Chapter 1

Vector Spaces

1.1 Vector spaces

Definition 1.1.1. *A non-empty set V is said to be a vector space over a field F, if V is an abelian group under an operation which we denote by '+', and if for every $\alpha \in F$, $v \in V$ there is defined an element, written $\alpha v \in V$ subject to*

i. $\alpha(v + w) = \alpha v + \alpha w$;

ii. $(\alpha + \beta)(v) = \alpha v + \beta v$;

iii. $\alpha(\beta v) = (\alpha\beta)v$;

iv. $1v = v$

for all $\alpha, \beta \in F$, $v, w \in V$ (where 1 corresponds to the unit element of F under multiplication).

Note that in Axiom i. above the '+' is that of V, whereas on the left-hand side of Axiom ii. it is that of F and on the right-hand side, that of V.

Problem 1.1.2. *Let V be the set of all 2×2 matrices with real entries and F be a field of real numbers. Prove that V is a vector space over F, under the binary operation '+' the usual matrix addition defined V and the multiplication of a matrix by a scalar as a scalar multiplication.*

Proof. Given V is the set of all 2×2 matrices with real entries and F is the set of all real numbers. That is,

$$V = \left\{ \left[\begin{array}{cc} a_1 & a_2 \\ a_3 & a_4 \end{array} \right] / a_1, a_2, a_3, a_4 \in \mathbb{R} \right\}.$$

We have to prove that V is a vector space. First we prove that, V is an abelian group under $'+'$.

For, let $v_1, v_2 \in V$. We have prove that $v_1 + v_2 \in V$. Let $v_1 = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ and $v_2 = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$ Then

$$v_1 + v_2 = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} + \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} a_1 + b_1 & a_2 + b_2 \\ a_3 + b_3 & a_4 + b_4 \end{bmatrix}.$$

Clearly, $a_1 + b_1, a_2 + b_2, a_3 + b_3$ and $a_4 + b_4$ is in $\mathbb{R}$, which implies $v_1 + v_2 \in V$. Thus V is closed under $'+'$.

Let $v_1, v_2, v_3 \in V$. We have prove that

$$v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3.$$

For, Let $v_1 = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ and $v_2 = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$ and $v_3 = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$. Consider,

$$\begin{aligned} v_1 + (v_2 + v_3) &= \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} + \left(\begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} + \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \right) \\ &= \begin{bmatrix} a_1 + (b_1 + c_1) & a_2 + (b_2 + c_2) \\ a_3 + (b_3 + c_3) & a_4 + (b_4 + c_4) \end{bmatrix} \\ &= \begin{bmatrix} (a_1 + b_1) + c_1 & (a_2 + b_2) + c_2 \\ (a_3 + b_3) + c_3 & (a_4 + b_4) + c_4 \end{bmatrix} \\ &= \left(\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} + \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} \right) + \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \\ &= (v_1 + v_2) + v_3. \end{aligned}$$

Thus '+' is associative.

If we let $e = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, then clearly e is the identity element and it is obvious that $-v$ is the inverse of v for any $v \in V$.

To prove the commutativity, let us consider $v_1 = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ and $v_2 = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$ in V. Then clearly,

$$v_1 + v_2 = \begin{bmatrix} a_1 + b_1 & a_2 + b_2 \\ a_3 + b_3 & a_4 + b_4 \end{bmatrix} = \begin{bmatrix} b_1 + a_1 & b_2 + a_2 \\ b_3 + a_3 & b_4 + a_4 \end{bmatrix} = v_2 + v_1.$$

Thus $(V, +)$ is abelian.

In follow we claim that the scalar multiplication defined by the multiplication of a matrix by a scalar(element of the field) is closed. For, let $\alpha \in \mathbb{R}$ and $v \in V$. We have to prove that $\alpha v \in V$. Let $v = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then

$$\alpha v = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{bmatrix}.$$

Thus, it can be seen clearly that $\alpha v \in V$. That is, the scalar multiplication is closed. Next we claim that, $\alpha(v_1+v_2) = \alpha v_1 + \alpha v_2$. For, let $v_1 = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ and $v_2 = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$. Then

$$\begin{aligned}
\alpha(v_1+v_2) &= \begin{bmatrix} \alpha(a_1+b_1) & \alpha(a_2+b_2) \\ \alpha(a_3+b_3) & \alpha(a_3+b_4) \end{bmatrix} \\
&= \begin{bmatrix} \alpha a_1 + \alpha b_1 & \alpha a_2 + \alpha b_2 \\ \alpha a_3 + \alpha b_3 & \alpha a_3 + \alpha b_4 \end{bmatrix} \\
&= \alpha \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} + \alpha \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \alpha v_1 + \alpha v_2.
\end{aligned}$$

We claim that, $(\alpha+\beta)v = \alpha v + \beta v$. For, let $v = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then

$$\begin{aligned}
(\alpha+\beta)v &= (\alpha+\beta) \begin{bmatrix} a & b \\ c & d \end{bmatrix} \\
&= \begin{bmatrix} (\alpha+\beta)a & (\alpha+\beta)b \\ (\alpha+\beta)c & (\alpha+\beta)d \end{bmatrix} \\
&= \begin{bmatrix} (\alpha a+\beta a) & (\alpha b+\beta b) \\ (\alpha c+\beta c) & (\alpha d+\beta d) \end{bmatrix} \\
&= \alpha \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \beta \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \alpha v + \beta v.
\end{aligned}$$

We claim that, $\alpha(\beta v) = (\alpha\beta)v$. For, let $v = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then we get

$$\begin{aligned}
\alpha(\beta v) &= \alpha \left(\beta \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) \\
&= \alpha \left(\begin{bmatrix} \beta a & \beta b \\ \beta c & \beta d \end{bmatrix} \right) \\
&= \begin{bmatrix} \alpha(\beta a) & \alpha(\beta b) \\ \alpha(\beta c) & \alpha(\beta d) \end{bmatrix} \\
&= \begin{bmatrix} (\alpha\beta)a & (\alpha\beta)b \\ (\alpha\beta)c & (\alpha\beta)d \end{bmatrix} \\
&= (\alpha\beta) \begin{bmatrix} a & b \\ c & d \end{bmatrix} = (\alpha\beta)\, v.
\end{aligned}$$

Finally, we claim that $1v = v$. For, let $v = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then we have

$$1v = \begin{bmatrix} 1a & 1b \\ 1c & 1d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = v.$$

Thus, V is a vector space over F. $\square$

Problem 1.1.3. *Let V be the set of all ordered n-tuples with entries from a field F. Then V is a vector space over F under the binary operation '+' defined by*

$$(a_1, \cdots, a_n) + (b_1, \cdots, b_n) = (a_1 + b_1, \cdots, a_n + b_n)$$

and the scalar multiplication defined by

$$\alpha(a_1, \cdots, a_n) = (\alpha a_1, \cdots, \alpha a_n).$$

Proof. For any two elements in $v_1, v_2 \in V$, clearly, $v_1 + v_2 \in V$. To prove associativity, let $v_1 = (a_1, \cdots, a_n), v_2 = (b_1, \cdots, b_n), v_3 = (c_1, \cdots, c_n) \in V$. Then we have

$$\begin{aligned}
v_1 + (v_2 + v_3) &= (a_1, \cdots, a_n) + ((b_1, \cdots, b_n) + (c_1, \cdots, c_n)) \\
&= (a_1, \cdots, a_n) + (b_1 + c_1, \cdots, b_n + c_n) \\
&= (a_1 + (b_1 + c_1), \cdots, a_n + (b_n + c_n)) \\
&= ((a_1 + b_1) + c_1, \cdots, (a_n + b_n) + c_n) \\
&= ((a_1 + b_1), \cdots, (a_n + b_n)) + (c_1, \cdots, c_n) \\
&= ((a_1, \cdots, a_n) + (b_1, \cdots, b_n)) + (c_1, \cdots, c_n) \\
&= (v_1 + v_2) + v_3.
\end{aligned}$$

Clearly, $(0, \cdots, 0)$ is the identity element and for any $v = (a_1, \cdots, a_n) \in V$, $-v = (-a_1, \cdots, -a_n)$ is the inverse of v. To prove the commutativity, Let $v_1 = (a_1, \cdots, a_n), v_2 = (b_1, \cdots, b_n) \in V$, then we have

$$\begin{aligned}
v_1 + v_2 &= (a_1, \cdots, a_n) + (b_1, \cdots, b_n) \\
&= ((a_1 + b_1), \cdots, (a_n + b_n)) \\
&= ((b_1 + a_1), \cdots, (b_n + a_n)) \\
&= (b_1, \cdots, b_n) + (a_1, \cdots, a_n) \\
&= v_2 + v_1.
\end{aligned}$$

From the definition of scalar multiplication given, clearly $\alpha v \in V$ for any $\alpha \in F$ and $v \in V$. Now we have to prove the remaining properties of the vector space. For, let $\alpha \in F$ and $v_1 = (a_1, \cdots, a_n), v_2 = (b_1, \cdots, b_n) \in V$, then

$$\begin{aligned}
\alpha(v_1 + v_2) &= \alpha((a_1, \cdots, a_n) + (b_1, \cdots, b_n)) \\
&= \alpha((a_1 + b_1), \cdots, (a_n + b_n)) \\
&= (\alpha(a_1 + b_1), \cdots, \alpha(a_n + b_n)) \\
&= ((\alpha a_1 + \alpha b_1), \cdots, (\alpha a_n + \alpha b_n)) \\
&= (\alpha a_1, \cdots, \alpha a_n) + (\alpha b_1, \cdots, \alpha b_n) \\
&= \alpha(a_1, \cdots, a_n) + \alpha(b_1, \cdots, b_n) \\
&= \alpha v_1 + \alpha v_2.
\end{aligned}$$

Now we claim that, $(\alpha+\beta)v = \alpha v + \beta v$. For, if we let $\alpha, \beta \in F$ and $v = (a_1, \cdots, a_n) \in V$, then we have

$$\begin{aligned}
(\alpha+\beta)v &= (\alpha+\beta)(a_1, \cdots, a_n) \\
&= ((\alpha+\beta)a_1, \cdots, (\alpha+\beta)a_n) \\
&= ((\alpha a_1 + \beta a_1), \cdots, (\alpha a_n + \beta a_n)) \\
&= (\alpha a_1, \cdots, \alpha a_n) + (\beta a_1, \cdots, \beta a_n) \\
&= \alpha v + \beta v.
\end{aligned}$$

Next, we claim that $\alpha(\beta v) = (\alpha\beta)v$. Let $\alpha, \beta \in F$ and $v = (a_1, \cdots, a_n) \in V$, then

$$\begin{aligned}
\alpha(\beta v) &= \alpha(\beta(a_1, \cdots, a_n)) \\
&= \alpha(\beta a_1, \cdots, \beta a_n) \\
&= (\alpha(\beta a_1), \cdots, \alpha(\beta a_n)) \\
&= ((\alpha\beta)a_1, \cdots, (\alpha\beta)a_n) \\
&= (\alpha\beta)(a_1, \cdots, a_n) \\
&= (\alpha\beta)v
\end{aligned}$$

Clearly, $1v = v$ for all $v \in V$. Thus V is a vector space over F. $\square$

Problems

- Let V be the set of all polynomials whose coefficients are from $\mathbb{R}$. Check whether V is vector space over $\mathbb{R}$, under the coefficientwise addition and the scalar multiplication as coefficientwise multiplication.

- Let V be the set of all functions from $\mathbb{R}$ to $\mathbb{R}$. Check whether V is vector space over $\mathbb{R}$, with pointwise addition as a binary operation on V and the scalar multiplication as usual multiplication of a scalar with a function that is $(\alpha f)(x) = \alpha(f(x))$.

- Let V be the set of all functions from $\mathbb{R}$ to $\mathbb{R}$, whose functional value is zero at zero. Check whether V is vector space over $\mathbb{R}$, with pointwise addition as a binary operation on V and the scalar multiplication as usual multiplication of a scalar with a function, i.e., $(\alpha f)(x) = \alpha(f(x))$.

- Let $V = \{(x, y)/x, y \in \mathbb{R}\}$ and F be a field of rational numbers. Check whether V is vector space over F, under the following operations

  i. $(a, b) + (c, d) = (a + c, 0)$ and $\alpha(a, b) = (\alpha a, 0)$

  ii. $(a, b) + (c, d) = (a + c, b + d)$ and $\alpha(a, b) = (|\alpha|a, |\alpha|b)$.

1.2 Subspaces

Definition 1.2.1. *Let V be a vector space over a field F and let W be a subset of V. Then W is said to be a subspace of V if W itself is a vector space over F with respect to the operations of vector addition and scalar multiplication in V.*

Theorem 1.2.2. *Let $W \subseteq V$. Then W is a subspace of a vector space $V(F)$ if and only if for any $w_1, w_2 \in W$ and $\alpha, \beta \in F$, $\alpha w_1 + \beta w_2 \in W$.*

Proof. If W is a subspace of V, then W must be closed under scalar multiplication and vector addition. Thus for any $\alpha, \beta \in F$ and $w_1, w_2 \in W$ we have $\alpha w_1 + \beta w_2 \in W$.

Conversely, suppose W is non-empty subset of V satisfying the given condition, that is for any $w_1, w_2 \in W$ and $\alpha, \beta \in F$, $\alpha w_1 + \beta w_2 \in W$. Then by taking $\alpha = \beta = 1$, we get $w_1 + w_2 \in W$. Thus W is closed under vector addition. If we let $\alpha = -1$ and $\beta = 0$, we see that $-1w + 0w = -w \in W$. Thus the additive inverse of any element $w \in W$ is in W. Suppose $\alpha = 0$ and $\beta = 0$, then it is easy to see that $0 \in W$. Thus the zero vector V is in W. Since the elements of W are the elements of V, clearly the vector addition is associative and commutative in W. Thus W is an abelian group under vector addition. Let $w_1 \in W$, then if we let $w_2 = 0$, then we have $\alpha w_1 + \beta 0 = \alpha w_1 \in W$. Thus W is closed under scalar multiplication. All other postulates of a vector space will hold in W as they hold in V of which W is a subset of V. Hence W is a subspace of V. $\square$

Theorem 1.2.3. *Let $W \subseteq V$. Then W is a subspace of a vector space $V(F)$ if and only if the following conditions are satisfied*

i. for any $w_1, w_2 \in W$, $w_1 - w_2 \in W$;

ii. for any $\alpha \in F$ and $w \in W$, $\alpha w \in W$.

Proof. Suppose W is subspace of V, then W is an abelian group with respect to vector addition, which implies for any $w_1, w_2 \in W$, $w_1 - w_2 \in W$ and for any $\alpha \in F$ and $w \in W$, $\alpha w \in W$.

Conversely, suppose W is a non-empty subset, that satisfies the two conditions. We claim that W is a subspace of V. First we claim that $0 \in W$ (that is the identity element of V is the required identity element of W). From the condition [i.], we have $\alpha - \alpha = 0 \in W$, which implies $0 \in W$. Now, since $0 \in W$ and by using [i.] again, we have $0 - \alpha = -\alpha \in W$. Thus the additive inverse of $w \in W$ is also in W. Using the above argument and using the condition [i.] again, for any two vectors $w_1, w_2 \in W$, we get

$$w_1 + w_2 = w_1 - (-w_2) \in W.$$

Thus W is closed under the vector addition.

Also, as elements of W are the elements of V, it is obvious that the vector addition is associative and commutative in W. Hence W is an abelian group. From condition [ii.], W is closed under scalar multiplication. The remaining postulates of a vector space will hold in W as they hold in V of which W is a subset of V. Hence W is a subspace of V. □

Note:

Let $V(F)$ be any vector space. Then V itself and the subset of V consisting of zero vector only are always subspaces of V. These two are called improper subspaces. If W is any other subspace of V, then W is a proper subspace of V.

Problem 1.2.4. *Let $V = V_3(F)$ be a vector space of ordered triads (a_1, a_2, a_3), where $a_1, a_2, a_3 \in F$ and let W be the subset of V consisting of all ordered triads $(a_1, a_2, 0)$, where $a_1, a_2 \in F$. Prove that W is subspace of V.*

Proof. Given $V = V_3(F)$ be a vector space of ordered triads (a_1, a_2, a_3), where $a_1, a_2, a_3 \in F$ and let W be the subset of V consisting of all ordered triads $(a_1, a_2, 0)$, where $a_1, a_2 \in F$. That is,

$$V = \{(a_1, a_2, a_3)/a_1, a_2, a_3 \in F\}$$

and

$$W = \{(a_1, a_2, 0)/a_1, a_2 \in F\}.$$

We claim that, W is a subspace of V. That is, we have to prove, for any $w_1, w_2 \in W$ and $\alpha, \beta \in F$, $\alpha w_1 + \beta w_2 \in W$.

Let $w_1 = (a_1, a_2, 0), w_2 = (b_1, b_2, 0) \in W$ and $\alpha, \beta \in F$. Then we have,

$$\begin{aligned} \alpha w_1 + \beta w_2 &= \alpha(a_1, a_2, 0) + \beta(b_1, b_2, 0) \\ &= (\alpha a_1, \alpha a_2, 0) + (\beta b_1, \beta b_2, 0) \\ &= (\alpha a_1 + \beta b_1, \alpha a_2 + \beta b_2, 0). \end{aligned}$$

Since $\alpha a_1 + \beta b_1, \alpha a_2 + \beta b_2 \in F$ and the last co-ordinate of this triad is zero, it follows that $\alpha w_1 + \beta w_2 \in W$. Hence W is as subspace of $V_3(F)$. $\square$

Problem 1.2.5. *Let $V(F)$ be a vector space of all $n \times 1$ matrices over a field F. Let A be an $m \times n$ matrix over(entries of the matrix are from) F. Then the set W of all $n \times 1$ matrices X over F such that $AX = \mathbf{0}_{m\times 1}$ is a subspace of V, where $\mathbf{0}_{m\times 1}$ is the zero matrix.*

Proof. Given $V(F)$ be a vector space of all $n \times 1$ matrices over a field F, that is,

$$V(F) = \left\{ B_{n\times 1} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} / b_i \in F \right\}.$$

And

$$W = \left\{ X_{n\times 1} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} / x_i \in F, A_{m\times n} X_{n\times 1} = \mathbf{0}_{m\times 1} \right\}.$$

Let $X, Y \in W$, then X and Y are $n \times 1$ matrices over F such that $AX = \mathbf{0}$ and $AY = \mathbf{0}$. Let $\alpha, \beta \in F$, then $\alpha X + \beta Y$ is also an $n \times 1$ matrix over F.

Also we have

$$\begin{aligned} \alpha X + \beta Y &= A(\alpha X + \beta Y) \\ &= A(\alpha X) + A(\beta Y) \\ &= \alpha(AX) + \beta(AY) \\ &= \alpha\mathbf{0} + \beta\mathbf{0} \\ &= \mathbf{0}. \end{aligned}$$

Thus, $\alpha X + \beta Y$ is in W and hence W is a subspace of V. □

Problems

1. Prove the following statements:

   - The necessary and sufficient condition for a non-empty subset W of a vector space V(F) to be a subspace of V is that W is closed under vector addition and scalar multiplication in V.

   - Let $W \subseteq V$. Then W is a subspace of a vector space $V(F)$ if and only if for any $w_1, w_2 \in W$ and $\alpha \in F$, $\alpha w_1 + w_2 \in W$.

2. Prove the following statements:

   - Let V be the vector space of all polynomials over a field F. Let W be a subset of V consisting of all polynomials of degree $\leq$ n. Then W is a subspace of V.

   - If a_1, a_2, a_3 are some fixed elements of a field F, then the set W of all ordered triads (x_1, x_2, x_3) of elements of F, such that

$$a_1x_1 + a_2x_2 + a_3x_3 = 0,$$

     is a subspace of $V_3(F)$.

3. Let $\mathbb{R}$ be the field of real numbers. Which of the following are subspaces of $V_3(\mathbb{R})$:

   i. $\{(x, 2y, 3z) : x, y, z \in \mathbb{R}\}$;

ii. $\{(x, x, x) : x \in \mathbb{R}\}$;

iii. $\{(x, y, z) : x, y, z$ are rational numbers$\}$.

4. Let V be the (real) vector space of all functions $f : \mathbb{R} \to \mathbb{R}$. Which of the following sets of functions are subspaces of V:

   i. all f such that $f(x^2) = (f(x))^2$;

   ii. all f such that $f(0) = f(1)$;

   iii. all f such that $f(3) = 1 + f(-5)$;

   iv. all f such that $f(-1) = 0$;

   v. set of all continuous functions.

5. Which of the following sets of vectors $v = (a_1, a_2, \cdots, a_n)$ in $\mathbb{R}^n$ are subspaces of $\mathbb{R}^n$, $n \geq 3$?

   i. all v such that $v_1 \geq 0$;

   ii. all v such that $a_1 + 3a_2 = a_2$;

   iii. all v such that $a_2 = a_1^2$;

   iv. all v such that $a_1 a_2 = 0$;

   v. all v such that a_2 is rational.

6. Let $\mathbb{C}$ be the field of complex numbers and let n be a positive integer ($n \geq 2$). Let V be the vector space of all $n \times n$ matrices over $\mathbb{C}$. Which of the following sets of matrices A is V are subspaces of V?

i. all invertible A;

ii. all non-invertible A;

iii. all A such that $AB = BA$, where B is some fixed matrix in V.

7. Let V be the vector space of all 2×2 matrices over the real field $\mathbb{R}$. Show that the subset of V consisting of all matrices A for which $A^2 = A$ is not a subspace of V.

Problem 1.2.6. *Prove that the intersection of two subspaces is a subspace.*

Proof. Let $V(F)$ be a vector space over a field F and W_1 and W_2 be two subspaces of $V(F)$.

Now since $0 \in W_1$ and $0 \in W_2$, we have $0 \in W_1 \cap W_2$. Thus $W_1 \cap W_2$ is non-empty. Let $\alpha, \beta \in F$ and $u, v \in W_1 \cap W_2$, then $u, v \in W_1$ and $u, v \in W_2$. Since both W_1 and W_2 are subspace of $V(F)$, it follows that, $\alpha u + \beta v \in W_1$ and $\alpha u + \beta v \in W_2$. Thus $\alpha u + \beta v \in W_1 \cap W_2$ and hence $W_1 \cap W_2$ is a subspace of $V(F)$. $\square$

Problem 1.2.7. *Prove that any arbitrary intersection of subspaces is a subspace.*

Smallest subspace containing any subset of $V(F)$:

Let $V(F)$ be a vector space over a field F. Let S be a subset of V. Let U be a subspace of $V(F)$ that contains S such that "if W is another subspace of $V(F)$ containing S, then $U \subset W$", then U is called a smallest subspace of $V(F)$ containing S.

The smallest subspace of $V(F)$ containing S is also called as the subspace of V generated or spanned by S. It can be seen easily that the intersection of all subspaces that contains S is the subspace of V generated or spanned by S.

1.3 Linear combination of vectors

Definition 1.3.1 (Linear combination)**.** *Let $V(F)$ be a vector space and let $\{v_1, v_2, v_3, \cdots, v_n\}$, then any vector v represented as*

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \cdots + \alpha_n v_n,$$

where $\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_n \in F$ is called a linear combination of vectors $v_1, v_2, v_3, \cdots, v_n$.

