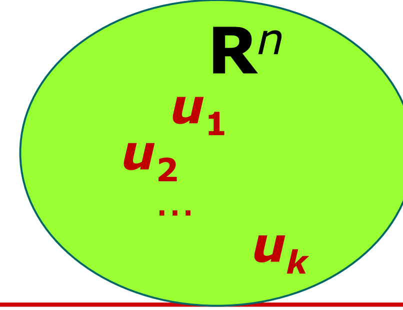
$$c_1\begin{pmatrix}1\\-2\\3\end{pmatrix} + c_2\begin{pmatrix}5\\6\\-1\end{pmatrix} + c_3\begin{pmatrix}3\\2\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \xrightarrow{\text{GE}} \overset{\text{REF}}{\left(\begin{array}{ccc|c}1 & 5 & 3 & 0\\0 & 16 & 8 & 0\\0 & 0 & 0 & 0\end{array}\right)}$$

homogeneous system in variables c_1, c_2, c_3

It has infinitely many solutions

What is linearly independence?

Definition 3.4.2.1



Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a set of vectors in $\mathbf{R}^n$.

If the vector equation

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k = \mathbf{0}$$

has only the trivial solution,

"Working" definition for linearly independence

i.e. the only possible scalars are:

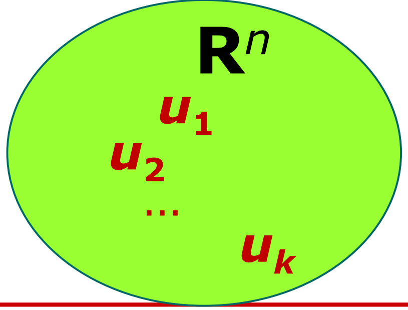
$$c_1 = 0, \; c_2 = 0, \ldots, c_k = 0$$

We say:

S is a linearly independent set and

$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are linearly independent

What is linearly dependence?



Definition 3.4.2.2

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a set of vectors in $\mathbf{R}^n$.

If the vector equation

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k = \mathbf{0}$$

has non-trivial solution,

"Working" definition for linearly dependence

i.e. there exists scalars $c_1, c_2, \ldots, c_n$, not all of them are zero

We say:
S is a linearly dependent set and
$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are linearly dependent

How to show that a set is linearly (in)dependent?

Example 3.4.3.1

Determine whether the vectors

(1, −2, 3), (5, 6, −1), (3, 2, 1)

are linearly independent.

Set up the vector equation:

$$c_1\begin{pmatrix}1\\-2\\3\end{pmatrix}+c_2\begin{pmatrix}5\\6\\-1\end{pmatrix}+c_3\begin{pmatrix}3\\2\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix} \xrightarrow[\text{matrix}]{\text{augmented}} \left(\begin{array}{ccc|c}1&5&3&0\\-2&6&2&0\\3&-1&1&0\end{array}\right)$$

$$\xrightarrow{\text{REF}} \left(\begin{array}{ccc|c}1&5&3&0\\0&16&8&0\\0&0&0&0\end{array}\right)$$

There are infinitely many solutions for c_1, c_2, c_3.

i.e. There exist non-trivial solutions.

So (1, −2, 3), (5, 6, −1), (3, 2, 1) are linearly dependent.

Example 3.4.3.2

Determine whether the vectors

$(1, 0, 0, 1), \; (0, 2, 1, 0), \; (1, -1, 1, 1)$

are linearly independent.

Set up the vector equation:

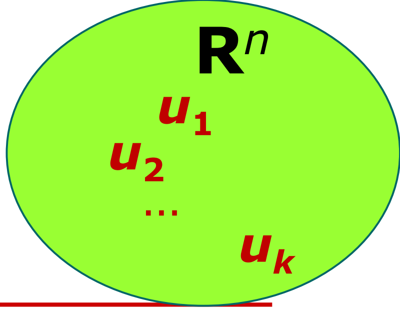
$$c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 2 \\ 1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \xrightarrow{\text{REF}} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Convert into augmented matrix

There is only one solution $c_1 = 0, \; c_2 = 0, \; c_3 = 0.$

So the vectors are linearly independent.

Intuitive meaning of linear dependence



Theorem 3.4.4.1

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ a set with at least 2 vectors

S is linearly dependent

if and only if

at least one vector $\boldsymbol{u}_i$ in S can be written as a linear combination of the other vectors in S

$$\boldsymbol{u}_i = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + \underbrace{c_{i-1}\boldsymbol{u}_{i-1} + c_{i+1}\boldsymbol{u}_{i+1}}_{\boldsymbol{u}_i \text{ is absent}} + \cdots + c_k\boldsymbol{u}_k$$

"redundant" vector

Remark 3.4.5.1

S is linearly dependent

$\Leftrightarrow$ there exists "redundant" vector in span(S)

Show a set is linearly dependent by finding a redundant vector from the set

Example 3.4.6.1

This method is not always easy

$S_1 = \{(1, 0), (0, 4), (2, 4)\} \in \mathbf{R}^2$.

(2, 4) is a linear combination of (1, 0) and (0, 4).

$$(2, 4) = 2(1, 0) + (0, 4)$$

(2,4) is a "redundant" vector:

$$\text{span}\{(1, 0), (0, 4), (2, 4)\} = \text{span}\{(1, 0), (0, 4)\}$$

So we can conclude that S_1 is linearly dependent.

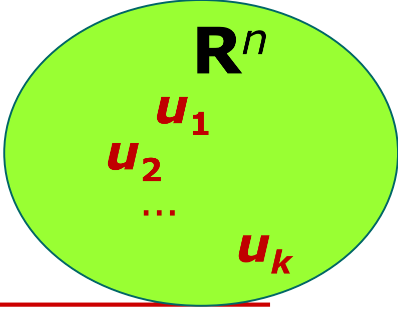
Verification:

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3 = \boldsymbol{0}$$

$$(0, 0) = 2(1, 0) + (0, 4) - (2, 4)$$

non-trivial scalars

Intuitive meaning of linear independence



Theorem 3.4.4.2

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ a set with at least 2 vectors

S is linearly independent
if and only if
no vector in S can be written
as a linear combination of other vectors in S

Remark 3.4.5.2

S is linearly independent

$\Leftrightarrow$ there is no "redundant" vector in span(S)

Show a set is linearly independent by showing there is no redundant vector from the set

Example 3.4.6.2

This is not an efficient method

$S_2 = \{(-1, 0, 0), (0, 3, 0), (0, 0, 7)\} \in \mathbf{R}^3$.

$(-1, 0, 0)$ not a lin. comb. of $(0, 3, 0)$ and $(0, 0, 7)$

$(0, 3, 0)$ not a lin. comb. of $(-1, 0, 0)$ and $(0, 0, 7)$

$(0, 0, 7)$ not a lin. comb. of $(-1, 0, 0)$ and $(0, 3, 0)$

We can conclude that S_2 is linearly independent.

There is no redundant vector in S_2:

$\boxed{\text{span}\{(-1, 0, 0), (0, 3, 0), (0, 0, 7)\}}$

$\neq$ $\neq$ $\neq$

$\boxed{\text{span}\{(0, 3, 0), (0, 0, 7)\}}$

$\boxed{\text{span}\{(-1, 0, 0), (0, 3, 0)\}}$ $\boxed{\text{span}\{(-1, 0, 0), (0, 0, 7)\}}$

A set with one vector

Example 3.4.3.3

The vector equation

$$c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k} = \mathbf{0}$$

has non-trivial solution/ only trivial solution

Let $S = \{\boldsymbol{u}\}$ be a set with one vector.

Is S linearly dependent / independent?

$c\boldsymbol{u} = \mathbf{0}$ for some nonzero c / only $c = 0$

If $\boldsymbol{u} = \mathbf{0}$, then c can be non-zero.
So $S = \{\boldsymbol{u}\}$ is linearly dependent

If $\boldsymbol{u} \neq \mathbf{0}$, then c must be zero.
So $S = \{\boldsymbol{u}\}$ is linearly independent

A set with two vectors

Example 3.4.3.4

The vector equation

$$c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k} = \boldsymbol{0}$$

has non-trivial solution / only trivial solution

Let $S = \{\boldsymbol{u}, \boldsymbol{v}\}$ be a set with two vectors.

Is S linearly dependent / independent ?

$c\boldsymbol{u} + d\boldsymbol{v} = \boldsymbol{0}$ for c, d not both 0 / c, d both 0

$\boldsymbol{u} = (-d/c)\boldsymbol{v}$ or $\boldsymbol{v} = (-c/d)\boldsymbol{u}$

$c \neq 0$ $d \neq 0$

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are scalar multiples of each other, S is linearly dependent

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are not scalar multiples of each other, S is linearly independent

A set with the zero-vector

Example 3.4.3.5

Let S be a finite subset of $\mathbf{R}^n$.

If $\mathbf{0} \in S$, then S is linearly dependent

Hint:

Consider the vector equation

$$c_1 \boxed{\mathbf{0}} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k} = \mathbf{0}$$

Show that this equation can have non-trivial solutions for $c_1, c_2, \ldots, c_k$

A sufficient condition for linear dependence

Theorem 3.4.7 & Example 3.4.8

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a set of vectors in $\mathbf{R}^n$.

If $k > n$, then S is linearly dependent.

If $S \subseteq \mathbf{R}^n$ and S has more than n elements, then S is linearly dependent.

1. In $\mathbf{R}^2$, a set of three or more vectors must be linearly dependent.

   {(1,2), (3,4), (5,6)} is linearly dependent

2. In $\mathbf{R}^3$, a set of four or more vectors must be linearly dependent.

   {(1,2,3), (3,4,5), (5,6,7), (7,8,9)} is linearly dependent

The proof

Theorem 3.4.7

$S = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_k}\}$ in $\mathbf{R}^n$

$$c_1\mathbf{u_1} + c_2\mathbf{u_2} + \cdots + c_k\mathbf{u_k} = \mathbf{0}$$ vector equation

$$\begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix} \quad \begin{pmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{pmatrix} \quad \begin{pmatrix} a_{k1} \\ a_{k2} \\ \vdots \\ a_{kn} \end{pmatrix}$$

$$\begin{cases} a_{11}c_1 + a_{21}c_2 + \ldots + a_{k1}c_k = 0 \\ a_{12}c_1 + a_{22}c_2 + \ldots + a_{k2}c_k = 0 \\ \vdots \qquad\qquad\qquad\qquad\qquad \vdots \\ a_{1n}c_1 + a_{2n}c_2 + \ldots + a_{kn}c_k = 0 \end{cases}$$

Homogeneous system of n linear equations in k variables $c_1, c_2, \ldots, c_k$

The proof

$$\begin{cases} a_{11}c_1 + a_{21}c_2 + \dots + a_{k1}c_k = 0 \\ a_{12}c_1 + a_{22}c_2 + \dots + a_{k2}c_k = 0 \\ \vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ a_{1n}c_1 + a_{2n}c_2 + \dots + a_{kn}c_k = 0 \end{cases}$$

Theorem 3.4.7

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_k}\}$

$$c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k} = \boldsymbol{0} \quad (*)$$

$k > n$ $\Rightarrow$ more variables than equations

$\Rightarrow$ the system has non-trivial solutions

$\Rightarrow$ equation $(*)$ has non-trivial scalars

$\Rightarrow$ S is linearly dependent.

Remark 1.5.4.2:
A homogeneous system with more unknowns than equations has infinitely many solutions

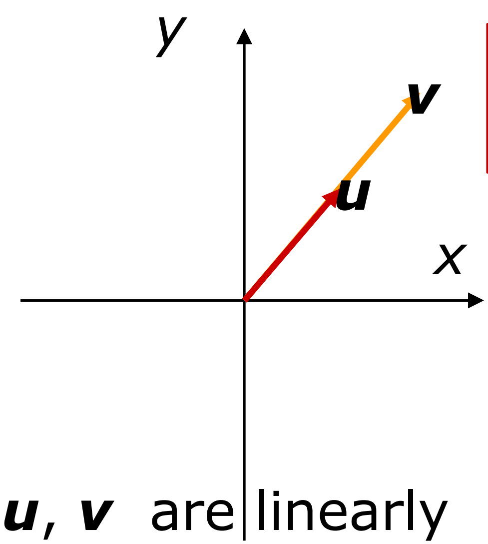
Homogeneous system of n linear equations in k variables $c_1, c_2, \dots, c_k$

Discussion 3.4.9.1 (for two vectors)

u and ***v*** in $\mathbf{R}^2$ are linearly independent if and only if span{***u***, ***v***} = $\mathbf{R}^2$

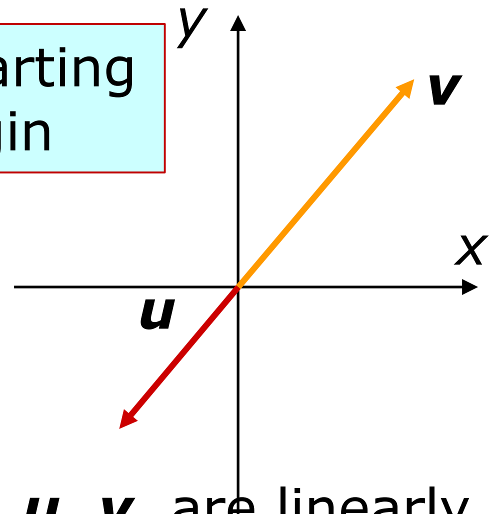
In $\mathbf{R}^2$ (or $\mathbf{R}^3$), two vectors ***u*** and ***v*** are linearly dependent if and only if they lie on the same line.

arrow starting from origin



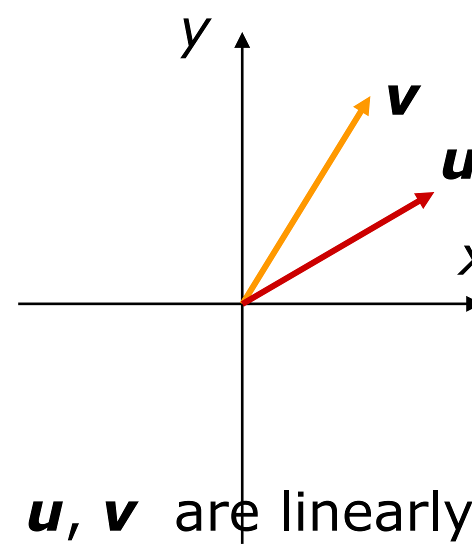
u, ***v*** are linearly dependent

span{***u***, ***v***} = line



u, ***v*** are linearly dependent

span{***u***, ***v***} = line



u, ***v*** are linearly independent

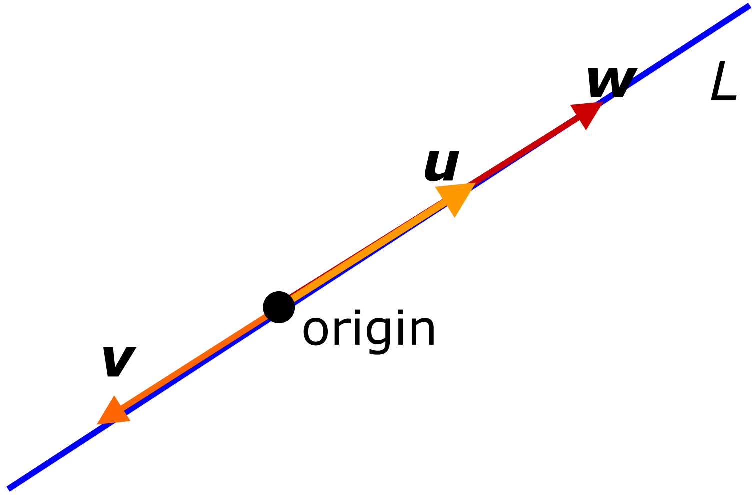
span{***u***, ***v***} = plane

Geometrical meaning of linear independence

Discussion 3.4.9.2 (for three vectors)

In $\mathbf{R}^3$,

three vectors $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ are linearly dependent if and only if they lie on the same line or same plane.



span{$\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$} = L
= span{$\boldsymbol{u}$}
= span{$\boldsymbol{v}$}
= span{$\boldsymbol{w}$}

$\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ are linearly dependent

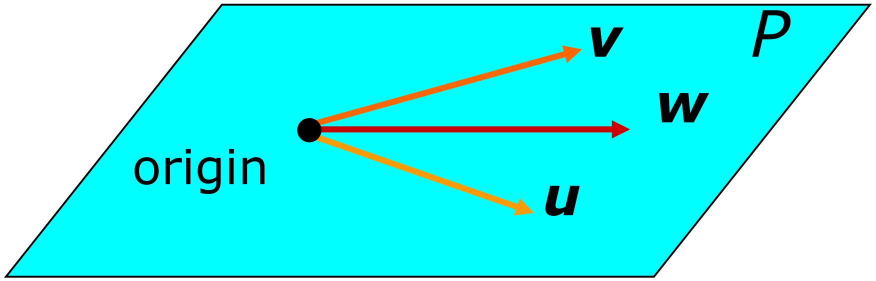
First case: same line

Geometrical meaning of linear independence

Discussion 3.4.9.2 (for three vectors)

In $\mathbf{R}^3$,

three vectors $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ are linearly dependent if and only if they lie on the same line or same plane.



$$\begin{aligned}\text{span}\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\} &= P \\ &= \text{span}\{\boldsymbol{u}, \boldsymbol{v}\} \\ &= \text{span}\{\boldsymbol{v}, \boldsymbol{w}\} \\ &= \text{span}\{\boldsymbol{u}, \boldsymbol{w}\}\end{aligned}$$

$\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ are linearly dependent

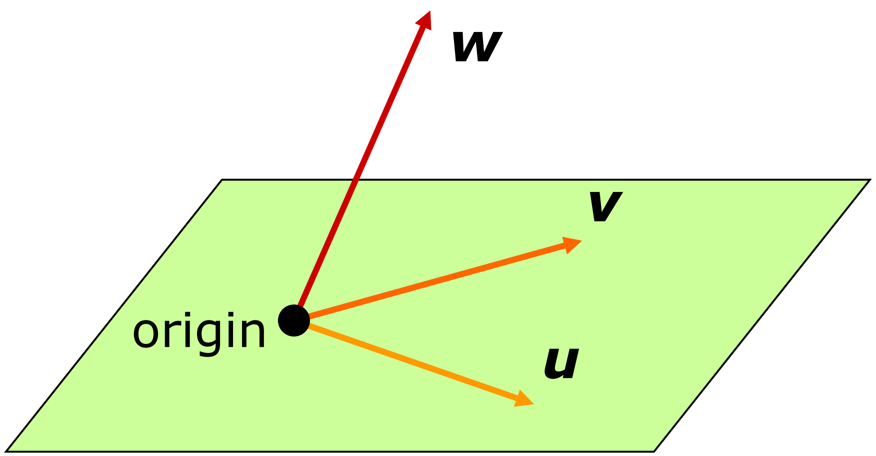
<u>Second case</u>: same plane

Geometrical meaning of linear independence

Discussion 3.4.9.2 (for three vectors)

In $\mathbf{R}^3$,

three vectors $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ are linearly dependent if and only if they lie on the same line or same plane.



$\text{span}\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\} = \mathbf{R}^3$
$\neq \text{span}\{\boldsymbol{u}, \boldsymbol{v}\}$
$\neq \text{span}\{\boldsymbol{v}, \boldsymbol{w}\}$
$\neq \text{span}\{\boldsymbol{u}, \boldsymbol{w}\}$

$\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ are linearly independent

$\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ in $\mathbf{R}^3$ are linearly independent if and only if $\text{span}\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\} = \mathbf{R}^3$

How to extend a linearly independent set?

Theorem 3.4.10

$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are linearly independent

$\boldsymbol{u}_{k+1}$ is not redundant

If $\boldsymbol{u}_{k+1}$ is not a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$

then $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}$ are linearly independent.

(This result gives us a way to add more vectors to a collection of linearly independent vectors.)

Outline of proof

Theorem 3.4.10

$\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}$ are linearly independent (I)

If $\boldsymbol{u_{k+1}}$ is not a linear combination of $\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}$ (II)

then $\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}, \boldsymbol{u_{k+1}}$ are linearly independent.

Prove by contradiction

Suppose $\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}, \boldsymbol{u_{k+1}}$ are linearly dependent

Then $c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_{k+1}\boldsymbol{u_{k+1}} = \boldsymbol{0}$ --(*)

for some $c_1, c_2, \cdots, c_{k+1}$ not all 0

Consider two cases: (i) $c_{k+1} = 0$ and (ii) $c_{k+1} \neq 0$

Case (i) (*) becomes $c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k} = \boldsymbol{0}$

This will contradict (I)

Case (ii) (*) becomes $c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k} = -c_{k+1}\boldsymbol{u_{k+1}}$

This will contradict (II)

Exercise (similar to Ex 3 Q27)

Given $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ are linearly independent

Are $\boldsymbol{u} + \boldsymbol{v}$, $\boldsymbol{u} + \boldsymbol{w}$, $\boldsymbol{v} + \boldsymbol{w}$ linearly independent?

Consider $a(\boldsymbol{u} + \boldsymbol{v}) + b(\boldsymbol{u} + \boldsymbol{w}) + c(\boldsymbol{v} + \boldsymbol{w}) = \mathbf{0}$ (*)

Does (*) have non-trivial scalars for a, b, c ?

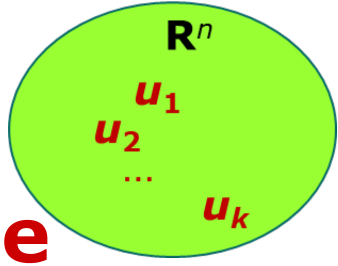
Rewrite (*): $(a+b)\boldsymbol{u} + (a+c)\boldsymbol{v} + (b+c)\boldsymbol{w} = \mathbf{0}$ (**)

(**) has only trivial scalars for a+b, a+c, b+c

$$\left.\begin{aligned} a + b &= 0 \\ a + c &= 0 \\ b + c &= 0 \end{aligned}\right\} \text{ Solve: } a = b = c = 0$$

So (*) has only trivial scalars for a, b, c

So $\boldsymbol{u} + \boldsymbol{v}$, $\boldsymbol{u} + \boldsymbol{w}$, $\boldsymbol{v} + \boldsymbol{w}$ are linearly independent



Linear span VS linear independence

Given that: $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is a subset of $\mathbf{R}^n$

To Show:
$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ spans $\mathbf{R}^n$

same as: span$(S) = \mathbf{R}^n$

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k = \boldsymbol{v} \quad \begin{pmatrix} x \\ y \\ \vdots \\ z \end{pmatrix}$$

$\boldsymbol{v}$ is any general vector in $\mathbf{R}^n$

check whether the system is always consistent

yes → spans

no → does not span

To Show:
$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is lin. indep.

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k = \boldsymbol{0} \quad \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$\boldsymbol{0}$ is the zero vector in $\mathbf{R}^n$

check whether the system has non-trivial solution

yes → lin.dep

no → lin.indep

Section 3.5

Bases

Objective

- What is a basis for a vector space?
- How to show that a set is a basis?
- How to find a basis for a vector space?
- What are coordinate vectors?

Discussion 3.5.1

A set V is called a vector space if:

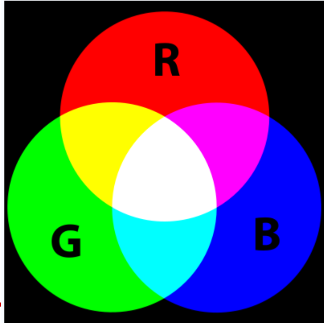
- either $V = \mathbf{R}^n$
- or V is a subspace of $\mathbf{R}^n$.

Examples

They are vector spaces

- $\mathbf{R}^3$ is a subspace of $\mathbf{R}^3$
- $\{\mathbf{0}\}$ is a subspace of $\mathbf{R}^3$
- span{ (1,2,3) } is a subspace of $\mathbf{R}^3$
- span{ (1,2,3), (2,1,4) } is a subspace of $\mathbf{R}^3$

Color mixing



Three primary colors: Red, Green, Blue (RGB)

Different color shade combination gives "all" colors

e.g. 20% Red + 45% Green + 30% Blue

The three primary colors span the color space:

- span{ Red, Green, Blue} = Color space

None of the three primary colors are redundant:

- {Red, Green, Blue} is linearly independent

What is a basis?

Example

e.g. $(2, 3, -5) = 2\boldsymbol{e}_1 + 3\boldsymbol{e}_2 - 5\boldsymbol{e}_3$

Standard basis vectors for $\mathbf{R}^3$

$\boldsymbol{e}_1 = (1, 0, 0)$, $\boldsymbol{e}_2 = (0, 1, 0)$, $\boldsymbol{e}_3 = (0, 0, 1)$

- $\text{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\} = \mathbf{R}^3$ building block

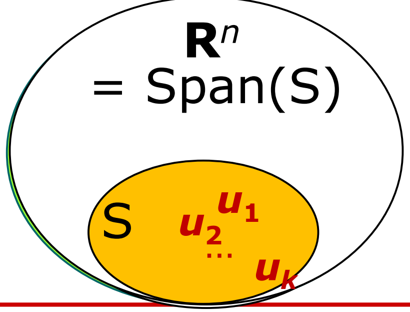
- $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is linearly independent

No redundant vectors

$S = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is called a basis for $\mathbf{R}^3$

S is a smallest possible subset of $\mathbf{R}^3$ so that every vector in $\mathbf{R}^3$ is a linear combination of the elements in S.

What is a basis for $\mathbf{R}^n$?



Definition 3.5.4

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a subset of $\mathbf{R}^n$.

Then S is called a basis for $\mathbf{R}^n$ if

1. S is linearly independent   no redundant vectors in S
2. S spans $\mathbf{R}^n$. $\text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} = \mathbf{R}^n$

Remark 3.5.6.1

A basis for $\mathbf{R}^n$ contains the smallest possible number of vectors that can span $\mathbf{R}^n$.

Remark 3.5.6.3

$\mathbf{R}^n$ has infinitely many bases.

$\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$

$\{(2, 0, 0), (0, 2, 0), (0, 0, 2)\}$

$\{(1, 2, 1), (2, 9, 0), (3, 3, 4)\}$

How to show that a set is a basis (for $\mathbf{R}^3$)?

Example 3.5.5.1

Show that $S = \{(1, 2, 1), (2, 9, 0), (3, 3, 4)\}$ is a basis for $\mathbf{R}^3$.

(i) S is linearly independent:

$$c_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 9 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 3 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Gaussian Elimination (details skipped) $\Rightarrow$ $c_1 = 0$, $c_2 = 0$ and $c_3 = 0$

The system only has the trivial solution.
So S is linearly independent.

How to show that a set is a basis (for $\mathbf{R}^3$)?

Example 3.5.5.1

Show $S = \{(1, 2, 1), (2, 9, 0), (3, 3, 4)\}$ is a basis for $\mathbf{R}^3$.

(ii) span(S) $=$ $\mathbf{R}^3$:

Let (x, y, z) be any (general) vector in $\mathbf{R}^3$.

$$c_1\begin{pmatrix}1\\2\\1\end{pmatrix} + c_2\begin{pmatrix}2\\9\\0\end{pmatrix} + c_3\begin{pmatrix}3\\3\\4\end{pmatrix} = \begin{pmatrix}x\\y\\z\end{pmatrix}$$

Gaussian Elimination $\Rightarrow$ system is consistent for any values of x, y, z.
(details skipped)

So span (S) = $\mathbf{R}^3$.

By (i) and (ii), we conclude S is a basis for $\mathbf{R}^3$.

A set that is not a basis (for $\mathbf{R}^4$)

Example 3.5.5.3

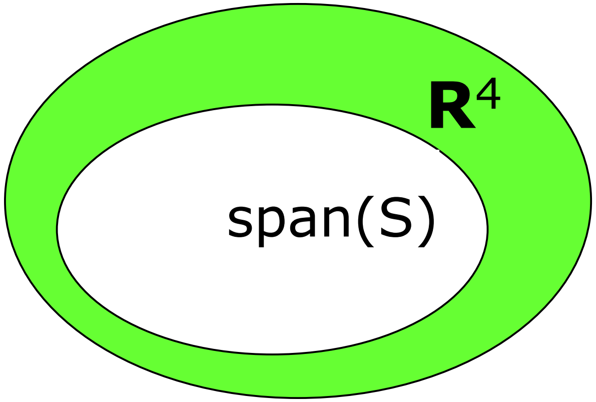
Is $S = \{(1, 1, 1, 1), (0, 0, 1, 2), (-1, 0, 0, 1)\}$ a basis for $\mathbf{R}^4$?

A basis for $\mathbf{R}^n$ must have n elements

S is linearly independent

span(S) $\neq \mathbf{R}^4$ $(|S| < 4)$

So S is not a basis for $\mathbf{R}^4$.



span(S) is a subspace of $\mathbf{R}^4$

S is a basis for this subspace span(S)

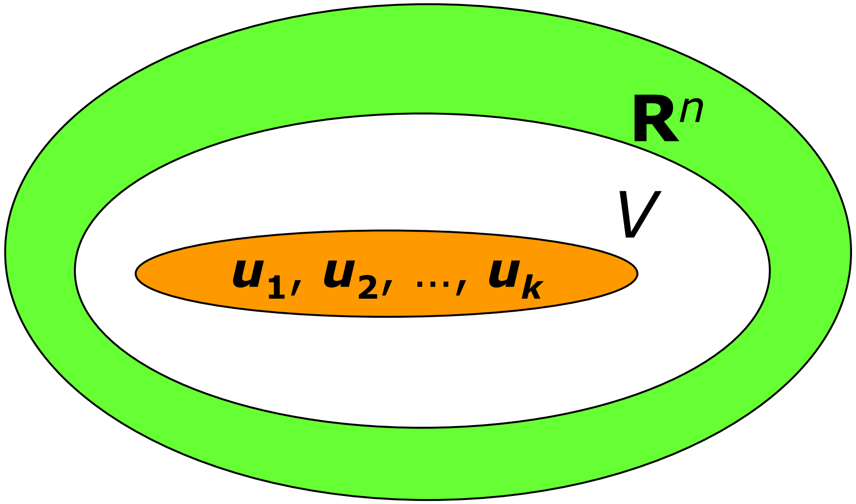
What is a basis for a subspace of $\mathbf{R}^n$?

Definition 3.5.4

Let V be a subspace of $\mathbf{R}^n$

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a subset of V.
Then S is called a basis for V if

1. S is linearly independent — no redundant vectors in S
2. S spans V. — $\text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} = V$



A set that is a basis for a subspace (of $\mathbf{R}^4$)

Example 3.5.5.2

Let V = span{(1,1,1,1), (1,-1,-1,1), (1,0,0,1)} and S = {(1, 1, 1, 1), (1, -1, -1, 1)}.

Show S a basis for V.

(i) S is linearly independent:

$$c_1 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Gaussian Elimination (details skipped) $\Rightarrow$ $c_1 = 0$ and $c_2 = 0$

The system only has the trivial solution.
So S is linearly independent.

Alternatively:
Just observe that (1, 1, 1, 1) and (1, -1, -1, 1) are not scalar multiple of each other, hence S is linearly indep.

A set that is a basis for a subspace (of $\mathbf{R}^4$)

Example 3.5.5.2

Let V = span{$\overset{\boldsymbol{u_1}}{(1,1,1,1)}$, $\overset{\boldsymbol{u_2}}{(1,-1,-1,1)}$, $\overset{\boldsymbol{u_3}}{(1,0,0,1)}$}
and S = {$\underset{\boldsymbol{u_1}}{(1, 1, 1, 1)}$, $\underset{\boldsymbol{u_2}}{(1, -1, -1, 1)}$}.

Show S is a basis for V.

show this vector is redundant

(ii) span(S) = V: $\leftrightarrow$ span{$\boldsymbol{u_1}$, $\boldsymbol{u_2}$ } = span{$\boldsymbol{u_1}$, $\boldsymbol{u_2}$, $\boldsymbol{u_3}$}

Just need to show (1,0,0,1) is a linear combination of (1, 1, 1, 1) and (1, -1, -1, 1)

We can easily get

$$(1,0,0,1) = \tfrac{1}{2}\,(1,1,1,1) + \tfrac{1}{2}\,(1,-1,-1,1)$$

So (1,1,1,1), (1,-1,-1,1), (1,0,0,1) ∈ sp
By Theorem 3.2.12,
span{(1,1,1,1), (1,-1,-1,1), (1,0,0,1)} ⊆ span(S)

Can be omitted

A set that is not a basis for a subspace (of $\mathbf{R}^3$)

Example 3.5.5.4

Let V = span(S) where

$S = \{(1, 1, 1), (0, 0, 1), (1, 1, 0)\}$

Is S a basis for V ?

S is linearly dependent $(1, 1, 1) = (0, 0, 1) + (1, 1, 0)$

So *S* is not a basis for V though span(*S*) = V

In general,

- if S is linearly dependent, then S is not a basis for span(S)
- if S is linearly independent, then S is a basis for span(S).

S spans span(S)

How to find a basis for a subspace?

Example

$V = \{(a, a+b, b) \mid a, b \text{ in } \mathbf{R}\}$ is a subspace of $\mathbf{R}^3$

Find a basis for *V*.

Write *V* in linear span form

$$\begin{pmatrix} a \\ a+b \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Express
a general vector in *V* as
a linear combination

(1) $V = \text{span}\{(1, 1, 0), (0, 1, 1)\}$

(2) $\{(1, 1, 0), (0, 1, 1)\}$ is linearly independent

the two vectors are not scalar multiples of each other

So $\{(1, 1, 0), (0, 1, 1)\}$ is a basis for *V*

How to show that a set is a basis for a subspace?

Example

$V = \{(a, a+b, b) \mid a, b \text{ in } \mathbf{R}\}$ is a subspace of $\mathbf{R}^3$

Show that $S = \{(1, 3, 2), (1, 2, 1)\}$ is a basis for V

Check S is linearly independent

(1, 3, 2) and (1, 2, 1) are not scalar multiples of each other

Check $\boxed{\text{span}(S) = V}$ $\qquad V = \text{span}\{(1, 1, 0), (0, 1, 1)\}$

$\text{span}\{(1,3,2), (1,2,1)\} = \text{span}\{(1,1,0), (0,1,1)\}$ (*)

To show (*), refer: Example 3.2.11

$$\left(\begin{array}{cc|c|c} 1 & 1 & 1 & 0 \\ 3 & 2 & 1 & 1 \\ 2 & 1 & 0 & 1 \end{array}\right) \qquad \left(\begin{array}{cc|c|c} 1 & 0 & 1 & 1 \\ 1 & 1 & 3 & 2 \\ 0 & 1 & 2 & 1 \end{array}\right)$$

Basis for the zero space

Remark 3.5.6.2

What is a basis for the zero space $\{\mathbf{0}\}$?

$\{\mathbf{0}\} = \text{span}\{\mathbf{0}\}$

$\{\mathbf{0}\}$ is linearly dependent

So $\{\mathbf{0}\}$ is not a basis for the zero space

We regard the empty set $\emptyset$ as the basis for $\{\mathbf{0}\}$.

Uniqueness expression in terms of basis

Theorem 3.5.7

S a basis for V

S spans V

S lin. indep.

Let $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$
be a basis for a vector space V.

subspace of $\mathbf{R}^n$

Every vector $\boldsymbol{v}$ in V
can be expressed in the form

consequence of S spans V

$$\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}$$

in exactly one way.

consequence of S is linearly indep.

i.e. there is a unique set of values for $c_1, c_2, \ldots, c_k$.

Example Suppose $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ is a basis for $\mathbf{R}^3$.

Then $3\boldsymbol{u_1} + 5\boldsymbol{u_2} + 2\boldsymbol{u_3} \neq 2\boldsymbol{u_1} + 4\boldsymbol{u_2} + 6\boldsymbol{u_3}$

Proof of uniqueness

Theorem 3.5.7

Every vector $\boldsymbol{v}$ in V can be expressed in the form

$$\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}$$

in exactly one way.

Express $\boldsymbol{v}$ as two linear combinations:

$$\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k} \quad (1)$$

$$\boldsymbol{v} = d_1\boldsymbol{u_1} + d_2\boldsymbol{u_2} + \cdots + d_k\boldsymbol{u_k} \quad (2)$$

(1) – (2):

$$\Rightarrow (c_1 - d_1)\boldsymbol{u_1} + (c_2 - d_2)\boldsymbol{u_2} + \cdots + (c_k - d_k)\boldsymbol{u_k} = \boldsymbol{0}$$

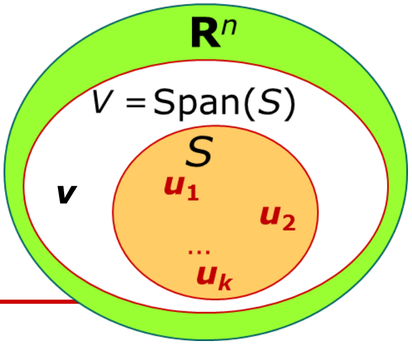
Given S is linearly independent

$$\Rightarrow c_1 - d_1 = 0, \; c_2 - d_2 = 0, \; \ldots, \; c_k - d_k = 0$$

$$\Rightarrow c_1 = d_1, \; c_2 = d_2, \; \ldots, \; c_k = d_k.$$

So the expression is unique.

What are coordinate vectors?



Definition 3.5.8

fixed order

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$: a basis for a vector space V

$\boldsymbol{v}$: a vector in V $\in \mathbf{R}^n$

$$\boldsymbol{v} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k$$ (unique expression)

$(\boldsymbol{v})_S = (c_1, c_2, \ldots, c_k)$ $\in \mathbf{R}^k$

$c_1, c_2, \ldots, c_k$ are called the coordinates of $\boldsymbol{v}$ relative to the basis S

Form the vector $(c_1, c_2, \ldots, c_k)$ in $\mathbf{R}^k$

This is called the coordinate vector of $\boldsymbol{v}$ relative to S

Denote this vector by $(\boldsymbol{v})_S$

depends on basis S

How to find coordinate vectors?

Example 3.5.9.1

Let $S = \{(1, 2, 1), (2, 9, 0), (3, 3, 4)\}$.

S is a basis for $\mathbf{R}^3$.

(a) Find the coordinate vector of $\boldsymbol{v} = (5, -1, 9)$ relative to S.

$$\boldsymbol{v} \longrightarrow (\boldsymbol{v})_S\,?$$

(b) Find a vector $\boldsymbol{w}$ in $\mathbf{R}^3$ such that $(\boldsymbol{w})_S = (-1, 3, 2)$.

$$\boldsymbol{w}\,? \longleftarrow (\boldsymbol{w})_S$$

How to find coordinate vectors?

$$\boldsymbol{v} \longrightarrow (\boldsymbol{v})_S$$

$$\boldsymbol{w} \longleftarrow (\boldsymbol{w})_S$$

Example 3.5.9.1

(a) Solving the equation

$$a(1, 2, 1) + b(2, 9, 0) + c(3, 3, 4) = (5, -1, 9)$$

set up LS and use GE etc

we obtain $a = 1$, $b = -1$, $c = 2$.

i.e. $\boldsymbol{v} = 1(1, 2, 1) - (2, 9, 0) + 2(3, 3, 4)$.

So $(\boldsymbol{v})_S = (1, -1, 2)$.

