Linear Algebra

AIMA 104

Department of Artificial Intelligence,
Kathmandu University, Dhulikhel, Kavre

Lecture notes by
Sandesh Thakuri
Sandesh.775509@cdmath.tu.edu.np

Text Book:

1. David C. Lay, Linear Algebra and Its Applications (4th edition), Pearson.
2. Gilbart Strang, Introduction to Linear Algebra, Wellesley-Cambridge Press.

2025

Contents

1 Vector Space

1.1 Introduction

The branch *linear algebra* is much younger than the branch *calculus*. To give a clear introduction of the course linear algebra, we need to look at it from at least three sides, and consequently we can say that we can approch it from the following three sides.

1. **Representation System**

   Linear Algebra is the algebra for Multivariable Calculus. The vectors and matrices are the building block of linear algebra. And linear algebra consists of the algebra of vectors and matrices, just like ordinary algebra is the algebra of numbers and their variables. Consider a multivariable function $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ defined by $f(x_1, x_2) = 3x_1 + 2x_2$. The consise representation of the multivariable functions give rise to vectors and matrices as follows:
   Let, $\vec{x} = (x_1, y_1)$ and $\vec{a} = (3, 2)$. Then the above function is f is represented as $f(\vec{x}) = \vec{a}'.\vec{x}$.

2. **Study of Linear System**

   Linear Algebra has its roots in solving a system of linear equations. Solving a linear system made a great contribution in Matrix theory and Determinant theory. Gauss elimination method and Gauss-Jordan method contributed to matrix theory while Cramers' method contributed to determinant theory.

3. **Study of Linear Map**

   Another approch to study linear algebra is study of linear maps. It can be said that it is from this linear algebra got its name. The linear maps eventually settles down to the maps by matrices, and the dimension of range is the dimension of column space of the matrix in the linear map and so on.

1.2 Vector Space

We saw the consise representation of multivariable functions leads to representation by vectors and matrices. The set of vectors gives rise to vector spaces. For a single variable calculus we need a **number-system** like $\mathbb{R}$. For a multivariable calculus we need a system of n-tuples, like the system of *order pairs*. This system is developed into a vector space. Just like a number system is need for a single variable calculus, a vector space needed for a mulltivariable calculus for its structures; vector space like $\mathbb{R}^2$. And the algebra need for this calculus is given by the linear algebra.

Definition 1 (**Vector Space**)**.** A vector space is a non-empty set V of objects, called vectors, over a scalr field $\mathbb{F}$, on which are defined two operations, called *addition* and *multiplication by scalars* subject to the **ten** axioms listed below. For all $u, v, w \in V$ and for all scalrs $\alpha, beta \in \mathbb{F}$.

i) $u + v \in V$

ii) $u + v = v + u$

iii) $(u + v) + w = u + (v + w)$

iv) $0 \in V \ : \ u + 0 = 0$

v) $-u \in V \ : \ u + (-u) = 0$

vi) $\alpha u \in V$

vii) $\alpha(u + v) = \alpha u + \alpha v$

viii) $(\alpha + \beta)u = \alpha u + \beta u$

ix) $\alpha(\beta u) = (\alpha\beta)u$

x) $1u = u$

Example 1. The space $\mathbb{R}^2, \mathbb{R}^3$ are some common examples of vector spaces over $\mathbb{R}$.

Example 2. The set of all functions defined on an interval $[a, b]$ in a real line forms a vector space over $\mathbb{R}$.

1.3 Systems of Linear Equations

A completely general system of m linear equations in n unknowns is of the following form:

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

The a_{1j} is the coefficient of x_j in the ith equation. The data for this system of equations are all the numbers a_{ij} and b_i. Now consider the four matrices.

$$
A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & & & \\ . & & & \\ . & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{bmatrix}, \qquad [A|b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{array}\right]
$$

In this context of a system of equations; A is called the coefficient matrix, x is called vector of unknowns, b is called righthandside vector, $[A|b]$ is called the augmented matrix. So, the **matrix notation** of the system of linear equations is $Ax = b$.
Homework: Solve the following using matrix notation:

$$\begin{aligned} x_1 - 3x_2 + 4x_3 &= -4 \\ x_1 + 3x_2 + 5x_3 &= -2 \\ x_1 + 7x_2 + 7x_3 &= 6 \end{aligned}$$

1.3.1 Elementary Row operations

Basically we have following three elementary row operations.

S.N	Operation	Description	Notation
1.	replacement	Add a multiple of one row to another.	$r_i \leftarrow r_i + ar_j \quad (i \neq j)$
2.	scale	Multiply a row by a nonzero factor.	$r_i \leftarrow cr_i \quad (c \neq 0)$
3.	swap	Interchange a pair of rows.	$r_i \leftrightarrow r_j$

1.4 Echelon Form

1.4.1 Reduced Row Echelon Form

With the help of the elementary row operations, any matrix can be transformed into a standard form called reduced row echelon form.

A matrix is in **reduced row echelon form** if

- All zero rows have been moved to the bottom of matrix.
- Each nonzero row has 1 as its leading nonzero entry, using left-to-right ordering. Each such leading 1 one is called a **pivot**.
- In each column containing a pivot, there is no other nonzero elements.
- The pivot in any row is farther to the right than the pivots in rows above.

