Linear Algebra

Vector Spaces

September 2015

Dr P Kathirgamanathan

vector space

- A **vector space** (also called a **linear space**) is a collection of objects called **vectors**, which may be addedtogether and multiplied ("scaled") by numbers, called *scalars* in this context. Scalars are often taken to be real numbers, but there are also vector spaces with scalar multiplication by complex numbers, rational numbers, or generally any field. The operations of vector addition and scalar multiplication must satisfy certain requirements, called *axioms*, listed below

The following defines the notion of a vector space V where K is the field of scalars.

DEFINITION: Let V be a nonempty set with two operations:

i. ***Vector Addition:*** This assigns to any $u, v \in V$ a *sum* $u + v$ in V

ii. ***Scalar Multiplication:*** This assigns to any $u \in V$, $k \in K$ a *product* $ku \in V$.

Then V is called a *vector space* (over the field K) if the following axioms hold for any vectors $u, v, w \in V$:

[A_1] $(u + v) + w = u + (v + w)$

[A_2] There is a vector in V, denoted by 0 and called the *zero vector*, such that, for any $u \in V$,
$u + 0 = 0 + u = u$

[A_3] For each $u \in V$, there is a vector in V, denoted by $-u$, and called the *negative* of u, such that
$u + (-u) = (-u) + u = 0.$

[A_4] $u + v = v + u.$

$[M_1]$ $k(u + v) = ku + kv$, for any scalar $k \in K$.

$[M_2]$ $(a + b)u = au + bu$, for any scalars $a, b \in K$.

$[M_3]$ $(ab)u = a(bu)$, for any scalars $a, b \in K$.

$[M_4]$ $1u = u$, for the unit scalar $1 \in K$.

EXAMPLE The Zero Vector Space

Let V consist of a single object, which we denote by $\mathbf{0}$, and define

$$\mathbf{0}+\mathbf{0}=\mathbf{0} \quad \text{and} \quad k\mathbf{0}=\mathbf{0}$$

for all scalars k. It is easy to check that all the vector space axioms are satisfied. We call this the *zero vector space*.

EXAMPLE R^n Is a Vector Space

Let $V = R^n$, and define the vector space operations on V to be the usual operations of addition and scalar multiplication of n-tuples; that is,

$$\begin{aligned} \mathbf{u}+\mathbf{v} &= (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n) = (u_1+v_1, u_2+v_2, \ldots, u_n+v_n) \\ k\mathbf{u} &= (ku_1, ku_2, \ldots, ku_n) \end{aligned}$$

The set $V = R^n$ is closed under addition and scalar multiplication because the foregoing operations produce n-tuples as their end result.

The Vector Space of Infinite Sequences of Real Numbers

Let V consist of objects of the form

$$\mathbf{u} = (u_1, u_2, \ldots, u_n, \ldots)$$

in which $u_1, u_2, \ldots, u_n, \ldots$ is an infinite sequence of real numbers. We define two infinite sequences to be *equal* if their corresponding components are equal, and we define addition and scalar multiplication componentwise by

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (u_1, u_2, \ldots, u_n, \ldots) + (v_1, v_2, \ldots, v_n \ldots) \\ &= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n, \ldots) \\ k\mathbf{u} &= (ku_1, ku_2, \ldots ku_n, \ldots) \end{aligned}$$

We leave it as an exercise to confirm that V with these operations is a vector space. We will denote this vector space by the symbol R^∞.

A Vector Space of 2 × 2 Matrices

Let V be the set of 2×2 matrices with real entries, and take the vector space operations on V to be the usual operations of matrix addition and scalar multiplication; that is,

$$\mathbf{u}+\mathbf{v}=\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}+\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}=\begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix}$$

$$k\mathbf{u}=k\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}=\begin{bmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{bmatrix}$$

Subspace of a Linear Vector Space

A subset M of a linear vector space V is a subspace of V if:

a. $x + y \in M$, whenever $x \in M$ and $y \in M$.
b. $\alpha x \in M$, whenever $x \in M$ and $\alpha \in R^1(C^1)$.

Example

The set R^n consisting of all n dimensional vectors is a linear vector space, with addition and scalar multiplication defined as:

$$x + y = [x_1 + y_1 \quad x_2 + y_2 \quad \cdots \quad x_n + y_n]^T$$

and

$$\alpha x = [\alpha x_1 \quad \alpha x_2 \quad \cdots \quad \alpha x_n]^T$$

where $x = [x_1 \quad x_2 \quad \ldots \quad x_n]^T$ and $y = [y_1 \quad y_2 \quad \ldots \quad y_n]^T$ are elements of V and α is a scalar. Note that the set of real numbers is by itself a linear vector space. The complex linear vector space C^n consists of all n dimensional vectors and is defined similarly to R^n.

EXAMPLE

Consider the vector space R^2. Let S be the set of all vectors in R^2 whose components are nonnegative that is:

$$S = \{v = [x \quad y]^T \in R^2 \mid x \geq 0 \text{ and } y \geq 0\}$$

S is not a subspace of R^2 since if $v \in S$, then αv is not in S for negative values of α.

Example

Let V be a vector space consisting of all 3×3 matrices. The set of all lower triangular 3×3 matrices is a subspace of vector space V. This can be easily verified since:

a. The sum of two lower triangular matrices is a lower triangular matrix, and
b. multiplying a lower triangular matrix by a scale factor results in another lower triangular matrix.

Consider set S, a subset of vector space R^3, defined by:

$$S = \{v = [x \quad y \quad z]^T \in R^3 \mid x = y\}$$

S is nonempty since $x = [1 \ 1 \ 0]^T \in S$. Furthermore:

i. If $v = [x \quad x \quad z] \in S$, then $\alpha v = [\alpha x \quad \alpha x \quad \alpha z] \in S$

ii. If $v_1 = [x_1 \quad x_1 \quad z_1] \in S$ and $v_2 = [x_2 \quad x_2 \quad z_2] \in S$ then the sum

$$v_1 + v_2 = [x_1 + x_2 \quad x_1 + x_2 \quad z_1 + z_2] \in S.$$

Since S is nonempty and satisfies the two closure conditions, S is a subspace of R^3.

Span of a Set of Vectors

A set of vectors can be combined to form a new space, called the span of a set of vectors. The span of a set of vectors can be defined in term of linear combinations.

Definition of linear combination: Vector V is a linear combination of vectors $V_1, V_2, \ldots, V_n$ if there are scalar constants $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that:

$$V = \alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_n V_n$$

Example

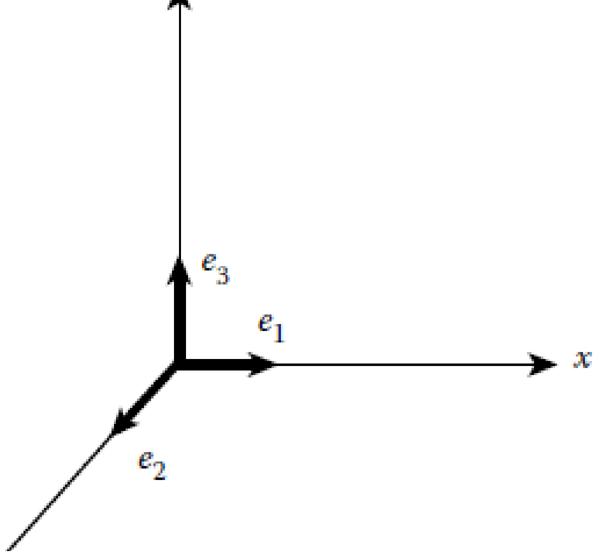
Show that vector $V = [1 \quad 8 \quad 11]^T$ is a linear combination of $V_1 = [1 \quad 2 \quad 1]^T$ and $V_2 = [-1 \quad 1 \quad 4]^T$

Solution:

$$V = \alpha_1 V_1 + \alpha_2 V_2 \rightarrow \begin{bmatrix} 1 \\ 8 \\ 11 \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix} \rightarrow \begin{cases} \alpha_1 - \alpha_2 = 1 \\ 2\alpha_1 + \alpha_2 = 8 \\ \alpha_1 + 4\alpha_2 = 11 \end{cases}$$

The above set of three equations with two unknowns has a unique solution $\alpha_1 = 3$, $\alpha_2 = 2$. Therefore, vector V is a linear combination of the two vectors V_1 and V_2.