(b) $(\boldsymbol{w})_S = (-1, 3, 2)$

substitution

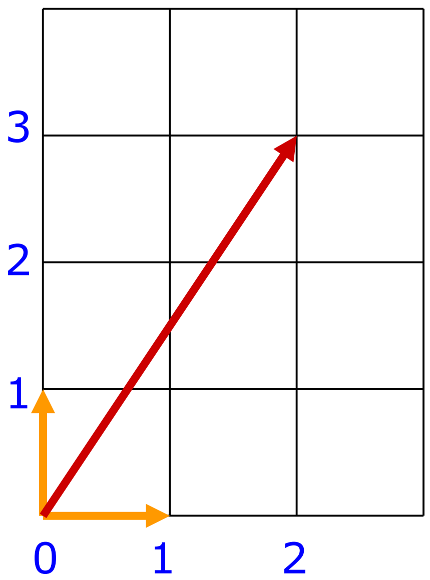
$$\boldsymbol{w} = a(1, 2, 1) + b(2, 9, 0) + c(3, 3, 4)$$

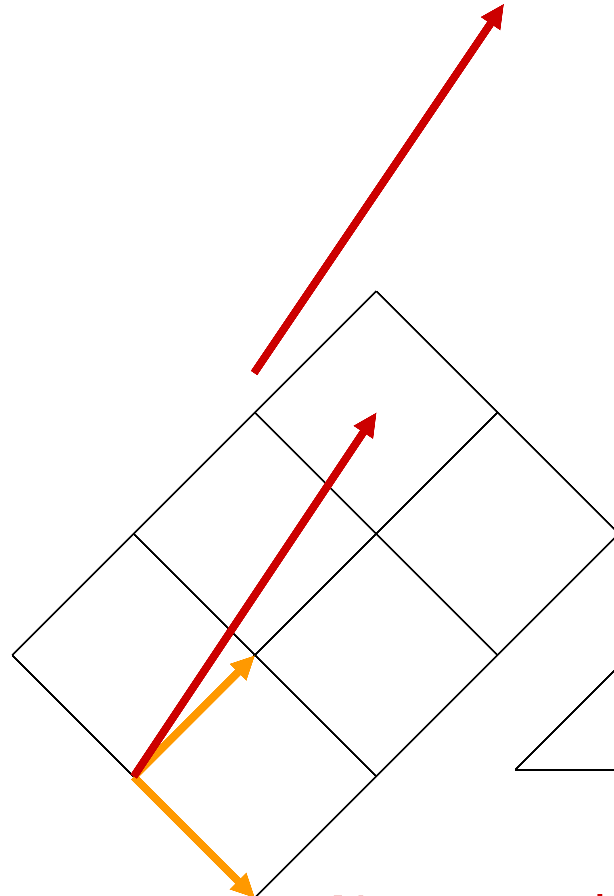
$$= (11, 31, 7).$$

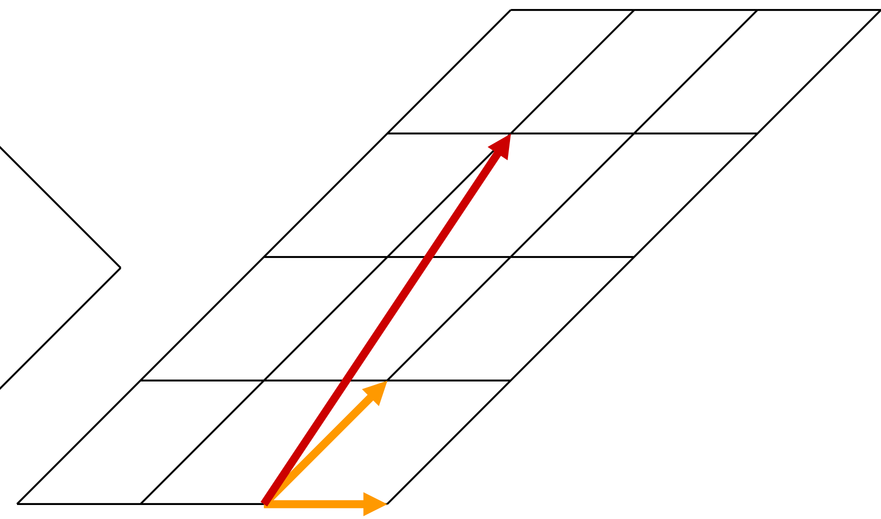
Geometrical meaning of coordinate vectors

Example 3.5.9.2

$\boldsymbol{v} = (2, 3)$







Standard basis

$S_1 = \{(1, 0), (0, 1)\}$

Non-standard bases

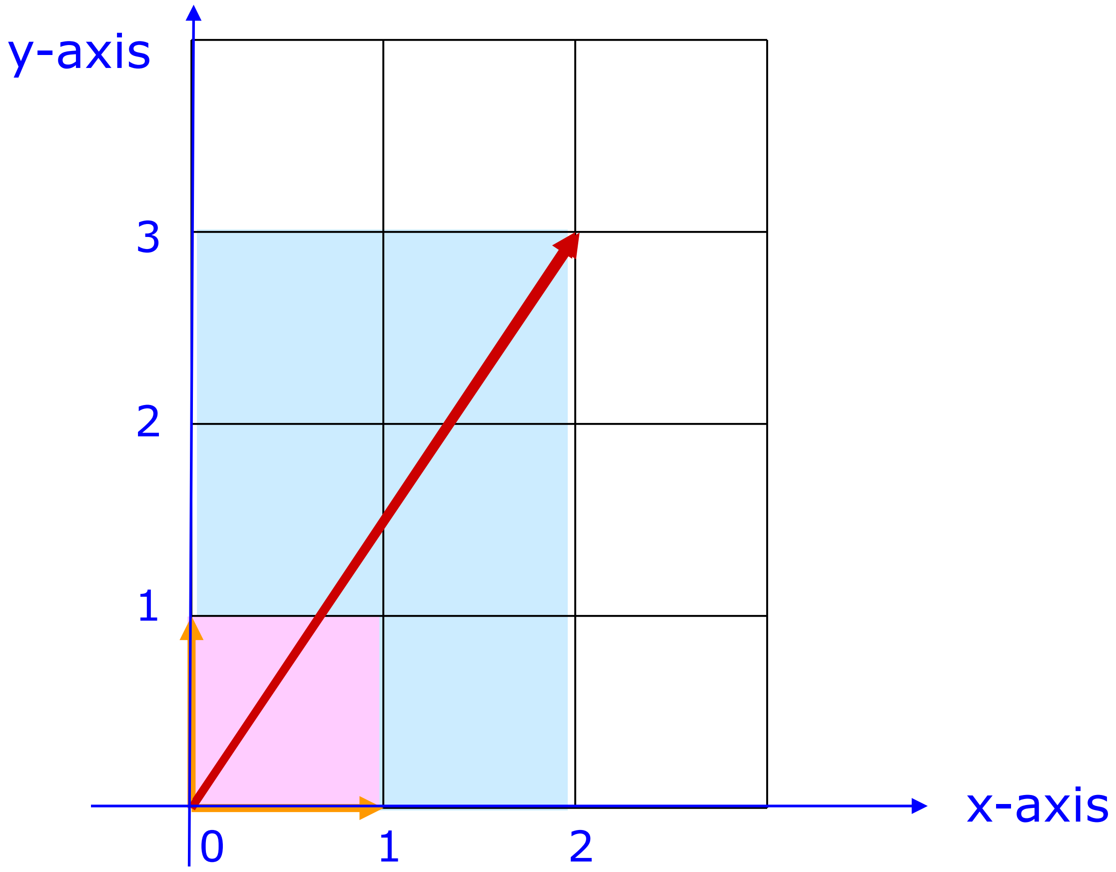
$S_2 = \{(1, -1), (1, 1)\}$

$S_3 = \{(1, 0), (1, 1)\}$

$S_1 = \{(1, 0), (0, 1)\}$

Example 3.5.9.2(a)

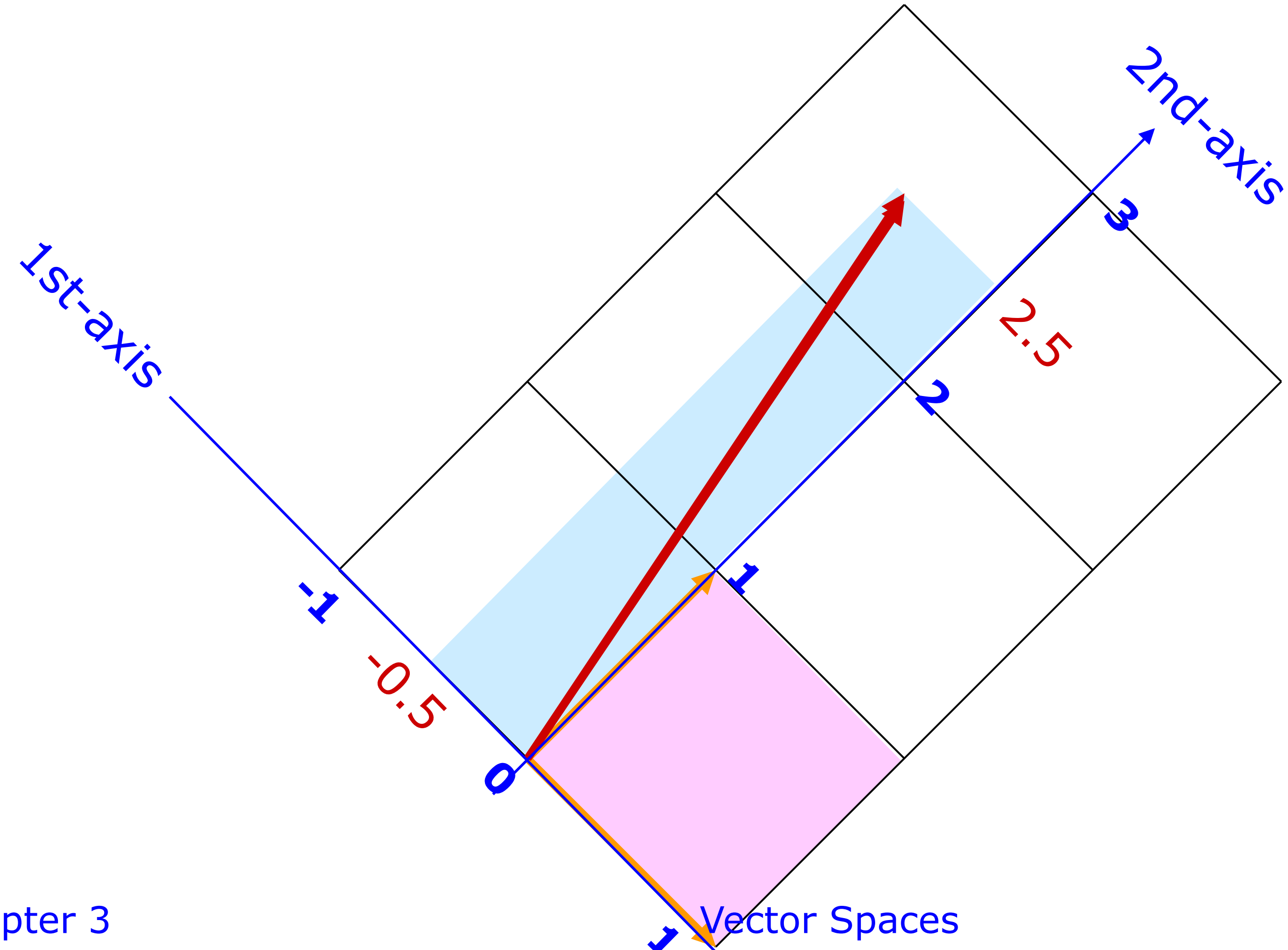
$$\boldsymbol{v} = (2, 3) = 2(1, 0) + 3(0, 1) \Rightarrow \quad (\boldsymbol{v})_{S_1} = (2, 3)$$



$S_2 = \{(1, -1), (1, 1)\}$

Example 3.5.9.2(b)

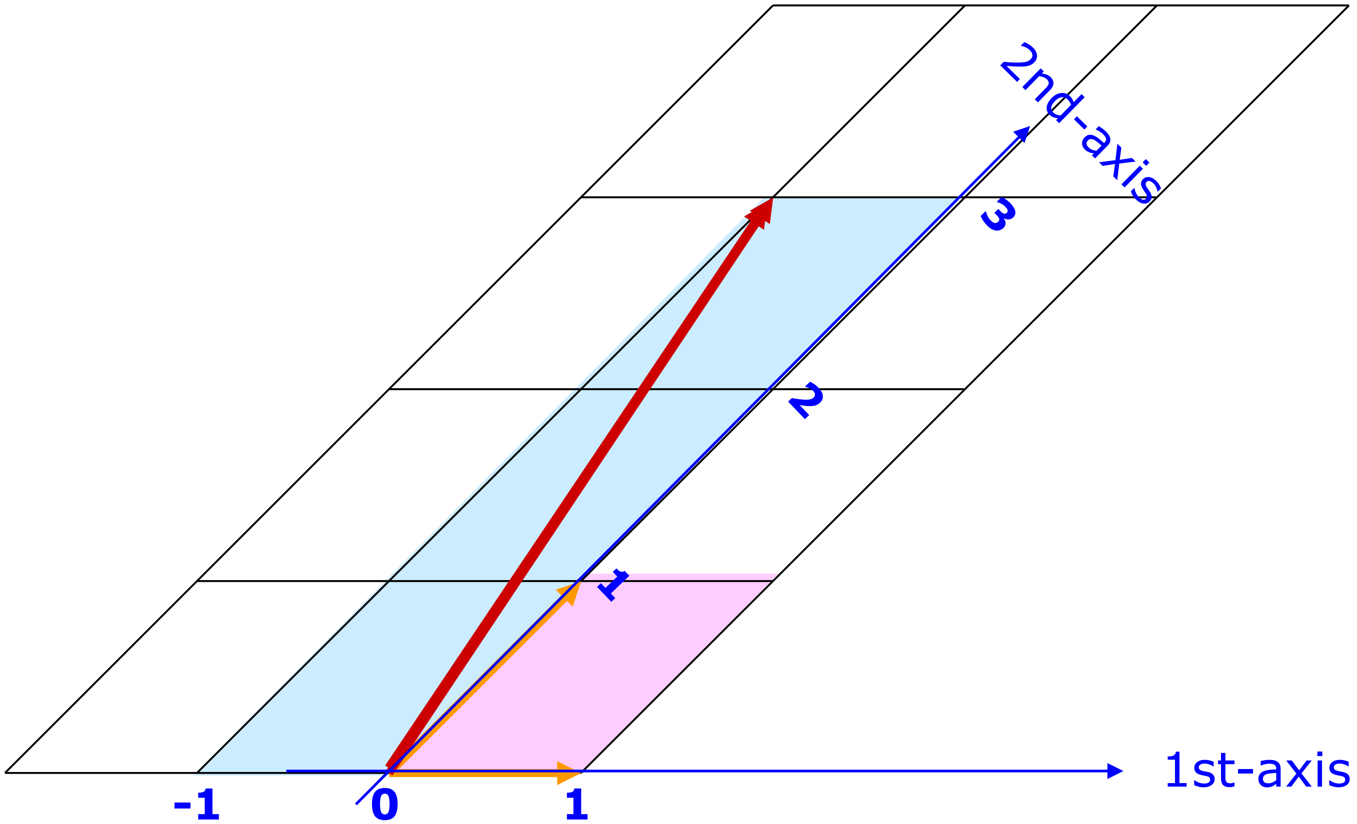
$$\boldsymbol{v} = (2, 3) = -\tfrac{1}{2}(1, -1) + \tfrac{5}{2}(1, 1) \Rightarrow (\boldsymbol{v})_{S_2} = \left(-\tfrac{1}{2}, \tfrac{5}{2}\right)$$



$S_3 = \{(1, 0), (1, 1)\}$

Example 3.5.9.2(c)

$$\boldsymbol{v} = (2, 3) = -(1, 0) + 3(1, 1) \Rightarrow \quad (\boldsymbol{v})_{S_3} = (-1, 3)$$



Coordinate vectors with respect to standard basis

Example 3.5.9.3 (Standard Basis for $\boldsymbol{R}^n$)

If S is the standard basis for $\boldsymbol{R}^n$, then $(\boldsymbol{u})_S = \boldsymbol{u}$

$$S = \{\boldsymbol{e_1}, \boldsymbol{e_2}, \dots, \boldsymbol{e_n}\}$$

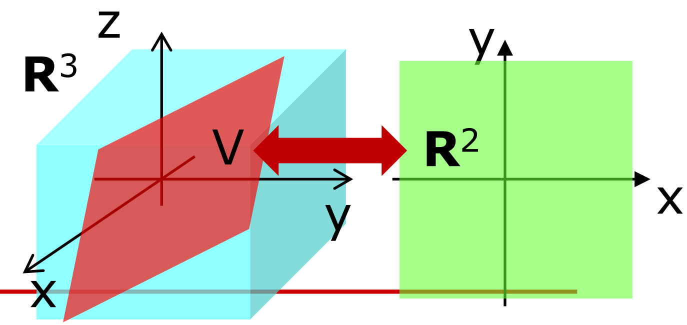
$$\boldsymbol{e_1} = (1, 0, 0, \dots, 0, 0)$$
$$\boldsymbol{e_2} = (0, 1, 0, \dots, 0, 0)$$
$$\dots$$
$$\boldsymbol{e_n} = (0, 0, 0, \dots, 0, 1)$$

For a vector $\boldsymbol{u} = (u_1, u_2, \dots, u_n)$ in $\mathbf{R}^n$

$$\boldsymbol{u} = u_1\boldsymbol{e_1} + u_2\boldsymbol{e_2} + \cdots + u_n\boldsymbol{e_n}$$

$$(\boldsymbol{u})_S = (u_1, u_2, \dots, u_n)$$

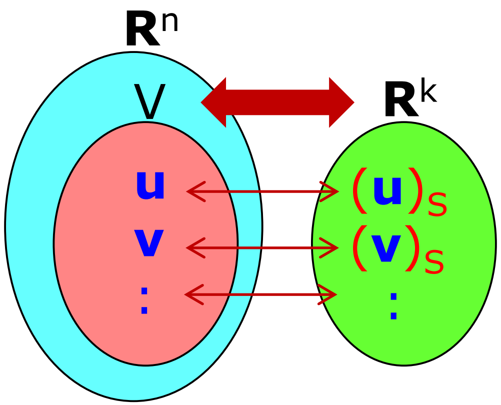
Properties of coordinate vectors



Remark 3.5.10

Some useful rules about coordinate vectors:

Let S be a basis for a vector space V.



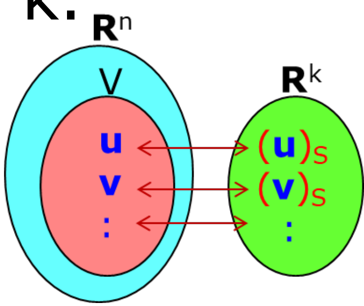
1. For any $\mathbf{u}, \mathbf{v} \in V$,
   $\mathbf{u} = \mathbf{v}$ if and only if $(\mathbf{u})_S = (\mathbf{v})_S$.

2. For any $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \in V$ and $c_1, c_2, \ldots, c_r \in \mathbf{R}$,
   $$(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_r\mathbf{v}_r)_S = c_1(\mathbf{v}_1)_S + c_2(\mathbf{v}_2)_S + \ldots + c_r(\mathbf{v}_r)_S.$$

coordinate vector of linear combination
= linear combination of coordinate vectors

Theorem 3.5.11

S be a basis for a vector space V with |S| = k.



Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \in V$. Then

1. $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ are linearly dependent (resp. independent) in V if and only if $(\mathbf{v}_1)_S, (\mathbf{v}_2)_S, \ldots, (\mathbf{v}_r)_S$ are linearly dependent (resp. independent) in $\mathbf{R}^k$;
2. $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\} = V$ if and only if $\text{span}\{(\mathbf{v}_1)_S, (\mathbf{v}_2)_S, \ldots, (\mathbf{v}_r)_S\} = \mathbf{R}^k$.

$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a basis for V
if and only if
$\{(\mathbf{v}_1)_S, (\mathbf{v}_2)_S, \ldots, (\mathbf{v}_r)_S\}$ is a basis for $\mathbf{R}^k$

Section 3.6

Dimensions

Objective

- What is the dimension of a vector space?
- How to compute dimension for a vector space?
- What are some equivalent conditions for a set to be a basis for a vector space?

Number of vectors in a basis

Theorem 3.6.1

any vector space also can dunnid to only be R

Let V be a vector space which has a basis $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$ with k vectors.

1. Any subset of V with more than k vectors is always linearly dependent.
2. Any subset of V with less than k vectors cannot span V.

Recall Thm 3.4.7:

both trying to say that if == n then can be a basis

Any subset of $\mathbf{R}^n$ with more than n vectors is linearly dep.

Recall Thm 3.2.7:

Any subset of $\mathbf{R}^n$ with less than n vectors cannot span $\mathbf{R}^n$.

Number of vectors in a basis

Theorem 3.6.1 & Remark 3.6.2

Let V be a vector space which has a basis $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ with k vectors.

1. Any subset of V with more than k vectors is always linearly dependent.
2. Any subset of V with less than k vectors cannot span V.

> k : too many vectors to be a basis

< k : too few vectors to be a basis

All bases for a vector space have the same number of vectors

What is dimension of a vector space?

Definition 3.6.3

The dimension of a vector space V
denoted by $\dim(V)$
is the number of vectors in a basis for V.

Recall:
The basis for zero space is defined to be the empty set.
The number of vector in this “basis” is 0.

$\dim(\{\mathbf{0}\}) = 0$

Geometrical meaning of dimension

Example 3.6.4.1-3

1. The dimension of $\mathbf{R}^n$ is n,

   i.e. $\dim(\mathbf{R}^n) = n$.

2. Except $\{\mathbf{0}\}$ and $\mathbf{R}^2$, all subspaces of $\mathbf{R}^2$ are

   lines through the origin span$\{\mathbf{u}\}$

   they are of dimension 1.

3. Except $\{\mathbf{0}\}$ and $\mathbf{R}^3$, all subspaces of $\mathbf{R}^3$ are

   either lines through the origin span$\{\mathbf{u}\}$

   they are of dimension 1,

   or planes containing the origin, span$\{\mathbf{u}, \mathbf{v}\}$

   they are of dimension 2.

Number of vectors in a basis → dimension of the vector space

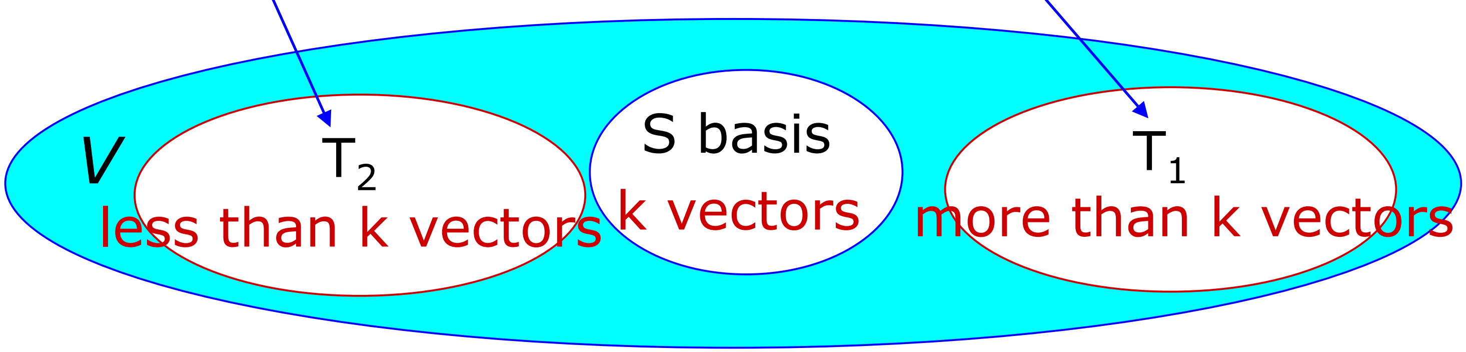
Theorem 3.6.1

Let V be a vector space which has a basis $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$ with k vectors.

dim V = k

1. Any subset of V with more than k vectors is always linearly dependent.

2. Any subset of V with less than k vectors cannot span V.



Finding dimension of a subspace

Example 3.6.4.4

Not the same as the "dimension" of the vectors in the subspace

Find a basis for and determine the dimension of the subspace $W = \{\boxed{(x, y, z)} \mid y = z\}$ of $\mathbf{R}^3$.

Note: $\dim(W) \neq 3$

Explicit: $(x, y, y) = x(1, 0, 0) + y(0, 1, 1)$

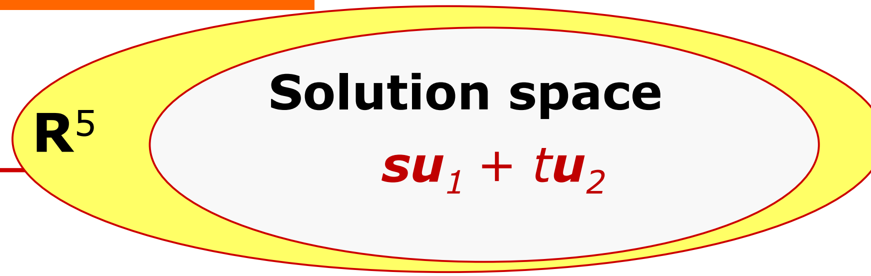
So $W = \text{span}\underbrace{\{(1, 0, 0), (0, 1, 1)\}}_{\text{linearly independent}}$

basis for W : $\{(1, 0, 0), (0, 1, 1)\}$

$\dim(W) = 2$

Dimension of solution space

Example 3.6.6



Find a basis for and determine the dimension of the solution space of the homogeneous system

$$\begin{cases} 2v + 2w - x + z = 0 \\ -v - w + 2x - 3y + z = 0 \\ x + y + z = 0 \\ v + w - 2x - z = 0 \end{cases} \xrightarrow{\text{general solution}} \begin{pmatrix} v \\ w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -s-t \\ s \\ -t \\ 0 \\ t \end{pmatrix} = s \underbrace{\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\boldsymbol{u}_1} + t \underbrace{\begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}}_{\boldsymbol{u}_2}$$

linearly indep.

solution space = span$\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$

basis for the solution space = $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$

dim(solution space) = 2

no. of parameters in the general solution

Discussion 3.6.5 (Example)

homogeneous system with 6 variables: *u, v, w, x, y, z*

Gaussian Elimination

general solution with 4 parameters: *s, t, r, q*

$$\begin{pmatrix} u \\ v \\ w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} s+5t-3r \\ s \\ t \\ 2r-3q \\ r \\ q \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 5 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + r\begin{pmatrix} -3 \\ 0 \\ 0 \\ 2 \\ 1 \\ 0 \end{pmatrix} + q\begin{pmatrix} 0 \\ 0 \\ 0 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

$$s\boldsymbol{u}_1 + t\boldsymbol{u}_2 + r\boldsymbol{u}_3 + q\boldsymbol{u}_4$$

linearly independent

the solution space= span$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$

$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$ is a basis for the solution space

dim(solution space) = 4

Discussion 3.6.5

homogeneous system $\longrightarrow$ row echelon form ***R***

number of non-pivot columns in ***R***

||

number of parameters in general solution

||

number of vectors in basis for solution space

||

the dimension of the solution space

Showing a set form a basis (alternative ways)

Theorem 3.6.7

Let V be a vector space of dimension k and S a subset of V.

The following are equivalent:

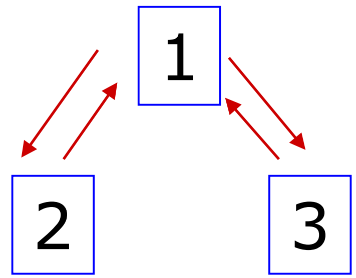
1. S is a basis for V
2. S is linearly independent and $|S| = k$ $= \dim(V)$
3. S spans V and $|S| = k$ $= \dim(V)$

To show S is a basis for V :

S lin. indep S spans V	or	S lin. indep $\lvert S \rvert = \dim V$	or	S spans V $\lvert S \rvert = \dim V$

The proof

1. S is a basis for V.
2. S is lin indep and $|S| = k$.
3. S spans V and $|S| = k$.



Theorem 3.6.7

"$1 \Rightarrow 2$" and "$1 \Rightarrow 3$" is immediate.

$2 \Rightarrow 1$: (prove by contradiction)

Assume that S is not a basis for V

Given S is linearly independent and $|S| = k$.

So $\text{span}(S) \neq V$.

There is a vector $\boldsymbol{u}$ in V and $\boldsymbol{u} \notin \text{span}(S)$.

$\boldsymbol{u}$ is not redundant in span(S)

Let $S' = S \cup \{\boldsymbol{u}\}$ $\Rightarrow S'$ is linearly indep.

$k + 1$ vectors

Contradiction

see Theorem 3.4.10

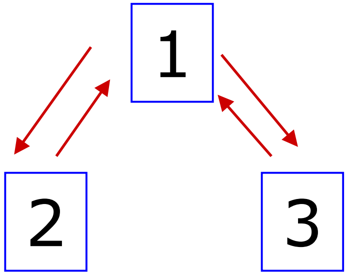
$\Rightarrow S'$ is linearly dep.

see Theorem 3.6.1.1

So S is a basis for V

The proof

1. S is a basis for V.
2. S is lin indep and $|S| = k$.
3. S spans V and $|S| = k$.



Theorem 3.6.7

$3 \Rightarrow 1$: (prove by contradiction)

Assume S not a basis for V

Given S spans V and $|S| = k$.

S is linearly dependent.

There is a redundant vector $\boldsymbol{v}$ in S.

see Theorem 3.2.12

Let $S'' = S - \{\boldsymbol{v}\}$ $\Rightarrow \text{span}(S'') = \text{span}(S) = V$

$k - 1$ vectors

Contradiction

$\Rightarrow \text{span}(S'') \neq V$ see Theorem 3.6.1.2

So S is a basis for V

Example 3.6.8

Show that

$\boldsymbol{u_1} = (2, 0, -1)$, $\boldsymbol{u_2} = (4, 0, 7)$ and $\boldsymbol{u_3} = (-1, 1, 4)$

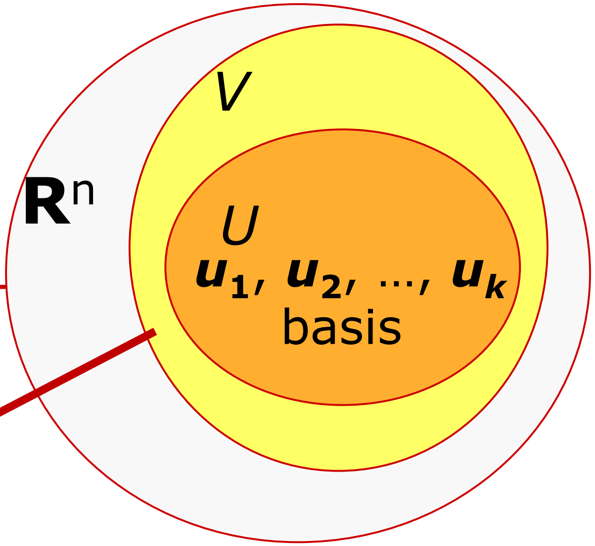
form a basis for $\mathbf{R}^3$.

Since dim $\mathbf{R}^3 = 3$,
we only need to show the set of 3 vectors is
either linear independent or spans $\mathbf{R}^3$.

If we don't know the dimension of a vector space *V*,
to show a set is a basis for *V*,
we still need to check
the set is both linear independent and spans *V*.

Dimensions give the “size” of subspaces of $\mathbf{R}^n$

Theorem 3.6.9



Let U and V be subspaces of $\mathbf{R}^n$

We say: U is a subspace of V.

(i) If $U \subseteq V$, then $\dim(U) \leq \dim(V)$

(ii) If $U \subseteq V$ and $U \neq V$, then $\dim(U) < \dim(V)$

For (i), $\dim(U) = k$

$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are k lin. indep. vectors in V

So $k \leq \dim(V)$

For (ii), suppose $\dim(U) = \dim(V)$

Then $\dim(V) = k$

So $V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} = U$.

contradiction

Dimensions give the "size" of subspaces of $\mathbf{R}^n$

Example 3.6.10

Given V a plane in $\mathbf{R}^3$ containing the origin.

Suppose U is a subspace of V such that $U \neq V$.
What can we say about U?

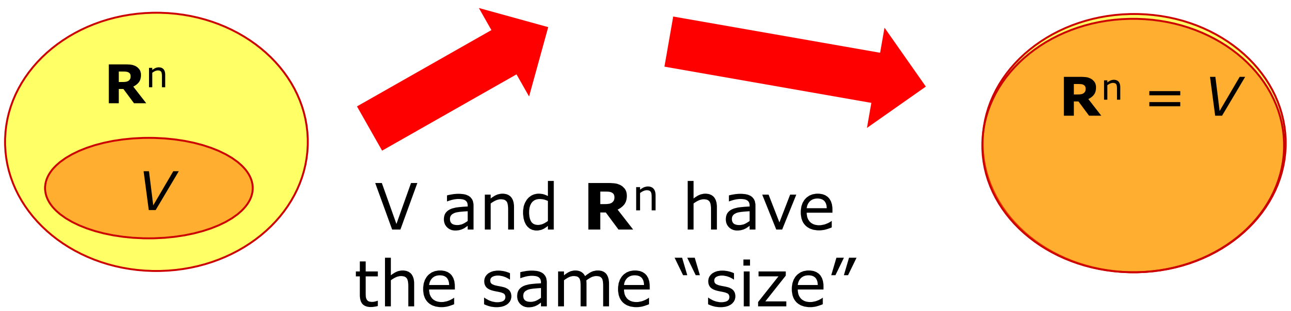
V is of dimension 2.
By Theorem 3.6.9, $\dim(U) < 2$.
So
either $\dim(U) = 0 \quad \Leftrightarrow U = \{\mathbf{0}\}$
or $\dim(U) = 1 \quad \Leftrightarrow U =$ a line through the origin

True or False

Let U and V be subspaces of $\mathbf{R}^n$

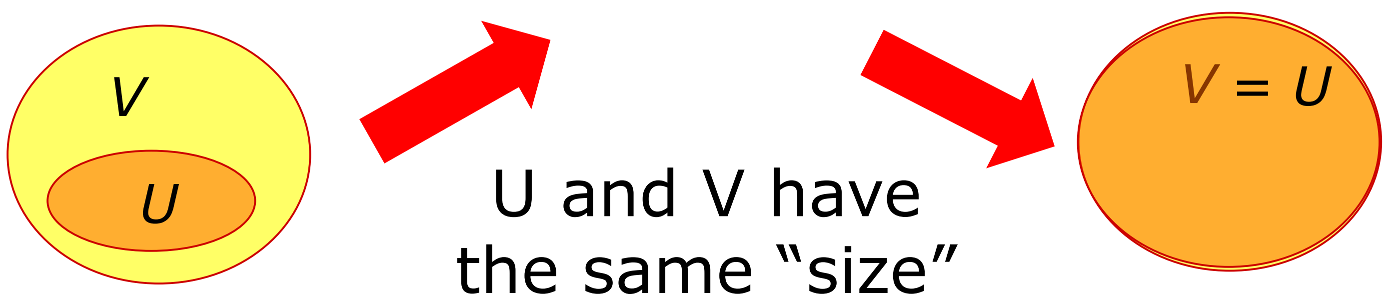
No subspace of $\mathbf{R}^n$ has dimension n, except $\mathbf{R}^n$ itself.

A. If $\dim(V) = n$, then $V = \mathbf{R}^n$ True



Theorem 3.6.9.2

B. If $U \subseteq V$ and $\dim(U) = \dim(V)$, then $U = V$ True



A very^3 important theorem (revisited)

Theorem 3.6.11

A is an n x n matrix.

The following statements are equivalent:

1. ***A*** is invertible.
2. The linear system $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution.
3. The reduced row-echelon form of ***A*** is an identity matrix.
4. ***A*** can be expressed as a product of elementary matrices.
5. $\det(\boldsymbol{A}) \neq 0$.
6. The rows of ***A*** form a basis for $\mathbf{R}^n$.
7. The columns of ***A*** form a basis for $\mathbf{R}^n$.

1. $\boldsymbol{A}$ is invertible

2. $\boldsymbol{Ax} = \mathbf{0}$ has only trivial solution

7. The columns of $\boldsymbol{A}$ form a basis for $\mathbf{R}^n$

Example

Suppose we know $\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$ is invertible.

Then we know that the linear system

$$\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

has only trivial solution

Write the linear system in vector equation form:

$$x\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + y\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

has only zero coefficients

$x = y = z = 0$

We conclude that $\left\{\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right\}$ is linearly independent

hence form a basis for $\mathbf{R}^3$

1. $\boldsymbol{A}$ is invertible
5. det $\boldsymbol{A} \neq 0$
7. The columns of $\boldsymbol{A}$ form a basis for $\mathbf{R}^n$
6. The rows of $\boldsymbol{A}$ form a basis for $\mathbf{R}^n$

Example

Suppose we know $\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$ is invertible.

Then we know that the determinant $\begin{vmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{vmatrix} \neq 0$

Then the transpose determinant $\begin{vmatrix} 1 & 3 & 2 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} \neq 0$

So $\begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ is invertible.

So the columns $\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\}$ form a basis for $\mathbf{R}^3$

So the rows $\{(1\ 2\ 1), (3\ 1\ 0), (2\ 0\ 1)\}$ form a basis for $\mathbf{R}^3$

Alternative method to check basis for $\mathbf{R}^n$

Example 3.6.12 (Determinant method)

$\boldsymbol{u_1} = (1, 1, 1),\ \boldsymbol{u_2} = (-1, 1, 2),\ \boldsymbol{u_3} = (1, 0, 1)$

Is $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ a basis for $\mathbf{R}^3$? YES

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & 1 & 2 \\ 1 & 0 & 1 \end{vmatrix} = 3 \neq 0$$

$\boldsymbol{u_1} = (1, 1, 1, 1),\ \boldsymbol{u_2} = (1, -1, 1, -1),$
$\boldsymbol{u_3} = (0, 1, -1, 0),\ \boldsymbol{u_4} = (2, 1, 1, 0)$

Is $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}, \boldsymbol{u_4}\}$ a basis for $\mathbf{R}^4$? NO

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

Cannot use this method to check basis for subspaces of $\mathbf{R}^n$

Section 3.7

Transition Matrices

Objective

- What is a transition matrix?
- How to compute transition matrices?
- What is the relation between coordinate vectors w.r.t. different bases?

From one basis to another

Example 3.7.4.1

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ basis for $\mathbf{R}^3$

$\boldsymbol{u}_1 = (1, 0, -1),\ \boldsymbol{u}_2 = (0, -1, 0),\ \boldsymbol{u}_3 = (1, 0, 2).$

$T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ basis for $\mathbf{R}^3$

$\boldsymbol{v}_1 = (1, 1, 1),\ \boldsymbol{v}_2 = (1, 1, 0),\ \boldsymbol{v}_3 = (-1, 0, 0).$

Given $(\boldsymbol{w})_S = (2, -1, 2)$. Find $(\boldsymbol{w})_T$.

2.3 meter $\longrightarrow$ 7.5 feet

$$\begin{tikzcd}
(\boldsymbol{w})_S \arrow[r] \arrow[rr, bend right] & \boldsymbol{w} \arrow[r] & (\boldsymbol{w})_T
\end{tikzcd}$$

Is there a direct method?

Coordinate vector notation: $(\boldsymbol{v})_S$ & $[\boldsymbol{v}]_S$

Notation 3.7.1

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$: a basis for a vector space V

$\boldsymbol{v}$: a vector in V

Write $\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}$

Then $(\boldsymbol{v})_S = (c_1, c_2, \ldots, c_k)$ row form of coordinate vector

$$[\boldsymbol{v}]_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$

column form of coordinate vector

We need to pre-multiply the coordinate-vector by a $k \times k$ matrix

From one basis to another

Discussion 3.7.2

$$[\boldsymbol{w}]_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix} \qquad [\boldsymbol{w}]_T = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{pmatrix}$$

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$ and $T = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}\}$ two bases for a vector space V.

Take a vector $\boldsymbol{w}$ in V

Relation between $[\boldsymbol{w}]_S$ and $[\boldsymbol{w}]_T$?

$\boldsymbol{w}$ in terms of $\boldsymbol{u}_i$

$\boldsymbol{w}$ in terms of $\boldsymbol{v}_i$

We will show that

does not depend on $\boldsymbol{w}$

$[\boldsymbol{w}]_T = \boldsymbol{P}\,[\boldsymbol{w}]_S$ for some fixed k x k matrix $\boldsymbol{P}$

transition matrix

Finding transition matrix from S to T

Definition 3.7.3

Read Discussion 3.7.2 to see why it works

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$ and $T = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}\}$

two bases for a vector space V.

1. Express each $\boldsymbol{u_i}$ as linear combination of $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}\}$
2. Form the (column) coordinate vectors w.r.t. T

$$[\boldsymbol{u}_1]_T = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{k1} \end{pmatrix} \quad [\boldsymbol{u}_2]_T = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{pmatrix} \quad \cdots \quad [\boldsymbol{u}_k]_T = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{kk} \end{pmatrix}$$

3. Form the matrix $\mathbf{P} = ([\boldsymbol{u_1}]_T \ [\boldsymbol{u_2}]_T \ \ldots \ [\boldsymbol{u_k}]_T)$

$$P = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}$$

transition matrix from S to T

4. $\mathbf{P}\,[\boldsymbol{w}]_S = [\boldsymbol{w}]_T$ for any vector $\boldsymbol{w}$ in V.

From one basis to another

Example 3.7.4.1

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ basis for $\mathbf{R}^3$

$\boldsymbol{u_1} = (1, 0, -1)$, $\boldsymbol{u_2} = (0, -1, 0)$, $\boldsymbol{u_3} = (1, 0, 2)$.

$T = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ basis for $\mathbf{R}^3$

$\boldsymbol{v_1} = (1, 1, 1)$, $\boldsymbol{v_2} = (1, 1, 0)$, $\boldsymbol{v_3} = (-1, 0, 0)$.

(a) Find the transition matrix from S to T.

$$P = \left([\boldsymbol{u}_1]_T \quad [\boldsymbol{u}_2]_T \quad [\boldsymbol{u}_3]_T\right)$$

(b) $\boldsymbol{w}$ a vector in $\mathbf{R}^3$ with $(\boldsymbol{w})_S = (2, -1, 2)$.