1. Here are four examples of matrices in reduced row echelon form:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 5 & 0 & -7 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

2. Here are four matrices not in reduced row echelon form.

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

1.4.2 Row echelon form

A matrix is in **row echelon form** if

- All zero rows have been moved to the bottom of matrix.
- The leading nonzero element in any row is farther to the right than the pivots in rows above.
- In each column containing a leading nonzero element, the entries below that leading nonzero element are zero.

1. Notice the **staircase pattern** of the pivot positions.
2. A row echelon form is obtain with less work than is required for the reduced row echelon form.
3. The reduced row echelon form of a matrix is *unique*, whereas a matrix may have many forms of row echelon forms.

Definition 2. A **pivot position** in a matrix is a location where a leading 1 (a pivot) appears in the reduced row echelon form of that matrix. In general, we do not know the pivot positions until we have found the reduced row echelon form of the matrix, or any row echelon form.

1.4.3 Algorithm for the Reduced Row Echelon Form

1. Interchange the rows if necessary to place all zero rows on the bottom.
2. Identify the leftmost nonzero column. Say it is pivot column j. Interchange rows to bring a nonzero element to the top row and jth column, which is the pivot position. Use the row replacement operation to create zeros in all positions in the pivot column below the pivot position.
3. Repeat Steps 1 and 2 on the remaining submatrix until there are no nonzero rows left. (*We have a row echelon form.*)
4. Beginning with the rightmost pivot, working upward and to the left, use row replacement operations to create zeros in all positions in the pivot column above the pivot position. Scale the entry in the pivot row to create a leading 1.
5. Repeat Step 4, ending with the unique reduced row echelon form of the given matrix.

1.4.4 Exercise

1. Solve: $3x_1 + 6x_2 + 6x_3 = 21, \quad 2x_1 + 4x_2 + 5x_3 = 16, \quad 2x_1 + 5x_2 + 4x_3 = 17$
2. Find all solutions: $x - y + z = 4, \quad 2x + y - 3z = 5, \quad -y + 7x - 3z = 22$

3. Find the general solution of the system, $x_1 + 3x_2 + 9x_3 = 6, \quad 2x_1 + 7x_2 + 3x_3 = -5, \quad x_1 + 4x_2 - x_3 = -11$
4. Solve: $x_2 + 4x_3 = -5, \quad x_1 + 3x_2 + 5x_3 = -2, \quad 3x_1 + 7x_2 + 7x_3 = 6.$
5. Find the value(s) of h such that the augmented matrix is of a consistent system.
$\begin{bmatrix} 1 & h & 4 \\ 3 & 6 & 8 \end{bmatrix} \qquad \begin{bmatrix} 2 & -3 & h \\ -6 & 9 & 5 \end{bmatrix}$

1.4.5 Uniquely and Parametrically represented solutions.

In section we discuss how to tell whether a system of linear equations has a unique solution or many solutions which are parametrically represented solutions or has no solution. The case of having no solution is the case of inconsistency of the system, based on its reduced echelon form.

Definition 3 (Rank of a matrix)**.** The rank of a matrix is the number of nonzero rows in its reduced row echelon form or row echelon form. We use the notation Rank(A) for this number. So, the rank of a matrix is equal to the number of its pivots.

So, it also defined as the number of pivot positions in the matrix.

Remark. It is not necessary to carry out the reduction to reduced row echelon form to determine the pivot positions in a matrix. Reduction to row echelon form is sufficient for this.

Remark. This rank of a matrix is also called its **row rank** or its **column rank**.

Let A be $m \times n$ matrix.

1. The rank of the coefficient matrix A equal to n then, it has a unique solution.
2. The rank of the coefficient matrix A equals to the rank of $[A : b]$, but is less than n then, it has more than one solutions.
3. The rank of A is less than the rank of $[A : b]$ then, the system of equation is inconsistent. This is the case where the rank of A is less than n but there is a pivot position in the last column of the augmented matrix, $[A : b]$.

More than one solution:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 20 \\ 4 & 5 & 6 & 47 \\ 7 & 8 & 9 & 74 \\ 10 & 11 & 12 & 101 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 11 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 11 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Inconsistent System

$$\left[\begin{array}{cc|c} 2 & -4 & 3 \\ 4 & -1 & 2 \\ 1 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{array}\right]$$

1.4.6 Exercise

1. Find the rank of coefficient and augmented matrix and check the consistency of the given system of equations:

   a). $2x - 6y + 8z = 2, \quad -4x + 13y + 3z = 6, \quad -6x + 20y + 14z = -2$

   b). $2x - 2y + 4z + 6w = 8, \quad -4x + 5y - 2z - 7w = -10, \quad 2x + y + 22z + 21w = 10, \quad -3x + 5y - 4z + 11w = 10$

2. Obtain the row rank and parametrically represented solution of the following systems.

   a). $-4x + 12y - 7z = 8, \qquad x - 3y + 2z = -1$

   b). $3x_1 - x_2 + 3x_3 = 5, \quad x_1 + 2x_2 - 3x_4 = -1, \quad 2x_1 + 5x_2 + 4x_3 + 2x_4 = 10.$

3. Choose h and k such that the system has (a) no solution (b) a unique solution (c) many solutions.
   $x_1 + hx_2 = 2, \quad 4x_1 + 8x_2 = k$

1.5 Linear Dependence and Independence

1. **Linear Combination of vectors**

   A linear combination of the vectors
   $\vec{u_1}, \vec{u_2}, ..., \vec{u_m}$ is a sum of the vectors multiplied by scalars, such as

   $$\alpha_1\vec{u_1} + \alpha_2\vec{u_2} + ... + \alpha_n\vec{u_m} = \sum_{j=1}^{m} \alpha_j\vec{u_j}$$

2. **Span of Vectors**

   The collection of all linear combinations of vectors in the given set is called the span of that set of vectors. If the set is S, its span is denoted by Span(S).

