1. If $\mathbf{u}, \mathbf{v} \in V$, then $\mathbf{u} + \mathbf{v} \in V$.
2. If $\mathbf{u} \in V$ and $k \in \mathbb{F}$, then $k\mathbf{u} \in V$.
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
4. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
5. There is an object $\mathbf{0}$ in V, called a **zero vector** for $V$, such that $\mathbf{u}+\mathbf{0} = \mathbf{0}+\mathbf{u} = \mathbf{u}$ for all $\mathbf{u}$ in $V$.
6. For each $\mathbf{u}$ in $V$, there is an object **-u** in $V$, called the **additive inverse** of $\mathbf{u}$, such that $\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{0}$;
7. $k(l\mathbf{u}) = (kl)\mathbf{u}$
8. $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$
9. $(k + l)\mathbf{u} = k\mathbf{u} + l\mathbf{u}$
10. $1\mathbf{u} = \mathbf{u}$



**Remark** The elements of the underlying field $\mathbb{F}$ are called scalars and the elements of the vector space are called vectors. Note also that we often restrict our attention to the case when $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

**Examples of Vector Spaces**

A wide variety of vector spaces are possible under the above definition as illustrated by the following examples. In each example we specify a nonempty set of objects $V$. We must then define two operations - addition and scalar multiplication, and as an exercise we will demonstrate that all the axioms are satisfied, hence entitling $V$ with the specified operations, to be called a vector space.

1. The set of all $n$-tuples with entries in the field $\mathbb{F}$, denoted $\mathbb{F}^n$ (especially note $\mathbb{R}^n$ and $\mathbb{C}^n$).

## 1.4 Linear Combinations of Vectors and Systems of Linear Equations

Have $m$ linear equations in $n$ variables:

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

Write in matrix form: $A\boldsymbol{x} = \boldsymbol{b}$.

$A = [a_{ij}]$ is the $m \times n$ coefficient matrix.

$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is the column vector of unknowns, and $\boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ is the column vector of RHS.



Note: $a_{ij}$, $b_j \in \mathbb{R}$ or $\mathbb{C}$.

### 1.4.1 Gaussian Elimination

To solve $A\boldsymbol{x} = \boldsymbol{b}$:

write *augmented matrix*: $[A|\boldsymbol{b}]$.

**1.** Find the left-most non-zero column, say column $j$.

**2.** Interchange top row with another row if necessary, so top element of column $j$ is non-zero. (The **pivot**.)

**3.** Subtract multiples of row 1 from all other rows so all entries in column $j$ below the top are then 0.

**4.** Cover top row; repeat **1** above on rest of rows.

Continue until all rows are covered, or until only $00\ldots0$ rows remain.

## 1.5 Linear Independence

In the previous section it was stated that a set of vectors $S$ spans a given vector space $V$ if every vector in $V$ is expressible as a linear combination of the vectors in $S$. In general, it is possible that there may be more than one way to express a vector in $V$ as a linear combination of vectors in a spanning set. This section will focus on the conditions under which each vector in $V$ is expressible as a unique linear combination of the spanning vectors. Spanning sets with this property play a fundamental role in the study of vector spaces.

**Definitions** If $S = \{v_1, v_2, \ldots, v_r\}$ is a nonempty set of vectors, then the vector equation

$$k_1\mathbf{v_1} + k_2\mathbf{v_2} + \cdots + k_r\mathbf{v_r} = \mathbf{0}$$

has at least one solution, namely

$$k_1 = 0, k_2 = 0, \ldots, k_r = 0$$



If this is the only solution, then $S$ is called a **linearly independent** set. If there are other solutions, then $S$ is called a **linearly dependent** set.

Examples

1. If $\mathbf{v_1} = (2, -1, 0, 3), v_2 = (1, 2, 5, -1)$ and $v_3 = (7, -1, 5, 8)$, then the set of vectors $S = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ is linearly dependent, since $3\mathbf{v_1} + \mathbf{v_2} - \mathbf{v_3} = \mathbf{0}$.

2. The polynomials

$$\mathbf{p_1} = 1 - x,\ \mathbf{p_2} = 5 + 3x - 2x^2,\ \mathbf{p_3} = 1 + 3x - x^2$$

   form a linearly dependent set in $P_2$ since $3\mathbf{p_1} - \mathbf{p_2} + 2\mathbf{p_3} = \mathbf{0}$

3. Consider the vectors $\mathbf{i} = (1, 0, 0), \mathbf{j} = (0, 1, 0), \mathbf{k} = (0, 0, 1)$ in $\mathbb{R}^3$. In terms of components the vector equation

$$k_1\mathbf{i} + k_2\mathbf{j} + k_3\mathbf{k} = \mathbf{0}$$

becomes

$$k_1(1,0,0) + k_2(0,1,0) + k_3(0,0,1) = (0,0,0)$$

or equivalently,

$$(k_1, k_2, k_3) = (0,0,0)$$

Thus the set $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is linearly independent. A similar argument can be used to extend S to a linear independent set in $\mathbb{R}^n$.

