IT 5845
Mathematics for Artificial Intelligence

Lecture 1

Dr. Thilini Piyatilake

Department of Computational Mathematics
University of Moratuwa

thilinisp@uom.lk

MSc in Artificial Intelligence

Course Details

Module Code: IT 5845

Module Title: Mathematics for Artificial Intelligence

No. of Credits: 3

Course Objectives

To provide a broad overview of influence areas of mathematics and statistics for Artificial Intelligence and thereby establish the mathematical and statistical basis required for studies in Artificial Intelligence.

Intended Learning Outcomes

On successful completion of this module, students should be able to;

- Recognize basic concepts and methods of vectors, matrix algebra and multivariate calculus.
- Apply techniques of vectors, matrix algebra and multivariate calculus to solve real world problems.
- Identify an appropriate probability distribution for a given discrete or continuous random variable and use its properties to calculate probabilities.
- Explain the rationale behind each method of estimation.
- Apply method of estimation to derive the estimators of the population parameters of the different types of distributions and calculate estimates.
- Use the latest tools to solve mathematical problems.

Outline Syllabus

- Vectors
  Vectors and vector quantities, Addition and subtraction of vectors, Magnitude and direction of a vector, Multiplication of a vector by a scalar and unit vectors, Basis vectors, Products of vectors, Vector equation of a line
- Linear Algebra
  Matrices, Transformations, Solving systems of equations, Gauss elimination, Inverse, Determinant and trace of a matrix, Eigenvectors and eigenvalues
- Functions of More than One Variable
  Functions of two variable, Partial differentiation and gradients, Changing variables: Chain rule, Total derivative along a path, Higher-order partial derivatives, Optimization: constraint and unconstrained, Double integrals

Outline Syllabus (Continue...)

- Elements of Probability
  Sample spaces and events, Event operations, Counting techniques, Axioms of probability, Methods for determining probability, Conditional probability, Rules of probability
- Distribution Theory
  Probability mass function, Probability density function, Cumulative distribution function, Descriptive properties of distributions, Models for discrete distributions, Models for continuous distributions
- Estimation
  Point estimation, Interval estimation

Assessment of Course Objectives

- The final grade of the course module will be based on continuous assignments, and a final examination. Assignment will be posted on the course MOODLE page and announced in class.
- Grading Policy
  - Continuous Assignments (40%)
  - End Examination (60%)
- **No late assignment/homework will be accepted.**
- **Make-Up Policy**
  There will be NO make-up/alternate exams for missed assignments.

Attendance Policy

It is student's responsibility to attend every class. Students are expected to arrive for class on time and to remain for the class entire period.

Chapter 1: Vector

Learning Outcomes

By the end of the lecture, students will be able to;

- define a vector.
- recognize algebraic laws in vectors.
- recognize the geometric representations of vectors.
- define and find the inner product and norm of a vector.
- recognize the properties of inner product and norm.
- find the angle between two vectors.

Vectors in $\mathbb{R}^n$

Vectors in $\mathbb{R}^n$

- At some time in our earlier mathematical studies, we learned the concept of ordered pairs (x, y) of real numbers, then ordered triplets (x, y, z), and perhaps ordered n-tuples $(x_1, x_2, \dots, x_n)$.
- This idea and these notations are simple and natural.
- In addition to being more important in theoretical applications (in defining a function, for example), these n-tuples serve to hold data and to carry information.
- For example, the ordered 4-tuple $(200219202757, 23, 1.7, 60)$ might describe a person with ID number 200219202757 who is 23 years old, 1.7 m tall, and weights 60 kg.
- In interpreting such data, it is important that we understand the order in which the information is given.
- Despite the generality of the idea we have just been discussing, we will consider only real numerical data for most of our study of linear algebra.

Definition

Definition

A **vector** is an ordered finite list of real numbers. A **row vector**, denoted by $[x_1 \ x_2 \ \ldots \ x_n]$, is an ordered finite list of numbers written horizontally. A **column vector**, denoted by $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, is an ordered finite list of real numbers written vertically. Each number x_i making up a vector is called a **component** of the vector.

Remark

We will use *lowercase* bold letters to denote vectors and all other lowercase letters (usually italicized) to represent real numbers.

For example,

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

An alternative, especially prevalent in applied courses, is to use arrows to denote vectors: $\vec{v}$. This form is also used in classroom presentations because it is difficult to indicate boldface with pens or pencils.