Find $(\boldsymbol{w})_T$.

$$[\boldsymbol{w}]_T = P\,[\boldsymbol{w}]_S$$

Finding transition matrix

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$
$T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$

Example 3.7.4.1(a)

$$\boldsymbol{u}_1 = a_{11}\boldsymbol{v}_1 + a_{21}\boldsymbol{v}_2 + a_{31}\boldsymbol{v}_3$$
$$\boldsymbol{u}_2 = a_{12}\boldsymbol{v}_1 + a_{22}\boldsymbol{v}_2 + a_{32}\boldsymbol{v}_3$$
$$\boldsymbol{u}_3 = a_{13}\boldsymbol{v}_1 + a_{23}\boldsymbol{v}_2 + a_{33}\boldsymbol{v}_3$$

find $a_{11}, a_{21}, \ldots, a_{33}$

Convert to three linear systems:

$$\begin{cases} a_{11} + a_{21} - a_{31} = 1 \\ a_{11} + a_{21} = 0 \\ a_{11} = -1 \end{cases} \quad \begin{cases} a_{12} + a_{22} - a_{32} = 0 \\ a_{12} + a_{22} = -1 \\ a_{12} = 0 \end{cases} \quad \begin{cases} a_{13} + a_{23} - a_{33} = 1 \\ a_{13} + a_{23} = 0 \\ a_{13} = 2 \end{cases}$$

$$\begin{array}{cccccc} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 & \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{array}$$

$$\left(\begin{array}{ccc|c|c|c} 1 & 1 & -1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 & 0 & 2 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \left(\begin{array}{ccc|c|c|c} 1 & 0 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & -1 & -1 & -1 \end{array}\right)$$

$[\boldsymbol{u}_1]_T [\boldsymbol{u}_2]_T [\boldsymbol{u}_3]_T$

transition matrix from S to T

Finding $(\boldsymbol{w})_T$ form $(\boldsymbol{w})_S$

Example 3.7.4.1(b)

$(\boldsymbol{w})_S = (2, -1, 2)$

$$[\boldsymbol{w}]_T = (\text{Transition matrix from S to T})[\boldsymbol{w}]_S$$

$$\boldsymbol{P} = \begin{pmatrix} -1 & 0 & 2 \\ 1 & -1 & -2 \\ -1 & -1 & -1 \end{pmatrix}$$

$$[\boldsymbol{w}]_T = \boldsymbol{P}[\boldsymbol{w}]_S = \begin{pmatrix} -1 & 0 & 2 \\ 1 & -1 & -2 \\ -1 & -1 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ -3 \end{pmatrix}$$

So $(\boldsymbol{w})_T = (2, -1, -3)$.

From S to T and from T to S

Example 3.7.4.2

$\boldsymbol{P}\,[\boldsymbol{w}]_S = [\boldsymbol{w}]_T$ for any vector $\boldsymbol{w}$

$\boldsymbol{Q}\,[\boldsymbol{w}]_T = [\boldsymbol{w}]_S$ for any vector $\boldsymbol{w}$

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ $\quad \boldsymbol{u}_1 = (1, 1),\ \boldsymbol{u}_2 = (1, -1).$

$T = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ $\quad \boldsymbol{v}_1 = (1, 0),\ \boldsymbol{v}_2 = (1, 1).$

two bases for $\mathbf{R}^2$

transition matrix from S to T

$$\begin{cases} \mathbf{u}_1 = 0\mathbf{v}_1 + \mathbf{v}_2 \\ \mathbf{u}_2 = 2\mathbf{v}_1 - \mathbf{v}_2 \end{cases}$$

$$[\mathbf{u}_1]_T = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad [\mathbf{u}_2]_T = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

transition matrix from T to S

$$\begin{cases} \mathbf{v}_1 = \frac{1}{2}\mathbf{u}_1 + \frac{1}{2}\mathbf{u}_2 \\ \mathbf{v}_2 = \mathbf{u}_1 + 0\mathbf{u}_2 \end{cases}$$

$$[\mathbf{v}_1]_S = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \quad [\mathbf{v}_2]_S = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\mathbf{P} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix} \longleftrightarrow \boldsymbol{Q} = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{pmatrix}$$

inverse of each other

The inverse of transition matrix

Theorem 3.7.5

S and T : two bases of a vector space

$\boldsymbol{P}$: the transition matrix from S to T.

1. $\boldsymbol{P}$ is invertible.
2. $\boldsymbol{P}^{-1}$ is the transition matrix from T to S.

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$, $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ bases

$\mathbf{P} = ([\boldsymbol{u}_1]_T \ [\boldsymbol{u}_2]_T \ \ldots \ [\boldsymbol{u}_k]_T) \quad \Rightarrow \boldsymbol{P}$ is invertible

$[\boldsymbol{u}_1]_T \ [\boldsymbol{u}_2]_T \ \ldots \ [\boldsymbol{u}_k]_T$ are linearly independent

Let $\boldsymbol{Q}$ be the transition matrix from T to S.

$\mathbf{Q} = ([\boldsymbol{v}_1]_S \ [\boldsymbol{v}_2]_S \ \ldots \ [\boldsymbol{v}_k]_S)$

To show $\boldsymbol{QP} = \boldsymbol{I}$

The proof: two observations

Theorem 3.7.5

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ basis

Obs. 1

$$[\boldsymbol{u}_1]_S = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad [\boldsymbol{u}_2]_S = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \cdots \quad [\boldsymbol{u}_k]_S = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

standard basis vectors

$$\boldsymbol{u}_1 = 1\boldsymbol{u}_1 + 0\boldsymbol{u}_2 + \ldots + 0\boldsymbol{u}_k$$

Obs. 2

any mxn matrix $\boldsymbol{A}$

$[\boldsymbol{u}_\mathbf{i}]_S$

$$\begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & a_{1i} & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,i-1} & a_{2,i} & a_{2,i+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{m,i-1} & a_{m,i} & a_{m,i+1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ 0 \\ \vdots \\ a_{mj} \end{pmatrix}$$

i^{th} column of $\boldsymbol{A}$

i^{th} coordinate

The proof

$$[\boldsymbol{u}_1]_S = \begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix} \quad [\boldsymbol{u}_2]_S = \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix} \quad \cdots \quad [\boldsymbol{u}_k]_S = \begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix}$$

Theorem 3.7.5

To show $\boldsymbol{QP} = \boldsymbol{I}$

Examine the i^{th} column of $\boldsymbol{QP}$ for $i = 1, 2, \ldots, k$

i^{th} column of $\boldsymbol{A} = \boldsymbol{A}\,[\boldsymbol{u}_i]_S$

$$i^{th} \text{ column of } \boldsymbol{QP} = \boldsymbol{QP}\,[\boldsymbol{u}_i]_S = \boldsymbol{Q}[\boldsymbol{u}_i]_T = [\boldsymbol{u}_i]_S = \begin{pmatrix}0\\0\\1\\\vdots\\0\end{pmatrix}$$

$\boldsymbol{P}$: transition matrix from S to T

$\boldsymbol{Q}$: transition matrix from T to S

$$\boldsymbol{QP} = \begin{pmatrix}1 & 0 & & 0\\0 & 1 & & 0\\0 & 0 & \ddots & 0\\\vdots & \vdots & & \vdots\\0 & 0 & & 1\end{pmatrix} = \boldsymbol{I}$$

So $\boldsymbol{P}$ is invertible and $\boldsymbol{P}^{-1} = \boldsymbol{Q}$

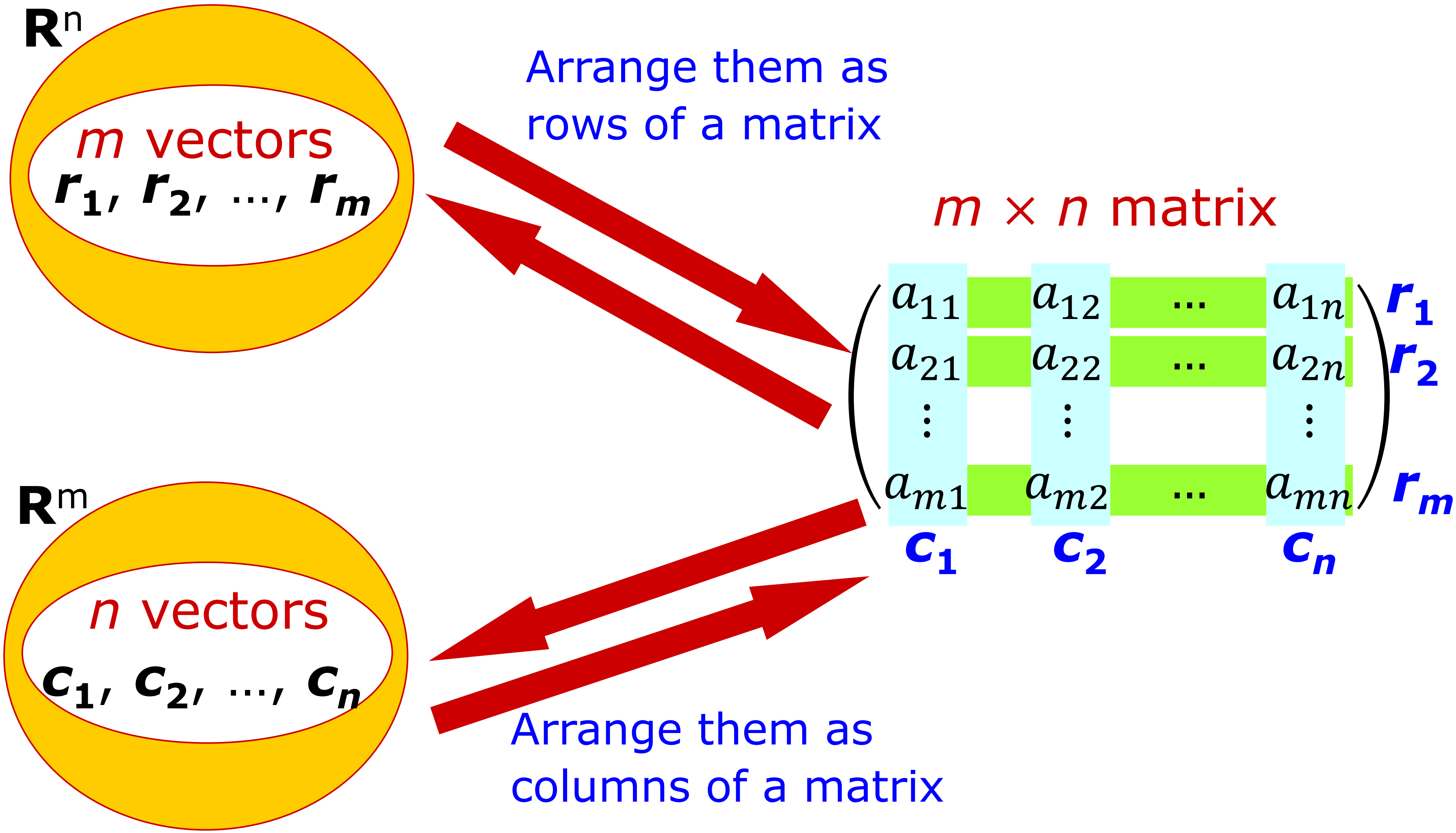
Section 4.1

Row Spaces and Column Spaces

Objectives

- What are row space and column space of a matrix?
- How to find bases for row /column spaces?
- How to use row /column spaces to find bases for vector spaces?
- How to extend a basis?
- What is the relation between column space and consistency of linear system?

Discussion 4.1.1



Example 4.1.4.1

$$\boldsymbol{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & -1 & 3 \\ -5 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

rows of $\boldsymbol{A}$

$\boldsymbol{r_1} = (2, -1, 0)$
$\boldsymbol{r_2} = (1, -1, 3)$
$\boldsymbol{r_3} = (-5, 1, 0)$
$\boldsymbol{r_4} = (1, 0, 1)$

We call it the row space of $\boldsymbol{A}$

span{(2, −1, 0), (1, −1, 3), (−5, 1, 0), (1, 0, 1)} a subspace of $\mathbf{R}^3$

span vector rows to be subspace

columns of $\boldsymbol{A}$

$$\mathbf{c}_1 = \begin{pmatrix} 2 \\ 1 \\ -5 \\ 1 \end{pmatrix} \quad \mathbf{c}_2 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{c}_3 = \begin{pmatrix} 0 \\ 3 \\ 0 \\ 1 \end{pmatrix}$$

We call it the column space of $\boldsymbol{A}$

$$\text{span}\left\{ \begin{pmatrix} 2 \\ 1 \\ -5 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\}$$ a subspace of $\mathbf{R}^4$

Row space and column space

Definition 4.1.2

$$A = \begin{array}{cc} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} & \begin{array}{c} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \\ \boldsymbol{r}_m \end{array} \\ \begin{array}{cccc} \boldsymbol{c}_1 & \boldsymbol{c}_2 & & \boldsymbol{c}_n \end{array} & \end{array}$$

an m x n matrix

The row space of $\boldsymbol{A}$ $= \text{span}\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_m\}$

$\neq$

$\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_m\}$

a subspace of $\mathbf{R}^n$

The column space of $\boldsymbol{A}$ $= \text{span}\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_n\}$

$\neq$

$\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_n\}$

a subspace of $\mathbf{R}^m$

Remark 4.1.3

row space of $\boldsymbol{A}$ = column space of $\boldsymbol{A}^T$
column space of $\boldsymbol{A}$ = row space of $\boldsymbol{A}^T$

$$\begin{array}{ccc} A = & \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} & \begin{array}{c} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \\ \boldsymbol{r}_m \end{array} \\ & \begin{array}{cccc} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \quad & \boldsymbol{c}_n \end{array} & \end{array}$$

$$\begin{array}{ccc} A^{\mathrm{T}} = & \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} & \begin{array}{c} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ \\ \boldsymbol{c}_n \end{array} \\ & \begin{array}{cccc} \boldsymbol{r}_1 & \boldsymbol{r}_2 & \quad & \boldsymbol{r}_m \end{array} & \end{array}$$

Some special matrices

Row (column) space of zero matrix ***0*** = zero space

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

Row (column) space of n x n identity matrix $\boldsymbol{I}_n$ = $\mathbf{R}^n$

$$\boldsymbol{I}_n = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

Bases for row space and column space

Example 4.1.4.2

$$\boldsymbol{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & -1 & 3 \\ -5 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Find a basis and the dimension for the row space

span$\{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{r}_4\}$

basis = $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{r}_4\}$?

Find a basis and the dimension for the column space

span$\{\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3\}$

basis = $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3\}$?

These sets may be linearly dependent

There may be redundant vectors

not necessary

Row equivalent matrices

Discussion 4.1.6

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be row equivalent matrices.

$$\boldsymbol{A} \to \to \dots \to \boldsymbol{B}$$

Row equivalence (r.e.) is an equivalence relation on matrices of the same size

- $\boldsymbol{A}$ is r.e. to itself
- If $\boldsymbol{A}$ is r.e. to $\boldsymbol{B}$, then $\boldsymbol{B}$ is r.e. to $\boldsymbol{A}$
- If $\boldsymbol{A}$ is r.e. to $\boldsymbol{B}$, and $\boldsymbol{B}$ is r.e. to $\boldsymbol{C}$, then $\boldsymbol{A}$ is r.e. to $\boldsymbol{C}$.

If two matrices **M** and **N** (of the same size) have the same reduced row echelon form, then **M** and **N** are row equivalent.

Row equivalent matrices have same row space

Theorem 4.1.7

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be row equivalent matrices.

Then

row space of $\boldsymbol{A}$ = row space of $\boldsymbol{B}$

elementary row operations

change the rows of a matrix

but do not change the row space of a matrix.

Idea of proof

Theorem 4.1.7

Let $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ be rows of a matrix.

We need to show that

1. $\text{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_i, \ldots, \boldsymbol{a}_n\}$
$= \text{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, c\boldsymbol{a}_i, \ldots, \boldsymbol{a}_n\}$

2. $\text{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_i, \ldots, \boldsymbol{a}_j, \ldots, \boldsymbol{a}_n\}$
$= \text{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_j, \ldots, \boldsymbol{a}_i, \ldots, \boldsymbol{a}_n\}$

3. $\text{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_i, \ldots, \boldsymbol{a}_n\}$
$= \text{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_i + c\boldsymbol{a}_j, \ldots, \boldsymbol{a}_n\}$

Row equivalent matrices have same row space

Example 4.1.8.1

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 4 \\ \frac{1}{2} & 1 & 2 \end{pmatrix} \quad \boldsymbol{B} = \begin{pmatrix} \frac{1}{2} & 1 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{pmatrix} \quad \boldsymbol{C} = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{pmatrix} \quad \boldsymbol{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{A} \xrightarrow{R_1 \leftrightarrow R_3} \boldsymbol{B} \xrightarrow{2R_1} \boldsymbol{C} \xrightarrow{R_1 - R_2} \boldsymbol{D}$$

A, ***B***, ***C***, ***D*** are row equivalent to one another

So their row spaces are all the same

In particular

span{(0, 0, 1), (0, 2, 4), (1/2, 1, 2)} row space of ***A***

= span{(1, 0, 0), (0, 2, 4), (0, 0, 1)}. row space of ***D***

Finding basis for row space

Example 4.1.8.2

$$A = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & -2 & 0 & -1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} R = \begin{array}{c} \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ \boxed{0 \quad 0} & \frac{3}{2} & -3 & \frac{3}{2} \\ \boxed{0 \quad 0 \quad 0} & & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{l} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \boldsymbol{r}_3 \\ \\ \end{array}$$

row echelon form

The row space of ***A*** **=** The row space of ***R***

span$\{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{0}\}$

span$\{(2,2,-1,0,1),\ (0,0,\frac{3}{2},-3,\frac{3}{2}),\ (0,0,0,3,0)\}$

The three non-zero rows $\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3$ of ***R*** are linearly indep.

So $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3\}$ is a basis for the row space of ***A***

Finding basis for row space

Remark 4.1.9

$\boldsymbol{A} \longrightarrow \boldsymbol{R}$ (row-echelon form)

The set of nonzero rows of $\boldsymbol{R}$ $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_k\}$

is a basis for the row space of $\boldsymbol{A}$.

spans the row space of $\boldsymbol{R}$

spans the row space of $\boldsymbol{A}$

linearly independent

Note that this basis may not contain the original rows of $\boldsymbol{A}$

Discussion 4.1.10

Can we take the non-zero columns of a row-echelon form to form a basis for the column space?

Is this a basis for the column space of ***A***?

$$\boldsymbol{A} = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & -2 & 0 & -1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \boldsymbol{R} = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ 0 & 0 & \frac{3}{2} & -3 & \frac{3}{2} \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

not **linearly indep**

BAD NEWS: Row equivalent matrices may have different column spaces

Discussion 4.1.10

Elementary row operations may not preserve the column space of a matrix.

$\boldsymbol{A} \to \to \ldots \to \boldsymbol{B}$

row sp A = row sp B
col. sp A ≠ col. sp B

For example, $\boldsymbol{A} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

A and ***B*** are row equivalent but their column spaces are different.

The column space of **A** = span$\left\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$

The column space of **B** = span$\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$

GOOD NEWS: Row equivalent matrices preserve linear dependency of the columns

Example 4.1.12.1

$$\boldsymbol{A} = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & -2 & 0 & -1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \boldsymbol{R} = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ 0 & 0 & \frac{3}{2} & -3 & \frac{3}{2} \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

1. The 1st, 3rd and 5th columns of $\boldsymbol{R}$ are linearly dependent.

   Correspondingly,
   the 1st, 3rd and 5th columns of $\boldsymbol{A}$ are linearly dependent.

GOOD NEWS: Row equivalent matrices preserve linear dependency of the columns

Example 4.1.12.2

$$\boldsymbol{A} = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & -2 & 0 & -1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \boldsymbol{R} = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ 0 & 0 & \frac{3}{2} & -3 & \frac{3}{2} \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

2. The 1st, 3rd and 4th columns of $\boldsymbol{R}$ are linearly independent.

   Correspondingly,
   the 1st, 3rd and 4th columns of $\boldsymbol{A}$ are linearly independent.

Row equivalent matrices preserve linear dependency of the columns

Theorem 4.1.11

$$A \xleftrightarrow{\text{row equivalent}} B$$

column space of $\boldsymbol{A}$ | may not be equal | column space of $\boldsymbol{B}$

A set of columns of $\boldsymbol{A}$ is linearly independent $\Leftrightarrow$ corresponding columns of $\boldsymbol{B}$ are linearly independent

linearly dependent | linearly dependent

a column of $\boldsymbol{A}$ is redundant | corresponding column of $\boldsymbol{B}$ is redundant

A set of columns of $\boldsymbol{A}$ form a basis for the column space of $\boldsymbol{A}$ $\Leftrightarrow$ corresponding columns of $\boldsymbol{B}$ form a basis for the column space of $\boldsymbol{B}$

Finding basis for column space

Example 4.1.12.2

linearly indep.
pivot columns

$$\boldsymbol{A} = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & -2 & 0 & -1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \boldsymbol{R} = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ 0 & 0 & \frac{3}{2} & -3 & \frac{3}{2} \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

non-pivot columns
redundant

The 1st, 3rd and 4th columns of $\boldsymbol{R}$ form a basis for the column space of $\boldsymbol{R}$.

Correspondingly,
the 1st, 3rd and 4th columns of $\boldsymbol{A}$ form a basis for the column space of $\boldsymbol{A}$.

Remark 4.1.13

$\boldsymbol{A} \longrightarrow \boldsymbol{R}$ (row-echelon form)

pivot columns

$\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_k\}$

$\{\boldsymbol{c}_1^*, \boldsymbol{c}_2^*, \ldots, \boldsymbol{c}_k^*\}$

columns of $\boldsymbol{A}$ that correspond to the pivot columns in $\boldsymbol{R}$.

basis for the column space of $\boldsymbol{A}$

basis for the column space of $\boldsymbol{R}$

may not be basis for the column space of $\boldsymbol{A}$

Idea of proof of Theorem 4.1.11

Remark

row operations preserve linear relations among columns

$$\boxed{C_1 + C_3 = C_4}$$

$$\begin{array}{cccc} C_1 & C_2 & C_3 & C_4 \\ \end{array}$$
$$\begin{pmatrix} a & \# & b & a+b \\ c & \# & d & c+d \\ e & \# & f & e+f \end{pmatrix} \xrightarrow[\text{Type 1}]{kR_1} \begin{pmatrix} ka & \# & kb & k(a+b) \\ c & \# & d & c+d \\ e & \# & f & e+f \end{pmatrix}$$

$R_1 \leftrightarrow R_3$ Type 2:

$$\begin{array}{cccc} C_1 & C_2 & C_3 & C_4 \\ \end{array}$$
$$\begin{pmatrix} e & \# & f & e+f \\ c & \# & d & c+d \\ a & \# & b & a+b \end{pmatrix}$$

$R_2 + kR_1$ Type 3:

$$\begin{array}{cccc} C_1 & C_2 & C_3 & C_4 \\ \end{array}$$
$$\begin{pmatrix} a & \# & b & a+b \\ c+ka & \# & d+kb & (c+d)+k(a+b) \\ e & \# & f & e+f \end{pmatrix}$$

Application: finding basis for linear span

Example 4.1.14.1

Find a basis for span$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4, \boldsymbol{u}_5, \boldsymbol{u}_6\}$

$$\boldsymbol{u}_1 = (1, 2, 2, 1)$$
$$\boldsymbol{u}_2 = (3, 6, 6, 3)$$
$$\boldsymbol{u}_3 = (4, 9, 9, 5)$$
$$\boldsymbol{u}_4 = (-2, -1, -1, 1)$$
$$\boldsymbol{u}_5 = (5, 8, 9, 4)$$
$$\boldsymbol{u}_6 = (4, 2, 7, 3)$$

Arrange the vectors as rows of a matrix

Row space method

Column space method

Arrange the vectors as columns of a matrix

Application: finding basis for linear span

Example 4.1.14.1 (Row space method)

Place the vectors in the form of rows in a 6 x 4 matrix.

$$\boldsymbol{A} = \begin{array}{c} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 3 & 6 & 6 & 3 \\ 4 & 9 & 9 & 5 \\ -2 & -1 & -1 & 1 \\ 5 & 8 & 9 & 4 \\ 4 & 2 & 7 & 3 \end{pmatrix} \end{array} \begin{array}{c} \boldsymbol{u_1} \\ \boldsymbol{u_2} \\ \boldsymbol{u_3} \\ \boldsymbol{u_4} \\ \boldsymbol{u_5} \\ \boldsymbol{u_6} \end{array} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

row space of $\boldsymbol{A}$ = span$\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}, \boldsymbol{u_4}, \boldsymbol{u_5}, \boldsymbol{u_6}\}$

$\{(1, 2, 2, 1), (0, 1, 1, 1), (0, 0, 1, 1)\}$ is a basis

not from the original rows

Example 4.1.14.1 (Column space method)

Place the vectors in the form of columns in a 4 x 6 matrix.

$$\boldsymbol{B} = \begin{array}{c} \begin{pmatrix} 1 & 3 & 4 & -2 & 5 & 4 \\ 2 & 6 & 9 & -1 & 8 & 2 \\ 2 & 6 & 9 & -1 & 9 & 7 \\ 1 & 3 & 5 & 1 & 4 & 3 \end{pmatrix} \\ \begin{matrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 & \boldsymbol{u}_4 & \boldsymbol{u}_5 & \boldsymbol{u}_6 \end{matrix} \end{array} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \begin{pmatrix} 1 & 3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

column space of $\boldsymbol{B}$ $= \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4, \boldsymbol{u}_5, \boldsymbol{u}_6\}$

Pivot columns: 1st, 3rd and 5th columns

{(1, 2, 2, 1), (4, 9, 9, 5), (5, 8, 9, 3)} is a basis

all from the original columns

Application: extend a set to a basis

Example 4.1.14.2

$S = \{(1, 4, -2, 5, 1), (2, 9, -1, 8, 2), (2, 9, -1, 9, 3)\}$

S is linearly independent.

Extend S to a basis for $\mathbf{R}^5$.

Different from finding a basis for $\mathbf{R}^5$

This means:
Add on non-redundant vectors to S
to form a basis for $\mathbf{R}^5$

Need two more vectors

Use row space method

Application: extend a set to a basis

Example 4.1.14.2

$S = \{(1, 4, -2, 5, 1), (2, 9, -1, 8, 2), (2, 9, -1, 9, 3)\}$

$$\boldsymbol{A} = \begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 2 & 9 & -1 & 8 & 2 \\ 2 & 9 & -1 & 9 & 3 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \boldsymbol{R} = \begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

1. Form a matrix ***A*** using the vectors in ***S*** as rows.
2. Reduce ***A*** to a row-echelon form ***R***.
3. Identify the non-pivot columns of ***R.***

   Look for columns without leading entries

   the 3rd and the 5th columns

Application: extend a set to a basis

form a basis for $\mathbf{R}^5$

complete $\boldsymbol{R}$ to a 5x5 matrix

Example 4.1.14.2

$$\boldsymbol{A} = \begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 2 & 9 & -1 & 8 & 2 \\ 2 & 9 & -1 & 9 & 3 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \boldsymbol{R} = \begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & x & * & * \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & y \end{pmatrix}$$

are not redundant in row space of $\boldsymbol{A}$ $\left\{ \begin{array}{l} \text{E.g. } (0\ 0\ 1\ 0\ 0) \\ \text{E.g. } (0\ 0\ 0\ 0\ 1) \end{array} \right.$

4. Form (row) vectors with leading entries at the non-pivot columns.

5. {Row vectors in $\boldsymbol{A}$} ∪ {vectors from Step 4 } form a basis for $\mathbf{R}^n$

$$\boxed{\{ (1, 4, -2, 5, 1), (2, 9, -1, 8, 2), (2, 9, -1, 9, 3), (0, 0, 1, 0, 0), (0, 0, 0, 0, 1) \}}$$

Revision on Bases

$S = \{(2, -1, 0), (1, -1, 3), (-5, 1, 0), (1, 0, 1)\}$

How to get a basis from S for $\mathbf{R}^3$?

Throw out redundant vectors from S

Arrange the vectors as columns of a matrix
Look for pivot columns of the REF

$$T = \left\{ \begin{pmatrix} 2 \\ 1 \\ -5 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\}$$

How to extend T to a basis for $\mathbf{R}^4$?

Add on non-redundant vectors to T

Arrange the vectors as rows of a matrix
Look for 'missing' leading entries of the REF

Solutions of linear system revisited

$\mathbf{Ax} = \mathbf{b}$

How do we tell whether this system has
(i) no solution, (ii) unique solution; (iii) infinite solutions ?

Approach 1: Form ($\mathbf{A}$ | $\mathbf{b}$) and look at REF

Approach 2: If $\mathbf{A}$ is a square matrix

$\mathbf{A}$ is invertible $\Rightarrow$ system has unique solution

$\mathbf{A}$ is singular $\Rightarrow$ system has no or infinite solutions

Approach 3: $\mathbf{A}$ is any matrix

$\mathbf{b}$ belongs to column space of $\mathbf{A}$
$\Rightarrow$ system has unique or infinite solutions

$\mathbf{b}$ does not belong to column space of $\mathbf{A}$
$\Rightarrow$ system has no solution

Discussion 4.1.15

matrix multiply with vector

general linear combination of the column vectors

$$\begin{pmatrix} 2 & -1 & 0 \\ 1 & -1 & 3 \\ -5 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ -2 \\ 3 \end{pmatrix} \Leftrightarrow 1\begin{pmatrix} 2 \\ 1 \\ -5 \\ 1 \end{pmatrix} + 3\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ -2 \\ 3 \end{pmatrix}$$

matrix equation form

vector equation form

system has a solution

actual linear combination of the column vectors

this vector belongs to the column space

Discussion 4.1.15

$$\begin{pmatrix} 2 & -1 & 0 \\ 1 & -1 & 3 \\ -5 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ -2 \\ 3 \end{pmatrix}$$

$$\boldsymbol{A} \quad \boldsymbol{x} = \boldsymbol{b}$$

system $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution

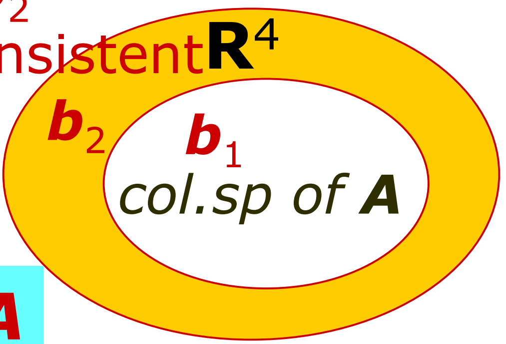
$$x\begin{pmatrix} 2 \\ 1 \\ -5 \\ 1 \end{pmatrix} + y\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 0 \\ 3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ -2 \\ 3 \end{pmatrix}$$

$$x\boldsymbol{C}_1 + y\boldsymbol{C}_2 + z\boldsymbol{C}_3 = \boldsymbol{b}$$

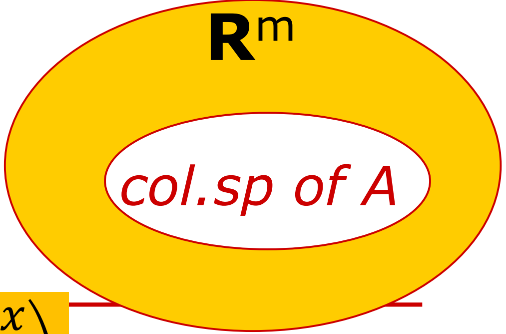
$\boldsymbol{b}$ can be written as a linear combination of the columns of $\boldsymbol{A}$

$\boldsymbol{b}$ belongs to the column space of $\boldsymbol{A}$

$\boldsymbol{Ax} = \boldsymbol{b}_2$
not consistent



$\boldsymbol{Ax} = \boldsymbol{b}_1$
consistent



Theorem 4.1.16

Let $\boldsymbol{A}$ be an $m \times n$ matrix.

$$\text{The column space of } \boldsymbol{A} = \{\ \boldsymbol{Au} \ \mid\ \boldsymbol{u} \in \mathbf{R}^n\ \}, \qquad \boldsymbol{u} = \begin{pmatrix} x \\ y \\ \vdots \\ z \end{pmatrix}.$$

$$\boldsymbol{A} = (\ \boldsymbol{c}_1 \mid \boldsymbol{c}_2 \mid \dots \mid \boldsymbol{c}_n\ ) \qquad \boldsymbol{Au} = x\boldsymbol{c}_1 + y\boldsymbol{c}_2 + \dots + z\boldsymbol{c}_n$$

$$\text{Span}\{\boldsymbol{c}_1, \boldsymbol{c}_2, \dots, \boldsymbol{c}_n\} = \left\{ \begin{array}{c} \text{all linear combination of} \\ \text{the column vectors of } \boldsymbol{A} \end{array} \right\}$$

A system of linear equation $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent if and only if $\boldsymbol{b}$ lies in the column space of $\boldsymbol{A}$.

Section 4.2

Ranks

Objectives

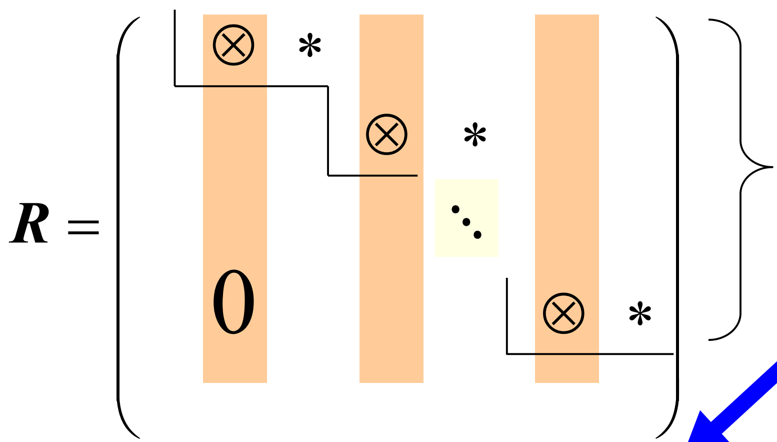
- What is the rank of a matrix?
- What is the relation between rank and invertibility of a matrix?
- What is the relation between rank and consistency of linear system?

Dimension of row space and column space

Theorem 4.2.1

The row space and column space of a matrix have the same dimension.

Let $\boldsymbol{A}$ be a matrix with row-echelon form $\boldsymbol{R}$.



dimension of row space of $\boldsymbol{A}$

= the number of nonzero rows

= the number of leading entries

dimension of column space of $\boldsymbol{A}$

= the number of pivot columns

= the number of leading entries

What is the rank of a matrix?

Definition 4.2.3

rank of a matrix :
dimension of its row space or column space.

Notation rank of matrix $\boldsymbol{A}$: rank($\boldsymbol{A}$)

If $\boldsymbol{R}$ is a row-echelon form of $\boldsymbol{A}$,

rank($\boldsymbol{A}$) = the number of nonzero rows of $\boldsymbol{R}$
= the number of leading entries in $\boldsymbol{R}$
= the number of pivot columns in $\boldsymbol{R}$

= largest number of linearly independent rows in $\boldsymbol{A}$

= largest number of linearly independent columns in $\boldsymbol{A}$

Ranks of some special matrices

Example 4.2.4.1

Row (column) space of zero matrix $\mathbf{0}$ = zero space

rank($\mathbf{0}$) = 0

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Row (column) space of $n \times n$ identity matrix $\boldsymbol{I_n}$ = $\mathbf{R}^n$

rank($\boldsymbol{I_n}$) = n

$$\boldsymbol{I}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Dimension is for vector space
Rank is for matrix

Example 4.2.4.3

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 3 & 6 & 6 & 3 \\ 4 & 9 & 9 & 5 \\ -2 & -1 & -1 & 1 \\ 5 & 8 & 9 & 4 \\ 4 & 2 & 7 & 3 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \begin{array}{c} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{l} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \\ \\ \\ \\ \end{array}$$

$$\begin{array}{ccc} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{array}$$

Basis for row space of $\boldsymbol{A} = \{\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\}$

Basis for column space of $\boldsymbol{A} = \{\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_3\}$

rank($\boldsymbol{A}$) = 3

DON'T Write: dim($\boldsymbol{A}$) = 3

Largest possible rank of a matrix

Example 4.2.4.4

What is the largest possible rank of a 5 x 3 matrix ?

The answer is 3

Find the largest possible number of pivot columns in a row-echelon form of a 5 x 3 matrix.

3 columns

What is the largest possible rank of a 3 x 5 matrix ?

The answer is 3

3 rows

Find the largest possible number of non-zero rows in a row-echelon form of a 3 x 5 matrix.

Largest possible rank of a matrix

Remark 4.2.5.1

For an $m \times n$ matrix $\boldsymbol{A}$, $\operatorname{rank}(\boldsymbol{A}) \leq \underbrace{\min\{m, n\}}$.

the smaller of the two numbers m and n

Example: $\boldsymbol{A}$ is 4×6

possible $\operatorname{rank}(\boldsymbol{A}) = 0, 1, 2, 3, 4$

$\boldsymbol{A}$ is full rank $\Leftrightarrow$ $\operatorname{rank}(\boldsymbol{A}) = 4$

An $m \times n$ matrix $\boldsymbol{A}$ with $\operatorname{rank}(\boldsymbol{A}) = \min\{m, n\}$ is said to be of full rank.

Remark 4.2.5.2-3

A square matrix $\boldsymbol{A}$ is of full rank
if and only if $\det(\boldsymbol{A}) \neq 0$.

$$\begin{tikzcd}
\boldsymbol{A}\text{ is of full rank} \arrow[r, leftrightarrow] \arrow[d, leftrightarrow] & \det(\boldsymbol{A}) \neq 0 \arrow[d, leftrightarrow] \\
\text{RREF of } \boldsymbol{A} \text{ is } \boldsymbol{I} \arrow[r, leftrightarrow] & \boldsymbol{A}\text{ is invertible}
\end{tikzcd}$$

$\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^T)$ for any matrix $\boldsymbol{A}$

row space of $\boldsymbol{A}$ = column space of $\boldsymbol{A}^T$

Relation between rank and consistency of system

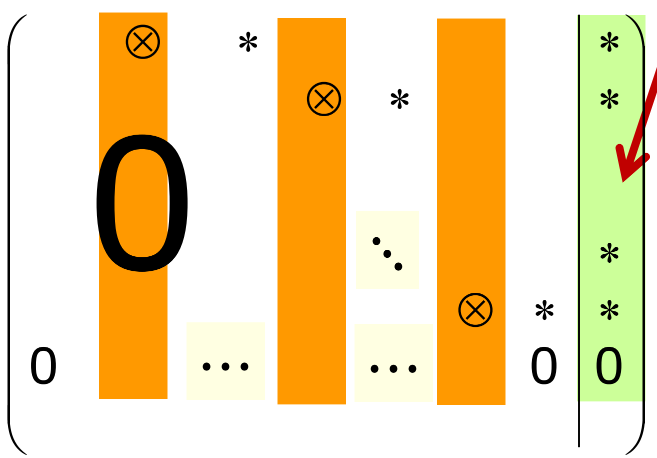
Remark 4.2.6

A system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent ⟺

Last lecture:
$\boldsymbol{b}$ ∈ column space of $\boldsymbol{A}$

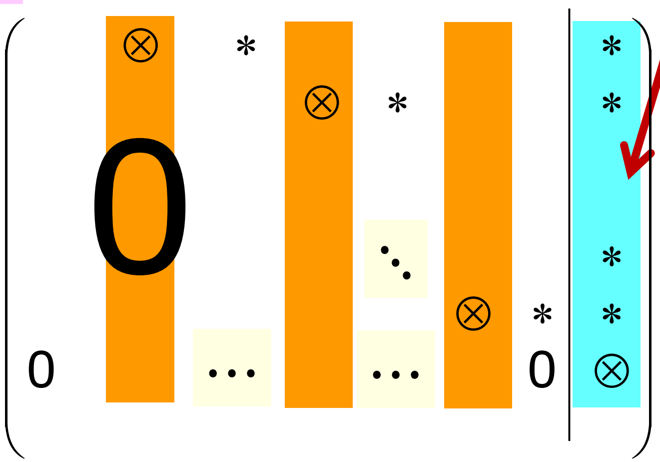
if and only if
the coefficient matrix $\boldsymbol{A}$
and the augmented matrix $(\boldsymbol{A} \mid \boldsymbol{b})$
have the same rank.

not a pivot column



system is consistent
rank($\boldsymbol{A}$) = rank($\boldsymbol{A}$ | $\boldsymbol{b}$)

a pivot column



system is inconsistent
rank($\boldsymbol{A}$) < rank($\boldsymbol{A}$ | $\boldsymbol{b}$)

Example 4.2.7

$$\begin{cases} 2x - y = 1 \\ x - y + 3z = 0 \\ -5x + y = 0 \\ x + z = 0 \end{cases}$$

coefficient matrix

$$\boldsymbol{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & -1 & 3 \\ -5 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

augmented matrix

$$(\boldsymbol{A} \mid \boldsymbol{b}) = \left(\begin{array}{ccc|c} 2 & -1 & 0 & 1 \\ 1 & -1 & 3 & 0 \\ -5 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 2 & -1 & 0 & 1 \\ 1 & -1 & 3 & 0 \\ -5 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gaussian}} \left(\begin{array}{ccc|c} \multicolumn{3}{c}{\overbrace{\hphantom{1 \quad 0 \quad 1}}^{\text{REF of } \boldsymbol{A}}} & \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

REF of $\boldsymbol{A}$ rank($\boldsymbol{A}$) = 3

REF of ($\boldsymbol{A}|\boldsymbol{b}$) rank($\boldsymbol{A}|\boldsymbol{b}$) = 4

The system is inconsistent.

