Chapter 1 Vector Spaces

1 Fields

The rational numbers $Q$, the real numbers $R$ and the complex numbers $C$, (but not the natural numbers $N$ or the integers $Z$) with the usual operations of addition and multiplication, provide examples of a mathematical structure called a field.

Definition 1.1. (Field) A field is a set $F$ together with two binary operations, addition and multiplication, with the following properties.

(1) Addition: To every pair $a$ and $b$ in $F$ there corresponds an element $a+b$ of $F$ such that

(i) addition is associative: $(a+b) + c = a + (b + c)$;

(ii) addition is commutative: $a + b = b + a$;

(iii) there exists an element (additive identity) 0 of $F$ such that $a + 0 = a$ for every $a$ in $F$;

(iv) for every $a$ in $F$ there is an element (additive inverse) $-a$ of $F$ such that $a + (-a) = 0$.

(2) Multiplication: To every $a$ and $b$ in $F$ there corresponds an element $ab$ of $F$ such that

(i) multiplication is associative: $(ab)c = a(bc)$;

(ii) multiplication is commutative: $ab = ba$;

(iii) there exists an element (multiplicative identity) 1 of $F$, with $1 \neq 0$, such that $a1 = a$ for all $a$ in $F$;

(iv) for every $a$ in $F$ other than 0 there is an element (multiplicative inverse) $a^{-1}$ of $F$ such that $aa^{-1} = 1$.

(3) Multiplication distributes over addition: $a(b+c) = ab + ac$. Note: Multiplication takes precedence over addition. Thus, $ab + ac$ is read as $(ab) + (ac)$. 

1
Remark (Notation) We write \( a - b \) for \( a + (-b) \); \( \frac{1}{a} \) or \( 1/a \) for \( a^{-1} \); \( \frac{a}{b} \) or \( a/b \) for \( ab^{-1} \); \( x^2 \) for \( xx \), \( a^3 \) for \( xxx \) and so on.

Exercises Prove that the following hold in any field \( F \). Below, \( a, b \) and \( c \) are elements of \( F \).

1. Additive and multiplicative identities are unique.
2. Additive and multiplicative inverses are unique.
3. If \( a + b = a + c \) then \( b = c \).
4. \( a0 = 0 \).
5. \( (-1)a = -a \).
6. \( (-a)(-b) = ab \).
7. If \( ab = 0 \) then \( a = 0 \) or \( b = 0 \).

2 Vector Space

Definition 2.1. A vector space over the field \( F \) (called the set of scalars ) is a set \( V \) (whose elements are called vectors ) together with two binary operations, vector addition and scalar multiplication, with the following properties.

1. Vector addition: To every \( x \) and \( y \) in \( V \) there corresponds an element \( x + y \) of \( V \) such that
   (i) vector addition is associative: \( (x + y) + z = x + (y + z) \);
   (ii) vector addition is commutative: \( x + y = y + x \);
   (iii) there exists an element (additive identity or origin) \( 0 \) of \( V \) such that \( x + 0 = x \) for every \( x \) in \( V \);
   (iv) for every \( x \) in \( V \) there is an element (additive inverse) \( -x \) of \( V \) such that \( x + (-x) = (-x) + x = 0 \). We may write \( x - y \) for \( x + (-y) \).

2. Scalar multiplication: To every \( a \) in \( F \) and \( y \) in \( V \) there corresponds an element \( ax \) of \( V \) such that
   (i) scalar multiplication is associative: \( a(bx) = (ab)x \);
   (ii) \( 1x = x \) for every vector \( x \).

3. Scalar multiplication distributes over vector addition: \( a(x + y) = ax + ay \).
Scalar multiplication distributes over addition of scalars: \((a + b)x = ax + ay\).

**Examples** Note that a field must have at least two elements (0 and 1), while a vector space needs only have one (the origin). An example of a vector space over a field \(F\) is the field \(F\) itself. A less interesting example is \(\{0\}\) (the singleton of 0 in \(F\)). A less trivial example can be constructed as follows. Let \(C(R)\) denote the set of all continuous functions \(f : R \to R\) with scalar multiplication and vector addition defined by

\[
(1) \quad (af)(t) = af(t)
\]

and

\[
(2) \quad (f + g)(t) = f(t) + g(t)
\]

for all real \(t\). Since the function \(af\) is continuous whenever \(f\) is continuous and the function \(f + g\) is continuous whenever \(f\) and \(g\) are continuous, \(C(R)\) with the operations defined in (1) and (2) is a vector space over \(R\). The zero function \(0\) is defined by \(0(t) = 0\) for all real \(t\). The additive inverse of \(f\) is \(-f\), defined by \((-f)(t) = -f(t)\) for all real \(t\). By replacing “continuous” with “differentiable”, or “Riemann integrable”, we obtain more examples of vector spaces over \(R\). Whether \(0\) means the scalar or the vector additive identity is resolved by context.

