Unit 1: Machine Learning
TCS 509

Best wishes to all readers ☺

Professor Jay Bhatnagar

7-2 Mean, Median, Mode, and Range

mean
median
mode
range
outlier

Helpful Hint

The mean is sometimes called the average.

The **mean** is the sum of the data values divided by the number of data items.

The **median** is the middle value of an odd number of data items arranged in order. For an even number of data items, the median is the average of the two middle values.

The **mode** is the value or values that occur most often. When all the data values occur the same number of times, there is no mode.

The **range** of a set of data is the difference between the greatest and least values. It is used to show the spread of the data in a data set.

Outlier

In Statistics, an outlier are data points that differ significantly from other observations (neighborhoods – DBSCAN density based spatial clustering of apps under noise, normalized deviation - z –score). An outlier may be due to variability in the measurement, an indication of novel data, or it may be the result of experimental error; the latter are sometimes excluded from the data set. An outlier can be an indication of exciting possibility but can also cause serious problems in statistical analyses.

Additional Example 1: Finding the Mean, Median, Mode, and Range of Data

Find the mean, median, mode, and range of the data set.

4, 7, 8, 2, 1, 2, 4, 2

mean:

$$\underbrace{4 + 7 + 8 + 2 + 1 + 2 + 4 + 2}_{\text{8 items}} = \underbrace{30}_{\text{sum}}$$

Add the values.

$$30 \div 8 = 3.75$$

Divide the sum by the number of items.

The mean is 3.75.

Additional Example 1 Continued

Find the mean, median, mode, and range of the data set.

4, 7, 8, 2, 1, 2, 4, 2

median:

~~1, 2, 2,~~ 2, 4, ~~4, 7, 8~~ *Arrange the values in order.*

$2 + 4 = 6$

$6 \div 2 = 3$

There are two middle values, so find the mean of these two values.

The median is 3.

Additional Example 1 Continued

Find the mean, median, mode, and range of the data set.

4, 7, 8, 2, 1, 2, 4, 2

mode:

1, 2, 2, 2, 4, 4, 7, 8 *The value 2 occurs three times.*

The mode is 2.

Check It Out: Example 1

Find the mean, median, mode, and range of the data set.

6, 4, 3, 5, 2, 5, 1, 8

mean:

$$\underbrace{6 + 4 + 3 + 5 + 2 + 5 + 1 + 8}_{\text{8 items}} = \underbrace{34}_{\text{sum}}$$

Add the values.

$$34 \div 8 = 4.25$$

Divide the sum by the number of items.

The mean is 4.25.

Check It Out: Example 1 Continued

Find the mean, median, mode, and range of the data set.

6, 4, 3, 5, 2, 5, 1, 8

median:

~~1, 2, 3,~~ 4, 5, ~~5, 6, 8~~

Arrange the values in order.

4 + 5 = 9

9 ÷ 2 = 4.5

There are two middle values, so find the mean of these two values.

The median is 4.5.

Check It Out: Example 1 Continued

Find the mean, median, mode, and range of the data set.

6, 4, 3, 5, 2, 5, 1, 8

mode:

1, 2, 3, 4, 5, 5, 6, 8 *The value 5 occurs two times.*

The mode is 5.

Check It Out: Example 1 Continued

Find the mean, median, mode, and range of the data set.

6, 4, 3, 5, 2, 5, 1, 8

range:

1, 2, 3, 4, ~~5, 5~~, 6, 8

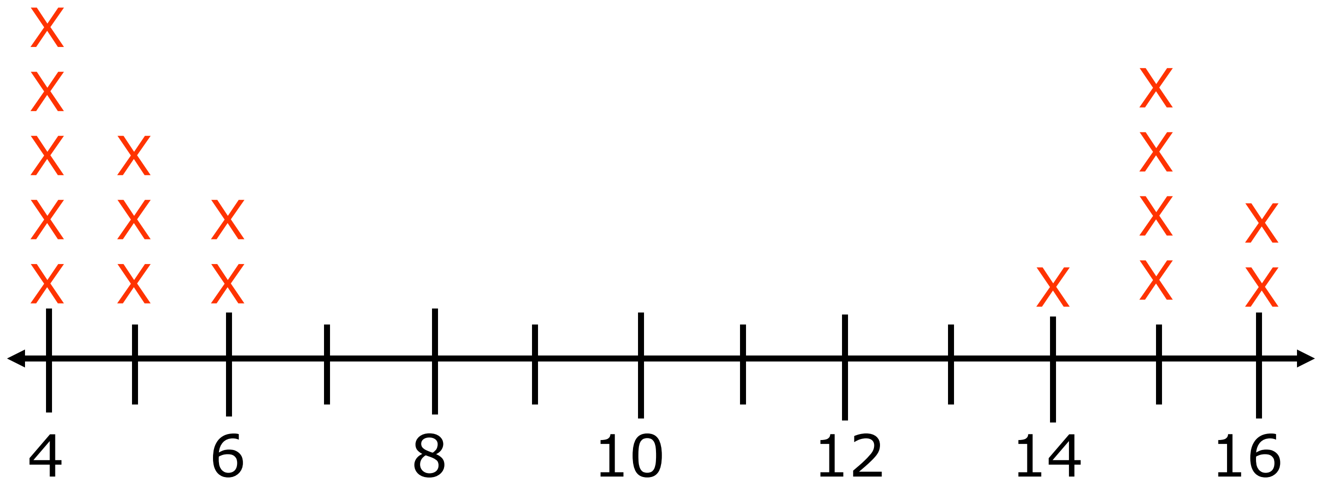
Subtract the least value from the greatest value.

$8 - 1 = 7$

The range is 7.

Additional Example 2: Choosing the Best Measure to Describe a Set of Data

The line plot shows the number of miles each of the 17 members of the cross-country team ran in a week. Which measure of central tendency best describes this data? Justify your answer.



Additional Example 2 Continued

The line plot shows the number of miles each of the 17 members of the cross-country team ran in a week. Which measure of central tendency best describes this data? Justify your answer.

mean:

$$\frac{4 + 4 + 4 + 4 + 4 + 5 + 5 + 5 + 6 + 6 + 14 + 15 + 15 + 15 + 15 + 16 + 16}{17}$$

$$= \frac{153}{17} = 9$$

The mean is 9. The mean best describes the data set because the data is clustered fairly evenly about two areas.

Additional Example 1 Continued

Find the mean, median, mode, and range of the data set.

4, 7, 8, 2, 1, 2, 4, 2

range:

1, 2, 2, 2, 4, 4, 7, 8

Subtract the least value from the greatest value.

$8 - 1 = 7$

The range is 7.

Additional Example 2 Continued

The line plot shows the number of miles each of the 17 members of the cross-country team ran in a week. Which measure of central tendency best describes this data? Justify your answer.

median:

4, 4, 4, 4, 4, 5, 5, 5, 6, 6, 14, 15, 15, 15, 15, 16, 16

The median is 6. The median does not best describe the data set because many values are not clustered around the data value 6.

Additional Example 2 Continued

The line plot shows the number of miles each of the 17 members of the cross-country team ran in a week. Which measure of central tendency best describes this data? Justify your answer.

mode:

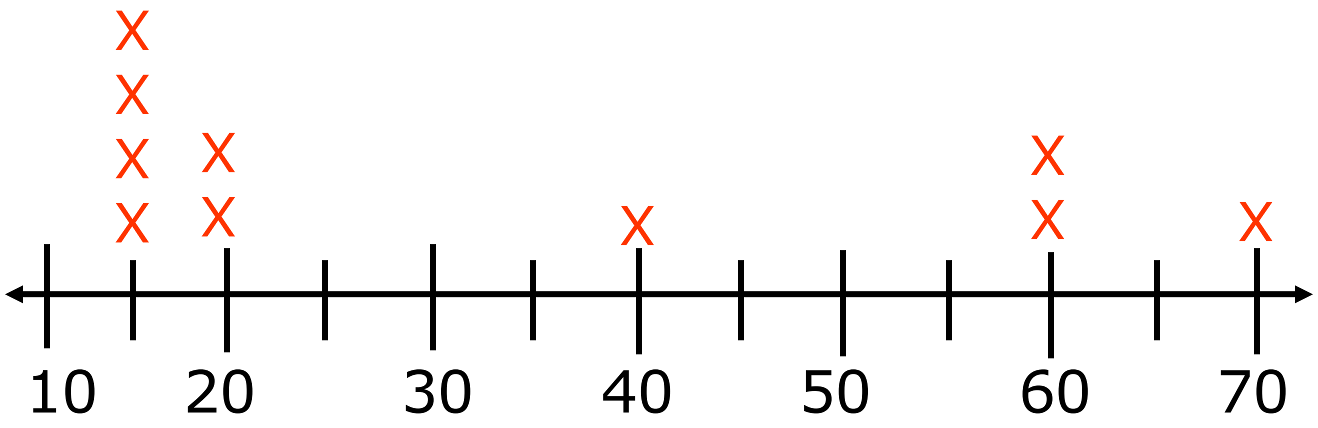
The greatest number of X’s occur above the number 4 on the line plot.

The mode is 4.

The mode focuses on one data value and does not describe the data set.

Check It Out: Example 2

The line plot shows the number of dollars each of the 10 members of the cheerleading team raised in a week. Which measure of central tendency best describes this data? Justify your answer.



Check It Out: Example 2 Continued

The line plot shows the number of dollars each of the 10 members of the cheerleading team raised in a week. Which measure of central tendency best describes this data? Justify your answer.

mean:

$$\frac{15 + 15 + 15 + 15 + 20 + 20 + 40 + 60 + 60 + 70}{10}$$

$$= \frac{330}{10} = 33$$

The mean is 33. Most of the cheerleaders raised less than $33, so the mean does not describe the data set best.

Check It Out: Example 2 Continued

The line plot shows the number of dollars each of the 10 members of the cheerleading team raised in a week. Which measure of central tendency best describes this data? Justify your answer.

median:

15, 15, 15, 15, 20, 20, 40, 60, 60, 70

The median is 20. The median best describes the data set because it is closest to the amount most cheerleaders raised.

Check It Out: Example 2 Continued

The line plot shows the number of dollars each of the 10 members of the cheerleading team raised in a week. Which measure of central tendency best describes this data? Justify your answer.

mode:

The greatest number of X's occur above the number 15 on the line plot.

The mode is 15.

The mode focuses on one data value and does not describe the data set.

7-2 Mean, Median, Mode, and Range

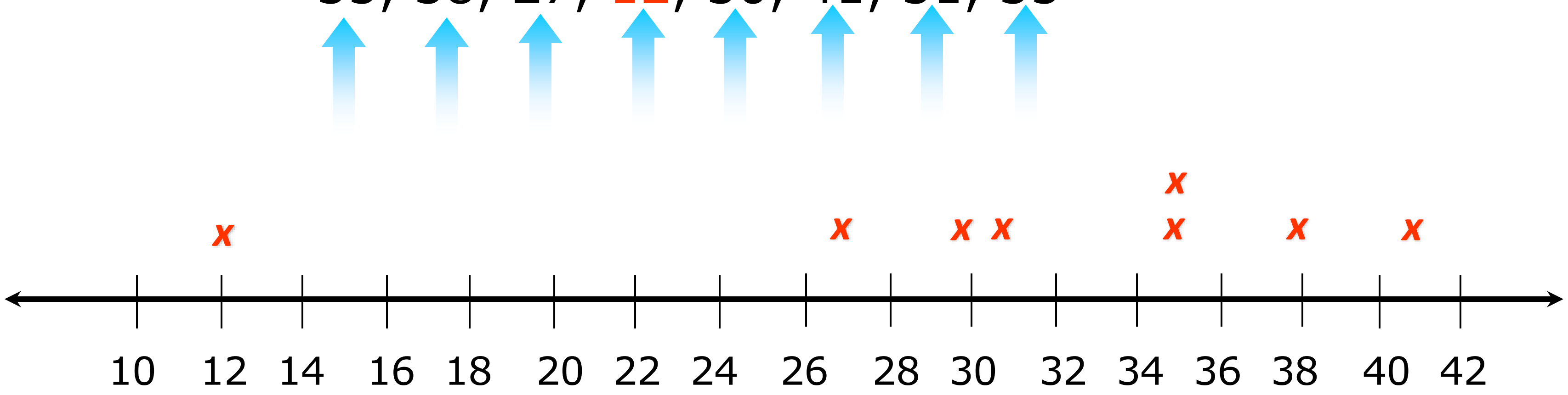
IMP

Measure	Most Useful When
mean	The data are spread fairly evenly
median	The data set has an outlier
mode	The data involve a subject in which many data points of one value are important, such as election results.

7-2 Mean, Median, Mode, and Range

In the data set below, the value 12 is much less than the other values in the set. An extreme value such as this is called an <u>outlier</u>.

35, 38, 27, 12, 30, 41, 31, 35



Additional Example 3: Exploring the Effects of Outliers on Measures of Central Tendency

The data shows Sara's scores for the last 5 math tests: 88, 90, 55, 94, and 89. Identify the outlier in the data set. Then determine how the outlier affects the mean, median, and mode of the data. Then tell which measure of central tendency best describes the data with the outlier.

55, 88, 89, 90, 94

outlier $\longrightarrow$ 55

Additional Example 3 Continued

With the Outlier

55, 88, 89, 90, 94

outlier → 55

mean:	median:	mode:
55+88+89+90+94 = 416 416 ÷ 5 = 83.2 The mean is 83.2.	55, 88, 89, 90, 94 The median is 89.	There is no mode.

Additional Example 3 Continued

Without the Outlier

~~55~~, 88, 89, 90, 94

mean:

88+89+90+94 = 361

361 ÷ 4 = 90.25

The mean is 90.25.

median:

88, $\frac{89+90}{2}$, 94

= 89.5

The median is 89.5.

mode:

There is no mode.

Since all the data values occur the same number of times, the set has no mode.

Additional Example 3 Continued

	Without the Outlier	With the Outlier
mean	90.25	83.2
median	89.5	89
mode	no mode	no mode

Adding the outlier decreased the mean by 7.05 and the median by 0.5.

The mode did not change.

The median best describes the data with the outlier.

Check It Out: Example 3

Identify the outlier in the data set. Then determine how the outlier affects the mean, median, and mode of the data. The tell which measure of central tendency best describes the data with the outlier.

63, 58, 57, 61, 42

42, 57, 58, 61, 63

outlier → 42

Check It Out: Example 3 Continued

With the Outlier

42, 57, 58, 61, 63

outlier → 42

mean:	median:	mode:
$42+57+58+61+63 = 281$	42, 57, 58, 61, 63	
$281 \div 5 = 56.2$		
The mean is 56.2.	The median is 58.	There is no mode.

Check It Out: Example 3 Continued

Without the Outlier

~~42~~, 57, 58, 61, 63

mean:	median:	mode:
57+58+61+63 = 239 239 ÷ 4 = 59.75 The mean is 59.75.	57, 58, 61, 63 $\frac{58+61}{2}$ = 59.5 The median is 59.5.	There is no mode.

Check It Out: Example 3 Continued

	Without the Outlier	With the Outlier
mean	59.75	56.2
median	59.5	58
mode	no mode	no mode

Adding the outlier decreased the mean by 3.55 and decreased the median by 1.5.

The mode did not change.

The median best describes the data with the outlier.

Lesson Quiz: Part I

1. Find the mean, median, mode, and range of the data set. 8, 10, 46, 37, 20, 8, and 11

mean: 20; median: 11; mode: 8; range: 38

Lesson Quiz: Part II

2. Identify the outlier in the data set, and determine how the outlier affects the mean, median, and mode of the data. Then tell which measure of central tendency best describes the data with and without the outlier. Justify your answer.

85, 91, 83, 78, 79, 64, 81, 97

The outlier is 64. Without the outlier the mean is 85, the median is 83, and there is no mode. With the outlier the mean is 82, the median is 82, and there is no mode. Including the outlier decreases the mean by 3 and the median by 1, there is no mode. Because they have the same value and there is no outlier, the median and mean describes the data with the outlier. The median best describes the data without the outlier because it is closer to more of the other data values than the mean.

Measurement of dispersion tendency:

Absolute Deviation, Squared deviation, Mean Square error/ MS deviation aka Variance

Population *vs*. Sample: Designing to Sample

Machine Learning

What is Learning?

- Herbert Simon: “Learning is any process by which a system improves performance from experience.”
- What is the task?
  - Classification
  - Categorization/clustering
  - Problem solving / planning / control
  - Prediction
  - others

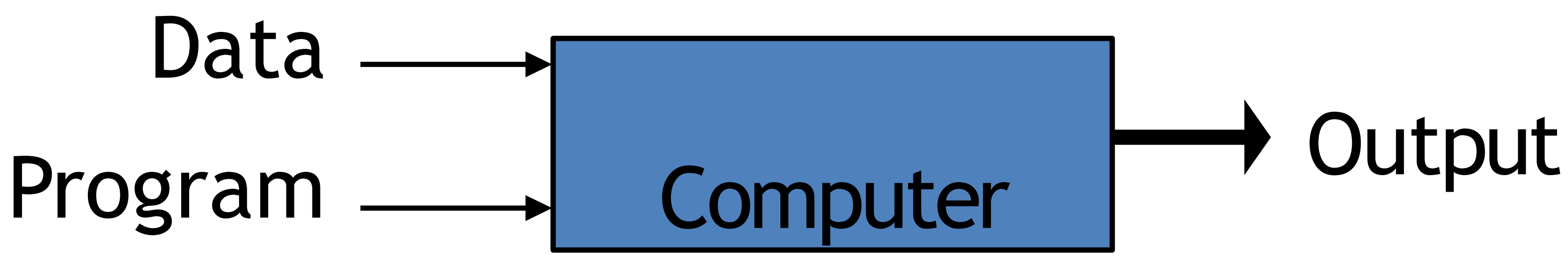
Why Study Machine Learning?
Developing Better Computing Systems

- Develop systems that are too difficult/expensive to construct manually because they require specific detailed skills or knowledge tuned to a specific task (***knowledge engineering bottleneck***).
- Develop systems that can automatically adapt and customize themselves to individual users.
  - Personalized news or mail filter
  - Personalized tutoring
- Discover new knowledge from large databases (***data mining***).
  - Market basket analysis (e.g. diapers and beer)
  - Medical text mining (e.g. migraines to calcium channel blockers to magnesium)

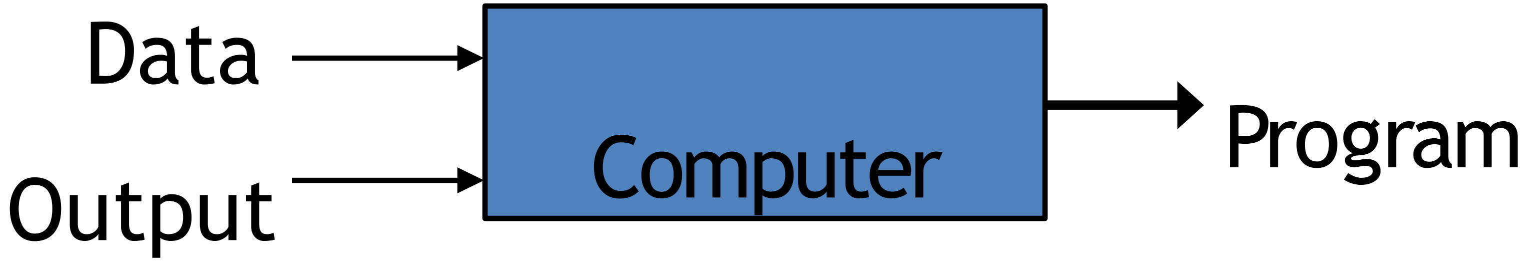
Related Disciplines

- Artificial Intelligence
- Data Mining
- Probability and Statistics
- Information theory
- Numerical optimization
- Computational complexity theory
- Control theory (adaptive)
- Psychology (developmental, cognitive)
- Neurobiology and many more

Traditional Programming



Machine Learning



Sample Applications

- Web search
- Computational biology
- Finance
- E-commerce
- Space exploration
- Robotics
- Information extraction
- Social networks
- Debugging software
- [Your favorite area]

Human Learning

- It is a process of gaining information through observation, to solve/deal with real world scenarios.
- Human learning happens in one of the three ways:
  - An expert in the subject directly teaches us
  - We build our own notion indirectly based on what we have learnt in the past
  - We learn ourselves, may be after multiple attempts.

Types of Human Learning

- Learning under expert guidance
- Learning guided by knowledge gained from experts
- Learning by self

MACHINE LEARNING

A computer program is said to learn from experience ‘E’ with respect to some class of tasks ‘T’ and performance measure ‘P’, if its performance at tasks in ‘T’, as measured by ‘P’, improves with experience ‘E’.