   Remark. The span of the vectors $\{\hat{i}, \hat{j}\}$ is the set $\{x\hat{i} + y\hat{j} \ : \ x, y \in \mathbb{R}\}$ which is $\mathbb{R}^2$. So, the geometry of the span of two vectors is a plane and the span of a single vector is line which will be discussed in the next section

3. **Linear Dependence**

   Consider a finite *indexed* set of vectors $\{u_1, u_2, ..., u_m\}$ in a vector space.

   - We say that the indexed set is **linearly dependent** if there exist scalars c_i such that $\sum_{i=1}^{m} c_i u_i = 0, \quad \sum_{i=1}^{m} |c_i| > 0.$
   - If the indexed set is not linearly dependent, we can say that it is linearly **independent**. The expression implies that at least one c_i is nonzero.

   There is a difference between an index set and a set. The set $\{\vec{i}, \vec{i}\}$ is linearly independent because it consist of only one vector $\vec{i}$. but the index set $\{\vec{i}, \vec{i}\}$ is linearly dependent because this set consists of two vectors both of which are same.

1.5.1 Exercise

1. Determine if b is a linear combination of a_1 and a_2, and a_3.
$$a_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \qquad a_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \qquad a_3 = \begin{bmatrix} 5 \\ -6 \\ 8 \end{bmatrix}, \qquad b = \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix}$$

2. Is the indexed set of rows in the matrix linearly dependent? $\begin{bmatrix} 2 & 5 & 7 \\ 4 & 1 & -5 \\ 2 & 5 & 7 \end{bmatrix}$

3. Determine whether this set of vectors is linearly dependent:
$$(3, 2, 7), \quad (4, 1, -3) \quad (6, -1, -23)$$

1.6 Basis and Dimension

1.6.1 Basis

In a vector space V, a linearly independent set of vectors that spans V is called a **basis** for V. In another words, it is the minimum collection of vectors of vectors that generate the given vector space.

Example 3. Any set of two linearly independent vector is a basis of $\mathbb{R}^2$, say $\mathcal{B} = \{(1, 1), (1, 3)\}$ is a basis of $\mathbb{R}^{\nvDash}$. The set $\{\hat{\imath}, \hat{\jmath}\}$ in $\mathbb{R}^2$, is a basis of $\mathbb{R}^2$, where $\hat{\imath} = (1, 0)$ and $\hat{\jmath} = (0, 1)$. As this set is linearly independent and any vector $\vec{x} \in \mathbb{R}^2$ can be expressed as a linear combination of the vectors $\hat{\imath}$ and $\hat{\jmath}$. This basis of $\mathbb{R}^2$ is called the **standard basis** of $\mathbb{R}^2$. Similarly the standard basis of $\mathbb{R}^3$ is $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$.

Question: What is the standard basis of $\mathbb{R}^4$?

Theorem 1. If a vector space has a finite basis, then all of its bases have the same number of elements. And this number is called the dimension of the vector space.

For instance, the dimension of a vector space having the number of basis vectors, three, has the dimension 3. So, the dimension of $\mathbb{R}^3$ is 3.

Definition 4. A vector space is **finite dimensional** if it has a finite basis; in that event, its **dimension** is the number of elements in any of its basis. Thus the vector space $\mathbb{R}^n$ is finite dimensional with the dimension n.

1.7 Vector Subspaces

A subspace of a vector space V is a subset H of V that is a vector space in itself with the same operations of V. Mathematically, a subset H of V is a vector subspace of V if

1. H is closed under vector addtion, i.e for all $u, v \in H, \quad u + v \in H$.

2. H is colsed under multiplication by scalars, i.e for all $u \in H$ and for each scalar α, $\alpha u \in H$.

Example 4. If $H = \{0\}$ then H is a subspace of V.

Example 5. The set $X = \{(x, 0) : x \in \mathbb{R}\}$ is a vector subspace of $\mathbb{R}^2$.

Example 6. The Column Space, the Null Space are the important vector subspaces that will discussed in the coming sections.

1.7.1 Null Space

The **null space** of a matrix A is the space $\{x : Ax = 0\}$. It is denoted by Null(A). Null space of a matrix A is also called the *kernal* of A which is denoted by Ker(A). The Null space of a matrix A **implicit** in nature, meaning there is no obvious relation between Null(A) and the entries in A. So, to produce the explicit description of Null(A), the equation $Ax = 0$ is solved.

1.7.2 Column Space and Null Space of a Matrix

1. The **column space** of a matrix A is the span of the set of columns in A. This is denoted by Col(A).

2. Let R be the reduced row echelon matrix of the matrix A. Let the columns of A whose corresponding columns in R have pivots be $c_1, c_2, ..., c_n$. Then Col(A)=Span$\{c_1, c_2, ..., c_n\}$. That is Col(A) is given by the span of the columns from A that have pivot positions.

3. The Column space of a matrix A is **explicit** in nature, meaning there is an obvious relation between Col(A) and the entries in A.