Transpose of a Vector

If we have a column vector $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, then its transpose, denoted by $\boldsymbol{x}^T$, is the row vector $[x_1 \ x_2 \ \ldots \ x_n]$.

Similarly, the transpose of a row vector, $\boldsymbol{w} = [w_1 \ w_2 \ \ldots \ w_n]$, denoted by $\boldsymbol{w}^T$ is the column vector $\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$.

Column Vectors and Row Vectors

Consider the following grade sheet for an advanced math class.

Student	Test 1	Test 2	Test 3
Javier	78	84	87
Matthew	85	76	89
Rosemary	93	88	94
Bao	89	94	95
Jennifer	62	79	87

- We can form various vectors from this table. For example, the column vector $\begin{bmatrix} 84 \\ 76 \\ 88 \\ 94 \\ 79 \end{bmatrix}$, formed from the second column of the table, gives all the grades on Test 2.
- Similarly, the row vector $[89 \ \ 94 \ \ 95]$ represents all Bao's test marks, the position (column) of a number indicating the number of the test.

Equality of Vectors

- In working with vectors, it is important that the components be ordered, that is, each component represents a specific piece of information and any rearrangement of the data represents a different vector.
- A more precise way of saying this is that **two vectors**, say $\boldsymbol{x} = [x_1 \ x_2 \ \ldots \ x_n]$ and $\boldsymbol{y} = [y_1 \ y_2 \ \ldots \ y_n]$ are **equal** *if and only if*

$$x_1 = y_1, \ x_2 = y_2, \ \ldots, \ x_n = y_n.$$

Euclidean n-Space

Definition

For a positive integer n, a set of all vectors $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ with components consisting of real numbers is called **Euclidean n-space**, and is denoted $\mathbb{R}^n$ by:

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} : x_i \in \mathbb{R} \text{ for } i = 1, 2, \ldots, n \right\}$$
$$= \left\{ \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T : x_i \in \mathbb{R} \text{ for } i = 1, 2, \ldots, n \right\}$$

Euclidean n-Space (Continue....)

- The positive integer n is called the dimension of the space.
- $\mathbb{R}^1$, or just $\mathbb{R}$, is the set of all real numbers, whereas $\mathbb{R}^2$ and $\mathbb{R}^3$ are familiar geometrical entities: the sets of all points in two-dimensional and three-dimensional "space", respectively.
- For values of $n > 3$, we lose the geometric interpretation, but we can still deal with $\mathbb{R}^n$ algebraically.

Vector Addition and Subtraction

- Suppose the components of vector $\boldsymbol{v_1} = \begin{bmatrix} 322 \\ 283 \\ 304 \\ 292 \end{bmatrix}$ in $\mathbb{R}^4$ represent the revenue (in rupees) from sales of a certain item in a store during weeks 1, 2, 3, and 4, respectively,
- Vector $\boldsymbol{v_2} = \begin{bmatrix} 187 \\ 203 \\ 194 \\ 207 \end{bmatrix}$ represents the revenue from sales of a different item during the same time period.
- If we combine (add) the two vectors of data in a component by component way, we get the total revenue generated by the two items in the same 4 week period:

$$\boldsymbol{v_1} + \boldsymbol{v_2} = \begin{bmatrix} 322 \\ 283 \\ 304 \\ 292 \end{bmatrix} + \begin{bmatrix} 187 \\ 203 \\ 194 \\ 207 \end{bmatrix} = \begin{bmatrix} 322+187 \\ 283+203 \\ 304+194 \\ 292+207 \end{bmatrix} = \begin{bmatrix} 509 \\ 486 \\ 498 \\ 499 \end{bmatrix}.$$

Vector Addition and Subtraction (Continue...)

Definition

If $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, and $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ are two vectors in $\mathbb{R}^n$, then the **sum** of $\boldsymbol{x}$ and $\boldsymbol{y}$ is the vector in $\mathbb{R}^n$, defined by $\boldsymbol{x} + \boldsymbol{y} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix}$, and the difference of $\boldsymbol{x}$ and $\boldsymbol{y}$ is the vector in $\mathbb{R}^n$, defined by $\boldsymbol{x} - \boldsymbol{y} = \begin{bmatrix} x_1-y_1 \\ x_2-y_2 \\ \vdots \\ x_n-y_n \end{bmatrix}$.