- *Tom Mitchell*

History of Machine Learning

- 1950s
  - Automation, Machines (mechanical then arrived computation)
- 1960s:
  - Neural networks: Perceptron
  - Pattern recognition – Cover, Bradley, Brieman, Schapire, Nilsson, Hastie, Friedman, Tibshirani, Vapnik, Rumelhart, Schimdhuber, Hinton, Bengio, Blum
  - Learning in the limit theory
  - Minsky and Papert prove limitations of Perceptron

1970s:
  - Symbolic concept induction
  - Winston's arch learner
  - Expert systems and the knowledge acquisition bottleneck
  - Quinlan's ID3
  - Michalski's AQ and soybean diagnosis

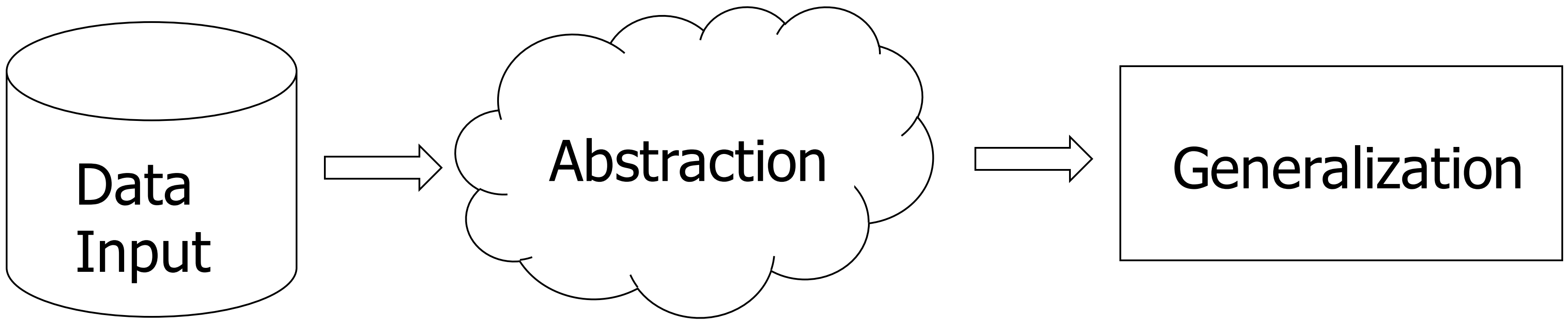
History of Machine Learning (cont.)

- 1980s:
  - Advanced decision tree and rule learning
  - Explanation-based Learning (EBL)
  - Learning and planning and problem solving
  - Utility problem, Analogy
  - Cognitive architectures
  - Resurgence of neural networks (connectionism, backpropagation)
- 1990s
  - Data mining
  - Adaptive software agents and web applications
  - Text learning
  - Reinforcement learning (RL) ….. 1950’s Mendel, Fu, Richard Sutton
  - Inductive Logic Programming (ILP)
  - Ensembles: Bagging, Boosting, and Stacking
  - Bayes Net learning

History of Machine Learning (cont.)

- 2000s
  - Support vector machines
  - Kernel & Graphical models
  - Statistical relational and Transfer learning
  - Sequence labeling
  - Collective classification and structured outputs
  - Computer Systems Applications
    - Compilers
    - Debugging
    - Graphics
    - Security (intrusion, virus, and worm detection)
  - E-mail management
  - Personalized assistants that learn
  - Learning in robotics and vision

How do Machines Learn?



1. Data Input: Past data or information is utilized as a basis for future decision-making
2. Abstraction: The input data is represented in a broader way through the underlying algorithm.
3. Generalization: The abstracted representation is generalized to form a framework for making decisions.

Abstraction

- The data, given as input, cannot be used in the original form.
- **Abstraction helps in deriving a conceptual map based on the** input data.
- The model may be in any one of forms:
  - Computational blocks like if/else rules
  - Mathematical equations
  - Specific data structures like trees or graphs
  - Logical groupings of similar observations

Abstraction Cont'd

- The choice of the model is human specific.
- Selection of model is based on:
  - The type of problem to be solved.
  - Nature of the input data.
  - Domain of the problem.

Generalization

- Generalization is used for taking the decisions after training the model.
- The model is trained for a limited set of data. If we want to apply the model to take decision on a set of unknown data, we may encounter following problems:
    - The trained model is aligned with training data too much, hence may not represent the actual trend.
    - The test data possess certain characteristics apparently unknown to the training data.

Well-posed learning problem

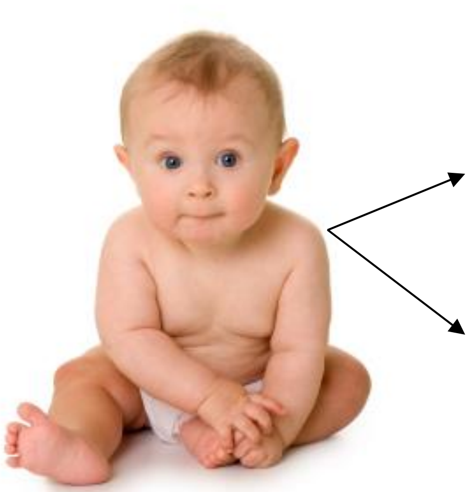
- A framework can be designed for deciding whether a problem can be solved using ML. The framework should answer:
  - What is the problem?
  - Why does the problem need to be solved?
  - How to solve the problem?

Types of Machine Learning

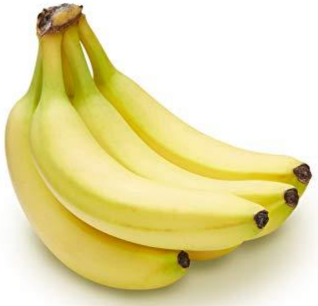
- Supervised learning: A machine predicts the class of unknown objects based on prior class-related information of similar objects. Also called predictive learning.
- Unsupervised/clustering learning: A machine finds patterns in unknown objects by grouping similar objects together. Also called descriptive learning.
- Reinforcement learning: A machine learns to act on its own to achieve the given goals.

SUPERVISED LEARNING

- The major motivation of supervised learning is to learn from past information.
- But how do machine learns?
  - By TRAINING DATA using labels.



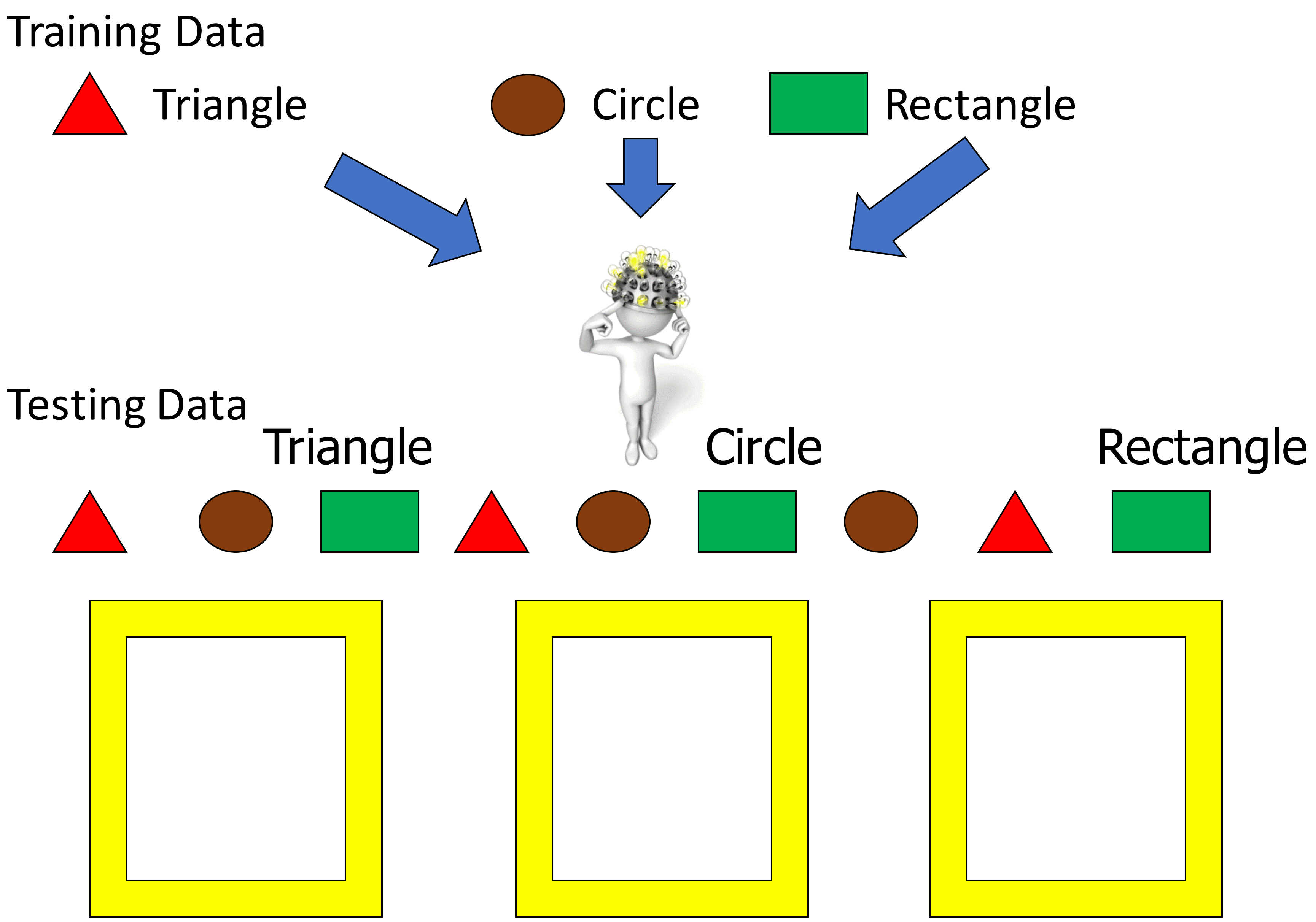
Apple



Banana

What’s this?

It’s an Apple


Training Data
Triangle
Circle
Rectangle
Testing Data
Triangle
Circle
Rectangle

Examples of Supervised Learning

- Predicting the results of a game.
- Predicting a tumour is malignant or benign.
- Predicting the piece of domains like real estate, stocks, etc.
- Classify texts such as classifying a set of emails as spam or non-spam.

Classification & Regression

- When we are trying to predict a categorical or nominal variable, the problem is known as a **classification** problem.
  - Ex: Identify Cat or Dog
- When we are trying to predict a real-valued variable, the problem falls under the category of **regression**.
  - Ex: Predict the value of a property.

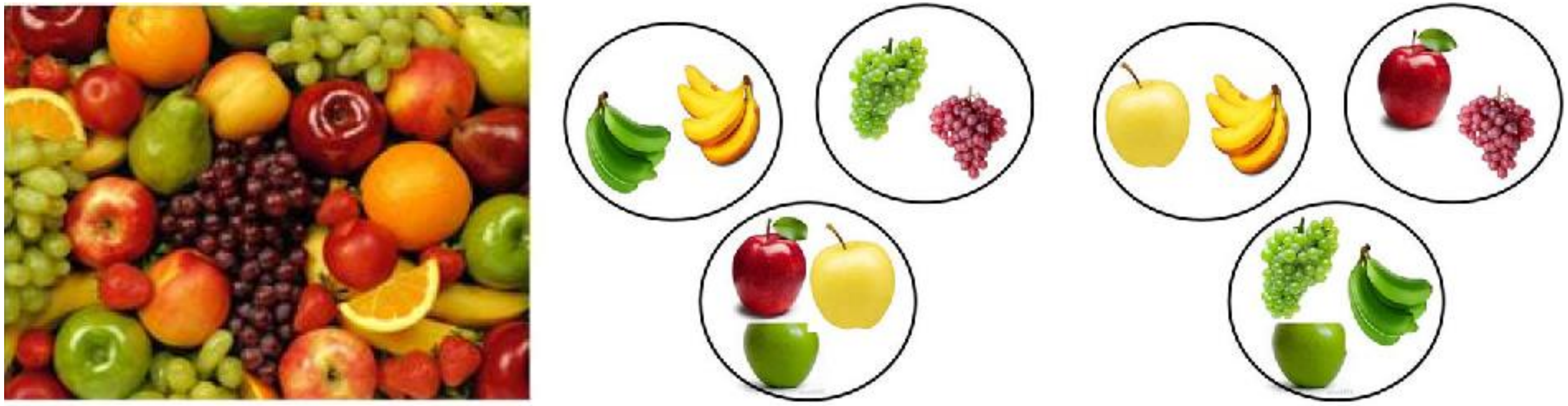
Unsupervised Learning

- Unsupervised learning(USL) is a type of self-organized learning that helps find previously unknown patterns in data set without pre-existing labels.
- The objective is to take a dataset as input and try to find natural grouping or patterns within the data elements or records.
- Hence, USL is termed as Descriptive Model and the process of USL is called Pattern or Knowledge Discovery.
- In unsupervised learning, the system is presented with unlabeled, uncategorized data and the system’s algorithms act on the data without prior training. The output is dependent upon the coded algorithms.

Unsupervised Learning

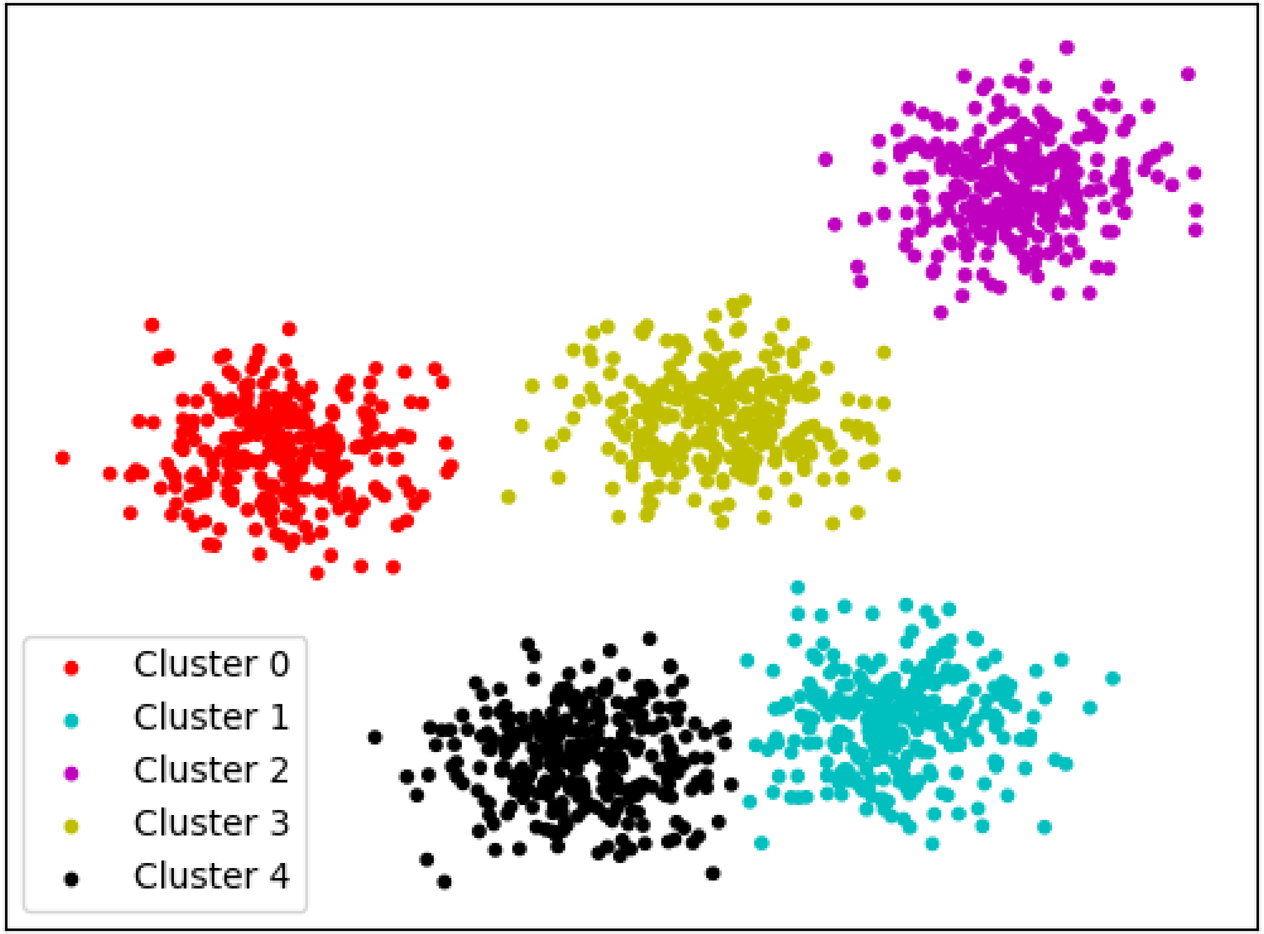
- Clustering is the main type of Unsupervised Learning.
- Clustering groups similar objects together.

Input Data Cluster by type Cluster by color

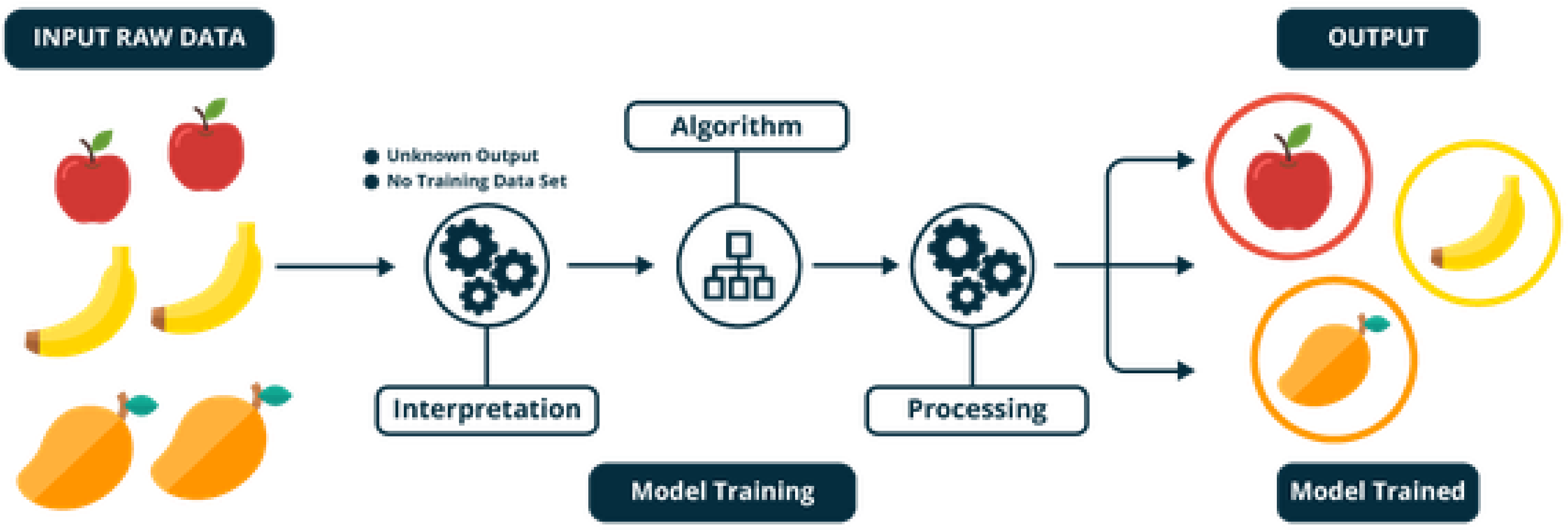


(a) (b) (c)

Sample Clustering



Unsupervised Learning



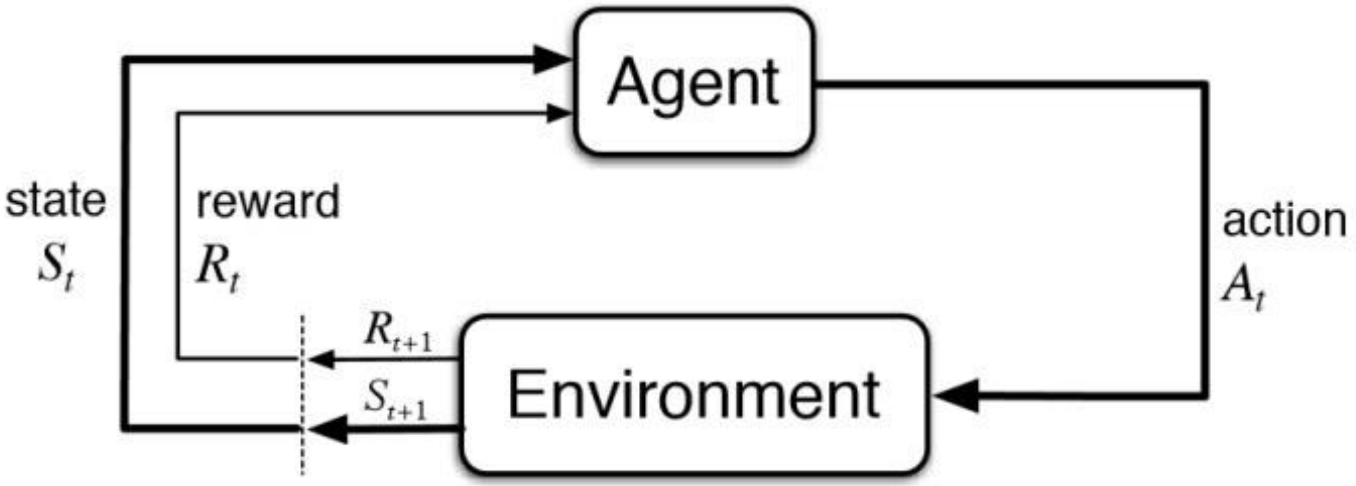
Unsupervised Learning Example: Categorize Cats and Dogs