Definition 1.3.2 (Linear span)**.** *Let $V(F)$ be a vector space and S be any non-empty set of V. Then the linear span of S is the set of all linear combinations of finite sets of elements of S and is denoted by $L(S)$. That is,*

$$L(S) = \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \ / \ \alpha_1, \alpha_2, \cdots, \alpha_n \in F, \quad v_1, v_2, \cdots, v_n \in S\}.$$

Theorem 1.3.3. *The linear span $L(S)$ of any subset S of a vector space $V(F)$ is a subspace of V generated by S.*

Proof. Let $a = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n, b = \beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_m b_m$ be any two elements of $L(S)$, where $a_i, b_i \in V$ and $\alpha_i, \beta_i \in F$. We have to prove that for any $\alpha, \beta \in F$, $\alpha a + \beta b \in L(S)$. For, consider

$$\begin{aligned}
\alpha a + \beta b &= \alpha(\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n) + \beta(\beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_m b_m) \\
&= \alpha(\alpha_1 a_1) + \alpha(\alpha_2 a_2) + \cdots + \alpha(\alpha_n a_n) \\
&\qquad + \beta(\beta_1 b_1) + \beta(\beta_2 b_2) + \cdots + \beta(\beta_m b_m) \\
&= (\alpha\alpha_1)a_1 + (\alpha\alpha_2)a_2 + \cdots + (\alpha\alpha_n)a_n \\
&\qquad + (\beta\beta_1)b_1 + (\beta\beta_2)b_2 + \cdots + (\beta\beta_m)b_m.
\end{aligned}$$

Thus $\alpha a + \beta b$ has been expressed as a linear combination of a finite set $\{a_1, a_2, \cdots, a_n, b_1, b_2, \cdots, b_n\}$ of elements of S. Consequently, $\alpha a + \beta b$ in $L(S)$, which implies $L(S)$ is a subspace V.

As any $a \in S$, can be represented as $a = 1a$; it clearly follows that $S \subset L(S)$. Hence $L(S)$ is a subspace of V containing S.

Now suppose W is any subspace of V containing S, then each element of $L(S)$ must be in W as W being a subspace must be closed under vector addition and scalar multiplication. Thus $L(S) \subseteq W$ and hence $L(S)$ is the smallest subspace of V containing S. □

Examples

1. The subset containing the single element $(1, 0, 0)$ of the vector space $V_3(F)$ generates the subspace which is the totality of the elements of

the form $(a, 0, 0)$.

2. The subset containing the elements $\{(1,0,0),(0,1,0)\}$ of the vector space $V_3(F)$ generates the subspace which is the totality of the elements of the form $(a, b, 0)$.

3. The subset $S = \{(1,0,0),(0,1,0),(0,0,1)\}$ of $V_3(F)$ generates or spans the entire vector space $V_3(F)$, that is $L(S) = V$.

   If (a, b, c) be any element of V, then

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

   which implies $(a, b, c) \in L(S)$. Thus $V \subseteq L(S)$ and hence $L(S) = V$.

4. Let V be the vector space of all polynomials over the field F. Let S be the subset of V consisting of the polynomials $f_0, f_1, f_2, \cdots$, defined by $f_n = x^n$, $n = 0, 1, 2, \cdots$. Then $V = L(S)$.

Linear sum of two subspaces:

Definition 1.3.4. *Let W_1 and W_2 be two subspaces of the vector space $V(F)$. Then the linear sum of the subspaces W_1 and W_2 denoted by $W_1 + W_2$, is the set of all sums $w_1 + w_2$ such that $w_1 \in W_1$ and $w_2 \in W_2$. That is,*

$$W_1 + W_2 = \{w_1 + w_2 / w_1 \in W_1, w_2 \in W_2\}.$$

Theorem 1.3.5. *If W_1 and W_2 are subspaces of vector space $V(F)$, then* *i.* $W_1 + W_2$ *is a subspace of* $V(F)$*;*

ii. $L(W_1 \cup W_2) = W_1 + W_2$.

Proof. To prove[i.], let a, b be any two elements of W_1+W_2. Then $a = a_1+a_2$ and $b = b_1 + b_2$, where $a_1, a_2 \in W_1$ and $b_1, b_2 \in W_2$. Let $\alpha, \beta \in F$. We claim that $\alpha a + \beta b \in W_1 + W_2$. Consider,

$$\begin{aligned}\alpha a + \beta b &= (\alpha a_1 + \alpha a_2) + (\beta b_1 + \beta b_2) \\ &= (\alpha a_1 + \beta b_1) + (\alpha a_2 + \beta b_2).\end{aligned}$$

Since W_1 and W_2 are subspaces of V, we have $\alpha a_1 + \beta b_1 \in W_1$ and $\alpha a_2 + \beta b_2 \in W_2$ and hence $(\alpha a_1 + \beta b_1) + (\alpha a_2 + \beta b_2) \in W_1 + W_2$. Hence $W_1 + W_2$ is a subspace of V.

To prove [ii.] We have to prove that $L(W_1 \cup W_2) = W_1 + W_2$, that is to prove that $L(W_1 \cup W_2) \subseteq W_1 + W_2$ and $L(W_1 \cup W_2) \supseteq W_1 + W_2$.

For, as any $a \in W_1$ and $b \in W_2$ can be written as $a = a+0$ and $b = b+0$, it is easy to see that $a, b \in W_1 + W_2$, which in turn implies that both W_1 and W_2 are subsets of $W_1 + W_2$. Therefore $W_1 \cup W_2 \subseteq W_1 + W_2$. Thus $W_1 + W_2$ is a subspace V that contains $W_1 \cup W_2$. But since $L(W_1 \cup W_2)$ is the smallest subspace V that contains $W_1 \cup W_2$, it follows that $L(W_1 \cup W_2) \subseteq W_1 + W_2$.

To prove the other side, let $a = a_1 + a_2 \in W_1 + W_2$. Then $a_1 \in W_1$ and $a_2 \in W_2$, which implies $a_1, a_2 \in W_1 \cup W_2$.Thus we can write a as

$$a = a_1 + a_2 = 1a_1 + 1a_2$$

and hence it follows that $L(W_1 \cup W_2) \supseteq W_1 + W_2$ as desired. $\square$

Problems

1. If S and T are subsets of $V(F)$, then prove that

   - Suppose $S \subseteq T$, then $L(S) \subseteq L(T)$;
   - $L(S \cup T) = L(S) + L(T)$;
   - If S is a subspace of V, then $L(S) = S$;
   - $L(L(S)) = L(S)$.

2. Show that the intersection of any collection of subspaces of a vector space is a subspace. Can you replace the 'intersection by 'union' in this problem.

3. Show that the union of two subspaces is a subspace if and only if one is contained in other.

4. Let V be the vector space of all functions from $\mathbb{R}$ to $\mathbb{R}$; let V_e be the subset of even functions, ie., $f(-x) = f(x)$; let V_o be the subset of odd functions, ie., $f(-x) = -f(x)$. Then prove that

   - Prove that V_e and V_o are subspaces of V;
   - Prove that $V_e + V_o = V$;
   - Prove that $V_e \cap V_o = \{\mathbf{0}\}$.

1.4 Linear dependence and linear independence of vectors

Definition 1.4.1 (Linear dependence).

Let $V(F)$ be a vector space. A finite set $\{v_1, v_2, \cdots, v_n\}$ of vectors of V is said to be linearly dependent over F if there exist scalars $\alpha_1, \alpha_2, \cdots, \alpha_n \in F$ not all of them are 0 such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + v_n \alpha_n = \mathbf{0}.$$

Definition 1.4.2 (Linear independence).

A finite set $\{v_1, v_2, \cdots, v_n\}$ of vectors of a vector space $V(F)$ is said to be linearly independent over F if it is not linearly dependent.

Any infinite set of vectors of V is said to be linearly independent if every finite subset of the infinite set is linearly independent, otherwise it is linearly dependent.

It is easy to see that, the set of vectors $\{(1,0,0),(0,1,0),(0,0,1)\}$ is linearly independent over the vector space of $\mathbb{R}^3$ over $\mathbb{R}$, whereas the set of vectors $\{(1,1,0),(3,1,3),(5,3,3)\}$ is linearly dependent. Note that the property of linear dependence not only depends on the given vector space but also it depends on over which field it is considered. For example, consider the set of vectors $\{1, i\}$ in the vector space of complex numbers over the field of real numbers. Then it is easy to see that the set is linearly independent whereas the same set of vectors is linearly dependent if we consider the set of complex numbers as a vector space over the field of complex numbers, as $i(1) + (-1)i = \mathbf{0}$.

Result 1.4.3. *Prove that if a set of two vectors is linearly dependent, one of them is a scalar multiple of the other.*

Proof. Let $\{u, v\}$ be a set linearly dependent set of vectors in a vector space V over a field F. Then there exist $\alpha, \beta \in F$ such that either $\alpha \neq 0$ or $\beta \neq 0$ and $\alpha u + \beta v = \mathbf{0}$. Suppose $\alpha \neq 0$, then

$$u = (\frac{-\beta}{\alpha})v,$$

which implies u is a scalar multiple of v. If $\beta \neq 0$, then

$$v = (\frac{-\alpha}{\beta})u,$$

which implies v is a scalar multiple of u. Thus one of the vectors u and v is a scalar multiple of the other. □

Example 1.4.4. *In the vector space $V_n(F)$, the set of n vectors $e_1 = (1, 0, \cdots, 0), e_2 = (0, 1, \cdots, 0), \cdots, e_n = (0, 0, \cdots, 1)$ is linearly independent where 1 denotes the unity of the field F.*

Proof. Suppose let $\alpha_1, \alpha_2, \cdots, \alpha_n$ be a set of scalars, such that

$$\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n = \mathbf{0}.$$

Then we have

$$\begin{aligned}
\mathbf{0} &= (0,0,\cdots,0) \\
&= \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n \\
&= \alpha_1(1,0,\cdots,0) + \alpha_2(0,1,\cdots,0) + \cdots + \alpha_n(0,0,\cdots,1) \\
&= (\alpha_1,0,\cdots,0) + (0,\alpha_2,\cdots,0) + \cdots + (0,0,\cdots,\alpha_n) \\
&= (\alpha_1,\alpha_2,\cdots,\alpha_n).
\end{aligned}$$

Thus it follows that $\alpha_1 = 0, \alpha_2 = 0, \cdots, \alpha_n = 0$, which implies the set of n vectors

$$e_1 = (1,0,\cdots,0), e_2 = (0,1,\cdots,0), \cdots, e_n = (0,0,\cdots,1)$$

is linearly independent. □

Result 1.4.5. *If the set $S = \{v_1, v_2, \cdots, v_n\}$ of vectors of $V(F)$ is linearly independent, then $v_i \neq \mathbf{0}$ for all i.*

Proof. Let v_r be a vector in S, such that $v_r = \mathbf{0}$. If we let $\alpha_i = 0$ for all $i \neq r$ and α_r be any non-zero scalar, then the zero vector $\mathbf{0}$ can be represented as

$$0v_1 + 0v_2 + \cdots + \alpha_r v_r + 0v_{r+1} + \cdots + \alpha_n v_n = \mathbf{0}.$$

Thus it follows that S is linearly dependent, which leads to a contradiction.

□

Result 1.4.6. *Every superset of a linearly dependent set of vectors is linearly dependent.*

Result 1.4.7. *If v is any non zero vector, in a vector space $V(F)$, then $\{v\}$ is linearly independent.*

Problem 1.4.8. *Show that*

$$S = \{(1,2,4),(1,0,0),(0,1,0),(0,0,1)\}$$

is a linearly dependent subset of the vector space $V_3(\mathbb{R})$ where $\mathbb{R}$ is the field of real numbers.

Proof. To prove the statement, we have to produce four scalars $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ not necessarily distinct and not all zero, such that

$$\alpha_1(1,2,4) + \alpha_2(1,0,0) + \alpha_3(0,1,0) + \alpha_4(0,0,1) = (0,0,0).$$

It is clear to see that, if we let $\alpha_1 = 1, \alpha_2 = -1, \alpha_3 = -2, \alpha_4 = -4$, then

$$1(1,2,4) + (-1)(1,0,0) + (-2)(0,1,0) + (-4)(0,0,1) = (0,0,0).$$

□

Problem 1.4.9. *In $\mathbb{R}^3$, where $\mathbb{R}$ is the field of reals, examine each of the following sets of vectors for linear dependence:*

1. $S = \{(2,1,2),(8,4,8)\}$;
2. $S = \{(1,2,0),(0,3,1),(-1,0,1)\}$;
3. $S = \{(-1,2,1),(3,0,-1),(-5,4,3)\}$;

4. $S = \{(2,3,5),(4,9,25)\}$*;*

5. $S = \{(1,3,2),(1,-7,-8),(2,1,-1)\}$*;*

Proof. [1.] Clearly, the set $S = \{(2,1,2),(8,4,8)\}$ is linearly dependent. For, if we let $\alpha_1 = 4$ and $\alpha_2 = -1$, then it is clear to see that $\alpha_1(2,1,2) + \alpha_2(8,4,8) = (0,0,0)$.

[2.] Suppose α, β, γ be scalars(reals) such that

$$\alpha(1,2,0) + \beta(0,3,1) + \gamma(-1,0,1) = (0,0,0).$$

[Our aim is to find the value of the scalars, if the value of all the scalars is zero then the set is linearly independent; suppose the value of one of the non-zero then the set is linearly dependent]. Then, we get that

$$(\alpha - \gamma, 2\alpha + 3\beta, \beta + \gamma) = (0,0,0).$$

Therefore

$$\begin{aligned} \alpha + 0\beta - \gamma &= 0; \\ 2\alpha + 3\beta + 0\gamma &= 0; \\ 0\alpha + \beta + \gamma &= 0. \end{aligned}$$

These equations will have a non-zero solution; that is a solution in which α, β, γ are not all zero if the rank of the co-efficient matrix is less than three (the number of unknowns). If the rank is 3, then the solution $\alpha = 0$; $\beta = 0$; $\gamma = 0$, is the only solution of the system of equations.

Coefficient matrix is given by

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

Thus the determinant of the matrix is given by

$$|A| = 1(3-0) - 2(0+1) = 1 \neq 0,$$

which implies the rank of the matrix is 3. Hence $\alpha = 0, \beta = 0, \gamma = 0$ is the only solution. Therefore the given set is linearly independent.

[3.] Let α, β, γ be scalars(reals) such that

$$\alpha(-1,2,1) + \beta(3,0,-1) + \gamma(-5,4,3) = (0,0,0).$$

[Our aim is to find the value of the scalars, if the value of all the scalars is zero then the set is linearly independent; suppose the value of one of the non-zero then the set is linearly dependent]. Then, we get that

$$(-\alpha + 3\beta - 5\gamma, 2\alpha + 0\beta + 4\gamma, \alpha - \beta + 3\gamma) = (0,0,0).$$

Therefore

$$\begin{aligned} -\alpha + 3\beta - 5\gamma &= 0; \\ 2\alpha + 0\beta + 4\gamma &= 0; \\ \alpha - \beta + 3\gamma &= 0. \end{aligned}$$

Coefficient matrix is given by

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

Thus the determinant of the matrix is given by $|A| = 0$, which implies the rank of the matrix is < 3. Therefore the system of equations posses a non-zero solution. In particular, $\alpha = -2, \beta = 1, \gamma = 1$ is a non-zero solution. Hence the set of vectors is linearly dependent.

[4.] Let α, β be scalars (reals) such that

$$\alpha(2,3,5) + \beta(4,9,25) = (0,0,0).$$

[Our aim is to find the value of the scalars, if the value of all the scalars is zero then the set is linearly independent; suppose the value of one of the non-zero then the set is linearly dependent]. Then, we get that

$$(2\alpha + 4\beta, 3\alpha + 9\beta, 5\alpha + 25\beta) = (0,0,0).$$

Therefore

$$\begin{aligned} 2\alpha + 4\beta &= 0 \\ 3\alpha + 9\beta &= 0 \\ 5\alpha + 25\beta &= 0. \end{aligned}$$

Coefficient matrix is given by

$$A = \begin{bmatrix} 2 & 4 \\ 3 & 9 \\ 5 & 25 \end{bmatrix}.$$

Obviously, rank $A = 2$ (equal to the number of unknowns). Therefore these equations have the only solution $a = 0$, $b = 0$. Hence the linearly set of vectors is linearly independent.

[5.] Let α, β, γ be scalars(reals) such that

$$\alpha(1,3,2) + \beta(1,-7,-8) + \gamma(2,1,-1) = (0,0,0).$$

[Our aim is to find the value of the scalars, if the value of all the scalars is zero then the set is linearly independent; suppose the value of one of the non-zero then the set is linearly dependent]. Then, we get that

$$(\alpha + \beta + 2\gamma, 3\alpha - 7\beta + \gamma, 2\alpha - 8\beta - \gamma) = (0,0,0).$$

Therefore

$$\alpha + \beta + 2\gamma = 0 \tag{1.1}$$

$$3\alpha - 7\beta + \gamma = 0 \tag{1.2}$$

$$2\alpha - 8\beta - \gamma = 0. \tag{1.3}$$

Eliminating γ in 1.1 and 1.2, we get $5\alpha - 15\beta = 0$, which implies $\alpha - 3\beta = 0$. Eliminating γ in 1.2 and 1.3, we get $5\alpha - 15\beta = 0$, which implies $\alpha - 3\beta = 0$. If we choose $\beta = 1$, then $\alpha = 3$. Also we get $\gamma = -2$. Hence the given set is linearly dependent. $\square$

Problems

1. If F is the field of real numbers, prove that the vectors (a_1, a_2) and (b_1, b_2) in $V_2(F)$ are linearly dependent if and only if $a_1 b_2 - a_2 b_1 = 0$.

2. If $u, v \in V(F)$ and let $\alpha, \beta \in F$. Show that the set $\{u, v, \alpha u + \beta v\}$ is linearly dependent.

3. Let v_1, v_2, v_3 be a set of vectors $V(F)$, $\alpha, \beta \in F$. Show that the set $\{v_1, v_2, v_3\}$ is linearly dependent if the set $\{v_1 + \alpha v_2 + \beta v_3, v_2, v_3\}$ is linearly dependent.

4. If v_1, v_2, v_3 be a linearly independent set of vectors $V(F)$, where F is any subfield of the field of complex numbers then so also are $v_1+v_2, v_2+v_3, v_3+v_1$.

5. In the vector space $F[x]$ of all polynomials over the field F the infinite set $S = \{1, x, x^2, x^3, \cdots\}$ is linearly independent.

Problem 1.4.10. *Check whether the vector* $(2, -5, 3)$ *is an element of the subspace spanned by the vectors* $(1, -3, 2)$, $(2, -4, -1)$ *and* $(1, -5, 7)$ *in* $\mathbb{R}^3$.

Proof. Let $u = (2, -5, 3)$, $v_1 = (1, -3, 2)$, $v_2 = (2, -4, -1)$ and $v_3(1, -5, 7)$. If u can be expressed as a linear combination of the vectors $v_1 = (1, -3, 2)$, $v_2 = (2, -4, -1)$ and $v_3(1, -5, 7)$, then it will be in the subspace of $\mathbb{R}^3$ spanned by these vectors otherwise it will not be.

Let $u = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$, where $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. Then we get

$$(2, -5, 3) = \alpha_1(1, -3, 2) + +\alpha_2(2, -4, -1) + \alpha_3(1, -5, 7).$$

That is,

$$(2, -5, 3) = (\alpha_1 + 2\alpha_2 + \alpha_3, -3\alpha_1 - 4\alpha_2 - 5\alpha_3, 2\alpha_1 - 1\alpha_2 + 7\alpha_3).$$

Therefore

$$\begin{aligned} \alpha_1 + 2\alpha_2 + \alpha_3 &= 2 && (1.4) \\ -3\alpha_1 - 4\alpha_2 - 5\alpha_3 &= -5 && (1.5) \\ 2\alpha_1 - 1\alpha_2 + 7\alpha_3 &= 3. && (1.6) \end{aligned}$$

Multiplying the equation 1.4 by 3 and adding to the equation 1.5, we get $2\alpha_2 - 2\alpha_3 = 1$, which implies $\alpha_2 - \alpha_3 = \frac{1}{2}$. Again multiplying the equation 1.4 by 2 and subtracting it from the equation 1.6, we get $-5\alpha_2 + 5\alpha_3 = -1$, that is $\alpha_2 - \alpha_3 = \frac{1}{5}$. Clearly, the system is inconsistent. Hence the vector u cannot be represented as a linear combination of the vectors $v_1 = (1, -3, 2)$, $v_2 = (2, -4, -1)$ and $v_3(1, -5, 7)$. Therefore the vector $(2, -5, 3)$ in the subspace of $\mathbb{R}^3$ spanned by the vectors $(1, -3, 2)$, $(2, -4, -1)$ and $(1, -5, 7)$. □

Theorem 1.4.11. *Let $V(F)$ be a vector space. If $v_1, v_2, \cdots, v_n$ are non-zero vectors in V, then either they are linearly independent or some v_k, $2 \leq k \leq n$, is a linear combination of the preceding ones $v_1, v_2, \cdots, v_{k-1}$.*

Proof. Let $v_1, v_2, \cdots, v_n$ be a set of non-zero vectors in V. Suppose the set of vectors are linearly independent, then there is nothing to prove.

So, we let the set of vectors $v_1, v_2, \cdots, v_n$ to be linearly dependent. Then there exist scalars $\alpha_1, \alpha_2, \cdots, \alpha_n$, not all zero, such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \mathbf{0}. \tag{1.7}$$

Let k be the largest integer for which $\alpha_k \neq 0$ and $\alpha_{k+1} = \alpha_{k+2} = \cdots = \alpha_n = 0$ (there is no harm in this assumption because atmost if $\alpha_n \neq 0$, then we can let $k = n$. Also k cannot be less than 2, as if $\alpha_1 = \alpha_2 = \alpha_3 = \cdots = \alpha_n = 0$, then obviously the set is linearly independent; on the other hand, if $\alpha_2 = \alpha_3 = \cdots = \alpha_n = 0$, then $\alpha_1 v_1 = \mathbf{0}$, which implies $v_1 = \mathbf{0}$, as α_1 cannot be zero in this case. This leads to a contradiction).