1.7.3 Exercise

1. Find the explicit description of Null(A), and hence find bases of Null(A), and Col(A).

   a). $A = \begin{bmatrix} 1 & -2 & 0 & 4 & 0 \\ 0 & 0 & 1 & -9 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ b). $A = \begin{bmatrix} 1 & 3 & 5 & 0 \\ 0 & 1 & 4 & -2 \end{bmatrix}$

2. Determine if $u = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$ belongs to the null space of $A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}$.

3. Find a simple description of the column space of the following matrix $\begin{bmatrix} 1 & 3 & 2 & 4 \\ 1 & 0 & 4 & -2 \\ 2 & 2 & 1 & 7 \\ 4 & 5 & 7 & 9 \end{bmatrix}$

1.7.4 Rank

Definition 5 (Rank of a matrix)**.** The rank of a matrix is the number of pivots in it. We use the notation Rank(A) for this number. So, it also defined as the number of pivot positions in the matrix.

Theorem 2 (Rank Nullity Theorem)**.** For any matrix, the number of columns equals the dimension of the column space plus the dimension of the null space. For a $m \times n$ matrix we have,

$$Dim(Col(A)) \; + \; Dim(Null(A)) \; = \; n.$$

In other words,

$$Rank(A) \; + \; Nullity(A) \; = \; n.$$

Because the dimension of column space of A is equivalent to the rank of A, as rank of A is the number of pivots in the row echelon form of A. And Nullity is the defined as the dimension of null space of A.

1.7.5 Exercise

1. Verify Rank Nullity Theorem: $\begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \\ 7 & 8 & 9 \end{bmatrix}$
2. If a 3×8 matrix A has rank 3, find $Dim(Col(A)), \;\; Dim(Null(A)), \;\; Dim(Row(A))$.
3. Suppose a 4×7 matrix A has four pivot columns. Is $Col(A) = \mathbb{R}^4$? Is $Nul(A) = \mathbb{R}^3$? Explain your answer.

1.8 Some Insights (Extra Material)

Let a system of linear equations be represented by a matrix equation $Ax = b$. Then,

1. the system is **homogeneous system**, if the right-hand side vector b is the zero vector. i.e $Ax = 0$. Note that this system is always consistent because this system always has a solution $x = 0$, which is called a trival solution. The question is *When does the system has a non-trival solution (a nonzero solution)?* This answered at the end of this subsection.
   It can latter be shown that whenever a homogeneous system has a non-trivial solution than it is the case of **infinitely many solutions**. And a homogeneous system of consisting of more variables than the number of equations always has *infinitely many solutions.*

2. the system is **non-homogeneous system**, if the right-hand side vector b is a non-zero vector. i.e $Ax = b, \;\; b \neq 0$.

1.8.1 Interpretation of Existence of a Solution of the system

Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$.

1.8.2 Interpretation of Existence of a Solution of the system

The given system $Ax = b$ is

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \tag{1.1}$$

If the system $Ax = b$ has a solution then there exists some values of x_1, x_2, x_3 such that the Equation-1.1 is true.

1. This means the right-hand side vector b is in the spanned by the columns of A. That means the right hand side vector b is in the column space of A, i.e $b \in Col(A)$.

2. if the vector b is the zero vector then the solution vector, then $Ax = b$ has a solution $x_0 \neq 0$ means the solution vector $x_0 \in Null(A)$. So, the homogeneous system has more nonzero solution when the nullity of A greater than 1, that happens when the rank of A is less than n, for a $m \times n$ matrix A.

1.9 Extra Exercises

pg 64, Qn 11. Find the genera flow pattern of the network.

2 Linear Transformation

2.1 Introduction

The terms *function*, *mapping*, *map*, and *transformation* are synonymous. Great part of the linear algebra is dedicated to the study of linear transformation.

Definition 6. A mapping T from a vector space V to a vector space W is **linear** if

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$$

for all vectors $x, y \in V$ and for all scalars α, β.
Equivalently, T **linear** if,

i) $T(x + y) = T(x) + T(y)$ for all vectors $x, y \in V$, and

ii) $T(\alpha x) = \alpha T(x)$ for any scalars α.

Example 7. Show that if A is a $m \times n$ matrix then T defined by $T(x) = Ax$ is a linear transformation from $\mathbb{R}^n \to \mathbb{R}^m$.

2.1.1 Exercise

1. Is the map $f : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $f(x_1, x_2, x_3) = (x_1 + (x_2, 3x_1 - x_2 + x_3, 5x_1 - x_3)$ a linear map? Explain.

2. Show that shift map is not a linear transformation.

3. Is there a linear transformation that maps $(1, 0)$ to $(5, 3, 4)$ and maps $(3, 0)$ to $(1, 3, 2)$?

4. Show that the transformation T defined by $T(x_1, x_2) = (4x_1 - 2x_2, 3|x_2|)$ is not linear.

5. T is defined by $T(x) = Ax$, find a vector x whose image under T is b, determine whether x is unique.

$$A = \begin{bmatrix} 1 & 0 & -2 \\ -2 & 1 & 6 \\ 3 & -2 & 5 \end{bmatrix}, \qquad b = \begin{bmatrix} -1 \\ 7 \\ -3 \end{bmatrix}$$

6. How many rows and columns must a matrix A have in order to define a mapping from $\mathbb{R}^4$ to $\mathbb{R}^5$ by the rule $T(x) = Ax$?