Rank of a product of two matrices

rank($\boldsymbol{AB}$) ≤ rank($\boldsymbol{A}$)

rank($\boldsymbol{AB}$) ≤ rank($\boldsymbol{B}$)

Theorem 4.2.8

rank($\boldsymbol{AB}$) ≤ min{ rank($\boldsymbol{A}$), rank($\boldsymbol{B}$) }

$\boldsymbol{A}$: m×n

$\boldsymbol{B}$: n×p

Proof

Let $\boldsymbol{B} = (\boldsymbol{b}_1\ \boldsymbol{b}_2\ \ldots\ \boldsymbol{b}_p)$

$\boldsymbol{AB} = (\boldsymbol{Ab}_1\ \boldsymbol{Ab}_2\ \ldots\ \boldsymbol{Ab}_p)$ see Notation 2.2.15

where $\boldsymbol{Ab}_i$ is the ith column of $\boldsymbol{AB}$.

$\boldsymbol{Ab}_i \in$ column space of $\boldsymbol{A}$ By Theorem 4.1.16

span{$\boldsymbol{Ab}_1$, $\boldsymbol{Ab}_2$, …, $\boldsymbol{Ab}_p$} ⊆ column space of $\boldsymbol{A}$

=

column space of $\boldsymbol{AB}$

⇓ By Theorem 3.2.10

dim(column space of $\boldsymbol{AB}$) ≤ dim(column space of $\boldsymbol{A}$)

rank($\boldsymbol{AB}$) rank($\boldsymbol{A}$)

Rank of a product of two matrices

$\mathrm{rank}(\boldsymbol{AB}) \le \mathrm{rank}(\boldsymbol{A})$

$\mathrm{rank}(\boldsymbol{AB}) \le \mathrm{rank}(\boldsymbol{B})$

Theorem 4.2.8

$$\mathrm{rank}(\boldsymbol{AB}) \le \min\{\ \mathrm{rank}(\boldsymbol{A}),\ \mathrm{rank}(\boldsymbol{B})\ \}$$

Proof

$\mathrm{rank}(\boldsymbol{AB}) \le \mathrm{rank}(\boldsymbol{A})$

Also need to show: $\mathrm{rank}(\boldsymbol{AB}) \le \mathrm{rank}(\boldsymbol{B})$

we have $\mathrm{rank}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}) \le \mathrm{rank}(\boldsymbol{B}^{\mathrm{T}})$

$$\mathrm{rank}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}) = \mathrm{rank}((\boldsymbol{AB})^{\mathrm{T}}) = \mathrm{rank}(\boldsymbol{AB}), \quad \mathrm{rank}(\boldsymbol{B}^{\mathrm{T}}) = \mathrm{rank}(\boldsymbol{B})$$

$$\boxed{\mathrm{rank}(\boldsymbol{AB}) \ \le\ \mathrm{rank}(\boldsymbol{B})}$$

Therefore

$\mathrm{rank}(\boldsymbol{AB}) \le \min\{\mathrm{rank}(\boldsymbol{A}), \mathrm{rank}(\boldsymbol{B})\}$.

column space of $\boldsymbol{AB} \subseteq$ column space of $\boldsymbol{A}$

From proof of thm 4.2.8

Quiz Time

row space of $\boldsymbol{AB} \subseteq$ row space of $\boldsymbol{B}$

column space of $(\boldsymbol{AB})^{\mathrm{T}} \subseteq$ column space of $\boldsymbol{B}^{\mathrm{T}}$

column space of $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} \subseteq$ column space of $\boldsymbol{B}^{\mathrm{T}}$

A. True
B. False

Section 4.3

Nullspaces and Nullities

Objectives

- What is the nullspace and nullity of a matrix?
- What is the Dimension Theorem?
- What is the relation between nullspace and solution set of a linear system?

What is the nullspace and nullity of a matrix?

Definition 4.3.1

$\boldsymbol{A}$: $m \times n$ matrix

nullspace of $\boldsymbol{A}$ subspace of $\mathbf{R}^n$

is the solution space of the homogeneous system of linear equations $\boldsymbol{Ax} = \mathbf{0}$

nullity of $\boldsymbol{A}$ a number $\leq n$

is the dimension of the nullspace of $\boldsymbol{A}$

denoted by nullity($\boldsymbol{A}$)

Number of parameters in the general solution

Nullspace of a matrix $\boldsymbol{A}$

same

Solution space of a linear system $\boldsymbol{Ax} = \mathbf{0}$

all the vectors in $\mathbf{R}^n$ that are "killed" by $\boldsymbol{A}$

all the vectors in $\mathbf{R}^n$ that satisfy $\boldsymbol{Ax} = \mathbf{0}$

Basis for the nullspace

Example 4.3.3.1

$$\text{span}\left\{\begin{pmatrix}-1\\1\\0\\0\\0\end{pmatrix},\begin{pmatrix}-1\\0\\-1\\0\\1\end{pmatrix}\right\}$$

Find a basis for the nullspace of the matrix

$$\boldsymbol{A}=\left(\begin{array}{ccccc|c}2&2&-1&0&1&0\\-1&-1&2&-3&1&0\\0&0&1&1&1&0\\1&1&-2&0&-1&0\end{array}\right)\xrightarrow{\text{G.E.}}\left(\begin{array}{ccccc|c}1&1&0&0&1&0\\0&0&1&0&1&0\\0&0&0&1&0&0\\0&0&0&0&0&0\end{array}\right)$$

write all vectors as columns

The general solution of $\boldsymbol{Ax}=\mathbf{0}$

$$\mathbf{x}=\begin{pmatrix}-s-t\\s\\-t\\0\\t\end{pmatrix}=s\begin{pmatrix}-1\\1\\0\\0\\0\end{pmatrix}+t\begin{pmatrix}-1\\0\\-1\\0\\1\end{pmatrix}$$

$\text{nullity}(\boldsymbol{A})=2$

basis for the nullspace of $\boldsymbol{A}$

Example 4.3.3.2

$$\boldsymbol{B} = \begin{pmatrix} 2 & 1 & -5 & 1 \\ -1 & -1 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \begin{pmatrix} 1 & 0 & 0 & -\frac{7}{9} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{4}{9} \end{pmatrix} \quad \text{rank}(\boldsymbol{B}) = 3$$

general solution of $\boldsymbol{Bx} = \mathbf{0}$

$$\mathbf{x} = \begin{pmatrix} \frac{7}{9}t \\ -\frac{1}{3}t \\ \frac{4}{9}t \\ t \end{pmatrix} = \frac{1}{9}t \begin{pmatrix} 7 \\ -3 \\ 4 \\ 9 \end{pmatrix}$$

nullity($\boldsymbol{B}$) = 1   basis for the nullspace of $\boldsymbol{B}$

rank($\boldsymbol{B}$) + nullity($\boldsymbol{B}$) = 3 + 1 = 4

= the number of columns of $\boldsymbol{B}$

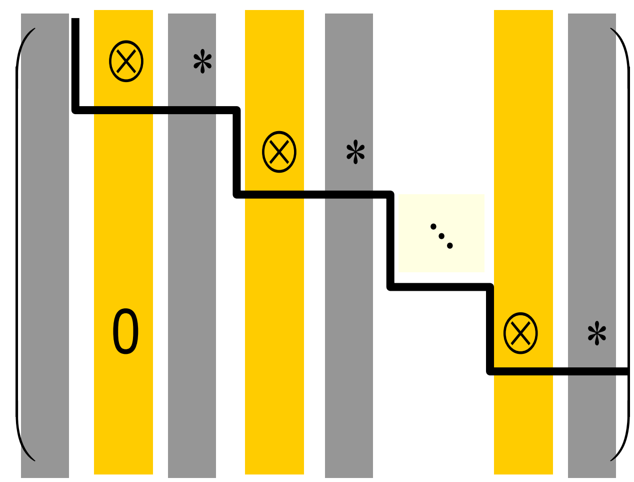
Dimension Theorem for Matrices

Theorem 4.3.4

If $\boldsymbol{A}$ is a matrix with n columns, then

$$\text{rank}(\boldsymbol{A}) + \text{nullity}(\boldsymbol{A}) = n.$$

row-echelon form

$\boldsymbol{R} =$ 

pivot columns

(correspond to basis for column space of $\boldsymbol{A}$) rank($\boldsymbol{A}$)

non-pivot columns

(correspond to parameters in general solutions) nullity($\boldsymbol{A}$)

Applying Dimension Theorem

Example 4.3.5.2

In each of the following cases,
find rank($\boldsymbol{A}$), nullity($\boldsymbol{A}$) and nullity($\boldsymbol{A}^T$).

Size of $\boldsymbol{A}$	# column of $\boldsymbol{A}$	# column of $\boldsymbol{A}^T$	rank($\boldsymbol{A}$) rank($\boldsymbol{A}^T$)	nullity($\boldsymbol{A}$)	nullity($\boldsymbol{A}^T$)
3×4	4	3	3	1	0
7×5	5	7	2	3	5
3×2	2	3	0	2	3

rank($\boldsymbol{A}$) + nullity($\boldsymbol{A}$) = # column of $\boldsymbol{A}$

rank($\boldsymbol{A}^T$) + nullity($\boldsymbol{A}^T$) = # column of $\boldsymbol{A}^T$

homogeneous linear system $\left\{\begin{array}{rcrcrcrcrcr} & & & & 2x_3 & + & 4x_4 & + & 2x_5 & = & 0 \\ x_1 & + & 2x_2 & + & 4x_3 & + & 5x_4 & + & 3x_5 & = & 0 \\ -2x_1 & - & 4x_2 & - & 5x_3 & - & 4x_4 & + & 3x_5 & = & 0 \end{array}\right\} (L_0)$

Example 1.4.7 (revisited)

Non-homogeneous linear system:

$$\left\{\begin{array}{rcrcrcrcrcr} & & & & 2x_3 & + & 4x_4 & + & 2x_5 & = & 8 \\ x_1 & + & 2x_2 & + & 4x_3 & + & 5x_4 & + & 3x_5 & = & -9 \\ -2x_1 & - & 4x_2 & - & 5x_3 & - & 4x_4 & + & 3x_5 & = & 6 \end{array}\right\} (L)$$

solutions of (L_0)

general solution of (L) not solutions of (L) a solution of (L)

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -29 - 2s + 3t \\ s \\ 8 - 2t \\ t \\ -4 \end{pmatrix} = \underbrace{s\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 3 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}}_{\text{general solution of } (L_0)} + \boxed{\begin{pmatrix} -29 \\ 0 \\ 8 \\ 0 \\ -4 \end{pmatrix}}$$

can be replaced by any other solution of (L)

Exercise 2 Q9

Suppose the homogeneous system $\boldsymbol{Ax} = \mathbf{0}$ has non-trivial solutions. ← $\boldsymbol{u}$ is a non-trivial solution

Show that the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ has either no solution or infinitely many solutions.

Idea of proof

We already know $\boldsymbol{Ax} = \boldsymbol{b}$ has either:

- No solution
- Exactly one solution ← $\boldsymbol{v}$ is a solution
- Infinitely many solutions

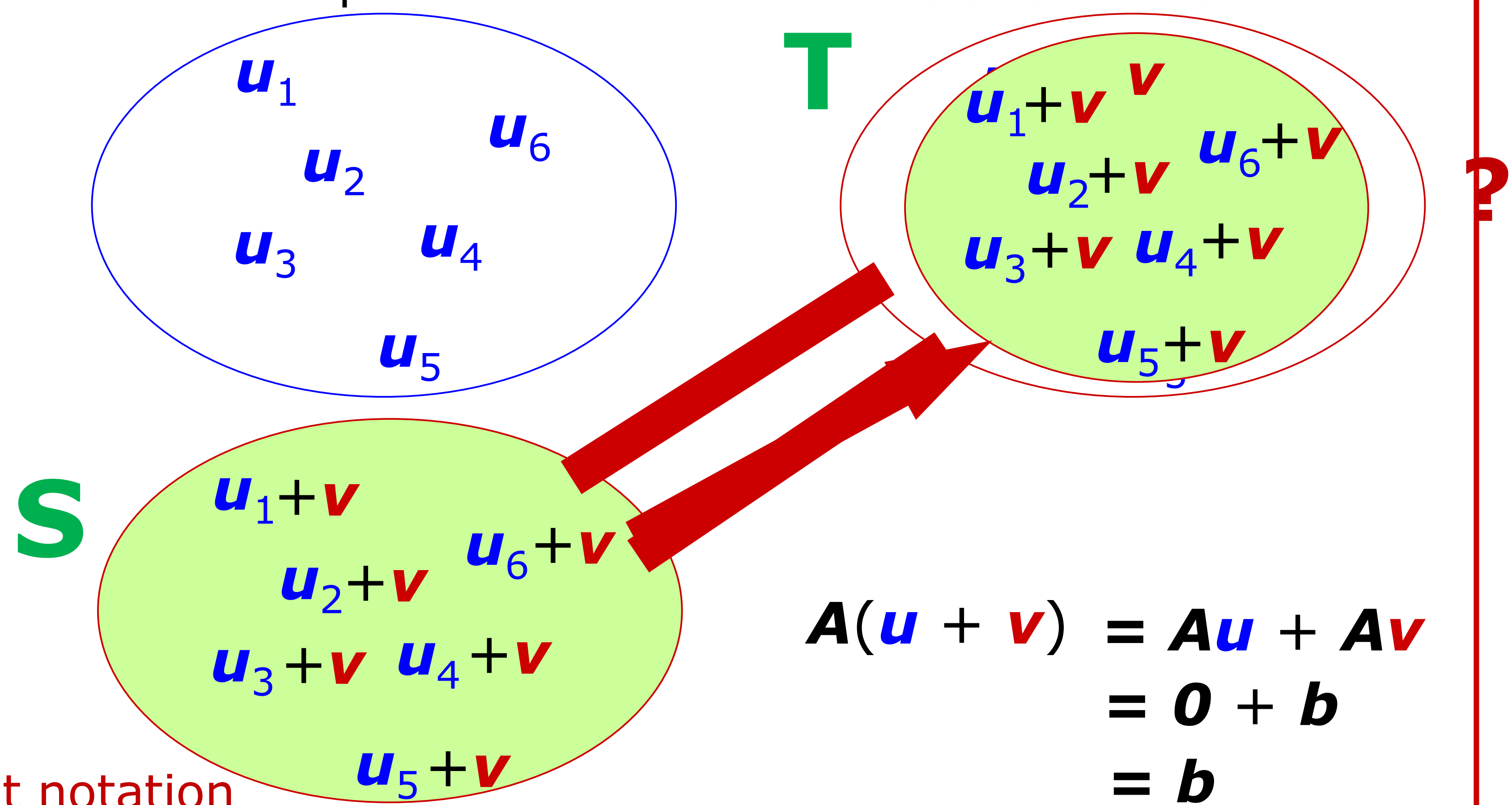
Not possible

$\boldsymbol{u} + \boldsymbol{v}$ is also a solution of $\boldsymbol{Ax} = \boldsymbol{b}$

Solution set of non-homogeneous system

Theorem 4.3.6 (Diagram version)

solution space of $\boldsymbol{Ax} = \mathbf{0}$ solution set of $\boldsymbol{Ax} = \boldsymbol{b}$



$$\boldsymbol{A}(\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{Au} + \boldsymbol{Av} = \mathbf{0} + \boldsymbol{b} = \boldsymbol{b}$$

set notation

$\{ \boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u}$ is a solution of the system $\boldsymbol{Ax} = \mathbf{0} \}$

Solution set of non-homogeneous system

Theorem 4.3.6

Suppose the system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$ has a (particular) solution $\boldsymbol{v}$.

The solution set of $\boldsymbol{Ax} = \boldsymbol{b}$

$$= \{ \underset{\text{vary}}{\boldsymbol{u}} + \underset{\text{fix}}{\boldsymbol{v}} \mid \boldsymbol{u} \text{ is an element of the nullspace of } \boldsymbol{A} \}$$

The general solution of $\boldsymbol{Ax} = \boldsymbol{b}$ can be given by

(the general solution of $\boldsymbol{Ax} = \boldsymbol{0}$) + $\boldsymbol{v}$

If we know the general solution of $\boldsymbol{Ax} = \boldsymbol{0}$
and one particular solution of $\boldsymbol{Ax} = \boldsymbol{b}$,
then we have the general solution for $\boldsymbol{Ax} = \boldsymbol{b}$.

Solution set of non-homogeneous system

Example 4.3.8

linear system $\boldsymbol{Ax} = \boldsymbol{b}$

one particular solution

$$\boldsymbol{A} = \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & -2 & 0 & -1 \end{pmatrix} \quad \boldsymbol{b} = \begin{pmatrix} 0 \\ 0 \\ 3 \\ -3 \end{pmatrix} \quad \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

By Example 4.3.3.1,

the nullspace of $\boldsymbol{A}$ = $\left\{ s\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \;\middle|\; s,t \text{ in } \mathbf{R} \right\}$

solution space of $\boldsymbol{Ax} = \boldsymbol{0}$

solution set of $\boldsymbol{Ax} = \boldsymbol{b}$ $\left\{ s\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; s,t \text{ in } \mathbf{R} \right\}$

The proof

Theorem 4.3.6

T = the solution set of $\boldsymbol{Ax} = \boldsymbol{b}$

$S = \{ \boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u}$ is an element of the nullspace of $\boldsymbol{A}\}$

We want to show: T = S

Need to show: T ⊆ S and S ⊆ T

T ⊆ S

Show every solution of $\boldsymbol{Ax} = \boldsymbol{b}$ has the form $\boldsymbol{u} + \boldsymbol{v}$

Next slide

S ⊆ T

Show every $\boldsymbol{u} + \boldsymbol{v}$ is a solution of $\boldsymbol{Ax} = \boldsymbol{b}$

Substitute $\boldsymbol{u} + \boldsymbol{v}$ for $\boldsymbol{x}$ in $\boldsymbol{Ax} = \boldsymbol{b}$

T = the solution set of $\boldsymbol{Ax} = \boldsymbol{b}$

The proof

$S = \{\, \boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u}$ is an element of the nullspace of $\boldsymbol{A}\}$

a solution of $\boldsymbol{Ax} = \boldsymbol{b}$

Theorem 4.3.6

To show $T \subseteq S$:

element-chasing method

Let $\boldsymbol{w} \in T$ Want to show: $\boldsymbol{w} \in S$

i.e. Given $\boldsymbol{Aw} = \boldsymbol{b}$

i.e. To show $\boldsymbol{w}$ can be written as $\boldsymbol{u} + \boldsymbol{v}$

We have $\boldsymbol{Av} = \boldsymbol{b}$

i.e. To show $\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}$

i.e. To show $\boldsymbol{w} - \boldsymbol{v} = \boldsymbol{u}$

$$\begin{aligned} &\boldsymbol{A}(\boldsymbol{w} - \boldsymbol{v}) \\ &= \boldsymbol{Aw} - \boldsymbol{Av} \\ &= \boldsymbol{b} - \boldsymbol{b} = \boldsymbol{0} \end{aligned}$$

i.e. To show $\boldsymbol{w} - \boldsymbol{v}$ is an element of the nullspace of $\boldsymbol{A}$

i.e. To show $\boldsymbol{A}(\boldsymbol{w} - \boldsymbol{v}) = \boldsymbol{0}$

Hence $T \subseteq S$.

Solution set of non-homogeneous system

Remark 4.3.7

Suppose the system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$ has a (particular) solution $\boldsymbol{v}$.

The solution set of $\boldsymbol{Ax} = \boldsymbol{b}$

$= \{\, \boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u}$ is an element of the nullspace of $\boldsymbol{A}\,\}$

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a consistent linear system. Then

$\boldsymbol{Ax} = \boldsymbol{b}$ has exactly one solution
if and only if
the nullspace of $\boldsymbol{A}$ is equal to $\{\boldsymbol{0}\}$

Section 5.1

Inner Products in $\mathbf{R}^n$

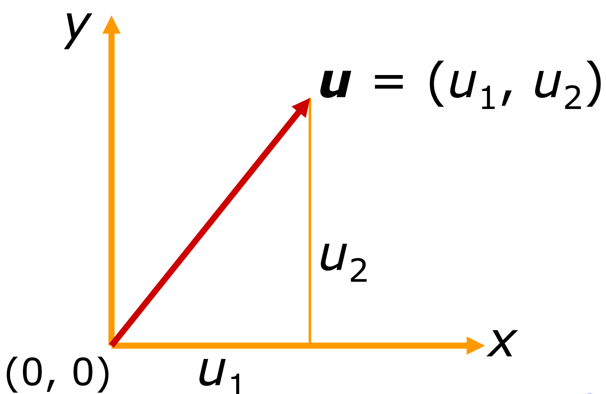
Objectives

- What are the algebraic representation of length, distance and angles in $\mathbf{R}^n$?
- What is the dot product of vectors?

Discussion 5.1.1

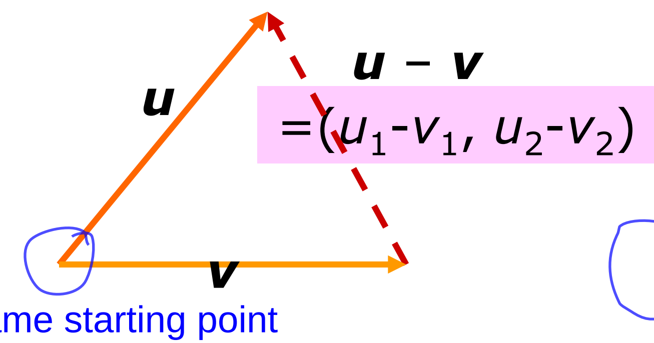
$\boldsymbol{u} = (u_1, u_2)$ and $\boldsymbol{v} = (v_1, v_2)$: vectors in $\mathbf{R}^2$

length of vector



$$\|\boldsymbol{u}\| = \sqrt{u_1^{\,2} + u_2^{\,2}}$$

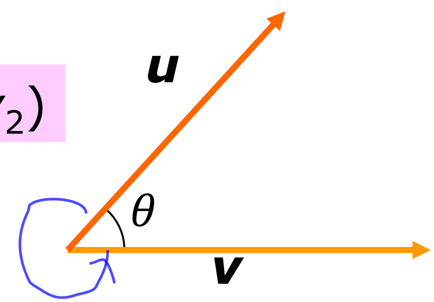
distance between two vectors



$$\|\boldsymbol{u} - \boldsymbol{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2}$$

angle between two vectors

$0 \leq \theta < \pi$



$$\cos\theta = \frac{u_1 v_1 + u_2 v_2}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|}$$

derived from cosine rule

$$\theta = \cos^{-1}\left(\frac{u_1 v_1 + u_2 v_2}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|}\right)$$

Similarly for $\mathbf{R}^3$ case

Definition 5.1.2

$\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$, $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ vectors in $\mathbf{R}^n$

	length $\|\|\boldsymbol{u}\|\|$ of vector	distance $\|\|\boldsymbol{u} - \boldsymbol{v}\|\|$ between two vectors	angle θ between two vectors
$\mathbf{R}^2$	$\sqrt{u_1^{\ 2} + u_2^{\ 2}}$	$\sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2}$	$\cos^{-1}\left(\dfrac{u_1v_1 + u_2v_2}{\|\|\boldsymbol{u}\|\|\,\|\|\boldsymbol{v}\|\|}\right)$
$\mathbf{R}^3$	$\sqrt{u_1^{\ 2} + u_2^{\ 2} + u_3^{\ 2}}$	$\sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}$	$\cos^{-1}\left(\dfrac{u_1v_1 + u_2v_2 + u_3v_3}{\|\|\boldsymbol{u}\|\|\,\|\|\boldsymbol{v}\|\|}\right)$
$\mathbf{R}^n$	$\sqrt{u_1^{\ 2} + u_2^{\ 2} + \ldots + u_n^{\ 2}}$	$\sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}$	$\cos^{-1}\left(\dfrac{u_1v_1 + u_2v_2 + \ldots + u_nv_n}{\|\|\boldsymbol{u}\|\|\,\|\|\boldsymbol{v}\|\|}\right)$

cumbersome

What is dot product?

Definition 5.1.2.1

$\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$, $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ vectors in $\mathbf{R}^n$

The dot product of $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined to be the value (scalar)

scalar product

$$\underbrace{\boxed{\boldsymbol{u} \cdot \boldsymbol{v}}}_{\text{product of two vectors}} = \underbrace{\boxed{u_1 v_1 + u_2 v_2 + \ldots + u_n v_n}}_{\text{scalar}}$$

inner product

In particular,

$$\boldsymbol{u} \cdot \boldsymbol{u} = u_1^{\,2} + u_2^{\,2} + \ldots + u_n^{\,2}$$

Definition 5.1.2 ($\mathbf{R}^n$ case)

$\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$, $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ vectors in $\mathbf{R}^n$

What for?

norm length $\|\|\boldsymbol{u}\|\|$ of vector	distance $\|\|\boldsymbol{u} - \boldsymbol{v}\|\|$ between two vectors	angle θ between two vectors
$\sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$	$\sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}$	$\cos^{-1}\left(\dfrac{u_1 v_1 + u_2 v_2 + \ldots + u_n v_n}{\|\|\boldsymbol{u}\|\|\,\|\|\boldsymbol{v}\|\|}\right)$
$\sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}$	$\sqrt{(\boldsymbol{u} - \boldsymbol{v}) \cdot (\boldsymbol{u} - \boldsymbol{v})}$	$\cos^{-1}\left(\boxed{\dfrac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\|\boldsymbol{u}\|\|\,\|\|\boldsymbol{v}\|\|}}\right)$

vectors of norm 1 are called unit vectors

Does this quotient have value between -1 and 1 ?

$\boldsymbol{u}$ is a unit vector $\Leftrightarrow$ $||\boldsymbol{u}|| = 1$

Dot product as matrix multiplication

Remark 5.1.3

$\boldsymbol{u} = (u_1 \;\; u_2 \;\; \dots \;\; u_n)$ and $\boldsymbol{v} = (v_1 \;\; v_2 \;\dots\; \; v_n)$

regarded as row matrix

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1v_1 + u_2v_2 + \dots + u_nv_n = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \boldsymbol{u}\boldsymbol{v}^T$$

$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}\mathbf{v}^T$ 1 x 1 matrix

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \qquad \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \qquad \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v}$$

regarded as column matrix

Properties of dot product

Theorem 5.1.5

Let c be a scalar and $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ vectors in $\mathbf{R}^n$.

1. $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$ — commutative law
2. $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$
   $\boldsymbol{w} \cdot (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{w} \cdot \boldsymbol{u} + \boldsymbol{w} \cdot \boldsymbol{v}$ — distributive law
3. $(c\boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot (c\boldsymbol{v}) = c(\boldsymbol{u} \cdot \boldsymbol{v})$ — scalar mult.
4. $||c\boldsymbol{u}|| = |c|\,||\boldsymbol{u}||$ — (not $c\,||\boldsymbol{u}||$)
5. (i) $\boldsymbol{u} \cdot \boldsymbol{u} \geq 0$
   (ii) $\boldsymbol{u} \cdot \boldsymbol{u} = 0$ if and only if $\boldsymbol{u} = \mathbf{0}$.

$u_1^2 + u_2^2 + \cdots + u_n^2 = 0 \Rightarrow u_1 = 0,\ u_2 = 0,\ \ldots,\ u_n = 0$

Additional example

$\boldsymbol{Av} = \mathbf{0}$ if and only if $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Av} = \mathbf{0}$

Proof

$$\boldsymbol{Av} = \mathbf{0} \Rightarrow \boldsymbol{A}^{\mathrm{T}}\boldsymbol{Av} = \boldsymbol{A}^{\mathrm{T}}\mathbf{0} \Rightarrow \boldsymbol{A}^{\mathrm{T}}\boldsymbol{Av} = \mathbf{0}$$

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Av} = \mathbf{0}$$

$$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Av} = \boldsymbol{v}^{\mathrm{T}}\mathbf{0}$$

$$(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}})\boldsymbol{Av} = \mathbf{0}$$

$$(\boldsymbol{Av})^{\mathrm{T}}\boldsymbol{Av} = \mathbf{0}$$

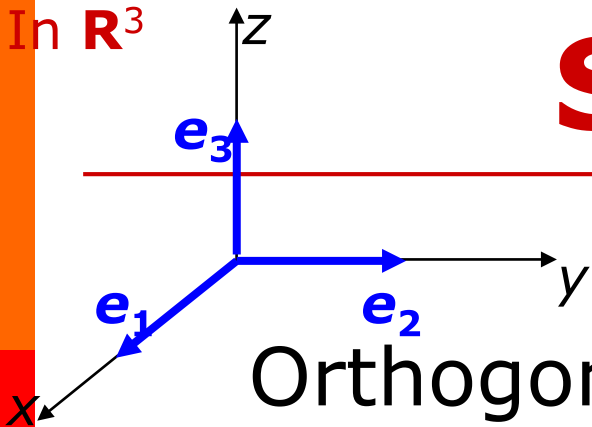
$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v}$ for column vectors

$$(\boldsymbol{Av})\cdot(\boldsymbol{Av}) = \mathbf{0}$$

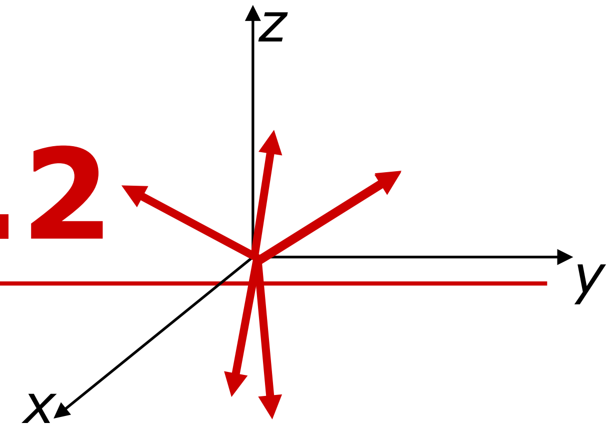
$$\boldsymbol{Av} = \mathbf{0}$$

$\boldsymbol{u} \cdot \boldsymbol{u} = 0$ if and only if $\boldsymbol{u} = \mathbf{0}$

In $\mathbf{R}^3$



Section 5.2



Orthogonal and Orthonormal Basis

Objectives

- What is an orthogonal/orthonormal set?
- How to normalize a vector?
- What are the properties of orthogonal sets?

Ortho- means: *straight, upright, right, correct*

What is an orthogonal/orthonormal set?

Definition 5.2.1

1. Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbf{R}^n$ are called orthogonal if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

   In $\mathbf{R}^2$ and $\mathbf{R}^3$, it means "perpendicular"

2. A set S of vectors in $\mathbf{R}^n$ is called orthogonal if every pair of distinct vectors in S are orthogonal.

   $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$

   $\boldsymbol{u}_1 \cdot \boldsymbol{u}_2 = 0, \ \boldsymbol{u}_1 \cdot \boldsymbol{u}_3 = 0, \ \ldots \ \boldsymbol{u}_{k-1} \cdot \boldsymbol{u}_k = 0$

   $||\boldsymbol{u}_1|| = ||\boldsymbol{u}_2|| = \ldots = ||\boldsymbol{u}_k|| = 1$

3. A set S of vectors in $\mathbf{R}^n$ is called orthonormal if S is orthogonal and every vector in S is a unit vector.

Angle between two orthogonal vectors

Remark 5.2.2

Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be two vectors in $\mathbf{R}^n$.

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal, $\Rightarrow \boldsymbol{u} \cdot \boldsymbol{v} = 0$

the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$:

$$\cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{||\boldsymbol{u}||\,||\boldsymbol{v}||}\right) = \cos^{-1}(0) = \frac{\pi}{2}$$

So $\boldsymbol{u}$ and $\boldsymbol{v}$ are perpendicular

An example of an orthogonal/orthonormal set

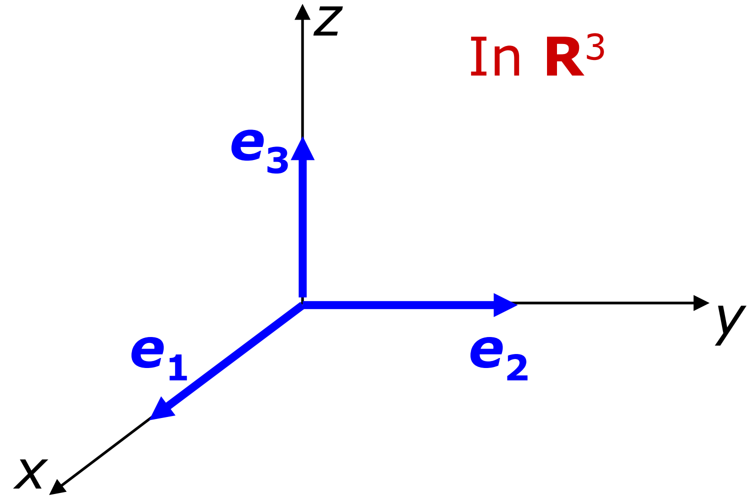
Example 5.2.3.3

Consider the standard basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ for $\mathbf{R}^n$.

$\boldsymbol{e}_1 = (1, 0, \ldots, 0)$

$\boldsymbol{e}_2 = (0, 1, \ldots, 0)$

$\boldsymbol{e}_n = (0, 0, \ldots, 1)$



For $i \neq j$, $\boldsymbol{e}_i \cdot \boldsymbol{e}_j = 0$.

So the standard basis is an orthogonal set

For $i = 1, 2, \ldots, n$, $\|\boldsymbol{e}_i\| = 1$.

So the standard basis is also an orthonormal set.

Another example of an orthogonal set

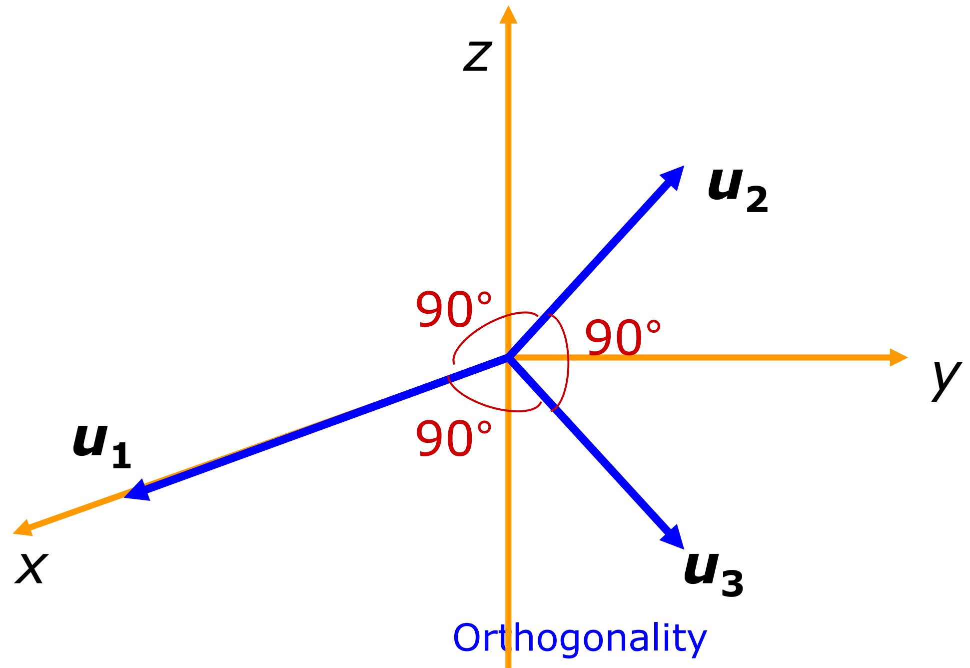
Example 5.2.3.2

$\boldsymbol{u_1} = (2, 0, 0)$, $\boldsymbol{u_2} = (0, 1, 1)$ and $\boldsymbol{u_3} = (0, 1, -1)$.

$$\boldsymbol{u_1} \cdot \boldsymbol{u_2} = 0, \quad \boldsymbol{u_1} \cdot \boldsymbol{u_3} = 0 \quad \text{and} \quad \boldsymbol{u_2} \cdot \boldsymbol{u_3} = 0$$

$\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ is an orthogonal set.

It is not an orthonormal set.



Converting orthogonal to orthonormal set

Example 5.2.3.2

$\boldsymbol{u_1} = (2, 0, 0)$ $\boldsymbol{u_2} = (0, 1, 1)$ $\boldsymbol{u_3} = (0, 1, -1)$

$$\left.\begin{aligned}
\boldsymbol{v}_1 &= \frac{1}{||\boldsymbol{u}_1||}\boldsymbol{u}_1 = \tfrac{1}{2}(2, 0, 0) = (1, 0, 0)\\
\boldsymbol{v}_2 &= \frac{1}{||\boldsymbol{u}_2||}\boldsymbol{u}_2 = \tfrac{1}{\sqrt{2}}(0, 1, 1) = (0, \tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}})\\
\boldsymbol{v}_3 &= \frac{1}{||\boldsymbol{u}_3||}\boldsymbol{u}_3 = \tfrac{1}{\sqrt{2}}(0, 1, -1) = (0, \tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}})
\end{aligned}\right\}\text{unit vectors}$$

$$||\boldsymbol{v}_i|| = \left\| \frac{1}{||\boldsymbol{u}_i||}\boldsymbol{u}_i \right\| = \frac{1}{||\mathbf{u}_i||}\,||\,\mathbf{u}_i\,|| = 1$$

For $i \neq j$,

$$\boldsymbol{v}_i \cdot \boldsymbol{v}_j = \left(\frac{1}{||\boldsymbol{u}_i||}\boldsymbol{u}_i\right)\cdot\left(\frac{1}{||\boldsymbol{u}_j||}\boldsymbol{u}_j\right) = \frac{1}{||\boldsymbol{u}_i||\,||\boldsymbol{u}_j||}(\boldsymbol{u}_i \cdot \boldsymbol{u}_j) = 0$$

So the set $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$ is orthonormal.

Normalizing a vector

Remark on Example 5.2.3.2

$$\boldsymbol{u} \xrightarrow{\text{normalizing}} \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u}$$

any non-zero vector — unit vector

Scalar multiple of the original vector

$$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\} \xrightarrow{\text{normalizing}} \left\{\frac{1}{\|\boldsymbol{u}_1\|}\boldsymbol{u}_1, \quad \frac{1}{\|\boldsymbol{u}_2\|}\boldsymbol{u}_2, \quad \frac{1}{\|\boldsymbol{u}_3\|}\boldsymbol{u}_3\right\}$$

an orthogonal set — an orthonormal set

orthogonal ⇒ linearly independent

Theorem 5.2.4

Let S be an orthogonal set of nonzero vectors in a vector space.

Then S is linearly independent.

Proof

Let $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}\}$ orthogonal set

$$c_1 \boldsymbol{u_1} + c_2 \boldsymbol{u_2} + \ldots + c_n \boldsymbol{u_n} = \mathbf{0}$$

Want to show:

$c_1 = 0, c_2 = 0, \ldots, c_n = 0$ is the only solution

Take dot product on both sides with $\boldsymbol{u_i}$ for every i.

orthogonal ⇒ linearly independent

$\boldsymbol{u_i} \cdot \boldsymbol{u_i} \neq 0$ for all i

Theorem 5.2.4

$\boldsymbol{u_j} \cdot \boldsymbol{u_i} = 0$ if $j \neq i$

Proof

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}\}$ orthogonal set
nonzero vectors

$$c_1 \boldsymbol{u_1} + c_2 \boldsymbol{u_2} + \ldots + c_n \boldsymbol{u_n} = \mathbf{0}$$

$$(c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}) \cdot \boldsymbol{u_1} = \mathbf{0} \cdot \boldsymbol{u_1}$$

$$c_1(\boldsymbol{u_1} \cdot \boldsymbol{u_1}) + c_2(\boldsymbol{u_2} \cdot \boldsymbol{u_1}) + \cdots + c_k(\boldsymbol{u_k} \cdot \boldsymbol{u_1}) = 0$$

$$c_1(\boldsymbol{u_1} \cdot \boldsymbol{u_1}) = 0$$

$$c_1 = 0$$

Similarly we can show $c_2 = 0, \ldots, c_n = 0$

What is an orthogonal/orthonormal basis?

Definition 5.2.5

1. A basis S for a vector space is called an orthogonal basis if S is orthogonal.

2. A basis S for a vector space is called an orthonormal basis if S is orthonormal.

$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is an orthogonal basis for $\mathbf{R}^3$

$\{(2,0,0), (0,1,1), (0,1,-1)\}$ is an orthogonal basis for $\mathbf{R}^3$

$\{(1,0,0), (1,1,0), (1,1,1)\}$ is not an orthogonal basis for $\mathbf{R}^3$

$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is an orthonormal basis for $\mathbf{R}^3$

$\{(2,0,0), (0,1,1), (0,1,-1)\}$ is not an orthonormal basis for $\mathbf{R}^3$

$\{(1,0,0), (1,1,0), (1,1,1)\}$ is not an orthonomal basis for $\mathbf{R}^3$

$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is a basis for $\mathbf{R}^3$

$\{(2,0,0), (0,1,1), (0,1,-1)\}$ is a basis for $\mathbf{R}^3$

$\{(1,0,0), (1,1,0), (1,1,1)\}$ is a basis for $\mathbf{R}^3$

Remark 5.2.6

A set S of nonzero vectors in a vector space V.