Remark

- We can combine vectors only if they have the same number of components, that is, only if both vectors come from $\mathbb{R}^n$ for the same value of n.
- Because the sum (or difference) of two vectors in $\mathbb{R}^n$ is again a vector in $\mathbb{R}^n$, we say that $\mathbb{R}^n$ is closed under addition (or closed under subtraction).
- In other words, we do not find ourselves outside the space $\mathbb{R}^n$ when we add or subtract two vectors in $\mathbb{R}^n$.

Scalar Multiplication

- Suppose the vector $\begin{bmatrix} 18.50 \\ 24 \\ 15.95 \end{bmatrix}$ in $\mathbb{R}^3$ represents the prices of three items in a store.
- If the store charges $7\frac{1}{4}$% sales tax, then we can compute the tax on the three items as follows, producing a "sales tax vector":

$$\text{Sales tax} = 0.0725 \begin{bmatrix} 18.50 \\ 24 \\ 15.95 \end{bmatrix} = \begin{bmatrix} 0.0725(18.50) \\ 0.0725(24) \\ 0.0725(15.95) \end{bmatrix} = \begin{bmatrix} 1.34 \\ 1.74 \\ 1.16 \end{bmatrix}$$

- We can find the "total cost" vector:

Total cost $=$ Price $+$ Sales tax

$$= \begin{bmatrix} 18.50 \\ 24 \\ 15.95 \end{bmatrix} + 0.0725 \begin{bmatrix} 18.50 \\ 24 \\ 15.95 \end{bmatrix} = \begin{bmatrix} 18.50 \\ 24 \\ 15.95 \end{bmatrix} + \begin{bmatrix} 1.34 \\ 1.74 \\ 1.16 \end{bmatrix} = \begin{bmatrix} 19.84 \\ 25.74 \\ 17.11 \end{bmatrix}$$

Scalar Multiplication (Continue...)

Definition

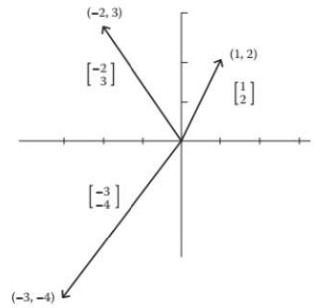
If k is a real number and $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, then **scalar multiplication** of the vector $\boldsymbol{x}$ by the number k is defined by $k\boldsymbol{x} = \begin{bmatrix} kx_1 \\ kx_2 \\ \vdots \\ kx_n \end{bmatrix}$. The number k in this case is called a **scalar** (or **scalar quantity**) to distinguish it from a vector.

Remark

A scalar multiple of a vector in $\mathbb{R}^n$ is again a vector in $\mathbb{R}^n$, we say that $\mathbb{R}^n$ is closed under scalar multiplication.

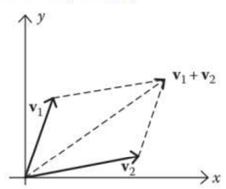
Geometric Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$

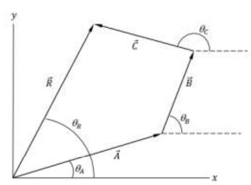
- A vector is a quantity that has both magnitude (size) and direction. Eg: Velocity, acceleration, and other forces
- A vector $\begin{bmatrix} x \\ y \end{bmatrix}$ is interpreted as a directed line segment, or arrow, from the origin to the point (x, y) in the usual Cartesian coordinate plane.



Geometric Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ (Continue...)

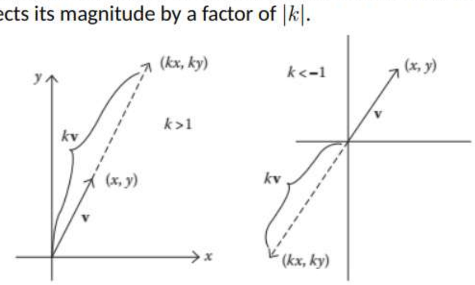
- The magnitude of a vector, a nonnegative quantity, is indicated by the length of the arrow.
- The direction of such a geometric vector is determined by the angle θ which the arrow makes with the positive x-axis.
- The addition of vectors is carried out according to the Parallelogram Law and Polygon Law, and the sum of two vectors is usually referred to as their resultant (vector).





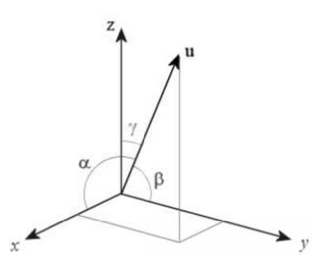
Geometric Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ (Continue...)