Reinforcement Learning

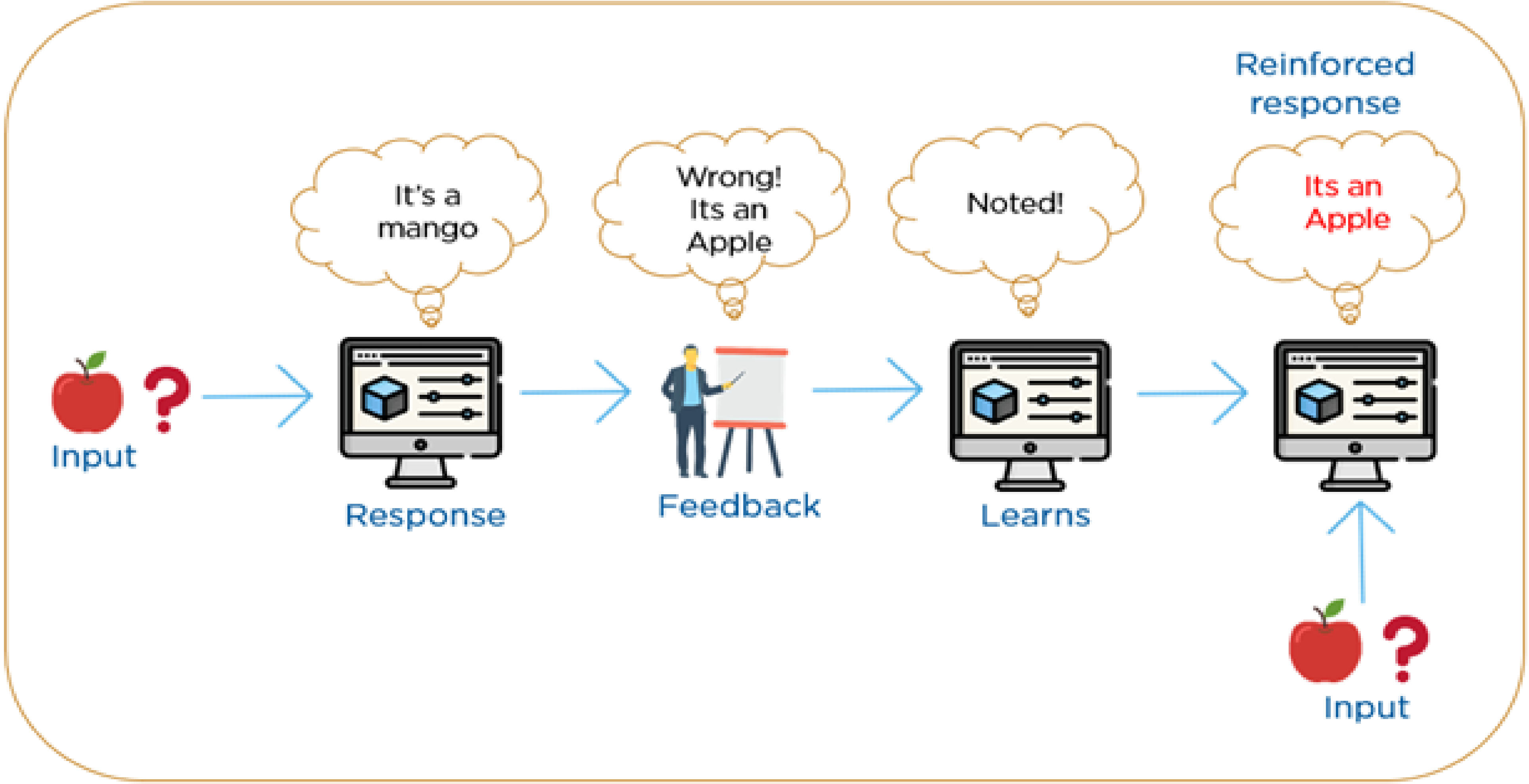
- It is about taking suitable action to maximize reward in a particular situation.
- It is employed by various software and machines to find the best possible behavior or path it should take in a specific situation.
- Close to human learning.

Cont'd

- Algorithm learns a policy of how to act in a given environment.
- Every action has some impact in the environment, and the environment provides rewards that guides the learning algorithm.



Reinforcement Learning



Different Varieties of Machine Learning

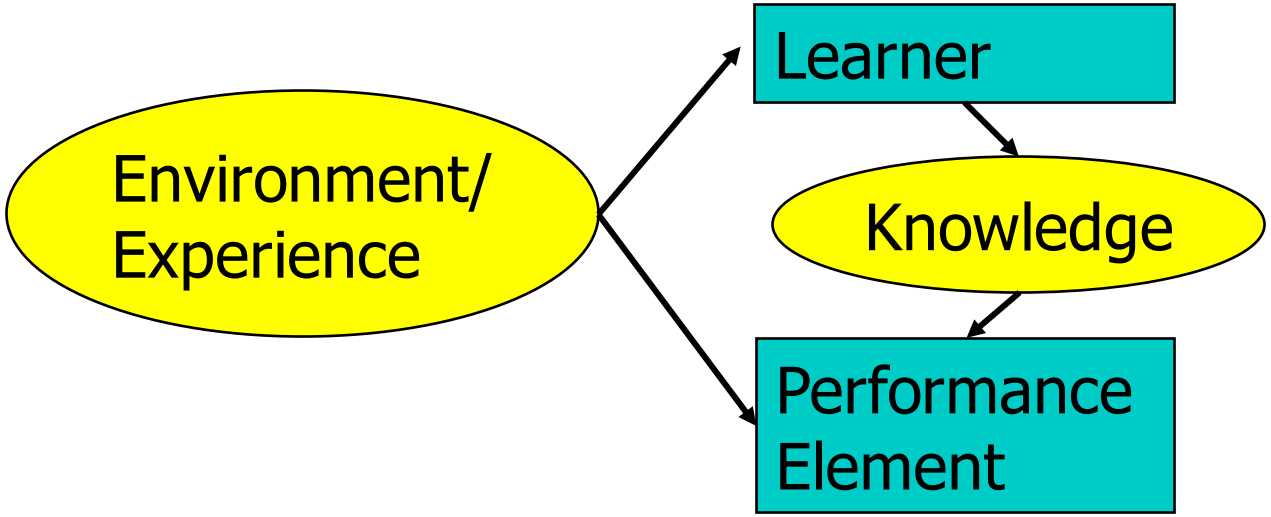
- Concept Learning
- Clustering Algorithms
- Connectionist Algorithms
- Genetic Algorithms
- Explanation-based and Transformation-based Learning
- Reinforcement and Case-based Learning
- Macro Learning
- Evaluation Functions
- Cognitive Learning Architectures
- Constructive Induction
- Discovery Systems

Languages or Tools for Machine Learning

- Python – Open source programming language adopted for machine learning.
- R – Open source software. Used for statistical computing and data analysis
- Matlab - Developed by MathWorks. Licensed version. Used for variety of applications.
- SAS – Statistical Analysis System, was developed and licensed by SAS Institute provides strong support for ML.
- Others-
  - SPSS(Statistical Package for the Social Sciences) – IBM
  - Julia – MIT(Massachusetts Institute of Technology)

Designing a Learning System

- Choose the training experience
- Choose exactly what is too be learned, i.e. the ***target function***.
- Choose how to represent the target function.
- Choose a learning algorithm to infer the target function from the experience.



Supervised Learning Classification

- Example: Cancer diagnosis

Patient ID	# of Tumors	Avg Area	Avg Density	Diagnosis
1	5	20	118	Malignant
2	3	15	130	Benign
3	7	10	52	Benign
4	2	30	100	Malignant

Training Set

- Use this training set to learn how to classify patients where diagnosis is not known:

Patient ID	# of Tumors	Avg Area	Avg Density	Diagnosis
101	4	16	95	?
102	9	22	125	?
103	1	14	80	?

Test Set

Input Data Classification

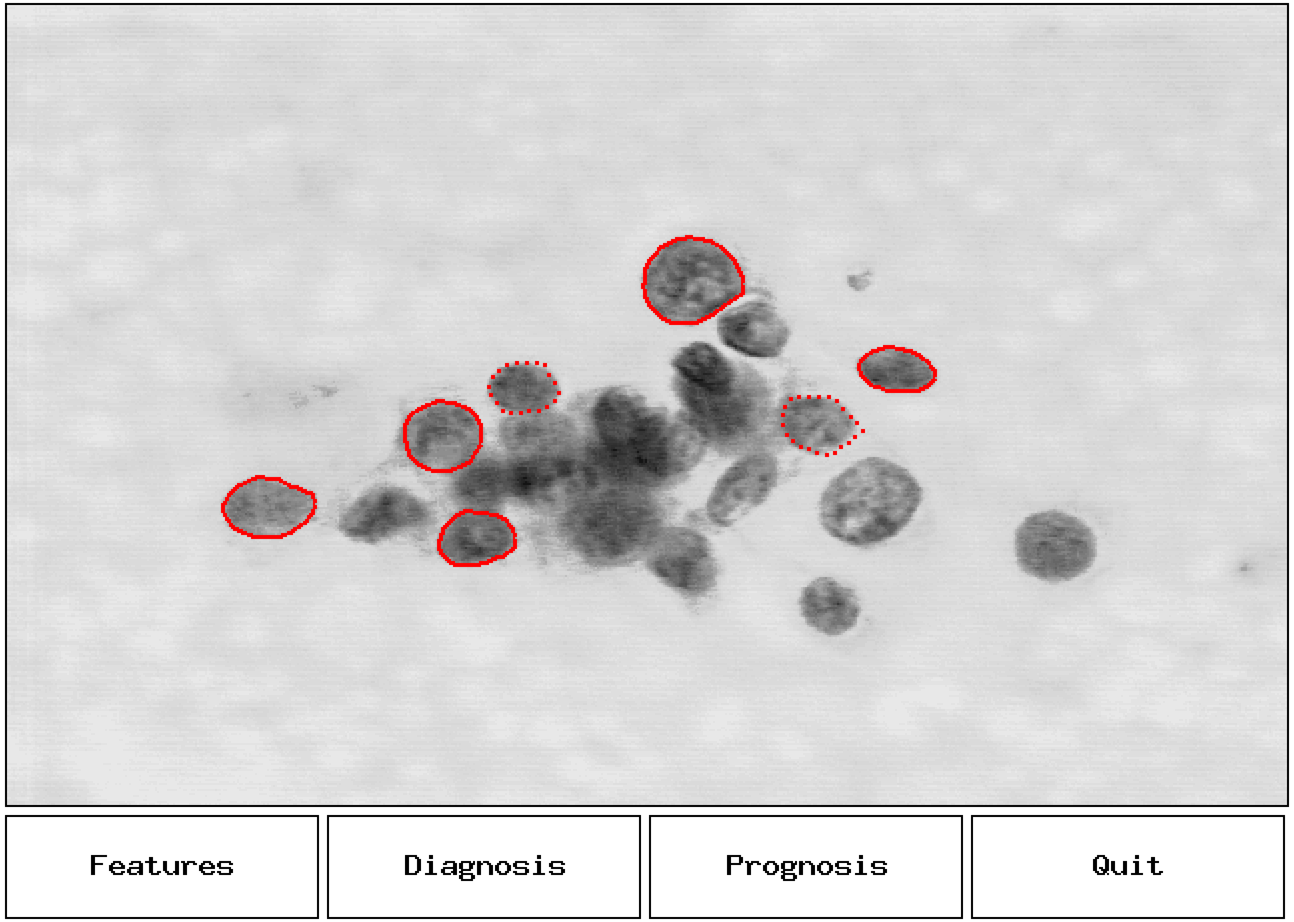
- The input data is often easily obtained, whereas the classification is not.

Classification Problem

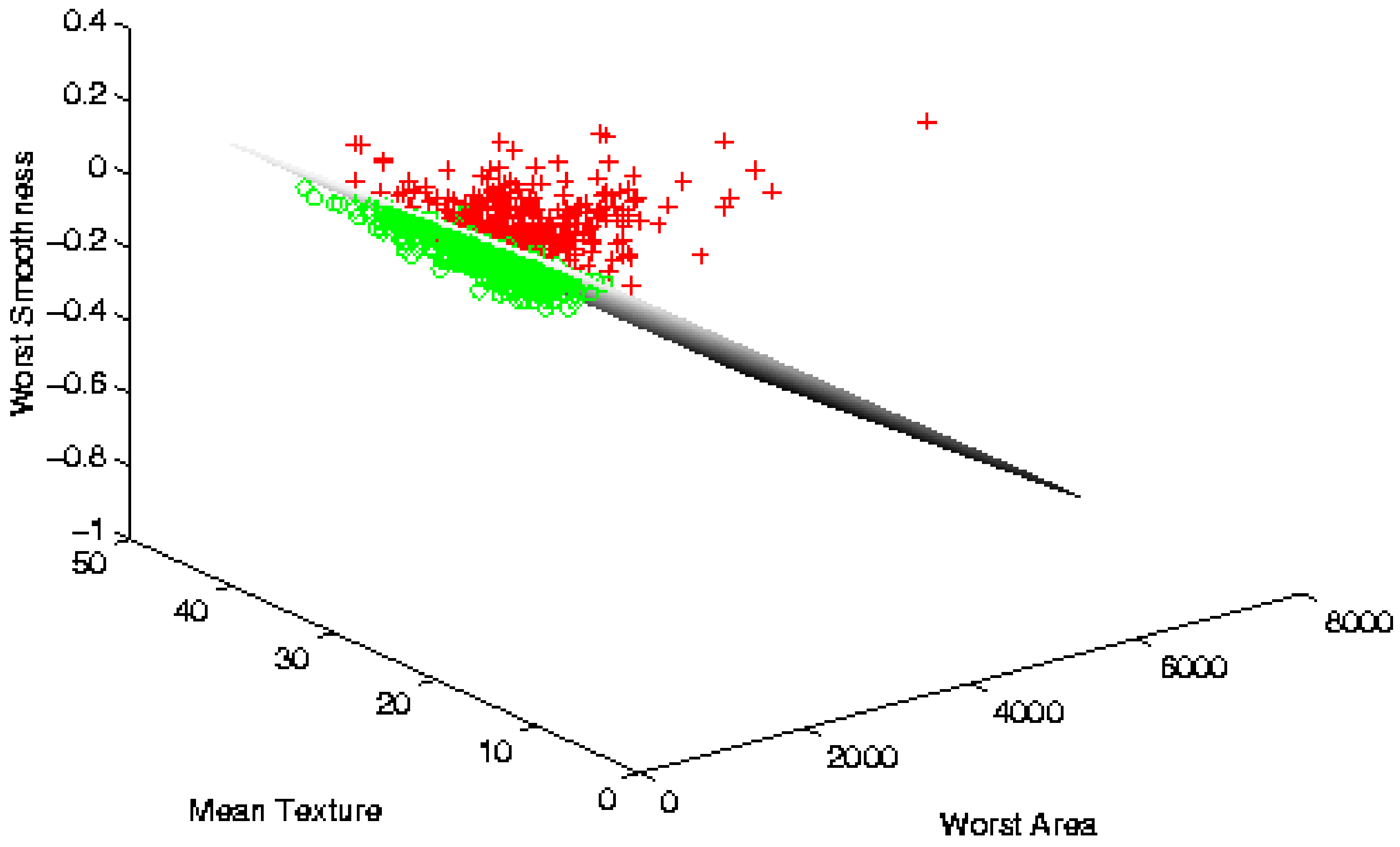
- Goal: Use training set + some learning method to produce a predictive model.
- Use this predictive model to classify new data.
- Sample applications:

Application	Input Data	Classification
Medical Diagnosis	Noninvasive tests	Results from invasive measurements
Optical Character Recognition	Scanned bitmaps	Letter A-Z
Protein Folding	Amino acid construction	Protein shape (helices, loops, sheets)
Research Paper Acceptance	Words in paper title	Paper accepted or rejected

Application: Cancer Diagnosis

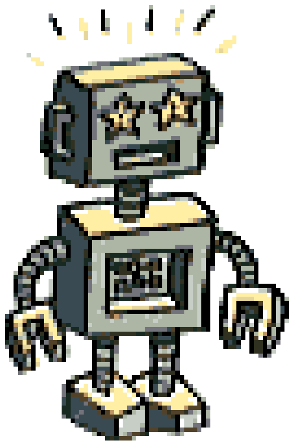


Cancer Diagnosis Separation



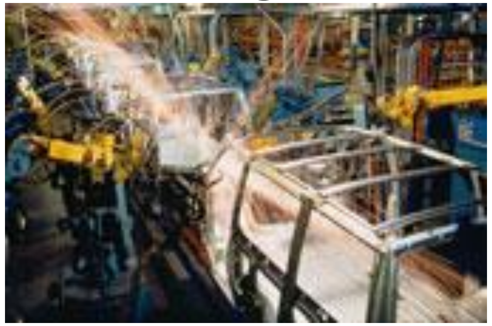
Robotics and ML

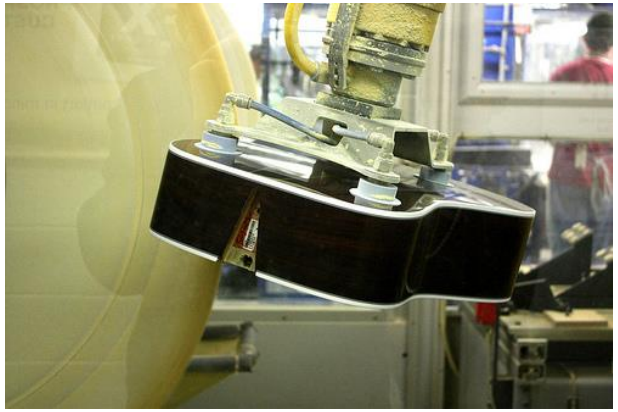
- Areas that robots are used:
  - Industrial robots
  - Military, government and space robots
  - Service robots for home, healthcare, laboratory
- Why are robots used?
  - Dangerous tasks or in hazardous environments
  - Repetitive tasks
  - High precision tasks or those requiring high quality
  - Labor savings
- Control technologies:
  - Autonomous (self-controlled), tele-operated (remote control)



Industrial Robots

- Uses for robots in manufacturing:
  - Welding
  - Painting
  - Cutting
  - Dispensing
  - Assembly
  - Polishing/Finishing
  - Material Handling
    - Packaging, Palletizing
    - Machine loading





Space Robots

- Mars Rovers – Spirit and Opportunity
  - Autonomous navigation features with human remote control and oversight

Service Robots

- Many uses…
  - Cleaning & Housekeeping
  - Humanitarian Demining
  - Rehabilitation
  - Inspection
  - Agriculture & Harvesting
  - Lawn Mowers
  - Surveillance
  - Mining Applications
  - Construction
  - Automatic Refilling
  - Fire Fighters
  - Search & Rescue

iRobot Roomba vacuum cleaner robot

Issues in Machine Learning

- What algorithms can approximate functions well and when?
  - How does the number of training examples influence accuracy
- Problem representation / feature extraction
- Intention/independent learning
- Integrating learning with systems
- What are the theoretical limits of learnability
- Transfer learning
- Continuous learning

Scaling issues in ML

- Number of
  - Inputs
  - Outputs
  - Batch vs realtime
  - Training vs testing

Machine Learning vs. Human Learning

- Some ML behavior can challenge the performance of human experts (e.g., playing chess)
- Although ML sometimes matches human learning capabilities, it is not able to learn as well as humans or in the same way that humans do
- There is no claim that machine learning can be applied in a truly creative way
- Formal theories of ML systems exist but are often lacking (why a method succeeds or fails is not clear)
- ML success is often attributed to manipulation of symbols (rather than mere numeric information)

Questions

- How does ML affect information science?
- Natural vs artificial learning – which is better?
- Is ML needed in all problems?
- What are the future directions of ML?

4.1 The vector Space R^n

Definition 1.

Let $(u_1, u_2, ..., u_n)$ be a sequence of n real numbers. The set of all such sequences is called ***n*-space (or *n*-dimensional. space)** and is denoted $\mathbf{R}^n$.

u_1 is the **first component** of $(u_1, u_2, ..., u_n)$.

u_2 is the **second component** and so on.

Example 1

- $\mathbf{R}^2$ is the collection of all sets of two ordered real numbers. For example, (0, 0) , (1, 2) and (-2, -3) are elements of $\mathbf{R}^2$.

- $\mathbf{R}^3$ is the collection of all sets of three ordered real numbers. For example, (0,0, 0) and (-1,3, 4) are elements of $\mathbf{R}^3$.

Definition 2.

Let $\mathbf{u} = (u_1, u_2, ..., u_n)$ and $\mathbf{v} = (v_1, v_2, ..., v_n)$ be two elements of $\mathbf{R}^n$.