Now from the equality 1.7, we get

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \mathbf{0},$$

where $\alpha_k \neq 0$. Therefore

$$-\alpha_1 v_1 - \alpha_2 v_2 - \cdots - \alpha_{k-1} v_{k-1} = \alpha_k v_k.$$

Thus it follow that,

$$\alpha_k^{-1}(-\alpha_1 v_1 - \alpha_2 v_2 - \cdots - \alpha_{k-1} v_{k-1}) = \alpha_k^{-1}(\alpha_k v_k).$$

Hence we have

$$-\alpha_k^{-1}(\alpha_1 v_1) - \alpha_k^{-1}(\alpha_2 v_2) - \cdots - \alpha_k^{-1}(\alpha_{k-1} v_{k-1}) = \alpha_k^{-1}(\alpha_k v_k).$$

This implies that,

$$(-\alpha_k^{-1}\alpha_1)v_1 + (-\alpha_k^{-1}\alpha_2)v_2 + \cdots + (-\alpha_k^{-1}\alpha_{k-1})v_{k-1} = (\alpha_k^{-1}\alpha_k)v_k = v_k.$$

Thus we have represented v_k as a linear combination of its preceding vectors as desired. $\square$

Problem 1.4.12. *Prove the following statements:*

- *The set of non-zero vectors $v_1, v_2, \cdots, v_n$ of $V(F)$ is linearly dependent if some v_k, $2 \leq k \leq n$, is a linear combination of the preceding ones.*

- *If in a vector space $V(F)$, a vector u is a linear combination of the set of vectors $v_1, v_2, \cdots, v_n$, then the set of vectors $u, v_1, v_2, \cdots, v_n$ is linearly dependent.*

- *Let S be a linearly independent subset of a vector space V. Suppose u is a vector in V which is not in the subspace spanned by S. Then the set obtained by adjoining u to S is linearly independent.*

- *The set of non-zero vectors $v_1, v_2, \cdots, v_n$ of $V(F)$ is linearly dependent if and only if one of these vectors is a linear combination of the remaining $(n-1)$ vectors.*

1.5 Basis of a vector space

Definition 1.5.1. *A subset S of a vector space $V(F)$ is said to be a basis of $V(F)$, if*

1. *S is linearly independent*
2. *S generates $V(F)$, that is, $L(S) = V$, that is, each vector in V is a linear combination of a finite number of elements of S.*

Example 1.5.2. *In the vector space $V_n(F)$, the set of n vectors $e_1 = (1, 0, \cdots, 0), e_2 = (0, 1, \cdots, 0), \cdots, e_n = (0, 0, \cdots, 1)$ is a basis of $V_n(F)$.*

Proof. First we have to check whether the set is linearly independent. Suppose let $\alpha_1, \alpha_2, \cdots, \alpha_n$ be a set of scalars, such that

$$\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n = \mathbf{0}.$$

Then we have

$$\begin{aligned} \mathbf{0} &= (0, 0, \cdots, 0) \\ &= \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n \\ &= \alpha_1(1, 0, \cdots, 0) + \alpha_2(0, 1, \cdots, 0) + \cdots + \alpha_n(0, 0, \cdots, 1) \\ &= (\alpha_1, 0, \cdots, 0) + (0, \alpha_2, \cdots, 0) + \cdots + (0, 0, \cdots, \alpha_n) \\ &= (\alpha_1, \alpha_2, \cdots, \alpha_n). \end{aligned}$$

Thus it follows that $\alpha_1 = 0, \alpha_2 = 0, \cdots, \alpha_n = 0$, which implies the set of n vectors

$$e_1 = (1, 0, \cdots, 0), e_2 = (0, 1, \cdots, 0), \cdots, e_n = (0, 0, \cdots, 1)$$

is linearly independent.

On the other hand, let $(v_1, v_2, \cdots, v_n)$ be an element in $V_n(F)$. Then it can be represented as

$$\begin{aligned}(v_1, v_2, \cdots, v_n) &= v_1e_1 + v_2e_2 + \cdots + v_ne_n \\ &= v_1(1, 0, \cdots, 0) + v_2(0, 1, \cdots, 0) + \cdots + v_n(0, 0, \cdots, 1) \\ &= (v_1, 0, \cdots, 0) + (0, v_2, \cdots, 0) + \cdots + (0, 0, \cdots, v_n) \\ &= (v_1, v_2, \cdots, v_n)\end{aligned}$$

Hence S is a basis of $V_n(F)$. $\square$

Note that the subset discussed in the above example is usually called as a standard basis.

Example 1.5.3.

The infinite set is a basis

$$S = \{1, x, x^2, \cdots\}$$

is a basis of the vector space $F[x]$ of polynomials over the field F. Note that no finite set of vectors forms a basis for $F[x]$.

1.6 Finite dimensional vector spaces

Definition 1.6.1. *The vector space $V(F)$ is said to be finite dimensional or finitely generated if there exists a finite subset S of V such that*

$$V = L(S).$$

Example 1.6.2. *The vector space $V_n(F)$ of n-tuples is a finite dimensional vector space.*

The vector space $F[x]$ of all polynomials over a field F is not finite dimensional; as there exists no finite subset S of $F[x]$ which spans $F[x]$. A vector space which is not finitely generated is referred as an infinite dimensional space.

Result 1.6.3. *There exists a basis for each finite dimensional vector space.*

Proof. Let $V(F)$ be a finite dimensional vector space and let $S = \{v_1, v_2, \cdots, v_n\}$ be a finite set of vectors in V such that $L(S) = V$. We may suppose that there $\mathbf{0}$ is not a member of S.

If S is linearly independent, then S itself is a basis of V. Suppose S is linearly dependent, then there exists a vector $v_i \in S$, which can be written as a linear combination of the preceding vectors $v_1, \cdots, v_{i-1}$.

Let $S' = S - \{v_i\}$, then clearly $L(S') = V$. For, $v \in V$, we have to prove that v can be written as a linear combination of finite number of elements of S'. But since $L(S) = V$, we have

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_i v_i + \cdots + \alpha_n v_n.$$

Also since $v_i \in S$, can be written as a linear combination $v_1, \cdots, v_{i-1}$, it follows that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_i(\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_{i-1} v_{i-1}) + \cdots + \alpha_n v_n.$$

as desired. In this way, any $v \in V$ can be expressed as a linear combination of the vectors in S' and therefore $L(S') = V$.

If S' is linearly independent, then S' is a required basis of V. If S' is linearly dependent, then as proceeding as above, we shall get a new set of $n-2$ vectors, that generates V. Continuing the process, up to finite number of steps, we get a linearly independent subset of S, which generates V and that forms a basis of V. $\square$

Theorem 1.6.4. *If $V(F)$ is a finite dimensional vector space, then any two bases of V have the same number of elements.*

Proof. Suppose $V(F)$ is a finite dimensional vector space. Then V definitely posses a basis. Let

$$S_1 = \{u_1, u_2, \cdots, u_m\}$$

and

$$S_2 = \{v_1, v_2, \cdots, v_n\}$$

be two bases of V. We shall prove that $m = n$. Now since $V = L(S_1)$ and $v_1 \in V$, it is possible to find scalars $\alpha_1, \alpha_2, \cdots, \alpha_m$, such that

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m = v_1.$$

Consequently the set $S_3 = \{v_1, u_1, u_2, \cdots, u_m\}$, which also generates V, is linearly dependent. Therefore there exists a member $u_i \neq v_1$ of this set S_3, such that u_i is a linear combination of the preceding vectors

$v_1, u_1, u_2, \cdots, u_{i-1}$. If we remove that vector from S_3, then it is clear to see that V can also be generated by the set containing the remaining vectors.

Let

$$S_4 = \{v_1, u_1, u_2, \cdots, u_{i-1}, u_{i+1}, \cdots, u_m\}.$$

Since $V = L(S_4)$ and $v_2 \in V$, it can be expressed as a linear combination of vectors belonging to S_4. Consequently the set

$$S_5 = \{v_2, v_1, u_1, u_2, u_{i-1}, u_{i+1}, \cdots, u_m\},$$

which also generates V, is linearly dependent. Therefore there exist a member $u_k \in S_5$, which can be expressed as a linear combination of preceding vectors.

Here note that, obviously the u_k must be different from v_1 and v_2, as the set $\{v_1, v_2\}$ is linearly independent. Therefore, if we exclude u_k from S_5, then the remaining set of vectors will also generate $V(F)$.

In the above process, it is easy note that, at each step there is an exclusion of a vector u_i from S_1 and an inclusion of an vector v_i to S_1.

By continuing the process, obviously the set S_1 of u_i's cannot be exhausted before the S_2 of v_i, otherwise V will be a linear span of a proper subset of S_2 and thus S_2 will become linearly dependent. Therefore we have $m \not< n$.

Interchanging the roles of S_1 and S_2, we get $n \not< m$. Hence $m = n$. $\square$

Example 1.6.5. *If we let $\mathbb{R}^3$ as a vector space over $\mathbb{R}$, then the sets*

$$S_1 = \{(1,0,0), (0,1,0), (0,0,1)\}$$

and

$$S_2 = \{(1,0,0),(1,1,0),(1,1,1)\}$$

both forms a basis for $\mathbb{R}^3$.

Dimension of a vector space:

Definition 1.6.6. *Let* $V(F)$ *be a finitely generated (finite dimensional) vector space. Then the dimension of* $V(F)$ *is defined as the number of elements in any basis of* $V(F)$. *We denote dimension of* $V(F)$ *as* $\dim V$.

The vector space $V_n(F)$ is of dimension n. If a field is regarded as a vector space over itself then the dimension of F is one.

Some theorems of finite dimensional vector space

- Every linearly independent subset of a finitely generated vector space $V(F)$ forms a part of a basis of V. In other words, every linearly independent subset of a finitely generated vector space $V(F)$ is either a basis of V or can be extended to form a basis.

- Each set of $(n+1)$ or more vectors of finite dimensional vector space $V(F)$ of dimension n is linearly dependent.

- Let V be a vector space which is spanned by a finite set of vectors $v_1, v_2, \cdots, v_m$. Then any linearly independent set of vectors in V is finite and contains no more than m vectors.

- If $V(F)$ is a finite dimensional vector space of dimension n, then any set of n linearly independent vectors in V forms a basis of V.

- If a set of n vectors of a finite dimensional vector space $V(F)$ of dimension n generates $V(F)$, then S is a basis of V.

- Let $S = \{v_1, v_2, \cdots, v_n\}$ be a basis of a finite dimensional vector space $V(F)$ of dimension n. Then every element $u \in V$ can be expressed as

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n,$$

  where $\alpha_i \in F$ for all $i = 1, 2, \cdots, n$.

- Each subspace W of a finite dimensional vector space $V(F)$ of dimension n is a finite dimensional space with $\dim m \leq n$.

- If W_1 and W_2 are two subspaces of a finite dimensional vector space $V(F)$, then

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Problems

1. Let V be the vector space of all 2×2 matrices over the field F. Prove that V has dimension 4 by exhibiting a basis for V that has 4 elements.

2. Show that if $S = \{v_1, v_2, v_3\}$ is a basis of $\mathbb{C}^3(\mathbb{C})$, then the set $S' = \{v_1 + v_2, v_2 + v_3, v_3 + v_1\}$ is also a basis of $\mathbb{C}^3(\mathbb{C})$.

3. If W_1 and W_2 are two subspaces of a finite dimensional vector space $V(F)$, with same dimension and if $W_1 \subseteq W_2$, then $W_1 = W_2$.

4. Show that the set of vectors $(1, 2, 1), (2, 1, 0), (1, -1, 2)$ forms a basis of $\mathbb{R}^3$.

5. Let V be the vector space of ordered pairs of complex numbers over the real field $\mathbb{R}$. Show that the set $\{(1, 0), (i, 0), (0, 1), (0, i)\}$ is a basis for V.

6. In the vector space $\mathbb{R}^3$, let $v_1 = (1, 2, 1)$, $v_2 = (3, 1, 5)$, $v_3 = (3, -4, 7)$. Show that the subspaces spanned by $S = \{v_1, v_2\}$ and $T = \{v_1, v_2, v_3\}$ are the same.

7. Prove that the space of all $m \times n$ matrices over the field F has dimension mn, by exhibiting a basis for this space.

8. If a vector space V is spanned by a finite set of m vectors, then show that any linearly independent set of vectors in V has atmost m elements.

1.7 Linear Transformations

Definition 1.7.1. *Let $V(F)$ and $V'(F)$ be two vector spaces. Then a mapping $T : V \to V'$ is called a homomorphism or a linear transformation of V to W if*

1. *$T(v_1 + v_2) = T(v_1) + T(v_2)$, for all $v_1, v_2 \in V$;*

2. *$T(\alpha v_1) = \alpha T(v_1)$, for all $\alpha \in F$ and $v_1 \in V$.*

The conditions (1) and (2) can be combined into a single condition

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2),$$

for all $v_1, v_2 \in V$ and for all $\alpha, \beta \in F$.

If T is a homomorphism of V onto V', then V' is called a homomorphic image of V.

Linear operator: Let $V(F)$ be a vector space. A linear operator on V is a function T from V to V such that

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$$

for all $\alpha, \beta \in F$ and for all $v_1, v_2 \in V$.

Theorem 1.7.2. *If T is a homomorphism of $V(F)$ to $V'(F)$, then*

1. *$T(\mathbf{0}) = \mathbf{0}'$, where $\mathbf{0}$ and $\mathbf{0}'$ are the zero vectors of V and V' respectively.*

2. *$T(-v) = -T(v)$, for all $v \in V$.*

Problems

Prove the following:

1. The mapping $T : V_3(F) \to V_2(F)$ defined by $T((a_1, a_2, a_3)) = (a_1, a_2)$ is a homomorphism of $V_3(F)$ onto $V_2(F)$.

2. Let $V(F)$ be the vector space of $m \times n$ matrices over the field F. Let P be a fixed $m \times m$ matrix over F, and let Q be a fixed $n \times n$ matrix over F, then T from V to V defined by

$$T(A) = PAQ$$

for all $A \in V$ is a linear operator on V.

3. Let $V(F)$ be the vector space of all polynomials over the field F. Let $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ in V be a polynomial of degree n in determinate x. Let us define

$$Df(x) = a_1 + 2a_2x + \cdots + na_nx^{n-1}$$

if $n > 1$ and $Df(x) = 0$ if $f(x)$ is a constant polynomial, then D from V into V is a linear operator on V.

4. Let $V(\mathbb{R})$ be the vector space of all continuous from $\mathbb{R}$ to $\mathbb{R}$. For any $f \in V$ define T by $(Tf)(x) = \int\limits_0^x f(t)dt$ for all $x \in \mathbb{R}$, then T is linear transformation from V to V.

Zero linear transformation: Let V and V' be two vector spaces over a field F. The function $T : V \to V'$ defined by $T(v) = \mathbf{0}'$ for all $v \in V$, is a linear transformation from V to V'.

Identity operator: Let $V(F)$ be a vector space. The function $I : V \to V$ defined by $I(v) = v$ for all $v \in V$, is a linear transformation from V to V.

Definition 1.7.3 (Range). *Let V and V' be two vector spaces over a field F and let T be a linear transformation from V to V'. Then the range of T written as $R(T)$ is the set of all vectors $v' \in V'$ such that $T(v) = v'$ for some $v \in V$. That is*

$$R(T) = \{T(v) \in V'/v \in V\}.$$

Definition 1.7.4 (Null space). *Let V and V' be two vector spaces over a field F and let T be a linear transformation from V to V'. Then the null space of T written as $N(T)$ is the set of all vectors $v \in V$ such that $T(v) = \mathbf{0}'$, where $\mathbf{0}'$ is the vector of V'. That is,*

$$N(T) = \{v \in V/T(v) = \mathbf{0}' \in V'\}.$$

Theorem 1.7.5. *Let $V(F)$ and $V'(F)$ be two vector spaces and let T be a linear transformation from V to V', then the kernel of T and the range of T are subspaces of V and V' respectively.*

Theorem 1.7.6. *Let T be a linear transformation from a vector space $V(F)$ into a vector space $V'(F)$. If V is finite dimensional, then the range of T is a finite dimensional subspace of V'.*

Rank and nullity of a linear transformation:

The dimension of range of T is called a rank of T and the dimension of kernel of T is called a nullity of T.

Theorem 1.7.7. *Let V and V' be vector spaces over the field F and let T be a linear transformation from V to V'. Suppose that V is finite dimensional. Then*

$$\dim V = rank(T) + nullity(T).$$

Proof. Let $\dim V = n$ and let $N(T)$ be the null space of T. Then $N(T)$ is a subspace of V. Since V is finite dimensional, clearly $N(T)$ is finite dimensional subspace.

Let $\dim N(T) = k$. Suppose let $S = \{v_1, v_2, \cdots, v_k\}$ be a basis for $N(T)$. Since S is a basis of $N(T)$, it is linearly independent.

But we know that any linear independent subset of a vector space can be extended as a basis. Therefore S can be extended to form some basis, say S' of V.

Let

$$S' = \{v_1, v_2, \cdots, v_k, v_{k+1}, v_{k+2}, \cdots, v_n\}.$$

Here note that the vectors $T(v_1), T(v_2), \cdots, T(v_k), T(v_{k+1}), \cdots, T(v_n)$ are in range of T. We claim that, the set of vectors $\{T(v_{k+1}), T(v_{k+2}), \cdots, T(v_n)\}$, forms a basis of $R(T)$.

First we shall prove that the vectors $T(v_{k+1}), \cdots, T(v_n)$ span the range of T. For let $v' \in R(T)$, then there exists $v \in V$ such that

$$T(v) = v'.$$

Now since $v \in V$, there exists $\alpha_1, \alpha_2, \cdots, \alpha_n \in F$, such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k + \alpha_{k+1} v_{k+1} + \alpha_{k+2} v_{k+2} + \cdots + \alpha_n v_n.$$

This implies

$$T(v) = T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k + \alpha_{k+1} v_{k+1} + \alpha_{k+2} v_{k+2} + \cdots + \alpha_n v_n).$$

Thus it results that,

$$T(v) = v' = T(\alpha_1 v_1) + T(\alpha_2 v_2) + \cdots + T(\alpha_k v_k) + T(\alpha_{k+1} v_{k+1}) + \cdots + T(\alpha_n v_n).$$

But since $T(\alpha_1 v_1) = T(\alpha_2 v_2) = \cdots = T(\alpha_k v_k) = \mathbf{0}'$, we have

$$T(v) = v' = \alpha_{k+1} T(v_{k+1}) + \cdots + \alpha_n T(v_n).$$

Therefore the set of vectors $T(v_{k+1}), T(v_{k+2}), \cdots, T(v_n)$ spans $R(T)$. We shall claim that the vectors $T(v_{k+1}), T(v_{k+2}), \cdots, T(v_n)$, are linearly independent. Suppose let $\alpha_{k+1}, \alpha_{k+2}, \cdots, \alpha_n \in F$ such that

$$\alpha_{k+1} T(v_{k+1}) + \cdots + \alpha_n T(v_n) = \mathbf{0}'.$$

Then we have

$$T(\alpha_{k+1} v_{k+1} + \alpha_{k+2} v_{k+2} + \cdots + \alpha_n v_n) = \mathbf{0}'.$$

Thus it follow that, the vector

$$\alpha_{k+1} v_{k+1} + \alpha_{k+2} v_{k+2} + \cdots + \alpha_n v_n \in N(T).$$

(Since each vector in $N(T)$ can be expressed as a linear combination of vectors $\{v_1, v_2, \cdots, v_k\}$ as the set forms a basis of $N(T)$.) Therefore there exist scalars $\beta_1, \beta_2, \cdots, \beta_k$ such that

$$\alpha_{k+1}v_{k+1} + \alpha_{k+2}v_{k+2} + \cdots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_k v_k.$$

Thus we get that,

$$(-\beta_1)v_1 + (-\beta_2)v_2 + \cdots + (-\beta_k)v_k + \alpha_{k+1}v_{k+1} + \alpha_{k+2}v_{k+2} + \cdots + \alpha_n v_n = \mathbf{0}'.$$

But since the vectors $v_1, v_2, \cdots, v_k, v_{k+1}, v_{k+2}, \cdots, v_n$ are linearly independent, it follow that the scalars

$$\beta_1 = \beta_2 = \cdots = \beta_k = \alpha_{k+1} = \alpha_{k+2} = \cdots = \alpha_n = 0.$$

Thus the set $\{T(v_{k+1}), T(v_{k+2}), \cdots, T(v_n)\}$ is linearly independent and hence the set

$$\{T(v_{k+1}), T(v_{k+2}), \cdots, T(v_n)\}$$

forms a basis for $R(T)$. Therefore

$$\dim R(T) = n - k,$$

which implies

$$n = \dim R(T) + k.$$

Thus we get that

$$\dim V = rank(T) + nullity(T).$$

$\square$

Problems

1. Show that the mapping $T : V_3(\mathbb{R}) \to V_2(\mathbb{R})$ defined as

$$T((a_1, a_2, a_3)) = (3a_1 - 2a_2 + a_3, a_1 - 3a_2 - 2a_2)$$

is a linear transformation from $V_3(\mathbb{R})$ to $V_2(\mathbb{R})$.

2. Show that the mapping $T : V_2(\mathbb{R}) \to V_3(\mathbb{R})$ defined as

$$T((a, b)) = (a + b, a - b, b)$$

is a linear transformation from $V_2(\mathbb{R})$ to $V_3(\mathbb{R})$. Find the range, rank, null space and nullity of T.

3. Let V be the vector space of all $n \times n$ matrices over a field F and let B be a fixed $n \times n$ matrix. If

$$T(A) = AB - BA \text{ for all } A \in V,$$

then verify that T is a linear transformation from V to V.

4. Let $V(F)$ and $V'(F)$ be two vector spaces and let T_1 and T_2 be two linear transformations from V to V'. Let α, β be given elements of F. Then the mapping T defined as

$$T(v) = \alpha T_1(v) + \beta T_2(v)$$

for all $v \in V$ is a linear transformation from V to V'.

5. Let V be a vector space and T a linear transformation from V to V. Prove that the following two statements about T are equivalent:

   (a) The intersection of the range of T and the null space of T is the zero subspace of V, that is $R(T) \cap N(T) = \{\mathbf{0}\}$.

   (b) $T(T(v)) = \mathbf{0}$ implies $T(v) = \mathbf{0}$.

6. Show that the mapping $T : V_3(\mathbb{R}) \to V_2(\mathbb{R})$ defined as

$$T(a_1, a_2, a_3) = (a_1 - a_2, a_1 - a_3)$$

   is a linear transformation.

7. Show that the mapping $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined as

$$T(a, b) = (a - b, b - a, -a)$$

   is a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$. Find the range, rank, null space and nullity of T.

8. Let F be a subfield of complex numbers and let T be the function from F^3 to F^3 defined by

$$T(a, b, c) = (a - b + 2c, 2a + b, -a - 2b + 2c).$$

   Show that T is a linear transformation. Find also the rank and the nullity of T.

9. Which of the following functions T from $\mathbb{R}^2$ to $\mathbb{R}^2$ are linear transformations?

   (a) T(a,b)=(1+a,b);

   (b) T(a,b)=(b,a);

   (c) T(a,b)=(a+b,a).

10. Let V be the space of $n \times 1$ matrices over a field F and let W be the space of $m \times 1$ matrices over F. Let A be a fixed $m \times n$ matrix over F and let T be the linear transformation from V to W defined by $T(X) = AX$. Prove that T is the zero transformation if and only if A is the zero matrix.

Note: Let $L(V, V')$ be the set of all linear transformations from a vector space $V(F)$ to $V'(F)$. Then $L(V, V')$ is a vector space over the same field F. For this purpose we shall have to suitably define addition in $L(V, V')$ and scalar multiplication in $L(V, V')$ over F.

Theorem 1.7.8. *Let V and V' be vector spaces over the field F. Let T_1 and T_2 be linear transformations from V to V'. The function $T_1 + T_2$ defined by*

$$(T_1 + T_2)(v) = T_1(v) + T_2(v), \text{ for all } v \in V$$

is a linear transformation from V to V'. If α is any element of F, the function αT defined by

$$(\alpha T)(v) = \alpha(T(v)), \text{ for all } v \in V$$

is a linear transformation. in addition, the set of all linear transformation from V to V', together with the addition and scalar multiplication defined above is a vector space over the field F.

Definition 1.7.9. *Let V and V' be vector spaces over the field F. Let T be linear transformations from V to V' such that T is one-one and onto. Then T is called invertible.*

Definition 1.7.10. *Let T be linear transformations from $V(F)$ to $V'(F)$, then T is said to be non-singular, if the null space of T consists of the zero vector alone. That is, if*

$$v \in V \text{ and } T(v) = \mathbf{0}' \text{ implies } v = \mathbf{0}.$$

If T is not non-singular, that is, if there exists

$$v \neq \mathbf{0} \in V \text{ such that } T(v) = \mathbf{0}',$$

then we say T to be singular.