Let e_1, e_2 and e_3 be three unit vectors in R^3 as sl



Then, the span of e_1 and e_2 is:

$$\text{span}(e_1, e_2) = \alpha_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ 0 \end{bmatrix}$$ =set of all vectors in $x-y$ plane. It is a subspace

of vector space R^3. Also, the span of e_1, e_2 and e_3 is:

$$\text{span}(e_1, e_2, e_3) = \alpha_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = R^3$$

Theorem If $V_1, V_2, \ldots, V_n$ are elements of a vector space V, then: span($V_1, V_2, \ldots, V_n$) is a subspace of V.

Proof: Let $V = \alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_n V_n$ be an arbitrary element of span($V_1, V_2, \ldots, V_n$), then for any scalar β.

$$\beta V = \beta(\alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_n V_n) = (\beta\alpha_1)V_1 + (\beta\alpha_2)V_2 + \cdots + (\beta\alpha_n)V_n$$

$$= \gamma_1 V_1 + \gamma_2 V_2 + \cdots + \lambda_n V_n$$

Therefore, $\beta V \in$ span($V_1, V_2, \ldots, V_n$). Next, let $V = \alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_n V_n$ and $W = \beta_1 V_1 + \beta_2 V_2 + \ldots + \beta_n V_n$, then

$$V + W = \alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_n V_n + \beta_1 V_1 + \beta_2 V_2 + \cdots + \beta_n V_n$$

$$= (\alpha_1 + \beta_1)V_1 + (\alpha_2 + \beta_2)V_2 + \cdots + (\alpha_n + \beta_n)V_n$$

$$= \gamma_1 V_1 + \gamma_2 V_2 + \cdots + \gamma_n V_n \in \operatorname{span}(V_1, V_2, \ldots, V_n)$$

Therefore, span($V_1, V_2, \ldots, V_n$) is a subspace of V.

Spanning Set of a Vector Space

Definition: The set $\{V_1, V_2, \ldots, V_n\}$ is a spanning set of vector space V if and only if every vector in vector space V can be written as a linear combination of $V_1, V_2, \ldots, V_n$.

It can be easily proved (see Problem 3.3) that if $V_1, V_2, \ldots, V_n$ spans the vector space V and one of these vectors can be written as a linear combination of the other $n - 1$ vectors, then these $n - 1$ vectors span the vector space V.

Linear Dependence

Definition: A set of vectors $V_1, V_2, \ldots, V_n$ is linearly dependent if there are scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ at least one of which is not zero, such that $\alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_n V_n = 0$. This means that a set of vectors is linearly dependent if one of them can be written as a linear combination of the others. A set of vectors that is not linearly dependent is linearly independent.

Example Consider the following two vectors $V_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $V_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ in R^2. Are V_1 and V_2 linearly independent?

Solution:

$$\alpha_1 V_1 + \alpha_2 V_2 = 0 \rightarrow$$

$$\alpha_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0 \rightarrow \begin{cases} \alpha_1 + \alpha_2 = 0 \\ \alpha_1 + 2\alpha_2 = 0 \end{cases} \rightarrow \begin{cases} \alpha_1 = 0 \\ \alpha_2 = 0 \end{cases}$$

Therefore, V_1 and V_2 are linearly independent.

Example

Are the following three vectors linearly independent?

$$V_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, V_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, V_3 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

Solution:

$$\alpha_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \alpha_3 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = 0 \rightarrow \begin{cases} \alpha_1 + \alpha_2 + 2\alpha_3 = 0 \\ 2\alpha_1 - \alpha_2 + \alpha_3 = 0 \\ 0\alpha_1 + \alpha_2 + \alpha_3 = \mathbf{0} \end{cases} \rightarrow \begin{cases} \alpha_1 = 1 \\ \alpha_2 = 1 \\ \alpha_3 = -1 \end{cases}$$

Since α_1, α_2, and $\alpha_3 \neq 0$, therefore V_1, V_2 and V_3 are not linearly independent. Theorem 3.2 provides the necessary and sufficient condition for linear independence of a set of vectors.

Theorem Let $X_1, X_2, \ldots, X_n$ be n vectors in R^n. Define the $n \times n$ matrix X as:

$$X = [X_1 \mid \quad X_2 \mid \quad \cdots \quad \mid X_n]$$

The vectors $X_1, X_2, \ldots, X_n$ are linearly dependent if and only if X is singular, that is:

$$\det(X) = 0$$

Example

Are the following three vectors linearly independent?

$$X_1 = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}, X_2 = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, X_3 = \begin{bmatrix} 2 \\ -5 \\ 3 \end{bmatrix}$$

Solution: Form the 3×3 matrix

$$X = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 3 & -5 \\ 3 & 1 & 3 \end{bmatrix}$$

The determinant of matrix X is:

$$\det(X) = 4\begin{vmatrix} 3 & -5 \\ 1 & 3 \end{vmatrix} - 2\begin{vmatrix} 2 & -5 \\ 3 & 3 \end{vmatrix} + 2\begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix} = 56 - 42 - 14 = 0$$

$\det(X) = 0$, therefore X_1, X_2 and X_3 are linearly dependent.

Example Linear Combinations

Consider the vectors $\mathbf{u} = (1, 2, -1)$ and $\mathbf{v} = (6, 4, 2)$ in R^3. Show that $\mathbf{w} = (9, 2, 7)$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$ and that $w' = (4, -1, 8)$ is *not* a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Solution In order for $\mathbf{w}$ to be a linear combination of $\mathbf{u}$ and $\mathbf{v}$, there must be scalars k_1 and k_2 such that $\mathbf{w} = k_1\mathbf{u} + k_2\mathbf{v}$; that is,

$$(9, 2, 7) = k_1(1, 2, -1) + k_2(6, 4, 2)$$

or

$$(9, 2, 7) = (k_1 + 6k_2, 2k_1 + 4k_2, -k_1 + 2k_2)$$

Equating corresponding components gives

$$\begin{aligned} k_1 + 6k_2 &= 9 \\ 2k_1 + 4k_2 &= 2 \\ -k_1 + 2k_2 &= 7 \end{aligned}$$

Solving this system using Gaussian elimination yields $k_1 = -3$, $k_2 = 2$, so

$$\mathbf{w} = -3\mathbf{u} + 2\mathbf{v}$$

Similarly, for $\mathbf{w}'$ to be a linear combination of $\mathbf{u}$ and $\mathbf{v}$, there must be scalars k_1 and k_2 such that $\mathbf{w}' = k_1\mathbf{u} + k_2\mathbf{v}$; that is,

$$(4, -1, 8) = k_1(1, 2, -1) + k_2(6, 4, 2)$$

or

$$(4, -1, 8) = (k_1 + 6k_2, 2k_1 + 4k_2, -k_1 + 2k_2)$$

Equating corresponding components gives

$$\begin{aligned} k_1 + 6k_2 &= 4 \\ 2k_1 + 4k_2 &= -1 \\ -k_1 + 2k_2 &= 8 \end{aligned}$$