4. In $M_{2\times 3}(\mathbb{R})$, the set

$$\left\{ \begin{pmatrix} 1 & -3 & 2 \\ -4 & 0 & 5 \end{pmatrix}, \begin{pmatrix} -3 & 7 & 4 \\ 6 & -2 & -7 \end{pmatrix}, \begin{pmatrix} -2 & 3 & 11 \\ -1 & -3 & 2 \end{pmatrix} \right\}$$

is linearly dependent since

$$5\begin{pmatrix} 1 & -3 & 2 \\ -4 & 0 & 5 \end{pmatrix} + 3\begin{pmatrix} -3 & 7 & 4 \\ 6 & -2 & -7 \end{pmatrix} - 2\begin{pmatrix} -2 & 3 & 11 \\ -1 & -3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$



The following two theorems follow quite simply from the definition of linear independence and linear dependence.

**Theorem 1.8.** *A set $S$ with two or more vectors is:*

*(a) Linearly dependent if and only if at least one of the vectors in $S$ is expressible as a linear combination of the other vectors in $S$.*

*(b) Linearly independent if and only if no vector in $S$ is expressible as a linear combination of the other vectors in $S$.*

**Example**

1. Recall that the vectors

$$\mathbf{v_1} = (2,-1,0,3),\ \mathbf{v_2} = (1,2,5,-1),\ \mathbf{v_3} = (7,-1,5,8)$$

**Example** Let $S = \{\mathbf{i}, \mathbf{j}\}$, $U = \{\mathbf{i}, 2\mathbf{j}\}$ and $V = \{\mathbf{i}+\mathbf{j}, \mathbf{j}\}$. Let the sets $S$,$U$ and $V$ be three sets of basis vectors. Let $P$ be the point $\mathbf{i} + 2\mathbf{j}$. The coordinates of $P$ relative to each set of basis vectors is:

$$S \rightarrow (1,2)$$

$$U \rightarrow (1,1)$$

$$T \rightarrow (1,1)$$

The following definition makes the preceding ideas more precise and enables the extension of a coordinate system to general vector spaces.

**Definition**

- If $V$ is any vector space and $S = \{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ is a set of vectors in $V$, then $S$ is called a **basis** for $V$ if the following two conditions hold:

  (a) $S$ is linearly independent

  (b) $S$ spans $V$.



A basis is the vector space generalization of a coordinate system in 2-space and 3-space. The following theorem will aid in understanding how this is so.

**Theorem 2.1.** *If $S = \{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ is a basis for a vector space $V$, then every vector $\mathbf{v}$ in $V$ can be expressed in the form $\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n}$ in exactly one way.*

*Proof.* Since $S$ spans $V$, it follows from the definition of a spanning set that every vector in $V$ is expressible as a linear combination of the vectors in $S$. To see that there is only one way to express a vector as a linear combination of the vectors in $S$, suppose that some vector $\mathbf{v}$ can be written as

$$\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n}$$

and also as

$$\mathbf{v} = k_1\mathbf{v_1} + k_2\mathbf{v_2} + \cdots + k_n\mathbf{v_n}$$

*(a) Every set with more than $n$ vectors is linearly dependent.*

*(b) No set with fewer than $n$ vectors spans $V$.*

*Proof.* (a) Let $S' = \{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}\}$ be any set of $m$ vectors in $V$, where $m > n$. It remains to be shown that $S'$ is linearly dependent. Since $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a basis for $V$, each $\mathbf{w_i}$ can be expressed as a linear combination of the vectors in $S$, say:

$$\mathbf{w_1} = a_{11}\mathbf{v_1} + a_{21}\mathbf{v_2} + \cdots + a_{n1}\mathbf{v_n}$$

$$\mathbf{w_2} = a_{12}\mathbf{v_1} + a_{22}\mathbf{v_2} + \cdots + a_{n2}\mathbf{v_n}$$

$$\vdots$$

$$\mathbf{w_m} = a_{1m}\mathbf{v_1} + a_{2m}\mathbf{v_2} + \cdots + a_{nm}\mathbf{v_n}$$

To show that $S'$ is linearly dependent, scalars $k_1, k_2, \ldots, k_m$ must be found, not all zero, such that

$$k_1\mathbf{w_1} + k_2\mathbf{w_2} + \cdots + k_m\mathbf{w_m} = \mathbf{0}$$



combining the above 2 systems of equations gives

$$\begin{aligned}(k_1a_{11} + k_2a_{12} + \cdots + k_ma_{1m})&\mathbf{v_1}\\ + (k_1a_{21} + k_2a_{22} + \cdots + k_ma_{2m})&\mathbf{v_2}\\ \ddots&\\ + (k_1a_{n1} + k_2a_{n2} + \cdots + k_ma_{nm})&\mathbf{v_n} = \mathbf{0}\end{aligned}$$