2.2 Matrix of a Linear Transformation

Theorem 3. Let A be an $m \times n$ matrix. The mapping $x \mapsto Ax$ is linear from $\mathbb{R}^n$ to $\mathbb{R}^m$. Conversely, for a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ there exists an $m \times n$ matrix A such that $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$. In fact, A is the $m \times n$ matrix whose jth column is the vector $T(e_j)$, where $\{e_j : j = 1, 2, ..., n\}$ is the the basis of the domain $\mathbb{R}^n$,i.e $A = [T(e_1) \; ... \; T(e_n)]$.

2.2.1 Exercise

Find the standard matrix of the linear transformation T.

1. For the dilation transformation $T(x) = 3x$, $x \in \mathbb{R}^2$.
2. Show that the transformation that rotates each point in $\mathbb{R}^2$, about the origin, through an angle ϕ counterclockwise, is a linear transformation by find the standard matrix of the transformation.
3. $T(x_1, x_2) = (2x_2 - 3x_1, x_1 - 4x_2, 0, x_2)$
4. $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ first reflects points through the x_1-axis then reflects points through the line $x_2 = x_1$
5. $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ that rotates each point through an angle ϕ, with counterclockwise rotation.
6. Describe the transformation of the following matrices geometrically:

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

Example 8. Describe the linear mapping that has the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Let the given matrix is denoted by A. A is a 2×2 So, the map is from $\mathbb{R}^2 \mapsto \mathbb{R}^2$. Then for every $\vec{x} = (x_1, x_2) \in \mathbb{R}^2$, Now $A\vec{x} = (-x_2, x_1)$ which is a **rotation** transformation that rotates every vector through the angle 90^0 counterclockwise.

2.2.2 Types of Transformation

Geometrically, linear transformation are basically of the following types:

1. Reflection
2. Rotation
3. Contraction and Expansion

4. Shears: Horizontal and Vertical

5. Projections

2.3 Kernel and Image of Linear Transformation

Let $T : V \mapsto W$ be a linear transformation.

- **Kernal**

  The kernal of T is the set of all vectors in V that maps to zero vector in W. It is denoted by *Ker(T)*. $Ker(T) = \{v \in V\,;\, T(v) = 0\}$.In another words, Kernal is the *Null Space* of T.

- **Range**

  The range of T is the set of all vectors in W which are the images of the vectors in V. It is denoted by $R(t)$. $R(T) = \{w \in W\ :\ w = T(v)\text{for some}v \in V\}$.

2.3.1 Exercise

1. Show that $Ker(T)$ is a subspace of V and $R(T)$ is a subspace of W.

2. Given the vector space V of all real-valued functions defined on an interval $[a, b]$ such that their first derivative functios are continous on $[a, b]$. Let W be the vector space of all continuous functions of $[a, b]$. Show that $D : V \mapsto V$ that maps $f \in V$ to $f' \in w$ is a linear transformation and find the kernal of D.

2.3.2 Facts:

1. The linear transformation $T : V \to W$ is one-to-one if, $dim(Ker(T) = 0$,
   i.e $Ker(T) = 0$.

2. The linear transformation $T : V \to W$ is onto if, $dim(R(T)) = dim(W)$.

2.4 Properties of Linear Transformation

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

1. Then $T(0) = 0$.

2. T is one-to-one if and only if the equation $T(x) = 0$ has only the trivial solution. (*Indirectly it has got to do with the nullity or nullspace of the matrix.*)

3. T is onto if and only if the columns of the corresponding standard matrix spans the $\mathbb{R}^m$. (*So it has got to do with the column space of the matrix.*)

Give an example of one-to-one and not one-to-one linear transformation.

3

Eigenvectors, Eigenvalues and Diagonalization

3.1 Introduction

Let A be any square matrix, real or complex. A number λ is an **eigenvalue** of A if the equation

$$Ax = \lambda x$$

is true for some nonzero vector x. The vector x is and **eigenvector** associated with the eigenvalue λ. Both the eigenvalue and the eigenvector may be complex.

3.1.1 Exercise

1. Is $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} -3 & 1 \\ -3 & -8 \end{bmatrix}$?

2. Is $\lambda = 2$ a eigenvalue of a. $\begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$?

Theorem 4. A scalar λ is an eigenvalue of a matrix A if and only if $Det(A - \lambda I) = 0$.

The equation $Det(A - \lambda I) = 0$ is called the **characteristic equation** of A. It is the equation from which we can compute the eigenvalues of A. The function $p : p(\lambda) = Det(A - \lambda I)$ is the **characteristic polynomial** of A.

3.1.2 Eigenspace

For an eigenvalue λ of a matrix A, the set $\{x : Ax = \lambda x\}$ forms a vector space. This forms a vector space because the vector x is a nonzero vector for it be an eigenvector. If x is a nonzero solution of $Ax = \lambda x \implies (A - \lambda I)x = 0$, which is a homogeneous system, then this homogeneous system has infinitely many solution. And this vector space is called eigenspace.