To check whether S is an orthonormal basis for V:

S is an orthogonal set

S spans V

S lin. indep

Theorem 5.2.4

Only need to check:

(i) S is orthonormal and

(ii) span(S) = V.

If we know dim V,

Only need to check:

(i) S is orthonormal and

(ii) $|S|$ = dim V.

How to check a set is an orthogonal basis?

Example 5.2.7.2

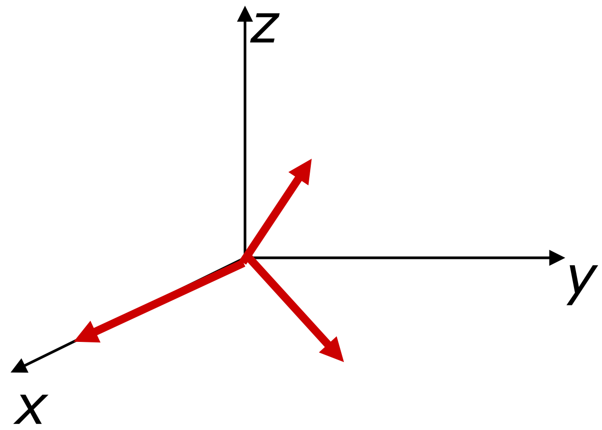
$$\boldsymbol{u}_1 = (2,\ 0,\ 0) \quad \boldsymbol{u}_2 = (0,\ 1,\ 1) \quad \boldsymbol{u}_3 = (0,\ 1, -1)$$

$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$

- an orthogonal set

- has three vectors $= \dim \mathbf{R}^3$

$\Rightarrow$ an orthogonal basis for $\mathbf{R}^3$.



Quiz Time

True or false

$$\mathbf{u}_1 = (1,\ -1,\ 1,\ -1) \quad \mathbf{u}_2 = (1,\ 1,\ 1,\ 1) \quad \mathbf{u}_3 = (0,\ 1, 0, -1)$$

$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is an orthogonal basis for

$V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$

Check:

$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is an orthogonal set

$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ spans V

$\Rightarrow \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is an orthogonal basis for V.

Coordinate vector w.r.t. orthogonal basis

Example 5.2.9.2

$\boldsymbol{u_1} = (1, 1, 1)$, $\boldsymbol{u_2} = (1, 0, -1)$ and $\boldsymbol{u_3} = (1, -2, 1)$.

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ is an orthogonal basis for $\mathbf{R}^3$.

Let $\boldsymbol{w} = (1, -1, 0)$. Find $(\boldsymbol{w})_S$

coordinate vector w.r.t. basis S

$$\boldsymbol{w} = c_1 \boldsymbol{u_1} + c_2 \boldsymbol{u_2} + c_3 \boldsymbol{u_3} \Rightarrow (\boldsymbol{w})_S = (c_1, c_2, c_3)$$

standard approach: need to solve linear system

Short cut formula (when S is orthogonal) :

$$(\boldsymbol{w})_S = \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\|\boldsymbol{u_1}\|^2}, \frac{\boldsymbol{w} \cdot \boldsymbol{u_2}}{\|\boldsymbol{u_2}\|^2}, \frac{\boldsymbol{w} \cdot \boldsymbol{u_3}}{\|\boldsymbol{u_3}\|^2} \right) = \left(0, \frac{1}{2}, \frac{1}{2}\right)$$

Coordinate vector w.r.t. orthogonal basis

Theorem 5.2.8.1

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$: an orthogonal basis for V

For any vector $\boldsymbol{w}$ in V,

$$\boldsymbol{w} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \ldots + c_k\boldsymbol{u_k}$$

$$c_1 = \frac{\boldsymbol{w}\cdot\boldsymbol{u_1}}{\|\boldsymbol{u_1}\|^2} \qquad c_2 = \frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2} \qquad c_k = \frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}$$

$$(\boldsymbol{w})_S = \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u_1}}{\|\boldsymbol{u_1}\|^2}, \frac{\boldsymbol{w}\cdot\boldsymbol{u_2}}{\|\boldsymbol{u_2}\|^2}, \cdots, \frac{\boldsymbol{w}\cdot\boldsymbol{u_k}}{\|\boldsymbol{u_k}\|^2}\right)$$

Theorem 5.2.8.2: orthonormal basis
Special case of part 1, with $\|\boldsymbol{u}_i\|^2 = 1$ for all i

The proof

$$\boldsymbol{u_j} \cdot \boldsymbol{u_i} = 0 \quad \text{if } j \neq i$$

$$\boldsymbol{u_i} \cdot \boldsymbol{u_i} = ||\boldsymbol{u_i}||^2$$

Theorem 5.2.8.1

WTS: $c_i = \dfrac{\mathbf{w} \cdot \mathbf{u}_i}{||\mathbf{u}_i||^2}$

Let $\boldsymbol{w} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}$

$$\boldsymbol{w} \cdot \boldsymbol{u_1} = (c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}) \cdot \boldsymbol{u_1}$$

$$= c_1(\boldsymbol{u_1} \cdot \boldsymbol{u_1}) + c_2(\boldsymbol{u_2} \cdot \boldsymbol{u_1}) + \cdots + c_k(\boldsymbol{u_k} \cdot \boldsymbol{u_1})$$

$$= c_1(\boldsymbol{u_1} \cdot \boldsymbol{u_1})$$

$$= c_1 ||\boldsymbol{u_1}||^2$$

So $c_1 = \dfrac{\mathbf{w} \cdot \mathbf{u}_1}{||\mathbf{u}_1||^2}$

Coordinate vector w.r.t. orthonormal basis

Example 5.2.9.1

$\boldsymbol{v}_1 = (\frac{3}{5}, \frac{4}{5})$ $\quad \boldsymbol{v}_2 = (\frac{4}{5}, -\frac{3}{5})$

$||\boldsymbol{v}_i||^2 = 1 \qquad \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0$

$S = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ is an orthonormal basis for $\mathbf{R}^2$.

Let $\boldsymbol{w} = (x, y)$ be any vector in $\mathbf{R}^2$.

Express $(\boldsymbol{w})_S$ in terms of x and y

$\boldsymbol{w} = (\boldsymbol{w} \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{w} \cdot \boldsymbol{v}_2)\boldsymbol{v}_2$

$$\left.\begin{aligned} \boldsymbol{w} \cdot \boldsymbol{v}_1 &= \frac{3x+4y}{5} \\ \boldsymbol{w} \cdot \boldsymbol{v}_2 &= \frac{4x-3y}{5} \end{aligned}\right\} \Rightarrow \quad \boldsymbol{w} = \frac{3x+4y}{5}\boldsymbol{v}_1 + \frac{4x-3y}{5}\boldsymbol{v}_2$$

$$(\boldsymbol{w})_S = \left(\frac{3x+4y}{5}, \frac{4x-3y}{5}\right)$$

Coordinate vector w.r.t. orthogonal basis

Example 5.2.9.2

$\boldsymbol{u_1} = (1, 1, 1)$, $\boldsymbol{u_2} = (1, 0, -1)$ and $\boldsymbol{u_3} = (1, -2, 1)$.

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ is an orthogonal basis for $\mathbf{R}^3$.

Let $\boldsymbol{w} = (1, -1, 0)$. Find $(\boldsymbol{w})_S$

coordinate vector w.r.t. basis S

$$\boldsymbol{w} = c_1 \boldsymbol{u_1} + c_2 \boldsymbol{u_2} + c_3 \boldsymbol{u_3} \Rightarrow (\boldsymbol{w})_S = (c_1, c_2, c_3)$$

Theorem 5.2.8 (when S is orthogonal) :

$$(\boldsymbol{w})_S = \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\|\boldsymbol{u_1}\|^2}, \frac{\boldsymbol{w} \cdot \boldsymbol{u_2}}{\|\boldsymbol{u_2}\|^2}, \frac{\boldsymbol{w} \cdot \boldsymbol{u_3}}{\|\boldsymbol{u_3}\|^2} \right) = \left(0, \frac{1}{2}, \frac{1}{2} \right)$$

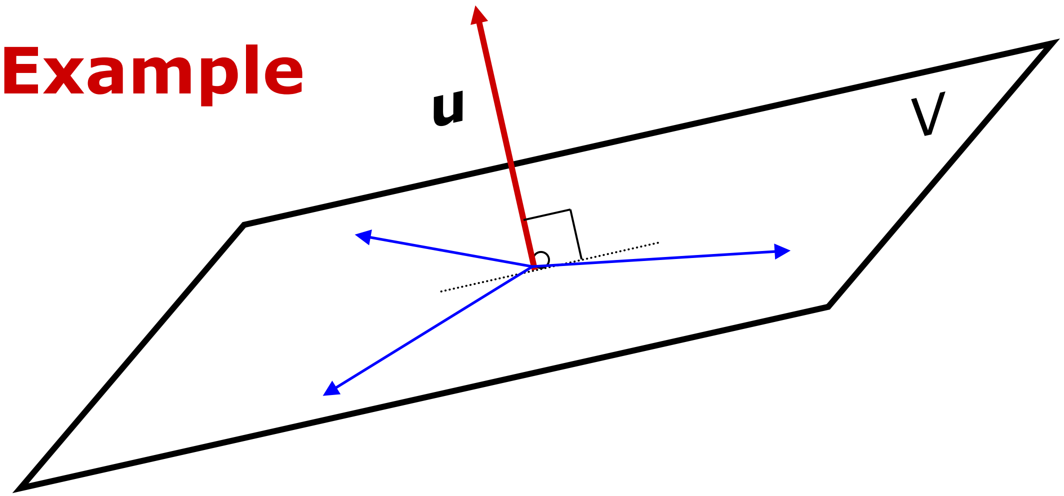
A vector orthogonal to a subspace

Definition 5.2.10

Let V be a subspace of $\mathbf{R}^n$.

A vector $\boldsymbol{u}$ is orthogonal to the subspace V if $\boldsymbol{u}$ is orthogonal to all vectors in V.

Example



A vector orthogonal to a plane

Example 5.2.11.1

$3x - 5y + 11z = 0$

V a plane in $\mathbf{R}^3$ with equation $ax + by + cz = 0$.

(3, -5, 11)

$\boldsymbol{n} = (a, b, c)$ Why it works?

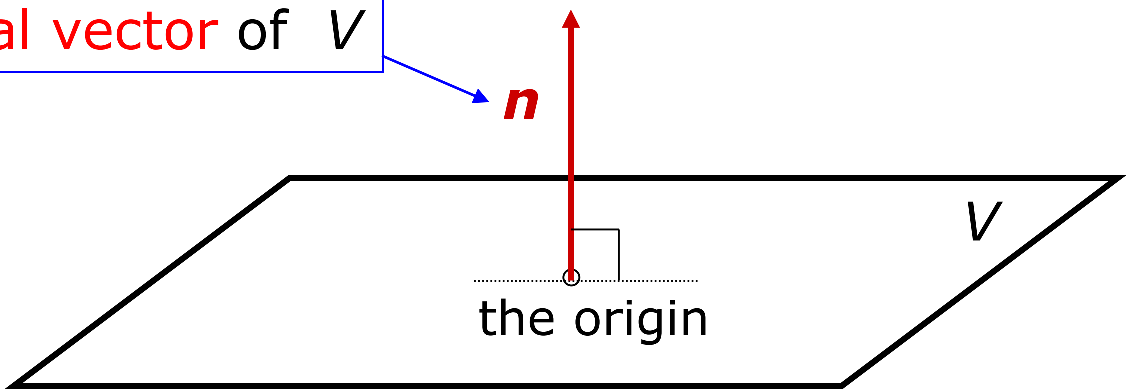
satisfies the equation

Take any $\boldsymbol{u} = (x_0, y_0, z_0)$ in V

Take the dot product $\boldsymbol{n} \cdot \boldsymbol{u} = ax_0 + by_0 + cz_0 = 0$

So $\boldsymbol{n}$ is orthogonal to every vector $\boldsymbol{u}$ in V

a normal vector of V



How to find vectors orthogonal to a subspace?

Example 5.2.11.2

$\boldsymbol{u_1} = (1, 1, 1, 0)$ and $\boldsymbol{u_2} = (0, -1, -1, 1)$

$V = \text{span}\{\boldsymbol{u_1}, \boldsymbol{u_2}\}$ a subspace of $\mathbf{R}^4$

Find all vectors that are orthogonal to V.

(w, x, y, z)

Let $\boldsymbol{v}$ be orthogonal to $V = \text{span}\{\boldsymbol{u_1}, \boldsymbol{u_2}\}$

$\Leftrightarrow$ $\boldsymbol{v}$ is orthogonal to $a\boldsymbol{u_1} + b\boldsymbol{u_2}$ for all a, b

$\Leftrightarrow$ $\boldsymbol{v} \cdot (a\boldsymbol{u_1} + b\boldsymbol{u_2}) = 0$ for all a, b

$\Leftrightarrow$ $\boldsymbol{v} \cdot \boldsymbol{u_1} = 0$ and $\boldsymbol{v} \cdot \boldsymbol{u_2} = 0$

$$\underbrace{w + x + y = 0 \quad \text{and} \quad -x - y + z = 0}_{\text{solve this homog. system}}$$

$\Rightarrow$ general solution

Section 5.2

Orthogonal and Orthonormal Basis

Objectives

- What is the projection of a vector onto a subspace?
- What is Gram-Schmidt Process?

Usage of the word "Orthogonal"

- A vector $\boldsymbol{u}$ is orthogonal to another vector $\boldsymbol{v}$ (same as: two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal)
- A set of vectors is orthogonal (same as: every pair of vectors in the set is orthogonal)
- A vector $\boldsymbol{u}$ is orthogonal to a subspace V (same as: $\boldsymbol{u}$ is orthogonal to every vector in subspace V)

Remark

To show a vector $\boldsymbol{v}$ is orthogonal to a subspace $V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ of $\mathbf{R}^n$

Show: $\boldsymbol{v} \cdot \boldsymbol{u}_1 = 0, \boldsymbol{v} \cdot \boldsymbol{u}_2 = 0, \ldots, \boldsymbol{v} \cdot \boldsymbol{u}_k = 0$

To find a vector $\boldsymbol{v}$ that is orthogonal to a subspace $V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ of $\mathbf{R}^n$

Let $\boldsymbol{v} = (x_1, x_2, \ldots, x_n)$ (unknowns)

Convert $\boldsymbol{v} \cdot \boldsymbol{u}_1 = 0, \boldsymbol{v} \cdot \boldsymbol{u}_2 = 0, \ldots, \boldsymbol{v} \cdot \boldsymbol{u}_k = 0$

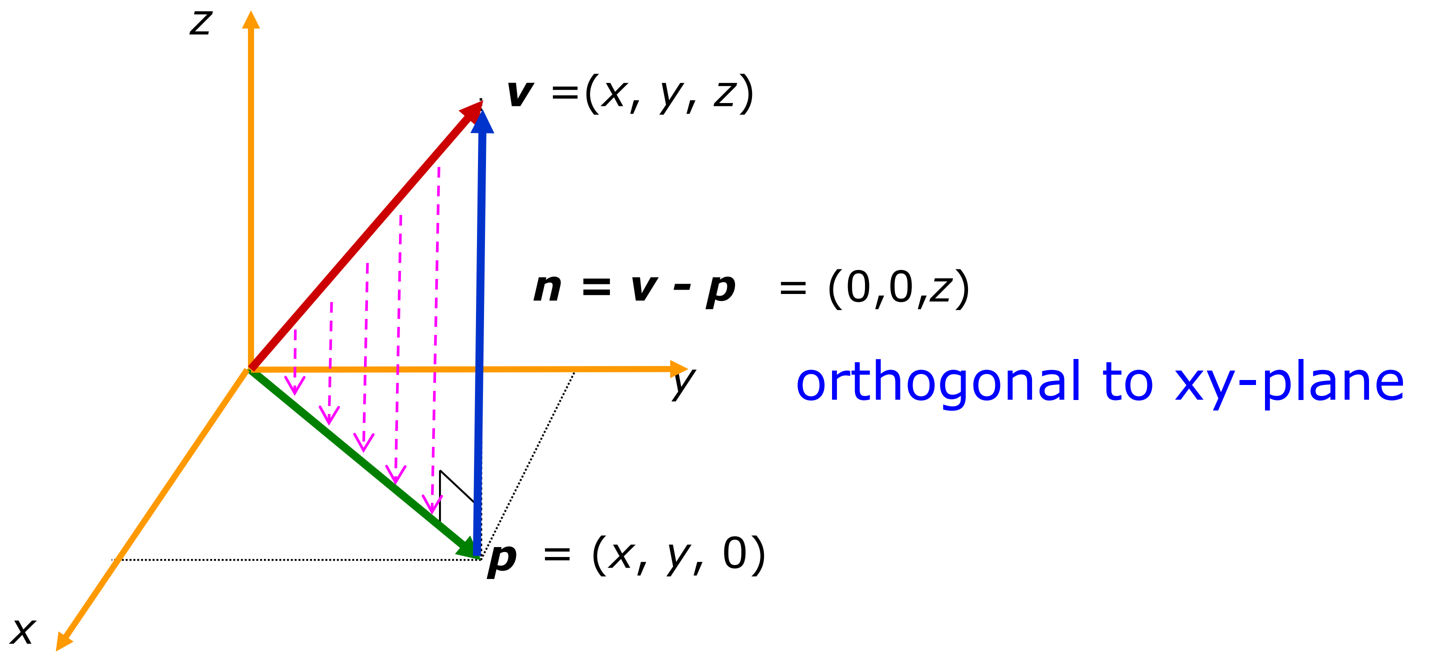
into a homogeneous system.

Solve the system.

Example 5.2.11.2

Example 5.2.14.2

The projection of $\boldsymbol{v} = (x, y, z)$ onto the xy-plane

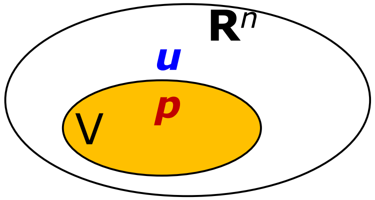


$\boldsymbol{p}$ is a projection of $\boldsymbol{v}$ onto the plane $\Longleftrightarrow$ $\boldsymbol{v} - \boldsymbol{p}$ is orthogonal to the plane

Projection of a vector onto a subspace of $\mathbf{R}^n$

Definition 5.2.13

Let V be a subspace of $\mathbf{R}^n$ and $\boldsymbol{u}$ a vector in $\mathbf{R}^n$.



Let $\boldsymbol{p}$ be a vector in V.

$\boldsymbol{p}$ is called the projection of $\boldsymbol{u}$ onto V

if $\boldsymbol{u} - \boldsymbol{p}$ is a vector orthogonal to V.

good for checking, but not finding projection.

Every vector has exactly one projection onto a given subspace.

see Ex5 Q18

How to find projection in general?

Example 5.2.16

This is the xz-plane

V = span{(1,0,1), (1,0,-1)}   a plane in $\mathbf{R}^3$

Find the projection $\boldsymbol{p}$ of $\boldsymbol{w}$ = (1, 1, 0) onto V

$\boldsymbol{u}_1 = (1, 0, 1)$ and $\boldsymbol{u}_2 = (1, 0, -1)$ orthogonal basis for V

$\boldsymbol{p}$ lies on V

$$\Rightarrow \boldsymbol{p} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 = \frac{1}{2}(1, 0, 1) + \frac{1}{2}(1, 0, -1)$$

Theorem 5.2.15

$$c_1 = \frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{||\boldsymbol{u}_1||^2} = \frac{1}{2} \qquad c_2 = \frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{||\boldsymbol{u}_2||^2} = \frac{1}{2}$$

$$= (1, 0, 0)$$

This is the projection of $\boldsymbol{w}$ onto V

Check: $\boldsymbol{w} - \boldsymbol{p}$ is orthogonal to V

How to find projection using orthogonal basis?

Theorem 5.2.15

Let V be a subspace of $\mathbf{R}^n$ and $\boldsymbol{w}$ a vector in $\mathbf{R}^n$.

1. $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$: an orthogonal basis for V, the projection $\boldsymbol{p}$ of $\boldsymbol{w}$ onto V is

Look familiar?

$$\mathbf{p} = \frac{\mathbf{w}\cdot\mathbf{u}_1}{||\mathbf{u}_1||^2}\mathbf{u}_1 + \frac{\mathbf{w}\cdot\mathbf{u}_2}{||\mathbf{u}_2||^2}\mathbf{u}_2 + \cdots + \frac{\mathbf{w}\cdot\mathbf{u}_k}{||\mathbf{u}_k||^2}\mathbf{u}_k$$

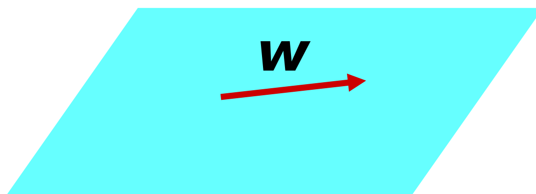
2. $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$: an orthonormal basis for V, the projection $\boldsymbol{p}$ of $\boldsymbol{w}$ onto V is

$$\boldsymbol{p} = (\boldsymbol{w}\cdot\boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{w}\cdot\boldsymbol{v}_2)\boldsymbol{v}_2 + \cdots + (\boldsymbol{w}\cdot\boldsymbol{v}_k)\boldsymbol{v}_k$$

Theorems 5.2.8 VS 5.2.15

Theorem 5.2.8

$\boldsymbol{w}$ a vector in V

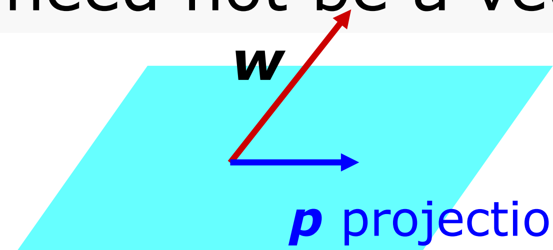


V a subspace

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$
orthogonal basis

Theorem 5.2.15

$\boldsymbol{w}$ need not be a vector in V



V a subspace

$S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$
orthogonal basis

$$\frac{\boldsymbol{w}\cdot\boldsymbol{u_1}}{||\boldsymbol{u_1}||^2}\boldsymbol{u_1} + \frac{\boldsymbol{w}\cdot\boldsymbol{u_2}}{||\boldsymbol{u_2}||^2}\boldsymbol{u_2} + \cdots + \frac{\boldsymbol{w}\cdot\boldsymbol{u_k}}{||\boldsymbol{u_k}||^2}\boldsymbol{u_k} = \begin{cases} \boldsymbol{w} & \text{if } \boldsymbol{w} \in V \\ \boldsymbol{p} & \text{if } \boldsymbol{w} \notin V \end{cases}$$

The proof

$\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$
an orthogonal basis for V

Theorem 5.2.15

Let $$\boldsymbol{p} = \frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{||\boldsymbol{u_1}||^2}\boldsymbol{u_1} + \frac{\boldsymbol{w} \cdot \boldsymbol{u_2}}{||\boldsymbol{u_2}||^2}\boldsymbol{u_2} + \cdots + \frac{\boldsymbol{w} \cdot \boldsymbol{u_k}}{||\boldsymbol{u_k}||^2}\boldsymbol{u_k}$$

Show $\boldsymbol{p}$ is the projection of $\boldsymbol{w}$ onto V

Just need to show $\boldsymbol{w} - \boldsymbol{p}$ is orthogonal to V

span$\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$

Just need to show $\boldsymbol{w} - \boldsymbol{p}$ is orthogonal to $\boldsymbol{u_i}$ for all i.

$$(\boldsymbol{w} - \boldsymbol{p}) \cdot \boldsymbol{u_1} \quad = \boldsymbol{w} \cdot \boldsymbol{u_1} - \boldsymbol{p} \cdot \boldsymbol{u_1}$$

$$\boldsymbol{p} = \frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{||\boldsymbol{u_1}||^2}\boldsymbol{u_1} + \frac{\boldsymbol{w} \cdot \boldsymbol{u_2}}{||\boldsymbol{u_2}||^2}\boldsymbol{u_2} + \cdots + + \frac{\boldsymbol{w} \cdot \boldsymbol{u_k}}{||\boldsymbol{u_k}||^2}\boldsymbol{u_k}$$

$$= \mathbf{w} \cdot \mathbf{u_1} - \frac{\mathbf{w} \cdot \mathbf{u_1}}{||\mathbf{u_1}||^2}\mathbf{u_1} \cdot \mathbf{u_1} \qquad = 0$$

How to convert a basis to an orthogonal basis?

Discussion 5.2.18.1

$\boldsymbol{u_1}$ $\quad$ $\boldsymbol{u_2} - \boldsymbol{p}$

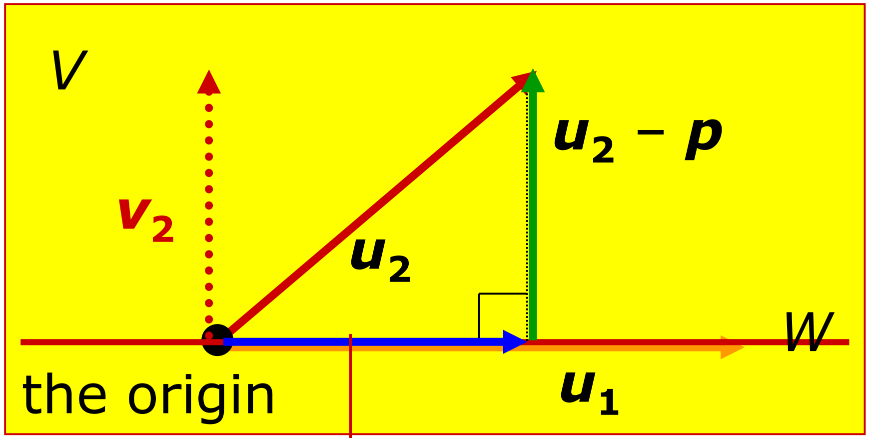
$\boldsymbol{v_1}, \boldsymbol{v_2}$

V = span{$\boldsymbol{u_1}, \boldsymbol{u_2}$} a plane

basis $\longrightarrow$ orthogonal basis

W = span{ $\boldsymbol{u_1}$} a line

projection of $\boldsymbol{u_2}$ onto W



$$\boldsymbol{p} = \frac{\boldsymbol{u}_2 \cdot \boldsymbol{u}_1}{||\boldsymbol{u}_1||^2}\boldsymbol{u}_1$$

Theorem 5.2.15

An orthogonal basis for V

$$\mathbf{v}_1 = \mathbf{u}_1$$

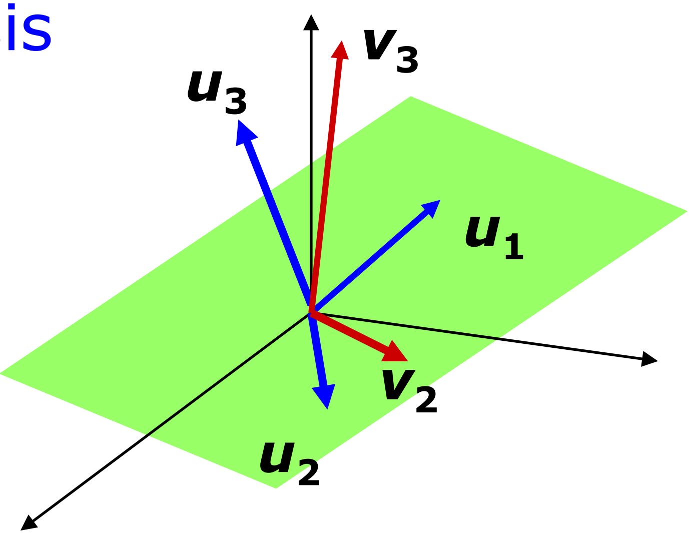
$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{||\mathbf{v}_1||^2}\mathbf{v}_1$$

How to convert a basis to an orthogonal basis?

Discussion 5.2.18.2

Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ be a basis for $\mathbf{R}^3$.

"Convert" to an orthogonal basis



$V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ a plane

$$\mathbf{v}_1 = \mathbf{u}_1$$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$$

an orthogonal basis for V

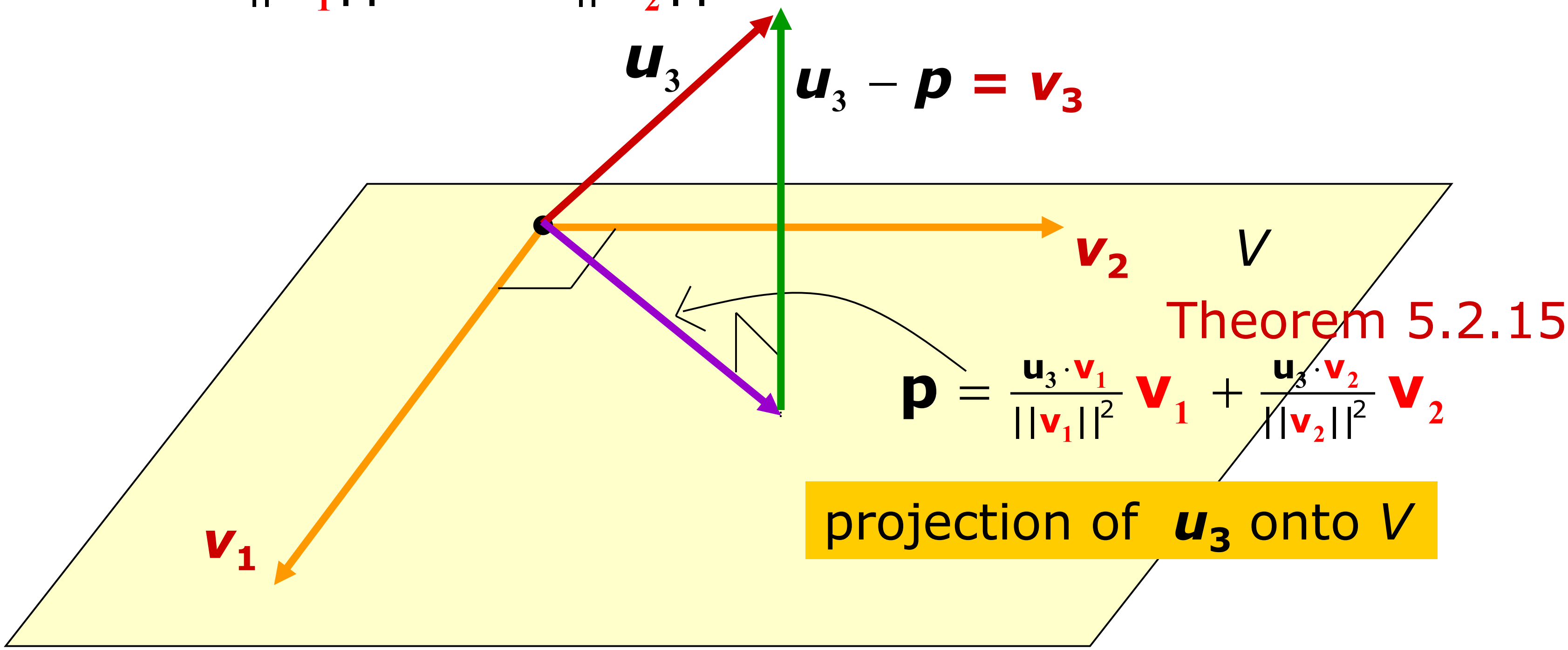
How to convert a basis to an orthogonal basis?

Discussion 5.2.18.2

$$\left.\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 \\
\mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1
\end{aligned}\right\} \text{orthogonal basis for } V$$

$$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$$

$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$: orthogonal basis for $\mathbf{R}^3$



Gram-Schmidt Process

Theorem 5.2.19

$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$: a basis for a vector space V.

Define $\mathbf{v}_1 = \mathbf{u}_1$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$$

orthogonal to $\boldsymbol{v}_1$

$$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$$

orthogonal to $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$

$$\vdots$$

$$\mathbf{v}_k = \mathbf{u}_k - \frac{\mathbf{u}_k \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\mathbf{u}_k \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \cdots - \frac{\mathbf{u}_k \cdot \mathbf{v}_{k-1}}{\|\mathbf{v}_{k-1}\|^2}\mathbf{v}_{k-1}$$

orthogonal to $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k-1}$

$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is an orthogonal basis for V.

Gram-Schmidt Process

$\{u_1, u_2, \ldots, u_k\}$
basis for V

$\{v_1, v_2, \ldots, v_k\}$
orthogonal basis for V

Theorem 5.2.19

$\{v_1, v_2, \ldots, v_k\}$ is an orthogonal basis for V.

Normalize this basis:

$$\mathbf{w}_1 = \frac{1}{||\mathbf{v}_1||}\mathbf{v}_1 \qquad \mathbf{w}_2 = \frac{1}{||\mathbf{v}_2||}\mathbf{v}_2 \qquad \cdots \qquad \mathbf{w}_k = \frac{1}{||\mathbf{v}_k||}\mathbf{v}_k$$

$\{w_1, w_2, \ldots, w_k\}$ is an orthonormal basis for V.

Example 5.2.20

$\boldsymbol{u}_1 = (1, -1, 2) \quad \boldsymbol{u}_2 = (2, 1, 0) \quad \boldsymbol{u}_3 = (0, 0, 1)$

$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is a basis for $\mathbf{R}^3$.

Apply the Gram-Schmidt Process to transform this basis into an orthonormal basis.

Gram-Schmidt Process

Example 5.2.20

$$\mathbf{u}_1 = (1, -1, 2)$$
$$\mathbf{u}_2 = (2, 1, 0)$$
$$\mathbf{u}_3 = (0, 0, 1)$$

Visualization tool

$$\mathbf{v}_1 = \mathbf{u}_1 = (1, -1, 2)$$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$$
$$= (2, 1, 0) - \tfrac{1}{6}(1, -1, 2) = (\tfrac{11}{6}, \tfrac{7}{6}, -\tfrac{1}{3})$$

$$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$$
$$= (0, 0, 1) - \tfrac{2}{6}(1, -1, 2) - \tfrac{-1/3}{29/6}(\tfrac{11}{6}, \tfrac{7}{6}, -\tfrac{1}{3})$$
$$= (-\tfrac{6}{29}, \tfrac{12}{29}, \tfrac{9}{29})$$

$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is an orthogonal basis for $\mathbf{R}^3$.

Gram-Schmidt Process

Example 5.2.20

$$\mathbf{w}_1 = \frac{1}{||\mathbf{v}_1||}\mathbf{v}_1 = \tfrac{1}{\sqrt{6}}(1,-1,2) = \left(\tfrac{1}{\sqrt{6}}, -\tfrac{1}{\sqrt{6}}, \tfrac{2}{\sqrt{6}}\right)$$

$$\mathbf{w}_2 = \frac{1}{||\mathbf{v}_2||}\mathbf{v}_2 = \tfrac{1}{\sqrt{29/6}}\left(\tfrac{11}{6}, \tfrac{7}{6}, -\tfrac{1}{3}\right) = \left(\tfrac{11}{\sqrt{174}}, \tfrac{7}{\sqrt{174}}, -\tfrac{2}{\sqrt{174}}\right)$$

$$\mathbf{w}_3 = \frac{1}{||\mathbf{v}_3||}\mathbf{v}_3 = \tfrac{1}{\sqrt{9/29}}\left(-\tfrac{6}{29}, \tfrac{12}{29}, \tfrac{9}{29}\right) = \left(-\tfrac{2}{\sqrt{29}}, \tfrac{4}{\sqrt{29}}, \tfrac{3}{\sqrt{29}}\right)$$

$\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}$ is an orthonormal basis for $\mathbf{R}^3$.

Section 5.3

Best Approximations

Objectives

• What is a Least Squares solution ?

• How to find the best approximate solution to inconsistent system?

Discussion 5.3.1

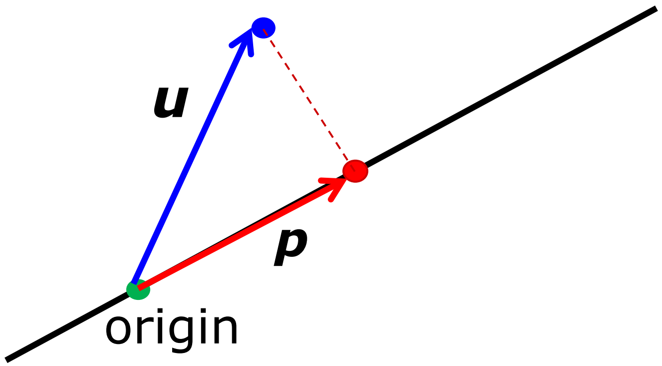
orthogonality $\xrightarrow{\text{applications}}$ study of approximations

Linear system

- consistent — get exact solutions
- inconsistent — get approximate solutions

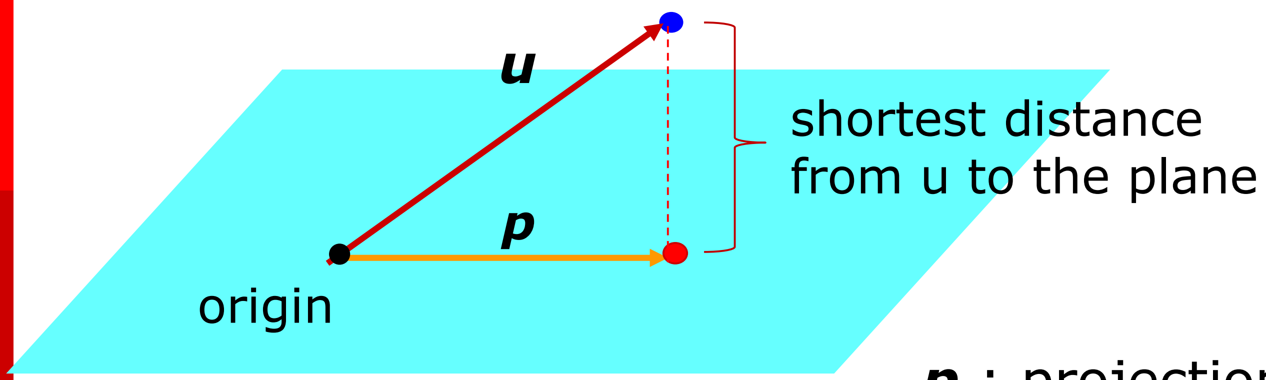
Finding the "best approximation" of a vector from a subspace

Nearest point



p : projection of ***u*** onto the line

We say: ***p*** is the best approximation of ***u*** in the line



Example 5.3.3

p : projection of ***u*** onto the plane

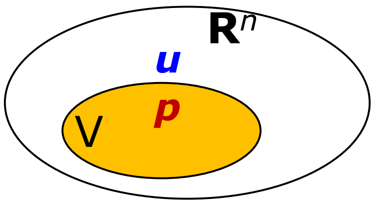
We say: ***p*** is the best approximation of ***u*** in the plane

Finding the "best approximation" of a vector from a subspace

Theorem 5.3.2

V : subspace in $\mathbf{R}^n$ and $\boldsymbol{u} \in \mathbf{R}^n$.

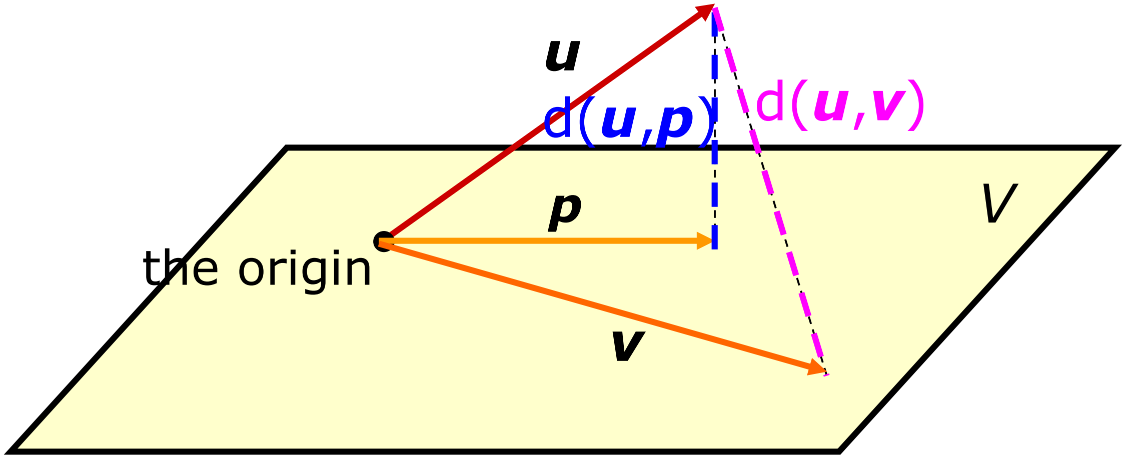
need not be a line or plane



$\boldsymbol{p}$: projection of $\boldsymbol{u}$ onto V

Then $d(\boldsymbol{u}, \boldsymbol{p}) \le d(\boldsymbol{u}, \boldsymbol{v})$ for any vector $\boldsymbol{v}$ in V

i.e. $\boldsymbol{p}$ is the best approximation of $\boldsymbol{u}$ in V.