Thus, from the linear independence of $S$, the problem of proving that $S'$ is a linearly dependent set reduces to showing there are scalars $k_1, k_2, \ldots, k_m$, not all zero, that satisfy

$$a_{11}k_1 + a_{12}k_2 + \cdots + a_{1m}k_m = 0$$

$$a_{21}k_1 + a_{22}k_2 + \cdots + a_{2m}k_m = 0$$

$$\vdots$$

**Examples**

1. Let $A$ be an $m \times n$ matrix, and $B$ be an $n \times p$ matrix, then $AB$ is an $m \times p$ matrix. Also $T_A : \mathbb{F}^n \to \mathbb{F}^m$, and $T_B : \mathbb{F}^p \to \mathbb{F}^n$ are both linear transformations. Then

$$\begin{aligned} T_A \circ T_B &= T_A(T_B(\mathbf{x})) \\ &= AB\mathbf{x} \\ &= (AB)\mathbf{x} \\ &= T_{AB}(\mathbf{x}) \end{aligned}$$

where $\mathbf{x} \in \mathbb{F}^p$. And therefore $T_A \circ T_B = T_{AB} : \mathbb{F}^p \to \mathbb{F}^m$.

2. If $V$ has a basis $\beta = \{\mathbf{v_1}, \mathbf{v_2}\}$ and $T : V \to V$ is a linear transformation given by

$$\begin{aligned} T(\mathbf{v_1}) &= 2\mathbf{v_1} + 3\mathbf{v_2} \\ T(\mathbf{v_2}) &= -7\mathbf{v_1} + 8\mathbf{v_2} \end{aligned}$$

To find $T \circ T(-\mathbf{v_1} + 3\mathbf{v_2})$ takes two steps as shown below.

$$\begin{aligned} T(-\mathbf{v_1} + 3\mathbf{v_2}) &= -T(\mathbf{v_1}) + 3T(\mathbf{v_2}) \\ &= -2\mathbf{v_1} - 3\mathbf{v_2} + 3(-7\mathbf{v_1} + 8\mathbf{v_2}) \\ &= -23\mathbf{v_1} + 21\mathbf{v_2} \end{aligned}$$

Hence

$$\begin{aligned} T \circ T(-\mathbf{v_1} + 3\mathbf{v_2}) &= T(-23\mathbf{v_1} + 21\mathbf{v_2}) \\ &= -23T(\mathbf{v_1}) + 21T(\mathbf{v_2}) \\ &= -23(2\mathbf{v_1} + 3\mathbf{v_2}) + 21(-7\mathbf{v_1} + 8\mathbf{v_2}) \\ &= -193\mathbf{v_1} + 99\mathbf{v_2} \end{aligned}$$

**Example**

Let $A$ be $m \times n$ and let $T = T_A$. Then $T_A : \mathbb{F}^n \to \mathbb{F}^m$. Let $\{\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_n}\}$ be the standard basis for $\mathbb{F}^n$. Then by the previous theorem it can be stated

$$\begin{aligned} Im(T_A) &= span(T_A(\mathbf{e_1}), T_A(\mathbf{e_2}), \ldots, T_A(\mathbf{e_n})) \\ &= span(A\mathbf{e_1}, A\mathbf{e_2}, \ldots, A\mathbf{e_n}) \\ &= span(col_1(A), col_2(A), \ldots, col_n(A)) \end{aligned}$$

## 4.4 Rank and Nullity

**Definitons** If $T : U \to V$ is a linear transformation,

- the dimension of the image of $T$ is called the **rank of $T$** and is denoted by rank($T$),
- the dimension of the kernel is called the **nullity of $T$** and is denoted by nullity(T).



**Example**

- Let $U$ be a vector space of dimension $n$, with basis $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$, and let $T : U \to U$ be a linear transformation defined by

$$T(\mathbf{u_1}) = \mathbf{u_2}, T(\mathbf{u_2}) = \mathbf{u_3}, \cdots, T(\mathbf{u_{n-1}}) = \mathbf{u_n} \text{ and } T(\mathbf{u_n}) = \mathbf{0}$$

Find bases for $ker(T)$ and $Im(T)$ and determine rank($T$) and nullity($T$).

**Theorem 4.6.** *If $T : U \to V$ is a linear transformation from an $n$-dimensional vector space $U$ to a vector space $V$, then*

$$rank(T) + nullity(T) = dim(U) = n$$

*Proof.* The proof is divided up into two cases.