3.1.3 Exercise

1. What are the characteristic equation and the eigenvalues of the following matrices? For each eigenvalue, find an eigenvector.

   a. $\begin{bmatrix} 2 & 4 & 6 \\ 0 & -3 & 5 \\ 0 & 0 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 4 & 1 & 1 \\ 2 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix}$ c. $\begin{bmatrix} 2 & -i & 0 \\ i & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

2. Find a basis for the eigenspace corresponding to each eigenvalue.

   a. $\begin{bmatrix} 5 & 0 \\ 2 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 10 & -9 \\ 4 & -2 \end{bmatrix}$ c. $\begin{bmatrix} 4 & 0 & 0 \\ -2 & -1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$

3. Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Find the eigenvalue-eigenvector pairs. Explore the geometric effect of letting $x^{(k)} = Ax^{(k-1)}$ and $k = 0, 1, 2, ...$

4. Prove that A and A^t have the same eigenvalues.

3.1.4 Important Results

Theorem 5. The eigenvalues of a triangular matrix are the entries on its main diagonal.

Theorem 6. If $v_1, ..., v_r$ are eigenvectors that corresponds to distinct eigenvalues $\lambda_1, ..., \lambda_r$of an $n \times n$ matrix A, then the set $\{v_1, ..., v_r\}$ is linearly independent.

Theorem 7. Zero is an eigenvalue of A if and only if A is not invertible.

3.2 Diagonalization

A square matrix A is said to be diagonalizable if, there exists an invertible matrix P and a diagonal matrix D, such that

$$A = PDP^{-1}.$$

If $A = PDP^{-1}$ then A is said to be similar to D.
This is standardize as if, A is similar to a diagonal matrix.

If $A = PDP^{-1}$ then prove that $A^k = PD^kP^{-1}$.

3.2.1 The Diagonalization Procedure

Suppose A is an $n \times n$ matrix with n different eigenvalues: $\lambda_1, \lambda_2, ..., \lambda_n$ so that the set of n corresponding eigenvectors $v_1, v_2, ..., v_n$ are linearly independent. Now, $Av_1 = \lambda_1 v_1, Av_2 = \lambda_2 v_2, ..., Av_n = \lambda_n v_n$. Let P be the matrix whose columns are the eigenvectors, i.e $P = [v_1,\ v_2,\ ...\ v_n]$. Then

$$AP = [\lambda_1 v_1,\ \lambda_2 v_2,\ ...\ \lambda_n v_n] = [v_1,\ v_2,\ ,...,\ v_n] \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & & & \\ 0 & 0 & ... & \lambda_1 \end{bmatrix} = PD \tag{3.1}$$

where $D = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & & & \\ 0 & 0 & ... & \lambda_1 \end{bmatrix}$, so that $AP = PD$ gives, $A = PDP^{-1}$. This is summarized as the following theorem.

Theorem 3 (**The Diagonalization Theorem**)
An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors of P.

Theorem 4. An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

3.2.2 Exercise

1. Diagonalize the matrices, if possible:

   i). $\begin{bmatrix} 3 & -1 \\ 1 & 5 \end{bmatrix}$ iii). $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$

   ii). $\begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$ iv). $\begin{bmatrix} 2 & 2 & -1 \\ 1 & 3 & -1 \\ -1 & -2 & 2 \end{bmatrix}$

2. Compute A^8, where $A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$

3.3 Matrices Whose Eigenvalues are not Distinct

If $n \times n$ A has n distinct eigenvalues, with corresponding eigenvectors, then A is automatically diagonalizable. Now, we will look at a case where A has fewer than n distinct eigenvalues, and it is still possible to diagonalize A.

Theorem 5. Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \lambda_2, ...\ \lambda_p$.

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

b. The matrix A is diagonalizable if and only if the sum of the dimensions of the distinct eigenspaces equals n. And this happens if and only if the dimension of the eigenspace for each λ_k equals to the multiplicity of λ_k.

c. If A is diagonalizable and $\mathcal{B}_k$ is a basis for the eigenspace corresponding to λ_k for each k, then the total collection of vectors in the sets $\mathcal{B}_1, \mathcal{B}_2, ..., \mathcal{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$.

3.3.1 Multiplicities of an eigenvalue

- The algebraic multiplicity of λ is the number of times $(\lambda - t)$ occurs as a factor of in its characteristic polynomial.
- The geometric multiplicity of an eigenvalue λ is the dimension of the corresponding eigenspace E_λ. i.e $dim\{E_\lambda\}$.

3.3.2 Exercise

Diagonalizable the following matrix, if possible.

1. $A = \begin{bmatrix} 4 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

2. $\begin{bmatrix} 4 & -9 \\ 4 & 8 \end{bmatrix}$

3.3.3 Similarity

If A an B are $n \times n$ matrices, then A **similar to** B if there is an invertible matrix P such that

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

or equivalently $B = P^{-1}AP$. Writing $Q = P^{-1}$, we can say that B is similar to A, and we simply say A and B are similar.

Theorem 6. If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

Proof. Given, $B = P^{-1}AP$ we have, $(B - \lambda I) = P^{-1}AP - \lambda P^{-1}P = P^{-1}(A - \lambda I)P$. Then

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

As, $\det(P^{-1}).\ \det(P) = \det(P^{-1}.P) = det(I) = 1$. Hence A and B has same characteristic polynomial. And the proof follows. □

3.4 Linear Transformation and Diagonalization

We will see that the transformation $x \to Ax$ is essentially same as the very simple mapping $x \to Dx$ when viewed from the proper perspective.

Theorem 7. Suppose $A = PDP^{-1}$, where D is a diagonal $n \times n$. If B is the basis for $\mathbb{R}$, formed from the columns of P, then D is the $\mathcal{B}$-matix for the transformation $x \mapsto Ax$.