Finding the "best approximation" of a vector from a subspace

Theorem 5.3.2

V : subspace in $\mathbf{R}^n$ and $\boldsymbol{u} \in \mathbf{R}^n$.

need not be a line or plane

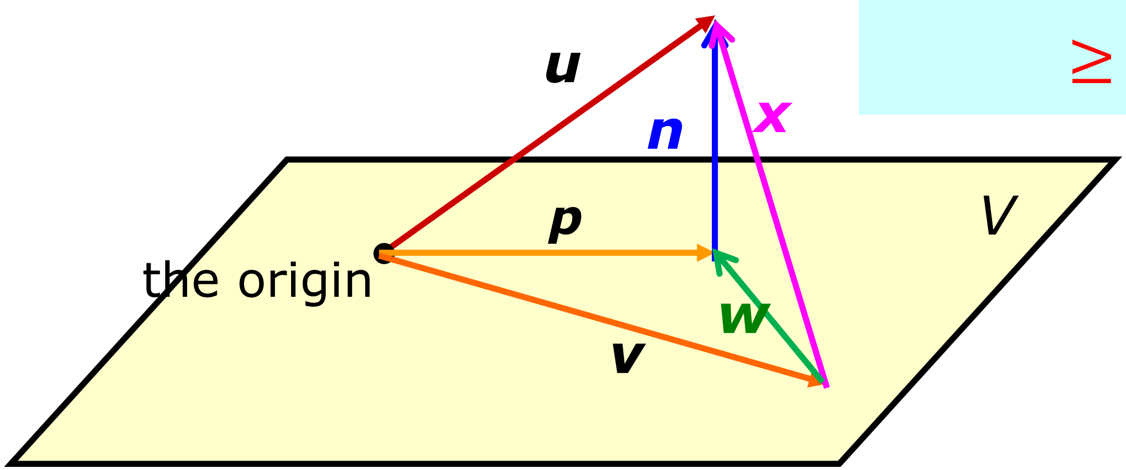
$\boldsymbol{p}$: projection of $\boldsymbol{u}$ onto V

Then $d(\boldsymbol{u}, \boldsymbol{p}) \le d(\boldsymbol{u}, \boldsymbol{v})$ for any vector $\boldsymbol{v}$ in V

$||\boldsymbol{n}|| \le ||\boldsymbol{x}||$

$$||\boldsymbol{x}||^2 = ||\boldsymbol{n} + \boldsymbol{w}||^2$$
$$= ||\boldsymbol{n}||^2 + ||\boldsymbol{w}||^2$$
$$\ge ||\boldsymbol{n}||^2$$

(see Ex 5 Q9)



Inconsistent system

$t = cr^2 + ds + e$

Example 5.3.5 experimental errors

6 equations
3 unknowns c, d, e

system $\boldsymbol{Ax} = \boldsymbol{b}$

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \\ 4 & 1 & 1 \\ 4 & 2 & 1 \end{pmatrix} \qquad \boldsymbol{b} = \begin{pmatrix} 0.5 \\ 1.6 \\ 2.8 \\ 0.8 \\ 5.1 \\ 5.9 \end{pmatrix}$$

This system is inconsistent $\boldsymbol{Ax} - \boldsymbol{b} \neq \boldsymbol{0}$

Find the best approximate solution see example 5.3.11.2

Find $\boldsymbol{x}_0$ such that $||\boldsymbol{Ax}_0 - \boldsymbol{b}||$ is the smallest

$\sqrt{\text{sum of squares}}$

Such an $\boldsymbol{x}_0$ is called
a least squares solution to the system $\boldsymbol{Ax} = \boldsymbol{b}$.

What is a least squares solution?

Definition 5.3.6

A least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$ ($\boldsymbol{A}$: m×n)

is a vector $\boldsymbol{u}$ in $\mathbf{R}^n$ that minimize $||\boldsymbol{b} - \boldsymbol{Ax}||$

i.e. $||\boldsymbol{b} - \boldsymbol{Au}|| \leq ||\boldsymbol{b} - \boldsymbol{Av}||$ for all $\boldsymbol{v}$ in $\mathbf{R}^n$

good for intuition, but not finding this approximation.

working definition

new linear system

is an actual solution of $\boxed{\boldsymbol{A}^{T}\boldsymbol{Ax} = \boldsymbol{A}^{T}\boldsymbol{b}}$

Theorem 5.3.10

Finding least squares solution

Exercise 5 Q24

$$\begin{cases} x + y + z = 1 \\ \quad\;\; y + z = 1 \\ x - y - z = 1 \\ \qquad\quad\; z = 1 \end{cases}$$

$\boldsymbol{Ax} = \boldsymbol{b}$ inconsistent

does not satisfy

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \mathbf{A}^T = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & 1 & -1 & 1 \end{pmatrix}$$

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 3 \\ 0 & 3 & 4 \end{pmatrix} \qquad \mathbf{A}^T\mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \qquad \boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b} \text{ consistent}$$

satisfies

$$\mathbf{u} = \begin{pmatrix} 1 \\ -\frac{2}{3} \\ 1 \end{pmatrix}$$

Theorem 5.3.10

u gives the least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$

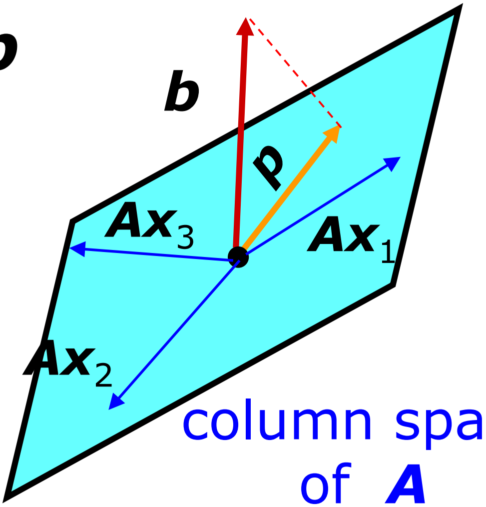
Projection of *b* onto the column space of *A*

Discussion 5.3.7

Find least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$

Find $\boldsymbol{u}$ that minimize $||\boldsymbol{b} - \boldsymbol{Ax}||$

the projection $\boldsymbol{p}$ of $\boldsymbol{b}$
onto the column space of $\boldsymbol{A}$



Find $\boldsymbol{u}$ such that $\boldsymbol{Au} = \boldsymbol{p}$

This system is always consistent

$\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix}$ $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$: columns of $\boldsymbol{A}$ $\mathbf{x} = \begin{pmatrix} c \\ d \\ e \end{pmatrix}$

$\Rightarrow \boldsymbol{Ax} = c\boldsymbol{u_1} + d\boldsymbol{u_2} + e\boldsymbol{u_3}$ linear comb of columns of $\boldsymbol{A}$

All $\boldsymbol{Ax}$ belong to column space of $\boldsymbol{A}$ Discussion 4.1.16

Least squares solutions and projection

Theorem 5.3.8

$\boldsymbol{u}$ is a least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$

$\Leftrightarrow$ $\boldsymbol{u}$ is a solution of $\boldsymbol{Ax} = \boldsymbol{p}$

$\Leftrightarrow$ $\boldsymbol{Au} = \boldsymbol{p}$

$\boldsymbol{p}$: projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$

Alternative way to find least squares solution:
If we know
the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$,
then
we can find the least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$.

Use projection to find least squares solution

Example 5.3.9

A(least squares solution) **=** projection

Find the least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Solve $\boldsymbol{Ax} = \boldsymbol{p}$

$$\mathbf{x} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

column space of ***A***

$$\mathrm{V} = \mathrm{span}\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$$

projection of ***b*** onto the column space of ***A***

$$\mathbf{p} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

see example 5.3.3

Use least squares solution to find projection

Example 5.3.11. $\boldsymbol{A}$(least squares solution) = projection

Find the projection of (1,1,1,1) onto

V = span{(1,-1,1,-1), (1,2,0,1), (2,1,1,0)}

Form matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$ $\quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

First find the least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$

Solve $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$ Theorem 5.3.10

$$\mathbf{x} = \begin{pmatrix} -t + \frac{2}{5} \\ -t + \frac{4}{5} \\ t \end{pmatrix} \longrightarrow \text{Take } \mathbf{u} = \begin{pmatrix} \frac{2}{5} \\ \frac{4}{5} \\ 0 \end{pmatrix} \longrightarrow \mathbf{Au} = \begin{pmatrix} \frac{6}{5} \\ \frac{6}{5} \\ \frac{2}{5} \\ \frac{2}{5} \end{pmatrix}$$

Solution of $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b} \Leftrightarrow$ least squares solution of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

Theorem 5.3.10

V = column space of $\boldsymbol{A}$

$\mathbf{A} = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3)$

$\boldsymbol{u}$ is the least squares solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

if and only if $\boldsymbol{u}$ is a solution of $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$

$\Leftrightarrow$ $\boldsymbol{A}\boldsymbol{u}$ is the projection of $\boldsymbol{b}$ onto V

$\Leftrightarrow$ $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{u}$ is orthogonal to V

$\Leftrightarrow$ $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{u}$ is orthogonal to $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$

$$\Leftrightarrow \mathbf{A}^T(\mathbf{b} - \mathbf{A}\mathbf{u}) = \begin{pmatrix} \mathbf{a}_1{}^T \\ \mathbf{a}_2{}^T \\ \mathbf{a}_3{}^T \end{pmatrix} (\mathbf{b} - \mathbf{A}\mathbf{u}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$\Leftrightarrow$ $\boldsymbol{A}^T\boldsymbol{b} - \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{u} = \mathbf{0}$

$\Leftrightarrow$ $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{u} = \boldsymbol{A}^T\boldsymbol{b}$

Solution of $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b} \Leftrightarrow$ least squares solution of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

Theorem 5.3.10

always exists

$\boldsymbol{u}$ is the least squares solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

$\Leftrightarrow$ $\boldsymbol{u}$ is a solution of $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$ always consistent

$\Leftrightarrow$ $\boldsymbol{u}$ is a solution of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{p}$ always consistent

where $\boldsymbol{p}$ is the projection of $\boldsymbol{b}$ onto column space of $\boldsymbol{A}$

Theorem 5.3.8

Section 5.4

Another usage of "orthogonal"

Orthogonal Matrices

Objective

- What is an orthogonal matrix?
- How is orthogonal matrix related to orthonormal basis?
- How is transition matrix related to orthogonal matrix?

What is an orthogonal matrix?

Definition 5.4.3 & Remark 5.4.4

A square matrix $\boldsymbol{A}$ is called an orthogonal matrix

if $\boldsymbol{A}^{-1} = \boldsymbol{A}^{T}$

Equivalently (and more easily),

if $\boldsymbol{A}\boldsymbol{A}^{T} = \boldsymbol{I}$ (or $\boldsymbol{A}^{T}\boldsymbol{A} = \boldsymbol{I}$).

See Ex 2.12

All orthogonal matrices are invertible.

What is an orthogonal matrix?

Example 5.4.5

These are orthogonal matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

rotation anticlockwise through angle θ

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

inverse of each other

(multiply them to check)

Their transposes:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

rotation clockwise through angle θ

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Their transposes are also orthogonal matrices

Orthogonal matrix vs orthonormal basis for $\mathbf{R}^n$

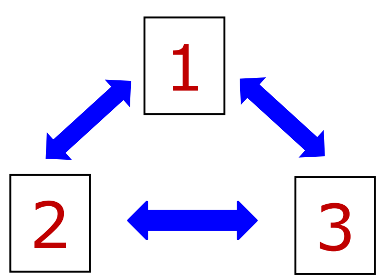
Theorem 5.4.6

Let $\boldsymbol{A}$ be a square matrix of order n.

The following statements are equivalent:

1. $\boldsymbol{A}$ is an orthogonal matrix.
2. The rows of $\boldsymbol{A}$ form an orthonormal basis for $\mathbf{R}^n$.
3. The columns of $\boldsymbol{A}$ form an orthonormal basis for $\mathbf{R}^n$.

Shall prove
(1) $\Leftrightarrow$ (2)
and
(1) $\Leftrightarrow$ (3)



The proof

1. A is orthogonal
2. The rows of A form an orthonormal basis for $\mathbf{R}^n$

Theorem 5.4.6 (1 ⇔ 2)

For $i = 1, 2, \ldots, n$, let $\boldsymbol{a}_i$ be the ith row of A.

$$\boldsymbol{A} = \begin{array}{c} \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix} \end{array} \begin{array}{c} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \boldsymbol{a}_3 \end{array} \qquad \boldsymbol{A}^T = \begin{array}{ccc} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \\ \boldsymbol{a}_1^T \quad \boldsymbol{a}_2^T \quad \boldsymbol{a}_3^T \end{array}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{A}\boldsymbol{A}^T = \begin{pmatrix} \boldsymbol{a}_1\boldsymbol{a}_1^T & \boldsymbol{a}_1\boldsymbol{a}_2^T & \boldsymbol{a}_1\boldsymbol{a}_3^T \\ \boldsymbol{a}_2\boldsymbol{a}_1^T & \boldsymbol{a}_2\boldsymbol{a}_2^T & \boldsymbol{a}_2\boldsymbol{a}_3^T \\ \boldsymbol{a}_3\boldsymbol{a}_1^T & \boldsymbol{a}_3\boldsymbol{a}_2^T & \boldsymbol{a}_3\boldsymbol{a}_3^T \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{a}_1 & \boldsymbol{a}_1 \cdot \boldsymbol{a}_2 & \boldsymbol{a}_1 \cdot \boldsymbol{a}_3 \\ \boldsymbol{a}_2 \cdot \boldsymbol{a}_1 & \boldsymbol{a}_2 \cdot \boldsymbol{a}_2 & \boldsymbol{a}_2 \cdot \boldsymbol{a}_3 \\ \boldsymbol{a}_3 \cdot \boldsymbol{a}_1 & \boldsymbol{a}_3 \cdot \boldsymbol{a}_2 & \boldsymbol{a}_3 \cdot \boldsymbol{a}_3 \end{pmatrix}$$

orthogonal

$\boldsymbol{a}_i \cdot \boldsymbol{a}_i = 1$ for all i ⇔ $\|\boldsymbol{a}_i\| = 1$

$\boldsymbol{a}_i \cdot \boldsymbol{a}_j = 0$ for $i \neq j$ ⇔ $\boldsymbol{a}_i$ and $\boldsymbol{a}_j$ orthogonal

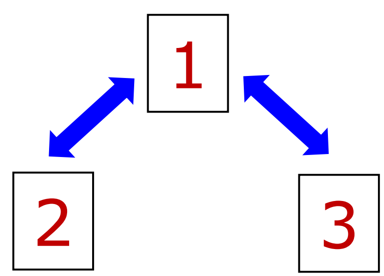
$\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$ is an orthonormal basis for $\mathbf{R}^n$

The proof

Theorem 5.4.6

1. $\boldsymbol{A}$ is an orthogonal matrix. $\Leftrightarrow \boldsymbol{A}^{\mathrm{T}}$ is orthogonal matrix
2. The rows of $\boldsymbol{A}$ form an orthonormal basis for $\mathbf{R}^n$.
3. The columns of $\boldsymbol{A}$ form an orthonormal basis for $\mathbf{R}^n$.

We have proven
(1) ⇔ (2)



Use $\boldsymbol{A}^{\mathrm{T}}$ to derive
(1) ⇔ (3)

Transition matrix revisited

Discussion 5.4.1

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ be two bases for a vector space V.

> Procedure to compute transition matrix $\boldsymbol{P}$ from S to T:
>
> (i) write each $\boldsymbol{u}_i$ as a linear combination of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$.
>
> (ii) use the coordinate vector $[\boldsymbol{u}_i]_T$ as the i th column $\boldsymbol{P}$.

$$\boldsymbol{P} = (\, [\boldsymbol{u}_1]_T \;\; [\boldsymbol{u}_2]_T \cdots [\boldsymbol{u}_k]_T \,)$$

For any vector $\boldsymbol{w}$ in V, $[\boldsymbol{w}]_T = \boldsymbol{P}\,[\boldsymbol{w}]_S$.

Example 5.4.2

$S = \{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\}$: the standard basis for $\mathbf{R}^3$

$$\mathbf{e}_1 = (1, 0, 0) \quad \mathbf{e}_2 = (0, 1, 0) \quad \mathbf{e}_3 = (0, 0, 1)$$

$T = \{\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}\}$: orthonormal basis for $\mathbf{R}^3$

$$\mathbf{u}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \quad \mathbf{u}_2 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right) \quad \mathbf{u}_3 = \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

Transition matrix from T to S

$$\begin{cases} \mathbf{u}_1 = \frac{1}{\sqrt{3}}\mathbf{e}_1 + \frac{1}{\sqrt{3}}\mathbf{e}_2 + \frac{1}{\sqrt{3}}\mathbf{e}_3 \\ \mathbf{u}_2 = \frac{1}{\sqrt{2}}\mathbf{e}_1 - \frac{1}{\sqrt{2}}\mathbf{e}_3 \\ \mathbf{u}_3 = \frac{1}{\sqrt{6}}\mathbf{e}_1 - \frac{2}{\sqrt{6}}\mathbf{e}_2 + \frac{1}{\sqrt{6}}\mathbf{e}_3 \end{cases}$$

$$\boldsymbol{P} = \begin{array}{c} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix} \\ \begin{array}{ccc} \mathbf{u}_1^T & \mathbf{u}_2^T & \mathbf{u}_3^T \end{array} \end{array}$$

Example 5.4.2

$S = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$: the standard basis for $\mathbf{R}^3$

$$\mathbf{e}_1 = (1, 0, 0) \quad \mathbf{e}_2 = (0, 1, 0) \quad \mathbf{e}_3 = (0, 0, 1)$$

$T = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$: orthonormal basis for $\mathbf{R}^3$

$$\mathbf{u}_1 = \left(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\right) \quad \mathbf{u}_2 = \left(\tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}}\right) \quad \mathbf{u}_3 = \left(\tfrac{1}{\sqrt{6}}, -\tfrac{2}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}\right)$$

Transition matrix from S to T

$$\begin{array}{l} \qquad\qquad \mathbf{e}_1\cdot\mathbf{u}_1 \qquad\quad \mathbf{e}_1\cdot\mathbf{u}_2 \qquad\quad \mathbf{e}_1\cdot\mathbf{u}_3 \\ \begin{cases} \mathbf{e}_1 = \frac{1}{\sqrt{3}}\mathbf{u}_1 + \frac{1}{\sqrt{2}}\mathbf{u}_2 + \frac{1}{\sqrt{6}}\mathbf{u}_3 \\ \mathbf{e}_2 = \frac{1}{\sqrt{3}}\mathbf{u}_1 - \frac{2}{\sqrt{6}}\mathbf{u}_3 \\ \mathbf{e}_3 = \frac{1}{\sqrt{3}}\mathbf{u}_1 - \frac{1}{\sqrt{2}}\mathbf{u}_2 + \frac{1}{\sqrt{6}}\mathbf{u}_3 \end{cases} \end{array}$$

$$Q = \begin{array}{cc} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} & \begin{array}{c} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \end{array} \end{array}$$

Theorem 5.2.8.2

Transition matrix between orthonormal bases

Example 5.4.2

Transition matrix from T to S

$$P = \begin{array}{c} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix} \\ \begin{array}{ccc} \mathbf{u}_1^T & \mathbf{u}_2^T & \mathbf{u}_3^T \end{array} \end{array}$$

Transition matrix from S to T

$$Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{array}{c} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \end{array}$$

By theorem 3.7.5

$$\left.\begin{array}{l} Q = P^{-1} \\ Q = P^{T} \end{array}\right\} \quad P^{-1} = P^{T}$$

S: orthonormal basis

T: orthonormal basis

So $\boldsymbol{P}$ is an orthogonal matrix

Transition matrix between orthonormal bases

Theorem 5.4.7

S and T: two orthonormal bases for a vector space.

The transition matrix $\boldsymbol{P}$ from S to T is orthogonal.

So $\boldsymbol{P}^T$ is the transition matrix from T to S.

Example 5.4.8.2

S : $\mathbf{u}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ $\quad \mathbf{u}_2 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$ $\quad \mathbf{u}_3 = \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$

T : $\mathbf{v}_1 = (0, 0, 1)$ $\quad \mathbf{v}_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)$ $\quad \mathbf{v}_3 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$

transition matrix from S to T

$$\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{2} & \frac{3}{\sqrt{12}} \\ \frac{2}{\sqrt{6}} & \frac{1}{2} & -\frac{1}{\sqrt{12}} \end{pmatrix}$$

take transpose

transition matrix from T to S

$$\mathbf{P}^T = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{6}} & \frac{3}{\sqrt{12}} & -\frac{1}{\sqrt{12}} \end{pmatrix}$$

The proof

Theorem 5.4.7

S and T are orthonormal bases

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$.

By Theorem 5.2.8.2

$$\begin{cases} \mathbf{u}_1 = \boxed{(\mathbf{u}_1 \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u}_1 \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{u}_1 \cdot \mathbf{v}_k)\mathbf{v}_k} \\ \mathbf{u}_2 = \boxed{(\mathbf{u}_2 \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u}_2 \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{u}_2 \cdot \mathbf{v}_k)\mathbf{v}_k} \\ \quad \vdots \\ \mathbf{u}_k = \boxed{(\mathbf{u}_k \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u}_k \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{u}_k \cdot \mathbf{v}_k)\mathbf{v}_k} \end{cases}$$

The transition matrix from S to T is

$$\mathbf{P} = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{v}_1 & \mathbf{u}_2 \cdot \mathbf{v}_1 & \cdots & \mathbf{u}_k \cdot \mathbf{v}_1 \\ \mathbf{u}_1 \cdot \mathbf{v}_2 & \mathbf{u}_2 \cdot \mathbf{v}_2 & \cdots & \mathbf{u}_k \cdot \mathbf{v}_2 \\ \vdots & \vdots & & \vdots \\ \mathbf{u}_1 \cdot \mathbf{v}_k & \mathbf{u}_2 \cdot \mathbf{v}_k & \cdots & \mathbf{u}_k \cdot \mathbf{v}_k \end{pmatrix}$$

The proof

Theorem 5.4.7

S and T are orthonormal bases

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$.

By Theorem 5.2.8.2

$$\begin{cases} \mathbf{v}_1 = \boxed{(\mathbf{v}_1 \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_1 \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}_1 \cdot \mathbf{u}_k)\mathbf{u}_k} \\ \mathbf{v}_2 = \boxed{(\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_2 \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}_2 \cdot \mathbf{u}_k)\mathbf{u}_k} \\ \qquad \vdots \\ \mathbf{v}_k = \boxed{(\mathbf{v}_k \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_k \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}_k \cdot \mathbf{u}_k)\mathbf{u}_k} \end{cases}$$

The transition matrix from T to S is

$$\mathbf{Q} = \begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{u}_1 & \mathbf{v}_2 \cdot \mathbf{u}_1 & \cdots & \mathbf{v}_k \cdot \mathbf{u}_1 \\ \mathbf{v}_1 \cdot \mathbf{u}_2 & \mathbf{v}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{v}_k \cdot \mathbf{u}_2 \\ \vdots & \vdots & & \vdots \\ \mathbf{v}_1 \cdot \mathbf{u}_k & \mathbf{v}_2 \cdot \mathbf{u}_k & \cdots & \mathbf{v}_k \cdot \mathbf{u}_k \end{pmatrix}$$

The proof

Theorem 5.4.7

S and T are orthonormal bases

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k\}$.

transition matrix from S to T $\longleftrightarrow$ inverse of each other $\longleftrightarrow$ transition matrix from T to S

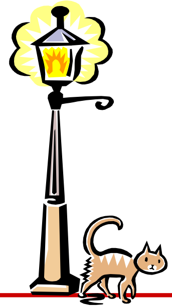
$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{u}_1 \cdot \boldsymbol{v}_1 & \boldsymbol{u}_2 \cdot \boldsymbol{v}_1 & \cdots & \boldsymbol{u}_k \cdot \boldsymbol{v}_1 \\ \boldsymbol{u}_1 \cdot \boldsymbol{v}_2 & \boldsymbol{u}_2 \cdot \boldsymbol{v}_2 & \cdots & \boldsymbol{u}_k \cdot \boldsymbol{v}_2 \\ \vdots & \vdots & & \vdots \\ \boldsymbol{u}_1 \cdot \boldsymbol{v}_k & \boldsymbol{u}_2 \cdot \boldsymbol{v}_k & \cdots & \boldsymbol{u}_k \cdot \boldsymbol{v}_k \end{pmatrix} \qquad \mathbf{Q} = \begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{u}_1 & \mathbf{v}_2 \cdot \mathbf{u}_1 & \cdots & \mathbf{v}_k \cdot \mathbf{u}_1 \\ \mathbf{v}_1 \cdot \mathbf{u}_2 & \mathbf{v}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{v}_k \cdot \mathbf{u}_2 \\ \vdots & \vdots & & \vdots \\ \mathbf{v}_1 \cdot \mathbf{u}_k & \mathbf{v}_2 \cdot \mathbf{u}_k & \cdots & \mathbf{v}_k \cdot \mathbf{u}_k \end{pmatrix}$$

We have $\boldsymbol{Q} = \boldsymbol{P}^T$

We also have $\boldsymbol{Q} = \boldsymbol{P}^{-1}$

So $\boldsymbol{P}^{-1} = \boldsymbol{P}^T$, i.e. $\boldsymbol{P}$ is orthogonal.

Rotation of *xy*-coordinates

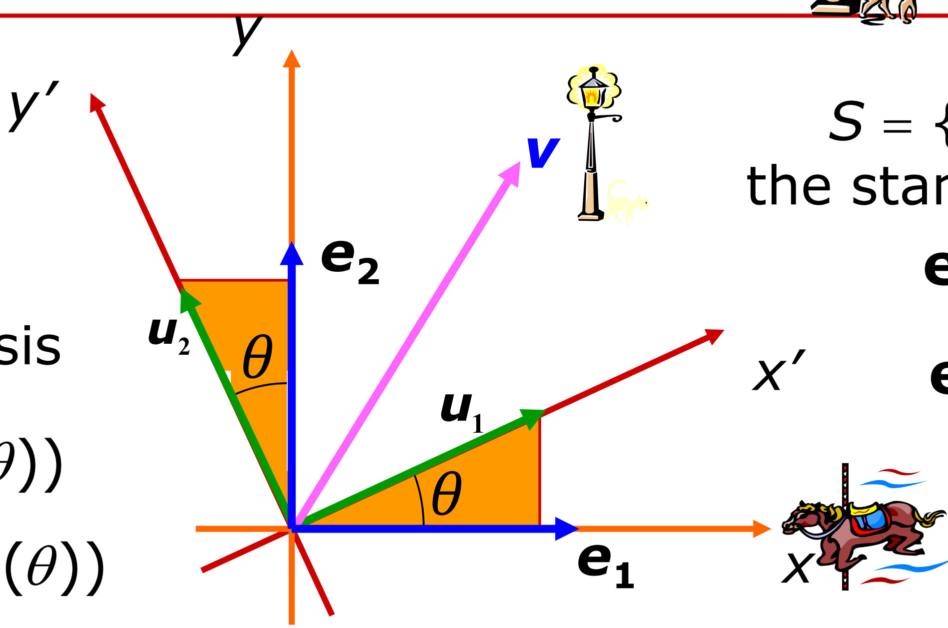


Example 5.4.8.1

$T = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$
an orthonormal basis

$$\boldsymbol{u}_1 = (\cos(\theta), \sin(\theta))$$

$$\boldsymbol{u}_2 = (-\sin(\theta), \cos(\theta))$$



$S = \{\mathbf{e}_1, \mathbf{e}_2\}$
the standard basis

$$\mathbf{e}_1 = (1, 0)$$

$$\mathbf{e}_2 = (0, 1)$$

new *x′y′*-coordinate system

What is the coordinate of $\boldsymbol{v}$ w.r.t. the new coordinate system? Ans: $[\boldsymbol{v}]_T$

What is the transition matrix between *S* and *T* ?

Rotation of *xy*-coordinates

Example 5.4.8.1

$\mathbf{u}_1 = (\cos(\theta), \sin(\theta))$

$\mathbf{u}_2 = (-\sin(\theta), \cos(\theta))$

$S = \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$
the standard basis

$T = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$
an orthonormal basis

transition matrix from T to S

$$\boldsymbol{P} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

transition matrix from S to T

$$\boldsymbol{P}^T = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

What is the coordinate of $\boldsymbol{v}$ w.r.t. the new coordinate system?

$$[\boldsymbol{v}]_T = \boldsymbol{P}^T[\boldsymbol{v}]_S$$

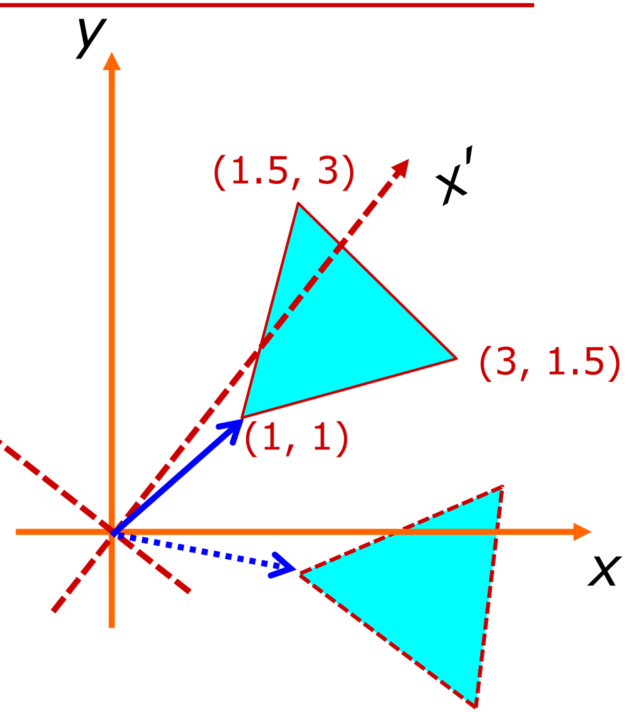
coordinates of $\boldsymbol{v}$ in the new x'y'-coordinate system

usual coordinates of $\boldsymbol{v}$

Rotation of *xy*-coordinates

Quiz Time

A new $x'y'$-coordinate system is obtained by rotating the xy-coordinate anti-clockwise by 60°.
What is the $x'y'$-coordinates of vector (1,1)?



$$\boxed{\begin{pmatrix}\cos(\frac{\pi}{3}) & \sin(\frac{\pi}{3})\\ -\sin(\frac{\pi}{3}) & \cos(\frac{\pi}{3})\end{pmatrix}}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}\frac{1+\sqrt{3}}{2}\\ \frac{1-\sqrt{3}}{2}\end{pmatrix} = \begin{pmatrix}1.366\\-0.366\end{pmatrix}$$

$$= \left(\begin{array}{c|c|c}\frac{1+\sqrt{3}}{2} & \frac{3\sqrt{3}+1.5}{2} & \frac{1.5+3\sqrt{3}}{2}\\ \frac{1-\sqrt{3}}{2} & \frac{1.5-3\sqrt{3}}{2} & \frac{3-1.5\sqrt{3}}{2}\end{array}\right)$$

Same effect as fixing the xy-coordinate and rotate the vector clockwise by 60°.

Section 6.1

Eigenvalues and Eigenvectors

Objectives

- What are Eigenvalues, Eigenvectors and Eigenspace?
- How to find eigenvalues and eigenvectors of a matrix?
- How is eigenvalue related to invertibility of matrix?

Google page rank

Google ranks webpages according to “hyperlinks”

e.g. we want to rank 4 webpages: A, B, C, D

Form a 4x4 matrix:

$$\begin{array}{cc} & \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 1 & \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{1}{2} & 0 \end{pmatrix} \end{array} \xrightarrow{\text{eigenvector}} \begin{pmatrix} 0.446 \\ 0.223 \\ 0.743 \\ 0.446 \end{pmatrix} \quad \begin{array}{l} \text{page rank} \\ 2 \text{ (tie)} \\ 4 \\ 1 \\ 2 \text{ (tie)} \end{array}$$

A has a link to C, but not to B and D

B has a link to A, C, D

Power of matrices revisited

Example 6.1.1

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} \qquad \boxed{\boldsymbol{A}^n = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}^n}$$

"Factorize" $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}}_{\boldsymbol{P}} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}}_{\boldsymbol{D}} \underbrace{\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1}}_{\boldsymbol{P}^{-1}}$$

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$$

$$\boldsymbol{A}^n = (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})^n \quad \neq \boldsymbol{P}^n\boldsymbol{D}^n\boldsymbol{P}^{-n}$$

$$= (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}) \cdots (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}) \quad (n \text{ times})$$

$$= \boldsymbol{P}\boldsymbol{D}\,\boldsymbol{D} \cdots \boldsymbol{D}\boldsymbol{P}^{-1}$$

$$= \boldsymbol{P}\boldsymbol{D}^n\boldsymbol{P}^{-1}$$

Power of matrices revisited

Example 6.1.1

$$\boldsymbol{A}^n = \boldsymbol{P}\boldsymbol{D}^n\boldsymbol{P}^{-1}$$

$$\boldsymbol{D}^n = \begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}^n = \begin{pmatrix} 1^n & 0 \\ 0 & 0.95^n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0.95^n \end{pmatrix}$$

$$\begin{aligned} \boldsymbol{A}^{100} &= \boldsymbol{P}\boldsymbol{D}^{100}\boldsymbol{P}^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0.95^{100} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0.2047 & 0.1988 \\ 0.7953 & 0.8012 \end{pmatrix} \end{aligned}$$

Remark 6.1.2

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix} \boxed{\begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}} \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1}$$

This is called "diagonalizing" a square matrix.

We need the concept of eigenvalues and eigenvectors.

What are eigenvalue and eigenvector?

Definition 6.1.3

$$\boldsymbol{A}\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}2\\4\\6\end{pmatrix}$$

relation?

Let $\boldsymbol{A}$ be a square matrix of order n.

Let $\boldsymbol{x}$ be a nonzero (column) vector in $\mathbf{R}^n$

If $\boldsymbol{Ax}$ = scalar multiple of $\boldsymbol{x}$ — $\boldsymbol{Ax}$ and $\boldsymbol{x}$ are parallel

$= \lambda\boldsymbol{x}$ for some scalar λ — lambda

then $\boldsymbol{x}$ is called an eigenvector of $\boldsymbol{A}$

The scalar λ is called an eigenvalue of $\boldsymbol{A}$

and $\boldsymbol{x}$ is said to be an eigenvector of $\boldsymbol{A}$

associated with the eigenvalue λ.

Visual

What are eigenvalue and eigenvector?

Example 6.1.4.1

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} \qquad \boldsymbol{x} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \qquad \boldsymbol{y} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\boldsymbol{Ax} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}\begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \mathbf{x}$$

$\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$ with the eigenvalue 1.

$$\boldsymbol{Ay} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0.95 \\ -0.95 \end{pmatrix} = 0.95\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0.95\boldsymbol{y}$$

$\boldsymbol{y}$ is an eigenvector of $\boldsymbol{A}$ with the eigenvalue 0.95.

$$\boxed{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1}}$$

What are eigenvalue and eigenvector?

Example 6.1.4.2

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \mathbf{z} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

$$\mathbf{Bx} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 3\mathbf{x}$$

$\boldsymbol{x}$ is an eigenvector associated with eigenvalue 3

$$\mathbf{B}(2\mathbf{x}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix} = 3\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 3(2\mathbf{x})$$

$2\boldsymbol{x}$ is an eigenvector associated with eigenvalue 3

$$\mathbf{B}(k\mathbf{x}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} k \\ k \\ k \end{pmatrix} = 3\begin{pmatrix} k \\ k \\ k \end{pmatrix} = 3(k\mathbf{x})$$

$k\boldsymbol{x}$ is an eigenvector associated with eigenvalue 3

$$\mathbf{B}=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}=\begin{pmatrix}1&1&1\\1&0&-2\\1&-1&1\end{pmatrix}\begin{pmatrix}3&0&0\\0&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&1\\1&0&-2\\1&-1&1\end{pmatrix}^{-1}$$

Example 6.1.4.2

$$\mathbf{B}=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}\qquad \mathbf{x}=\begin{pmatrix}1\\1\\1\end{pmatrix}\qquad \mathbf{y}=\begin{pmatrix}1\\0\\-1\end{pmatrix}\qquad \mathbf{z}=\begin{pmatrix}1\\-2\\1\end{pmatrix}$$

$$\mathbf{Bx}=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}3\\3\\3\end{pmatrix}=3\begin{pmatrix}1\\1\\1\end{pmatrix}=3\mathbf{x}$$

$\boldsymbol{x}$ is an eigenvector associated with eigenvalue 3

$$\mathbf{By}=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}\begin{pmatrix}1\\0\\-1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}=0\begin{pmatrix}1\\0\\-1\end{pmatrix}=0\mathbf{y}$$

$\boldsymbol{y}$ is an eigenvector associated with eigenvalue 0

$$\mathbf{Bz}=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}\begin{pmatrix}1\\-2\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}=0\begin{pmatrix}1\\-2\\1\end{pmatrix}=0\mathbf{z}$$

$\boldsymbol{z}$ is an eigenvector associated with eigenvalue 0

Eigenvalues of triangular matrices

Theorem 6.1.9 & Example 6.1.10

If ***A*** is a triangular matrix, in particular, diagonal matrix

the eigenvalues of ***A*** are the diagonal entries of ***A***.

$$\begin{pmatrix} -1 & 3.5 & 14 \\ 0 & 5 & -26 \\ 0 & 0 & 2 \end{pmatrix}$$ The eigenvalues are -1, 5 and 2.

$$\begin{pmatrix} -2 & 0 & 0 \\ 99 & 0 & 0 \\ 10 & -4.5 & 10 \end{pmatrix}$$ The eigenvalues are -2, 0 and 10.

The proof

Theorem 6.1.9

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ 0 & & & a_{nn} \end{pmatrix} \quad \lambda\boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ & \lambda - a_{22} & \cdots & -a_{2n} \\ & & \ddots & \vdots \\ 0 & & & \lambda - a_{nn} \end{pmatrix}$$

characteristic polynomial of $\boldsymbol{A}$

$$\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = (\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$$

this polynomial is completely factorized

roots of the polynomial: $a_{11}, a_{22}, \ldots, a_{nn}$

eigenvalues of $\boldsymbol{A}$ $\Longleftrightarrow$ diagonal entries of $\boldsymbol{A}$

How to find eigenvalues?

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

Remark 6.1.5

Let $\boldsymbol{A}$ be a square matrix of order n.

λ is an eigenvalue of $\boldsymbol{A}$

$\Leftrightarrow$ $\boldsymbol{Ax} = \lambda\,\boldsymbol{x}$ for some nonzero column vector $\boldsymbol{x}$

$\Leftrightarrow$ $\lambda\,\boldsymbol{x} - \boldsymbol{Ax} = \mathbf{0}$ for some nonzero column vector $\boldsymbol{x}$

$\Leftrightarrow$ $\underbrace{(\lambda \boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{x} = \mathbf{0}}_{\text{homog. system}}$ for some nonzero column vector $\boldsymbol{x}$

homog. system has non-trivial solutions

$\Leftrightarrow$ $\underbrace{\det(\lambda\,\boldsymbol{I} - \boldsymbol{A})}_{\text{a polynomial in } \lambda} = 0$

Solve this equation to find the eigenvalues of $\boldsymbol{A}$

a polynomial in λ

How to find eigenvalues?

Example 6.1.7.1

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$$

The eigenvalues of $\boldsymbol{A}$ are 1 and 0.95.

$$\begin{aligned}\det(\lambda \boldsymbol{I} - \boldsymbol{A}) &= \det\left(\lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}\right) \\ &= \begin{vmatrix} \lambda - 0.96 & -0.01 \\ -0.04 & \lambda - 0.99 \end{vmatrix} \\ &= (\lambda - 0.96)(\lambda - 0.99) - (-0.01)(-0.04) \\ &= \lambda^2 - 1.95\lambda + 0.95 \quad \text{polynomial of degree 2} \\ &= (\lambda - 1)(\lambda - 0.95) \quad \text{factorize the polynomial}\end{aligned}$$

$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ if and only if $\lambda = 1$ or 0.95

What is characteristic polynomial?

Definition 6.1.6

Let $\boldsymbol{A}$ be a square matrix of order n.

The polynomial $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ degree n
is called the characteristic polynomial of $\boldsymbol{A}$.