We say that $\mathbf{u}$ and $\mathbf{v}$ are **equal** if $u_1 = v_1, \ldots, u_n = v_n$.

Thus two elements of $\mathbf{R}^n$ are equal if their **corresponding components** are equal.

Definition 3.

Let $\mathbf{u} = (u_1, u_2, ..., u_n)$ and $\mathbf{v} = (v_1, v_2, ..., v_n)$ be elements of $\mathbf{R}^n$ and let c be a scalar. **Addition** and **scalar multiplication** are performed as follows:

Addition: $\mathbf{u} + \mathbf{v} = (u_1 + v_1, ..., u_n + v_n)$

Scalar multiplication : $c\mathbf{u} = (cu_1, ..., cu_n)$

► The set $\mathbf{R}^n$ with operations of componentwise addition and scalar multiplication is an example of a **vector space,** and its elements are called **vectors.**

We shall henceforth interpret $\boldsymbol{R}^n$ to be a vector space.

(We say that $\mathbf{R}^n$ is **closed** under addition and scalar multiplication).

► In general, if **u** and **v** are vectors in the same vector space, then **u** + **v** is the diagonal of the **parallelogram** defined by **u** and **v**.

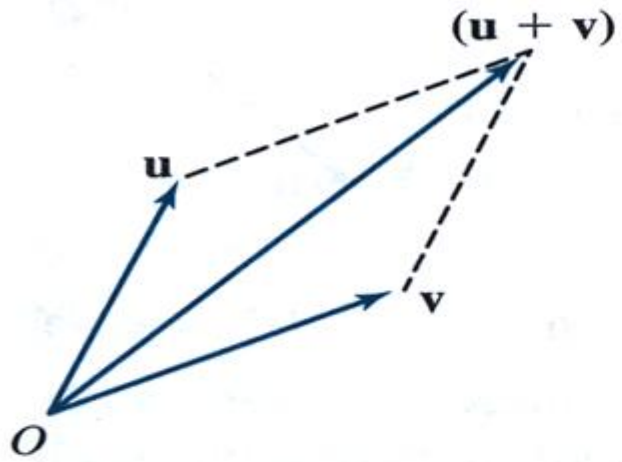


Figure 4.1

Example 2

Let **u** = (–1, 4, 3) and **v** = (–2, –3, 1) be elements of $\mathbf{R}^3$.

Find **u** + **v** and 3**u**.

Solution: **u** + **v** = (–1, 4, 3) + (– 2, –3, 1) = (-3 ,1 ,4)

3**u** = **3** (–1, 4, 3) = (-3 ,12 ,9)

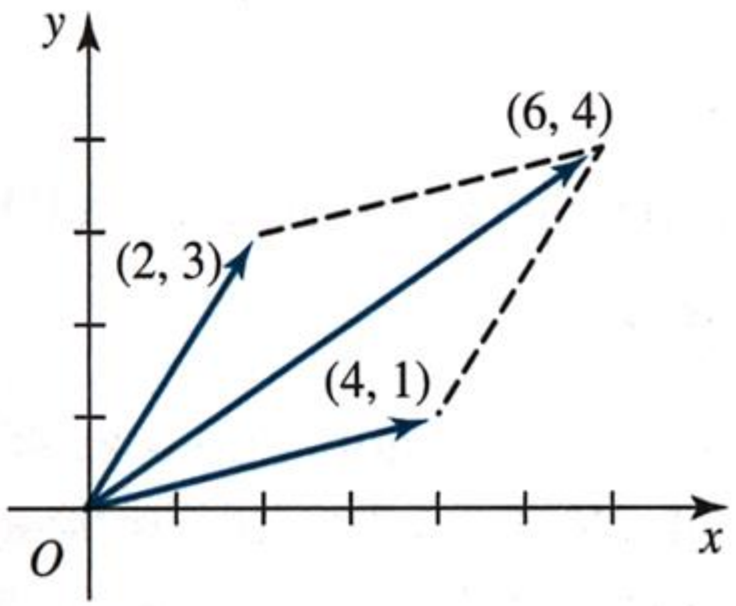
Example 3

In R^2 , consider the two elements (4, 1) and (2, 3).

Find their sum and give a geometrical interpretation of this sum.

we get (4, 1) + (2, 3) = (6, 4).

The vector (6, 4), the sum, is the diagonal of the parallelogram.

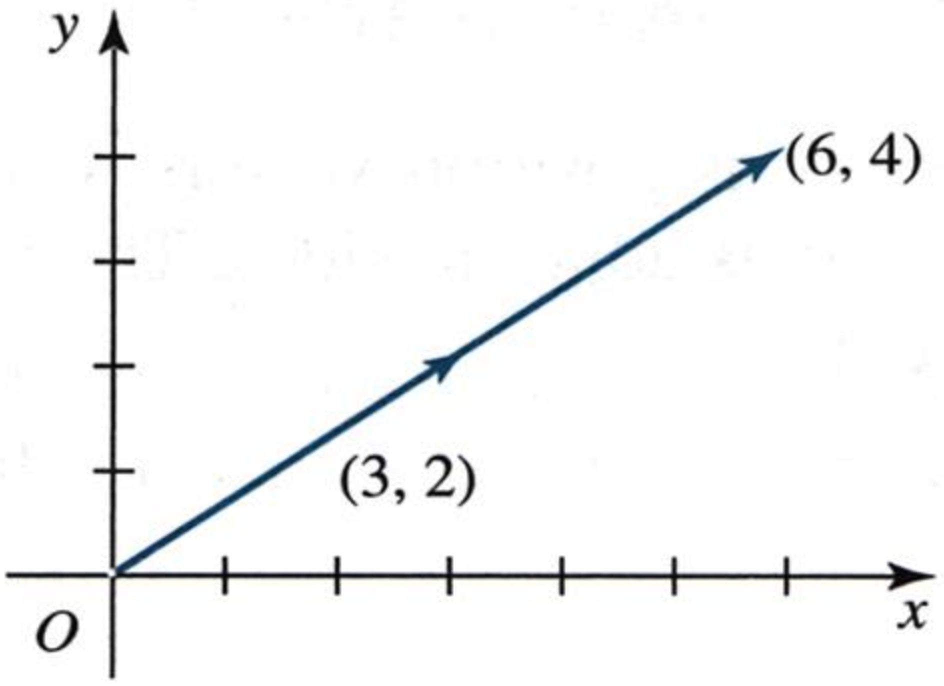


Example 4

Consider the scalar multiple of the vector (3, 2) by 2, we get

$$2(3, 2) = (6, 4)$$

Observe in Figure 4.3 that (6, 4) is a vector in the same direction as (3, 2), and 2 times it in length.

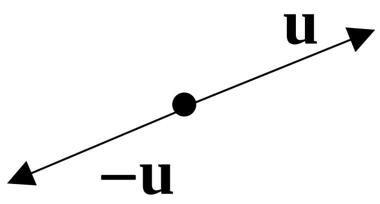


Zero Vector

The vector (0, 0, …, 0), having n zero components, is called the **zero vector** of $\mathbf{R}^n$ and is denoted $\mathbf{0}$.

Negative Vector

The vector $(-1)\mathbf{u}$ is writing $-\mathbf{u}$ and is called **the negative of u**. It is a vector having the same length (or magnitude) as $\mathbf{u}$, but lies in the opposite direction to $\mathbf{u}$.



Subtraction

Subtraction is performed on element of $\mathbf{R}^n$ by subtracting corresponding components.

Theorem 4.1

Let **u**, **v**, and **w** be vectors in $\mathbf{R}^n$ and let c and d be scalars.

(a) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

(b) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$

(c) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$

(d) $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$

(e) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$

(f) $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$

(g) $c(d\mathbf{u}) = (cd)\mathbf{u}$

(h) $1\mathbf{u} = \mathbf{u}$

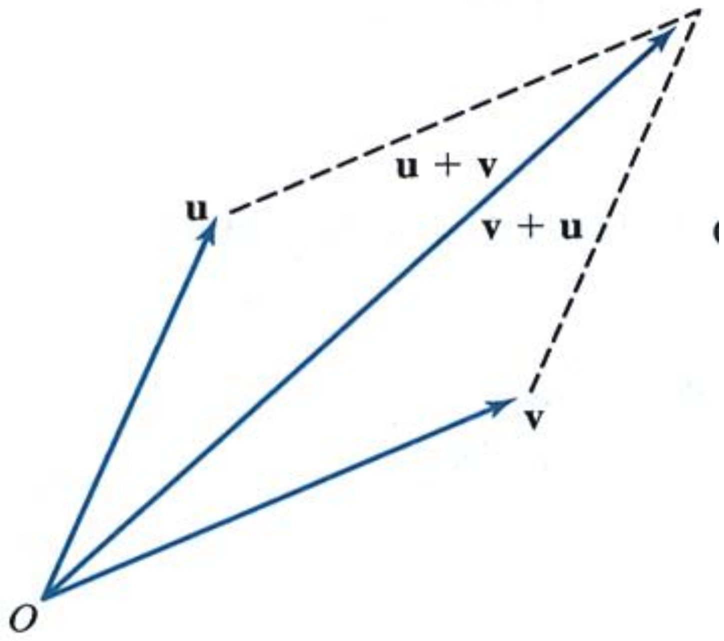


Figure 4.4

Commutativity of vector addition

$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

Example 5

Let **u** = (2, 5, –3), **v** = (–4, 1, 9), **w** = (4, 0, 2) in the vector space R^3. Determine the vector 2**u** – 3**v** + **w**.

Solution

$$\begin{aligned}
\mathbf{2u} - \mathbf{3v} + \mathbf{w} &= 2(2, 5, -3) - 3(-4, 1, 9) + (4, 0, 2) \\
&= (4, 10, -6) - (-12, 3, 27) + (4, 0, 2) \\
&= (4 + 12 + 4, 10 - 3 + 0, -6 - 27 + 2) \\
&= (20, 7, -31)
\end{aligned}$$

Column Vectors

Row vector: $\mathbf{u} = (u_1, u_2, ..., u_n)$

Column vector: $\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$

We defined addition and scalar multiplication of column vectors in $\mathbf{R}^n$ in a componentwise manner:

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} \quad \text{and} \quad c\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} cu_1 \\ \vdots \\ cu_n \end{bmatrix}$$

4.2 Dot Product, Norm, Angle, and Distance

Definition

Let $\mathbf{u} = (u_1, u_2, ..., u_n)$ and $\mathbf{v} = (v_1, v_2, ..., v_n)$ be two vectors in $\mathbf{R}^n$.
The **dot product** of **u** and **v** is denoted **u.v** and is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n$$

The dot product assigns a real number to each pair of vectors.

Example 1

Find the dot product of

u = (1, –2, 4) and **v** = (3, 0, 2)

Solution

$$\mathbf{u} \cdot \mathbf{v} = (1 \times 3) + (-2 \times 0) + (4 \times 2)$$
$$= 3 + 0 + 8$$
$$= 11$$

Properties of the Dot Product

Let **u**, **v**, and **w** be vectors in $\mathbf{R}^n$ and let c be a scalar. Then

1. **u.v = v.u**
2. **(u + v).w = u.w + v.w**
3. c**u.v** = c**(u.v)** = **u.**c**v**
4. **u.u** ≥ 0, and **u.u** = 0 if and only if **u = 0**

Proof

1. Let $\mathbf{u} = (u_1, u_2, ..., u_n)$ and $\mathbf{v} = (v_1, v_2, ..., v_n)$. We get

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= u_1v_1 + \cdots + u_nv_n \\ &= v_1u_1 + \cdots + v_nu_n \quad \text{by the commutative property of real numbers} \\ &= \mathbf{v}\cdot\mathbf{u}\end{aligned}$$

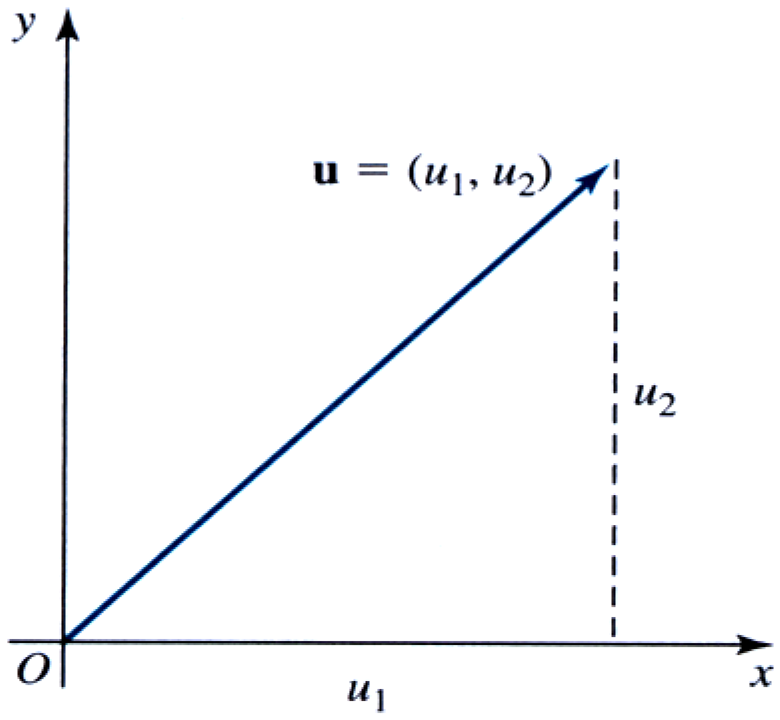
4. $\mathbf{u}\cdot\mathbf{u} = u_1u_1 + \cdots + u_nu_n = (u_1)^2 + \cdots + (u_n)^2$

$(u_1)^2 + \cdots + (u_n)^2 \geq 0$, thus $\mathbf{u}\cdot\mathbf{u} \geq 0$.

$(u_1)^2 + \cdots + (u_n)^2 = 0$, if and only if $u_1 = 0, \cdots, u_n = 0$.

Thus $\mathbf{u}\cdot\mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

Norm of a Vector in $\mathbf{R}^n$



length of **u**

Definition

The **norm** (**length** or **magnitude**) of a vector $\mathbf{u} = (u_1, \ldots, u_n)$ in $\mathbf{R}^n$ is denoted $\|\mathbf{u}\|$ and defined by

$$\|\mathbf{u}\| = \sqrt{(u_1)^2 + \cdots + (u_n)^2}$$

Note:
The norm of a vector can also be written in terms of the dot product $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$

Example 2

Find the norm of each of the vectors $\mathbf{u} = (1, 3, 5)$ of $\mathbf{R}^3$ and $\mathbf{v} = (3, 0, 1, 4)$ of $\mathbf{R}^4$.

Solution

$$\|\mathbf{u}\| = \sqrt{(1)^2 + (3)^2 + (5)^2} = \sqrt{1 + 9 + 25} = \sqrt{35}$$

$$\|\mathbf{v}\| = \sqrt{(3)^2 + (0)^2 + (1)^2 + (4)^2} = \sqrt{9 + 0 + 1 + 16} = \sqrt{26}$$

Definition

A **unit vector** is a vector whose norm is 1.

If $\mathbf{v}$ is a nonzero vector, then the vector

$$\boxed{\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}}$$

is a unit vector in the direction of $\mathbf{v}$.

This procedure of constructing a unit vector in the same direction as a given vector is called **normalizing** the vector.

Example 3

(a) Show that the vector (1, 0) is a unit vector.

(b) Find the norm of the vector (2, –1, 3). Normalize this vector.

Solution

(a) $\|(1,0)\| = \sqrt{1^2+0^2} = 1$. Thus (1, 0) is a unit vector. It can be similarly shown that (0, 1) is a unit vector in $\mathbf{R}^2$.

(b) $\|(2,-1,3)\| = \sqrt{2^2+(-1)^2+3^2} = \sqrt{14}$. The norm of (2, –1, 3) is $\sqrt{14}$.

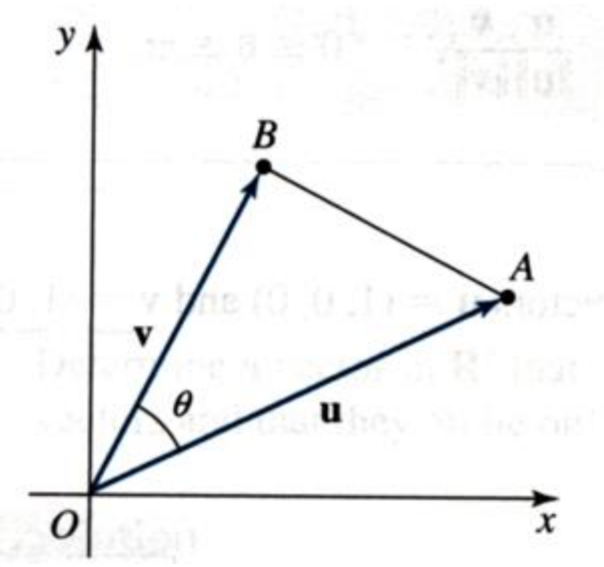
The normalized vector is

$$\frac{1}{\sqrt{14}}(2,-1,3)$$

The vector may also be written $\left(\frac{2}{\sqrt{14}}, \frac{-1}{\sqrt{14}}, \frac{3}{\sqrt{14}}\right)$.

This vector is a unit vector in the direction of (2, –1, 3).

Angle between Vectors (in R^2)



The law of cosines gives:

$$\boxed{\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}}$$

Angle between Vectors (in R ^{n})

Definition

Let $\mathbf{u}$ and $\mathbf{v}$ be two nonzero vectors in $\mathbf{R}^n$.

The **cosine of the angle θ** between these vectors is

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} \qquad 0 \le \theta \le \pi$$

Example 4

Determine the angle between the vectors $\mathbf{u} = (1, 0, 0)$ and $\mathbf{v} = (1, 0, 1)$ in $\mathbf{R}^3$.

Solution $\mathbf{u}\cdot\mathbf{v} = (1,0,0)\cdot(1,0,1) = 1$

$$\|\mathbf{u}\| = \sqrt{1^2+0^2+0^2} = 1 \qquad \|\mathbf{v}\| = \sqrt{1^2+0^2+1^2} = \sqrt{2}$$

Thus $\cos\theta = \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \dfrac{1}{\sqrt{2}}$, the angle between $\mathbf{u}$ and $\mathbf{v}$ is 45°.

Orthogonal Vectors

Definition

Two nonzero vectors are **orthogonal** if the angle between them is a right angle .

Theorem 4.2

Two nonzero vectors **u** and **v** are orthogonal *if and only if* $\mathbf{u}.\mathbf{v} = \mathbf{0}$.

Proof

$$\mathbf{u}, \mathbf{v} \text{ are orthogonal} \quad \Leftrightarrow \quad \cos\theta = 0 \quad \Leftrightarrow \quad \mathbf{u} \cdot \mathbf{v} = 0$$

Example 5

Show that the following pairs of vectors are orthogonal.

(a) (1, 0) and (0, 1).

(b) (2, –3, 1) and (1, 2, 4).

Solution

(a) $(1, 0).(0, 1) = (1 \times 0) + (0 \times 1) = 0.$

The vectors are orthogonal.

(b) $(2, -3, 1).(1, 2, 4) = (2 \times 1) + (-3 \times 2) + (1 \times 4) = 2 - 6 + 4 = 0.$

The vectors are orthogonal.

Note

- (1, 0), (0,1) are <u>orthogonal</u> <u>unit</u> vectors in $\mathbf{R}^2$.
- (1, 0, 0), (0, 1, 0), (0, 0, 1) are orthogonal unit vectors in $\mathbf{R}^3$.
- (1, 0, …, 0), (0, 1, 0, …, 0), …, (0, …, 0, 1) are orthogonal unit vectors in $\mathbf{R}^n$.

Example 6

Determine a vector in $\mathbf{R}^2$ that is orthogonal to (3, –1). Show that there are many such vectors and that they all lie on a line.

Solution

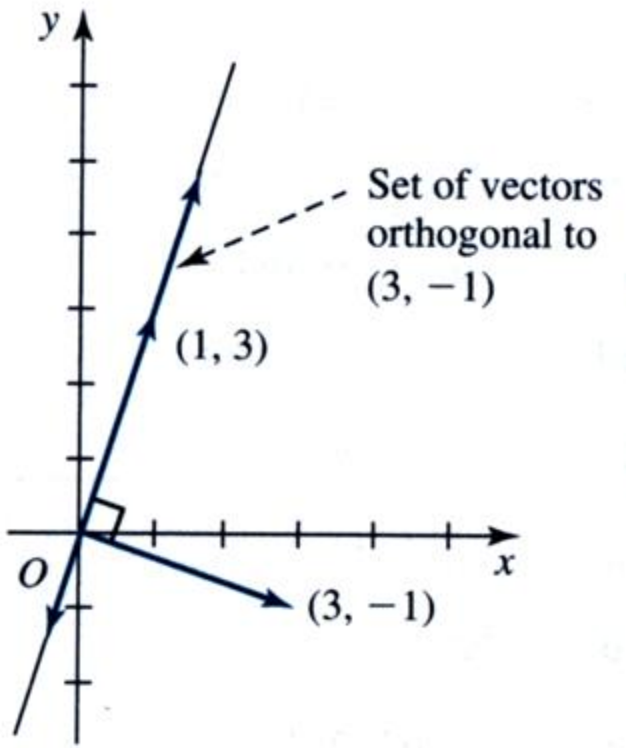
Let the vector (a, b) be orthogonal to (3, –1)

We get
$$(a,b)\cdot(3,-1)=0$$
$$(a\times 3)+(b\times(-1))=0$$
$$3a-b=0$$
$$b=3a$$

Thus any vector of the form $(a, 3a)$ is orthogonal to the vector (3, –1).