If we drop continuity, we still have a vector space. The set \(R^R\) of all function from \(R\) to \(R\) with the operations defined by (1) and (2) is a vector space over \(R\). (We can replace \(R\) with any field \(F\).) The set \(R^N\) of all functions \(f : N \to R\), with scalar multiplication and vector addition still defined by (1) and (2), is another example of a vector space. Replace \(f\) and \(g\) with \(x\) and \(y\) and write \(x = (\xi_0, \xi_1, \cdots)\) and \(y = (\eta_0, \eta_1, \cdots)\), where \(\xi_n = x(n)\) and \(\eta_n = y(n)\) for \(n \in N\). Then, (1) and (2) can be restated as \(ax = (a\xi_0, a\xi_1, \cdots)\) and \(x + y = (\xi_0 + \eta_0, \xi_1 + \eta_1, \cdots)\). Further, \(0 = (0, 0, 0, \cdots)\) and \(-x = (-\xi_0, -\xi_1, \cdots)\). We can replace \(R\) with any field \(F\) to obtain \(F^N\), the vector space of all infinite sequences in \(F\).

Consider \(R^n\) for \(n \in Z\), with the operations of scalar multiplication and vector addition still defined by (1) and (2). It is a vector space.

### 3 Linear Dependence and Independence

**Definition 3.1.** A subset \(S = \{x_1, x_2, \cdots, x_n\}\) of a vector space \(V\) is **linearly dependent** if there exist scalars \(a_1, a_2, \cdots, a_n\), not all zero, such that

\[a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0.\]

Otherwise \(S\) is called **linearly independent**.
Any sum of the form $a_1 x_1 + \cdots + a_n x_n$ is called a **linear combination** of the vectors in $S$. The collection of all linear combinations of the vectors in $S$ (whether $S$ is linearly independent or not) is called the span of $S$ and is denoted by $\text{span } S$. In other words,

$$\text{span } S = \{ x \in V \mid x = a_1 x_1 + \cdots + a_n x_n, \text{ for scalars } a_1, \ldots, a_n \}.$$  

**Remark 3.1.** The condition that $S = \{ x_1, x_2, \cdots, x_n \}$ is linearly dependent can be written

$$\exists a_i (a_1 x_1 + \cdots + a_n x_n = 0 \land a_i \neq 0).$$

In other words, $S$ is linearly independent iff for all scalars $a_1, \cdots, a_n$,

$$a_1 x_1 + \cdots + a_n x_n = 0$$

implies

$$a_1 = \cdots = a_n = 0.$$  

**Remark 3.2.** If $S = \{ x_1, x_2, \cdots, x_n \}$ is linearly independent then for all $k$, $1 \leq k \leq n$, $\{ x_1, x_2, \cdots, x_k \}$ is linearly independent.

Indeed, if $a_1 x_1 + \cdots + a_k x_k = 0$ then, by choosing $a_{k+1} = \cdots = a_n = 0$, we have

$$0 = a_1 x_1 + \cdots + a_k x_k = a_1 x_1 + \cdots + a_k x_k + a_{k+1} x_{k+1} + \cdots + a_n x_n,$$

which implies $a_1 = \cdots = a_n = 0$, since $S$ is linearly independent.