- Multiplication of a vector by a positive scalar k does not change the direction of the vector, but affects its magnitude by a factor of k. Multiplication of a vector by a negative scalar reverses the vector's direction and affects its magnitude by a factor of $|k|$.



Geometric Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ (Continue...)

- In $\mathbb{R}^3$, a vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ can be interpreted as an arrow connecting the origin $(0, 0, 0)$ to a point (x, y, z) in the usual $x - y - z$ plane. The operations of addition and scalar multiplication have the same geometric meaning as in $\mathbb{R}^2$.



Algebraic Properties

Theorem - Properties of Vector Operations

If $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ are any elements of $\mathbb{R}^n$ and k, k_1, and k_2 are real numbers, then

1. **Commutativity of Addition:** $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$
2. **Associativity of Addition:** $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$
3. **Additive Identity:** There is a **zero vector**, denoted by $\mathbf{0}$, such that $\boldsymbol{x} + \mathbf{0} = \mathbf{0} + \boldsymbol{x} = \boldsymbol{x}$ for every $\boldsymbol{x} \in \mathbb{R}^n$.
4. **Additive Inverse of $\boldsymbol{x}$:** For any vector $\boldsymbol{x}$, there exists a vector denoted by $-\boldsymbol{x}$ such that $\boldsymbol{x} + (-\boldsymbol{x}) = (-\boldsymbol{x}) + \boldsymbol{x} = \mathbf{0}$, where $\mathbf{0}$ denotes the zero vector.
5. **Associative Property of Scalar Multiplication:** $(k_1k_2)\boldsymbol{x} = k_1(k_2\boldsymbol{x})$.
6. **Distributivity of Scalar Multiplication over Vector Addition:** $k(\boldsymbol{x} + \boldsymbol{y}) = k\boldsymbol{x} + k\boldsymbol{y}$.
7. **Distributivity of Scalar Multiplication over Scalar Addition:** $(k_1 + k_2)\boldsymbol{x} = k_1\boldsymbol{x} + k_2\boldsymbol{x}$
8. **Identity Element for Scalar Multiplication:** $1 \cdot \boldsymbol{x} = \boldsymbol{x} \cdot 1 = \boldsymbol{x}$

The Inner Product and Norm

A Product of Vectors

- Suppose that the vector $[1235 \;\; 985 \;\; 1050 \;\; 3460]^T$ holds four prices expressed in Euros, British pounds, Australian dollars, and Mexican pesos, respectively.
- On a particular day, we know that 1 Euro $=$ \$1.46420, 1 British pound $=$ \$1.83637, 1 Australian dollar $=$ \$0.83580, and 1 Mexican peso $=$ \$0.09294.
- How can we use vectors to find the total of the four prices in U.S. dollars?

A Product of Vectors (Continue...)

Arithmetically, we can calculate the total U.S. dollars as follows:

$$\text{Total} = 1235(1.46420) + 985(1.83637) + 1050(0.83580) + 3460(0.09294) = \$4816.27$$

We can interpret this answer as the result of combining two vectors.

$$\overbrace{\begin{bmatrix} 1235 \\ 985 \\ 1050 \\ 3460 \end{bmatrix}}^{\text{Price vector}} \cdot \overbrace{\begin{bmatrix} 1.46420 \\ 1.83637 \\ 0.83580 \\ 0.09294 \end{bmatrix}}^{\text{Conversion rate vector}}$$
$$= 1235(1.46420) + 985(1.83637) + 1050(0.83580) + 3460(0.09294)$$
$$= \overbrace{\$4816.27}^{\text{Total price (a scalar)}}$$

Dot Product or Scalar Product

Definition

If $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, then the **(Euclidean) inner product** (or **dot product**) of $\boldsymbol{x}$ and $\boldsymbol{y}$, denoted $\boldsymbol{x} \cdot \boldsymbol{y}$, is defined as follows:

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

We can describe the dot product as a function from the set $\mathbb{R}^n \times \mathbb{R}^n$ into the set $\mathbb{R}$.