This system of equations is inconsistent (verify), so no such scalars k_1 and k_2 exist. Consequently, $\mathbf{w}'$ is not a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Example The Standard Unit Vectors Span R^n

Recall that the standard unit vectors in R^n are

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0),\ \ \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots,\ \ \mathbf{e}_n = (0, 0, 0, \ldots 1)$$

These vectors span R^n since every vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in R^n can be expressed as

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n$$

which is a linear combination of $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Thus, for example, the vectors

$$\mathbf{i} = (1, 0, 0),\ \ \mathbf{j} = (0, 1, 0),\ \ \mathbf{k} = (0, 0, 1)$$

span R^3 since every vector $\mathbf{v} = (a, b, c)$ in this space can be expressed as

$$\mathbf{v} = (a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}\,.$$

Testing for Spanning

Determine whether $\mathbf{v}_1 = (1, 1, 2)$, $\mathbf{v}_2 = (1, 0, 1)$, and $v_3 = (2, 1, 3)$ span the vector space R^3.

Solution We must determine whether an arbitrary vector $\mathbf{b} = (b_1, b_2, b_3)$ in R^3 can be expressed as a linear combination

$$\mathbf{b} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3$$

of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. Expressing this equation in terms of components gives

$$(b_1, b_2, b_3) = k_1(1, 1, 2) + k_2(1, 0, 1) + k_3(2, 1, 3)$$

or

$$(b_1, b_2, b_3) = (k_1 + k_2 + 2k_3,\ k_1 + k_3,\ 2k_1 + k_2 + 3k_3)$$

or

$$\begin{aligned} k_1 + k_2 + 2k_3 &= b_1 \\ k_1 \qquad + k_3 &= b_2 \\ 2k_1 + k_2 + 3k_3 &= b_3 \end{aligned}$$

Thus, our problem reduces to ascertaining whether this system is consistent for all values of b_1, b_2, and b_3.

the system is consistent if and only if its coefficient matrix

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

has a nonzero determinant. But this is *not* the case here; we leave it for you to confirm that $\det(A) = 0$, so $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ do not span R^3.

Example

Consider the vector space $\mathbf{M} = \mathbf{M}_{2,2}$ consisting of all 2×2 matrices, and consider the following four matrices in $\mathbf{M}$:

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Then clearly any matrix A in $\mathbf{M}$ can be written as a linear combination of the four matrices. For example,

$$A = \begin{bmatrix} 5 & -6 \\ 7 & 8 \end{bmatrix} = 5E_{11} - 6E_{12} + 7E_{21} + 8E_{22}$$

Accordingly, the four matrices $E_{11}, E_{12}, E_{21}, E_{22}$ span $\mathbf{M}$.

Example

We claim that the following vectors form a spanning set of $\mathbf{R}^3$:

$$e_1 = (1,0,0), \qquad e_2 = (0,1,0), \qquad e_3 = (0,0,1)$$

Specifically, if $v = (a, b, c)$ is any vector in $\mathbf{R}^3$, then

$$v = ae_1 + be_2 + ce_3$$

For example, $v = (5, -6, 2) = 5e_1 - 6e_2 + 2e_3$.

Example

We claim that the following vectors also form a spanning set of $\mathbf{R}^3$:

$$w_1 = (1, 1, 1), \qquad w_2 = (1, 1, 0), \qquad w_3 = (1, 0, 0)$$

Specifically, if $v = (a, b, c)$ is any vector in $\mathbf{R}^3$, then

$$v = (a, b, c) = cw_1 + (b - c)w_2 + (a - b)w_3$$

For example, $v = (5, -6, 2) = 2w_1 - 8w_2 + 11w_3$.

DEFINITION

If $\mathbf{f}_1 = f_1(x), \mathbf{f}_2 = f_2(x), \ldots, \mathbf{f}_n = f_n(x)$ are functions that are $n-1$ times differentiable on the interval $(-\infty, \infty)$, then the determinant

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

is called the ***Wronskian*** of $f_1, f_2, \ldots, f_n$.

Example

Let $C^{n-1}[a, b]$ represent the set of all continuous real-valued differentiable functions defined on the closed interval $[a, b]$. It is assumed that these functions are $n-1$ times differentiable. Let vectors $f_1, f_2, f_3, \ldots, f_n \in C^{n-1}[a, b]$. If $f_1, f_2, f_3, \ldots, f_n$ are linearly independent, then:

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) = 0$$

has a trivial solution $c_1 = c_2 = \cdots = c_n = 0$. Differentiation of the above equation $n-1$ times yields:

$$c_1 f_1'(x) + c_2 f_2'(x) + \cdots + c_n f_n'(x) = 0$$

$$c_1 f_1''(x) + c_2 f_2''(x) + \cdots + c_n f_n''(x) = 0$$

$$\vdots$$

$$c_1 f_1^{n-1}(x) + c_2 f_2^{n-1}(x) + \cdots + c_n f_n^{n-1}(x) = 0$$

Writing these equations in matrix form, we have:

$$\begin{bmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{n-1}(x) & f_2^{n-1}(x) & \cdots & f_n^{n-1}(x) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Let us define:

$$W(f_1, f_2, \ldots, f_n) = \begin{bmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{n-1}(x) & f_2^{n-1}(x) & \cdots & f_n^{n-1}(x) \end{bmatrix} \text{and } c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

Then, $Wc = 0$ has a unique trivial solution $c = 0$ if the $n \times n$ matrix $W(f_1, f_2, \ldots, f_n)$ is nonsingular, that is:

$$\det[W(f_1, f_2, \ldots, f_n)] \neq 0$$

Theorem 3.3: Let $f_1, f_2, \ldots, f_n \in C^{n-1}[a, b]$. If there exists a point x_0 in $[a, b]$ such that $|W[f_1(x_0), f_2(x_0), \ldots, f_n(x_0)]| \neq 0$, then $f_1, f_2, \ldots, f_n$ are linearly independent.

Example

Let $f_1(x) = e^x$ and $f_2(x) = e^{-x}$. $f_1(x), f_2(x) \in C(-\infty, +\infty)$. Since:

$$\det(W[f_1, f_2]) = \begin{vmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{vmatrix} = -e^x e^{-x} - e^x e^{-x} = -2 \neq 0$$

Hence, $f_1(x), f_2(x)$ are linearly independent.

Example

Use the Wronskian to show that $\mathbf{f}_1 = x$ and $\mathbf{f}_2 = \sin x$ are linearly independent.

Solution The Wronskian is

$$W(x) = \begin{vmatrix} x & \sin x \\ 1 & \cos x \end{vmatrix} = x\cos x - \sin x$$

This function is not identically zero on the interval $(-\infty, \infty)$ since, for example,

$$W\left(\frac{\pi}{2}\right) = \frac{\pi}{2}\cos\left(\frac{\pi}{2}\right) - \sin\left(\frac{\pi}{2}\right) = \frac{\pi}{2}$$

Thus, the functions are linearly independent.

Example

Use the Wronskian to show that $\mathbf{f}_1 = 1$, $\mathbf{f}_2 = e^x$, and $\mathbf{f}_3 = e^{2x}$ are linearly independent.