Result 1.7.11. *Let T be a linear transformation from a vector space $V(F)$ to $V(F)$, then T is non-singular if and only if T is invertible.*

Problem 1.7.12. *Describe explicitly the linear transformation T from F^2 to F^2 such that $T((e_1)) = (a, b)$ and $T((e_2)) = (c, d)$ where $e_1 = (1, 0)$ and $e_2 = (0, 1)$.*

Solution: Let (x_1, x_2) be any element in F^2. Given that $T((1,0)) = (a,b)$ and $T((0,1)) = (c,d)$. Now since $\{e_1, e_2\}$ forms a basis for the vector space F^2, any element in F^2 can be expressed as the linear combination of e_1 and e_2. Therefore we have

$$(x_1, x_2) = x_1(1,0) + x_2(0,1) = x_1e_1 + x_2e_2.$$

Thus we get that

$$\begin{aligned} T(x_1, x_2) &= T(x_1e_1 + x_2e_2) \\ &= x_1T(e_1) + x_2T(e_2) \\ &= x_1(a,b) + x_2(c,d) \\ &= (x_1a + x_2c, x_1b + x_2d). \end{aligned}$$

Problems

1. Describe explicitly the linear transformation T from $\mathbb{R}^2$ to $\mathbb{R}^2$ such that $T((2,3)) = (4,5)$ and $T((1,0)) = (0,0)$.

2. Find a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T((1,0)) = (1,1)$ and $T((0,1)) = (-1,2)$. Prove that T maps the square with vertices $(0,0)$, $(1,0)$, $(1,1)$ and $(0,1)$ into a parallelogram.

3. Describe explicitly a linear transformation from $V_3(\mathbb{R})$ to $V_3(\mathbb{R})$ which has its range the subspace spanned by $(1,0,-1)$ and $(1,2,2)$.

4. Let T be a linear transformation $V_3(\mathbb{R})$ to $V_3(\mathbb{R})$ defined by

$$T(a,b,c) = (3a, a-b, 2a+b+c)$$

for all $(a,b,c) \in V_3(\mathbb{R})$. Is T invertible? If so, find T^{-1}. Also prove that $(T^2 - I)(T - 3I) = \widehat{\mathbf{0}}$.

5. If A is a linear transformation on a vector space V such that $A^2 - A + I = \widehat{\mathbf{0}}$, then A is invertible.

6. Let V be a vector space over the field F and T be linear operator on V. If $T^2 = \widehat{\mathbf{0}}$, what can you say about the relation of the range of T to the null space of T? Give an example of a linear operator T on $V_2(\mathbb{R})$ such that $T^2 = \widehat{\mathbf{0}}$ but $T = \widehat{\mathbf{0}}$.

7. Let V be a finite dimensional vector space and T be a linear operator on V. Suppose $rank(T^2) = rank(T)$. Prove that the range and null space of T have only the zero vector in common.

Definition 1.7.13. *Let $V(F)$ and $V'(F)$ be two vector spaces. Then a mapping $T : V \to V'$ is called an isomorphism of V to V' if T is one-to-one. Two vector spaces V and V' are said to be isomorphic if there exists an isomorphism V onto V'. Symbolically, we write $V \cong V'$.*

Problem 1.7.14. *Let $T : V \to V'$ be a linear transformation. Then kernel of T is a subspace of V and that T is an isomorphism if and only if its kernel is $\{\mathbf{0}'\}$.*

Theorem 1.7.15. *Two finite dimensional vector spaces over the same field are isomorphic if only if they are of the same dimension.*

Theorem 1.7.16. *Every n-dimensional vector space $V(F)$ is isomorphic to $V_n(F)$.*

Proof. Let $\{v_1, v_2, \cdots, v_n\}$ be any basis of $V(F)$. Then every vector $v \in V$ can be uniquely expressed as $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$, where $\alpha_i \in F$. Here note that the n-tuple $(\alpha_1, \alpha_2, \cdots, \alpha_n) \in V_n(F)$. Let $T : V(F) \to V_n(F)$ defined by

$$T(v) = T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n).$$

Since in the representation of v as a linear combination of $v_1, v_2, \cdots, v_n$, the scalars $\alpha_1, \alpha_2, \cdots, \alpha_n$ are unique, therefore $T(v)$ is a unique element of $V_n(F)$ and thus the mapping T is well defined. We wish to show that T is an isomorphism of $V(F)$ onto $V_n(F)$.

First we claim that T is one-to-one.

Let $u, v \in V(F)$, such that $T(u) = T(v)$. Since $u, v \in V(F)$, we have $u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ and $v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$. Now consider,

$T(u) = T(v)$

$$\begin{aligned}
&\Rightarrow \quad T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = T(\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n) \\
&\Rightarrow \quad (\alpha_1, \alpha_2, \cdots, \alpha_n) = (\beta_1, \beta_2, \cdots, \beta_n) \\
&\Rightarrow \quad \alpha_1 = \beta_1, \alpha_2 = \beta_2, \cdots, \alpha_n = \beta_n \\
&\Rightarrow \quad u = v.
\end{aligned}$$

Therefore f is one-to-one. Clearly, T is onto, as for any $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ in $V_n(F)$, there exists an element $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ in $V(F)$ such that

$$T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n).$$

Finally we claim that, T is a linear transformation. For, let $\alpha, \beta \in F$ and $u, v \in V(F)$, then we have

$T(\alpha u + \beta v)$

$$\begin{aligned}
&= \quad T(\alpha(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) + \beta(\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n)) \\
&= \quad T((\alpha\alpha_1 + \beta\beta_1)v_1 + (\alpha\alpha_2 + \beta\beta_2)v_2 + \cdots + (\alpha\alpha_n + \beta\beta_n)v_n) \\
&= \quad (\alpha\alpha_1 + \beta\beta_1, \alpha\alpha_2 + \beta\beta_2, \cdots, \alpha\alpha_n + \beta\beta_n) \\
&= \quad (\alpha\alpha_1, \alpha\alpha_2, \cdots, \alpha\alpha_n) + (\beta\beta_1, \beta\beta_2, \cdots, \beta\beta_n) \\
&= \quad \alpha(\alpha_1, \alpha_2, \cdots, \alpha_n) + \beta(\beta_1, \beta_2, \cdots, \beta_n) \\
&= \quad \alpha T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) + \beta T(\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n) \\
&= \quad \alpha T(u) + \beta T(v).
\end{aligned}$$

Thus T is a linear transformation and hence T is an isomorphism of $V(F)$ onto $V_n(F)$. That is, $V(F) \cong V_n(F)$. $\square$

Problem 1.7.17. *If $V(F)$ is a finite dimensional vector space and T is an isomorphism of V to V, prove that T just maps V onto V.*

Proof. Let $V(F)$ be finite dimensional vector space of dimension n. Let T be an isomorphism of V to V. That is, T is a linear transformation and T is one-to-one.

We have to prove that, T is onto. For, let $S = \{v_1, v_2, \cdots, v_n\}$ be a basis of V. We shall first prove that $S' = \{T(v_1), T(v_2), \cdots, T(v_n)\}$, is also a basis of V; that is, we claim that, S' is linearly independent and $L(S') = V$.

Suppose there exists $\alpha_1, \alpha_2, \cdots, \alpha_n$, such that

$$\alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n) = \mathbf{0}.$$

Then we get

$$\begin{aligned} T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = \mathbf{0} \\ \Rightarrow \quad \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \mathbf{0} \\ \Rightarrow \quad \alpha_1 = \alpha_2 = \cdots = \alpha_n = \mathbf{0}. \end{aligned}$$

Therefore S' is linearly independent. Now as V is of dimension n and S' is a linearly independent subset of V containing n vectors. Therefore S' must be a basis of V. Therefore each vector in V can be expressed as a linear combination of the vectors belonging to S'.

Now we shall show that T is onto. Let v be any element in V. Then there exist scalars $\alpha_1, \alpha_2, \cdots, \alpha_n$ such that

$$v = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n).$$

But as T is a linear transformation, we get

$$v = T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n).$$

Thus the image of $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in V$ under T is v, which implies T is onto. $\square$

Problems

1. Let $V(\mathbb{R})$ be the vector space of all complex numbers $a + ib$ over the field of reals $\mathbb{R}$ and let T be the mapping from $V(\mathbb{R}) \to V_2(\mathbb{R})$ defined as $T(a + ib) = (a, b)$. Show T is an isomorphism.

2. Let $V(F)$ and $W(F)$ be two finite dimensional vector spaces whose dimensions are same. If T is an isomorphism of V into W. Prove that T maps V onto W.

3. If V is a finite dimensional vector space and T is a homomorphism of V onto V, then prove that T is one-to-one.

4. If V is finite dimensional and T is a homomorphism of V to itself which is not onto, then prove that there is some $v \neq \mathbf{0}$ in V such that $T(v) = \mathbf{0}$.

Quotient Space:

Let W be any subspace of a vector space $V(F)$. Let $v \in V$. Then the set

$$W + v = \{w + v/w \in W\}$$

is called a right coset of W in V generated by v. Similarly,

$$v + W = \{v + w/w \in W\}$$

is called a left coset of W in V generated by v.

Obviously $W + v$ and $v + W$ are both subsets of V. Since addition in V is commutative, therefore we have $W + v = v + W$. Hence we shall call $W + v$ is simply a coset of W in V generated by v.

Result 1.7.18. *If W is a subspace of a vector space $V(F)$, then the set V/W of all coset $W + v$, where $v \in V$, is a vector space over the field F, with the addition and scalar multiplication defined as follows:*

$$(W + v) + (W + v') = W + (v + v')$$

for all $v, v' \in V$ and

$$\alpha(W + v) = W + \alpha v$$

for all $v \in V$ and $\alpha \in F$.

Dimension of Quotient Space:

Theorem 1.7.19. *If W is a subspace of a finite dimensional vector space $V(F)$, then* $\dim(V/W) = \dim V - \dim W$.

Fundamental Theorem of Vector Space Homomorphism

Theorem 1.7.20. *Let V and V' be vector spaces over a field F and let T be a linear transformation from V onto V'. Let W be the kernel of T, then $V/W \cong V'$.*

Proof. First we have to prove that W is a subspace of V. That is to prove that for any $w_1, w_2 \in W$ and $\alpha, \beta \in F$, $\alpha w_1 + \beta w_2 \in W$. For, let $w_1, w_2 \in W$ and $\alpha, \beta \in F$. Then clearly $T(w_1) = T(w_2) = \mathbf{0}'$, where $\mathbf{0}'$ is the additive identity of V'. Now consider,

$$\begin{aligned} T(\alpha w_1 + \beta w_2) &= \alpha T(w_1) + \beta T(w_2) \\ &= \alpha \mathbf{0}' + \beta \mathbf{0}' \\ &= \mathbf{0}'. \end{aligned}$$

Therefore $\alpha w_1 + \beta w_2 \in W$ and thus W is a subspace of V. (Thus talking about V/W is meaningful).

Now define a map $\varphi : V/W \to V'$ by

$$\varphi(W + v) = T(v).$$

Then clearly φ is well defined. For, if we let $W + u = W + v$, for some $u, v \in V$ then clearly, $u \in W + v$ and $v \in W + u$. Therefore u can be written as $u = w_1 + v$ for some $w_1 \in W$ and in a similar way v can be written as

$v = w_2 + u$ for some $w_2 \in W$. Now consider,

$$\begin{aligned}\varphi(W+u) &= T(u)\\ &= T(w_1+v)\\ &= T(w_1)+T(v)\\ &= \mathbf{0}'+T(v)\\ &= T(v)\\ &= \varphi(W+v).\end{aligned}$$

Now we claim that φ is one-one. For, let $W+u$, $W+v$ be in V/W, such that $\varphi(W+u) = \varphi(W+v)$. Then we have to prove that $W+u = W+v$. Now consider

$$\begin{aligned}&\varphi(W+u) = \varphi(W+v)\\ &\qquad \Rightarrow \; T(u) = T(v)\\ &\qquad \Rightarrow \; T(u) - T(v) = \mathbf{0}'\\ &\qquad \Rightarrow \; T(u-v) = \mathbf{0}'\end{aligned}$$

This implies $u - v \in ker(T) = W$ and hence $W+u = W+v$.

Next we claim that, φ is onto. Let $v' \in V'$. Since T is onto there must exists a $v \in V$ such that $T(v) = v'$ but by the definition of φ, we have $\varphi(W+v) = T(v)$. Thus for any $v' \in V'$ there exists a $W+v \in V/W$ such that $\varphi(W+v) = v'$.

Finally, we prove that φ is a linear transformation. For, let $W+u$, $W+v$ be in V/W, then we have

$$\begin{aligned}
\varphi((W+u)+(W+v)) &= \varphi(W+(u+v)) \\
&= T(u+v) \\
&= T(u)+T(v) \\
&= \varphi(W+u)+\varphi(W+v).
\end{aligned}$$

Also, we have

$$\begin{aligned}
\varphi(\alpha(W+u)) &= \varphi(W+\alpha u) \\
&= T(\alpha u) \\
&= \alpha T(u) \\
&= \alpha\varphi(W+u).
\end{aligned}$$

Thus φ is an isomorphism from V/W to V'. □

Row and column rank of a matrix:

Let $A=[a_{ij}]_{m\times n}$ be an $m\times n$ matrix over a field F. The row vectors of A are the vectors $v_1, v_2, \cdots, v_m \in V_n(F)$ defined by $v_i=(a_{i1}, a_{i2}, \cdots, a_{in})$, $1\le i\le m$.

The row space of A is the subspace of $V_n(F)$ spanned by these vectors. The row rank of A is the dimension of the row space of A.

The column vectors of A are the vectors $u_1, u_2, \cdots, u_n \in V_m(F)$ defined by $u_j=(a_{1j}, a_{2j}, \cdots, a_{mj})$, $1\le j\le m$.

The column space of A is the subspace of $V_m(F)$ spanned by these vectors. The column rank of A is the dimension of the column space of A.

Here let us recall that, if A is a non-zero row reduced echelon matrix, then the non-zero row vectors of A are linearly independent and therefore they form a basis for the row space of A; in order to find the row rank of a matrix A, it is enough to find the reduced row echelon form of A by elementary row operations and the number of non zero rows gives us the row rank of A.

Result 1.7.21. *The row and the column ranks of any matrix are equal.*

Proof. Let $A = (a_{ij})$ be an $m \times n$ matrix. Let $R_1, R_2, \cdots, R_m$ denote the rows of A, then clearly, $R_i = (a_{i1}, a_{i2}, \cdots, a_{in})$.

Suppose the row rank of A is r. Then the dimension of the row space is r. Let $v_1 = (b_{11}, b_{12,} \cdots, b_{1n}), v_2 = (b_{21}, b_{22,} \cdots, b_{2n}), \cdots, v_r = (b_{r1}, b_{r2,} \cdots, b_{rn})$ be a basis for the row space of A.

Then each row is a linear combination of the vectors $v_1, v_2, \cdots, v_r$. Let

$$\begin{aligned}
R_1 &= k_{11}v_1 + k_{12}v_2 + \cdots + k_{1r}v_r \\
R_2 &= k_{21}v_1 + k_{12}v_2 + \cdots + k_{1r}v_r \\
\vdots &= \vdots \qquad \vdots \\
R_m &= k_{m1}v_1 + k_{m2}v_2 + \cdots + k_{mr}v_r
\end{aligned}$$

where $k_{ij} \in F$. Equating the i^{th} component of each of the above equations, we get

$$\begin{aligned}
a_{1i} &= k_{11}b_{1i} + k_{12}b_{2i} + \cdots + k_{1r}b_{ri} \\
a_{2i} &= k_{21}b_{1i} + k_{22}b_{2i} + \cdots + k_{2r}b_{ri} \\
\vdots &= \vdots \qquad \vdots \\
a_{mi} &= k_{m1}b_{1i} + k_{m2}b_{2i} + \cdots + k_{mr}b_{ri}
\end{aligned}$$

Hence

$$\begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} = b_{1i} \begin{pmatrix} k_{11} \\ k_{21} \\ \vdots \\ k_{m1} \end{pmatrix} + b_{2i} \begin{pmatrix} k_{12} \\ k_{22} \\ \vdots \\ k_{m2} \end{pmatrix} + \cdots + b_{ri} \begin{pmatrix} k_{1r} \\ k_{2r} \\ \vdots \\ k_{mr} \end{pmatrix}.$$

Thus each column of A is a linear combination of r vectors and hence the dimension of column space $\leq r$. Therefore,

$$\text{column rank of } A \leq r = \text{ row rank of } A.$$

Similarly, we can prove that

$$\text{row rank of } A \leq \text{ column rank of } A.$$

Hence the row rank and the column rank of A are equal. $\square$

Matrix of a linear transformation:

Let U be an n-dimensional vector space over the field F and let V be an m-dimensional vector space over F. Let

$$B = \{u_1, u_2, \cdots, u_n\} \text{ and } B' = \{v_1, v_2, \cdots, v_m\}.$$

be ordered basis for U and V respectively.

Suppose T is a linear transformation from U to V. We know that T is determined by its action on the vectors u_j belonging to the basis B of U and each of the n vectors $T(u_j)$ can be uniquely expressed as a linear combination of $v_1, v_2, \cdots, v_m$, as $T(u_j) \in V$ and these m vectors form a basis for V. That is,

$$T(u_j) = a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{mj}v_m = \sum_{i=1}^{m} a_{ij}v_i,$$

for all $j = 1, 2, \cdots, n$.

The scalars $a_{1j}, a_{2j}, \cdots, a_{mj}$ are the coordinates of $T(u_j)$ with respect to the ordered basis B'.

The $m \times n$ matrix whose jth column consists of these coordinates is called **the matrix of the linear transformation T relative to the pair of ordered bases B and B'**. We shall denote it by the symbol $[T; B; B']$ or simply by $[T]$ if the bases are understood.

Note:

Let T be a linear transformation from an n-dimensional vector space $V(F)$ into itself. Then in order to represent T by a matrix, it is most convenient to use the same ordered basis in each case, i.e., to take $B = B'$. The representing matrix will then be called the matrix of T relative to ordered basis B and will be denoted by $[T; B]$ or sometimes also by $[T]_B$.

Problem 1.7.22. *Let T be a linear transformation on the vector space $V_2(F)$*

defined by $T(a,b) = (a,0)$. *Write the matrix of* T *relative to the standard ordered basis* $V_2(F)$.

Solution: Let $B = \{v_1, v_2\}$ be the standard ordered basis for $V_2(F)$. Then $v_1 = (1,0)$ and $v_2 = (0,1)$. Also we have $T(v_1) = T((1,0)) = (1,0)$ and $T(v_2) = T((0,1)) = (0,0)$.

Now if we express $T(v_1)$ and $T(v_2)$ as a linear combination of vectors in B, then we get

$$T(v_1) = T((1,0)) = (1,0) = 1(1,0) + 0(0,1) = 1(v_1) + 0(v_2)$$

and

$$T(v_2) = T((0,1)) = (0,0) = 0(1,0) + 0(0,1) = 0(v_1) + 0(v_2).$$

Thus $1,0$ are the coordinates of $T(v_1)$ with respect to the ordered basis B and $0,0$ are the coordinates of $T(v_2)$ with respect to the ordered basis B. Hence $1,0$ forms the first column and $0,0$ forms the second column of matrix of T relative to ordered basis B. That is,

$$[T]_B = [T;B] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Problem 1.7.23. *Let* $V(\mathbb{R})$ *be the vector space of all polynomials in* x *with co-efficients in* $\mathbb{R}$ *of the form*

$$f(x) = a_0x^0 + a_1x^1 + a_2x^2 + a_3x^3,$$

i.e., the space of polynomials of degree three or less. The difference operator D is a linear transformation on V. The set $B = \{x^0, x^1, x^2, x^3\}$, is an ordered basis for V. Write the matrix of D relative to the ordered basis B.

Solution: Given the set $B = \{x^0, x^1, x^2, x^3\}$, is an ordered basis for V and D is a linear transformation. Thus we have

$$D(x^0) = 0 = 0x^0 + 0x^1 + 0x^2 + 0x^3;$$

$$D(x^1) = x^0 = 1x^0 + 0x^1 + 0x^2 + 0x^3;$$

$$D(x^2) = 2x^1 = 0x^0 + 1x^1 + 0x^2 + 0x^3;$$

$$D(x^3) = 3x^2 = 0x^0 + 0x^1 + 2x^2 + 0x^3.$$

Therefore the matrix of D relative to the ordered basis B is given by

$$[D; B] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Problems

1. Find the matrix of the linear transformation T on $V_3(\mathbb{R})$ to $V_3(\mathbb{R})$ defined as $T(a, b, c) = (2b + c, a - 4b, 3a)$, with respect to the ordered basis B and also with respect to the ordered basis B' where

   - $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$;
   - $B' = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$.

2. Let T be the linear operator on $\mathbb{R}^3$ defined by

$$T(x_1, x_2, x_3) = (3x_1 + x_2, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3).$$

Find the matrix of T with respect to the ordered basis $\{(1,0,1), (-1,2,1), (2,1,1)\}$.

3. Consider the vector space $V(\mathbb{R})$ of all 2×2 matrices over the field $\mathbb{R}$ of real numbers. Let T be the linear transformation on V that sends each matrix X onto AX, where $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Find the matrix of T with respect to the ordered basis $B = \{v_1, v_2, v_3, v_4\}$ for V where $v_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $v_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $v_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $v_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

4. If the matrix of a linear transformation T on $V_2(\mathbb{C})$, with respect to the ordered basis $B = \{(1,0), (0,1)\}$ is $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Find the matrix of T with respect to the ordered basis $B' = \{(1,1), (1,-1)\}$.

5. Find the matrix, relative to the basis $\{(\frac{2}{3}, \frac{2}{3}, \frac{-1}{3}), (\frac{1}{3}, \frac{-2}{3}, \frac{-2}{3}), (\frac{2}{3}, \frac{-1}{3}, \frac{2}{3})\}$ of $\mathbb{R}^3$, of the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ whose matrix relative to the standard basis is

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

6. Let V be the space of all 2×2 matrices over the field F and let P be a fixed 2×2 matrix over F. Let T be the linear operator on V defined by $T(A) = PA$ for all $A \in V$. Prove that $trace(T) = 2\, trace(P)$.

Problems

1. Find the linear transformation determined by the following matrices

   (a) $T : V_3(\mathbb{R}) \to V_3(\mathbb{R})$ given by $\begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & 3 \\ 1 & 1 & 4 \end{bmatrix}$ w.r.t the standard basis $\{e_1, e_2, e_3\}$.

   Solution. From the given matrix, we get

$$\begin{aligned} T(e_1) &= e_1 + 2e_2 + e_3 = (1, 2, 1); \\ T(e_2) &= 0e_1 + e_2 + e_3 = (0, 1, 1); \\ T(e_3) &= -e_1 + 3e_2 + 4e_3 = (-1, 3, 4). \end{aligned}$$

   Now any vector (a, b, c) in $V_3(\mathbb{R})$ can be written as

$$\begin{aligned} (a, b, c) &= a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) \\ &= ae_1 + be_2 + ce_3. \end{aligned}$$

   Therefore

$$\begin{aligned} T(a, b, c) &= T(ae_1 + be_2 + ce_3) \\ &= aT(e_1) + bT(e_2) + cT(e_3) \\ &= a(1, 2, 1) + b(0, 1, 1) + c(-1, 3, 4). \end{aligned}$$

   Thus $T(a, b, c) = (a - c, 2a + b + 3c, a + b + 4c)$ is the required linear transformation.

(b) $T : V_2(\mathbb{R}) \to V_2(\mathbb{R})$ given by $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ w.r.t the standard basis.

(c) $T : V_3(\mathbb{R}) \to V_3(\mathbb{R})$ given by $\begin{bmatrix} a & b & c \\ b & c & a \\ c & a & b \end{bmatrix}$ w.r.t the standard basis.

(d) $T : V_2(\mathbb{R}) \to V_3(\mathbb{R})$ given by $\begin{bmatrix} 2 & 1 \\ 1 & 1 \\ -1 & -1 \end{bmatrix}$ w.r.t the standard basis.