Any vector of this form can be written
$$a(1, 3)$$

The set of all such vectors lie on the line defined by the vector (1, 3).



Theorem 4.3

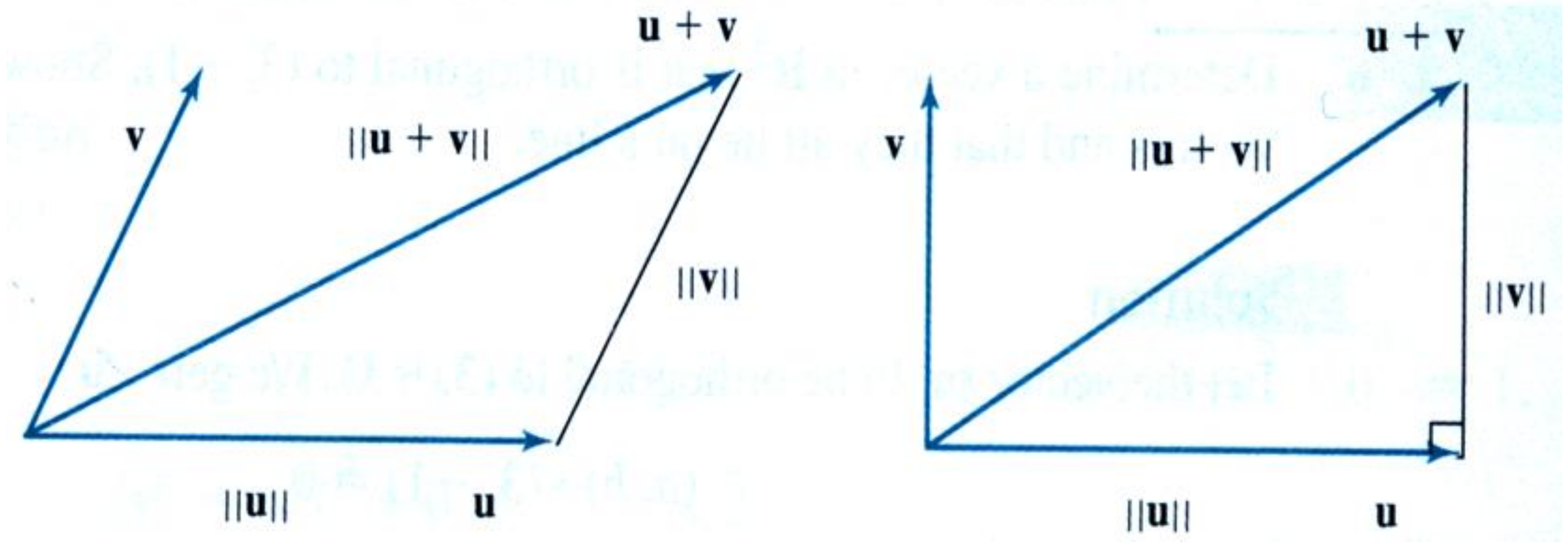
Let **u** and **v** be vectors in $\mathbf{R}^n$.

(a) **Triangle Inequality:**

$$\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|.$$

(a) **Pythagorean theorem :**

$$\text{If } \mathbf{u}.\mathbf{v} = 0 \text{ then } \|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$



Distance between Points

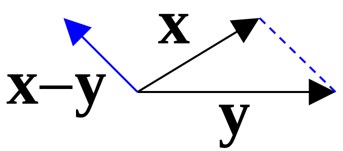
Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ be two points in $\mathbf{R}^n$.

The **distance** between **x** and **y** is denoted $d(\mathbf{x}, \mathbf{y})$ and is defined by

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$

Note: We can also write this distance as follows.

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$



Note: *It is clear that* $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (*the symmetric property*)

Example 7. Determine the distance between the points $\mathbf{x} = (1, -2, 3, 0)$ and $\mathbf{y} = (4, 0, -3, 5)$ in $\mathbf{R}^4$.

Solution

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(1-4)^2 + (-2-0)^2 + (3+3)^2 + (0-5)^2}$$

$$= \sqrt{9 + 4 + 36 + 25}$$

$$= \sqrt{74}$$

Homework

- Exercise set 1.5 pages 47 to 48:
  3, 7, 8, 9, 11, 13, 16, 17, 26.

Exercise 36

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbf{R}^n$.
Prove that $\|\mathbf{u}\| = \|\mathbf{v}\|$ *if and only if* $\mathbf{u} + \mathbf{v}$ and $\mathbf{u} - \mathbf{v}$ are orthogonal.

4.3 General Vector Spaces

Our aim in this section will be to focus on the algebraic properties of $\mathbf{R}^n$.

Definition

A **vector space** is a set *V* of elements called **vectors,** having operations of ***addition*** and ***scalar multiplication*** defined on it that satisfy the following conditions.

Let u, v, and w be arbitrary elements of V, and c and d are scalars.

- **Closure Axioms**

1. The sum $\mathbf{u} + \mathbf{v}$ exists and is an element of *V*. (*V* is closed under addition.)
2. $c\mathbf{u}$ is an element of *V*. (*V* is closed under scalar multiplication.)

Example 1

(1) V={ …, –3, –1, 1, 3, 5, 7, … }
V is not closed under addition because $1+3=4 \notin V$.

(2) Z={ …, –2, –1, 0, 1, 2, 3, 4, … }
Z is closed under addition because
for any $a, b \in Z$, $a + b \in Z$.
Z is not closed under scalar multiplication because
½ is a scalar, for any odd $a \in Z$, $(½)a \notin Z$.

Definition of Vector Space (continued)

- **Addition Axioms**

3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (commutative property)
4. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ (associative property)
5. There exists an element of V, called the **zero vector**, denoted $\mathbf{0}$, such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
6. For every element $\mathbf{u}$ of V there exists an element called the **negative** of $\mathbf{u}$, denoted $-\mathbf{u}$, such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.

- **Scalar Multiplication Axioms**

7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$
10. $1\mathbf{u} = \mathbf{u}$

A Vector Space in R^3

Let $W = \{ a(1,0,1) \mid a \in R \}$. Prove that W is a vector space.

Proof

Let $\mathbf{u} = a(1,0,1)$ and $\mathbf{v} = b(1,0,1) \in W$, for some $a, b \in \mathbf{R}$.

Axiom 1: $\mathbf{u} + \mathbf{v} = a(1,0,1) + b(1,0,1) = (a+b)(1,0,1)$

$\therefore \mathbf{u} + \mathbf{v} \in W$. Thus W is <u>closed under addition</u>.

Axiom 2: $c\mathbf{u} = ca(1,0,1) \in W$.

Thus W is <u>closed under scalar multiplication</u>.

Axiom 5: Let $\mathbf{0} = (0, 0, 0) = 0(1,0,1)$,

then $\mathbf{0} \in W$ and $\mathbf{0}+\mathbf{u} = \mathbf{u}+\mathbf{0} = \mathbf{u}$ for any $\mathbf{u} \in W$.

Axiom 6: For any $\mathbf{u} = a(1,0,1) \in W$. Let $-\mathbf{u} = -a(1,0,1)$,

then $-\mathbf{u} \in W$ and $(-\mathbf{u})+\mathbf{u} = \mathbf{0}$.

Axiom 3,4 and 7~10: trivial

Vector Spaces of Matrices ($\mathbf{M}_{mn}$)

Let $M_{22} = \{ \begin{bmatrix} p & q \\ r & s \end{bmatrix} \mid p, q, r, s \in R \}$. Prove that $\boldsymbol{M}_{22}$ is a vector space.

Proof

Let $\mathbf{u} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} \in M_{22}$.

Axiom 1:

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$$

$\mathbf{u} + \mathbf{v}$ is a 2×2 matrix. Thus $\boldsymbol{M}_{22}$ is closed under addition.

► ***Question***: Prove Axiom 2 and Axiom 7 .

Axiom 3 and 4:

From our previous discussions we know that 2×2 matrices are commutative and associative under addition (Theorem 2.2).

Axiom 5:

The 2×2 zero matrix is $\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, since

$$\mathbf{u} + \mathbf{0} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \mathbf{u}$$

Axiom 6:

If $\mathbf{u} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $-\mathbf{u} = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}$, since

$$\mathbf{u} + (-\mathbf{u}) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix} = \begin{bmatrix} a-a & b-b \\ c-c & d-d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0}$$

In general: The set of $m \times n$ matrices, $\boldsymbol{M}_{mn}$, is a vector space.

Is the set $W = \{ \begin{bmatrix} p & q \\ r & s \end{bmatrix} \mid p, q, r, s > 0\}$ a vector space?

Vector Spaces of Functions

Prove that $\boldsymbol{F} = \{ f \mid f: R \to R \}$ is a vector space.

Let $f, g \in \boldsymbol{F}$, $c \in R$.

For example: $f: R \to R$, $f(x)=2x$,
$g: R \to R$, $g(x)=x^2+1$.

Axiom 1:

$f + g$ is defined by $(f+g)(x) = f(x) + g(x)$.

$\Rightarrow f + g : R \to R$

$\Rightarrow f + g \in \boldsymbol{F}$. Thus $\boldsymbol{F}$ is closed under addition.

Axiom 2:

cf is defined by $(cf)(x) = c \cdot f(x)$.

$\Rightarrow cf : R \to R$

$\Rightarrow cf \in \boldsymbol{F}$. Thus $\boldsymbol{F}$ is closed under scalar multiplication.

Vector Spaces of Functions (continued)

Axiom 5:

Let $\mathbf{0}$ be the function such that $\mathbf{0}(x) = 0$ for every $x \in R$.

$\mathbf{0}$ is called the **zero function**.

We get $(f + \mathbf{0})(x) = f(x) + \mathbf{0}(x) = f(x) + 0 = f(x)$ for every $x \in R$.

Thus $f + \mathbf{0} = f$. ($\mathbf{0}$ is the **zero vector**.)

Axiom 6:

Let the function $-f$ defined by $(-f)(x) = -f(x)$.

$$\begin{aligned}[f+(-f)](x) &= f(x)+(-f)(x) \\ &= f(x)-[f(x)] \\ &= 0 \\ &= \mathbf{0}(x)\end{aligned}$$

Thus $[f + (-f)] = \mathbf{0}$, $-f$ is the negative of f.

Is the set $\mathcal{F} = \{ f \mid f(x) = ax^2 + bx + c$ for some $a,b,c \in R\}$ *a vector space*?

Theorem 4.4 (useful properties)

Let V be a vector space, $\mathbf{v}$ a vector in V, $\mathbf{0}$ the **zero vector** of V, c a scalar, and 0 the zero scalar. Then

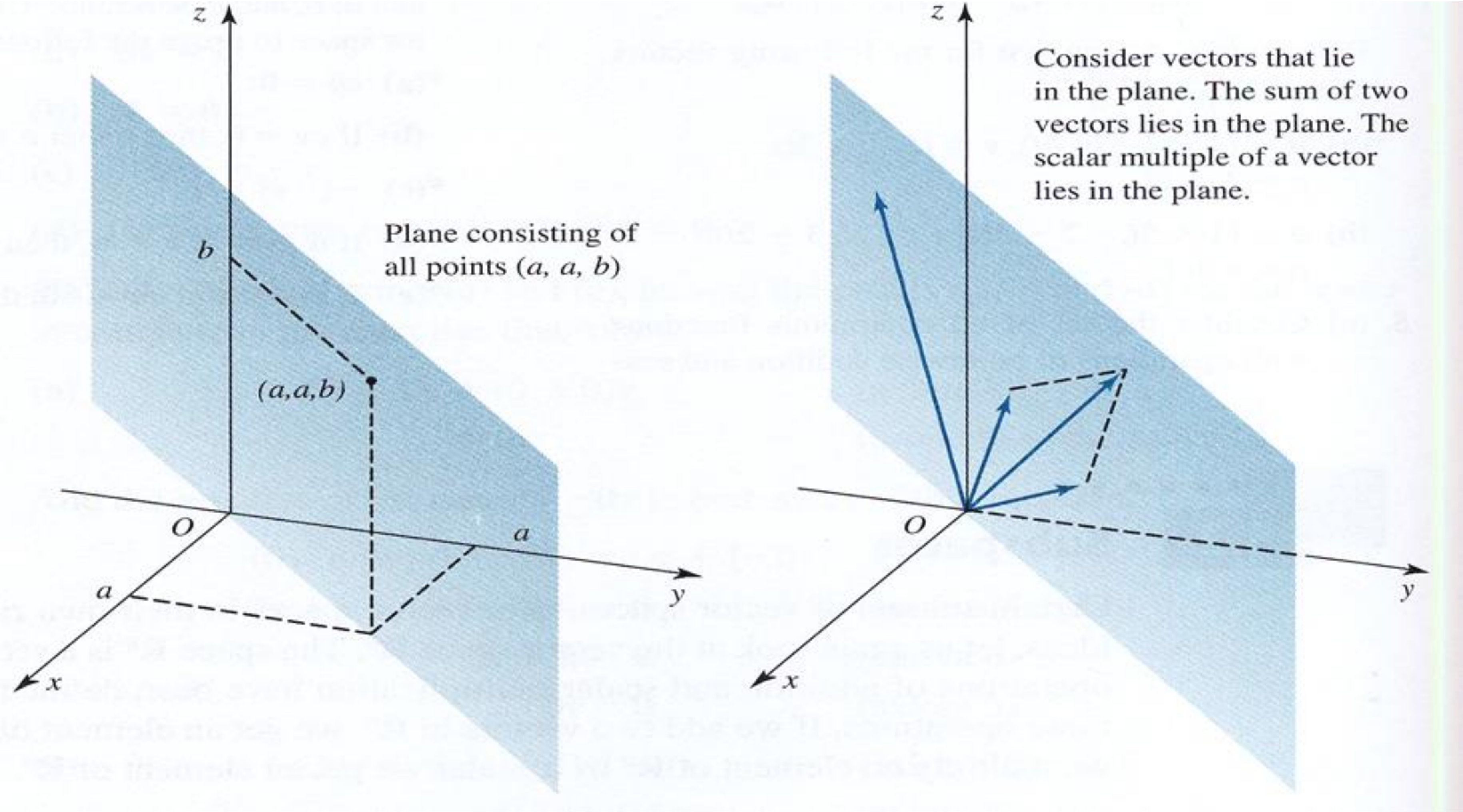
(a) $0\mathbf{v} = \mathbf{0}$

(b) $c\mathbf{0} = \mathbf{0}$

(c) $(-1)\mathbf{v} = -\mathbf{v}$

(d) If $c\mathbf{v} = \mathbf{0}$, then either $c = 0$ or $\mathbf{v} = \mathbf{0}$.

4.4 Subspaces



Note:

► In general, a subset of a vector space may or may not satisfy the closure axioms.

► However, any subset that is closed under both of these operations satisfies all the other vector space properties.

Definition

Let V be a vector space and U be a *nonempty* subset of V.

U is said to be a **subspace** of V if it is closed under addition and under scalar multiplication.

Example 1

Let U be the subset of $\mathbf{R}^3$ consisting of all vectors of the form $(a, 0, 0)$ (with zeros as second and third components and $a \in \mathbf{R}$),
i.e., $U = \{(a, 0, 0) \in \mathbf{R}^3\}$.
Show that U is a subspace of $\mathbf{R}^3$.

Solution

Let $(a, 0, 0), (b, 0, 0) \in U$, and let $k \in \mathbf{R}$.
We get

$$(a, 0, 0) + (b, 0, 0) = (\mathrm{a} + b, 0, 0) \in U$$

$$k(a, 0, 0) = (k\, a, 0, 0) \in U$$

The sum and scalar product are in U.
Thus U is a subspace of $\mathbf{R}^3$. #

Geometrically, *U is the set of vectors that lie on the x-axis.*

Example 2

Let V be the set of vectors of of $\mathbf{R}^3$ of the form (a, a^2, b), namely

$V = \{(a, a^2, b) \in \mathbf{R}^3\}$.

Show that V is not a subspace of $\mathbf{R}^3$.

Solution

Let $(a, a^2, b), (c, c^2, d) \in V$.

$$\begin{aligned}(a, a^2, b) + (c, c^2, d) &= (a + c, a^2 + c^2, b + d) \\ &\neq (a + c, (a + c)^2, b + d),\end{aligned}$$

since $a^2 + c^2 \neq (a + c)^2$.

Thus $(a, a^2, b) + (c, c^2, d) \notin V$.

V is <u>not closed</u> under addition.

V is <u>not a subspace</u>.

Example 3

Prove that the set W of 2×2 diagonal matrices is a subspace of the vector space $\boldsymbol{M}_{22}$ of 2×2 matrices.

Solution

(+) Let $\mathbf{u} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix} \in W.$

We get $\mathbf{u} + \mathbf{v} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix} = \begin{bmatrix} a+p & 0 \\ 0 & b+q \end{bmatrix}$

$\Rightarrow \mathbf{u} + \mathbf{v} \in W.$

$\Rightarrow W$ is <u>*closed under addition*</u>.

$(\cdot)$ Let $c \in \mathbf{R}$. We get $c\mathbf{u} = c\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = \begin{bmatrix} ca & 0 \\ 0 & cb \end{bmatrix}$

$\Rightarrow c\mathbf{u} \in W.$

$\Rightarrow W$ is <u>*closed under scalar multiplication*</u>.

$\Rightarrow W$ is a subspace of $\boldsymbol{M}_{22}$.

The vector space of polynomials (P_n)

Example 4. ***Let P_n denoted the set of real polynomial functions of degree $\leq$ n. Prove that P_n is a vector space if addition and scalar multiplication are defined on polynomials in a pointwise manner.***

Solution

Let f and $g \in \boldsymbol{P_n}$, where

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \text{ and}$$

$$g(x) = b_n x^n + b_{n-1} x^{n-1} + \ldots + b_1 x + b_0$$

► (+) $(f+g)(x)$

$$= f(x) + g(x)$$

$$= [a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0] + [b_n x^n + b_{n-1} x^{n-1} + \ldots + b_1 x + b_0]$$

$$= (a_n + b_n) x^n + (a_{n-1} + b_{n-1}) x^{n-1} + \ldots + (a_1 + b_1) x + (a_0 + b_0)$$

$(f+g)(x)$ is a polynomial of degree $\leq n$.

Thus $f + g \in \boldsymbol{P_n}$.

Then $\boldsymbol{P_n}$ is <u>closed under addition</u>.

▶ $(\cdot)$ Let $c \in \mathbf{R}$, $(cf)(x) = c[f(x)]$

$$= c[a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0]$$

$$= ca_n x^n + ca_{n-1}x^{n-1} + \ldots + ca_1 x + ca_0$$

$(cf)(x)$ is a polynomial of degree $\le n$.

So $cf \in \boldsymbol{P_n}$.

Then $\boldsymbol{P_n}$ is closed under scalar multiplication.

In conclusion : By $(+)$ and $(\cdot)$, $\boldsymbol{P_n}$ is a subspace of the vector space $\boldsymbol{F}$ of functions.

Therefore $\boldsymbol{P_n}$ is itself a vector space.

Theorem 4.5 (Very important condition)

Let U be a subspace of a vector space V.

U contains the zero vector of V.

Note. Let $\mathbf{0}$ be the zero vector of V.

If $\mathbf{0} \notin U \Rightarrow U$ is not a subspace of V.

If $\mathbf{0} \in U \Rightarrow (+)(\cdot)$ hold $\Rightarrow U$ is a subspace of V.

$(+)(\cdot)$ failed $\Rightarrow U$ is not a subspace of V.

Caution. *This condition is necessary but not sufficient.*

(See, for instance, *Example 2* above and *Example 5* below)

Example 5

Let W be the set of vectors of the form $(a, a, a+2)$. Show that W is not a subspace of $\mathbf{R}^3$.

Solution

If $(a, a, a+2) = (0, 0, 0)$, then $a = 0$ and $a + 2 = 0$.

This system is inconsistent it has no solution.

Thus $(0, 0, 0) \notin W$. (The necessary condition does not hold)

$\Rightarrow W$ is not a subspace of $\mathbf{R}^3$.

Homework

- Exercise set 4.1, pages 207-208:
  19, 21, 23, 25, 27, 29, 31, 33.
- Exercise set 1.3, page 32: 11, 12, 13, 15.

Exercise

Let $\boldsymbol{F} = \{ f \mid f\colon \mathbf{R} \to \mathbf{R} \}$ the vector space of functions on $\mathbf{R}$.
Which of the following are subspaces of $\boldsymbol{F}$?