Proof. Let the order of matrix A be n. Since A is diagonalizable, A has n linearly independent eigenvectors: $\mathcal{B} = \{v_1, ..., v_n\}$, that forms a basis of $\mathbb{R}^n$. Suppose for $x = x_1e_1 + ... + x_ne_n$ we have $x = x'_1v_1 + ... + x'_nv_n$.

$$
\begin{aligned}
Ax = A(x'_1v_1 + ... + x'_nv_n) &= x'_1Av_1 + ... + x'_nAv_n \\
&= [Av_1 \; ... \; Av_n](x'_1 \; ... \; x'_n)^t \\
&= [\lambda_1v_1 \; ... \; \lambda_nv_n](x'_1 \; ... \; x'_n)^t \\
&= \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & & & \\ 0 & 0 & ... & \lambda_1 \end{bmatrix} x \qquad (3.2) \\
&= Dx
\end{aligned}
$$

Here, $[\lambda_1v_1 \; ... \; \lambda_nv_n] = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & & & \\ 0 & 0 & ... & \lambda_1 \end{bmatrix}$. This is because for $(x'_1 \; ... \; x'_n)^t$, the vectors $(v_1, ..., v_n)$ are the unit vectors, just like $e_1, ..., e_n$ are for $x_1, ..., x_n$. ☐

Remark. This proof is not need for the students.

3.4.1 Exercise

Find a basis $\mathcal{B}$ for $\mathbb{R}^2$ with the property that the transformation matrix of the transformation given by the following matrix is Diagonal.

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

4

Orthogonality and Least Squares

4.1 Inner Product

Let $u = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ . \\ v_n \end{bmatrix}$ be any two vectors in $\mathbb{R}^n$. Then the number $u^T v$ is called the **inner product** of u and v. This inner product is also commonly known as **dot product** and denoted by **u**.**v**.

4.1.1 Properties of Inner Product

1. $u.v = v.u$
2. $(u + v).w = u.w + v.w$
3. $(\alpha u).v = \alpha(u.v) = u.(cv)$
4. $u.u \geq 0$ and $u.u = 0 \iff u = 0$

4.1.2 The Length of a Vector

The length of a vector v is called the **norm** of v.
It is denoted by $\|v\|$ and defined by $\|v\| = \sqrt{v_1^2 + v_2^2 + ... + v_n^2}$ so that, $\|v\|^2 = v.v$ There are several kinds of norms actually, this particular norm is called **Euclidean norm**. For any scalar α, $\|\alpha v\| = |\alpha|\|v\|$. A vector whose length is unity is called a **unit vector**. If we divide a nonzero vector v by its length, we obtain a unit vector u. This process is called **normalizing** of the vector v.

Distance between vectors

For u and v in a vector space V, the distance between them is written as dist(u,v) and is defined as $dist(u, v) = \|u - v\|$.

4.1.3 Exercise

1. Compute the following for the given vectors:
$u = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, v = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, w = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, x = \begin{bmatrix} 6 \\ -2 \\ 3 \end{bmatrix}$

   a). $u.u,\ \ v.u,\ \ \dfrac{v.u}{u.u},\ \ \|v\|$ b). $w.w,\ \ x.w,\ \ \dfrac{x.w}{w.w},\ \ \|x\|$

2. Find the distance between u and v, and w and x.

3. Use matrix product and transpose definition to verify, property-2 and 3 of the inner product.

4. Explain why $u.u \geq 0$. When is $u.u = 0$.?

4.2 Orthogonal Vectors

The two vectors u and v are orthogonal vectors if their dot product is zero,i.e $u.v = 0$. Observe that the zero vector is orthogonal to every vector as $0^T v = 0$ for all v.

Theorem 4 (The Pythagorean Theorem)
Two vectors u and v are orthogonal if and only if $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

4.2.1 Orthogonal Complement

- If a vector z is orthogonal to every vectors in a subspace W then, z is said to be orthogonal to W.

- The set of all vectors that are orthogonal to W is called the **orthogonal complement** of W. It is denoted by $W^\perp$. $W^\perp = \{z : \forall v \in W\ z.v = 0\}$

Theorem 5 1. A vector x is in $W^\perp$ if and only if x is orthogonal to every vector in a set that is spans W.

2. $W^\perp$ is also a subspace.

3. Row space is orthogonal complement of the Null space for a matrix.

4.2.2 Exercise

1. Verify parallelogram law: $\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$.

2. Suppose y is orthogonal to u and v. Show that y is orthogonal to every w in $Span\{u, v\}$.

3. Let W be a subspace of $\mathbb{R}^n$, then show that $W^\perp$ a subspace of $\mathbb{R}^n$.

4. Show that if x is in both W and $W^\perp$, then $x = 0$.

4.2.3 Orthogonal Sets

A set of vectors $\{u_1, ...u_p\}$ in a vector space V is said to be **orthogonal set** if each pair of distinct vectors from the set is orthogonal,i.e for all $u_i, u_j \in V$ we have $u_i.u_j = 0$ whenever $i \neq j$.

Theorem 3. Any orthogonal set is a linearly independent set.

Definition 7. An orthogonal basis for a vector space V is a basis for V that is an orthogonal set.

Theorem 4. Let $\{u_1, ..., u_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each y in W, the coordinates of y with respect to the orthogonal basis : $y = c_1u_1 + \ ... \ c_nu_n$ are given by $c_j = \dfrac{y.u_j}{u_j.u_j}$.