Dot Product and Linear Equations

- Consider the following equation of a straight line in $\mathbb{R}^2$.

$$3x + 4y = 5.$$

- This can be written using the dot product:

$$\begin{bmatrix} -3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 5.$$

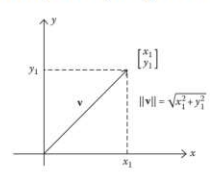
- More generally, any linear equation, $a_1 x_1 + a_2 x_2 + \ldots a_n x_n = b$, can be represented as

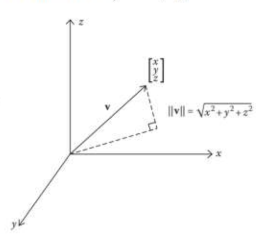
$$\boldsymbol{a} \cdot \boldsymbol{x} = b,$$

where $\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n$ is a vector of coefficients $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a vector of variables (or a vector of unknowns), and b is a scalar.

Introduction to the Norm of a Vector

- If $\boldsymbol{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ is a vector in $\mathbb{R}^2$, then its length, often called the norm of $\boldsymbol{v}$ and written as $\|\boldsymbol{v}\|$.
- This corresponds to the length of the hypotenuse of a right triangle and is given by the Pythagorean theorem as $\sqrt{x^2 + y^2}$.





- Similarly, the length of $\boldsymbol{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in $\mathbb{R}^3$ is $\sqrt{x^2 + y^2 + z^2}$.

Relation between Norm and Dot Product

- In $\mathbb{R}^2$:

$$\|\boldsymbol{v}\| = \left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\| = \sqrt{x^2 + y^2} = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$$

- In $\mathbb{R}^3$:

$$\|\boldsymbol{v}\| = \left\| \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right\| = \sqrt{x^2 + y^2 + z^2} = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$$

Exercise 1

Find the norms of the following vectors:

1. $\boldsymbol{v} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$
2. $\boldsymbol{w} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$

Norm

Definition

If $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is an element of $\mathbb{R}^n$, then we define the **(Euclidean) norm** (or **length**) of $\boldsymbol{x}$ as follows:

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}}$$

The norm is a *nonnegative scalar quantity*.

Properties of the Inner Product and the Norm

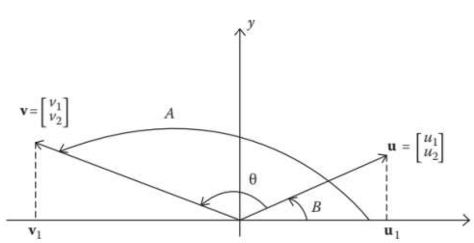
Theorem - Properties of the Inner Product and the Norm

If $\boldsymbol{v}$, $\boldsymbol{v_1}$, $\boldsymbol{v_2}$, and $\boldsymbol{v_3}$ are any elements of $\mathbb{R}^n$ and k is a real number, then

1. $\boldsymbol{v_1} \cdot \boldsymbol{v_2} = \boldsymbol{v_2} \cdot \boldsymbol{v_1}$
2. $\boldsymbol{v_1} \cdot (\boldsymbol{v_2} + \boldsymbol{v_3}) = \boldsymbol{v_1} \cdot \boldsymbol{v_2} + \boldsymbol{v_1} \cdot \boldsymbol{v_3}$ and $(\boldsymbol{v_1} + \boldsymbol{v_2}) \cdot \boldsymbol{v_3} = (\boldsymbol{v_1} \cdot \boldsymbol{v_3}) + (\boldsymbol{v_2} \cdot \boldsymbol{v_3})$
3. $(k\boldsymbol{v_1}) \cdot \boldsymbol{v_2} = \boldsymbol{v_1} \cdot (k\boldsymbol{v_2}) = k(\boldsymbol{v_1} \cdot \boldsymbol{v_2})$
4. $\|k\boldsymbol{v}\| = |k| \, \|\boldsymbol{v}\|$

The Angle between Vectors

Given two vectors $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \neq 0$ and $\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \neq 0$ with the angle θ between them, where the positive direction of measurement is counterclockwise.



Using standard trigonometric formula gives us

$$\cos(\theta) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\| \, \|\boldsymbol{v}\|}.$$

Remark

If $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^2$ and θ is the angle between these two vectors. Then it follows that,

- $-1 \leq \dfrac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\| \, \|\boldsymbol{v}\|} \leq 1.$
- $|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\| \, \|\boldsymbol{v}\|.$

The Cauchy-Schwarz Inequality

Theorem - The Cauchy-Schwarz Inequality

If $\boldsymbol{x} = [x_1 \ x_2 \ \ldots \ x_n]^T$ and $\boldsymbol{y} = [y_1 \ y_2 \ \ldots \ y_n]^T$ are vectors in $\mathbb{R}^n$, then

$$|\boldsymbol{x} \cdot \boldsymbol{y}| \leq \|\boldsymbol{x}\| \, \|\boldsymbol{y}\| .$$

- Equality holds in the Cauchy-Schwarz inequality if and only if one of the vectors is a scalar multiple of the other: $\boldsymbol{x} = c\boldsymbol{y}$ or $\boldsymbol{y} = k\boldsymbol{x}$ for scalars c and k.