Solution. From the given matrix, we get

$$\begin{aligned} T((1,0)) &= 2(1,0,0) + (0,1,0) + (0,0,1) = (2,1,-1); \\ T((0,1)) &= 1(1,0,0) + 1(0,1,0) - 1(0,0,1) = (1,1,-1). \end{aligned}$$

Now any vector (a, b) in $V_2(\mathbb{R})$ can be written as $(a, b) = a(1, 0)+b(0, 1)$. Therefore

$$\begin{aligned} T(a,b) &= aT(1,0) + bT(0,1); \\ &= a(2,1,-1) + b(1,1,-1). \end{aligned}$$

Thus $T(a, b) = (2a + b, a + b, -a - b)$ is the required linear transformation.

2. Obtain the matrices for the following linear transformations

(a) $T : V_2(\mathbb{R}) \to V_2(\mathbb{R})$ given by $T(a, b) = (-b, a)$ w.r.t.

a. standard basis.

b. the basis $\{(1, 2), (1, -1)\}$ for both domain and range.

(b) $T : V_3(\mathbb{R}) \to V_2(\mathbb{R})$ given by $T(a, b, c) = (a + b, 2c - a)$ w.r.t.

a. standard basis.

b. the basis $\{(1, 0, -1), (1, 1, 1), (1, 0, 0)\}$ as a basis for $V_3(\mathbb{R})$ and $\{(0, 1), (1, 0)\}$ for $V_2(\mathbb{R})$.

(c) $T : V_3(\mathbb{R}) \to V_3(\mathbb{R})$ given by $T(a, b, c) = (3a+c, -2a+b, a+2b+4c)$ w.r.t.

a. standard basis.

b. the basis $\{(1, 0, 1), (-1, 2, 1), (2, 1, 1)\}$ as a basis for domain and range

3. Let V be the set of all polynomials of degree $\leq n$ in $\mathbb{R}[x]$. Obtain the matrix of $T : V \to V$ denoted by $T(f) = \frac{df}{dx}$ w.r.t the basis $\{1, x, x^2, \cdots, x^n\}$.

1.8 Eigen Values

Definition 1.8.1. *Let T be a linear operator on an n-dimensional vector space V over the field F. Then a scalar $\lambda \in F$ is called a characteristic value of T if there is a non-zero vector $v \in V$ such that $Tv = \lambda v$.*

Also if λ is a characteristic value of T, then any non-zero vector $v \in V$ such that $Tv = \lambda v$ is called a characteristic vector of T belonging to the characteristic value λ.

Characteristic values are sometimes also called proper values, eigen values or spectral values. Similarly, characteristic vectors are sometimes also called proper vectors, eigen vectors or spectral vectors. The set of all characteristic values of T is called the spectrum of T.

Theorem 1.8.2. *If v is an eigen vector of T corresponding to the eigen value λ, then αv is also an eigen vector of T corresponding to the same eigen value λ. Here α is non-zero scalar.*

Proof. Since v is an eigen value of T corresponding to the eigen value λ, obviously it follows that $v \neq \mathbf{0}$ and $Tv = \lambda v$. Let α be any scalar, then clearly $\alpha v \neq \mathbf{0}$. Also

$$T(\alpha v) = \alpha T(v) = \alpha(\lambda v) = (\alpha\lambda)v = (\lambda\alpha)v = \lambda(\alpha v).$$

Thus αv is an eigen vector of T corresponding to the eigen value λ. $\square$

Theorem 1.8.3. *If v is a characteristic vector of T, then v cannot correspond to more than one characteristic value of T.*

Proof. Let v be a characteristic vector of T corresponding to two distinct characteristic value λ_1 and λ_2 of T. Then we have

$$Tv = \lambda_1 v \text{ and } Tv = \lambda_2 v.$$

Thus we have $\lambda_1 v = \lambda_2 v$. This implies, $\lambda_1 v - \lambda_2 v = \mathbf{0}$ and hence $(\lambda_1 - \lambda_2)v = \mathbf{0}$. Therefore $\lambda_1 = \lambda_2$. $\square$

Theorem 1.8.4. *Let T be a linear operator on a finite dimensional vector space V and let λ be a characteristic value of T. Then the set*

$$W_\lambda = \{v \in V/Tv = \lambda v\}$$

is a subspace of V.

Proof. Let $u, v \in W_\lambda$. Then $Tu = \lambda u$ and $Tv = \lambda v$. If $\alpha, \beta \in F$, then

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v) = \alpha(\lambda u) + \beta(\lambda v) = \lambda(\alpha u + \beta v).$$

Therefore $(\alpha u + \beta v) \in W_\lambda$. Thus W_λ is a subspace of V. □

Note: The set W_λ is nothing but the set of all eigen vectors of T corresponding to the eigen value λ provided; we included the zero vector in this set. In other words w_λ is the null space of the linear operator $T - \lambda I$, where I is the identity operator. The subspace W_λ of V is called as the characteristic space (eigen space) of the characteristic value (eigen value) λ of the linear operator T. It is also called as the space of characteristic vectors of T associated with the characteristic value λ.

Theorem 1.8.5. *Distinct eigen vectors of T corresponding to distinct eigen values of T are linearly independent.*

Proof. Let $\lambda_1, \lambda_2, \cdots, \lambda_m$ distinct characteristic values of T and let $v_1, v_2, \cdots, v_m$ be the characteristic vectors of T corresponding to these characteristic values respectively. Then we have

$$Tv_i = \lambda_i v_i, 1 \leq i \leq m.$$

Let $S = \{v_1, v_2, \cdots, v_m\}$, then we have to prove that the set S is linearly independent. We shall prove the theorem by induction on m, the number of vectors in S.

If $m = 1$, then S is linearly independent as S contains only one non-zero vector. (Note that a characteristic vector cannot be $\mathbf{0}$ by its definition.)

Now suppose that the set

$$S_1 = \{v_1, v_2, \cdots, v_k\},$$

is linearly independent.

Consider the set $S_2 = \{v_1, v_2, \cdots, v_k, v_{k+1}\}$. We have to show that S_2 is linearly independent. Let $\alpha_1, \alpha_2, \cdots, \alpha_{k+1} \in F$ and let

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{k+1} v_{k+1} = \mathbf{0}. \tag{1.8}$$

Then

$$\begin{aligned}
&T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{k+1} v_{k+1}) = T(\mathbf{0}) \\
&\qquad \Rightarrow \quad \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_{k+1} T(v_{k+1}) = \mathbf{0} \\
&\qquad \Rightarrow \quad \alpha_1(\lambda_1 v_1) + \alpha_2(\lambda_2 v_2) + \cdots + \alpha_{k+1}(\lambda_{k+1} v_{k+1}) = \mathbf{0} \qquad (1.9)
\end{aligned}$$

Multiplying (1.8) by the scalar λ_{k+1} and subtracting it from (1.9), we get

$$\alpha_1(\lambda_1 - \lambda_{k+1})v_1 + \cdots + \alpha_k(\lambda_k - \lambda_{k+1})v_k = \mathbf{0}.$$

But since $v_1, v_2, \cdots, v_k$ are linearly independent vectors by our assumption and $\lambda_1, \lambda_2, \cdots, \lambda_{k+1}$ are all distinct, we have

$$\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0.$$

Thus from (1.8) it follows that

$$\alpha_{k+1} v_{k+1} = \mathbf{0},$$

which implies that $\alpha_{k+1} = 0$, as $v_{k+1} \neq \mathbf{0}$, as desired. Hence S_2 is linearly independent and this completes the proof by induction. □

Corollary 1.8.6. *If T is a linear operator on an n-dimensional vector space V, then T cannot have more than n distinct eigen values.*

Similarity of matrices:

Let A and B be square matrices of order n over the field F. Then B is said to be similar to A if there exists an $n \times n$ invertible square matrix C with elements in F such that

$$B = C^{-1}AC.$$

Results

1. The relation of similarity is an equivalence relation in the set of all $n \times n$ matrices over the field F.

2. Similar matrices have the same determinant.

Similarity of Linear Transformations

Definition 1.8.7. *Let A and B be linear transformations on a vector space $V(F)$. Then B is said to be similar to A if there exists an invertible linear transformation C on V such that*

$$B = CAC^{-1}.$$

Results

1. The relation of similarity is an equivalence relation in the set of all linear transformations on a vector space $V(F)$.

2. Let T be a linear operator on an n-dimensional vector space $V(F)$ and let B and B' be two ordered bases for V. Then the matrix of T relative to B' is similar to the matrix of T relative to B.

3. Suppose B and B' are two ordered bases for an n-dimensional vector space $V(F)$. Let T be a linear operator on V. Suppose A is a matrix of T relative to B and C is the matrix of T relative to B'. If P is the transition matrix from the basis B to the basis B', then $C = P^{-1}AP$.

Characteristic values of a matrix:

Let $A = [a_{ij}]_{n\times n}$ be a square matrix of order n over the field F. An element λ in F is called a characteristic value of A if

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

where I is the unit matrix of order n.

Note: Now suppose T is a linear operator on an n-dimensional vector space V and A is the matrix of T with respect to any ordered basis B. Then λ is characteristic value of T if and only if λ is a characteristic value of the matrix A. Therefore the definition of characteristic values of the matrix is sensible.

Characteristic vector of a matrix:

If λ is a characteristic value of an $n \times n$ matrix A then an $n \times 1$ non-zero matrix X, is said to be a characteristic vector of A corresponding to the characteristic value λ if $AX = \lambda X$.

Characteristic equation of a matrix:

Let A be a square matrix of order n over the field F. Consider the matrix $A - \lambda I$. The elements of this matrix are polynomials in λ of degree at most 1. If we evaluate $det(A - \lambda I)$, then it will be a polynomial in λ of degree n. Let us denote this polynomial as $f(\lambda)$.

Then $f(\lambda) = det(A - \lambda I)$ is called the characteristic polynomial of the matrix A. The equation $f(\lambda) = 0$ is the characteristic equation of the matrix A. Now λ is the characteristic value of the matrix A if and only if $det(A - \lambda I) = 0$. That is, $f(\lambda) = 0$ if and only if λ is the root of the characteristic equation of A. Thus in order to find the characteristic values of a matrix, we should first obtain its characteristic equation and then we should find the roots of this equation.

Theorem 1.8.8. *Let T be a linear operator on an n-dimensional vector space V and A be the matrix of T relative to any ordered basis B. Then a vector v in V is an eigen vector of T corresponding to its eigen value λ if and only if its co-ordinate vector X relative to the basis B is an eigen vector of A corresponding to its eigen value λ.*

Proof. We have $[T - \lambda I]_B = [T]_B - \lambda[I]_B = A - \lambda I$. If $v \neq 0$, then the co-ordinate vector X of v is also non-zero. Now

$$[(T - \lambda I)(v)]_B = [T - \lambda I]_B[v]_B = (A - \lambda I)X.$$

Therefore

$$(T - \lambda I)(v) = \mathbf{0} \text{ if and only if } (A - \lambda I)X = O.$$

That is,

$$T(v) = \lambda v \text{ if and only if } AX = \lambda X.$$

That is, v is an eigen vector of T if and only if X is eigen vector of A. $\square$

Theorem 1.8.9. *Similar matrices A and B have the same characteristic polynomial and hence the same eigen values. If X is an eigen vector of A corresponding to the eigen value λ, then $P^{-1}X$ is an eigen vector of B corresponding to the eigen value λ where $B = P^{-1}AP$.*

Proof. Suppose A and B are similar matrices. Then there exists an invertible matrix P such that

$$B = P^{-1}AP.$$

This implies

$$B - \lambda I = P^{-1}AP - \lambda I = P^{-1}AP - P^{-1}(\lambda I)P = P^{-1}(A - \lambda I)P.$$

(Since $P^{-1}(\lambda I)P = \lambda(P^{-1}IP) = \lambda I$). Therefore

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

Thus the matrices A and B have the same characteristic polynomial and consequently they will have the same characteristic values.

Moreover, if λ is an eigen value of A and X is a corresponding eigen vector, then we have $AX = \lambda X$. Now consider,

$$B(P^{-1}X) = (P^{-1}AP)(P^{-1}X) = P^{-1}AX = P^{-1}\lambda X = \lambda(P^{-1}X).$$

Thus $P^{-1}X$ is an eigen vector of B corresponding to λ. This completes the proof. □

Now suppose that T is a linear operator on an n-dimensional vector space V. If B_1 and B_2 are any two ordered bases for V, then we know that the matrices $[T]_{B_1}$ and $[T]_{B_2}$ are similar. Also similar matrices have the same characteristic polynomial. The enables us to define sensibly the characteristic polynomial of T as follows:

Characteristic Polynomial of a Linear Operator

Definition 1.8.10. *Let T be a linear operator on an n-dimensional vector space V. The characteristic polynomial of any $n \times n$ matrix which represents T in some ordered basis for V. On account of the above discussion the characteristic polynomial of T as defined by us will be unique.*

If B is any ordered basis for V and A is the matrix of T with respect to B, then

$$\begin{aligned} \det(T - \lambda I) &= \det[T - \lambda I]_B \\ &= \det([T]_B - \lambda [I]_B) \\ &= \det(A - \lambda I). \end{aligned}$$

That is, the characteristic polynomial of A and so also that of T. Therefore the characteristic polynomial of $T = \det(T - \lambda I)$. The equation $\det(T - \lambda I) = 0$ is called the characteristic equation of T.

Theorem 1.8.11. *Let T be a linear operator on a finite dimensional vector space V. Then the following are equivalent*

1. *λ is an eigen value of T;*

2. *The operator $T - \lambda I$ is singular (not invertible);*

3. *$det(T - \lambda I) = 0$.*

Example 1.8.12. *Consider the linear operator T on $V_2(\mathbb{R})$ which is represented in the standard ordered basis by the matrix*

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

Then the characteristic polynomial for T (or for A) is

$$\det(A - \lambda I) = \det \begin{bmatrix} 0-\lambda & -1 \\ 1 & 0-\lambda \end{bmatrix} = \begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 + 1.$$

The polynomial equation $\lambda^2 + 1 = 0$ has no roots in $\mathbb{R}$. Therefore T has no characteristic values.

However if we consider T as a linear operator on $V_2(\mathbb{C})$, then the characteristic equation of T has two distinct roots i and $-i$ in $\mathbb{C}$. In this case T has two characteristic values i and $-i$.

Algebraic Multiplicity and Geometric Multiplicity of a characteristic value:

Let T be a linear operator on an n-dimensional vector space V and let λ be a characteristic value of T. By geometric multiplicity of λ we mean the dimension of the characteristic space W_λ of λ. By algebraic multiplicity of λ we mean the multiplicity of λ as root of the characteristic equation of T.

Method of finding the characteristic values and the corresponding the characteristic vectors of a linear operator T:

Let T be a linear operator on an n-dimensional vector space V over the field F. Let B be any ordered basis for V and let $A = [T]_B$. The roots of the equation $\det(A - \lambda I) = 0$ will give the characteristic values of A or also of T. Let λ be a characteristic value of T. Then $\mathbf{0} \neq v$ will be a characteristic vector corresponding to the characteristic value if

$$(T - \lambda I)v = \mathbf{0};$$

that is, if

$$[T - \lambda I]_B [v]_B = [\mathbf{0}]_B;$$

that is, if

$$(A - \lambda I)X = O,$$

where $X = [v]_B$ is a column matrix of the type $n \times 1$ and O is the null matrix of the type $n \times 1$. Thus to find the coordinate matrix of v with respect to B, we should solve the matrix equation $(A - \lambda I)X = O$ for X.

Matrix Polynomials:

An expression of the form

$$f(\lambda) = A_0 + A_1\lambda + A_2\lambda^2 + \cdots + A_m\lambda^m,$$

where $A_0, A_1, A_2, \cdots, A_m$ are all square matrices of order m, is called a matrix polynomial of degree m provided A_m is not a null matrix.

Theorem 1.8.13. *Let T be a linear operator on an n-dimensional vector space $V(F)$. Then T satisfies its characteristic equation (i.e.) $f(\lambda)$ be the characteristic polynomial of T, then $f(T) = 0$.*

or

Every square matrix satisfies its characteristic equation.

Proof. Let T be a linear operator on an n-dimensional vector space V over the field F. Let B be any ordered basis for V and A be the matrix of T relative to B. (i.e.)$A = [T]_B$. The characteristic polynomial of T is the same as the characteristic polynomial of A. If $A = [a_{ij}]_{n\times n}$, then the characteristic polynomial $f(\lambda)$ of A is given by

$$\begin{aligned} f(\lambda) &= det(A-\lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} \\ &= a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n (say) \end{aligned}$$

where the a_i are in F. Then the characteristic equation of A is $f(\lambda) = 0$. That is,

$$a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n = 0.$$

Since the elements of the matrix $A - \lambda I$ are polynomials at most of the first degree in λ, therefore the elements of the matrix $adj(A-\lambda I)$ are ordinary polynomials in λ of degree $n-1$ or less. Note that the elements of the matrix $adj(A - \lambda I)$ are the cofactors of the elements $A - \lambda I$. Therefore $adj(A - \lambda I)$ can be written as a matrix polynomial in λ in the form

$$adj(A - \lambda I) = B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_{n-1}\lambda^{n-1},$$

where the B_i's are square matrices of order n over F with elements independent of λ. Now by the property of adjoints, we know that

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

Therefore,

$$(A-\lambda I)(B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_{n-1}\lambda^{n-1}) = (a_0+a_1\lambda+a_2\lambda^2+\cdots+a_n\lambda^n)I.$$

Equating the coefficients of like powers of λ on both sides we get

$$\begin{aligned}
AB_0 &= a_0 I \\
AB_1 - IB_0 &= a_1 I \\
AB_2 - IB_1 &= a_2 I \\
\vdots \quad &\vdots \quad \vdots \\
AB_{n-1} - IB_{n-2} &= a_{n-1} I \\
IB_{n-1} &= a_n I
\end{aligned}$$

Premultiplying these equations successively by $I, A, A^2, \ldots, A^n$ and adding, we get

$$a_0 + a_1 A + a_2 A^2 + \cdots + a_n A^n = O,$$

where O is the null matrix of order n. Thus $f(A) = O$.

Now since

$$f(T) = a_0 + a_1 T + a_2 T^2 + \cdots + a_n T^n,$$

we have

$$\begin{aligned}[f(T)]_B &= [a_0 + a_1T + a_2T^2 + \cdots + a_nT^n]_B \\ &= a_0[I]_B + a_1[T]_B + a_2[T^2]_B + \cdots + a_n[T^n]_B \\ &= f(A).\end{aligned}$$

But as $f(A) = O$, we have

$$[f(T)]_B = O = [\hat{\mathbf{0}}]_B.$$

Therefore

$$f(T) = \mathbf{0}.$$

This implies that, $a_0 + a_1T + a_2T^2 + \cdots + a_nT^n = \hat{\mathbf{0}}$. $\square$

Note:

Since $f(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n = det(A - \lambda I)$, we have $f(0) = a_0 = detA = detT$. Moreover, if T is non-singular, then T is invertible and $detT \neq 0$, that is, $a_0 \neq 0$. Thus

$$a_0 = -(a_1T + a_2T^2 + \cdots a_nT^n)$$

and hence

$$I = -\left(\frac{a_1}{a_0}I + \frac{a_2}{a_0}T^2 + \cdots \frac{a_n}{a_0}T^{n-1}\right).$$

Therefore

$$T^{-1} = -\left(\frac{a_1}{a_0}T + \frac{a_2}{a_0}T^2 + \cdots \frac{a_n}{a_0}T^{n-1}\right).$$

Problem 1.8.14. *Let V be an n-dimensional vector space over F. What is the characteristic polynomial of (i) the identity operator on V (ii) the zero operator on V.*

Solution:

Let$B$ be any ordered basis for V.

Proof(i): If I is the identity operator on V, then $[I]_B = I$. Thus the characteristic polynomial of

$$I = \det(I - \lambda I) = \begin{vmatrix} 1-\lambda & 0 & \cdots & 0 \\ 0 & 1-\lambda & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1-\lambda \end{vmatrix} = (1-\lambda)^n.$$

proof(ii): If $\hat{\mathbf{0}}$ is the zero operator on V, the $[\hat{\mathbf{0}}]_B = O$, that is the null matrix of order n. Therefore the characteristic polynomial of

$$\hat{\mathbf{0}} = \det(O - \lambda I) = \begin{vmatrix} -\lambda & 0 & \cdots & 0 \\ 0 & -\lambda & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & -\lambda \end{vmatrix} = (-1)^n(\lambda)^n.$$

■

Problem 1.8.15. *Find all the eigen values and eigen vectors of the following matrices*

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

2. $\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$

3. $\begin{bmatrix} 1 & 1 \\ 0 & i \end{bmatrix}$

Solution: Let

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

Then we have

$$A - \lambda I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -\lambda & 1 \\ 0 & -\lambda \end{bmatrix}.$$

Therefore the characteristic equation of A is $\det(A - \lambda I) = 0$, that is $\lambda^2 = 0$. Thus 0 is the only characteristic value of A.

Now let x_1, x_2 be the components of a characteristic vector v corresponding to this characteristic value. Let X be the coordinate matrix of v. Then

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Now X will be given a non-zero solution of equation $(A - 0I)X = O$; that is,

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

That is,

$$\begin{bmatrix} x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thus $x_2 = 0$ and $x_1 = k$, where k is any non-zero complex number. Therefore $X = \begin{bmatrix} k \\ 0 \end{bmatrix}$, where k is any non-zero complex number is the required eigen vector corresponding to the eigen value 0.

Problem 1.8.16. *Obtain the eigen values, eigen vectors and the corresponding eigen spaces of the matrix*

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

Solution:

The characteristic polynomial of A is

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (\lambda^2 - 1)(\lambda - 1).$$

Therefore the eigen values of A are $1, -1$.

The eigen space (the set of eigen vectors) corresponding to the eigen value 1 is nothing but the solution space of the linear system

$$(A - \lambda I)(X) = (A - I)(X) = O,$$

that is

$$(A - I)(X) = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Let W_1 be the solution space of that linear system, then

$$W_1 = \{(k, k, l)/k, l \in \mathbb{R}\}.$$

It is clear that any element in W_1 can be expressed as a linear combination of the set vectors $\{(1, 1, 0), (0, 0, 1)\}$, which is linearly independent. Hence forms a basis of W_1 and therefore $\dim(W_1) = 2$.

The eigen space (the set of eigen vectors) corresponding to the eigen value 1 is nothing but the solution space of the linear system

$$(A - \lambda I)(X) = (A - I)(X) = O,$$

that is

$$(A - I)(X) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Let W_2 be the solution space of that linear system, then

$$W_2 = \{(k, -k, 0)/k \in \mathbb{R}\}.$$

It is clear that any element in W_2 can be expressed as a linear combination of the set vectors $\{(1, -1, 0)\}$, which is linearly independent. Hence forms a basis of W_2 and therefore $\dim(W_2) = 1$.

Algebraic multiplicity of 1 is 2 and algebraic multiplicity of -1 is 1. Further, geometric multiplicity of 1 is 2 and geometric multiplicity of -1 is 1.

Problems

1. Determine the dimension of the indicated subspace of $\mathbb{R}^4$.

   - The set of all vectors of the form $(a, b, c, 0)$.
   - The set of all vectors of the form (a, a, a, a).