(a) $\boldsymbol{W}_1 = \{ f \mid f\colon \mathbf{R} \to \mathbf{R},\ f(0)=0 \}$.

(b) $\boldsymbol{W}_2 = \{ f \mid f\colon \mathbf{R} \to \mathbf{R},\ f(0)=3 \}$.

(c) $\boldsymbol{W}_3 = \{ f \mid f\colon \mathbf{R} \to \mathbf{R}$, for some $c \in \mathbf{R}$, $f(x)=c$ for every $x\}$.

4.5 Linear Combinations of Vectors

$W=\{(a, a, b) \mid a,b \in \mathbf{R}\} \subseteq \mathbf{R}^3$

$(a, a, b) = a\,(1,1,0) + b\,(0,0,1)$

$\therefore W$ is generated by (1,1,0) and (0,0,1).

e.g., (2, 2, 3) = 2 (1,1,0) + 3 (0,0,1)

(–1, –1, 7) = –1 (1,1,0) + 7 (0,0,1).

Definition

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ be vectors in a vector space V.

We say that $\mathbf{v}$, a vector of V, is a **linear combination** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$, if there exist scalars $c_1, c_2, \ldots, c_m$ such that $\mathbf{v}$ can be written $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_m\mathbf{v}_m$.

Example 1

The vector (5, 4, 2) is a linear combination of the vectors (1, 2, 0), (3, 1, 4), and (1, 0, 3), since it can be written

$$(5, 4, 2) = (1, 2, 0) + 2(3, 1, 4) - 2(1, 0, 3)$$

Example 2

Determine whether or not the vector (–1, 1, 5) is a linear combination of the vectors (1, 2, 3), (0, 1, 4), and (2, 3, 6).

Solution

Suppose $c_1(1,2,3)+c_2(0,1,4)+c_3(2,3,6)=(-1,1,5)$

$$(c_1, 2c_1, 3c_1)+(0, c_2, 4c_2)+(2c_3, 3c_3, 6c_3)=(-1,1,5)$$

$$(c_1+2c_3, 2c_1+c_2+3c_3, 3c_1+4c_2+6c_3)=(-1,1,5)$$

$$\Rightarrow \begin{cases} c_1 \qquad\quad + 2c_3 = -1 \\ 2c_1 + \ c_2 + 3c_3 = \ 1 \\ 3c_1 + 4c_2 + 6c_3 = \ 5 \end{cases} \Rightarrow c_1 = 1, c_2 = 2, c_3 = -1$$

Thus (–1, 1, 5) is a linear combination of (1, 2, 3), (0, 1, 4), and (2, 3, 6), where $(-1,1,5)=(1,2,3)+2(0,1,4)-1(2,3,6)$.

Example 3

Express the vector (4, 5, 5) as a linear combination of the vectors (1, 2, 3), (–1, 1, 4), and (3, 3, 2).

Solution

Suppose $c_1(1,2,3)+c_2(-1,1,4)+c_3(3,3,2)=(4,5,5)$

$$(c_1, 2c_1, 3c_1)+(-c_2, c_2, 4c_2)+(3c_3, 3c_3, 2c_3)=(4,5,5)$$
$$(c_1-c_2+3c_3, 2c_1+c_2+3c_3, 3c_1+4c_2+2c_3)=(4,5,5)$$
$$\Rightarrow \begin{cases} c_1 - \ \ c_2+3c_3=4 \\ 2c_1+\ \ c_2+3c_3=5 \\ 3c_1+4c_2+2c_3=5 \end{cases} \Rightarrow c_1=-2r+3, c_2=r-1, c_3=r$$

Thus (4, 5, 5) can be expressed in many ways as a linear combination of (1, 2, 3), (–1, 1, 4), and (3, 3, 2):

$$(4,5,5)=(-2r+3)(1,2,3)+(r-1)(-1,1,4)+r(2,3,6)$$

Example 4

Show that the vector (3, –4, –6) cannot be expressed as a linear combination of the vectors (1, 2, 3), (–1, –1, –2), and (1, 4, 5).

Solution

Suppose

$$c_1(1,2,3)+c_2(-1,\ -1,-2)+c_3(1,4,5)=(3,-4,-6)$$

$\Rightarrow$

$$\begin{cases} c_1 - \ c_2 + \ c_3 = \ 3 \\ 2c_1 - \ c_2 + 4c_3 = -4 \\ 3c_1 - 2c_2 + 5c_3 = -6 \end{cases}$$

This system has no solution.

Thus (3, –4, –6) is not a linear combination of the vectors (1, 2, 3), (–1, –1, –2), and (1, 4, 5).

Example 5

Determine whether the matrix $\begin{bmatrix} -1 & 7 \\ 8 & -1 \end{bmatrix}$ is a linear combination of the matrices $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$, $\begin{bmatrix} 2 & -3 \\ 0 & 2 \end{bmatrix}$, and $\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$ in the vector space $\boldsymbol{M}_{22}$ of 2×2 matrices.

Solution

Suppose

$$c_1\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} + c_2\begin{bmatrix} 2 & -3 \\ 0 & 2 \end{bmatrix} + c_3\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 7 \\ 8 & -1 \end{bmatrix}$$

Then

$$\begin{bmatrix} c_1 + 2c_2 & -3c_2 + c_3 \\ 2c_1 + 2c_3 & c_1 + 2c_2 \end{bmatrix} = \begin{bmatrix} -1 & 7 \\ 8 & -1 \end{bmatrix}$$

$$\begin{cases} c_1 + 2c_2 = -1 \\ -3c_2 + c_3 = 7 \\ 2c_1 + 2c_3 = 8 \\ c_1 + 2c_2 = -1 \end{cases}$$

This system has the unique solution $c_1 = 3$, $c_2 = -2$, $c_3 = 1$. Therefore

$$\begin{bmatrix} -1 & 7 \\ 8 & -1 \end{bmatrix} = 3\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} - 2\begin{bmatrix} 2 & -3 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$$

Example 6

Determine whether the function $f(x) = x^2 + 10x - 7$ is a linear combination of the functions $g(x) = x^2 + 3x - 1$ and $h(x) = 2x^2 - x + 4.$

Solution

Suppose $c_1 g + c_2 h = f.$

Then

$$c_1(x^2 + 3x - 1) + c_2(2x^2 - x + 4) = x^2 + 10x - 7$$

$$(c_1 + 2c_2)x^2 + (3c_1 - c_2)x - c_1 + 4c_2 = x^2 + 10x - 7$$

$$\Rightarrow \begin{cases} c_1 + 2c_2 = 1 \\ 3c_1 - c_2 = 10 \\ -c_1 + 4c_2 = -7 \end{cases} \quad \Rightarrow c_1 = 3,\ c_2 = -1 \quad \Rightarrow f = 3g - h.$$

Spanning Sets

Definition

The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are said to **span** a vector space if every vector in the space can be expressed as a *linear combination* of these vectors.

In this case $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ is called a **spanning set**.

Example 7

Show that the vectors (1, 2, 0), (0, 1, –1), and (1, 1, 2) span $\mathbf{R}^3$.

Solution

Let (x, y, z) be an arbitrary element of $\mathbf{R}^3$.

Suppose $(x, y, z) = c_1(1, 2, 0) + c_2(0, 1, -1) + c_3(1, 1, 2)$

$$\Rightarrow \quad (x, y, z) = (c_1 + c_3, 2c_1 + c_2 + c_3, -c_2 + 2c_3)$$

$$\Rightarrow \begin{cases} c_1 \qquad + \quad c_3 = x \\ 2c_1 + c_2 + \quad c_3 = y \\ \qquad -c_2 + 2c_3 = z \end{cases} \quad \Rightarrow \quad \begin{cases} c_1 = 3x - y - z \\ c_2 = -4x + 2y + z \\ c_3 = -2x + y + z \end{cases}$$

$$\Rightarrow (x, y, z) = (3x - y - z)(1,2,0) + (-4x + 2y + z)(0, 1, -1) + (-2x + y + z)(1, 1, 2)$$

$\Rightarrow$The vectors (1, 2, 0), (0, 1, –1), and (1, 1, 2) span $\mathbf{R}^3$.

Example 8

Show that the following matrices span the vector space M_{22} of 2×2 matrices.

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

Solution

Let $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{22}$ (an arbitrary element).

We can express this matrix as follows:

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

proving the result.

Theorem 4.6

Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be vectors in a vector space V. Let U be the set consisting of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_m$.
Then U is a <u>subspace</u> of V spanned by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$.
U is said to be the vector space **generated** by $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Proof

(+) Let $\mathbf{u}_1 = a_1\mathbf{v}_1 + \ldots + a_m\mathbf{v}_m$ and $\mathbf{u}_2 = b_1\mathbf{v}_1 + \ldots + b_m\mathbf{v}_m \in U$.

Then $\mathbf{u}_1 + \mathbf{u}_2 = (a_1\mathbf{v}_1 + \ldots + a_m\mathbf{v}_m) + (b_1\mathbf{v}_1 + \ldots + b_m\mathbf{v}_m)$
$= (a_1 + b_1)\,\mathbf{v}_1 + \ldots + (a_m + b_m)\,\mathbf{v}_m$

$\Rightarrow \mathbf{u}_1 + \mathbf{u}_2$ is a <u>linear combination</u> of $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

$\Rightarrow \mathbf{u}_1 + \mathbf{u}_2 \in U$.

$\Rightarrow U$ is <u>closed under vector addition</u>.

(·)Let $c \in \mathbf{R}$. Then

$$c\mathbf{u}_1 = c(a_1\mathbf{v}_1 + \ldots + a_m\mathbf{v}_m)$$
$$= ca_1\mathbf{v}_1 + \ldots + ca_m\mathbf{v}_m$$

$\Rightarrow c\mathbf{u}_1$ is <u>a linear combination</u> of $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

$\Rightarrow c\mathbf{u}_1 \in U$.

$\Rightarrow U$ is <u>closed under scalar multiplication</u>.

Thus U is a subspace of V.

By the definition of U, every vector in U can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Thus $\mathbf{v}_1, \ldots, \mathbf{v}_m$ span U.

Example 9

Consider the vector space $\mathbf{R}^3$.

The vectors $(-1, 5, 3)$ and $(2, -3, 4)$ are in $\mathbf{R}^3$.

Let U be the subset of $\mathbf{R}^3$ consisting of all vectors of the form

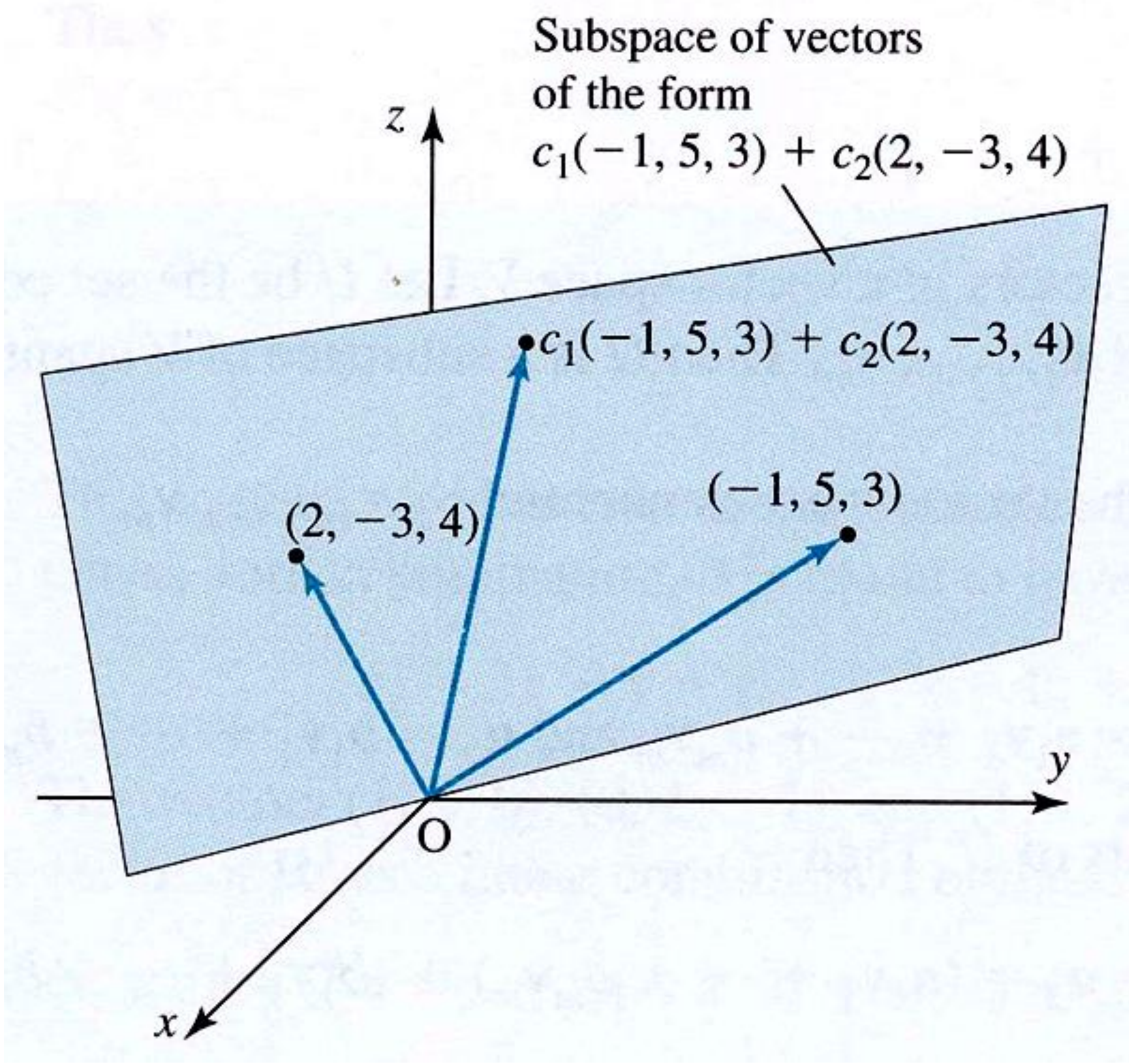
$$c_1(-1, 5, 3) + c_2(2, -3, 4)$$

Then U is a subspace of $\mathbf{R}^3$ spanned by $(-1, 5, 3)$ and $(2, -3, 4)$.

The following are examples of some of the vectors in U, obtained by given c_1 and c_2 various values.

$$c_1 = 1, c_2 = 0; \text{vector } (-1, 5, 3)$$
$$c_1 = 0, c_2 = 1; \text{vector } (2, -3, 4)$$
$$c_1 = 0, c_2 = 0; \text{vector } (0, 0, 0)$$
$$c_1 = 2, c_2 = 3; \text{vector } (4, 1, 18)$$

We can visualize U. U is made up of all vectors in the plane defined by the vectors $(-1, 5, 3)$ and $(2, -3, 4)$.

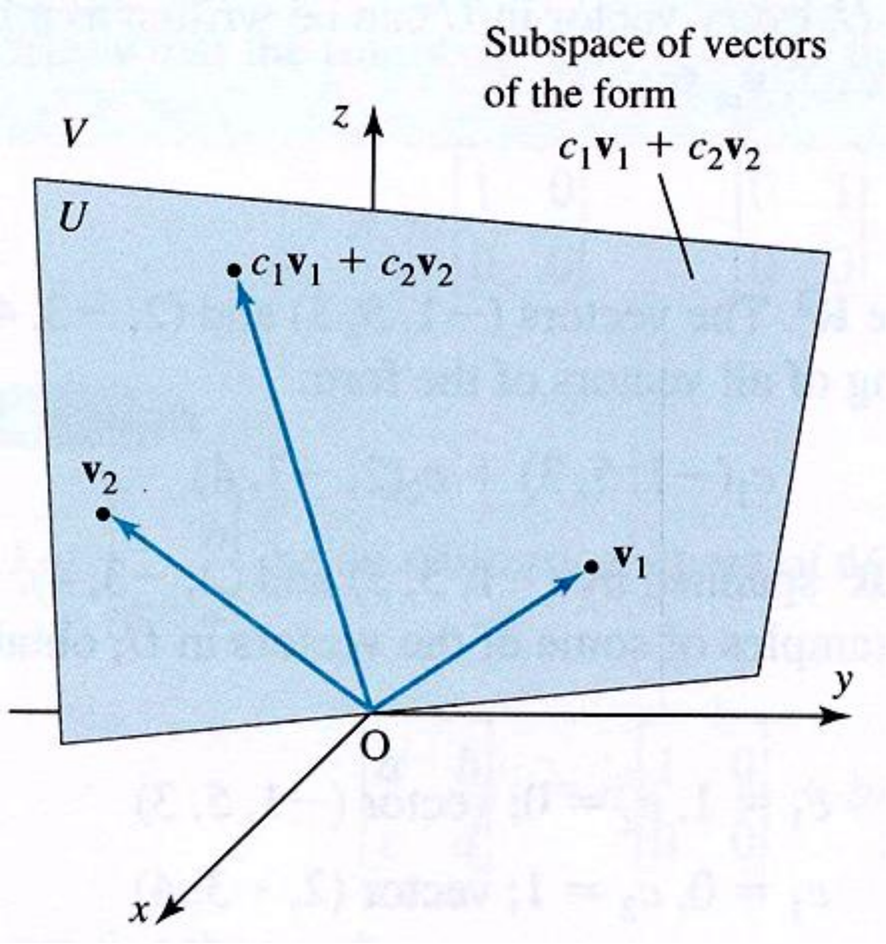


Subspace of vectors
of the form
$c_1(-1, 5, 3) + c_2(2, -3, 4)$
z
$c_1(-1, 5, 3) + c_2(2, -3, 4)$
$(2, -3, 4)$
$(-1, 5, 3)$
y
O
x

We can generalize this result. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be vectors in the space $\mathbf{R}^3$.

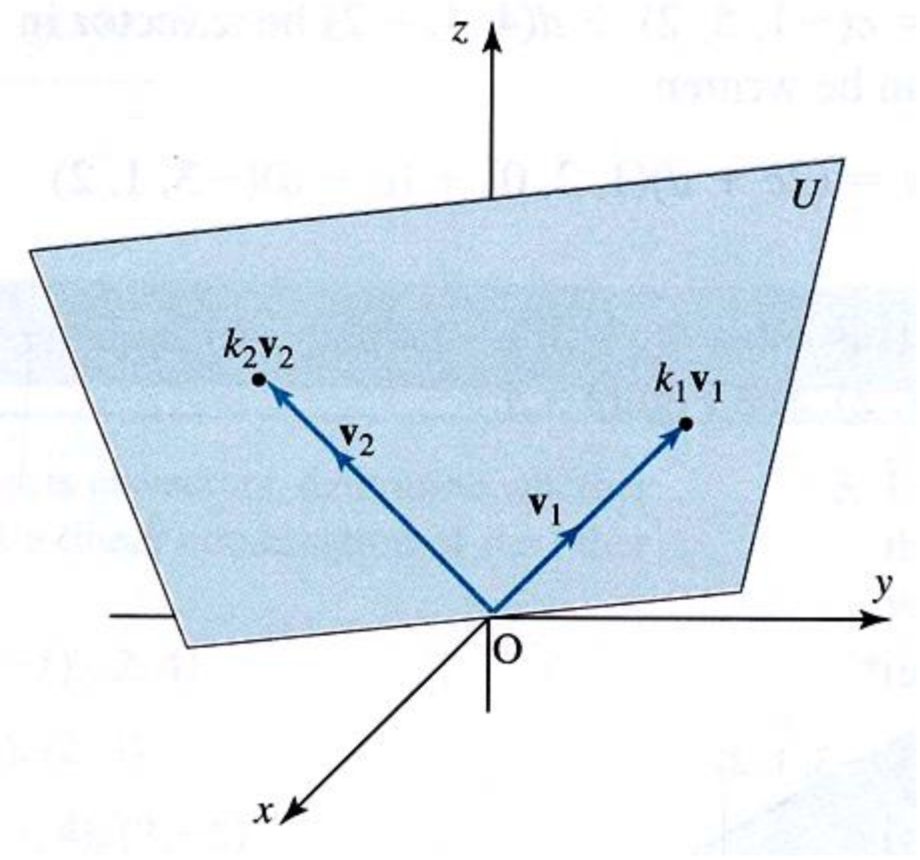
The subspace U generated by $\mathbf{v}_1$ and $\mathbf{v}_2$ is the set of all vectors of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$.

If $\mathbf{v}_1$ and $\mathbf{v}_2$ are not colinear, then U is the plane defined by $\mathbf{v}_1$ and $\mathbf{v}_2$.



If $\mathbf{v}_1$ and $\mathbf{v}_2$ are vectors in $\mathbf{R}^3$ that are not colinear, then we can visualize U as a plane in three dimensions.

$k_1\mathbf{v}_1$ and $k_2\mathbf{v}_2$ will be vectors on the same lines as $\mathbf{v}_1$ and $\mathbf{v}_2$.