**Examples**

**Remark 3.3.** Note that the empty set $\emptyset$ is linearly independent and that any set containing the zero vector is linearly dependent. If any $x_k$ in $S = \{ x_1, \cdots, x_n \}$ is a linear combination of the other vectors, then

$$\text{span } \{ x_1, \cdots, x_n \} = \text{span } \{ x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n \}.$$  

In fact, if $x_k = \sum_{i=1}^n b_i x_i$, then

$$\sum_{i=1}^n a_i x_i = \sum_{i=1, i \neq k}^n a_i x_i + a_k \sum_{i=1, i \neq k}^n b_i x_i = \sum_{i=1, i \neq k}^n (a_i + a_k b_i) x_i.$$  

If $x \in \text{span } X$, where $X = \{ x_1, x_2, \cdots, x_n \}$ is linearly independent, then the representation

$$x = a_1 x_1 + \cdots + a_n x_n$$

is unique, i.e.: $x$ uniquely determines $a_1, \cdots, a_n$. Indeed, if we also have $x = b_1 x_1 + \cdots + b_n x_n$, then from the above, we obtain $(a_1 - b_1) x_1 + \cdots + (a_n - b_n) x_n = 0$, which implies, thanks to the linear independence of $X$, that $a_i = b_i$ for $i = 1, \cdots, n$.  

4
Theorem 3.1. A set of vectors is linearly dependent iff one of the vectors is a linear combination of the others.

Proof. If \( S = \{x_1, \ldots, x_n\} \) is linearly dependent, then there exists \( a_k \neq 0 \) for some \( k \) such that \( a_1x_1 + \cdots + a_kx_k + \cdots + a_nx_n = 0 \). Therefore,

\[
x_k = \sum_{i=1, i \neq k}^{n} \left( -\frac{a_i}{a_k} \right)x_i.
\]

If, on the other hand, one of the vectors in \( S \) is a linear combination of the others, we have

\[
x_k = \sum_{i=1, i \neq k}^{n} a_i x_i,
\]

for some collection \( a_1, \ldots, a_{k-1}, a_{k+1}, \ldots, a_n \) of scalars. Then, by choosing \( a_k = -1 \) we have \( a_1x_1 + \cdots + a_nx_n = 0 \), where not all of the \( a_i \) are zero, showing that \( S \) is linearly dependent. \( \square \)

4 Bases and Dimension

Definition 4.1. A linearly independent subset \( X \) of a vector space \( V \) is called a basis in \( V \) if every vector in \( V \) is a linear combination of the vectors in \( X \). A vector space is finite-dimensional if it has a finite basis.

Examples

Every vector \( x \) in \( \mathbb{R}^3 \) is of the form \( x = (\xi_1, \xi_2, \xi_3) \). Therefore, \( x = (\xi_1, \xi_2, \xi_3) = \xi_1(1, 0, 0) + \xi_2(0, 1, 0) + \xi_3(0, 0, 1) \), showing that \( X = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} \) spans \( \mathbb{R}^3 \). Is \( X \) linearly independent? It is, since \( 0 = a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) = (a_1, a_2, a_3) \) implies \( a_1 = a_2 = a_3 = 0 \).

For a natural number \( n \), let \( P_n(R) \) denote the set of all polynomials of the form

\[
p(t) = a_0 + a_1t + \cdots + a_nt^n
\]

in the real variable \( t \), where \( a_0, \ldots, a_n \) are real numbers. In other words, \( P_n(R) \) is the subset of \( \mathbb{R}^k \) consisting of all functions \( p : R \to R : t \to p(t) \) of the form polynomials. It is easy to verify that \( P_n(R) \) is a vector space over \( R \).

For \( k = 0, \ldots, n \), define \( x_k : R \to R : t \to t^k \). Clearly, each \( x_k \) is an element of \( P_n(R) \). Moreover,

\[
p(t) = a_0x_0(t) + a_1x_1(t) + \cdots + a_nx_n(t)
\]

for all real \( t \), showing that \( X = \{x_0, \ldots, x_n\} \) spans \( P_n(R) \).

If \( p(t) = a_0 + a_1t + \cdots + a_nt^n = 0 \) for all \( t \) then \( p \) is the zero polynomial, i.e.: \( a_0 = \cdots = a_n = 0 \). (A fact we do not prove here.) This establishes the linear independence of \( X \). Hence, \( X \) is a basis in \( P_n(R) \).
Theorem 4.1. If $V$ is a finite-dimensional vector space and $Y$ is any linearly independent subset of $V$ then $Y \subset S$ for some basis $S$ in $V$. In other words, any linearly independent set of vectors in a finite-dimensional vector space that is not already a basis can be augmented in such a way that the augmented set is a basis.

Proof. Suppose that $V$ is a finite-dimensional vector space and that $Y = \{y_1, y_2, \cdots, y_k\}$ is a linearly independent subset of $V$. If the span of $Y$ is $V$, we take $S = Y$ and we are finished. Otherwise, we proceed inductively.