Solution The Wronskian is

$$W(x) = \begin{vmatrix} 1 & e^x & e^{2x} \\ 0 & e^x & 2e^{2x} \\ 0 & e^x & 4e^{2x} \end{vmatrix} = 2e^{3x}$$

This function is obviously not identically zero on $(-\infty, \infty)$, so $\mathbf{f}_1$, $\mathbf{f}_2$, and $\mathbf{f}_3$ form a linearly independent set.

Example

Let $f_1(x) = 1$, $f_2(x) = x$, $f_3(x) = x^2$ and $f_4(x) = x^3$ be elements of a linear vector space consisting of polynomials of degree less than or equal to 3. Then:

$$W[f_1, f_2, f_3, f_4] = \begin{vmatrix} 1 & x & x^2 & x^3 \\ 0 & 1 & 2x & 3x^2 \\ 0 & 0 & 2 & 6x \\ 0 & 0 & 0 & 6 \end{vmatrix} = 12 \neq 0$$

Therefore, the set f_1, f_2, f_3, f_4 is linearly independent.

Basis Vectors

The minimal spanning set of a vector space forms the basis vector set for that space. A vector space V is called finite dimensional if it has a basis consisting of a finite number of vectors. The dimension of V, denoted by $\dim(V)$, is the number of vectors in a basis for V. A vector space that has no finite basis is called infinite dimensional. From the above, it is obvious that the vectors $V_1, V_2, \ldots, V_n$ form a basis for a finite dimensional vector space V if and only if:

i. $V_1, V_2, \ldots, V_n$ are linearly independent.
ii. $V_1, V_2, \ldots, V_n$ spans V.

Example

Given R^3, the standard basis set is $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

Since:

$$A = [e_1 \;\vdots\; e_2 \;\vdots\; e_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and}$$

$$\det(A) = 1$$

Therefore, the set of vectors e_1, e_2, e_3 are linearly independent.

Example

Consider the vector space V consisting of all 2×2 matrices. Show that the set $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ given by:

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

forms basis vectors for the vector space V.

Solution: E_{11}, E_{12}, E_{21}, and E_{22} are linearly independent since $c_1E_{11} + c_2E_{12} + c_3E_{21} + c_4E_{22} = 0$ implies that:

$$c_1\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + c_2\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + c_3\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} + c_4\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

$$\begin{bmatrix}c_1 & 0\\0 & 0\end{bmatrix} + \begin{bmatrix}0 & c_2\\0 & 0\end{bmatrix} + \begin{bmatrix}0 & 0\\c_3 & 0\end{bmatrix} + \begin{bmatrix}0 & 0\\0 & c_4\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

$$\begin{bmatrix}c_1 & c_2\\c_3 & c_4\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

E_{11}, E_{12}, E_{21}, E_{22} are linearly independent and form a basis for vector space V, i.e. if $A \in V$, A can be written as a linear combination of E_{11}, E_{12}, E_{21}, E_{22}. That is:

$$A = \begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix} = a_{11}\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + a_{12}\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + a_{21}\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} + a_{22}\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$$

$$= a_{11}E_{11} + a_{12}E_{12} + a_{21}E_{21} + a_{22}E_{22}$$

Change of Basis Vectors

Let V be an n dimensional linear vector space with basis $E = [v_1 \mid v_2 \mid \cdots \mid v_n]$. Assume that $v \in V$ is an element of vector space V, then v can be written as linear combination of $v_1, v_2, \ldots, v_n$. This implies that:

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = Ec$$

where $c = [c_1 \quad c_2 \quad \cdots \quad c_n]^T$. The set of scalars $c_1, c_2, \ldots, c_n$ are called coordinates of vector v with respect to the basis $v_1, v_2, \ldots, v_n$. Now assume that we would like to change the basis vectors to $F = [u_1 \mid u_2 \mid \cdots \mid u_n]$, then:

$$v = d_1 u_1 + d_2 u_2 + \cdots + d_n u_n = Fd$$

where $d = [d_1 \quad d_2 \quad \cdots \quad d_n]^T$.

Comparing Equations we have: $Fd = Ec$ $\qquad$ $d = F^{-1}Ec = Sc$

The matrix $S = F^{-1}E$ is called a transition matrix.

Example

Let $X = 3v_1 + 2v_2 - v_3$ with respect to basis vectors

$$E = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 5 \\ 1 & 2 & 4 \end{bmatrix}$$

Find the coordinates of vector X with respect to a new basis defined by:

$$F = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

Solution:

$$d = F^{-1}Ec = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 5 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -3 \\ -1 & -1 & 0 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 8 \\ -5 \\ 3 \end{bmatrix}$$

Therefore, vector X can be written as:

$$X = 8u_1 - 5u_2 + 3u_3.$$

Example

Let:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; u_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}; u_3 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

a. Find the transition matrix corresponding to the change of basis from (e_1, e_2, e_3) to (u_1, u_2, u_3).

b. Find the coordinate of $[3 \quad 2 \quad 5]^T$ with respect to the basis (u_1, u_2, u_3).

Solution:

a. $S = F^{-1}E = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -1 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$

b. The coordinates of vector $[3 \quad 2 \quad 5]^T$ with respect to the new basis are:

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

Normed Vector Spaces

A normed linear vector space is a vector space V together with a real-valued function on V (called the norm), such that:

a. $\|x\| \geq 0$, for every $x \in V$. $\|x\| = 0$ if and only if $x = 0$.
b. $\|\alpha x\| = |\alpha| \, \|x\|$, for every $x \in V$ and every scalar α.
c. $\|x + y\| \leq \|x\| + \|y\|$, for every x, $y \in V$. This is called triangular inequality.

The norm is a generalization of the concept of the length of a vector in 2-d or 3-d space. Given a vector x in V, the nonnegative number $\|x\|$ can be thought of as the length of vector x or the distance of x from the origin of the vector space.

Example

Consider the linear vector space R^n or C^n. Let us define the norm as:

$$\|x\|_p = [|x_1|^p + |x_2|^p + \cdots + |x_n|^p]^{\frac{1}{p}}$$

where p is an integer $1 \le p < \infty$. Different values of p result in different normed vector spaces. If $p = 1$, it is called L_1 norm, and when $p = 2$, it is called L_2 norm and they are given by:

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

and

$$\|x\|_2 = [|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2]^{\frac{1}{2}}$$

The MATLAB® command to compute the p norm of an n dimensional real or complex vector is **norm(*x*,*p*)**.

Example Compute L_1, L_2 and L_∞ norms of the following two vectors in R^3.

$$x = [1 \quad -2 \quad 3 \quad -4]^T,\; y = [3 \quad 0 \quad 5 \quad -2]^T.$$

Solution:

$$\|x\|_1 = |x_1| + |x_2| + |x_3| + |x_4| = 10$$

$$\|x\|_2 = [|x_1|^2 + |x_2|^2 + |x_3|^2 + |x_4|^2]^{\frac{1}{2}} = 5.4772$$

$$\|x\|_\infty = \max_i |x_i| = 4$$

$$\|y\|_1 = |y_1| + |y_2| + |y_3| + |y_4| = 10$$

$$\|y\|_2 = [|y_1|^2 + |y_2|^2 + |y_3|^2 + |y_4|^2]^{\frac{1}{2}} = 6.1644$$

$$\|y\|_\infty = \max_i |y_i| = 5$$

Example

Compute L_1, L_2 and L_∞ norms of complex vector $x \in C^3$.

$$x = [1 - j \quad j \quad 3 + j4]^T$$

Solution:

$$\|x\|_1 = |x_1| + |x_2| + |x_3| = \sqrt{2} + 1 + 5 = 6 + \sqrt{2} = 7.4142$$

$$\|x\|_2 = [|x_1|^2 + |x_2|^2 + |x_3|^2]^{\frac{1}{2}} = \sqrt{28} = 5.2915$$

$$\|x\|_\infty = \max_i |x_i| = 5$$

Distance Function

Given two vectors x and y in a normed vector space, the metric defined by:

$$d(x, y) = \|x - y\|_p$$

is called the p-norm distance between vector x and vector y. The distance function defined by the above equation satisfies the following properties:

a. $d(x, y) \geq 0$, for every x and $y \in V$. $d(x, y) = 0$ if and only if $x = y$.

b. $d(x, y) = d(y, x)$.

c. $d(x, z) \leq d(x, y) + d(y, z)$, for every x, y, and $z \in V$.