λ is an eigenvalue of $\boldsymbol{A}$ $\Leftrightarrow$ $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$

$\Leftrightarrow$ λ is a root of the characteristic polynomial

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & -a_{n1} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

Finding eigenvalues from characteristic polynomial

Example 6.1.7.3

$$\boldsymbol{C} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

The eigenvalues of $\boldsymbol{C}$ are 1, $\sqrt{2}$ and $-\sqrt{2}$

$$\det(\lambda \boldsymbol{I} - \boldsymbol{C}) = \begin{vmatrix} \lambda & 1 & 0 \\ 0 & \lambda & -2 \\ -1 & -1 & \lambda - 1 \end{vmatrix}$$

characteristic polynomial of $\boldsymbol{C}$

$$= \lambda^3 - \lambda^2 - 2\lambda + 2$$

one factor is $(\lambda - 1)$

$$= (\lambda - 1)(\lambda^2 - 2)$$

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

guess one root

$\lambda = 1$

$\det(\lambda \boldsymbol{I} - \boldsymbol{C}) = 0$ if and only if $\lambda = 1$, $\sqrt{2}$ or $-\sqrt{2}$

A very^3 important theorem (revisited)

Theorem 6.1.8

1,2,3,4,5,6,7,8

9

$\boldsymbol{A}$ is an n x n matrix

The following statements are equivalent:

1. $\boldsymbol{A}$ is invertible.
2. The linear system $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution.
3. The reduced row-echelon form of $\boldsymbol{A}$ is $\boldsymbol{I}$.
4. $\boldsymbol{A}$ can be expressed as a product of elementary matrices.
5. $\det(\boldsymbol{A}) \neq 0$.
6. The rows of $\boldsymbol{A}$ form a basis for $\mathbf{R}^n$.
7. The columns of $\boldsymbol{A}$ form a basis for $\mathbf{R}^n$.
8. $\text{rank}(\boldsymbol{A}) = n$.

9. 0 is not an eigenvalue of $\boldsymbol{A}$.

5. $\det(\boldsymbol{A}) \neq 0$

9. 0 is not an eigenvalue of $\boldsymbol{A}$

The proof

Theorem 6.1.8

We are going to show "5 $\Leftrightarrow$ 9".

Statement 9 0 is not an eigenvalue of $\boldsymbol{A}$

$\Leftrightarrow$ 0 is not a root of the char. poly. $\det(\lambda\boldsymbol{I} - \boldsymbol{A})$

$\Leftrightarrow$ $\det(0\boldsymbol{I} - \boldsymbol{A}) \neq 0$

$\Leftrightarrow$ $\det(-\boldsymbol{A}) \neq 0$

$\Leftrightarrow$ $(-1)^n \det(\boldsymbol{A}) \neq 0$

$\Leftrightarrow$ $\det(\boldsymbol{A}) \neq 0$ Statement 5

How to find eigenvectors?

Remark 6.1.5

eigenvector

Let $\boldsymbol{A}$ be a square matrix of order n.

λ is an eigenvalue of $\boldsymbol{A}$

$\Leftrightarrow$ $\boldsymbol{Ax} = \lambda \boldsymbol{x}$ for some nonzero column vector $\boldsymbol{x}$

$\Leftrightarrow$ $\lambda \boldsymbol{x} - \boldsymbol{Ax} = \boldsymbol{0}$ for some nonzero column vector $\boldsymbol{x}$

$\Leftrightarrow$ $\underbrace{(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}}_{\text{homog. system}}$ for some nonzero column vector $\boldsymbol{x}$

homog. system has non-trivial solutions

$\Leftrightarrow$ $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$

by solving this linear system

its solution space contains all the eigenvectors associated to λ

What is an eigenspace of a matrix?

Definition 6.1.11 (Eigenspace)

$\boldsymbol{A}$: square matrix of order n

λ : an eigenvalue of $\boldsymbol{A}$

The solution space of the linear system

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

has nontrivial solutions

is called the eigenspace of $\boldsymbol{A}$

associated with the eigenvalue λ

denoted by E_λ

If $\boldsymbol{u}$ is a nonzero vector in E_λ,

then $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{A}$ associated with the eigenvalue λ.

Eigenspace of a matrix

Example 6.1.12.1

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$$

By Example 6.1.8.1,

the eigenvalues of $\boldsymbol{A}$ are 1 and 0.95.

$\boldsymbol{A}$ has two eigenspaces E_1 and $E_{0.95}$

How to find eigenspace?

Example 6.1.12.1 (Find E_1)

For $\lambda = 1$,

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} 1-0.96 & -0.01 \\ -0.04 & 1-0.99 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

General solution

$$\begin{pmatrix} x \\ y \end{pmatrix} = t\begin{pmatrix} 0.25 \\ 1 \end{pmatrix}$$

t an arbitrary parameter

$$E_1 = \text{span}\boxed{\left\{\begin{pmatrix} 0.25 \\ 1 \end{pmatrix}\right\}}$$

any non-zero scalar multiple of $\begin{pmatrix} 0.25 \\ 1 \end{pmatrix}$ is an eigenvector of $\boldsymbol{A}$ associated with the eigenvalue 1

Basis for the eigenspace E_1

How to find eigenspace?

Example 6.1.12.1 (Find $E_{0.95}$)

For $\lambda = 0.95$,

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} 0.95 - 0.96 & -0.01 \\ -0.04 & 0.95 - 0.99 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

General solution

$$\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

t an arbitrary parameter

$$E_{0.95} = \text{span}\boxed{\left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}}$$

any non-zero scalar multiple of $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ is an eigenvector of $\boldsymbol{A}$ associated with the eigenvalue 0.95

Basis for the eigenspace $E_{0.95}$

Example 6.1.12.2

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

By Example 6.1.8.2,
the eigenvalues of $\boldsymbol{B}$ are 3 and 0.

$\boldsymbol{B}$ has two eigenspaces E_3 and E_0

How to find eigenspace?

Example 6.1.12.2 (Find E_0)

For $\lambda = 0$,

$$(\lambda \boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} 0-1 & -1 & -1 \\ -1 & 0-1 & -1 \\ -1 & -1 & 0-1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

General solution

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

s, t are arbitrary parameters

$$E_0 = \operatorname{span}\boxed{\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}}$$

Basis for the eigenspace E_0

any non-zero linear combination of $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ is an eigenvector of $\boldsymbol{B}$ associated with the eigenvalue 0

Section 6.2

Diagonalization

Objective

- What is a diagonalizable matrix?
- How to determine if a matrix is diagonalizable?
- How to diagonalize a matrix?
- How to compute powers of matrix using diagonalization?
- How to solve linear recurrence relation using diagonalization?

A 2x2 diagonalizable matrix

Example 6.2.2.1

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} \qquad \boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}$$

diagonalizable

diagonalizes $\boldsymbol{A}$

$$\boxed{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1}}$$

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1}}_{\boldsymbol{P}^{-1}}\underbrace{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}}_{\boldsymbol{A}}\underbrace{\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}}_{\boldsymbol{P}} = \begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}$$

$\Rightarrow$ diagonal

bring over to diagonalize the matrix

What is a diagonalizable matrix?

Definition 6.2.1

A square matrix $\boldsymbol{A}$ is called diagonalizable if there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a diagonal matrix.

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \boxed{\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}}$$

diagonal

diagonalizable

We say: the matrix $\boldsymbol{P}$ diagonalizes $\boldsymbol{A}$

A 3x3 diagonalizable matrix

Example 6.2.2.2

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

diagonalizable

$$\boldsymbol{P} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix}$$

diagonalizes $\boldsymbol{B}$

$$\underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix}^{-1}}_{\boldsymbol{P}^{-1}} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}}_{\boldsymbol{B}} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix}}_{\boldsymbol{P}} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

A non-diagonalizable matrix

Example 6.2.2.3

We will introduce a systematic way to determine whether a matrix is diagonalizable

$$\boldsymbol{M} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$$ not diagonalizable

Cannot find a matrix $\boldsymbol{P}$ that diagonalizes $\boldsymbol{M}$.

Prove by contradiction

Suppose there exist an invertible $\boldsymbol{P}$ such that

$\boldsymbol{P}^{-1}\boldsymbol{MP}$ = Diagonal matrix.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

Derive that: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ * & * \end{pmatrix}$ contradicts that $\boldsymbol{P}$ is invertible

How to tell whether a matrix is diagonalizable?

Example 6.2.2

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$$

diagonalizable

two eigenvalues : 1 and 0.95

two eigenvectors : $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$ $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

linearly independent

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

diagonalizable

two eigenvalues : 3 and 0

three eigenvectors : $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$

linearly independent

$$\boldsymbol{M} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$$

not diagonalizable

one eigenvalue : 2

only one eigenvector : $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

linearly independent

How to tell whether a matrix is diagonalizable?

Theorem 6.2.3

Let $\boldsymbol{A}$ be a square matrix of order n.

$\boldsymbol{A}$ is diagonalizable

if and only if

$\boldsymbol{A}$ has $\underbrace{n \text{ linearly independent eigenvectors}}_{\text{may be associated to the same eigenvalues}}$

Two observations

$$\boldsymbol{AB} = \boldsymbol{A}(\boldsymbol{b}_1 \ \ \boldsymbol{b}_2 \ \ \cdots \ \ \boldsymbol{b}_n) = (\boldsymbol{Ab}_1 \ \ \boldsymbol{Ab}_2 \ \ \cdots \ \ \boldsymbol{Ab}_n)$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{array}{c} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{pmatrix} \\ \boldsymbol{b}_1 \ \ \boldsymbol{b}_2 \ \ \boldsymbol{b}_3 \end{array} = \begin{array}{c} \begin{pmatrix} 2 & 3 & 4 \\ 2 & 5 & 8 \\ 2 & 4 & 6 \end{pmatrix} \\ \boldsymbol{Ab}_1 \ \boldsymbol{Ab}_2 \ \boldsymbol{Ab}_3 \end{array}$$

$$\boldsymbol{BD} = (\boldsymbol{b}_1 \ \ \boldsymbol{b}_2 \ \ \cdots \ \ \boldsymbol{b}_n)\, \boldsymbol{D} = (d_1\boldsymbol{b}_1 \ \ d_2\boldsymbol{b}_2 \ \ \cdots \ \ d_n\boldsymbol{b}_n)$$

diagonal matrix with
diagonal entries d_i

$$\begin{array}{c} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{pmatrix} \\ \boldsymbol{b}_1 \ \ \boldsymbol{b}_2 \ \ \boldsymbol{b}_3 \end{array} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \begin{array}{c} \begin{pmatrix} 2 & 3 & 4 \\ 2 & 6 & 12 \\ 2 & 9 & 20 \end{pmatrix} \\ 2\boldsymbol{b}_1 \ 3\boldsymbol{b}_2 \ 4\boldsymbol{b}_3 \end{array}$$

The proof

$\boldsymbol{A}$ diagonalizable

$\boldsymbol{A}$ has n linearly independent eigenvectors

Theorem 6.2.3 $(\Leftarrow)$

Suppose $\boldsymbol{A}$ has n linearly independent eigenvectors.

$\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}$

associating eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$

Define the invertible matrix $\boldsymbol{P} = (\boldsymbol{u_1} \;\; \boldsymbol{u_2} \;\; \cdots \;\; \boldsymbol{u_n})$

$$\begin{aligned}
\boldsymbol{AP} &= (\boldsymbol{Au_1} \;\; \boldsymbol{Au_2} \;\; \cdots \;\; \boldsymbol{Au_n}) \\
&= (\lambda_1\boldsymbol{u_1} \;\; \lambda_2\boldsymbol{u_2} \;\; \cdots \;\; \lambda_n\boldsymbol{u_n}) \\
&= (\boldsymbol{u_1} \;\; \boldsymbol{u_2} \;\; \cdots \;\; \boldsymbol{u_n})\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = \boldsymbol{P}\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}
\end{aligned}$$

So $\boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$ $\quad$ $\boldsymbol{A}$ is diagonalizable.

The proof

$\boldsymbol{A}$ diagonalizable
$\boldsymbol{A}$ has n linearly independent eigenvectors

Theorem 6.2.3 $(\Rightarrow)$

$\boldsymbol{A}$ is diagonalizable $\Rightarrow \boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{D}$

Let $\boldsymbol{P} = (\boldsymbol{u_1} \;\; \boldsymbol{u_2} \;\; \cdots \;\; \boldsymbol{u_n})$

$$\boxed{\boldsymbol{AP}} = \boxed{\boldsymbol{PD}}$$

$$\boldsymbol{A}(\boldsymbol{u_1} \;\; \boldsymbol{u_2} \;\; \cdots \;\; \boldsymbol{u_n}) \qquad (\boldsymbol{u_1} \;\; \boldsymbol{u_2} \;\; \cdots \;\; \boldsymbol{u_n})\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

$$(\boldsymbol{Au_1} \;\; \boldsymbol{Au_2} \;\; \cdots \;\; \boldsymbol{Au_n}) \qquad (\lambda_1\boldsymbol{u_1} \;\; \lambda_2\boldsymbol{u_2} \;\; \cdots \;\; \lambda_n\boldsymbol{u_n})$$

Compare each column on LHS and RHS

linearly independent

So $\boldsymbol{Au_i} = \lambda_i\boldsymbol{u_i}$ for all i $\Rightarrow \boldsymbol{u_i}$ are eigenvectors of $\boldsymbol{A}$ with eigenvalues λ_i

Algorithm 6.2.4 (Diagonalization)

Step 1: Solve the characteristic equation

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

to find all distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$.

Step 2: For each λ_i, find a basis S_{λ_i} for the eigenspace E_{λ_i}.

Step 3: Let $S = S_{\lambda_1} \cup S_{\lambda_2} \cup \cdots \cup S_{\lambda_k}$.

(a) If $|S| < n$, then $\boldsymbol{A}$ is not diagonalizable.

(b) If $|S| = n$, then $\boldsymbol{A}$ is diagonalizable.

Say, $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}\}$, then the square matrix $\boldsymbol{P} = (\boldsymbol{u_1} \ \ \boldsymbol{u_2} \ \ \cdots \ \ \boldsymbol{u_n})$ diagonalizes $\boldsymbol{A}$.

How to diagonalize a matrix?

Example 6.2.6.1

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Step 1: By solving characteristic polynomial, the eigenvalues are 3 and 0.

Step 2: For $\lambda = 3$, solve $(3\boldsymbol{I} - \boldsymbol{B})\,\boldsymbol{x} = 0$

For $\lambda = 0$, solve $(0\boldsymbol{I} - \boldsymbol{B})\,\boldsymbol{x} = 0$

$S_3 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ a basis for E_3 $S_0 = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ a basis for E_0

Step 3: $|S| = |S_3| + |S_0| = 1 + 2$ = order of $\boldsymbol{B}$

So $\boldsymbol{B}$ is diagonalizable

How to diagonalize a matrix?

Example 6.2.6.1

Step 3:

$$\boldsymbol{P} = \begin{array}{c} \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \\ \begin{matrix} \boldsymbol{u_1} & \boldsymbol{u_2} & \boldsymbol{u_3} \end{matrix} \end{array} \quad \text{Then} \quad \underbrace{\boldsymbol{P}^{-1}\boldsymbol{BP}}_{} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

you do not need to multiply this out!!!

P is not unique

$$\boldsymbol{P} = \begin{array}{c} \begin{pmatrix} 2 & -7 & 1 \\ 2 & 7 & 0 \\ 2 & 0 & -1 \end{pmatrix} \\ \begin{matrix} 2\boldsymbol{u_1} & 7\boldsymbol{u_2} & -\boldsymbol{u_3} \end{matrix} \end{array}$$

$$\mathbf{Q} = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{Then} \quad \mathbf{Q}^{-1}\mathbf{BQ} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

How to show a matrix is not diagonalizable?

Example 6.2.6.3

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{pmatrix}$$

Step 1: The eigenvalues are 1 and 2.

Step 2: For $\lambda = 1$, solve $(\boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{x} = 0$

For $\lambda = 2$, solve $(2\boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{x} = 0$

$S_1 = \left\{ \begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix} \right\}$ a basis for E_1 $\qquad$ $S_2 = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ a basis for E_2

Step 3: $|S| = |S_1| + |S_2| = 1 + 1 <$ order of $\boldsymbol{A}$

Only have two linearly independent eigenvectors, so $\boldsymbol{A}$ is not diagonalizable.

Matrix with no eigenvalue

Remark 6.2.5.1

Not in scope!

The characteristic polynomial $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ may have complex roots.

$$\boldsymbol{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^2 + 1$$

roots: $\lambda = \pm\, i$

i.e. the matrix has eigenvalues that are not real numbers but complex numbers.

We can still use the algorithm to diagonalize the matrix.

However, to discuss the theory, we need the concept of vector space over complex numbers.

Upper bound of dimension of eigenspace

Remark 6.2.5.2

$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - 1)^1(\lambda - 2)^3(\lambda - 4)^2$

Characteristic polynomial

multiplicity

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}$$

Then $\dim(E_{\lambda_i}) \le r_i$

$\dim(E_2) \le 3$

$\dim(E_4) \le 2$

$\dim(E_1) = 1$

The number of basis vectors in each eigenspace cannot be more than the multiplicity of the eigenvalue in the characteristic polynomial.

A is diagonalizable

if and only if

$\dim(E_{\lambda_i}) = r_i$ for all λ_i

Union of bases of eigenspaces

A = {(1,1,1), (1,2,3)}

B = {(2,2,2), (1,2,3)}

A ∪ B is linearly dependent

Remark 6.2.5.3

The set S is always linearly independent. Ex 6 Q22

$$S = \overbrace{\underbrace{S_{\lambda_1}}_{\text{linearly independent}} \cup \underbrace{S_{\lambda_2}}_{\text{linearly independent}} \cup \cdots \cup \underbrace{S_{\lambda_k}}_{\text{linearly independent}}}$$

bases: $S_{\lambda_1} \uparrow E_{\lambda_1}$, $S_{\lambda_2} \uparrow E_{\lambda_2}$, $S_{\lambda_k} \uparrow E_{\lambda_k}$

In particular

If $\mathbf{u}_1 \in E_{\lambda_1}, \mathbf{u}_2 \in E_{\lambda_2}, \cdots, \mathbf{u}_k \in E_{\lambda_k}$

then $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is linearly independent

Matrix with maximum number of eigenvalues

Theorem 6.2.7

Let $\boldsymbol{A}$ be a square matrix of order n.

If $\boldsymbol{A}$ has n distinct eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_n$

then $\boldsymbol{A}$ is diagonalizable.

$\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}$

linearly independent

Proof

We can find one eigenvector for each eigenvalue.

Hence we have n eigenvectors.

By Remark 6.2.5.3, these eigenvectors are linearly independent.

By Theorem 6.2.3, $\boldsymbol{A}$ is diagonalizable.

Matrix with maximum number of eigenvalues

Example 6.2.8

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

$\boldsymbol{A}$ has 4 distinct eigenvalues 1, 2, 3, 4.

So $\boldsymbol{A}$ is diagonalizable.

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

diagonal matrices are diagonalizable

$\boldsymbol{B}$ has only 2 distinct eigenvalues 1, 2.

And $\boldsymbol{B}$ is also diagonalizable.

Matrix with maximum number of eigenvalues

Remark 6.2.9

The converse of Theorem 6.2.7 is not true.

If $\boldsymbol{A}$ is an n x n diagonalizable matrix,

$\boldsymbol{A}$ need not have n distinct eigenvalues.

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

$\boldsymbol{B}$ has only 2 distinct eigenvalues 1, 2.

And $\boldsymbol{B}$ is also diagonalizable.

How to find powers of a matrix?

Discussion 6.2.10

Let $\boldsymbol{A}$ be a diagonalizable matrix of order n

$\boldsymbol{P}$ an invertible matrix such that

$$(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^m = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}^m = \begin{pmatrix} \lambda_1^m & & & 0 \\ & \lambda_2^m & & \\ & & \ddots & \\ 0 & & & \lambda_n^m \end{pmatrix}$$

$$\boldsymbol{P}^{-1}\boldsymbol{A}^m\boldsymbol{P}$$

Then $\boldsymbol{A}^m = \boldsymbol{P} \begin{pmatrix} \lambda_1^m & & & 0 \\ & \lambda_2^m & & \\ & & \ddots & \\ 0 & & & \lambda_n^m \end{pmatrix} \boldsymbol{P}^{-1}$

How to find powers of a matrix?

Example 6.2.11.1

invertible

$$\boldsymbol{A} = \begin{pmatrix} -4 & 0 & -6 \\ 2 & 1 & 2 \\ 3 & 0 & 5 \end{pmatrix}$$

Use Algorithm 6.2.4 to find the eigenvalues and eigenvectors

We have

$$\boldsymbol{P} = \begin{pmatrix} -2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad \boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

obtain this diagonal matrix from eigenvalues, not matrix multiplication!

$$\mathbf{A}^{m} = \mathbf{P}\begin{pmatrix} (-1)^{m} & 0 & 0 \\ 0 & 1^{m} & 0 \\ 0 & 0 & 2^{m} \end{pmatrix}\mathbf{P}^{-1}$$

$$\mathbf{A}^{-1} = \mathbf{P}\begin{pmatrix} (-1)^{-1} & 0 & 0 \\ 0 & 1^{-1} & 0 \\ 0 & 0 & 2^{-1} \end{pmatrix}\mathbf{P}^{-1}$$

Some applications

- Weather forecast (Markov chain)
- Population growth
- Cards shuffling
- Genetics
- Linear recurrence relation

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \to \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \to \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \cdots \to \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}$$

stage 0 (initial) | stage 1 | stage 2 | stage n

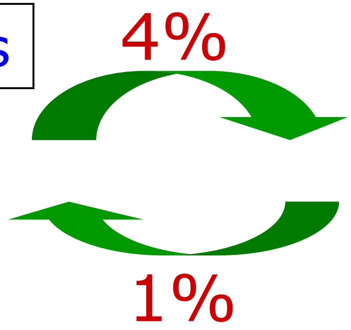
$\mathbf{x_0}$ $\quad \mathbf{x_1} = \mathbf{A}\mathbf{x_0}$ $\quad \mathbf{x_2} = \mathbf{A}\mathbf{x_1}$ $\quad \mathbf{x_n} = \mathbf{A}\mathbf{x_{n-1}}$

Application to modeling

Example 6.1.1 (Population)

Population after n years

$\boxed{a_n}$ rural population



Urban population $\boxed{b_n}$

Long term effect ?

Ans: ~ 20% rural population, ~ 80% urban population

$$a_n = 0.96a_{n-1} + 0.01b_{n-1}$$
$$b_n = 0.04a_{n-1} + 0.99b_{n-1}$$

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix}$$

Application to modeling

Example 6.1.1

$$\underbrace{\begin{pmatrix} a_n \\ b_n \end{pmatrix}}_{\boldsymbol{x_n}} = \underbrace{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}}_{\boldsymbol{A}} \underbrace{\begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix}}_{\boldsymbol{x_{n-1}}} \Rightarrow \begin{pmatrix} a_n \\ b_n \end{pmatrix} = \underbrace{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}^n}_{\mathbf{A}^n} \underbrace{\begin{pmatrix} a_0 \\ b_0 \end{pmatrix}}_{\mathbf{x}_0}$$

$$\boldsymbol{x_n} = \boldsymbol{A}\boldsymbol{x_{n-1}} = \boldsymbol{A}^2\boldsymbol{x_{n-2}} = \boldsymbol{A}^3\boldsymbol{x_{n-3}} = \cdots = \boldsymbol{A}^n\boldsymbol{x_0}$$

current population

long term effect → a_n and b_n for large n

→ $\boldsymbol{x}_n$ for large n

→ $\boldsymbol{A}^n$ for large n

$$\mathbf{A}^{(\text{big } n)} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \underbrace{0.95^{(\text{big } n)}}_{\to 0} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1} \approx \begin{pmatrix} 0.2 & 0.2 \\ 0.8 & 0.8 \end{pmatrix}$$

$$\begin{pmatrix} a_{(\text{big } n)} \\ b_{(\text{big } n)} \end{pmatrix} \approx \begin{pmatrix} 0.2 & 0.2 \\ 0.8 & 0.8 \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} 0.2(a_0 + b_0) \\ 0.8(a_0 + b_0) \end{pmatrix}$$

Fibonacci Numbers

0 1 1 2 3 5 8 13 21 34

What's the 100th number?

How to solve recurrence relation?

Example 6.2.11.2

Denote the Fibonacci numbers by $a_0, a_1, a_2, \ldots$

$a_0 = 0$ $\quad$ $a_1 = 1$ $\quad$ $a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$

initial conditions $\quad$ recurrence relation

What is the value of a_n ?

$$\boxed{a_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n}$$

due to eigenvalues

Example: $a_{100} = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{100} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{100}$

354224848179261915075

How to find recurrence matrix?

$a_0 = 0,\ a_1 = 1,$
$a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$

Example 6.2.11.2

Form the vector: $\boldsymbol{x_n} = \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix}$ $\quad \boldsymbol{x_{n-1}} = \begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}$

$$\mathbf{x}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \mathbf{x}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \mathbf{x}_3 = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \ldots$$

The recurrence matrix $\boldsymbol{A}$: $\quad \boldsymbol{x_n} = \boldsymbol{A}\boldsymbol{x_{n-1}}$ for all n

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}$$

Compare coefficients $\quad a_n = 0\, a_{n-1} + 1\, a_n$

Recurrence relation $\quad a_{n+1} = 1\, a_{n-1} + 1\, a_n$

$a_{n+1} = a_n + a_{n-1}$

$$\begin{aligned} a_n &= \boxed{}\, a_{n-1} + \boxed{}\, a_n \\ a_{n+1} &= \boxed{}\, a_{n-1} + \boxed{}\, a_n \end{aligned}$$

Example (Additional)

$$a_0 = 1,\ a_1 = 3,$$
$$a_n = 3a_{n-1} + 5a_{n-2} \quad \text{for } n \geq 2$$

What is the recurrence matrix ?

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 5 & 3 \end{pmatrix}$$

In general,

$$a_0 = \mathrm{s},\ a_1 = \mathrm{t},$$
$$a_n = \mathrm{p}a_{n-1} + \mathrm{q}a_{n-2} \quad \text{for } n \geq 2$$

The recurrence matrix $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ \mathrm{q} & \mathrm{p} \end{pmatrix}$

How to find the explicit formula?

Example 6.2.11.2

$a_0 = 0,\ a_1 = 1,$
$a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$

$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ has two eigenvalues $\lambda_1 = \frac{1+\sqrt{5}}{2}$ $\lambda_2 = \frac{1-\sqrt{5}}{2}$

So $\boldsymbol{A}$ is diagonalizable

Diagonalized by $\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}$

$$\boldsymbol{x_n} = \boldsymbol{A}^n \boldsymbol{x_0}$$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \mathbf{P} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} \mathbf{P}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n \\ \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+1} \end{pmatrix}$$

$$a_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

$a_0 = 0,\ a_1 = 1,\ a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$

Solving linear recurrence relation

$a_0 = u \qquad a_1 = v \qquad a_n = pa_{n-1} + qa_{n-2}$ for $n \geq 2$

Form the recurrence matrix $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

Find the eigenvalues of $\boldsymbol{A}$

$$\lambda_1 = \tfrac{1+\sqrt{5}}{2} \qquad \lambda_2 = \tfrac{1-\sqrt{5}}{2}$$

If $\boldsymbol{A}$ is diagonalizable, find the matrix $\boldsymbol{P}$ that diagonalizes $\boldsymbol{A}$

$$\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}$$

Set up $\boldsymbol{x_n} = \boldsymbol{A}^n \boldsymbol{x}_0$

and diagonalize $\boldsymbol{A}^n$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \boldsymbol{P} \begin{pmatrix} \lambda_1{}^n & 0 \\ 0 & \lambda_2{}^n \end{pmatrix} \boldsymbol{P}^{-1} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$$

Multiply out the RHS and equate the first component

$$a_n = s\left(\lambda_1\right)^n + t\left(\lambda_2\right)^n$$

Section 6.3

Orthogonal Diagonalization

Objective

- What is orthogonal diagonalization?
- When is a matrix orthogonally diagonalizable?
- How to orthogonally diagonalize a symmetric matrix?

What is an orthogonally diagonalizable matrix

Definition 6.3.2

Recall: Section 6.2

A square matrix **A** is called
diagonalizable
if there exists an invertible matrix **P** such that
$\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix.
We say the matrix **P** diagonalizes **A**.

A square matrix **A** is called
orthogonally diagonalizable
if there exists an orthogonal matrix **P** such that
$\mathbf{P}^{T}\mathbf{AP}$ is a diagonal matrix.
We say the matrix **P** orthogonally diagonalizes **A**.

Theorem 6.3.4

$$\begin{pmatrix} 0 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & -2 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

A square matrix is orthogonally diagonalizable
if and only if
it is symmetric.

beyond the scope of this course

A is orthogonally diagonalizable

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D}$$

$$\Rightarrow \mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T$$

$$\Rightarrow \mathbf{A}^T = (\mathbf{P}\mathbf{D}\mathbf{P}^T)^T$$

$$\Rightarrow \mathbf{A}^T = (\mathbf{P}^T)^T\mathbf{D}^T(\mathbf{P}^T)$$

$$\Rightarrow \mathbf{A}^T = \mathbf{P}\mathbf{D}\mathbf{P}^T = \mathbf{A}$$

So **A** is symmetric

How to orthogonally diagonalize a symmetric matrix

Algorithm 6.3.5 **A**: symmetric matrix

Step 1: Find all distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$.

Step 2: For each eigenvalue λ_i,

Step 2a: find a basis S_{λ_i} for the eigenspace E_{λ_i}

Step 2b: use the Gram-Schmidt Process (Theorem 5.2.19) to transform S_{λ_i} to an orthonormal basis T_{λ_i} .

Step 3: Let $T = T_{\lambda_1} \cup T_{\lambda_2} \cup \ldots \cup T_{\lambda_k}$

say $T = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$.

The square matrix $\mathbf{P} = (\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_n)$ is an orthogonal matrix that diagonalizes **A**.

Eigenvalues of symmetric matrix

Remark 6.3.6.1 **A**: symmetric matrix

In Step 1, the eigenvalues of a symmetric matrix are always real numbers.

Idea:

Let λ be an eigenvalue of a symmetric matrix

Write $\lambda = a + ib$ (a, b are real)

Conjugate $\bar{\lambda} = a - ib$ also an eigenvalue of the matrix

Try to show $\lambda = \bar{\lambda}$, which implies λ is real.

Dimension of Eigenspaces

In general
$\dim(E_{\lambda_i}) \le r_i$

Remark 6.3.6.2 **A**: symmetric matrix

Suppose the characteristic polynomial of **A** is

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \ldots(\lambda - \lambda_k)^{r_k}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of **A**.

Then for each eigenvalue λ_i,

$$\dim(E_{\lambda_i}) = r_i.$$

number of basis vectors in the eigenspace for λ_i

multiplicity of λ_i in the characteristic polynomial

$r_1 + r_2 + \cdots + r_k =$ degree of polynomial = order of **A**

$\dim E_{\lambda_1} + \dim E_{\lambda_2} + \cdots + \dim E_{\lambda_k} =$ no.lin.indep.eigenvectors

A symmetric matrix is always diagonalizable.

T is an orthonormal set

Step 2b: use the Gram-Schmidt Process to transform S_{λ_i} to an orthonormal basis T_{λ_i}

Remark 6.3.6

$$\begin{array}{ccccccc} & S_{\lambda_1} & & S_{\lambda_2} & & \ldots & S_{\lambda_k} \\ T = & T_{\lambda_1} & \cup & T_{\lambda_2} & \cup \ldots \cup & & T_{\lambda_k} \end{array}$$

3. In Step 3, the set T is always orthonormal.

   Not immediate

4. Since T is always orthonormal, the square matrix **P** in Step 3 is always orthogonal.

   Immediate from Theorem 5.4.6

Ex6 Q26 Proof later

Let **A** be a symmetric matrix.

If **u** and **v** are two eigenvectors of **A** associated with eigenvalues λ and μ, resp. where $\lambda \neq \mu$,

Then $\mathbf{u} \cdot \mathbf{v} = 0$.

OD a 2x2 symmetric matrix

Example 6.3.7.1

$$\mathbf{A} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$$

Step 1: The eigenvalues are 1/2 and 3/2.

Step 2a: Bases for $E_{1/2}$ and $E_{3/2}$: $\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$

Step 2b: Orthonormal bases: $\left\{\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\right\}$

No need Gram-Schmidt

Step 3: $\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ and $\mathbf{P}^T\mathbf{AP} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}$

OD a 3x3 symmetric matrix

Discussion 6.3.1

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Step 1: The eigenvalues are 3 and 0.

Step 2a: Bases for E_3 and E_0: $\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}$

Step 2b: Orthonormal bases: $\left\{\begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}\right\}$

just normalize

Gram-Schmidt

Step 3: $\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$ and $\mathbf{P}^T\mathbf{B}\mathbf{P} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Section 7.1

Linear Transformations from $\mathbf{R}^n$ to $\mathbf{R}^m$

Objective

- What is a linear transformation?
- How are linear transformations related to matrices?
- What are the conditions of a linear transformation?
- How to use basis to determine linear transformation?

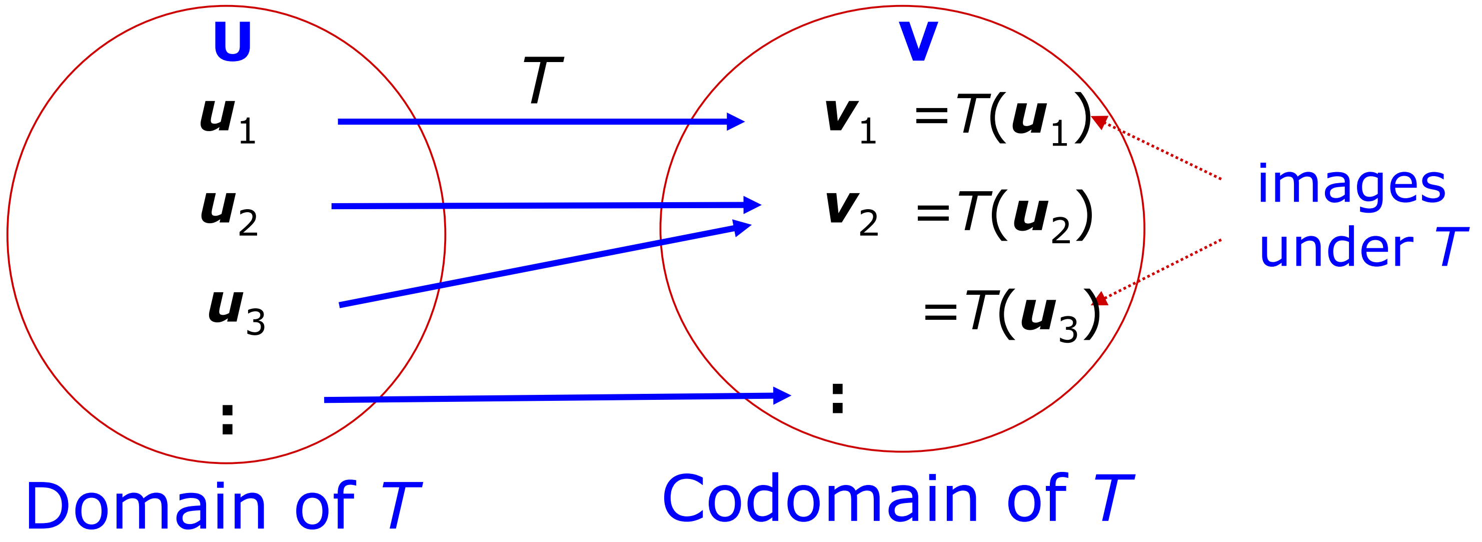
In this chapter, we shall always write vectors in $\mathbf{R}^n$ as column vectors.

Mapping $T : \mathbf{U} \rightarrow \mathbf{V}$

Let **U** and **V** be two sets

A mapping from **U** to **V**

assigns every element of **U** with an element of **V**

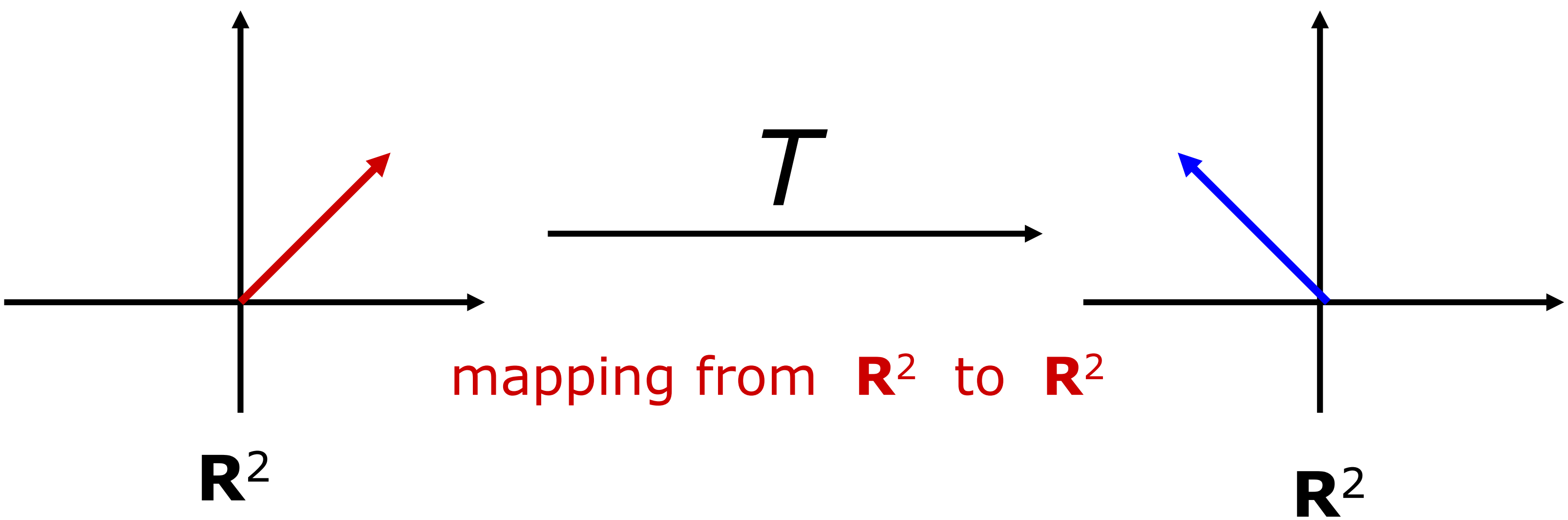


We call a mapping defined this way a linear transformation.

Matrix as a mapping

$$\mathbf{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \longrightarrow \boldsymbol{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \longrightarrow \mathbf{Au} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

input   output



Notation: $T : \mathbf{R}^2 \to \mathbf{R}^2$

defined by $T(\boldsymbol{u}) = \boldsymbol{Au}$ for all $\boldsymbol{u}$ in $\mathbf{R}^2$

$T : \mathbf{R}^2 \to \mathbf{R}^2$

defined by $T(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u}$

Matrix as a mapping

$$\boldsymbol{u} = \begin{pmatrix} x \\ y \end{pmatrix} \longrightarrow \boldsymbol{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \longrightarrow \boldsymbol{A}\boldsymbol{u} = \begin{pmatrix} -y \\ x \end{pmatrix}$$

input output

Formula of $T : \mathbf{R}^2 \to \mathbf{R}^2$ is given by

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} -y \\ x \end{pmatrix} \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbf{R}^2$$

Geometrical meaning

Rotation anticlockwise $90°$

What is a linear transformation?

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

Definition 7.1.1

$$\mathbf{u} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \Rightarrow A = m \times n \text{ matrix} \Rightarrow A\mathbf{u} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

input
Domain $\mathbf{R}^n$

output
Codomain $\mathbf{R}^m$

$T : \mathbf{R}^n \to \mathbf{R}^m$

defined by $T(\boldsymbol{u}) = \boldsymbol{Au}$ for all $\boldsymbol{u} \in \mathbf{R}^n$

T is called a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$

$\boldsymbol{A}$ is called the standard matrix of the linear transformation

Formula of T

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

An example of linear transformation

Example 7.1.2.3

$T : \mathbf{R}^2 \to \mathbf{R}^3$ defined by formula

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+y \\ 2x \\ -3y \end{pmatrix} = x\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + y\begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 0 & -3 \end{pmatrix}}_{\boldsymbol{A}}\begin{pmatrix} x \\ y \end{pmatrix}$$

Is T a linear transformation?

$T(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u}$ for some $\boldsymbol{A}$?

So T is a linear transformation with standard matrix $\boldsymbol{A}$

An example of non-linear transformation

Example 7.1.5.1

$T_1 : \mathbf{R}^2 \to \mathbf{R}^2$ defined by formula

$$T_1\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \boxed{+1} \\ y \boxed{+3} \end{pmatrix}$$

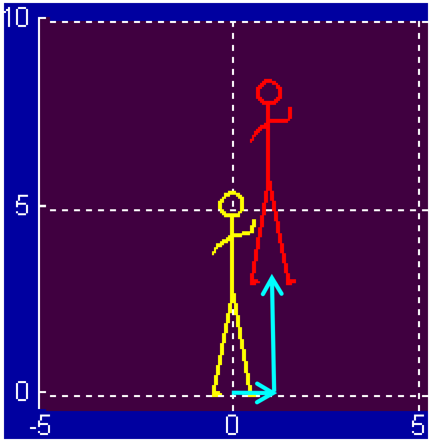
Can't have constant terms in the formula

Why?