Since $V$ is finite-dimensional, it has a finite basis, say $X = \{x_1, \cdots, x_n\}$. Define $X_1 = X \setminus Y = \{x_{1,1}, \cdots, x_{1,k}\}$, the set obtained from $X$ by removing those vectors that happen to be in both sets. Then, $k_1 < n$. Define $S_1 = Y \cup X_1$. Clearly, span $S_1 = V$. If $S_1$ is linearly independent, we are finished: Take $S = S_1$.

Otherwise, we claim that one of the vectors in $X_1$ is a linear combination of the others. In fact, let $x_{1,l}$ be in $X_1$ such that $\{y_1, \cdots, y_k, x_{1,1}, \cdots, x_{1,l}\}$ is linearly dependent. Then there exists scalars $a_1, \cdots, a_k, b_1, \cdots, b_l$, not all zero, such that

$$a_1y_1 + \cdots + a_ky_k + b_1x_{1,1} + \cdots + b_lx_{1,l} = 0.$$ 

Since $\{y_1, \cdots, y_k, x_{1,1}, \cdots, x_{1,l-1}\}$ is linearly independent, we deduce that $b_l \neq 0$. Therefore $x_{1,l}$ is a linear combination of the others. This proves the claim.

Remove from $X_1$ one of the vectors that is a linear combination of the other elements in $S_1$ to obtain a proper subset $X_2 = \{x_{2,1}, \cdots, x_{2,k_2}\}$ of $X_1$. Then, $k_2 < k_1$. Define $S_2 = Y \cup X_2$. Nothing has been lost in the sense that the span of $S_2$ is the span of $S_1$, which is $V$.

By repeating this process, at some point, we arrive at a set $S_m = Y \cup X_m$, where $X_m$ is a nonempty proper subset of $X$, the span of $S_m$ is still $V$ and $S_m$ is linearly independent. We take $S = S_m$.

\[\Box\]

Theorem 4.2. If a vector space $V$ is spanned by a subset of $n$ vectors then every subset of $V$ that contains $n + 1$ vectors is linearly dependent.

Proof. Proof Suppose the proposition is false. Then, there is a vector space $V$ spanned by some subset $X_0 = \{x_1, \cdots, x_n\}$ and there is a linearly independent subset $Y = \{y_1, \cdots, y_{n+1}\}$ of $V$. Starting with $X_0$, we construct inductively sets $X_1, \cdots, X_n$ such that span $X_0 = \text{span} X_1 = \cdots = \text{span} X_n = V$ and such that each $X_i$ is obtained from $X_0$ by replacing $i$ of its elements with elements of $Y$.

To construct $X_1$, we proceed as follows. Since $X_0$ spans $V$, $y_1$ is in the span of $X_0$. Hence, there are scalars $a_1 = 1$ and $b_1, \cdots, b_n$ such that

$$a_1y_1 + \sum_{k=1}^{n} b_kx_k = 0.$$
Suppose $b_1 = \cdots = b_n = 0$ then $a_1 = 0$, a contradiction. Hence some $b_k \neq 0$ and $x_k$ in $X_0$ is a linear combination of the other members of the set $Z_0 = X_0 \cup \{y_1\}$. We may as well suppose that this element $x_k$ is $x_n$ (by renaming the elements of $X_0$, if necessary). Remove $x_n$ from $Z_0$: Define $X_1 = Z_0 \setminus \{x_n\}$. Then, $\text{span } X_1 = \text{span } X_0$. Note that

\[ X_1 = \{y_1\} \cup ((X_0 \setminus \{x_n\}) = \{y_1, x_1, \ldots, x_{n-1}\}. \]

Suppose that for $0 \leq i < n$ a set $X_i$ has been defined which spans $V$ and which consists of all $y_j$ with $1 \leq j \leq i$ and of $n - i$ elements of $X_0$, say $x_1, \ldots, x_{n-i}$, i.e.: $X_i = \{y_1, \ldots, y_i, x_1, \ldots, x_{n-i}\}$. Since $X_i$ span $V$, $y_{i+1}$ is in the span of $X_i$. Hence, there are scalars $a_1, \ldots, a_{i+1}$ and $b_1, \ldots, b_{n-i}$, with $a_{i+1} = 1$, such that

\[ a_{i+1}y_{i+1} + \sum_{j=1}^{i} a_jy_j + \sum_{k=1}^{n-i} b_kx_k = 0. \]

If $b_1 = \cdots = b_{n-i} = 0$, then, since $Y$ is linearly independent, $a_1 = \cdots = a_{i+1} = 0$, a contradiction. Therefore, some $x_k$ in $X_i$ is a linear combination of the other elements of $Z_i = X_i \cup \{y_{i+1}\}$. We may as well suppose that this element $x_k$ is $x_{n-i}$ (by renaming the elements of $X_0$, if necessary).