2. Determine the dimension of the subspace of M_{31} consisting of vectors of the form
$$\begin{bmatrix} a-2b \\ a+b \\ 3b \end{bmatrix}.$$

3. Determine the matrix representation of the given linear transformation relative to the standard basis.

   - $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2x-3y+4z \\ 5x-y+2z \\ 4x+7y \end{bmatrix}$
   - $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2y+z \\ x-4y \\ 3x \end{bmatrix}$
   - $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 3x-y \\ 2y \end{bmatrix}$
   - $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+4y \\ 3x-2y \end{bmatrix}.$

4. Find the coordinate vector for v relative to the given basis $B = \{u_1, u_2\}$ of $\mathbb{R}^2$.

   - $\begin{bmatrix} 8 \\ -4 \end{bmatrix}; B = \left\{\begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}\right\}.$

- $\begin{bmatrix} 20 \\ 14 \end{bmatrix}$; $B = \left\{ \begin{bmatrix} -1 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$.

5. Find the eigen values and bases for the eigen spaces of the following matrices:

   - $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix}$
   - $\begin{bmatrix} -2 & 6 & -6 \\ 0 & 3 & -5 \\ 0 & -3 & 1 \end{bmatrix}$
   - $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$
   - $\begin{bmatrix} -2 & -6 & 19 \\ 0 & -2 & 5 \\ 1 & 0 & -4 \end{bmatrix}$
   - $\begin{bmatrix} 3 & 2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & -1 \end{bmatrix}$
   - $\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -3 & 5 & 0 \\ 6 & \frac{4}{9} & -8 \end{bmatrix}$.

6. Let T be a linear operator on a finite dimensional vector space V. Then show that $\mathbf{0}$ is a characteristic value of T if and only if T is not invertible.

7. Two linear transformations A and B on $V(F)$ are similar, then show that A^2 and B^2 are similar and if A, B are invertible, then A^{-1} and B^{-1} are similar.

8. If A and B are linear transformations on $V(F)$ and if one of them is invertible then AB and BA are similar.

9. Prove that similar matrices have same trace.

10. Prove the following:

    - If λ is a characteristic value of an invertible transformation T, then show that λ^{-1} is a characteristic value of T^{-1}.

    - If λ is an eigenvalue of A with corresponding eigenvector X, then λ^k is an eigenvalue of A^k with corresponding eigenvector X for k a positive integer.

    - A and A^T have the same eigenvalues.

    - If A is a triangular matrix, the eigenvalues of A are the same as the elements on its diagonal.

    - If A is a diagonal matrix, the eigenvalues of A are the same as the elements on its diagonal.

11. Prove that $\det(A)$ is equal to the product of diagonal entries.

12. Prove that $trace(A)$ is equal to the product of diagonal entries.

Problem 1.8.17. *Let T be a linear operator on a finite dimensional vector space V and let λ be a characteristic value of T. Show that the characteristic space of λ, that is W_λ is invariant under T. (If T is a linear operator on V*

and if W is a subspace of V, we say that W is invariant under T if $v \in W$ implies $T(v) \in W$).

Solution: We have by the definition of W_λ,

$$W_\lambda = \{v \in V : Tv = \lambda v\}.$$

Let $v_0 \in W_\lambda$, then we have $Tv_0 = \lambda v_0$. Since W_λ is a subspace, for any scalar $\alpha \in F$ and $v \in W_\lambda$ it follows that $\alpha v \in W_\lambda$. Thus in particular, $\lambda v_0 \in W_\lambda$, which implies $Tv_0 \in W_\lambda$. Since v_0 is arbitrary, it follows that W_λ is invariant under T. ■

Problem 1.8.18. *Suppose S and T are two linear operators on a finite dimensional vector space V. If S and T have the same characteristic polynomial, then* $\det S = \det T$.

Solution: Let $\dim V = n$. Let B be any ordered basis for V. If A is the matrix of S relative to B, then the characteristic polynomial $f(\lambda)$ of S is given by $f(\lambda) = \det(A - \lambda I)$. If we let

$$f(\lambda) = a_0 + a_1 x + \cdots + a_n x^n = \det(A - \lambda I),$$

then the constant term in this polynomial is

$$a_0 = f(0) = \det A.$$

Similarly, the constant term in the characteristic polynomial of $T = \det C$, where C is the matrix of T relative to B.

Since S and T have the same characteristic polynomial, therefore the two constant terms must be equal. Thus $\det A = \det C$, which implies $\det[S]_B = \det[T]_B$ and hence $\det S = \det T$. ■

Problems

1. If T be a linear operator on a finite dimensional vector space V and λ be a characteristic value of T, then show that the characteristic space of λ, (i.e., W_λ) is the null space of the operator $T - \lambda I$. Hint: to prove, $W_\lambda = \{v \in V/(T - \lambda I)v = \mathbf{0}\}$.

2. Show that the characteristic values of a diagonal matrix are precisely the elements in the diagonal. Hence show that if a matrix B is similar to a diagonal matrix D, then the diagonal elements of D are the characteristic values of B.

3. Let T be a linear operator on a finite dimensional vector space V. Then show that $\mathbf{0}$ is a characteristic value of T if and only if T is not invertible.

4. If λ is a characteristic value of an invertible transformation T, then show that λ^{-1} is a characteristic value of T^{-1}.

5. If $\lambda \in F$ is a characteristic value of a linear operator T on a vector space $V(F)$, then for any polynomial $p(x)$ over F, $p(\lambda)$ is a characteristic value of $p(T)$.

Hint: First prove that $T^k v = \lambda^k v$, for positive integer k.

6. If $\lambda \in F$ is a characteristic value of a square matrix A of order n, then for any polynomial $p(x)$ over F, $p(\lambda)$ is a characteristic value of $p(A)$.

7. Prove that similar matrices have same characteristic polynomial

Diagonalizable Matrix:

Definition 1.8.19. *A matrix A over a field F is said to be diagonalizable if it is similar to a diagonal matrix over the field F. Thus a matrix A is diagonalizable if there exists an invertible matrix P such that $P^{-1}AP = D$, where D is a diagonal matrix. Also the matrix P is then said to diagonalize A or transform A to diagonal form.*

Theorem 1.8.20. *If A be an $n \times n$ matrix over a field F. Then A is diagonalizable if and only if A has n linearly independent eigen vectors in $V_n(F)$.*

Proof. Suppose A is diagonalizable, then there exist an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$ or equivalently,

$$AP = PD.$$

Let $X_1, X_2, \cdots, X_n$ be column vectors of P and $\lambda_1, \lambda_2, \cdots \lambda_n$ be diagonal entries of D. Now

$$AP = A \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} = \begin{bmatrix} AX_1 & AX_2 & \cdots & AX_n \end{bmatrix}$$

and

$$PD = \begin{bmatrix} \lambda_1 X_1 & \lambda_2 X_2 & \cdots & \lambda_n X_n \end{bmatrix}$$

Thus, we have $AX_1 = \lambda_1 X_1$, $AX_2 = \lambda_2 X_2$, $\cdots$, $AX_n = \lambda_n X_n$. Since P is invertible, its column vectors $X_1, X_2, \cdots, X_n$ are linearly independent (and hence nonzero). Therefore, these n column vectors are eigen vectors of A.

Suppose A has n linearly independent eigen vectors, $X_1, X_2, \cdots, X_n$ and $\lambda_1, \lambda_2, \cdots \lambda_n$ are the corresponding eigen values.

If we let $P = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}$ and D to be the diagonal matrix that has $\lambda_1, \lambda_2, \cdots \lambda_n$ as its successive diagonal entries, then

$$\begin{aligned} AP &= A \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \\ &= \begin{bmatrix} AX_1 & AX_2 & \cdots & AX_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 X_1 & \lambda_2 X_2 & \cdots & \lambda_n X_n \end{bmatrix} \\ &= PD. \end{aligned}$$

Since the column vectors of P are linearly independent, P is invertible. Thus we have $P^{-1}AP = D$ and hence A is diagonalizable. □

Theorem 1.8.21. *If A be an $n \times n$ matrix over a field F. Then A is diagonalizable if and only if all algebraic multiplicity and geometric multiplicity of all eigen values A are equal.*

Definition 1.8.22. *A linear operator T on an n-dimensional vector space $V(F)$ is diagonalizable if its matrix A relative to any ordered basis B of V is* *diagonalizable.*

If A is diagonalizable, the process of finding for A is as follows:

1. Find the eigenvalues for A.

2. Determine the corresponding linearly independent eigenvectors.

3. Construct P as a matrix whose columns are the eigenvectors.

4. Calculate P^{-1}.

5. Calculate $D = P^{-1}AP$.

Problem 1.8.23. *Let T be the linear operator on $\mathbb{R}^3$ which is represented in the standard ordered basis by the matrix*

$$\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}.$$

Prove that T is diagonalizable.

Solution. Let

$$A = \begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}.$$

The characteristic equation of A is given by

$$\begin{vmatrix} -9-x & 4 & 4 \\ -8 & 3-x & 4 \\ -16 & 8 & 7-x \end{vmatrix} = 0$$

which is equivalent to

$$\begin{vmatrix} -1-x & 4 & 4 \\ -1-x & 3-x & 4 \\ -1-x & 8 & 7-x \end{vmatrix} = 0, \quad C_1 \to C_1 + C_2 + C_3$$

that is equivalent to

$$-(1+x)\begin{vmatrix} 1 & 4 & 4 \\ 1 & 3-x & 4 \\ 1 & 8 & 7-x \end{vmatrix} = 0$$

which is equivalent to

$$(1+x)\begin{vmatrix} 1 & 4 & 4 \\ 0 & -1-x & 0 \\ 0 & 4 & 3-x \end{vmatrix} = 0, \quad R_2 \to R_2 - R_1, \quad R_3 \to R_3 - R_1.$$

Thus it follows that

$$(1+x)(1+x)(3-x) = 0.$$

Therefore the roots of this equation are -1, -1, 3.

The characteristic vectors X of A corresponding to the eigen value -1 are given by the equation $(A-(-1)I)X = O$. That is,

$$(A+I)X = \begin{bmatrix} -8 & 4 & 4 \\ -8 & 4 & 4 \\ -16 & 8 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

These equations are equivalent to the equations

$$\begin{bmatrix} -8 & 4 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \; R_2 \to R_2 - R_1 \; R_3 \to R_3 - 2R_1.$$

The matrix of coefficient of these equations has rank 1. Therefore these equations will have non-zero solutions. In particular, these equations will have $3-1=2$ linearly independent set of solutions (G. M of -1). We see that these equations reduce to the single equation

$$-2x_1+x_2+x_3=0.$$

If we let $x_3=k$ and $x_2=l$, then $x_1=\frac{k+l}{2}$. Therefore the eigen space corresponding to the eigen value -1 is given by

$$\left\{\begin{bmatrix}\frac{k+l}{2}\\ l\\ k\end{bmatrix}/k,l\in\mathbb{R}\right\}.$$

Obviously,

$$X_1=\begin{bmatrix}1\\1\\1\end{bmatrix},\ X_2=\begin{bmatrix}0\\1\\-1\end{bmatrix}$$

are two linearly independent vectors of the eigen space. Therefore X_1 and X_2 are two linearly independent eigen vectors of A corresponding to the eigen value -1, which forms the basis of the respective eigen space.

Now the eigen vectors of A corresponding to the eigen value 3 are given by the equation $(A-3I)X=O$. That is,

$$\begin{bmatrix}-12&4&4\\-8&0&4\\-16&8&4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

These equations are equivalent to the equations

$$\begin{bmatrix}-12&4&4\\4&-4&0\\-4&4&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}\quad R_2\to R_2-R_1,\ R_3\to R_3-R_1$$

The matrix of coefficient of these equations has rank 2. Therefore these equations will have a non-zero solution. Also these equations will have

$3 - 2 = 1$ linearly independent set of solution. These equations can be written as

$$-12x_1 + 4x_2 + 4x_3 = 0 \tag{1.10}$$

$$4x_1 - 4x_2 = 0 \tag{1.11}$$

$$-4x_1 + 4x_2 = 0 \tag{1.12}$$

From these, we get $x_1 = x_2 = k$, say, Then $x_3 = 2$. Thus the eigen space corresponding to the eigen value 3 is

$$\left\{ \begin{bmatrix} k \\ k \\ 2k \end{bmatrix} \Big/ k \in \mathbb{R} \right\}.$$

Therefore,

$$X_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

is the linearly independent eigen vector of corresponding to the eigen value 3, which forms the basis of the respective eigen space.

Now if we let

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

then we have $detP = 1 \neq 0$. Therefore the matrix P is invertible. Therefore the columns of P are linearly independent vectors belonging to $\mathbb{R}^3$. Since the matrix A has three linearly independent eigen vectors in $\mathbb{R}^3$, Therefore it is diagonalizable. Also the diagonal form D of A is given by

$$P^{-1}AP = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} = D.$$

■

Problem 1.8.24. *Is the matrix*

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

similar over the field $\mathbb{R}$ *to a diagonal matrix? Is* A *similar over the field* $\mathbb{C}$ *to a diagonal matrix?*

Solution. The characteristic equation of A is

$$\begin{aligned} \begin{vmatrix} 1-x & 1 \\ -1 & 1-x \end{vmatrix} &= 0 \\ (1-x)^2 + 1 &= 0 \\ x^2 - 2x + 2 &= 0. \end{aligned}$$

The roots of this equation are $1+i$, $1-i$. Since the characteristic equation of A has no roots in $\mathbb{R}$, therefore the matrix A has no eigen value if we regard it as a matrix over $\mathbb{R}$. Consequently, A has no eigen vector in $\mathbb{R}^2$. Therefore the matrix A is not diagonalizable over the field $\mathbb{R}$.

If we regard A as a matrix over $\mathbb{C}$, then it has two eigen values. (*i.e.*) $1+i$, $1-i$.

Since A has two distinct eigen values, therefore it will have two linearly independent set of eigen vectors. Consequently, A is diagonalizable.

The eigen vectors of A corresponding to the eigen values $1+i$, $1-i$ are given by the system of equations

$$\begin{bmatrix} -i & 1 \\ -1 & i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

respectively.

From these we get

$$X_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}, \; X_2 \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

as the eigen vectors of A corresponding to the eigen values $1+i$, $1-i$ respectively.

If $P = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$, then $P^{-1}AP = \begin{bmatrix} 1+i & 0 \\ 0 & 1-i \end{bmatrix}$ gives the diagonal form of A.

Problem 1.8.25. *Prove that the matrix*

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

is not diagonalizable over the field $\mathbb{C}$.

Solution. The characteristic equation of A is

$$\begin{aligned} \begin{vmatrix} 1-x & 2 \\ 0 & 1-x \end{vmatrix} &= 0 \\ (1-x)^2 &= 0, \end{aligned}$$

The roots of this equation are $1, 1$. Therefore the only distinct eigen value of A is 1. The eigen vectors of A corresponding to that eigen value are given by

$$\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which implies $0x_1 + 2x_2 = 0$ and hence $x_2 = 0$.

If we let $x_1 = k$(say), then the eigen space corresponding to the eigen value 1 is given by

$$\left\{ \begin{bmatrix} k \\ 0 \end{bmatrix} / k \in \mathbb{C} \right\}.$$

Therefore the equation can have only one linearly independent solution and therefore A can have only one linearly independent set of eigen vector. Thus it is not diagonalizable. ■

Problems

1. Let T be a linear operator on the n-dimensional vector space V and suppose that T has n distinct characteristic values. Prove that T is diagonalizable.

2. If an $n \times n$ matrix A has n distinct eigen values, then A is diagonalizable.

3. Prove that the matrix
$$\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$$
is diagonalizable.

Minimal Polynomial of a linear operator

Definition 1.8.26. *Suppose T is a linear operator on a finite dimensional vector space over the field F and $f(x)$ is a polynomial over F. If $f(T) = \hat{\mathbf{0}}$, then we say that the polynomial $f(x)$ annihilates the linear operator T.*

Similarly, suppose A is square matrix of order n over the field F and $f(x)$ is a polynomial over F. If $f(A) = O$, then we say that the polynomial $f(x)$ annihilates the matrix A.

Note that the set of all polynomials that annihilates, a linear operator T on a finite dimensional vector space $V(F)$ is nonempty.

Definition 1.8.27. *A polynomial in x over a field F is called a monic polynomial if the coefficient of the highest power of x in it is unity.*

For example, $x^3 - 2x^2 + \frac{5}{7}x + 5$ is a monic polynomial of degree 3 over the field of rational numbers.

Definition 1.8.28. *Suppose T is a linear operator on an n-dimensional vector space $V(F)$. The monic polynomial of lowest degree over the field F that annihilates T is called the minimal polynomial of T. Also, if $f(x)$ is the minimal polynomial of T, the equation $f(x) = 0$ is called the minimal equation of the linear operator T.*

Similarly we can define the minimal polynomial of a matrix. Suppose A is a square matrix of order n over the field F. The monic polynomial of lowest degree over the field F that annihilates A is called the minimal polynomial of A.

Now suppose T is a linear operator on an n-dimensional vector space $V(F)$ and A is the matrix of T in some ordered basis B. If $f(x)$ is any polynomial over F, then $[f(T)]_B = f(A)$. Therefore $f(T) = \hat{\mathbf{0}}$ if and only if $f(A) = O$. Then $f(x)$ annihilates T iff it annihilates A. Therefore if $f(x)$ is the polynomial of lowest degree that annihilates T, then it is also the polynomial of lowest degree that annihilates A and conversely. Hence T and A have the same minimal polynomial. Further the characteristic polynomial of the matrix A is of degree n. Since the characteristic polynomial of A annihilates A, therefore the minimal polynomial of A cannot be of degree greater than n, its degree must be less than or equal to n.

Results

1. The minimal polynomial of a matrix or of a linear operator is unique.

2. The minimal polynomial of a matrix(linear operator) is a divisor of every polynomial that annihilates matrix(linear operator).

3. Let T be a linear operator on an n-dimensional vector space V [or, let A be an $n \times n$ matrix]. The characteristic and minimal polynomial for T [for A] have the same roots, except for multiplicities.

4. Let T be a diagonalizable linear operator and let $c_1, c_2, \cdots, c_n$ be the distinct characteristic values of T. Then the minimal polynomial for T is the polynomial

$$p(x) = (x - c_1)(x - c_2) \cdots (x - c_n).$$

We know that the monic polynomial x annihilates the zero operator $\hat{\mathbf{0}}$ and it is the polynomial of lowest degree that annihilates $\hat{\mathbf{0}}$. Hence x is the minimal polynomial for $\hat{\mathbf{0}}$.

Problem 1.8.29. *Let V be a finite-dimensional vector space. What is the minimal polynomial for the identity operator on V?*

Solution. We know that $I = 1.I$. (i.e.)$I - 1.I = \hat{\mathbf{0}}$. Thus the monic polynomial $x-1$ annihilates the identity operator I and it is the polynomial of lowest degree that annihilates I. Hence $x-1$ is the minimal polynomial for I.

Problem 1.8.30. *Let V be an n-dimensional vector space and let T be a linear operator on V. Suppose that there exists some positive integer k so that $T^k = \hat{\mathbf{0}}$. Prove that $T^n = \hat{\mathbf{0}}$.*

Solution. Since $T^k = \hat{\mathbf{0}}$, therefore the polynomial x^k annihilates T. So the minimal polynomial for T is a divisor of x^k. Let x^r be the minimal polynomial for T where $r \leq n$. Then $T^r = \hat{\mathbf{0}}$. Now

$$T^n = T^{n-r}T^r = T^{n-r}.\hat{\mathbf{0}} = \hat{\mathbf{0}}.$$

Problem 1.8.31. *Find the minimal polynomial for the real matrix*

$$A = \begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}$$

Solution.

$$
\begin{aligned}
|A - xI| &= \begin{vmatrix} 7-x & 4 & -1 \\ 4 & 7-x & -1 \\ -4 & -4 & 4-x \end{vmatrix} \\
&= \begin{vmatrix} 7-x & 4 & -1 \\ 4 & 7-x & -1 \\ 0 & 3-x & 3-x \end{vmatrix} \quad R_3 \to R_3 + R_2 \\
&= (3-x)\begin{vmatrix} 7-x & 4 & -1 \\ 4 & 7-x & -1 \\ 0 & 1 & 1 \end{vmatrix} \\
&= (3-x)\begin{vmatrix} 7-x & 4 & -1 \\ 4 & 7-x & -1 \\ 0 & 1 & 0 \end{vmatrix} \quad C_3 \to C_3 - C_2 \\
&= -(3-x)\begin{vmatrix} 7-x & -5 \\ 4 & x-8 \end{vmatrix} \quad \text{expanding along third row} \\
&= -(3-x)\begin{vmatrix} 3-x & 3-x \\ 4 & x-8 \end{vmatrix} \quad R_1 \to R_1 - R_2 \\
&= -(3-x)^2\begin{vmatrix} 1 & 1 \\ 4 & x-8 \end{vmatrix} \\
&= -(3-x)^2(x-12).
\end{aligned}
$$

Therefore the roots of the equation $|A - xI| = 0$ are $x = 3, 3, 12$. These are the characteristic roots of A.

Let us now find the minimal polynomial of A. We know that each characteristic root of A is also a root of its minimal polynomial. So if $m(x)$ is the minimal polynomial for A, then both $(x-3)$ and $(x-12)$ are factors of $m(x)$.

Let $h(x) = (x-3)(x-12) = x^2 - 15x + 36$, then we can easily verify that $A^2 - 15A + 36I = O$. Therefore $h(x)$ annihilates A. Thus $h(x)$ is the monic polynomial of lowest degree which annihilates A and hence $h(x)$ is the minimal polynomial for A.

Problem 1.8.32. *Show that similar matrices have the same minimal polynomial.*

Solution. Suppose A and B are two similar matrices. Then there exists a non-singular matrix P such that

$$B = P^{-1}AP.$$

Now we have to show that the matrices A and PAP^{-1} have the same monic polynomial. First we shall show that a monic polynomial $f(x)$ annihilates A if and only if it annihilates $P^{-1}AP$. We have

$$(P^{-1}AP)^2 = P^{-1}APP^{-1}AP = P^{-1}A^2P.$$

Proceeding in this way we can show that $(P^{-1}AP)^2 = P^{-1}A^2P$, where k is any positive integer.

Let $f(x) = x^r + a_1x^{r-1} + \cdots + a_{r-1}x + a_r$. Then

$$f(A) = A^r + a_1A^{r-1} + \cdots + a_{r-1}A + a_rI.$$

Also,

$$
\begin{aligned}
f(P^{-1}AP) &= (P^{-1}AP)^r + \cdots + a_{r-1}(P^{-1}AP) + a_r I \\
&= P^{-1}A^r P + \cdots + a_{r-1}(P^{-1}AP) + a_r I \\
&= P^{-1}(A^r + \cdots + a_{r-1}A + a_r I)P \\
&= P^{-1}f(A)P.
\end{aligned}
$$

Since P is non-singular, therefore

$$P^{-1}f(A)P = O \text{ if and only if } f(A) = O.$$

Thus $f(x)$ annihilates A if and only if it annihilates $P^{-1}AP$.

Therefore if $f(x)$ is the polynomial of lowest degree that annihilates A, then it is also the polynomial of lowest degree that annihilates $P^{-1}AP$ and conversely. Hence A and $P^{-1}AP$ have the same minimal polynomial.

Problems

1. Find the characteristic roots of the matrix

$$\begin{bmatrix} 5 & 6 & 8 \\ 0 & 7 & 2 \\ 0 & 0 & 4 \end{bmatrix}$$

2. Write the characteristic polynomial and the minimal polynomial of the matrix

$$\begin{bmatrix} 4 & 3 & 0 \\ 2 & 1 & 0 \\ 5 & 7 & 9 \end{bmatrix}$$

3. State whether the following statement is true or false: The matrices

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

and

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

have the same characteristic roots.

4. Show that the characteristic equation of the complex matrix

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

is $x^3 - ax^2 - bx - c = 0$.