Example

Compute the 2-norm distance $d(x, y)$ between the following two vectors:

$$x = [1 - j \quad j \quad 3 + j4]^T \text{ and } y = [1 + j \quad 2 \quad 1 + j3]^T$$

Solution:

$$x - y = [-j2 \quad -2 + j \quad 2 + j]^T$$

$$d = \|x - y\|_2 = [|-j2|^2 + |-2 + j|^2 + |2 + j|^2]^{\frac{1}{2}} = \sqrt{4 + 5 + 5} = 3.7416$$

Equivalence of Norms

The L_1, L_2 and L_∞ norms on R^n (or C^n) are equivalent in the sense that if a converging sequence x_m is converging to $x_\infty = \lim_{x\to\infty} x_m$, as determined by one of the norms, then it converges to x_∞ as determined in all three norms. It can also be proven that the L_1, L_2 and L_∞ norms on R^n (or C^n) satisfy the following inequalities

$$\frac{\|x\|_2}{\sqrt{n}} \le \|x\|_\infty \le \|x\|_2$$

$$\|x\|_2 \le \|x\|_1 \le \sqrt{n}\,\|x\|_2$$

$$\frac{\|x\|_1}{n} \le \|x\|_\infty \le \|x\|_1$$

Example

Consider the linear vector space consisting of the set of all continuous functions defined over the closed interval [a b]. This vector spaced is a normed vector space with the L_p norm defined as:

$$\|f(x)\|_p = \left[\int_a^b |f(x)|^p dx \right]^{\frac{1}{p}}, \quad 1 \le p < \infty$$

Similar to R^n, we have:

$$\|f(x)\|_1 = \int_a^b |f(x)|\, dx$$

$$\|f(x)\|_2 = \left[\int_a^b |f(x)|^2\, dx \right]^{\frac{1}{2}}$$

and

$$\|f(x)\|_\infty = \sup_{a \le x \le b} |f(x)|$$

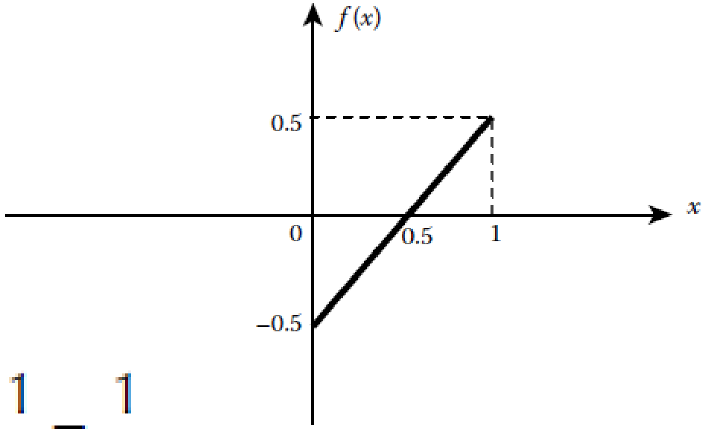
where "sup" stands for supremum, which is the generalization of the maximum

Definition of supremum: Given set S, a subset of real numbers, the smallest number a such that $a \geq x$ for every $x \in S$ is called the supremum (sup) of set S. Therefore:

$$a = \sup(S) \quad \text{if } a \geq x \text{ for every } x \in S$$

Example

Compute L_1, L_2 and L_∞ norm of the function $f(x) = x - 0.5$ defined over the closed interval [0 1]. The function $f(x)$ is shown in Figure



$$\|f(x)\|_1 = \int_a^b |f(x)|\,dx = \int_0^1 |x-0.5|\,dx = \int_0^{0.5} (-x+0.5)dx + \int_{0.5}^1 (x-0.5)\,dx = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

$$\|f(x)\|_2 = \left[\int_a^b |f(x)|^2\,dx\right]^{\frac{1}{2}} = \left[\int_0^1 (x-0.5)^2\,dx\right]^{\frac{1}{2}} = \frac{1}{\sqrt{12}}$$

$$\|f(x)\|_\infty = \sup_{a \le x \le b} |f(x)| = 0.5$$

Example

Consider the linear vector space consisting of the set of all continuous complex valued functions defined over the closed interval [0 T]. Compute L_1, L_2 and L_∞ norm of an element of this vector space given by:

$$f(t) = \sqrt{P}\exp\left(jk\frac{2\pi}{T}t\right)$$

where k is an integer.

Solution:

$$\|f(t)\|_1 = \int_a^b |f(t)|\,dt = \int_0^T \sqrt{P}dt = \sqrt{P}T$$

$$\|f(t)\|_2 = \left[\int_a^b |f(t)|^2\,dt\right]^{\frac{1}{2}} = \left[\int_0^T Pdt\right]^{\frac{1}{2}} = \sqrt{PT}$$

$$\|f(t)\|_\infty = \sup_{a\le t\le b} |f(t)| = \sqrt{P}$$

Inner Product Spaces

An inner product space is a linear vector space V with associated field R^1 or C^1, together with a function on $V \times V \to R^1$ (C^1) called the inner product (or dot product):

$$< x, y >: V \times V \to R^1 (C^1)$$

Such that:

a. $< x + y, z > = < x, z > + < y, z >$

b. $< x, y > = \overline{< x, y >}$ (bar stands for complex conjugate)

c. $< \alpha x, y > = \alpha < x, y >$

d. $< x, x > \geq 0$ ($=0$ if and only if $x = 0$)

further properties

Notice that, as a result of the above properties, $<x, y>$ has further properties such as:

$$< x, y+z > = < x, y > + < x, z >$$

$$< x, \alpha y > = \overline{\alpha} < x, y >$$

Example

Find the dot product between $x=[1-j \quad 2-j \quad 3]^T$ and $y=[1 \quad 2j \quad 4]^T$.

Solution:

$$< x, y > = \sum_{i=1}^{n} x_i \overline{y}_i = (1-j)\times 1 + (2-j)\times(-j2) + 3\times 4 = 11 - j5$$

EXAMPLE

Let V be a real inner product space. Then, by linearity,

$$\begin{aligned}\langle 3u_1 - 4u_2,\ 2v_1 - 5v_2 + 6v_3\rangle &= 6\langle u_1, v_1\rangle - 15\langle u_1, v_2\rangle + 18\langle u_1, v_3\rangle \\ &\quad - 8\langle u_2, v_1\rangle + 20\langle u_2, v_2\rangle - 24\langle u_2, v_3\rangle\end{aligned}$$

$$\begin{aligned}\langle 2u - 5v,\ 4u + 6v\rangle &= 8\langle u, u\rangle + 12\langle u, v\rangle - 20\langle v, u\rangle - 30\langle v, v\rangle \\ &= 8\langle u, u\rangle - 8\langle v, u\rangle - 30\langle v, v\rangle\end{aligned}$$

Observe that in the last equation we have used the symmetry property that $\langle u, v\rangle = \langle v\ u\rangle$

Norm of a Vector

$$\|u\| = \sqrt{\langle u, u\rangle}$$

EXAMPLE

Let $u = (1, 3, -4, 2)$, $\upsilon = (4, -2, 2, 1)$, $w = (5, -1, -2, 6)$ in $\mathbf{R}^4$.