Define $X_{i+1} = Z_i \setminus \{x_{n-i}\}$. Then, $\text{span } X_{i+1} = \text{span } Z_i = V$ and $X_{i+1}$ has the properties stated for $X_i$ with $i$ replaced by $i + 1$. We have $X_n = \{y_1, \ldots, y_n\}$ with $y_{n+1}$ not in the span of $X_n$, since $Y$ is linearly independent, a contradiction, since $X_n$ spans $V$.

**Corollary 4.3.** If $V$ is a finite-dimensional vector space then any two bases have the same number of elements.

**Proof.** Suppose $V$ is a finite-dimensional vector space with bases $X$, with $m$ elements, and $Y$, with $n$ elements. If $m > n$, then $X$ is linearly dependent, a contradiction. Hence, $m \leq n$. If $n > m$ then $Y$ is linearly dependent, another contradiction. Hence, $n \leq m$. We conclude that $m = n$. \hfill \Box

**Definition 4.2.** The number of elements in a basis of a finite-dimensional vector space $V$ is called the dimension of $V$ and is denoted by $\dim X$.

**Theorem 4.4.** In an $n$-dimensional vector space any set of $n$ linearly independent vectors is a basis.

**Proof.** Let $V$ be an $n$-dimensional vector space. Any set $X$ of $n$ linearly independent vectors in $V$ can be extended to a basis $S$ such that $X \subseteq S$. The number of elements of $S$ is $n$, since $V$ is $n$-dimensional. Hence, $X = S$, since $X$ also has $n$ elements and $X \subseteq S$. Therefore, $X$ is a basis. \hfill \Box
Exercises

1. Given the vectors \((1,1,1)\) and \((1,0,1)\) in \(C^3\) find a third vector such that the three vectors are a basis.

2. Suppose \(\{x, y, z\}\) is a linearly independent set. Is the set \(\{x + y, x + z, y + z\}\) linearly independent also? Can you give a criterion to check the independence for general case?

3. Show that the set \(\{x_0, x_1, x_2\}\) in \(P(R)\) defined by \(x_0(t) = 1, x_1(t) = t^2\) and \(x_2(t) = 1 + t + t^2\) is linearly independent.

4. Consider the subset of \(R^3\) given by \(S = \{(1, a, a^2), (1, b, b^2), (1, c, c^2)\}\). Impose conditions on \(a, b\) and \(c\) so that \(S\) is linearly independent.

5. Prove that \(C(R)\) is not finite-dimensional.

5 Subspaces

Definition 5.1. A nonempty subset \(W\) of a vector space \(V\) is called a subspace of \(V\) if, for any \(x, y \in W\) and scalar \(a\), one has

(i) \(x + y \in W\); and

(ii) \(ax \in W\).

Theorem 5.1. Suppose \(W\) is a subset of \(V\). Then \(W\) is a subspace of \(V\) iff, for all \(x, y \in V\), and all scalars \(a, b\),

\[ x, y \in W \Rightarrow ax + by \in W. \]

Examples.

Suppose \(V\) is an \(n\)-dimensional vector space and \(W\) is a subspace of \(V\). Since \(V\) cannot contain \(n + 1\) linearly independent vectors, neither can \(W\). Hence, no basis in \(W\) can have more that \(n\) elements. This establishes that if \(W\) has a basis then the dimension of \(W\) cannot exceed the dimension of \(V\), but does not establish that \(W\) has a basis. In other words, to say that \(W\) is \(m\)-dimensional, we must show that \(W\) has a basis with \(m\) elements. Where does such a basis come from?

This issue is resolved below.

Theorem 5.2. If \(W\) is a subspace of the finite-dimensional vector space \(V\) then \(\dim W \leq \dim V\):
Proof. Suppose that $V$ is a vector space of dimension $n$ and that $W$ is a subspace of $V$. We already know that a subspace of a vector space is also a vector space. Therefore $W$ is a vector space. We have to find a basis for it with at most $n$ elements.