5. Show that the minimal polynomial of the real matrix

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

is $x^2 + 1 = 0$

6. Show that the minimal polynomial of the real matrix

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

is $(x - 1)(x - 2) = 0$.

7. What is the minimal polynomial for the zero operator?

8. Find all(complex) eigen values and eigen vectors of the following matrices

   a. $\begin{bmatrix} 2 & 4 \\ 3 & 13 \end{bmatrix}$

   b. $\begin{bmatrix} 2 & -3 \\ -3 & 1 \end{bmatrix}$

   c. $\begin{bmatrix} 3 & -2 \\ 2 & 1 \end{bmatrix}$

   d. $\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$

9. What are the eigen values and eigen vectors of the identity matrix?

10. For each of the following matrices over the field $\mathbb{C}$, find the diagonal form and a diagonalizing matrix P.

    a. $\begin{bmatrix} 20 & 18 \\ -27 & -25 \end{bmatrix}$

    b. $\begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}$

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

    d. $\begin{bmatrix} -17 & 18 & -6 \\ -18 & 19 & -6 \\ -9 & 9 & 2 \end{bmatrix}$

11. Show that distinct eigen vectors of a matrix A corresponding to distinct eigen values of A are linearly independent.

12. Let T be the linear operator on $\mathbb{R}^3$ which is represented in the standard ordered basis by the matrix.

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

Find the characteristic values of A and prove that T is diagonalizable.

13. Is the matrix

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

similar over the field $\mathbb{R}$ to a diagonal matrix? Is A similar over the field $\mathbb{C}$ to a diagonal matrix?

14. Is the matrix

$$A = \begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}$$

similar over the field $\mathbb{R}$ to a diagonal matrix? Is A similar over the field $\mathbb{C}$ to a diagonal matrix?

Chapter 2

INNER PRODUCT SPACES

2.1 Inner Product Spaces

Throughout our study, we let K to be either the field of real numbers or the field of complex numbers.

Definition 2.1.1. *Let V be a vector space over K. An inner product on V is a function $\langle\ ,\ \rangle$ from $V \times V$ to K such that for all $u, v, w \in V$ and $k \in K$, we have*

1. *Symmetry axiom :* $\langle u, v\rangle = \overline{\langle v, u\rangle}$
2. *Additivity axiom :* $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$
3. *Homogeneity axiom :* $\langle ku, v\rangle = k\langle u, v\rangle$
4. *Positivity axiom :* $\langle v, v\rangle \geq 0$ *and* $\langle v, v\rangle = 0$ *if and only if* $v = \mathbf{0}$.

A vector space with an inner product is called an inner product space.

We call a vector space over $\mathbb{R}$ with an inner product as a real inner product space and a vector space over $\mathbb{C}$ with an inner product as a complex inner product space.

Problem 2.1.2. *Prove that if V is a complex inner product space, then $\langle u, kv\rangle = \bar{k}\langle u, v\rangle$.*

Problem 2.1.3. *Let P_n be the vector space of all polynomials of degree less then or equal to n over the field of real numbers. Suppose*

$$\begin{aligned} p(x) &= a_0 + a_1x + \cdots + a_nx^n, \\ q(x) &= b_0 + b_1x + \cdots + b_nx^n \end{aligned}$$

are two polynomials in P_n and $x_0, x_1, \cdots, x_n$ are distinct real numbers. Prove that the function

$$\langle p, q\rangle = p(x_0)q(x_0) + p(x_1)q(x_1) + \cdots + p(x_n)q(x_n)$$

is an inner product on P_n.

Proof. Let $p(x) = a_0 + a_1x + \cdots + a_nx^n$, $q(x) = b_0 + b_1x + \cdots + b_nx^n$ and $r(x) = c_0 + c_1x + \cdots + c_nx^n$.

Symmetry axiom:

$$\begin{aligned} \langle p, q\rangle &= p(x_0)q(x_0) + p(x_1)q(x_1) + \cdots + p(x_n)q(x_n) \\ &= q(x_0)p(x_0) + q(x_1)p(x_1) + \cdots + q(x_n)p(x_n) \\ &= \langle q, p\rangle \end{aligned}$$

Additivity axiom:

$$\begin{aligned}
\langle p+q, r\rangle &= (p(x_0)+q(x_0))r(x_0) + (p(x_1)+q(x_1))r(x_1) + \cdots \\
&\qquad\qquad + (p(x_n)+q(x_n))r(x_n) \\
&= p(x_0)r(x_0) + p(x_1)r(x_1) + \cdots + p(x_n)r(x_n) + \\
&\qquad q(x_0)r(x_0) + q(x_1)r(x_1) + \cdots + q(x_n)r(x_n) \\
&= \langle p, r\rangle + \langle q, r\rangle
\end{aligned}$$

Homogeneity axiom:

$$\begin{aligned}
\langle kp, q\rangle &= kp(x_0)q(x_0) + kp(x_1)q(x_1) + \cdots + kp(x_n)q(x_n) \\
&= k(p(x_0)q(x_0) + p(x_1)q(x_1) + \cdots + p(x_n)q(x_n)) \\
&= k\langle p, q\rangle
\end{aligned}$$

Positivity axiom:

$$\langle p, p\rangle = (p(x_0))^2 + (p(x_1))^2 + \cdots + (p(x_n))^2 \geq 0.$$

$$\begin{aligned}
\langle p, p\rangle = 0 &\Leftrightarrow (p(x_0))^2 + (p(x_1))^2 + \cdots + (p(x_n))^2 = 0 \\
&\Leftrightarrow p(x_0) = p(x_1) = \cdots = p(x_n) = 0 \\
&\Leftrightarrow p = 0.
\end{aligned}$$

Thus the given function is an inner product. $\square$

Problem 2.1.4. *Prove that the function* $\langle f, g\rangle = \int\limits_a^b f(x)g(x)dx$ *is an inner product on the vector space* $C[a,b]$ *over* $\mathbb{R}$, *where* $C[a,b]$ *is the set of all continuous functions on* $[a,b]$.

Proof. Let $f = f(x)$, $g = g(x)$ and $h = h(x)$ in $C[a,b]$.

Symmetry axiom:

$$\langle f, g\rangle = \int\limits_a^b f(x)g(x)dx = \int\limits_a^b g(x)f(x) = \langle g, f\rangle$$

Additivity axiom:

$$\begin{aligned} \langle f+g, h\rangle &= \int\limits_a^b (f(x)+g(x))h(x)dx \\ &= \int\limits_a^b f(x)h(x)dx + \int\limits_a^b g(x)h(x)dx \\ &= \langle f, h\rangle + \langle g, h\rangle \end{aligned}$$

Homogeneity axiom:

$$\langle kf, g\rangle = \int\limits_a^b kf(x)g(x)dx = k\int\limits_a^b f(x)g(x)dx = k\langle f, g\rangle$$

Positivity axiom:

$$\langle f, f\rangle = \int\limits_a^b f^2(x) \geq 0$$

$$\begin{aligned} \langle f, f\rangle = 0 &\Leftrightarrow \int\limits_a^b f^2(x) = 0 \\ &\Leftrightarrow f(x) = 0 \text{ for all } x \in [a,b] \\ &\Leftrightarrow f = 0 \end{aligned}$$

Thus the given function is an inner product. $\square$

Problems

1. Prove the following:

   - Let V be a vector space of all $n \times n$ real matrices over the field $\mathbb{R}$. Then the function

     $$\langle U, V \rangle = tr(U^T V)$$

     defines an inner product on V.

   - Let $\mathbb{R}^n$ be a vector space over the field $\mathbb{R}$. Suppose $u = (u_1, u_2, \cdots, u_n)$, $v = (v_1, v_2, \cdots, v_n)$ and $w_1, w_2, \cdots, w_n$ are positive real numbers. Then the function

     $$\langle u, v \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \cdots + w_n u_n v_n$$

     is an inner product on $\mathbb{R}^n$ called weighted Euclidean inner product. If $w_1 = w_2 = \cdots = w_n = 1$, then the above inner product is called as Euclidean inner product or standard inner product on $\mathbb{R}^n$.

   - Let $\mathbb{R}^n$ be a vector space over the field $\mathbb{R}$. Suppose $u = (u_1, u_2, \cdots, u_n)$, $v = (v_1, v_2, \cdots, v_n)$ and A is an invertible $n \times n$ matrix. Then the function

     $$\langle u, v \rangle = Au.Av$$

is an inner product on $\mathbb{R}^n$, where $Au.Av$ is given by the standard inner product of two vectors in $\mathbb{R}^n$.

2. Let P_n be the vector space of all polynomials of degree less then or equal to n over the field of real numbers. Suppose

$$\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_n x^n, \\ q(x) &= b_0 + b_1 x + \cdots + b_n x^n \end{aligned}$$

are two polynomials in P_n and $x_0, x_1, \cdots, x_n$ are distinct real numbers. Prove that the function

$$\langle p, q \rangle = a_0 b_0 + a_1 b_1 + \cdots + a_n b_n$$

is an inner product on P_n (Standard Inner Product).

3. Let $u = (u_1, u_2)$ and $v = (v_1, v_2)$. Prove that $\langle u, v \rangle = 3u_1 v_1 + 5u_2 v_2$ defines an inner product on $\mathbb{R}^2$ by showing that the inner product axioms hold.

4. What conditions must k_1 and k_2 satisfy for $\langle u, v \rangle = k_1 u_1 v_1 + k_2 u_2 v_2$ to define an inner product on $\mathbb{R}^2$? Justify your answer.

5. Show that the expressions does not define an inner product on $\mathbb{R}^3$ and list all inner product axioms that fail to hold.

   (a.) $\langle u, v \rangle = u_1^2 v_1^2 + u_2^2 v_2^2 + u_3^2 v_3^2$

   (b.) $\langle u, v \rangle = u_1 v_1 - u_2 v_2 + u_3 v_3$.

2.2 Length

Definition 2.2.1. *If V is a real inner product space, then the norm (or length) of a vector v in V is denoted by $\|v\|$ is defined as the non-negative square root of $\sqrt{\langle v, v\rangle}$.*

The distance between two vectors is denoted by $d(u, v)$ and is defined by

$$d(u, v) = \|u - v\| = \sqrt{\langle u - v, u - v\rangle}$$

A vector of norm 1 is called unit vector.

Definition 2.2.2. *If V is an inner product space, then the set of points in V that satisfy*

$$\|u\| = 1$$

is called the unit sphere or sometimes the unit circle in V.

Problem 2.2.3. *Sketch the unit circle in an xy-coordinate system in $\mathbb{R}^2$ using the inner product $\langle u, v\rangle = \frac{1}{9}u_1v_1 + \frac{1}{4}u_2v_2$.*

Solution If $u = (x, y)$, then $\|u\| = \langle u, u\rangle^{\frac{1}{2}} = \sqrt{\frac{1}{9}x^2 + \frac{1}{4}y^2}$. So the equation of the unit circle is $\sqrt{\frac{1}{9}x^2 + \frac{1}{4}y^2} = 1$. Squaring on both sides, we get

$$\frac{1}{9}x^2 + \frac{1}{4}y^2 = 1.$$

Theorem 2.2.4. *If u and v are vectors in a real inner product space V and if k is a scalar, then*

- $\|v\| \geq 0$ *and* $\|v\| = 0$ *if and only if* $v = \mathbf{0}$.

- $\|kv\| = |k|\|v\|$.

- $d(u, v) = d(v, u)$

- $d(u, v) \geq 0$ *with equality if and only if* $u = v$.

Theorem 2.2.5. *If u, v and w are vectors in a real inner product space V and if k is a scalar, then*

- $\langle \mathbf{0}, v \rangle = \langle v, \mathbf{0} \rangle = 0$

- $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$

- $\langle u, v - w \rangle = \langle u, v \rangle - \langle u, w \rangle$

- $\langle u - v, w \rangle = \langle u, w \rangle - \langle v, w \rangle$

- $k\langle u, v \rangle = \langle u, kv \rangle$

Proof. To prove (a): consider

$$\begin{aligned} \langle \mathbf{0}, v \rangle &= \langle 0.\mathbf{0}, v \rangle \\ &= 0.\langle \mathbf{0}, v \rangle \text{ [By Homogeneity]} \\ &= 0. \end{aligned}$$

By symmetry, $\langle \mathbf{0}, v \rangle = \langle v, \mathbf{0} \rangle = 0$.

To prove (b): consider

$$\begin{aligned}\langle u, v+w\rangle &= \langle v+w, u\rangle \text{ [By symmetry]}\\ &= \langle v, u\rangle + \langle w, u\rangle \text{ [By additivity]}\\ &= \langle u, v\rangle + \langle u, w\rangle \text{ [By symmetry].}\end{aligned}$$

To prove (e): consider

$$\begin{aligned}\langle u, kv\rangle &= \langle kv, u\rangle \text{ [By symmetry]}\\ &= k\langle v, u\rangle \text{ [By homogeneity]}\\ &= k\langle u, v\rangle \text{ [By symmetry]}\end{aligned}$$

To prove (c): consider

$$\begin{aligned}\langle u, v-w\rangle &= \langle u, v\rangle + \langle u, -w\rangle \text{ [By b]}\\ &= \langle u, v\rangle - \langle u, w\rangle \text{ [By e]}\end{aligned}$$

To prove (d): consider

$$\begin{aligned}\langle u-v, w\rangle &= \langle w, u-v\rangle \text{ [By symmetry]}\\ &= \langle w, u\rangle - \langle w, v\rangle \text{ [By c]}\\ &= \langle w, u\rangle - \langle v, w\rangle \text{ [By symmetry]}\end{aligned}$$

□

Problem 2.2.6. *Prove that* $\langle u-2v, 3u+4v\rangle = 3\|u\|^2 - 2\langle u, v\rangle - 8\|v\|^2$.

Proof.

$$\begin{aligned}
\langle u - 2v, 3u + 4v\rangle &= \langle u, 3u + 4v\rangle - \langle 2v, 3u + 4v\rangle \\
&= \langle u, 3u\rangle + \langle u, 4v\rangle - \langle 2v, 3u\rangle - \langle 2v, 4v\rangle \\
&= 3\langle u, u\rangle + 4\langle u, v\rangle - 6\langle v, u\rangle - 8\langle v, v\rangle \\
&= 3\|u\|^2 + 4\langle u, v\rangle - 6\langle u, v\rangle - 8\|v\|^2 \\
&= 3\|u\|^2 - 2\langle u, v\rangle - 8\|v\|^2
\end{aligned}$$

□

Problems

1. Let the vector space P_3 have the inner product

$$\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)dx$$

Find the following for $p = 2x^3$ and $q = 1 - x^3$

a. $\langle p, q\rangle$

b. $d(p, q)$

c. $\|p\|$

d. $\|q\|$

2. Sketch the unit circle in an xy-coordinate system in $\mathbb{R}^2$ using the Euclidean inner product $\langle u, v\rangle = u_1v_1 + u_2v_2$.

3. Prove that $\langle u, v\rangle = \frac{1}{4}\|u+v\|^2 - \frac{1}{4}\|u-v\|^2$.

4. Prove that $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$

5. Suppose that u and v are vectors in an inner product space. Rewrite the given expression in terms of $\langle u, v\rangle$, $\|u\|^2$ and $\|v\|^2$.

   a. $\langle 2v - 4u, u - 3v\rangle$

   b. $\langle 5u + 6v, 4v - 3u\rangle$

6. Use the inner product

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx$$

   on C[0, 1] to compute $\langle f, g\rangle$.

   a. $f = \cos 2\pi x,\ g = \sin 2\pi x$

   b. $f = x,\ g = e^x$

2.3 Angle and Orthogonality

Theorem 2.3.1 (Cauchy-Schwarz inequality)**.** *If u and v are vectors in a real inner product space V, then*

$$|\langle u, v\rangle| \leq \|u\|\|v\|$$

Proof. If $u = \mathbf{0}$, then the two sides of the inequality are equal, since $\langle u, v\rangle$ and $\|u\|$ are both zero. In a similar way, we can prove the case if $v = \mathbf{0}$.

Consider the case $u \neq \mathbf{0}$ and $v \neq \mathbf{0}$. let $a = \langle u, u\rangle$, $b = 2\langle u, v\rangle$, $c = \langle v, v\rangle$ and let t be any real number. Then by the positivity axiom, we have

$$\begin{aligned} 0 \leq \langle tu + v, tu + v\rangle &= \langle u, u\rangle t^2 + 2\langle u, v\rangle t + \langle v, v\rangle \\ &= at^2 + bt + c \end{aligned}$$

This inequality implies that the quadratic polynomial $at^2 + bt + c$ has either no real roots or a repeated real root. Therefore, its discriminant must satisfy the inequality $b^2 - 4ac \leq 0$. Expressing the coefficients a, b and c in terms of the vectors u and v gives $4\langle u, v\rangle^2 - 4\langle u, u\rangle\langle v, v\rangle \leq 0$ or equivalently,

$$\langle u, v\rangle^2 \leq \langle u, u\rangle\langle v, v\rangle.$$

Taking square root on both sides

$$|\langle u, v\rangle| \leq \|u\|\|v\|.$$

$\square$

Note: $-1 \leq \frac{|\langle u,v\rangle|}{\|u\|\|v\|} \leq 1$.

Definition 2.3.2. *Let V be a inner product space and $u, v \in V$. The angle θ between u and v to be*

$$\theta = \cos^{-1}\left(\frac{|\langle u, v\rangle|}{\|u\|\|v\|}\right).$$

Problem 2.3.3. *Let $M_2(\mathbb{R})$ be a vector space with the inner product*

$$\langle U, V\rangle = tr(U^T V).$$

Find the cosine of the angle between the vectors

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

Solution: $\|u\| = \sqrt{30}$, $\|v\| = \sqrt{14}$ and $\langle u, v\rangle = 16$.

$$\cos\theta = \frac{\langle u, v\rangle|}{\|u\|\|v\|} = \frac{16}{\sqrt{30}\sqrt{14}} = 0.78.$$

Theorem 2.3.4. *If u, v and w are vectors in a real inner product space V and if k is any scalar, then*

a. $\|u+v\| \leq \|u\| + \|v\|$ *[Triangle inequality for vectors]*

b. $d(u,v) \leq d(u,w) + d(w,v)$ *[Triangle inequality for distances]*

Proof. To prove (a) consider,

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle \\ &\leq \langle u,u\rangle + 2|\langle u,v\rangle| + \langle v,v\rangle \\ &\leq \langle u,u\rangle + 2\|u\|\|v\| + \langle v,v\rangle \\ &= \|u\| + 2\|u\|\|v\| + \|v\| \\ &= (\|u\| + \|v\|)^2 \end{aligned}$$

Taking square root we get

$$\|u+v\| = \|u\| + \|v\|.$$

The proof of (b) follows from (a). □

2.4 Orthogonal sets

Definition 2.4.1. *Two vectors u and v in an inner product space V called orthogonal if $\langle u, v\rangle = 0$.*

Example 2.4.2.

1. *The vectors $u = (1, 1)$ and $v = (1, -1)$ are orthogonal with respect to the Euclidean inner product on $\mathbb{R}^2$ since $\langle u, v\rangle = (1)(1) + (1)(-1) = 0$.*

   However, they are not orthogonal with respect to the inner product $\langle u, v\rangle = 3u_1v_1 + 2u_2v_2$, since $\langle u, v\rangle = 3(1)(1) + 2(1)(-1) = 1 \neq 0$

2. *The matrices $U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $V = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$ are orthogonal with respect to the standard inner product, since*

$$\langle U, V\rangle = 1(0) + 0(2) + 1(0) + 1(0) = 0.$$

3. *The polynomials $p(x) = x$ and $q(x) = x^2$ are orthogonal with respect to an inner product*

$$\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)dx$$

   since $\langle p, q\rangle = \int_{-1}^{1} x.x^2dx = 0$.

Theorem 2.4.3 (Generalized Theorem of Pythagoras)**.** *If u and v are orthogonal vectors in a real inner product space, then*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Proof. The orthogonality of u and v implies that $\langle u, v\rangle = 0$, so

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \|u\| + 2\langle u, v\rangle + \|v\|^2 \\ &= \|u\|^2 + \|v\|^2.\end{aligned}$$

□

Definition 2.4.4. *If W is a subspace of a real inner product space V, then the set of all vectors in V that are orthogonal to every vector in W is called the orthogonal complement of W and is denoted by the symbol $W^\perp$.*

Theorem 2.4.5. *If W is a subspace of a real inner product space V, then*

a. $W^\perp$ *is a subspace of* V

b. $W \cap W^\perp = \{\mathbf{0}\}$

Proof. (a.) The set $W^\perp$ contains at least the zero vector, since $\langle 0, w\rangle = 0$ for every vector w in W. It remains to show that $W^\perp$ is closed under addition and scalar multiplication. suppose that u and v are vectors in $W^\perp$. Then for every vector w in W we have $\langle u, w\rangle = 0$ and $\langle v, w\rangle = 0$. From the additivity and homogeneity axioms of inner products we have

$$\begin{aligned}\langle u+v, w\rangle &= \langle u, w\rangle + \langle v, w\rangle = 0 + 0 = 0 \\ \langle ku, w\rangle &= k\langle u, w\rangle = k.0 = 0\end{aligned}$$

Thus $u + v$ and ku are in $W^\perp$ and hence $W^\perp$ is a subspace of V.

b. If v is any vector in both W and $W^{\perp}$, then v is orthogonal to itself; that is, $\langle v, v\rangle = 0$. By the positivity axiom for inner products, we get $v = \mathbf{0}$.

$\square$

Theorem 2.4.6. *If W is a subspace of a real finite-dimensional inner product space V, then the orthogonal complement of $W^{\perp}$ is W; that is, $(W^{\perp})^{\perp} = W$.*

Problem 2.4.7. *Let W be the subspace of $\mathbb{R}^6$ spanned by the vectors*

$$w_1 = (1, 3, -2, 0, 2, 0) \qquad w_2 = (2, 6, -5, -2, 4, -3)$$

$$w_3 = (0, 0, 5, 10, 0, 15) \qquad w_4 = (2, 6, 0, 8, 4, 18)$$

Find a basis for the orthogonal complement of W.

Solution The subspace W is the same as the row space of the matrix

$$A = \begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 18 \end{bmatrix}$$

Since the row space and null space of A are orthogonal complements, our problem reduces to finding a basis for the null space of this matrix.

The null space of A is the solution space of the homogeneous linear system $Ax = 0$. (ie.)

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The solutions are

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} -3r-4s-2t \\ r \\ -2s \\ s \\ t \\ 0 \end{bmatrix} = r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Thus $v_1 = (-3, 1, 0, 0, 0, 0)$, $v_2 = (-4, 0, -2, 1, 0, 0)$, $v_3 = (-2, 0, 0, 0, 1, 0)$ are basis vectors for the null space of A.

Problems

1. Prove that equality holds in the Cauchy Schwarz inequality if and only if u and v are linearly dependent.

2. Find the cosine of the angle between A and B with respect to the standard inner product on M_{22}.

   a. $A = \begin{bmatrix} 2 & 6 \\ 1 & -3 \end{bmatrix}$, $A = \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$

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

3. Let W be the line in R^2 with equation $y = 2x$. Find an equation for $W^{\perp}$.

4. Let W be the y-axis in an xyz-coordinate system in R^3. Describe the subspace $W^{\perp}$.

5. Let $C[0, 1]$ be the inner product space with standard inner product.

Show that the vectors

$$p(x) = 1 \text{ and } q(x) = \frac{1}{2} - x$$

are orthogonal and satisfy the Theorem of Pythagoras.

6. Let V be an inner product space. Show that if u and v are orthogonal unit vectors in V, then $\|u - v\| = \sqrt{2}$.

7. Do there exist scalars k and l such that the vectors $p_1 = 2 + kx + 6x^2$, $p_2 = l + 5x + 3x^2$, $p^3 = 1 + 2x + 3x^2$ are mutually orthogonal with respect to the standard inner product on P_2?