Show $\langle 3u - 2\upsilon, w\rangle = 3\langle u, w\rangle - 2\langle \upsilon, w\rangle$

By definition,

$\langle u, w\rangle = 5 - 3 + 8 + 12 = 22$ and $\langle \upsilon, w\rangle = 20 + 2 - 4 + 6 = 24$

Note that $3u - 2\upsilon = (-5, 13, -16, 4)$. Thus,

$\langle 3u - 2\upsilon, w\rangle = -25 - 13 + 32 + 24 = 18$

As expected, $3\langle u, w\rangle - 2\langle \upsilon, w\rangle = 3(22) - 2(24) = 18 = \langle 3u - 2\upsilon, w\rangle$

Euclidean n-Space $\mathbf{R}^n$

Consider the vector space $\mathbf{R}^n$. The *dot product* or *scalar product* in $\mathbf{R}^n$ is defined by

$$u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

where $u = (a_i)$ and $v = (b_i)$. This function defines an inner product on $\mathbf{R}^n$. The norm $\|u\|$ of the vector $u = (a_i)$ in this space is as follows:

$$\|u\| = \sqrt{u \cdot u} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

EXAMPLE

Let $u = (1, 3, -4, 2)$, $\upsilon = (4, -2, 2, 1)$, $w = (5, -1, -2, 6)$ in $\mathbf{R}^4$.

Normalize u and υ:

By definition,

$$\|u\| = \sqrt{1+9+16+4} = \sqrt{30} \qquad \text{and} \qquad \|v\| = \sqrt{16+4+4+1} = 5$$

We normalize u and υ to obtain the following unit vectors in the directions of u and υ, respectively:

$$\hat{u} = \frac{1}{\|u\|}u = \left(\frac{1}{\sqrt{30}}, \frac{3}{\sqrt{30}}, \frac{-4}{\sqrt{30}}, \frac{2}{\sqrt{30}}\right) \qquad \text{and} \qquad \hat{v} = \frac{1}{\|v\|}v = \left(\frac{4}{5}, \frac{-2}{5}, \frac{2}{5}, \frac{1}{5}\right)$$

Function Space C[a, b] and Polynomial Space P(t)

The notation $C\,[a, b]$ is used to denote the vector space of all continuous functions on the closed interval $[a, b]$—that is, where $a \le t \le b$. The following defines an inner product on $C\,[a, b]$, where $f\,(t)$ and $g\,(t)$ are functions in $C\,[a, b]$:

$$\langle f, g \rangle = \int_a^b f(t)g(t)\,dt$$

It is called the *usual inner product* on $C\,[a, b]$.

EXAMPLE

Consider $f(t) = 3t - 5$ and $g(t) = t^2$ in the polynomial space **P** (t) with inner product

$$\langle f, g \rangle = \int_0^1 f(t)g(t)\, dt.$$

Find $\langle f, g \rangle$.

We have $f(t)g(t) = 3t^3 - 5t^2$. Hence,

$$\langle f, g \rangle = \int_0^1 (3t^3 - 5t^2)\, dt = \tfrac{3}{4}t^4 - \tfrac{5}{3}t^3 \Big|_0^1 = \frac{3}{4} - \frac{5}{3} = -\frac{11}{12}$$

EXAMPLE

Consider $f(t) = 3t - 5$ and $g(t) = t^2$ in the polynomial space $\mathbf{P}(t)$ with inner product

$$\langle f, g \rangle = \int_0^1 f(t)g(t)\,dt.$$

Find $\|f\|$ and $\|g\|$.

We have $[f(t)]^2 = f(t)f(t) = 9t^2 - 30t + 25$ and $[g(t)]^2 = t^4$. Then

$$\|f\|^2 = \langle f, f \rangle = \int_0^1 (9t^2 - 30t + 25)\,dt = 3t^3 - 15t^2 + 25t\Big|_0^1 = 13$$

$$\|g\|^2 = \langle g, g \rangle = \int_0^1 t^4\,dt = \tfrac{1}{5}t^5\Big|_0^1 = \tfrac{1}{5}$$

Therefore, $\|f\| = \sqrt{13}$ and $\|g\| = \sqrt{\tfrac{1}{5}} = \tfrac{1}{5}\sqrt{5}.$

Orthogonality

Definition: Two vectors x and y in an inner product space are said to be orthogonal if the inner product between x and y is zero. This means that:

$$< x, y > = 0$$

A set S in an inner product vector space is called an orthonormal set if:

$$< x, y >= \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases} \quad \text{for every pair } x, y \in S$$

EXAMPLE

The set $S = \{e_1 \; e_2 \; e_3\} \in R^3$ where $e_1 = [1 \quad 0 \quad 0]^T$, $e_2 = [0 \quad 1 \quad 0]^T$, and $e_3 = [0 \quad 0 \quad 1]^T$ is an orthonormal set since:

$$< e_i , e_j > = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

EXAMPLE

a. Consider the vectors $u = (1, 1, 1)$, $\upsilon = (1, 2, -3)$, $w = (1, -4, 3)$ in $\mathbf{R}^3$. Then

$$\langle u, v\rangle = 1 + 2 - 3 = 0, \qquad \langle u, w\rangle = 1 - 4 + 3 = 0, \qquad \langle v, w\rangle = 1 - 8 - 9 = -16$$

Thus, u is orthogonal to υ and w, but υ and w are not orthogonal.

b. Consider the functions $\sin t$ and $\cos t$ in the vector space $C[-\pi, \pi]$ of continuous functions on the closed interval $[-\pi, \pi]$. Then

$$\langle \sin t, \cos t\rangle = \int_{-\pi}^{\pi} \sin t \cos t \, dt = \tfrac{1}{2} \sin^2 t\Big|_{-\pi}^{\pi} = 0 - 0 = 0$$

Thus, $\sin t$ and $\cos t$ are orthogonal functions in the vector space $C[-\pi, \pi]$.

Remark: A vector $w = (x_1, x_2; \ldots, x_n)$ is orthogonal to $u = (a_1, a_2, \ldots, a_n)$ in R^n if

$$\langle u, w\rangle = a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

That is, w is orthogonal to u if w satisfies a homogeneous equation whose coefficients are the elements of u.

EXAMPLE

Find a nonzero vector w that is orthogonal to $u_1 = (1, 2, 1)$ and $u_2 = (2, 5, 4)$ in R^3.

Let $w = (x, y, z)$. Then we want $\langle u_1, w \rangle = 0$ and $\langle u_2, w \rangle = 0$. This yields the homogeneous system

$$\begin{aligned} x + 2y + z &= 0 \\ 2x + 5y + 4z &= 0 \end{aligned}$$

Here z is the only free variable in the echelon system. Set $z = 1$ to obtain $y = -2$ and $x = 3$. Thus, $w = (3, -2, 1)$ is a desired nonzero vector orthogonal to u_1 and u_2.