Example 10

Let U be the subspace of $\mathbf{R}^3$ generated by the vectors (1, 2, 0) and (–3, 1, 2). Let V be the subspace of $\mathbf{R}^3$ generated by the vectors (–1, 5, 2) and (4, 1, –2). Show that $U = V$.

Solution

$(U \subseteq V)$ Let $\mathbf{u}$ be a vector in U. Let us show that $\mathbf{u}$ is in V.
Since $\mathbf{u}$ is in U, there exist scalars a and b such that

$$\mathbf{u} = a(1, 2, 0) + b(-3, 1, 2) = (a - 3b, 2a + b, 2b)$$

Let us see if we can write $\mathbf{u}$ as a linear combination of (–1, 5, 2) and (4, 1, –2).

$$\mathbf{u} = p(-1, 5, 2) + q(4, 1, -2) = (-p + 4q, 5p + q, 2p - 2q)$$

Such p and q would have to satisfy

$$\begin{aligned} -p + 4q &= a - 3b \\ 5p + q &= 2a + b \\ 2p - 2q &= 2b \end{aligned}$$

This system of equations has unique solution $p = \frac{a+b}{3}, q = \frac{a-2b}{3}$.
Thus $\mathbf{u}$ can be written

$$\mathbf{u} = \frac{a+b}{3}(-1, 5, 2) + \frac{a-2b}{3}(4, 1, -2)$$

Therefore, $\mathbf{u}$ is a vector in V.

$(V \subseteq U)$ Let $\mathbf{v}$ be a vector in V. Let us show that $\mathbf{v}$ is in U.
Since $\mathbf{v}$ is in V, there exist scalars c and d such that

$$\mathbf{v} = c(-1, 5, 2) + d(4, 1, -2)$$

It can be shown that

$$\mathbf{v} = (2c+d)(1, 2, 0) + (c-d)(-3, 1, 2)$$

Therefore, $\mathbf{v}$ is a vector in U and hence $U=V$.

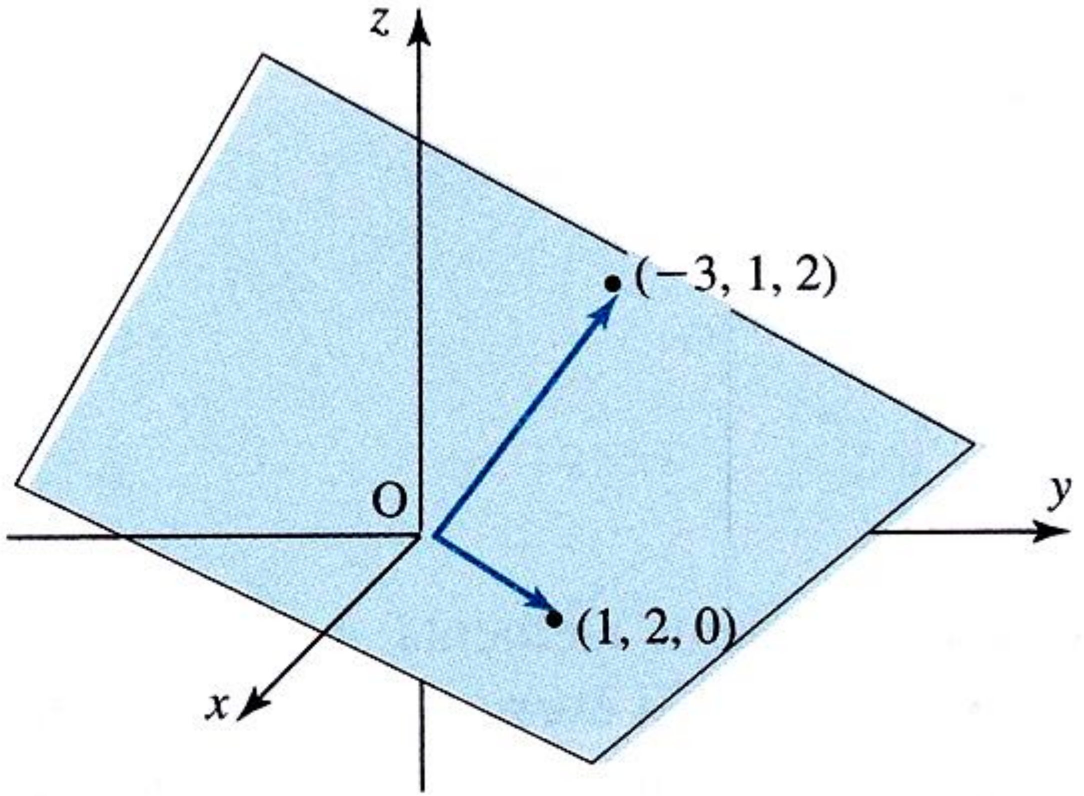


Figure 4.13

Example 11

Let U be the vector space generated by the functions $f(x) = x + 1$ and $g(x) = 2x^2 – 2x + 3$. Show that the function $h(x) = 6x^2–10x+5$ lies in U.

Solution

h will be in the space generated by f and g if there exist scalars a and b such that

$$a(x + 1) + b(2x^2 – 2x + 3) = 6x^2 – 10x + 5$$

This given

$$2bx^2 + (a – 2b)x + a + 3b = 6x^2 – 10x + 5$$

$$\Rightarrow \quad \begin{aligned} 2b &= 6 \\ a - 2b &= -10 \\ a + 3b &= 5 \end{aligned}$$

This system has the unique solution $a = – 4$, $b = 3$.

Thus $\quad – 4(x + 1) + 3(2x^2 – 2x + 3) = 6x^2 – 10x + 5$

4.6 Linear Dependence and Independence

The concepts of dependence and independence of vectors are useful tools in constructing "efficient" spanning sets for vector spaces – sets in which there are no redundant vectors.

Definition

(a) The set of vectors $\{ \mathbf{v}_1, \ldots, \mathbf{v}_m \}$ in a vector space V is said to be **linearly dependent** if there exist scalars $c_1, \ldots, c_m$, not all zero, such that $c_1\mathbf{v}_1 + \ldots + c_m\mathbf{v}_m = 0$

(b) The set of vectors $\{ \mathbf{v}_1, \ldots, \mathbf{v}_m \}$ is **linearly independent** if $c_1\mathbf{v}_1 + \ldots + c_m\mathbf{v}_m = 0$ can only be satisfied when $c_1 = 0, \ldots, c_m = 0$.

Example 1

Show that the set {(1, 2, 3), (–2, 1, 1), (8, 6, 10)} is linearly dependent in $\mathbf{R}^3$.

Solution

Suppose $c_1(1,2,3)+c_2(-2,1,1)+c_3(8,6,10)=\mathbf{0}$

$$\Rightarrow \quad (c_1, 2c_1, 3c_1)+(-2c_2, c_2, c_2)+(8c_3, 6c_3, 10c_3)=\mathbf{0}$$

$$(c_1-2c_2+8c_3, 2c_1+c_2+6c_3, 3c_1+c_2+10c_3)=\mathbf{0}$$

$$\Rightarrow \begin{cases} c_1-2c_2+8c_3=0 \\ 2c_1+c_2+6c_3=0 \\ 3c_1+c_2+10c_3=0 \end{cases} \Rightarrow \quad \begin{aligned} c_1 &= 4 \\ c_2 &= -2 \\ c_3 &= -1 \end{aligned}$$

Thus $4(1,2,3)-2(-2,1,1)-(8,6,10)=\mathbf{0}$

The set of vectors is linearly dependent.

Example 2

Show that the set {(3, –2, 2), (3, –1, 4), (1, 0, 5)} is linearly independent in $\mathbf{R}^3$.

Solution

Suppose $c_1(3,-2,2)+c_2(3,-1,4)+c_3(1,0,5)=\mathbf{0}$

$$\Rightarrow \quad (3c_1,-2c_1,2c_1)+(3c_2,-c_2,4c_2)+(c_3,0,5c_3)=\mathbf{0}$$
$$(3c_1+3c_2+c_3,-2c_1-c_2,2c_1+4c_2+5c_3)=\mathbf{0}$$

$$\Rightarrow \begin{cases} 3c_1+3c_2+c_3=0 \\ -2c_1-c_2=0 \\ 2c_1+4c_2+5c_3=0 \end{cases}$$

This system has the unique solution $c_1 = 0$, $c_2 = 0$, and $c_3 = 0$.
Thus the set is linearly independent.

Example 3

Consider the functions $f(x) = x^2 + 1$, $g(x) = 3x - 1$, $h(x) = -4x + 1$ of the vector space P_2 of polynomials of degree ≤ 2.
Show that the set of functions $\{ f, g, h \}$ is linearly independent.

Solution

Suppose

$$c_1 f + c_2 g + c_3 h = \mathbf{0}$$

Since for any real number x,

$$c_1(x^2+1) + c_2(3x-1) + c_3(-4x+1) = \mathbf{0}$$

Consider three convenient values of x. We get

$$x = 0: c_1 - c_2 + c_3 = 0$$
$$x = 1: 2c_1 + 2c_2 - 3c_3 = 0$$
$$x = -1: 2c_1 - 4c_2 + 5c_3 = 0$$

It can be shown that this system of three equations has the unique solution

$$c_1 = 0, c_2 = 0, c_3 = 0$$

Thus $c_1 f + c_2 g + c_3 h = \mathbf{0}$ implies that $c_1 = 0$, $c_2 = 0$, $c_3 = 0$.
The set $\{ f, g, h \}$ is linearly independent.

Theorem 4.7

A set consisting of two or more vectors in a vector space is linearly dependent *if and only if* it is possible to express one of the vectors as a linearly combination of the other vectors.

Example 4

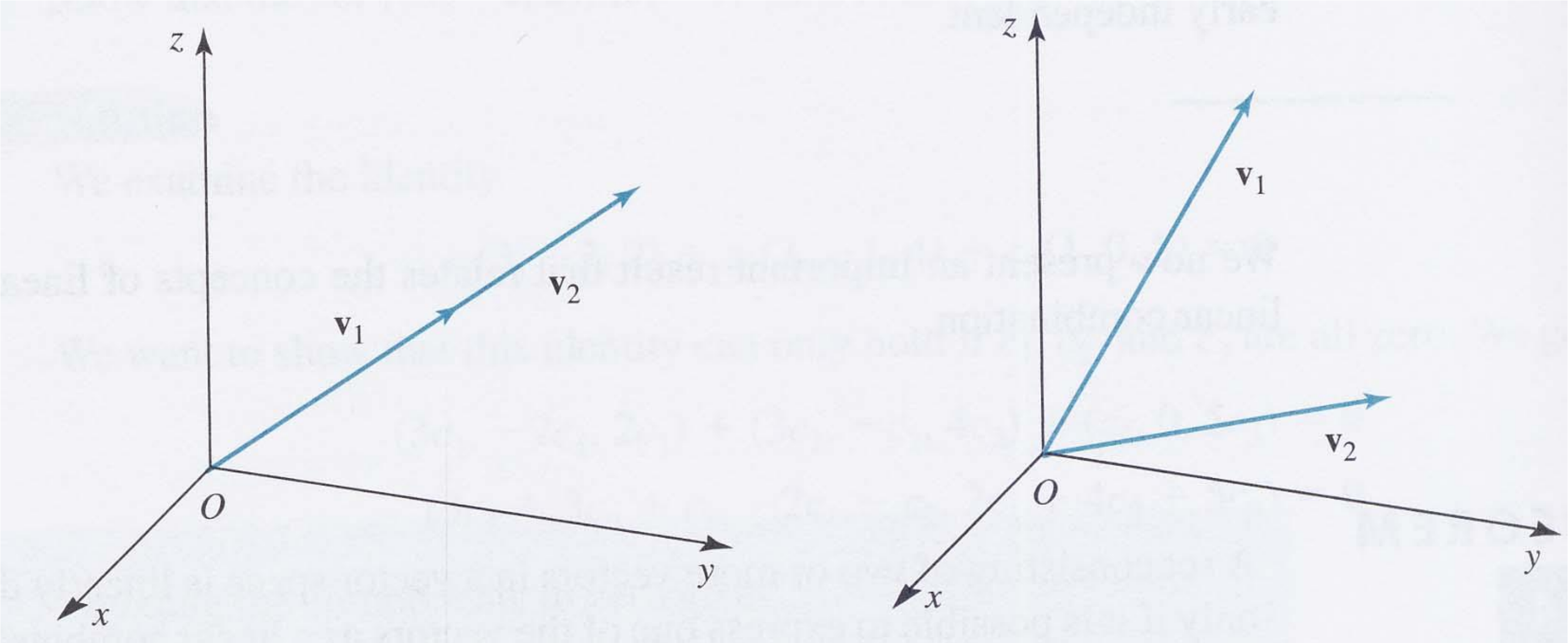
The set of vectors $\{\mathbf{v}_1=(1, 2, 1)\ , \mathbf{v}_2=(-1, -1, 0)\ , \mathbf{v}_3 = (0, 1,1)\}$

is linearly dependent, since $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$.

Thus, $\mathbf{v}_3$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

m

Linear Dependence of $\{v_1, v_2\}$

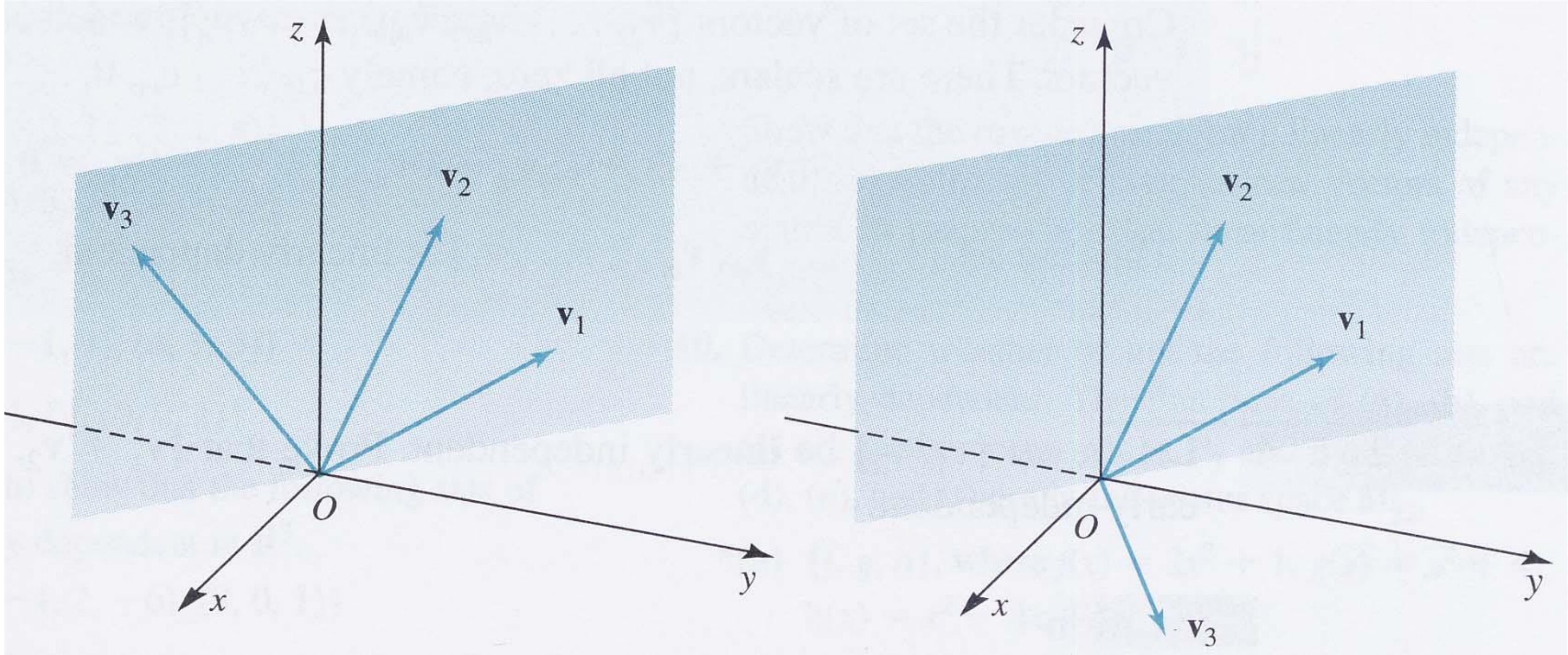


$\{\mathbf{v}_1, \mathbf{v}_2\}$ linearly dependent; vectors lie on a line

$\{\mathbf{v}_1, \mathbf{v}_2\}$ linearly independent; vectors do not lie on a line

Figure 4.14 Linear dependence and independence of $\{\mathbf{v}_1, \mathbf{v}_2\}$ in $\mathbf{R}^3$.

Linear Dependence of {v_1, v_2, v_3}



{$\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$} linearly dependent; vectors lie in a plane

{$\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$} linearly independent; vectors do not lie in a plane

Figure 4.15 Linear dependence and independence of {$\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$} in $\mathbf{R}^3$.

Theorem 4.8

Let V be a vector space. Any set of vectors in V that contains the zero is linearly dependent.

Proof

Consider the set $\{ \mathbf{0}, \mathbf{v}_2, \ldots, \mathbf{v}_m \}$, which contains the zero vectors. Let us examine the identity

$$c_1\mathbf{0} + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$$

We see that the identity is true for $c_1 = 1$, $c_2 = 0$, …, $c_m = 0$ (not all zero).

Thus the set of vectors is linearly dependent, proving the theorem.

Theorem 4.9

Let the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ be linearly dependent in a vector space V. Any set of vectors in V that contains these vectors will also be linearly dependent.

Example 5

The set of vectors

$\{\mathbf{v}_1=(1, 2, 1)\ , \mathbf{v}_2 =(-1, -1, 0)\ , \mathbf{v}_3 = (0, 1,1)\ , \mathbf{v}_4 =(1, 1, 1)\ \}$

is linearly dependent, since it contains the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ which are linearly dependent.

Homework

- Exercise set 4.3 pages 219-220:
  1, 3, 7, 8, 9, 13, 15, 17.

4.7 Bases and Dimension

Definition

A finite set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is called a **basis** for a vector space V if the set spans V and is linearly independent.

Standard Basis

The set of n vectors

$$\{(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, \ldots, 1)\}$$

is a basis for $\mathbf{R}^n$. This basis is called the **standard basis** for $\mathbf{R}^n$.

How to prove it?

Example 1

Show that the set $\{(1, 0, -1), (1, 1, 1), (1, 2, 4)\}$ is a basis for $\mathbf{R}^3$.

Solution

(span)

Let (x_1, x_2, x_3) be an arbitrary element of $\mathbf{R}^3$.

Suppose

$$(x_1, x_2, x_3) = a_1(1, 0, -1) + a_2(1, 1, 1) + a_3(1, 2, 4)$$

$$\Rightarrow \begin{cases} a_1 + a_2 + a_3 = x_1 \\ a_2 + 2a_3 = x_2 \\ -a_1 + a_2 + 4a_3 = x_3 \end{cases} \Rightarrow \begin{cases} a_1 = 2x_1 - 3x_2 + x_3 \\ a_2 = -2x_1 + 5x_2 - 2x_3 \\ a_3 = x_1 - 2x_2 + x_3 \end{cases}$$

Thus the set spans the space.

(linearly independent)

Consider the identity

$$b_1(1, 0, -1) + b_2(1, 1, 1) + b_3(1, 2, 4) = (0, 0, 0)$$

The identity leads to the system of equations

$$\begin{aligned} b_1 + b_2 + b_3 &= 0 \\ b_2 + 2b_3 &= 0 \\ -b_1 + b_2 + 4b_3 &= 0 \end{aligned}$$

$\Rightarrow$ $b_1 = 0$, $b_2 = 0$, and $b_3 = 0$ is the unique solution.
Thus the set is linearly independent.

$\Rightarrow$ $\{(1, 0, -1), (1, 1, 1), (1, 2, 4)\}$ spans $\mathbf{R}^3$ and is linearly independent.

$\Rightarrow$ It forms a basis for $\mathbf{R}^3$.

Example 2

Show that $\{ f, g, h \}$, where $f(x) = x^2 + 1$, $g(x) = 3x - 1$, and $h(x) = -4x + 1$ is a basis for P_2.

Solution

(linearly independent) see Example 3 of the previous section.

(span). Let p be an arbitrary function in P_2.

p is thus a polynomial of the form

$$p(x) = bx^2 + cx + d$$

Suppose $\quad p(x) = a_1 f(x) + a_2 g(x) + a_3 h(x)$

for some scalars a_1, a_2, a_3.

This gives

$$\begin{aligned} bx^2 + cx + d &= a_1(x^2 + 1) + a_2(3x - 1) + a_3(-4x + 1) \\ &= a_1 x^2 + (3a_2 - 4a_3)x + (a_1 - a_2 + a_3) \end{aligned}$$

$$\Rightarrow \begin{cases} a_1 = b \\ 3a_2 - 4a_3 = c \\ a_1 - a_2 + a_3 = d \end{cases} \Rightarrow \begin{cases} a_1 = b \\ a_2 = 4b - 4d - c \\ a_3 = 3b - 3d - c \end{cases}$$

Thus the polynomial p can be expressed

$$p(x) = a_1 f(x) + a_2 g(x) + a_3 h(x)$$

The functions f, g, and h span P_2.