Remark

As a consequence of the Cauchy-Schwarz inequality, the formula for the cosine of the angle between two vectors can be extended in a natural way to $\mathbb{R}^3$ and generalized without the geometric visualization to the case of any two vectors in $\mathbb{R}^n$.

Because the graph of $y = \cos(\theta)$ for $0 \leq \theta \leq \pi$ shows that for any real number $r \in [-1, 1]$ there is a unique real number θ such that $\cos(\theta) = r$, we see that there is a unique real number θ such that $\cos(\theta) = \dfrac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\| \, \|\boldsymbol{v}\|}$, $0 \leq \theta \leq \pi$, for any nonzero vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^n$.

The Angle between Two Vectors in $\mathbb{R}^n$

Definition

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are nonzero elements of $\mathbb{R}^n$, then we define the **angle** θ between $\boldsymbol{u}$ and $\boldsymbol{v}$ as the unique angle between 0 and π inclusive, satisfying

$$\cos(\theta) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\| \, \|\boldsymbol{v}\|}$$

Exercise 2

Find the angle between the two vectors $\boldsymbol{u} = \begin{bmatrix} 1 \\ \sqrt{3} \\ 2 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 2\sqrt{3} \\ 2 \\ \sqrt{3} \end{bmatrix}$ in $\mathbb{R}^3$.

Orthogonal Vectors

Definition

Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^n$, are called **orthogonal** (or **perpendicular** if $n = 2, 3$) if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$. In this case, we write $\boldsymbol{u} \perp \boldsymbol{v}$.

- The symbol $\boldsymbol{u} \perp \boldsymbol{v}$ we read this as "$\boldsymbol{u}$ perp $\boldsymbol{v}$".

Exercise 3

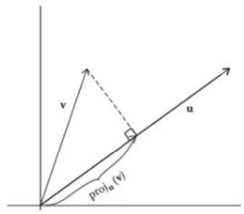
Show that the two vectors $\boldsymbol{u} = \begin{bmatrix} 1 \\ -2 \\ 3 \\ -4 \\ 5 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 10 \\ -4 \\ 1 \\ -1 \\ -5 \end{bmatrix}$ are orthogonal.

Orthogonal Projection

Definition

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors in $\mathbb{R}^n$, and $\boldsymbol{u} \neq \boldsymbol{0}$, then the **orthogonal projection** of $\boldsymbol{v}$ onto $\boldsymbol{u}$ is the vector $\text{proj}_{\boldsymbol{u}}(\boldsymbol{v})$ defined by

$$\text{proj}_{\boldsymbol{u}}(\boldsymbol{v}) = \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\boldsymbol{u} \cdot \boldsymbol{u}}\right) \boldsymbol{u}$$



Exercise 4

Find the vector projection of vector $\boldsymbol{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ onto vector $\boldsymbol{u} = \begin{bmatrix} 5 \\ -12 \end{bmatrix}$.

Cross Product or Vector Product

- The dot product is a multiplication of two vectors that results in a scalar.
- Now, we introduce a product of two vectors that generates a third vector orthogonal to the first two.
- Let $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ be nonzero vectors. We want to find a vector $\boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ orthogonal to both $\boldsymbol{u}$ and $\boldsymbol{v}$.
- That is, we want to find $\boldsymbol{w}$ such that $\boldsymbol{u} \cdot \boldsymbol{w} = 0$ and $\boldsymbol{v} \cdot \boldsymbol{w} = 0$.
- Therefore, the vector $\boldsymbol{w} = \begin{bmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{bmatrix}$.

Cross Product (Continue...)

Definition

Let $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ be nonzero vectors. Then, the **cross product** $\boldsymbol{u} \times \boldsymbol{v}$, is defined as follows:

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{bmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{bmatrix}.$$

Exercise 5

Let $\boldsymbol{u} = \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix}$. Find $\boldsymbol{u} \times \boldsymbol{v}$.