Any multiple of w will also be orthogonal to u_1 and u_2. Normalizing w, we obtain the following unit vector orthogonal to u_1 and u_2:

$$\hat{w} = \frac{w}{\|w\|} = \left(\frac{3}{\sqrt{14}}, -\frac{2}{\sqrt{14}}, \frac{1}{\sqrt{14}}\right)$$

Example Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set in $\mathbb{R}^3$ if

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Solution We must show that every pair of vectors from this set is orthogonal. This is true, since

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_2 &= 2(0) + 1(1) + (-1)(1) = 0 \\ \mathbf{v}_2 \cdot \mathbf{v}_3 &= 0(1) + 1(-1) + (1)(1) = 0 \\ \mathbf{v}_1 \cdot \mathbf{v}_3 &= 2(1) + 1(-1) + (-1)(1) = 0 \end{aligned}$$

Theorem If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then these vectors are linearly independent.

Proof If $c_1, \ldots, c_k$ are scalars such that $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$, then

$$(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) \cdot \mathbf{v}_i = \mathbf{0} \cdot \mathbf{v}_i = 0$$

or, equivalently,

$$c_1(\mathbf{v}_1 \cdot \mathbf{v}_i) + \cdots + c_i(\mathbf{v}_i \cdot \mathbf{v}_i) + \cdots + c_k(\mathbf{v}_k \cdot \mathbf{v}_i) = 0 \qquad (1)$$

Since $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is an orthogonal set, all of the dot products in Equation (1) are zero, except $\mathbf{v}_i \cdot \mathbf{v}_i$. Thus, Equation (1) reduces to

$$c_i(\mathbf{v}_i \cdot \mathbf{v}_i) = 0$$

Now, $\mathbf{v}_i \cdot \mathbf{v}_i \neq 0$ because $\mathbf{v}_i \neq \mathbf{0}$ by hypothesis. So we must have $c_i = 0$. The fact that this is true for all $i = 1, \ldots, k$ implies that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a linearly independent set.

Orthogonal Complements

Let S be a subset of an inner product space V. The orthogonal complement of S, denoted by $S^{\perp}$ (read "S perp") consists of those vectors in V that are orthogonal to every vector $u \in S$, that is,

$$S^{\perp} = \{v \in V : \langle v, u \rangle = 0 \text{ for every } u \in S\}$$

In particular, for a given vector u in V, we have

$$u^{\perp} = \{v \in V : \langle v, u \rangle = 0\}$$

that is, $u^{\perp}$ consists of all vectors in V that are orthogonal to the given vector u.

Example

The set $S = \{e_1\ e_2\ e_3\} \in R^3$ where $e_1 = [1 \quad 0 \quad 0]^T$, $e_2 = [0 \quad 1 \quad 0]^T$, and $e_3 = [0 \quad 0 \quad 1]^T$ is an orthonormal set since:

$$< e_i, e_j > = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Gram–Schmidt Orthogonalization Process

The Gram–Schmidt orthogonalization process is used to convert a set of independent or dependent vectors in a given vector space to an orthonormal set. Given a set of n nonzero vectors $S = \{x_1\ x_2\ \ldots\ x_n\}$ in vector space V, we would like to find an orthonormal set of vectors $\hat{S} = \{u_1\ u_2\ \ldots\ u_m\}$ with the same span as S. It is obvious that $m \leq n$. If the vectors forming set S are linearly independent, then $m = n$. This is accomplished by using the Gram–Schmidt orthogonalization process. The steps of the process as explained below:

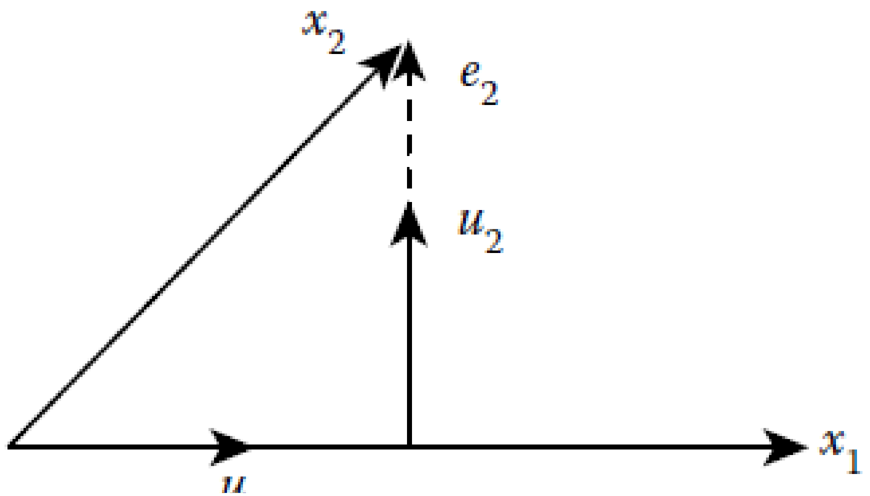
1. Determine u_1 by normalizing the first vector x_1 that is:

$$u_1 = \frac{x_1}{\|x_1\|}$$

2. Form the error vector e_2 by finding the difference between the projection of x_2 onto u_1 and x_2. If $e_2 = 0$, discard x_2; otherwise, obtain u_2 by normalizing e_2

$$e_2 = x_2 - \langle x_2, u_1 \rangle u_1$$

$$u_2 = \frac{e_2}{\| e_2 \|}$$



It is easy to see that u_1 and u_2 are orthogonal and each have unit norm.

3. Now, we need to find a unit norm vector u_3 that is orthogonal to both u_1 and u_2. This is done by finding the difference between x_3 and the projection of x_3 onto the subspace formed by the span of u_1 and u_2. This difference vector is normalized to produce u_3.

$$e_3 = x_3 - < x_3 , u_1 > u_1 - < x_3 , u_2 > u_2$$

Thus:

$$u_3 = \frac{e_3}{\| e_3 \|}$$

4. This is continued until the vector x_n has been processed, i.e. $k = 4, 5, \ldots, n$. At the k^{th} stage, vector e_k is formed as:

$$e_k = x_k - \sum_{i=1}^{k-1} < x_k , u_i > u_i$$

$$u_k = \frac{e_k}{\| e_k \|}$$

Example

Consider the set of polynomials $S = \{1 \;\; t \;\; t^2\}$ defined over the interval of $-1 \le t \le 1$. Using the Gram–Schmidt orthogonalization process, obtain an orthonormal set.

Solution:

$$u_1(t) = \frac{x_1(t)}{\|x_1(t)\|} = \frac{1}{\sqrt{\int_{-1}^{1} 1^2 dt}} = \frac{1}{\sqrt{2}}$$

$$e_2(t) = x_2(t) - < x_2(t), u_1(t) > u_1(t) = t - \frac{1}{\sqrt{2}} \int_{-1}^{1} t \frac{1}{\sqrt{2}} dt = t - \frac{1}{2} \int_{-1}^{1} t dt = t - 0 = t$$

$$u_2(t) = \frac{e_2(t)}{\|e_2(t)\|} = \frac{t}{\sqrt{\int_{-1}^{1} t^2 dt}} = \frac{t}{\sqrt{\frac{2}{3}}} = \sqrt{\frac{3}{2}}\, t$$

$$e_3(t) = x_3(t) - \langle x_3(t), u_1(t) \rangle u_1(t) - \langle x_3(t), u_2(t) \rangle u_2(t)$$

$$= t^2 - \frac{1}{\sqrt{2}}\int_{-1}^{1} t^2 \frac{1}{\sqrt{2}} dt - \sqrt{\frac{3}{2}} t \int_{-1}^{1} t^2 \sqrt{\frac{3}{2}} t\, dt = t^2 - \frac{1}{\mathbf{2}}\int_{-1}^{1} t^2 dt - \frac{3}{2} t \int_{-1}^{1} t^3 dt = t^2 - \frac{1}{3}$$

$$u_3(t) = \frac{e_3(t)}{\|e_3(t)\|} = \frac{t^2 - \frac{1}{3}}{\sqrt{\int_{-1}^{1} (t^2 - \frac{1}{3})^2 dt}} = \frac{t^2 - \frac{1}{3}}{\sqrt{\frac{8}{45}}} = \frac{1}{2}\sqrt{\frac{5}{2}}(3t^2 - 1)$$

Example

Consider the set $S = \{x_1\ x_2\ x_3\ x_4\} \in R^3$ where:

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, x_2 = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix}, x_4 = \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}$$

Use the Gram–Schmidt orthogonalization process to obtain an orthonormal set.