They form a basis for P_2.

Theorem 4.10

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V.
If $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ is a set of more than n vectors in V, then this set is linearly dependent.

Proof

Suppose

$$c_1\mathbf{w}_1 + \cdots + c_m\mathbf{w}_m = 0 \qquad (1)$$

We will show that values of $c_1, \ldots, c_m$ are not all zero.

The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for V. Thus each of the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_m$ can be expressed as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Let

$$\begin{gathered}\mathbf{w}_1 = a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \cdots + a_{1n}\mathbf{v}_n \\ \vdots \\ \mathbf{w}_m = a_{m1}\mathbf{v}_1 + a_{m2}\mathbf{v}_2 + \cdots + a_{mn}\mathbf{v}_n\end{gathered}$$

Substituting for $\mathbf{w}_1, \ldots, \mathbf{w}_m$ into Equation (1) we get

$$c_1(a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \cdots + a_{1n}\mathbf{v}_n) + \cdots + c_m(a_{m1}\mathbf{v}_1 + a_{m2}\mathbf{v}_2 + \cdots + a_{mn}\mathbf{v}_n) = \mathbf{0}$$

Rearranging, we get

$$(c_1a_{11} + c_2a_{21} + \cdots + c_ma_{m1})\mathbf{v}_1 + \cdots + (c_1a_{1n} + c_2a_{2n} + \cdots + c_ma_{mn})\mathbf{v}_n = \mathbf{0}$$

Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linear independent,

$$\begin{aligned} a_{11}c_1 + a_{21}c_2 + \cdots + a_{m1}c_m &= 0 \\ &\vdots \\ a_{1n}c_1 + a_{2n}c_2 + \cdots + a_{mn}c_m &= 0 \end{aligned}$$

Since $m > n$, there are many solutions in this system.

Thus the set $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ is linearly dependent.

Theorem 4.11

Any two bases for a vector space V consist of the same number of vectors.

Proof

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ be two bases for V.

By Theorem 4.10,

$$m \leq n \text{ and } n \leq m$$

Thus $n = m$.

Definition

If a vector space V has a basis consisting of n vectors, then the **dimension** of V is said to be n. We write **dim(V)** for the dimension of V.

- V is **finite dimensional** if such a finite basis exists.
- V is **infinite dimensional** otherwise.

Example 3

Consider the set {{1, 2, 3), (-2, 4, 1)} of vectors in $\mathbf{R}^3$.
These vectors generate a subspace V of $\mathbf{R}^3$ consisting of all vectors of the form

$$\mathbf{v} = c_1(1, 2, 3) + c_2(-2, 4, 1)$$

The vectors (1, 2, 3) and (-2, 4, 1) span this subspace.

Furthermore, since the second vector is not a scalar multiple of the first vector, the vectors are linearly independent.
Therefore {{1, 2, 3), (-2, 4, 1)} is a basis for V.
Thus dim(V) = 2.
We know that V is, in fact, a plane through the origin.

Theorem 4.12

(a) The origin is a subspace of $\mathbf{R}^3$. The dimension of this subspace is defined to be zero.

(b) The one-dimensional subspaces of $\mathbf{R}^3$ are lines through the origin.

(c) The two-dimensional subspaces of $\mathbf{R}^3$ are planes through the origin.

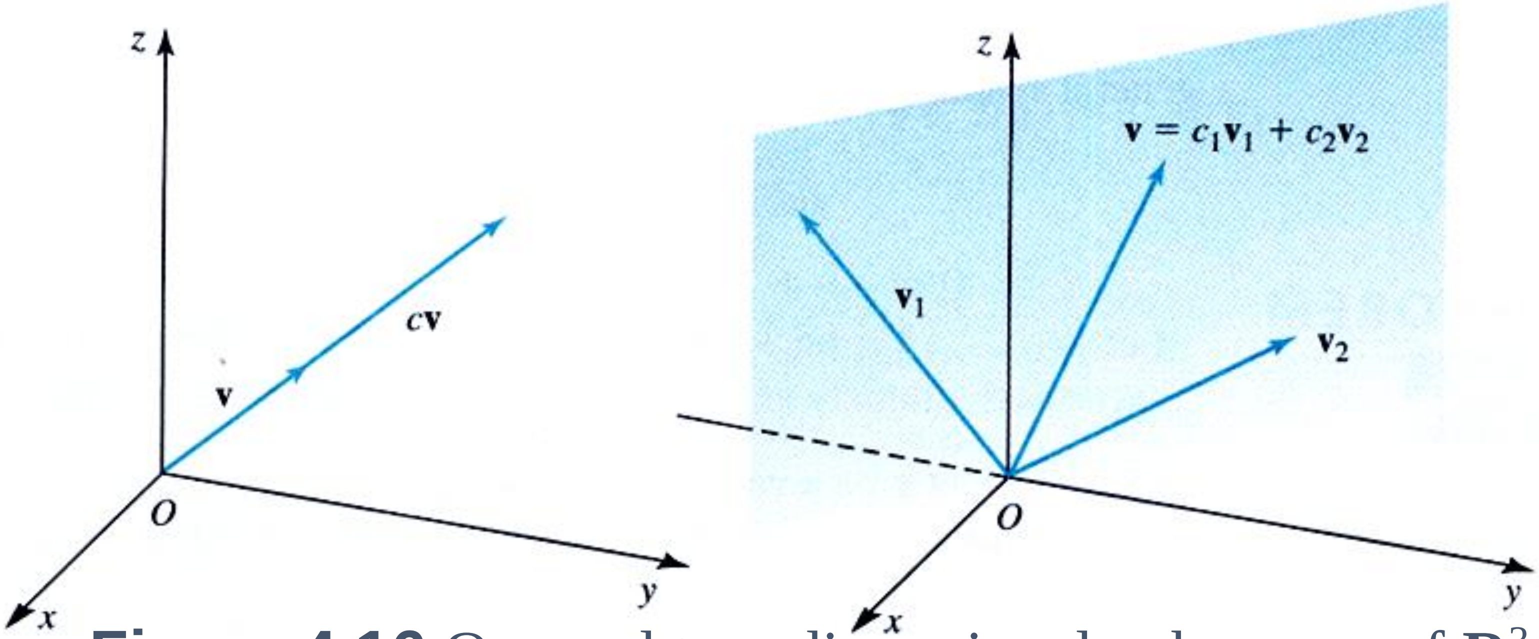


Figure 4.16 One and two-dimensional subspaces of $\mathbf{R}^3$

Proof

(a) Let V be the set $\{(0, 0, 0)\}$, consisting of a single elements, the zero vector of $\mathbf{R}^3$. Let c be the arbitrary scalar. Since

$$(0, 0, 0) + (0, 0, 0) = (0, 0, 0) \text{ and } c(0, 0, 0) = (0, 0, 0)$$

V is closed under addition and scalar multiplication. It is thus a subspace of $\mathbf{R}^3$. The dimension of this subspaces is defined to be zero.

(b) Let $\mathbf{v}$ be a basis for a one-dimensional subspace V of $\mathbf{R}^3$. Every vector in V is thus of the form $c\mathbf{v}$, for some scalar c. We know that these vectors form a line through the origin.

(c) Let $\{\mathbf{v}_1, \mathbf{v}_2\}$be a basis for a two-dimensional subspace V of $\mathbf{R}^3$. Every vector in V is of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. V is thus a plane through the origin.

Theorem 4.13

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V. Then each vector in V can be expressed uniquely as a linear combination of these vectors.

Proof

Let $\mathbf{v}$ be a vector in V. Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis, we can express $\mathbf{v}$ as a linear combination of these vectors. Suppose we can write

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \text{ and } \mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$$

Then

$$a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$$

giving

$$(a_1 - b_1)\mathbf{v}_1 + \cdots + (a_n - b_n)\mathbf{v}_n = \mathbf{0}$$

Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent. Thus $(a_1 - b_1) = 0, \ldots, (a_n - b_n) = 0$, implying that $a_1 = b_1, \ldots, a_n = b_n$. There is thus only one way of expressing $\mathbf{v}$ as a linear combination of the basis.

Theorem 4.14

Let V be a vector space of dimension n.

(a) If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of n linearly independent vectors in V, then S is a basis for V.

(b) If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of n vectors V that spans V, then S is a basis for V.

Let V be a vector space, $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of vectors in V.

(a) $\dim(V) = |S|$.

(b) S is a linearly independent set.

(c) S spans V.

S is a basis of V.

Example 4

Prove that the set B={(1, 3, -1), (2, 1, 0), (4, 2, 1)} is a basis for $\mathbf{R}^3$.

Solution

Since $\dim(\mathbf{R}^3)=|B|=3$. It suffices to show that this set is linearly independent or it spans $\mathbf{R}^3$.

Let us check for linear independence. Suppose

$$c_1(1,3,-1)+c_2(2,1,0)+c_3(4,2,1)=(0,0,0)$$

This identity leads to the system of equations

$$\begin{aligned} c_1+2c_2+4c_3&=0\\ 3c_1+c_2+2c_3&=0\\ -c_1+c_3&=0 \end{aligned}$$

This system has the unique solution $c_1 = 0$, $c_2 = 0$, $c_3 = 0$.

Thus the vectors are linearly independent.

The set {(1, 3, -1), (2, 1, 0), (4, 2, 1)} is therefore a basis for $\mathbf{R}^3$.

Theorem 4.15

Let V be a vector space of dimension n. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ be a set of m linearly independent vectors in V, where $m < n$.

Then there exist vectors $\mathbf{v}_{m+1}, \ldots, \mathbf{v}_n$ such that

$\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$ is a basis of V.

Example 5

State (with a brief explanation) whether the following statements are true or false.

(a) The vectors (1, 2), (–1, 3), (5, 2) are linearly dependent in $\mathbf{R}^2$.

(b) The vectors (1, 0, 0), (0, 2, 0), (1, 2, 0) span $\mathbf{R}^3$.

(c) {(1, 0, 2), (0, 1, -3)} is a basis for the subspace of $\mathbf{R}^3$ consisting of vectors of the form $(a, b, 2a - 3b)$.

(d) Any set of two vectors can be used to generate a two-dimensional subspace of $\mathbf{R}^3$.

Solution

(a) True: The dimension of $\mathbf{R}^2$ is two. Thus any three vectors are linearly dependent.

(b) False: The three vectors are linearly dependent. Thus they cannot span a three-dimensional space.

(c) True: The vectors span the subspace since

$$(a, b, 2a - 3b) = a(1, 0, 2) + b(0, 1, -3)$$

The vectors are also linearly independent since they are not colinear.

(d) False: The two vectors must be linearly independent.

Inverse of Matrix

Prepared by Vince Zaccone

For Campus Learning
Assistance Services at UCSB

The inverse of a square matrix A is another matrix with the following properties:

$$A \cdot A^{-1} = A^{-1} \cdot A = I$$

Here I represents the identity matrix of the same size as A and A^{-1}.

Note that A^{-1} must be a square matrix of the same size as A.

The inverse of a square matrix A is another matrix with the following properties:

$$A \cdot A^{-1} = A^{-1} \cdot A = I$$

Here I represents the identity matrix of the same size as A and A^{-1}.

Note that A^{-1} must be a square matrix of the same size as A.

Here is a system of linear equations. To solve it, we can put it into matrix format and try to find the inverse of the coefficient matrix.

Let's see how that works.

$$x - y + z = 2$$
$$x + y = 5$$
$$x + 2y - z = 4$$

Prepared by Vince Zaccone

For Campus Learning Assistance Services at UCSB

The inverse of a square matrix A is another matrix with the following properties:

$$A \cdot A^{-1} = A^{-1} \cdot A = I$$

Here I represents the identity matrix of the same size as A and A^{-1}.

Note that A^{-1} must be a square matrix of the same size as A.

Here is a system of linear equations. To solve it, we can put it into matrix format and try to find the inverse of the coefficient matrix.

Let's see how that works.

$$\begin{aligned} x - y + z &= 2 \\ x + y &= 5 \\ x + 2y - z &= 4 \end{aligned}$$

$$A \cdot \vec{x} = \vec{b}$$

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & -1 \end{bmatrix}; \vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}; \vec{b} = \begin{bmatrix} 2 \\ 5 \\ 4 \end{bmatrix}$$

Prepared by Vince Zaccone

For Campus Learning Assistance Services at UCSB

To find the inverse, for an augmented matrix with the coefficient matrix on the left, and the corresponding identity matrix on the right.

Next, row reduce until you have the identity on the left, and the inverse will be on the right.

$$[\mathrm{A} \quad \mathrm{I}] \Rightarrow \begin{array}{c}\text{row} \\ \text{reduction}\end{array} \Rightarrow [\mathrm{I} \quad \mathrm{A}^{-1}]$$

Prepared by Vince Zaccone

For Campus Learning Assistance Services at UCSB

To find the inverse, for an augmented matrix with the coefficient matrix on the left, and the corresponding identity matrix on the right.

Next, row reduce until you have the identity on the left, and the inverse will be on the right.

$$[A \quad I] \Rightarrow \begin{matrix}\text{row}\\ \text{reduction}\end{matrix} \Rightarrow [I \quad A^{-1}]$$

Here is the method, applied to our example:

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

Prepared by Vince Zaccone

For Campus Learning Assistance Services at UCSB

To find the inverse, for an augmented matrix with the coefficient matrix on the left, and the corresponding identity matrix on the right.

Next, row reduce until you have the identity on the left, and the inverse will be on the right.

$$[A \quad I] \Rightarrow \begin{matrix}\text{row}\\ \text{reduction}\end{matrix} \Rightarrow [I \quad A^{-1}]$$

Here is the method, applied to our example:

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right] \begin{matrix} \\ R_2^* = R_2 - R_1 \\ R_3^* = R_3 - R_1 \end{matrix} \quad \Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 3 & -2 & -1 & 0 & 1 \end{array}\right]$$

To find the inverse, for an augmented matrix with the coefficient matrix on the left, and the corresponding identity matrix on the right.

Next, row reduce until you have the identity on the left, and the inverse will be on the right.

$$[A \quad I] \Rightarrow \begin{matrix}\text{row}\\ \text{reduction}\end{matrix} \Rightarrow [I \quad A^{-1}]$$

Here is the method, applied to our example:

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right] \begin{matrix} \\ R_2^* = R_2 - R_1 \\ R_3^* = R_3 - R_1 \end{matrix} \Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 3 & -2 & -1 & 0 & 1 \end{array}\right] \begin{matrix} \\ \\ R_3^* = 3R_2 - 2R_3 \end{matrix}$$

Prepared by Vince Zaccone

For Campus Learning Assistance Services at UCSB

To find the inverse, for an augmented matrix with the coefficient matrix on the left, and the corresponding identity matrix on the right.

Next, row reduce until you have the identity on the left, and the inverse will be on the right.

$$[A \quad I] \Rightarrow \begin{matrix}\text{row}\\ \text{reduction}\end{matrix} \Rightarrow [I \quad A^{-1}]$$

Here is the method, applied to our example:

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right] \begin{matrix} \\ R_2^* = R_2 - R_1 \\ R_3^* = R_3 - R_1 \end{matrix} \Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 3 & -2 & -1 & 0 & 1 \end{array}\right] \begin{matrix} \\ \\ R_3^* = 3R_2 - 2R_3 \end{matrix}$$

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

To find the inverse, for an augmented matrix with the coefficient matrix on the left, and the corresponding identity matrix on the right.

Next, row reduce until you have the identity on the left, and the inverse will be on the right.

$$[A \quad I] \Rightarrow \begin{matrix}\text{row}\\ \text{reduction}\end{matrix} \Rightarrow [I \quad A^{-1}]$$

Here is the method, applied to our example:

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right] \begin{matrix} \\ R_2^* = R_2 - R_1 \\ R_3^* = R_3 - R_1 \end{matrix} \Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 3 & -2 & -1 & 0 & 1 \end{array}\right] \begin{matrix} \\ \\ R_3^* = 3R_2 - 2R_3 \end{matrix}$$

$$\Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \end{array}\right] R_2^* = R_2 + R_3$$

To find the inverse, for an augmented matrix with the coefficient matrix on the left, and the corresponding identity matrix on the right.

Next, row reduce until you have the identity on the left, and the inverse will be on the right.

$$[A \quad I] \Rightarrow \begin{matrix}\text{row}\\\text{reduction}\end{matrix} \Rightarrow [I \quad A^{-1}]$$

Here is the method, applied to our example:

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right] \begin{matrix} \\ R_2^* = R_2 - R_1 \\ R_3^* = R_3 - R_1 \end{matrix} \quad \Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 3 & -2 & -1 & 0 & 1 \end{array}\right] \begin{matrix} \\ \\ R_3^* = 3R_2 - 2R_3 \end{matrix}$$

$$\Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \end{array}\right] R_2^* = R_2 + R_3 \quad \Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & -2 & 4 & -2 \\ 0 & 0 & 1 & -1 & 3 & -2 \end{array}\right]$$

Prepared by Vince Zaccone

For Campus Learning Assistance Services at UCSB

To find the inverse, for an augmented matrix with the coefficient matrix on the left, and the corresponding identity matrix on the right.

Next, row reduce until you have the identity on the left, and the inverse will be on the right.

$$[A \quad I] \Rightarrow \begin{matrix}\text{row}\\ \text{reduction}\end{matrix} \Rightarrow [I \quad A^{-1}]$$

Here is the method, applied to our example:

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right] \begin{matrix} \\ R_2^* = R_2 - R_1 \\ R_3^* = R_3 - R_1 \end{matrix} \Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 3 & -2 & -1 & 0 & 1 \end{array}\right] R_3^* = 3R_2 - 2R_3$$

$$\Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \end{array}\right] R_2^* = R_2 + R_3 \Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & -2 & 4 & -2 \\ 0 & 0 & 1 & -1 & 3 & -2 \end{array}\right] R_2^* = \tfrac{1}{2} R_2$$

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

Prepared by Vince Zaccone

For Campus Learning Assistance Services at UCSB

To find the inverse, for an augmented matrix with the coefficient matrix on the left, and the corresponding identity matrix on the right.

Next, row reduce until you have the identity on the left, and the inverse will be on the right.

$$[A \quad I] \Rightarrow \begin{matrix} \text{row} \\ \text{reduction} \end{matrix} \Rightarrow [I \quad A^{-1}]$$

Here is the method, applied to our example:

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right] \begin{matrix} \\ R_2^* = R_2 - R_1 \\ R_3^* = R_3 - R_1 \end{matrix} \quad \Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 3 & -2 & -1 & 0 & 1 \end{array}\right] \begin{matrix} \\ \\ R_3^* = 3R_2 - 2R_3 \end{matrix}$$

$$\Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \end{array}\right] R_2^* = R_2 + R_3 \quad \Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & -2 & 4 & -2 \\ 0 & 0 & 1 & -1 & 3 & -2 \end{array}\right] R_2^* = \tfrac{1}{2} R_2$$

$$\Rightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & -1 & 3 & -2 \end{array}\right] \begin{matrix} R_1^* = R_1 + R_2 - R_3 \\ \\ \\ \end{matrix} \Rightarrow \left[\begin{array}{ccc|ccc} \multicolumn{3}{c|}{\overbrace{\hphantom{1 \quad 0 \quad 0}}^{I}} & \multicolumn{3}{c}{\overbrace{\hphantom{1 \quad -1 \quad 1}}^{A^{-1}}} \\ 1 & 0 & 0 & 1 & -1 & 1 \\ 0 & 1 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & -1 & 3 & -2 \end{array}\right]$$

Prepared by Vince Zaccone

For Campus Learning Assistance Services at UCSB

2 x 2 Inverse Matrix Shortcut

- **Theorem 4:** Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, If $ad - bc \neq 0$, then

  A is invertible and

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

  If $ad - bc = 0$, then *A* is not invertible.

- The quantity $ad - bc$ is called the determinant of *A*, and we write $\det A = ad - bc$
- This theorem says that a 2×2 matrix *A* is invertible if and only if $\det A \neq 0$

- **Theorem 8:** Let *A* be a square matrix. Then the following statements are equivalent. That is, for a given *A*, the statements are either all true or all false.

  a. *A* is an invertible matrix.

  b. *A* is row equivalent to the $n \times n$ identity matrix.

  c. *A* has *n* pivot positions.

  d. The equation $A\mathrm{x} = 0$ has only the trivial solution.

  e. The columns of *A* form a linearly independent set.

  f. The linear transformation $\mathrm{x} \mapsto A\mathrm{x}$ is one-to-one.

  g. The equation $A\mathrm{x} = \mathrm{b}$ has at least one solution for each b in $\mathbb{R}^n$.

  h. The columns of *A* span $\mathbb{R}^n$.

  i. The linear transformation $\mathrm{x} \mapsto A\mathrm{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

  j. There is an $n \times n$ matrix *C* such that $CA = I$.

  k. There is an $n \times n$ matrix *D* such that $AD = I$.

  l. A^T is an invertible matrix.

m. INVERSE OF 3 x 3 matrix ?

Prepared by Vince Zaccone

For Campus Learning Assistance Services at UCSB