There is no 2 x 2 matrix **A** such that $T_1\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix}$

T_1 is not a linear transformation.

T_1 represent a translation in xy-plane



Example 7.1.5.2

$T_2 : \mathbf{R}^3 \rightarrow \mathbf{R}^2$ defined by formula

$$T_2\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x^2 \\ yz \end{pmatrix}$$

Can't have non-linear terms in the formula

This is not a linear transformation. Why?

Identity transformation

Example 7.1.2.1

$I : \mathbf{R}^n \to \mathbf{R}^n$: the identity transformation

$I(\boldsymbol{u}) = \boldsymbol{u}$ for all $\boldsymbol{u}$ in $\mathbf{R}^n$. Do-nothing mapping

Is I a linear transformation?

$I(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u}$ for some $\boldsymbol{A}$?

$$I(\boldsymbol{u}) = \boldsymbol{I}_n\boldsymbol{u} \qquad \mathbf{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$ identity matrix

So I is a linear transformation with standard matrix $\boldsymbol{I}_n$

Formula of I

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \qquad T_1\left(\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}\right) = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

Zero transformation

Example 7.1.2.2

$O : \mathbf{R}^n \to \mathbf{R}^m$: the zero transformation

$$O(\boldsymbol{u}) = \mathbf{0} \quad \text{for all } \boldsymbol{u} \text{ in } \mathbf{R}^n.$$ Kill-everything mapping

Is O a linear transformation?

$O(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u}$ for some $\boldsymbol{A}$?

$$O(\boldsymbol{u}) = \boldsymbol{0}_{\mathrm{m\times n}}\boldsymbol{u} \qquad \mathbf{0}_{m\times n} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \text{ zero matrix}$$

Formula of O

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \qquad O\left(\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

So O is a linear transformation with standard matrix $\boldsymbol{0}_{\mathrm{m\times n}}$

scalar multiplication $2\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}1\\2\\3\end{pmatrix}(2)$ matrix multiplication

Ex 7 Q7 (Tutorial 11)

$P: \mathbf{R}^n \to \mathbf{R}^n$ defined by $P(\mathbf{x}) = \mathbf{x} - (\mathbf{n} \cdot \mathbf{x})\,\mathbf{n}$

n is some fixed vector

Show that P is a linear transformation.

Hint: Show $P(\mathbf{x}) = \mathbf{A}\mathbf{x}$ for some matrix $\mathbf{A}$

$$(\mathbf{n} \cdot \mathbf{x})\,\mathbf{n} = \mathbf{n}\,(\mathbf{n} \cdot \mathbf{x}) = \mathbf{n}\,(\mathbf{n}^T\mathbf{x}) = (\mathbf{n}\mathbf{n}^T)\,\mathbf{x}$$

Theorem 7.1.4

If $T : \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation, then

1. $T(\mathbf{0}) = \mathbf{0}$ $\quad$ $\boldsymbol{A}\mathbf{0} = \mathbf{0}$ $\quad$ T preserves zero vector

2. $T(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k)$ $\quad$ T preserves linear combinations

a linear combination in $\mathbf{R}^n$

$= c_1T(\boldsymbol{u}_1) + c_2T(\boldsymbol{u}_2) + \cdots + c_kT(\boldsymbol{u}_k)$

a linear combination in $\mathbf{R}^m$

$\boldsymbol{A}(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k)$

$= c_1\boldsymbol{A}\boldsymbol{u}_1 + c_2\boldsymbol{A}\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{A}\boldsymbol{u}_k$

Linearity condition

Remark 7.1.3

Formal definition of Linear Transformation

A linear transformation $T : \mathbf{R}^n \to \mathbf{R}^m$

is a mapping from $\mathbf{R}^n$ to $\mathbf{R}^m$

that satisfies the following condition:

For all vectors $\boldsymbol{u}, \boldsymbol{v}$ in $\mathbf{R}^n$ and scalars a, b

$$T(a\boldsymbol{u} + b\boldsymbol{v}) = aT(\boldsymbol{u}) + bT(\boldsymbol{v})$$

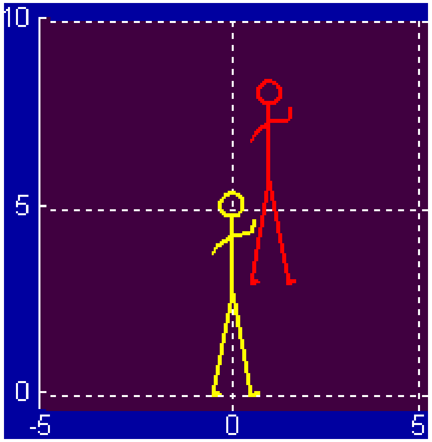
Linearity conditions of T

1. $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$ — T preserves addition
2. $T(c\boldsymbol{u}) = cT(\boldsymbol{u})$ — T preserves scalar multiplication

How to show a mapping is not linear transformation?

Example 7.1.5.1 revisited

$$T_1 : \mathbf{R}^2 \to \mathbf{R}^2 \quad T_1\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+1 \\ y+3 \end{pmatrix}$$



Check the image of zero vector ***0***:

$$T_1\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The property $T(\mathbf{0}) = \mathbf{0}$ is violated

Thus T_1 is not a linear transformation.

How to show a mapping is not linear transformation?

Example 7.1.5.2 revisited

$T_2 : \mathbf{R}^3 \to \mathbf{R}^2$ $\quad T_2\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x^2 \\ yz \end{pmatrix}$

The linearity condition
$T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$
is violated

Check the image of zero vector $\boldsymbol{0}$:

$$T_2\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Does not violate $T(\mathbf{0}) = \mathbf{0}$

$$T_2\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) = T_2\left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$T_2\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) + T_2\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$T_2\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$T_2\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Thus T_2 is not a linear transformation.

What is a linear operator?

Definition 7.1.1

If a linear transformation $T : \mathbf{R}^n \rightarrow \mathbf{R}^n$

maps from $\mathbf{R}^n$ to itself,

we say T is a linear operator on $\mathbf{R}^n$

Domain of T = Codomain of T

In this case, the standard matrix for T is a square matrix.

In example 7.1.2,
I is a linear operator;
O is a linear operator if domain = codomain;
T is not a linear operator.

LT without formula

$T : \mathbf{R}^3 \to \mathbf{R}^2$: the linear transformation such that

$$T\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right) = \begin{pmatrix}1\\3\end{pmatrix} \quad T\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right) = \begin{pmatrix}-1\\2\end{pmatrix} \quad T\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right) = \begin{pmatrix}4\\-1\end{pmatrix}$$

basis for $\mathbf{R}^3$

If the formula /standard matrix of T is NOT given, can we find the image of every vector in $\mathbf{R}^3$ under T ?

YES ! Provided …

Example 7.1.7

$T : \mathbf{R}^3 \to \mathbf{R}^2$: the linear transformation such that

$$\mathrm{T}\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right)=\begin{pmatrix}1\\3\end{pmatrix} \quad \mathrm{T}\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right)=\begin{pmatrix}-1\\2\end{pmatrix} \quad \mathrm{T}\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)=\begin{pmatrix}4\\-1\end{pmatrix}$$

basis for $\mathbf{R}^3$

(a) Find the image of $\begin{pmatrix}-1\\4\\6\end{pmatrix}$ under T.

(b) Find the formula of T.

How to determine LT from basis?

Example 7.1.7.1

$$T\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right)=\begin{pmatrix}1\\3\end{pmatrix}\quad T\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right)=\begin{pmatrix}-1\\2\end{pmatrix}\quad T\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)=\begin{pmatrix}4\\-1\end{pmatrix}$$

$$\begin{pmatrix}-1\\4\\6\end{pmatrix}=3\begin{pmatrix}1\\1\\1\end{pmatrix}+1\begin{pmatrix}0\\1\\1\end{pmatrix}+-2\begin{pmatrix}2\\0\\-1\end{pmatrix}$$

use Gaussian elimination to find the coefficients

$$T\left(\begin{pmatrix}-1\\4\\6\end{pmatrix}\right)=\boxed{T\left(3\begin{pmatrix}1\\1\\1\end{pmatrix}+\begin{pmatrix}0\\1\\1\end{pmatrix}-2\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)} \rightarrow$$

this step can be skipped

Linearity condition

$$=3T\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right)+T\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right)-2T\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)=3\begin{pmatrix}1\\3\end{pmatrix}+\begin{pmatrix}-1\\2\end{pmatrix}-2\begin{pmatrix}4\\-1\end{pmatrix}=\begin{pmatrix}-6\\13\end{pmatrix}$$

Images under LT in terms of basis

Discussion 7.1.6

$\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}\}$: a basis for $\mathbf{R}^n$

Any $\boldsymbol{v}$ in $\mathbf{R}^n$

$$\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n}$$

for some scalar $c_1, c_2, \ldots, c_n$

Suppose $T : \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation.

Linearity condition

$$T(\boldsymbol{v}) = T(\, c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n} \,)$$

image of a general vector $\boldsymbol{v}$

$$= c_1T(\boldsymbol{u_1}) + c_2T(\boldsymbol{u_2}) + \cdots + c_nT(\boldsymbol{u_n})$$

images of the basis vectors

Images under LT in terms of basis

Discussion 7.1.6

$\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}\}$: a basis for $\mathbf{R}^n$

Any $\boldsymbol{v}$ in $\mathbf{R}^n$

$$T(\boldsymbol{v}) = c_1 T(\boldsymbol{u_1}) + c_2 T(\boldsymbol{u_2}) + \cdots + c_n T(\boldsymbol{u_n})$$

Knowing the images $T(\boldsymbol{u_1}), T(\boldsymbol{u_2}), \ldots, T(\boldsymbol{u_n})$

is enough to determine

the image $T(\boldsymbol{v})$ of any vector $\boldsymbol{v}$ in the domain $\mathbf{R}^n$.

The linear transformation T

is completely determined by the images

$T(\boldsymbol{u_1}), T(\boldsymbol{u_2}), \ldots, T(\boldsymbol{u_n})$ of the basis.

How to determine LT from basis?

Example 7.1.7

$T : \mathbf{R}^3 \to \mathbf{R}^2$: the linear transformation such that

$$\mathrm{T}\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right)=\begin{pmatrix}1\\3\end{pmatrix}\quad \mathrm{T}\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right)=\begin{pmatrix}-1\\2\end{pmatrix}\quad \mathrm{T}\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)=\begin{pmatrix}4\\-1\end{pmatrix}$$

Find the formula of T.

Method 1: Direct Gaussian elimination

Method 2: Find $\mathrm{T}(\mathbf{e}_1)$, $\mathrm{T}(\mathbf{e}_2)$, $\mathrm{T}(\mathbf{e}_3)$

Method 3: Stacking matrices

How to determine LT from basis?

Method 1

Example 7.1.7.2

$$T\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right)=\begin{pmatrix}1\\3\end{pmatrix}\quad T\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right)=\begin{pmatrix}-1\\2\end{pmatrix}\quad T\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)=\begin{pmatrix}4\\-1\end{pmatrix}$$

Find the formula of T

$$T\left(\begin{pmatrix}x\\y\\z\end{pmatrix}\right)=\begin{pmatrix}2x-y\\x-y+3z\end{pmatrix}$$

$$\begin{pmatrix}x\\y\\z\end{pmatrix}=c_1\begin{pmatrix}1\\1\\1\end{pmatrix}+c_2\begin{pmatrix}0\\1\\1\end{pmatrix}+c_3\begin{pmatrix}2\\0\\-1\end{pmatrix}$$

$(x-2y+2z)$ $(-x+3y-2z)$ $(y-z)$

use Gaussian elimination to find the coefficients

$$\left(\begin{array}{ccc|c}1&0&2&x\\1&1&0&y\\1&1&-1&z\end{array}\right)$$

$$T\left(\begin{pmatrix}x\\y\\z\end{pmatrix}\right)=c_1T\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right)+c_2T\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right)+c_3T\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)=\boxed{c_1}\begin{pmatrix}1\\3\end{pmatrix}+\boxed{c_2}\begin{pmatrix}-1\\2\end{pmatrix}+\boxed{c_3}\begin{pmatrix}4\\-1\end{pmatrix}$$

in terms of x, y, z

Images of standard basis and standard matrix

Discussion 7.1.8

$T : \mathbf{R}^n \to \mathbf{R}^m$: any linear transformation

$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$: the standard basis for $\mathbf{R}^n$

$$\boldsymbol{A} = \begin{pmatrix} \boxed{\begin{matrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{matrix}} & \begin{matrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{matrix} & \begin{matrix} \cdots \\ \cdots \\ \\ \cdots \end{matrix} & \begin{matrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{matrix} \end{pmatrix}$$

the standard matrix of T

$$\boxed{T(\mathbf{e}_1) = \mathbf{A}\mathbf{e}_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}}$$

The image $T(\boldsymbol{e}_j)$ = the jth column of $\boldsymbol{A}$

$$\boldsymbol{A} = (T(\mathbf{e}_1) \;\; T(\mathbf{e}_2) \;\; \cdots \;\; T(\mathbf{e}_n))$$

Images of standard basis and standard matrix

Example 7.1.9

Method 2

$$T\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right)=\begin{pmatrix}1\\3\end{pmatrix} \quad T\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right)=\begin{pmatrix}-1\\2\end{pmatrix} \quad T\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)=\begin{pmatrix}4\\-1\end{pmatrix}$$

$$\mathbf{A} = (\ T(\mathbf{e}_1)\ T(\mathbf{e}_2)\ T(\mathbf{e}_3)\ )$$

$$\mathbf{A} = \begin{pmatrix}2 & -1 & 0\\ 1 & -1 & 3\end{pmatrix}$$

Find $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$, $T(\mathbf{e}_3)$

Find $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in terms of $\begin{pmatrix}1\\1\\1\end{pmatrix}$ $\begin{pmatrix}0\\1\\1\end{pmatrix}$ $\begin{pmatrix}2\\0\\-1\end{pmatrix}$

$$\left(\begin{array}{ccc|c|c|c}1 & 0 & 2 & 1 & 0 & 0\\ 1 & 1 & 0 & 0 & 1 & 0\\ 1 & 1 & -1 & 0 & 0 & 1\end{array}\right) \xrightarrow[\text{elimination}]{\text{Gauss-Jordan}} \left(\begin{array}{ccc|c|c|c}1 & 0 & 0 & 1 & -2 & 2\\ 0 & 1 & 0 & -1 & 3 & -2\\ 0 & 0 & 1 & 0 & 1 & -1\end{array}\right)$$

$$\mathbf{e}_1 = \begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}0\\1\\1\end{pmatrix} \qquad \mathbf{e}_2 = -2\begin{pmatrix}1\\1\\1\end{pmatrix} + 3\begin{pmatrix}0\\1\\1\end{pmatrix} + \begin{pmatrix}2\\0\\1\end{pmatrix} \qquad \mathbf{e}_3 = 2\begin{pmatrix}1\\1\\1\end{pmatrix} - 2\begin{pmatrix}0\\1\\1\end{pmatrix} - \begin{pmatrix}2\\0\\1\end{pmatrix}$$

$$T(\mathbf{e}_1) = T\begin{pmatrix}1\\1\\1\end{pmatrix} - T\begin{pmatrix}0\\1\\1\end{pmatrix} \qquad T(\mathbf{e}_2) = -2T\begin{pmatrix}1\\1\\1\end{pmatrix} + 3T\begin{pmatrix}0\\1\\1\end{pmatrix} + T\begin{pmatrix}2\\0\\1\end{pmatrix} \qquad T(\mathbf{e}_3) = 2T\begin{pmatrix}1\\1\\1\end{pmatrix} - 2T\begin{pmatrix}0\\1\\1\end{pmatrix} - T\begin{pmatrix}2\\0\\1\end{pmatrix}$$

Stacking the matrix

Method 3

$$T\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right)=\begin{pmatrix}1\\3\end{pmatrix}\quad T\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right)=\begin{pmatrix}-1\\2\end{pmatrix}\quad T\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)=\begin{pmatrix}4\\-1\end{pmatrix}$$

$$\mathbf{A}=\begin{pmatrix}2 & -1 & 0\\1 & -1 & 3\end{pmatrix}$$

$$\left(\begin{array}{ccc|c|c|c}1 & 0 & 2 & 1 & 0 & 0\\1 & 1 & 0 & 0 & 1 & 0\\1 & 1 & -1 & 0 & 0 & 1\end{array}\right)\xrightarrow[\text{elimination}]{\text{Gauss-Jordan}}\left(\begin{array}{ccc|c|c|c}1 & 0 & 0 & 1 & -2 & 2\\0 & 1 & 0 & -1 & 3 & -2\\0 & 0 & 1 & 0 & 1 & -1\end{array}\right)$$

$$\mathbf{A}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}1\\3\end{pmatrix},\ \mathbf{A}\begin{pmatrix}0\\1\\1\end{pmatrix}=\begin{pmatrix}-1\\2\end{pmatrix},\ \mathbf{A}\begin{pmatrix}2\\0\\-1\end{pmatrix}=\begin{pmatrix}4\\-1\end{pmatrix}$$

$$\mathbf{A}\begin{pmatrix}1 & 0 & 2\\1 & 1 & 0\\1 & 1 & -1\end{pmatrix}=\begin{pmatrix}1 & -1 & 4\\3 & 2 & -1\end{pmatrix}$$

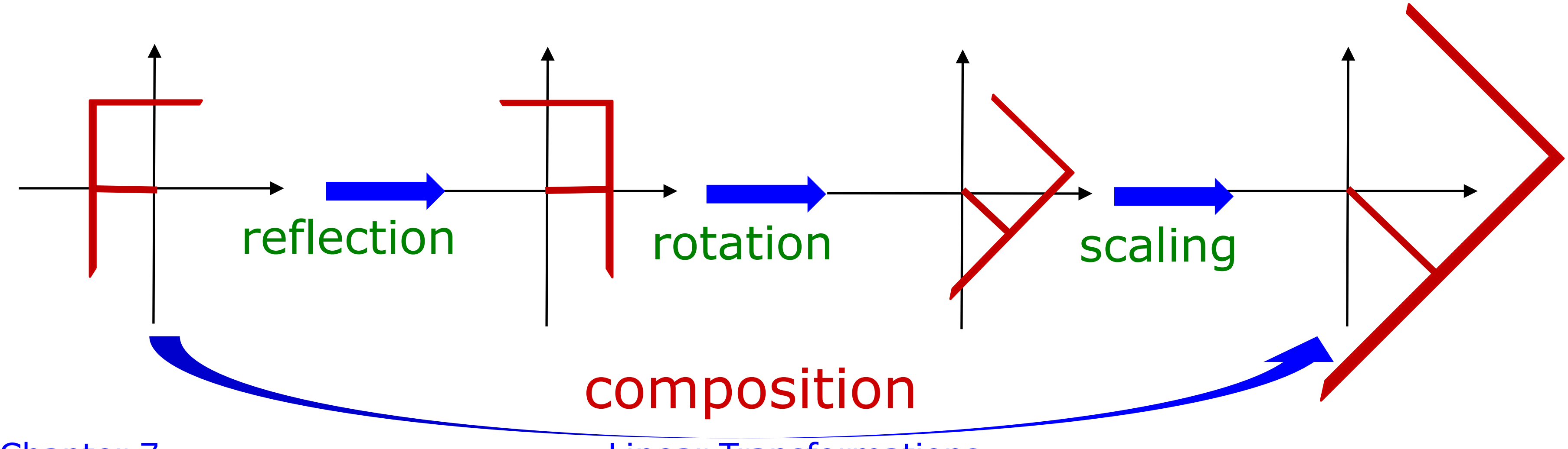
$$\mathbf{A}=\begin{pmatrix}1 & -1 & 4\\3 & 2 & -1\end{pmatrix}\begin{pmatrix}1 & 0 & 2\\1 & 1 & 0\\1 & 1 & -1\end{pmatrix}^{-1}=\begin{pmatrix}1 & -1 & 4\\3 & 2 & -1\end{pmatrix}\begin{pmatrix}1 & -2 & 2\\-1 & 3 & -2\\0 & 1 & -1\end{pmatrix}$$

Section 7.1

Linear Transformations from $\mathbf{R}^n$ to $\mathbf{R}^m$

Objective

- What is the composition of linear transformations?



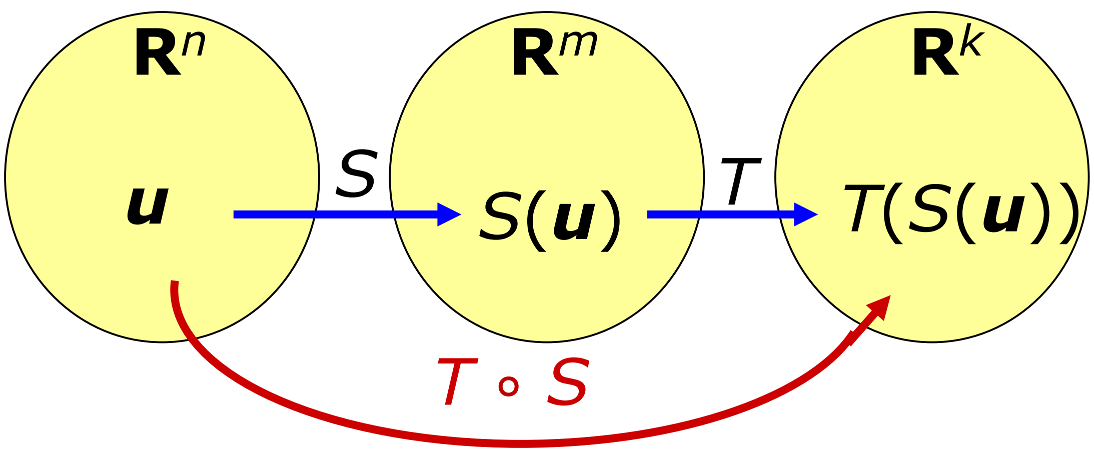
Composition of LT's

Definition 7.1.10

Let $S : \mathbf{R}^n \to \mathbf{R}^m$ and $T : \mathbf{R}^m \to \mathbf{R}^k$ be linear transformations.

The composition of T with S, denoted by $T \circ S$ First S, then T is a mapping from $\mathbf{R}^n$ to $\mathbf{R}^k$ such that

$$(T \circ S)(\boldsymbol{u}) = T(S(\boldsymbol{u})) \text{ for all } \boldsymbol{u} \text{ in } \mathbf{R}^n.$$



Composition of LT's

Example 7.1.12

S: $\mathbf{R}^3 \to \mathbf{R}^2$: the linear transformation defined by

$$S\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x+y \\ z \end{pmatrix} \quad \text{for all } \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ in } \mathbf{R}^3.$$

T : $\mathbf{R}^2 \to \mathbf{R}^3$: the linear transformation defined by

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ y \\ x \end{pmatrix} \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \text{ in } \mathbf{R}^2.$$

Find the composition of T with S.

Composition of LT's

Example 7.1.12

$$S\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x + y \\ z \end{pmatrix}$$

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ y \\ x \end{pmatrix}$$

$T \circ S$ is a mapping from $\mathbf{R}^3$ to $\mathbf{R}^3$:

$$(T \circ S)\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = T\left(S\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right)\right) = T\left(\begin{pmatrix} x + y \\ z \end{pmatrix}\right) = \begin{pmatrix} z \\ z \\ x + y \end{pmatrix}$$

Not recommended; alternative approach later

Is $T \circ S$ a linear transformation ?

Standard matrix of composition of LT's

Theorem 7.1.11

If $S : \mathbf{R}^n \to \mathbf{R}^m$ and $T : \mathbf{R}^m \to \mathbf{R}^k$ are linear transformations

S, T have standard matrices $\boldsymbol{A}$, $\boldsymbol{B}$ respectively

then $T \circ S : \mathbf{R}^n \to \mathbf{R}^k$ is again a linear transformation.

$T \circ S$ has standard matrix $\boldsymbol{BA}$

The proof

Theorem 7.1.11

linear transformation	standard matrix
$S : \mathbf{R}^n \to \mathbf{R}^m$	$\boldsymbol{A}$
$T : \mathbf{R}^m \to \mathbf{R}^k$	$\boldsymbol{B}$
$T \circ S : \mathbf{R}^n \to \mathbf{R}^k$	$\boldsymbol{BA}$

For all $\boldsymbol{u}$ in $\mathbf{R}^n$,

$$(T \circ S)(\boldsymbol{u}) = T(S(\boldsymbol{u})) = T(\boldsymbol{Au}) = \boldsymbol{B}(\boldsymbol{Au}) = (\boldsymbol{BA})\boldsymbol{u}$$

$T \circ S$ is a linear transformation

Standard matrix of composition of LT's

Example 7.1.12

$$S\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x+y \\ z \end{pmatrix} \qquad T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ y \\ x \end{pmatrix}$$

$$(T \circ S)\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} z \\ z \\ x+y \end{pmatrix}$$

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \boldsymbol{B} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\boldsymbol{BA} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$ standard matrix of $T \circ S$

$$(T \circ S)\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \mathbf{BA}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z \\ z \\ x+y \end{pmatrix}$$

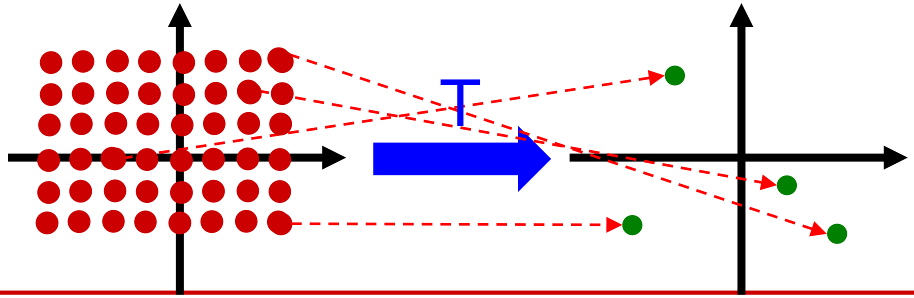
Section 7.2

Ranges and Kernel

Objective

- What are the range and kernel of a linear transformation?
- What are the rank and nullity of a linear transformation?
- What is the Dimension Theorem of linear transformation?

Visualization



$T : \mathbf{R}^2 \to \mathbf{R}^2$ linear transformation

Three possibilities:

- Images under T fill up the whole xy-plane ($\mathbf{R}^2$)
- Images under T all lie on a line
- Images under T all are the same point

range of T

$S : \mathbf{R}^3 \to \mathbf{R}^3$ linear transformation

Four possibilities:

- Images under S fill up the whole xyz-space ($\mathbf{R}^3$)
- Images under S all lie on a plane
- Images under S all lie on a line
- Images under S all are the same point

range of S

What is the range of a LT?

Definition 7.2.1

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

The range of T, denoted by $R(T)$,

is the set of images of T.

$R(T)$ = {images of T }

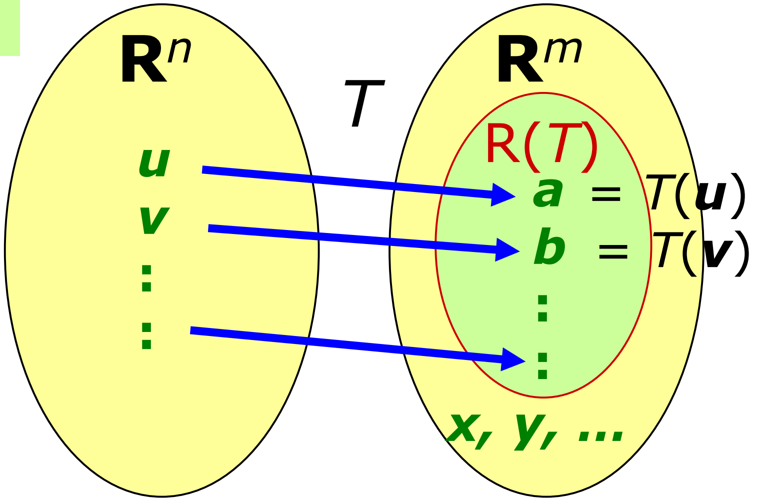
$R(T) = \{T(\boldsymbol{u}) \mid \boldsymbol{u} \in \mathbf{R}^n \}$

explicit set notation

$R(T)$ is a subset of $\mathbf{R}^m$

$R(T)$ may not be equal to $\mathbf{R}^m$

range of T ⊆ codomain of T



What is the range of a LT?

Example 7.2.2

$\mathrm{R}(T) = \{T(\boldsymbol{u}) \mid \boldsymbol{u} \in \mathbf{R}^n\}$

$T : \mathbf{R}^2 \to \mathbf{R}^3$: the linear transformation defined by

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+y \\ y \\ x \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + y\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \text{ in } \mathbf{R}^2.$$

What is $\mathrm{R}(T)$?

$$\mathrm{R}(T) = \left\{ \begin{pmatrix} x+y \\ y \\ x \end{pmatrix} \middle| \; x, y \in \boldsymbol{R} \right\} = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

explicit set notation

linear span form

a plane in $\mathbf{R}^3$

Example 7.2.2

$T : \mathbf{R}^2 \to \mathbf{R}^3$: the linear transformation defined by

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+y \\ y \\ x \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + y\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \text{ for all } \begin{pmatrix} x \\ y \end{pmatrix} \text{ in } \mathbf{R}^2.$$

What is R(T)?

standard matrix $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$

$$\mathrm{R}(T) = \left\{ \begin{pmatrix} x+y \\ y \\ x \end{pmatrix} \middle| \; x, y \in \boldsymbol{R} \right\} = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

explicit set notation

linear span form

column space of $\boldsymbol{A}$

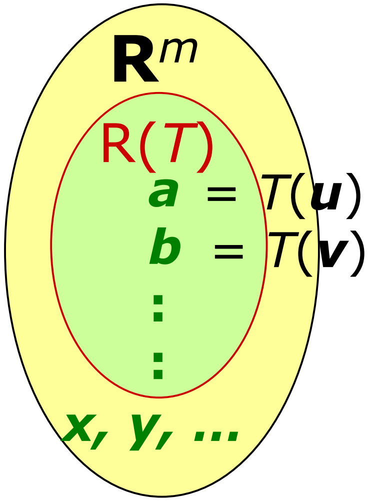
R(T) is the column space of standard matrix

Theorem 7.2.4

T : $\mathbf{R}^n \to \mathbf{R}^m$: a linear transformation

$\boldsymbol{A}$ the standard matrix for T

Then R(T) = span{ columns of $\boldsymbol{A}$ }
= the column space of $\boldsymbol{A}$



R(T) is a subspace of $\mathbf{R}^m$

R(T) is a subset of $\mathbf{R}^m$

What is the rank of a LT?

Definition 7.2.5

Let T be a linear transformation.

The dimension of $R(T)$ = dimension of column space of ***A***

called the rank of T denoted by rank(T)

A the standard matrix for T rank(T) = rank(***A***)

Example 7.2.2:

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+y \\ y \\ x \end{pmatrix} \quad R(T) = \text{span}\underbrace{\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right\}}_{\text{basis}} \quad \text{rank}(T) = 2$$

How to find a basis for R(T)?

Example 7.2.6

$T : \mathbf{R}^4 \to \mathbf{R}^4$: a linear transformation defined by

$$T\left(\begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x + 2y + z \\ x + 3y \\ x + 4y - z \\ y - z \end{pmatrix} \quad \text{for all } \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} \in \boldsymbol{R}^4$$

Find a basis for the range of T and determine the rank of T.

Let $\boldsymbol{A}$ be the standard matrix for T

Same as to find:
a basis for column space of $\boldsymbol{A}$ and rank($\boldsymbol{A}$).

R(T) in terms of basis

Discussion 7.2.3

$T : \mathbf{R}^n \to \mathbf{R}^m$ a linear transformation

$\mathrm{R}(T)$ = span{ columns of $\boldsymbol{A}$ }

= span { $T(\boldsymbol{e_1}), T(\boldsymbol{e_2}), \ldots, T(\boldsymbol{e_n})$ }

If $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}\}$ is any basis for $\mathbf{R}^n$

then $\boxed{\mathrm{R}(T) = \text{span}\,\{\, T(\boldsymbol{u_1}), T(\boldsymbol{u_2}), \ldots, T(\boldsymbol{u_n}) \,\}}$ **?**

$$\Downarrow \qquad\qquad\qquad\qquad \Uparrow$$

$$T(\boldsymbol{v}) \qquad c_1T(\boldsymbol{u_1}) + c_2T(\boldsymbol{u_2}) + \cdots + c_nT(\boldsymbol{u_n})$$

$$\Downarrow \qquad\qquad\qquad\qquad \Uparrow$$

$$T(\, c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n} \,)$$

We can write $\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n}$

Finding range R(T) and $\boxed{\text{its basis}}$

T: $\mathbf{R}^n \to \mathbf{R}^m$

I. if formula of T is given

- R(T) = {formula in $x_1, x_2, \ldots, x_n \mid x_1, x_2, \ldots, x_n \in \mathbf{R}$ }

II. if standard matrix **A** is given

- R(T) = span{ columns of **A** }

  or part I above

Find basis for column space of **A**

III. if image of a basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ for $\mathbf{R}^n$ is given

- R(T) = span$\{T(\mathbf{u}_1), T(\mathbf{u}_2), \ldots, T(\mathbf{u}_n)\}$

Throw out the redundant vectors in the span (use column space method if necessary)

Visualization

$T : \mathbf{R}^2 \to \mathbf{R}^2$ linear transformation

- Images under T fill up the whole xy-plane ($\mathbf{R}^2$)
- Images under T all lie on a line
- Images under T all are the same point

range of T

Some information is lost

kernel of T (or S)

$S : \mathbf{R}^3 \to \mathbf{R}^3$ linear transformation

- Images under S fill up the whole xyz-space ($\mathbf{R}^3$)
- Images under S all lie on a plane
- Images under S all lie on a line
- Images under S all are the same point

range of S

What is the kernel of a LT?

Definition 7.2.7

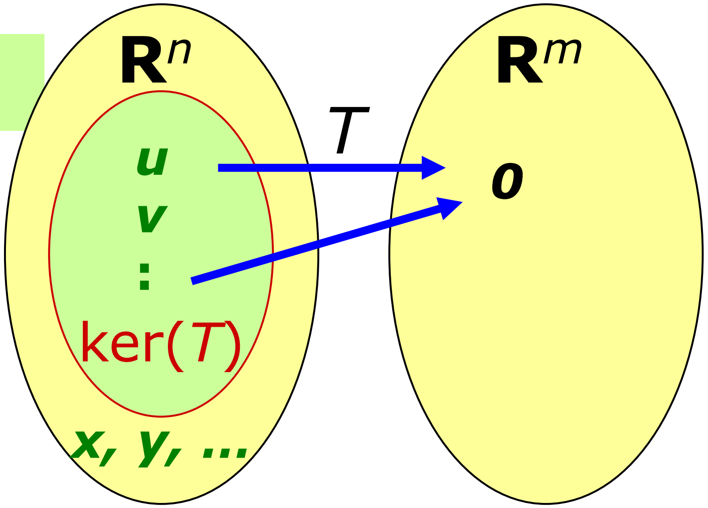
Let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation.

The kernel of T, denoted by $\ker(T)$, is the set of vectors in $\mathbf{R}^n$ whose image is the zero vector in $\mathbf{R}^m$.

$\ker(T)$ = {vectors that map to **0** under T }

$\ker(T) = \{\, \boldsymbol{u} \in \mathbf{R}^n \mid T(\boldsymbol{u}) = \mathbf{0} \,\}$

implicit set notation



$\ker(T)$ is a subset of $\mathbf{R}^n$

$\ker(T)$ may not be equal to $\mathbf{R}^n$

How to find kernel of a LT?

Example 7.2.8.1

$\ker(T) = \{\boldsymbol{u} \in \mathbf{R}^3 \mid \boxed{T(\boldsymbol{u}) = \mathbf{0}}\}$

$T : \mathbf{R}^3 \to \mathbf{R}^4$: a linear transformation defined by

$$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 2x - y \\ x - y + 3z \\ -5x + y \\ x - z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ for all } \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \boldsymbol{R}^3$$

homog. system → only trivial solution

What is the kernel of T ?

$$\ker(T) = \left\{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right\}$$

Find all $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \boldsymbol{R}^3$ that satisfy this hom. system.

the zero space

How to find kernel of a LT?

Example 7.2.8.2

$\ker(T) = \{\boldsymbol{u} \in \mathbf{R}^3 \mid \boxed{T(\boldsymbol{u}) = \mathbf{0}}\}$

Let $T : \mathbf{R}^3 \to \mathbf{R}^3$ be a linear transformation defined by

$$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} z - y \\ 0 \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{for all } \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \boldsymbol{R}^3$$

solve for x, y, z

we get $z = y$ and $x = 0$

$$\ker(T) = \left\{ \begin{pmatrix} 0 \\ y \\ y \end{pmatrix} \;\middle|\; y \in \mathbf{R} \right\} = \operatorname{span}\left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

a subspace of dimension 1

Theorem 7.2.9

$T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation

$\boldsymbol{A}$ the standard matrix for T $\quad$ $T(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u}$

$\ker(T)$ = all $\boldsymbol{u}$ such that $T(\boldsymbol{u}) = \mathbf{0}$

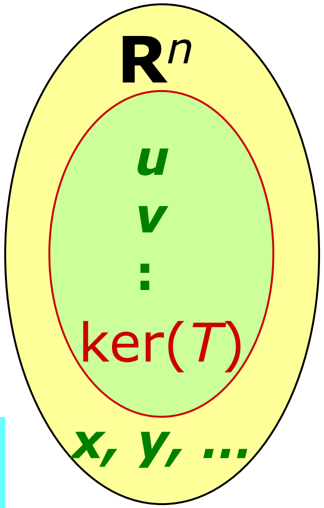
= all $\boldsymbol{u}$ such that $\boldsymbol{A}\boldsymbol{u} = \mathbf{0}$

= the solution space of $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$

= the nullspace of $\boldsymbol{A}$

$\ker(T)$ is a subspace of $\mathbf{R}^n$

$\ker(T)$ is a subset of $\mathbf{R}^n$



What is the nullity of a LT?

Definition 7.2.10

Let T be a linear transformation.

The dimension of $\ker(T)$

called the nullity of T

denoted by $\text{nullity}(T)$

$\ker(T)$ = the nullspace of standard matrix $\boldsymbol{A}$

$$\text{nullity}(T) = \text{nullity}(\boldsymbol{A})$$

How to find a basis for ker(T)?

Example 7.2.11.1

In example 7.2.8.1,

$$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 2x - y \\ x - y + 3z \\ -5x + y \\ x - z \end{pmatrix}$$

$$\ker(T) = \left\{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right\}$$

the nullity of T is 0

In example 7.2.8.2,

$$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x - z \\ 0 \\ y \end{pmatrix}$$

$$\ker(T) = \operatorname{span}\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right\}$$

the nullity of T is 1

Dimension Theorem for LT

Theorem 7.2.12

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

$$\text{rank}(T) + \text{nullity}(T) = n$$

By Thm 4.3.4. $\text{rank}(\boldsymbol{A}) + \text{nullity}(\boldsymbol{A}) = n$ (number of columns)

Proof

The standard matrix $\boldsymbol{A}$ of T is of size m x n

Range and kernel in proof

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

$\mathrm{Ker}(T) = \{ \mathbf{v} \in \mathbf{R}^n \mid T(\mathbf{v}) = \mathbf{0}\}$

if you want to show:

In a proof, if you start with: $\mathbf{v} \in \ker(T)$,

try to show:

you should follow by: $T(\mathbf{v}) = \mathbf{0}$.

$\mathcal{R}(T) = \{ T(\mathbf{v}) \mid \mathbf{v} \in \mathbf{R}^n\}$

if you want to show:

In a proof, if you start with: $\mathbf{v} \in R(T)$,

try to show:

you should follow by: $\mathbf{v} = T(\mathbf{u})$ for some $\mathbf{u} \in \mathbf{R}^n$.

Ex 7 Q17

S: $\mathbf{R}^n \to \mathbf{R}^m$ and T: $\mathbf{R}^m \to \mathbf{R}^k$ linear transformations

$\mathrm{Ker}(S) \subseteq \mathrm{Ker}(T \circ S)$

Hint: Take $\mathbf{u} \in \ker(S)$. Show that $\mathbf{u} \in \ker(T \circ S)$.

$S(\mathbf{u}) = \mathbf{0}$

$(T \circ S)(\mathbf{u}) = \mathbf{0}$

$R(T \circ S) \subseteq R(T)$

Hint: Take $\mathbf{u} \in R(T \circ S)$. Show that $\mathbf{u} \in R(T)$.

$\mathbf{u} = (T \circ S)(\mathbf{v})$ for some $\mathbf{v} \in \mathbf{R}^n$

$\mathbf{u} = T(\mathbf{w})$ for some $\mathbf{w} \in \mathbf{R}^m$