Solution:

$$u_1 = \frac{x_1}{\|x_1\|} = \frac{x_1}{3} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}^T = [0.3333 \quad 0.6667 \quad -0.6667]^T$$

$$e_2 = x_2 - < x_2, u_1 > u_1 = x_2 - u_1 = \begin{bmatrix} -\frac{4}{3} & \frac{7}{3} & \frac{5}{3} \end{bmatrix}^T$$

$$u_2 = \frac{e_2}{\|e_2\|} = \frac{e_2}{\sqrt{10}} = \left[-\frac{4}{3\sqrt{10}} \quad \frac{7}{3\sqrt{10}} \quad \frac{5}{3\sqrt{10}}\right]^T = [-0.4216 \quad 0.7379 \quad 0.527]^T$$

$$e_3 = x_3 - < x_3, u_1 > u_1 - < x_3, u_2 > u_2$$

$$= x_3 + \frac{13}{3} u_1 - \frac{7}{3\sqrt{10}} u_2 = [2.7556 \quad 0.3444 \quad 1.7222]^T$$

$$u_3 = \frac{e_3}{\|e_3\|} = \frac{e_3}{3.2677} = [0.8433 \quad 0.1054 \quad 0.527]^T$$

$$e_4 = x_4 - < x_4, u_1 > u_1 - < x_4, u_2 > u_2 - < x_4, u_3 > u_3$$

$$= x_4 - 3.333\, u_1 - 0.527\, u_2 - 3.6893\, u_3 = 0$$

Since $e_4=0$, we discard x_4. This means that the dimension of the space spanned by S is three as expected, since S is a subspace of R^3 and its dimension cannot exceed three.

MATLAB Code Implementation of Gram–Schmidt Orthogonalization Algorithm

```
function [V]=gramschmidt(A)
% A=a matrix of size mxn containing n vectors.
% The dimension of each vector is m.
% V=output matrix: Columns of V form an orthonormal set.
[m,n]=size(A);
V=[A(:,1)/norm(A(:,1))];
for j=2:1:n
 v=A(:,j);
 for i=1:size(V,2)
  a=v'*V(:,i);
  v=v-a*V(:,i);
 end
 if (norm(v))^4>=eps
  V=[V v/norm(v)];
 else
 end
end
```

Null Space

Let A be an $m \times n$ matrix. The set of vectors in R^n for which $Ax = 0$ is called the null space of A. It is denoted by $N(A)$. Therefore:

$$N(A) = \{x \in R^n \quad \text{such that} \quad Ax = 0\}$$

The null space of A is a subspace of the vector space R^n.

Example

Find the null space of the 2×3 matrix A

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

Solution: The null space of A is obtained by solving the system $Ax = 0$. We first start by generating the augmented matrix A_a and then simplifying it by utilizing the Gauss–Jordan technique. This is shown as:

$$Ax = \begin{bmatrix} 1 & 3 & 2 \\ 2 & -3 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 3x_2 + 2x_3 \\ 2x_1 - 3x_2 - 5x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$A_a = \begin{bmatrix} 1 & 3 & 2 & 0 \\ 2 & -3 & -5 & 0 \end{bmatrix} \xrightarrow{R_2 \leftarrow R_2 - 2R_1} \begin{bmatrix} 1 & 3 & 2 & 0 \\ 0 & -9 & -9 & 0 \end{bmatrix} \xrightarrow{R_2 \leftarrow R_2 / -9} \begin{bmatrix} 1 & 3 & 2 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$x_2 + x_3 = 0 \Rightarrow x_2 = -x_3$$

$$1x_1 + 3x_2 + 2x_3 = 0 \Rightarrow x_1 = -3x_2 - 2x_3 = -3(-x_3) - 2x_3 = x_3$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Therefore, the null space of A is the set of vectors in R^3 given by $x = [\alpha \quad -\alpha \quad \alpha]^T$.

Example

Find the null space of the 3×3 matrix A

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

Solution: The null space of A is obtained by solving the equation $Ax=0$. Since, in this case, $\det(A)=1\neq 0$, matrix A is nonsingular and $Ax=0$ has a unique solution $x=0$ since $x=A^{-1}0=0$. Therefore, the null space of matrix A is empty. That is $N(A)=\{\phi\}$.

Column Space C(A)

The space spanned by the columns of the $m \times n$ matrix A is called the column space of A. It is also called the range of matrix A and is denoted by $C(A)$. Therefore:

$$C(A) = \{y \in R^m \quad \text{such that} \quad y = Ax \text{ for } x \in R^n\}$$

Given that A is an $m \times n$ matrix, the range of A is a subspace of R^m. If $y_1 \in C(A)$, then y_1 is a linear combinations of the independent columns of A.

Example Find the range of the 2×3 matrix A

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

Solution: The range of A is the subspace spanned by the columns of A and is obtained by:

$$y=Ax=\begin{bmatrix}2 & -3 & 4\\ 1 & 3 & -2\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix}=\begin{bmatrix}2a-3b+4c\\ a+3b-2c\end{bmatrix}$$

$$=(2a-3b+4c)\begin{bmatrix}1\\ 0\end{bmatrix}+(a+3b-2c)\begin{bmatrix}0\\ 1\end{bmatrix} = \alpha\begin{bmatrix}1\\ 0\end{bmatrix}+\beta\begin{bmatrix}0\\ 1\end{bmatrix}$$

Hence, $C(A)=R^2$, which means that any arbitrarily vector in R^2 is in the range space of matrix A.

Row Space $R(A)$

The row space of an $m \times n$ matrix A is the column space (range) of A^T. The dimension of the column space of A represents the rank of matrix A and is less than or equal to the minimum of n and m. Similarly, the row space of matrix A is the linear combination of the rows of A. The dimension of the row space of A is also equal to the rank of matrix A.

Example Find the row space of the 2×3 matrix A

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

Solution: The row space of A is the subspace spanned by the columns of A^T and is obtained by:

$$y=A^Tx=\begin{bmatrix}1 & 1\\ -2 & 3\\ 3 & 4\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}=\begin{bmatrix}a+b\\ -2a+3b\\ 3a+4b\end{bmatrix}=a\begin{bmatrix}1\\ -2\\ 3\end{bmatrix}+b\begin{bmatrix}1\\ 3\\ 4\end{bmatrix}$$

Therefore, the row space of A is:

$$R(A)=\left\{\begin{bmatrix}1\\ -2\\ 3\end{bmatrix},\quad\begin{bmatrix}1\\ 3\\ 4\end{bmatrix}\right\}$$

Rank of a Matrix

The dimension of the column space (or the row space) of the matrix A is called the rank of A and is denoted by r. Thus:

$$r = \dim(R(A)) = \dim(R(A^T))$$

The rank of a matrix represents the number of independent rows or columns in the matrix. Therefore, for an $m \times n$ matrix:

$$r \leq \min(m, n)$$

Definition: An $m \times n$ matrix is said to be full rank if

$$r = \min(m, n)$$

A matrix which is not full rank is said to be rank deficient.