4.2.4 Orthogonal Projection

The coordinate c_j of y in Theorem 4 is actually orthogonal projection of y into the vector u_j. This can be generalized. For any given vector u. The orthogonal projection of a vector y on u is given by the formula $\hat{y} = \dfrac{y.u}{u.u}\ u$

Or, it can be derived as follows using the inner-product. $(y - \alpha u)$ and u are orthogonal so, $(y - \alpha u).u = 0$. This gives us $\alpha = \dfrac{y.u}{u.u}$. For two dimensional vectors, another orthogonal component z can be easily obtained as by subtracting the projection from the vector y. $z = y - \hat{y}$.

4.3 Orthonormal Sets, Important

Remark. A vector having length unity is called unit vector. Given a vector v if we divide it by its magnitude which its length we obtain a unit vector along v, i.e $\dfrac{v}{\|v\|}$.

A set $\{u_1, ..., u_p\}$ is an orthonormal set if it is an orthogonal set of unit vectors. And a basis of orthonormal set is a orthonormal basis. The simplest orthonormal basis is $\{e_1, \ ... \ e_n$ for $\mathbb{R}^n$.

Matrices whose columns form an orthonormal set are important in applications and in computer algorithms for matrix computations.

4.3.1 Unitary Matrix

A real matrix U is said to be unitary if $U^tU = I$. If U is complex then, $\overline{U}^{\,t}U = I$, where $\overline{U}^{\,t}$ is denoted by $U^\dagger$.

Theorem 5. An $m \times n$ matrix U has orthonormal columns if and only if $U^tU = I$

4.3.2 Exercise

1. Determine whether the following sets of vectors are orthogonal or not.

   (a) $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}$

2. Show that $\{u_1, u_2\}$ is an orthogonal basis of $\mathbb{R}^2$ and find the coordinates of x in terms of this basis.

   (a) $u_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, u_2 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}, x = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$

   (b) $u_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} 2 \\ -6 \end{bmatrix}, x = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$

   (c) Compute the orthogonal projection of $\begin{bmatrix} 1 \\ 7 \end{bmatrix}$ onto to the line through $\begin{bmatrix} -4 \\ 2 \end{bmatrix}$ and the origin.

   (d) Compute the distance of $y \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ to the line through $u = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$ and the origin.

3. Take a square unitary matrix of order 2 and a vector $x \in \mathbb{R}^2$ and verify that $\|Ux\| = \|x\|$. What can you infer from this property of a unitary matrix?

4. Show that the Hadamard quantum gate defined by $H = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$ is a unitary matrix.

4.4 The Gram-Schmidt Process

The Gram-Schmidt Process is a simple algorithm for producing an orthogonal or orthonormal basis for any nonzero subspace of $\mathbb{R}^n$. It uses the concept of **projection**. First we describe this process when a basis consisting of two vectors $\{x_1, x_2\}$ are given. Now we construct an orthogonal basis $\{v_1, v_2\}$ using the two vectors: x_1 and x_2 as follows:

1. Let $v_1 = x_1$.

2. Draw the orthogonal projection from x_2 to v_1, which is given by $\dfrac{x_2.v_1}{v_1.v_1} v_1$. Now, from the theory of orthogonal projection given above, the vector perpendicular to v_1 is given by $v_2 = x_2 - \dfrac{x_2.v_1}{v_1.v_1} v_1$.

For a given basis of p vectors $\{x_1, x_2, ..., x_p\}$, we can continue the process of drawing the orthogonal projections and subtraction to get orthogonal vectors as follows:

1. To obtain v_3 draw the orthogonal projection from x_3 onto the Spanv_1, v_2 and subtract this from x_3: $v_3 = x_3 - \dfrac{x_3.v_1}{v_1.v_1}\, v_1 - \dfrac{x_3.v_2}{v_2.v_2}\, v_2$.
2. $v_p = x_p - \dfrac{x_p.v_1}{v_1.v_1}\, v_1 - \dfrac{x_p.v_2}{v_2.v_2}\, v_2 - ... - \dfrac{x_p.v_{p-1}}{v_1.v_{p-1}}\, v_{p-1}$.

Remark. After constructing orthogonal basis orthonormal basis can be easily constructed from the orthogonal basis. *How?*

4.5 Exercise

1. Find an orthonormal basis from the given vectors:

   a). $\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 8 \\ 5 \\ 6 \end{bmatrix}$ b). $\begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$

2. Find an orthogonal basis for the column space of the matrix:

   a). $\begin{bmatrix} 3 & -5 & 1 \\ 1 & 1 & 1 \\ -1 & 5 & -2 \\ 3 & -7 & 8 \end{bmatrix}$ b). $\begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$

3. Find the QR factorization of the matrices given in the previous question.

4.5.1 QR Factorization

Theorem 6. If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an **orthonormal basis for** $Col(A)$ and R is an $n \times n$ **upper triangular matrix with positive entries on its diagonal**.

Proof. Given linearly independent columns of the matrix A, we can construct Q from the Gram-Schmidt process or any process. Then Q being real unitary matrix we have $Q^t\, Q = I$. Then $Q^t\, A = Q^t\, QR \implies R = Q^t\, A$. □

4.6 Extra Materials

The angle between the two vectors in plane and space can be generalized by the inner product. For two vectors $u, v \in \mathbb{R}^n$ we have, $u.v = \|u\|\, \|v\|\, cos\theta$, where θ is the angle between the two vectors: u, v.