8. Let P_3 have the standard inner product, and let

$$p = -1 - x + 2x^2 + 4x^3$$

Determine whether p is orthogonal to the subspace spanned by the polynomials $w_1 = 2 - x^2 + x^3$ and $w_2 = 4x - 2x^2 + 2x^3$.

2.5 Orthonormal Basis

Definition 2.5.1. *A set of two or more vectors in a real inner product space is said to be orthogonal if all pairs of distinct vectors in the set are orthogonal. An orthogonal set in which each vector has norm* 1 *is said to be orthonormal.*

Example 2.5.2. *Let* $V = \mathbb{R}^3$ *be an inner product space with Euclidean inner product and* $v_1 = (0, 1, 0)$*,* $v_2 = (1, 0, 1)$*,* $v_3 = (1, 0, -1)$ *be vectors in* V*. Then*

$\langle v_1, v_2\rangle = \langle v_1, v_3\rangle = \langle v_2, v_3\rangle = 0$. *Thus the set of vectors* $S = \{v_1, v_2, v_3\}$ *is orthogonal*

Normalizing a vector: A simple way to convert an orthogonal set of nonzero vectors into an orthonormal set is to multiply each vector v in the orthogonal set by the reciprocal of its length to create a vector of norm 1 (called a unit vector). This process of multiplying a vector v by the reciprocal of its length is called normalizing v.

Theorem 2.5.3. *If* $S = \{v_1, v_2, \cdots, v_n\}$ *is an orthogonal set of nonzero vectors in an inner product space, then* S *is linearly independent.*

Proof. Let $S = \{v_1, v_2, \cdots, v_n\}$ be an orthogonal set of nonzero vectors. Suppose $k_1v_1 + k_2v_2 + \cdots + k_nv_n = \mathbf{0}$, then we have to prove that $k_1 = k_2 = \cdots = k_n = 0$. For each $v_i \in S$, we have

$$\langle k_1v_1 + k_2v_2 + \cdots + k_nv_n, v_i\rangle = \langle \mathbf{0}, v_i\rangle = 0,$$

which implies

$$k_1\langle v_1, v_i\rangle + k_2\langle v_2, v_i\rangle + \cdots + k_n\langle v_n, v_i\rangle = 0.$$

From the orthogonality of S, we get $\langle v_j, v_i\rangle = 0$, for all $j \neq i$. Also as $k_i\langle v_i, v_i\rangle = 0$ and since the vectors in S are nonzero, we have $k_i = 0$ for all $i = 1, 2, \cdots, n$ as desired. $\square$

Definition 2.5.4. *In an inner product space, a basis consisting of orthonormal vectors is called an orthonormal basis and a basis consisting of orthogonal vectors is called an orthogonal basis.*

Example 2.5.5.

1. $\mathbb{R}^n$ *with Euclidean inner product has a standard basis* $\{(1,0,0,\cdots,0),(0,1,0,\cdots,0),\cdots,(0,0,0,\cdots,1)\}$ *which is an orthonormal one.*

2. P_n *with standard inner product has a standard basis* $\{1,x,x^2,\cdots,x^n\}$ *which is an orthonormal one.*

Theorem 2.5.6.

1. *If* $S=\{v_1,v_2,\cdots,v_n\}$ *is an orthogonal basis for an inner product space* V *and suppose* u *is any vector in* V*, then*

$$u=\frac{\langle u,v_1\rangle}{\|v_1\|^2}v_1+\frac{\langle u,v_2\rangle}{\|v_2\|^2}v_2+\cdots+\frac{\langle u,v_n\rangle}{\|v_n\|^2}v_n$$

2. *If* $S=\{v_1,v_2,\cdots,v_n\}$ *is an orthonormal basis for an inner product space* V *and suppose* u *is any vector in* V*, then*

$$u=\langle u,v_1\rangle v_1+\langle u,v_2\rangle v_2+\cdots+\langle u,v_n\rangle v_n.$$

Proof.

1. Since $S = \{v_1, v_2, \cdots, v_n\}$ is a basis for V, every vector u in V can be expressed in the form

$$u = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n$$

It is enough to prove that $k_i = \frac{\langle u, v_i \rangle}{\|v_i\|^2}$ for $i = 1, 2, \cdots, n$. For each i, we have

$$\begin{aligned}\langle u, v_i \rangle &= \langle k_1 v_1 + k_2 v_2 + \cdots + k_n v_n, v_i \rangle \\ &= k_1 \langle v_1, v_i \rangle + k_2 \langle v_2, v_i \rangle + \cdots + k_n \langle v_n, v_i \rangle\end{aligned}$$

Since S is an orthogonal set, $\langle v_j, v_i \rangle = 0$, when $j \neq i$. Thus

$$\langle u, v_i \rangle = k_i \langle v_i, v_i \rangle = k_i \|v_i\|^2$$

which implies

$$k_i = \frac{\langle u, v_i \rangle}{\|v_i\|^2}.$$

2. Since $\|v_1\| = \|v_2\| = \cdots = \|v_n\| = 1$ and by (a), we get

$$u = \langle u, v_1 \rangle v_1 + \langle u, v_2 \rangle v_2 + \cdots + \langle u, v_n \rangle v_n.$$

□

Note: The coordinate vector of a vector u in V relative to an orthogonal basis $S = \{v_1, v_2, \cdots, v_n\}$ is

$$(u)_S = \left(\frac{\langle u, v_1 \rangle}{\|v_1\|^2}, \frac{\langle u, v_2 \rangle}{\|v_2\|^2}, \cdots, \frac{\langle u, v_n \rangle}{\|v_n\|^2} \right)$$

and relative to an orthonormal basis $S = \{v_1, v_2, \cdots, v_n\}$ is

$$(u)_S = (\langle u, v_1\rangle, \langle u, v_2\rangle, \cdots, \langle u, v_n\rangle).$$

Problem 2.5.7. *Let*

$$S = \left\{(0,1,0), (\frac{-4}{5}, 0, \frac{3}{5}), (\frac{3}{5}, 0, \frac{4}{5})\right\}$$

be an orthonormal basis for $\mathbb{R}^3$ with the Euclidean inner product. Express the vector $u = (1,1,1)$ as a linear combination of the vectors in S and find the coordinate vector $(u)_S$.

Solution: The coordinate vector of u relative to S is

$$(u)_S = (\langle u, v_1\rangle, \langle u, v_2\rangle, \cdots, \langle u, v_n\rangle) = \left(1, \frac{-1}{5}, \frac{7}{5}\right).$$

Problems

1. Find vectors x and y in $\mathbb{R}^2$ that are orthonormal with respect to the inner product $\langle u, v\rangle = 3u_1v_1 + 2u_2v_2$ but are not orthonormal with respect to the Euclidean inner product.

2. Verify that the set of vectors $(1,0), (0,1)$ is orthogonal with respect to the inner product $\langle u, v\rangle = 4u_1v_1 + u_2v_2$ on $\mathbb{R}^2$; then convert it to an orthonormal set by normalizing the vectors.

3. Determine whether the set of vectors

$$p_1(x) = \frac{2}{3} - \frac{2}{3}x + \frac{1}{3}x^2,\ p_2(x) = \frac{2}{3} + \frac{1}{3}x - \frac{2}{3}x^2,\ p_3(x) = \frac{1}{3} + \frac{2}{3}x + \frac{2}{3}x^2$$

is orthogonal with respect to the standard inner product on P_2

4. show that the following column vectors of A form an orthogonal basis for the column space of A with respect to the Euclidean inner product, and then find an orthonormal basis for that column space.

   a. $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 5 \\ -1 & 2 & 0 \end{bmatrix}$

   b. $\begin{bmatrix} \frac{1}{5} & \frac{-1}{2} & \frac{1}{3} \\ \frac{1}{5} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{5} & 0 & \frac{-2}{3} \end{bmatrix}$

2.6 Orthogonal projections

Theorem 2.6.1. *(Projection Theorem) If W is a finite-dimensional subspace of an inner product space V, then every vector u in V can be expressed in exactly one way as*

$$u = w_1 + w_2$$

where w_1 is in W and w_2 is in $W^{\perp}$.

In the above theorem, the vectors w_1 and w_2 are commonly denoted by

$$w_1 = proj_W u \text{ and } w_2 = proj_{W^{\perp}} u$$

These are called the orthogonal projection of u on W and the orthogonal projection of u on $W^{\perp}$ respectively. The vector w_2 is also called the component of u orthogonal to W. Thus u can be written as

$$u = proj_W u + proj_{W^{\perp}} u.$$

Also since $proj_{W^\perp} = u - proj_W u$, $u = proj_W u + (u - proj_W u)$.

The following theorem provides formulae for calculating orthogonal projections.

Theorem 2.6.2. *Let W be a finite-dimensional subspace of an inner product space V.*

a. If $\{v_1, v_2, \cdots, v_n\}$ is an orthogonal basis for W, and u is any vector in V, then

$$proj_W u = \frac{\langle u, v_1\rangle}{\|v_1\|^2} v_1 + \frac{\langle u, v_2\rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle u, v_n\rangle}{\|v_r\|^2} v_n.$$

b. If $\{v_1, v_2, \cdots, v_n\}$ is an orthonormal basis for W, and u is any vector in V, then

$$proj_W u = \langle u, v_1\rangle v_1 + \langle u, v_2\rangle v_2 + \cdots + \langle u, v_n\rangle v_n.$$

Proof.

(a.) The vector u can be expressed in the form $u = w_1 + w_2$, where $w_1 = proj_W u$ is in W and w_2 is in $W^\perp$. The component $proj_W u = w_1$ can be expressed in terms of the basis vectors for W as

$$proj_W u = w_1 = \frac{\langle w_1, v_1\rangle}{\|v_1\|^2} v_1 + \frac{\langle w_1, v_2\rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle w_1, v_n\rangle}{\|v_r\|^2} v_n.$$

Since w_2 is orthogonal to W, $\langle w_2, v_1\rangle = \langle w_2, v_2\rangle = \cdots = \langle w_2, v_n\rangle = 0$. Thus

$$proj_W u = w_1 = \frac{\langle w_1 + w_2, v_1\rangle}{\|v_1\|^2} v_1 + \frac{\langle w_1 + w_2, v_2\rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle w_1 + w_2, v_n\rangle}{\|v_r\|^2} v_n.$$

which implies

$$proj_W u = w_1 = \frac{\langle u, v_1\rangle}{\|v_1\|^2}v_1 + \frac{\langle u, v_2\rangle}{\|v_2\|^2}v_2 + \cdots + \frac{\langle u, v_n\rangle}{\|v_r\|^2}v_n$$

(b.) Since $\|v_1\| = \|v_2\| = \cdots = \|v_n\| = 1$,

$$proj_W u = \langle u, v_1\rangle v_1 + \langle u, v_2\rangle v_2 + \cdots + \langle u, v_n\rangle v_n.$$

$\square$

Example 2.6.3. *Let $\mathbb{R}^3$ have the Euclidean inner product, and let W be the subspace spanned by the orthonormal vectors $v_1 = (0, 1, 0)$ and $v_2 = \left(\frac{-4}{5}, 0, \frac{3}{5}\right)$.*

The orthogonal projection of $u = (1, 1, 1)$ on W is

$$\begin{aligned} proj_W u &= \langle u, v_1\rangle v_1 + \langle u, v_2\rangle v_2 \\ &= (1)(0, 1, 0) + \frac{-1}{5}\left(\frac{-4}{5}, 0, \frac{3}{5}\right) \\ &= \left(\frac{4}{25}, 1, \frac{-3}{25}\right). \end{aligned}$$

The component of u orthogonal to W is

$$proj_{W^\perp} u = u - proj_W u = (1, 1, 1) - \left(\frac{4}{25}, 1, \frac{-3}{25}\right) = \left(\frac{21}{25}, 0, \frac{28}{25}\right).$$

Problems

1. Let $\mathbb{R}^3$ have the Euclidean inner product.

a. Find the orthogonal projection of u onto the plane spanned by the vectors v_1 and v_2.

b. Find the component of u orthogonal to the plane spanned by the vectors v_1 and v_2, and confirm that this component is orthogonal to the plane.

i $u = (4, 2, 1);\ v_1 = \left(\frac{1}{3}, \frac{2}{3}, \frac{-2}{3}\right),\ v_2 = \left(\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right)$

ii $u = (3, -1, 2);\ v_1 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}\right),\ v_2 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$.

2. The vectors v_1 and v_2 are orthogonal with respect to the Euclidean inner product on $\mathbb{R}^4$. Find the orthogonal projection of $b = (1, 2, 0, -2)$ on the subspace W spanned by these vectors.

   a. $v_1 = (1, 1, 1, 1),\ v_2 = (1, 1, -1, -1)$

   b. $v_1 = (0, 1, -4, -1),\ v_2 = (3, 5, 1, 1)$.

2.7 Gram-Schmidt process

Theorem 2.7.1. *Every nonzero finite-dimensional inner product space has an orthonormal basis.*

Proof. Let W be any nonzero finite-dimensional subspace of an inner product space and suppose that $\{u_1, u_2, \cdots, u_r\}$ is any basis for W. It suffices to show that W has an orthogonal basis since the vectors in that basis can be normalized to obtain an orthonormal basis.

The following sequence of steps will produce an orthogonal basis $\{v_1, v_2, \cdots, v_r\}$ for W :

Step 1. Let $v_1 = u_1$.

Step 2. We can obtain a vector v_2 that is orthogonal to v_1 by computing the component of u_2 that is orthogonal to the space W_1 spanned by v_1.

$$v_2 = u_2 - proj_{W_1} u_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1$$

If $v_2 = \mathbf{0}$, then $u_2 = \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} u_1$, which implies that u_2 is a multiple of u_1, contradicting the linear independence of the basis $\{u_1, u_2, \cdots, u_r\}$.

Step 3. To construct a vector v_3 that is orthogonal to both v_1 and v_2, we compute the component of u_3 orthogonal to the space W_2 spanned by v_1 and v_2, let

$$v_3 = u_3 - proj_{W_2} u_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2.$$

As in Step 2, the linear independence of $\{u_1, u_2, \cdots, u_r\}$ ensures that $v_3 \neq \mathbf{0}$. We leave the details for you.

Step 4. To determine a vector v_4 that is orthogonal to v_1, v_2, and v_3 we compute the component of u_4 orthogonal to the space W_3 spanned by v_1, v_2, and v_3, let

$$v_4 = u_4 - proj_{W_3} u_4 = u_4 - \frac{\langle u_4, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_4, v_2 \rangle}{\|v_2\|^2} v_2 - \frac{\langle u_4, v_3 \rangle}{\|v_3\|^2} v_3$$

Continuing in this way we will produce after r steps an orthogonal set of nonzero vectors $\{v_1, v_2, \cdots, v_r\}$. Since such sets are linearly independent, we will have produced an orthogonal basis for the r-dimensional space W. By normalizing these basis vectors we can obtain an orthonormal basis. $\square$

The step-by-step construction of an orthogonal (or orthonormal) basis given in the foregoing proof is called the Gram-Schmidt process.

The Gram-Schmidt Process

To convert a basis $\{u_1, u_2, \cdots, u_r\}$ into an orthogonal basis $\{v_1, v_2, \cdots, v_r\}$, perform the following computations:

Step 1 Let $v_1 = u_1$.

Step 2 $v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1$

Step 3 $v_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2$.

Step 4 $v_4 = u_4 - \frac{\langle u_4, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_4, v_2 \rangle}{\|v_2\|^2} v_2 - \frac{\langle u_4, v_3 \rangle}{\|v_3\|^2} v_3$

$\vdots$

(continue for r steps)

Optional Step. To convert the orthogonal basis into an orthonormal basis, say $\{q_1, q_2, \cdots, q_r\}$, normalize the orthogonal basis vectors.

Problem 2.7.2. *Assume that the vector space $\mathbb{R}^3$ has the Euclidean inner product. Apply the Gram-Schmidt process to transform the basis vectors*

$$u_1 = (1, 1, 1), u_2 = (0, 1, 1), u_3 = (0, 0, 1)$$

into an orthogonal basis $\{v_1, v_2, v_3\}$, and then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{q_1, q_2, q_3\}$.

Solution

Step 1 $v_1 = u_1 = (1, 1, 1)$.

Step 2

$$\begin{aligned} v_2 &= u_2 - proj_{W_1} u_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\|v_1\|^2} v_1 \\ &= (0, 1, 1) - \frac{2}{3}(1, 1, 1) \\ &= \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right) \end{aligned}$$

Step 3

$$\begin{aligned} v_3 &= u_3 - \frac{\langle u_3, v_1\rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2\rangle}{\|v_2\|^2} v_2 \\ &= (0, 0, 1) - \frac{1}{3}(1, 1, 1) - \frac{1/3}{2/3}\left(\frac{-2}{3}, \frac{1}{3}, \frac{1}{3}\right) \\ &= \left(0, \frac{-1}{2}, \frac{1}{2}\right) \end{aligned}$$

Thus clearly,

$$v_1 = (1, 1, 1),\ v_2 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right),\ v_3 = \left(0, \frac{-1}{2}, \frac{1}{2}\right)$$

forms an orthogonal basis for $\mathbb{R}^3$. The norms of these vectors are

$$\|v_1\| = \sqrt{3},\ v_2 = \frac{\sqrt{6}}{3},\ \|v_3\| = \frac{1}{\sqrt{2}}.$$

Therefore the required orthonormal basis for $\mathbb{R}^3$ is given by

$$\begin{aligned} q_1 &= \frac{v_1}{\|v_1\|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \\ q_2 &= \frac{v_2}{\|v_2\|} = \left(\frac{-2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) \\ q_3 &= \frac{v_3}{\|v_3\|} = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right). \end{aligned}$$

Theorem 2.7.3. *If W is a finite-dimensional inner product space, then*

a. *Every orthogonal set of nonzero vectors in W can be enlarged to an orthogonal basis for W.*

b. *Every orthonormal set in W can be enlarged to an orthonormal basis for W.*

Proof. Proof of (a) follows directly. To prove (b.), let $S = \{v_1, v_2, \cdots, v_s\}$ be an orthonormal set of vectors in W. Then S is linearly independent.

We know that every linearly independent set in a subspace is either a basis for that subspace or can be extended to a basis for it. we can enlarge S to some basis $S' = \{v_1, v_2, \cdots, v_s, v_{s+1}, \cdots, v_k\}$ for W.

If we now apply the Gram-Schmidt process to the set S, then the vectors $v_1, v_2, \cdots, v_s$ will not be affected since they are already orthonormal, and the resulting set

$$S'' = \{v_1, v_2, \cdots, v_s, v'_{s+1}, \cdots, v'_k\}$$

will be an orthonormal basis for W.

$\square$

Problems

1. Let P_2 have the inner product

$$\langle p, q\rangle = \int_0^1 p(x)q(x)dx$$

Apply the GramSchmidt process to transform the standard basis $S = \{1, x, x^2\}$ into an orthonormal basis.

2. let $\mathbb{R}^2$ have the Euclidean inner product and use the Gram-Schmidt process to transform the basis $\{u_1, u_2\}$ into an orthonormal basis. Draw both sets of basis vectors in the xy-plane.

   a. $u_1 = (1, -3),\ u_2 = (2, 2)$

   b. $u_1 = (1, 0),\ u_2 = (3, -5)$

3. Let $\mathbb{R}^3$ have the Euclidean inner product. Find an orthonormal basis for the subspace spanned by $(0, 1, 2), (-1, 0, 1), (-1, 1, 3)$.

2.8 QR Decomposition

Theorem 2.8.1 (QR Decomposition)**.** *If A is an $m \times n$ matrix with linearly independent column vectors, then A can be factored as*

$$A = QR$$

where Q is an $m \times n$ matrix with orthonormal column vectors and R is an $n \times n$ invertible upper triangular matrix.

Proof. Let A be an $m \times n$ matrix with linearly independent column vectors, and let Q be the matrix that results by applying the Gram-Schmidt process to the column vectors of A.

Suppose that the column vectors of A are $u_1, u_2, \cdots, u_n$ and that Q has orthonormal column vectors $q_1, q_2, \cdots, q_n$. Thus A and Q can be written in partitioned form as

$$A = \begin{bmatrix} u_1| & u_2| & \cdots & |u_n \end{bmatrix} \text{ and } Q = \begin{bmatrix} q_1| & q_2| & \cdots & |q_n \end{bmatrix}.$$

We know that, the column vectors $u_1, u_2, \cdots, u_n$ can be expressible in terms of the vectors $q_1, q_2, \cdots, q_n$ as

$$\begin{aligned} u_1 &= \langle u_1, q_1\rangle q_1 + \langle u_1, q_2\rangle q_2 + \cdots + \langle u_1, q_n\rangle q_n \\ u_2 &= \langle u_2, q_1\rangle q_1 + \langle u_2, q_2\rangle q_2 + \cdots + \langle u_2, q_n\rangle q_n \\ &\vdots \\ u_n &= \langle u_n, q_1\rangle q_1 + \langle u_n, q_2\rangle q_2 + \cdots + \langle u_n, q_n\rangle q_n \end{aligned}$$

The jth column vector of a matrix product is a linear combination of the column vectors of the first factor with coefficients coming from the jth column of the second factor. Therefore the above relationships can be expressed in the matrix form as

$$\begin{bmatrix} u_1| & u_2| & \cdots & |u_n \end{bmatrix} = \begin{bmatrix} q_1| & q_2| & \cdots & |q_n \end{bmatrix} \begin{bmatrix} \langle u_1, q_1\rangle & \langle u_2, q_1\rangle & \cdots & \langle u_n, q_1\rangle \\ \langle u_1, q_2\rangle & \langle u_2, q_2\rangle & \cdots & \langle u_n, q_2\rangle \\ \vdots & \vdots & & \vdots \\ \langle u_1, q_n\rangle & \langle u_2, q_n\rangle & \cdots & \langle u_n, q_n\rangle \end{bmatrix}$$

which implies

$$A = QR$$

where R is the second factor in the product. However, it is a property of the Gram-Schmidt process that for $j \geq 2$, the vector q_j is orthogonal to $u_1, u_2, \cdots, u_{j-1}$. Thus, all entries below the main diagonal of R are zero, and R has the form

$$\begin{bmatrix} \langle u_1, q_1 \rangle & \langle u_2, q_1 \rangle & \cdots & \langle u_n, q_1 \rangle \\ 0 & \langle u_2, q_2 \rangle & \cdots & \langle u_n, q_2 \rangle \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \langle u_n, q_n \rangle \end{bmatrix}$$

Since all the diagonal entries are nonzero, R is invertible. $\square$

Note: Every invertible matrix has a QR-decomposition.

Example 2.8.2. *Find a QR-decomposition of the matrix*

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

Solution. *The column vectors the matrix are given by*

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \; u_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, u_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Applying the Gram-Schmidt process with normalization to these column vectors yields the orthonormal vectors

$$q_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \; q_2 = \begin{bmatrix} \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, q_3 = \begin{bmatrix} 0 \\ \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Therefore R is given by

$$R = \begin{bmatrix} \langle u_1, q_1 \rangle & \langle u_2, q_1 \rangle & \langle u_3, q_1 \rangle \\ 0 & \langle u_2, q_2 \rangle & \langle u_3, q_2 \rangle \\ 0 & 0 & \langle u_3, q_3 \rangle \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{3}} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Hence the QR-decomposition of A is given by

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{-2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{3}{\sqrt{3}} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

■

Problems

1. Find a QR-decomposition of the following matrices

   a. $\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$

   b. $\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 4 \end{bmatrix}$

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