1. (a) *the vectors* $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_s}$ *defined by*

$$\mathbf{u_j} = x_{1j}\mathbf{v_1} + x_{2j}\mathbf{v_2} + \cdots + x_{nj}\mathbf{v_n}$$

*will be a basis for the kernel of* $T$.

(b) *the vectors* $T(\mathbf{v_{c_1}}), T(\mathbf{v_{c_2}}), \ldots, T(\mathbf{v_{c_r}})$ *form a basis for the image of* $T$.

2. *If* $N(A) = \{\mathbf{0}\}$, *then* $ker(T) = \{\mathbf{0}\}$

*If* $C(A) = \{\mathbf{0}\}$, *then* $Im(T) = \{\mathbf{0}\}$

3. *Thus it follows that rank(T) = rank(A) and nullity(T) = nullity(A).*

A corollary to our result is that for a vector space $V$ over $\mathbb{F}$, $V$ is isomorphic to $\mathbb{F}^n$ if and only if $\dim(V) = n$.

This formalises our association of $n$ dimensional vector spaces with $\mathbb{F}^n$ as I hinted at when we looked at standard bases.

Another consequence is that we can associate $\ell(V, W)$ with $M_{m\times n}$.

## 4.9 Change of Basis

A basis of a vector space is a set of vectors that specify the coordinate system. A vector space may have an infinite number of bases but each basis contains the same number of vectors. The number of vectors in the basis is called the dimension of the vector space. The coordinate vector or coordinate matrix of a point changes with any change in the basis used. If the basis for a vector space is changed from some old bases $\beta$ to some new bases $\gamma$, how is the old coordinate vector $[\mathbf{v}]_\beta$ of a vector $\mathbf{v}$ related to the new coordinate vector $[\mathbf{v}]_\gamma$? The following theorem answers that question.



**Theorem 4.10.** *If the basis for a vector space is changed from some old basis* $\beta = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ *to some new basis* $\gamma = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$*, then the old coordinate vector* $[\mathbf{w}]_\beta$ *is related to the new coordinate vector* $[\mathbf{w}]_\gamma$ *of the same vector* $\mathbf{w}$ *by the equation*

$$[\mathbf{w}]_\gamma = P[\mathbf{w}]_\beta$$

*where the columns of* $P$ *are the coordinate vectors of the old basis vectors relative to the new basis; that is, the column vectors of* $P$ *are*

$$[\mathbf{u_1}]_\gamma, [\mathbf{u_2}]_\gamma, \ldots, [\mathbf{u_n}]_\gamma$$

$P$ *is called the* **change of basis matrix** *or the* **change of coordinate matrix**.

*Proof.* Let $V$ be a vector space with a basis $\beta = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ and a new basis

$\gamma = \{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$. Let $\mathbf{w} \in V$. Therefore $\mathbf{w}$ can be expressed as:

$$\mathbf{w} = a_1\mathbf{u_1} + a_2\mathbf{u_2} + \cdots + a_n\mathbf{u_n}$$

Thus we have

$$[\mathbf{w}]_\beta = (a_1, a_2, \dots, a_n)$$

As $\gamma$ is also a basis of $V$ the elements of $\beta$ can be expressed as follows

$$\mathbf{u_1} = p_{11}\mathbf{v_1} + p_{21}\mathbf{v_2} + \cdots + p_{n1}\mathbf{v_n}$$

$$\mathbf{u_2} = p_{12}\mathbf{v_1} + p_{22}\mathbf{v_2} + \cdots + p_{n2}\mathbf{v_n}$$

$$\vdots$$

$$\mathbf{u_n} = p_{1n}\mathbf{v_1} + p_{2n}\mathbf{v_2} + \cdots + p_{nn}\mathbf{v_n}$$

Combining this system of equations with the above expression for $\mathbf{w}$ gives



$$\mathbf{w} = (p_{11}a_1 + p_{12}a_2 + \cdots + p_{1n}a_n)\mathbf{v_1}$$

$$+(p_{21}a_1 + p_{22}a_2 + \cdots + p_{2n}a_n)\mathbf{v_2}+$$

$$\vdots$$

$$+(p_{n1}a_1 + p_{n2}a_2 + \cdots + p_{nn}a_n)\mathbf{v_n}+$$

and thus it can be seen that

$$[\mathbf{w}]_\gamma = \begin{bmatrix} p_{11}a_1 + p_{12}a_2 + \cdots + p_{1n}a_n \\ p_{21}a_1 + p_{22}a_2 + \cdots + p_{2n}a_n \\ \vdots \\ p_{n1}a_1 + p_{n2}a_2 + \cdots + p_{nn}a_n \end{bmatrix}$$

which can be written as

$$[\mathbf{w}]_\gamma = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$