If $W = \{0\}$ then $W$ is zero-dimensional. Otherwise, $W$ contains a nonzero vector $x_1$. If $x_1$ spans $W$ then $W$ is one-dimensional. Otherwise, $W$ contains a vector $x_2$ not in the span of $x_1$ (so that $x_1$ and $x_2$ are linearly independent). If $W = \text{span} \{x_1, x_2\}$ then $W$ is two-dimensional.

We repeat this process by finding $m$ linearly independent vectors $x_1, \ldots, x_m$ in $W$, with $m \leq n$. If $W = \text{span} \{x_1, \ldots, x_m\}$ then $W$ is $m$-dimensional.

This process cannot last more than $n$ steps, for once we obtain a linearly independent subset $S = \{x_1, \ldots, x_m\}$ of $W$ (which must be also a subset of $V$) if $S$ does not span $W$ then there is a vector $x_{m+1}$ in $W$ with $x_{m+1}$ not in the span of $S$. If $m = n$ then $V$ contains $n + 1$ linearly independent vectors, which is a contradiction. We have shown that $W$ contains $m$ linearly independent vectors, $m \leq n$ which span $W$. Hence $\dim W \leq \dim V$.

Corollary 5.3. If $V$ is a vector space of dimension $n$ and $W$ is a subspace of dimension $m$ then there is a basis $\{x_1, \ldots, x_m, x_{m+1}, \ldots, x_n\}$ in $V$ such that $\{x_1, \ldots, x_m\}$ is a basis in $W$.

Proof. According to the proposition, there is a basis $\{x_1, \ldots, x_m\}$ in $W$ for some $m \leq n$. Extend this linearly independent subset of $V$ to a basis in $V$.

Remark 5.1. 1. If $X$ and $Y$ are subspaces of the vector space $V$, is $X \cup Y$ a subspace of $V$?

2. It is not clear whether the intersection $X \cap Y$ of two subspaces is a subspace. This is clarified below.

6 Intersection and Sum of Subspaces

Theorem 6.1. An intersection of subspaces of a vector space $V$ is also a subspace of $V$.

Proof. Let $V$ be a vector space. Let $\{X_i \mid i \in I\}$ be an arbitrary nonempty collection of subspaces of $V$, where $I$ is a index set. Let

$$X = \bigcap_{i \in I} X_i.$$ 

Every $X_i$ is a vector space. Hence, $0 \in X_i$ for all $i \in I$. Therefore, $0 \in X$. This shows that the intersection $X$ is not empty.

Suppose that $x$ and $y$ are in $X$. Then $x$ and $y$ also belong to $X_i$ for all $i \in I$. Therefore, for any pair of scalars $a$ and $b$, the vector $ax + by$ also belong to every $X_i$. 


since each is a vector space. It follows that $ax + by$ belongs to $X$. We conclude that $X$ is a subspace of $V$. 

**Definition 6.1.** If $X$ and $Y$ are subspaces of the vector space $V$, define the vector space

$$X + Y = \{x + y \in V \mid x \in X \text{ and } y \in Y\}.$$

In other words, $X + Y = \text{span}(X \cup Y)$.

**Definition 6.2.** Two subspaces $X$ and $Y$ of a vector space $V$ are called disjoint if $X \cap Y = \{0\}$. If $X$ and $Y$ are disjoint and $X + Y = V$ then $Y$ is called a complement of $X$.

**Remark 6.1.** (a) Note that in the above definition $X$ and $Y$ are not disjoint sets ($X \cap Y \neq \emptyset$)

(b) Do not confuse the relative complement of a set with the complement of a subspace relative to the whole space. A subspace of a vector space can have many complements. For example, consider the vector space $\mathbb{R}^2$ and the subspace $X$ of all vectors of the form $(\xi_1, 0)$ (the horizontal axis). Other than the horizontal axis, every line $Y$ containing the origin is a complement of $X$ and $\mathbb{R}^2 = X + Y$. To be specific, for real $a$ let $Y_a$ denote the subspace consisting of all vectors of the form $(\xi_1, a\xi_1)$. Then, $Y_0 = X$. Suppose that $a \neq 0$. Then, $Y_a$ is a complement of $X$.

**Example**

In $\mathbb{C}^3$, the vector space of all triples $(\xi_1, \xi_2, \xi_3)$ of complex numbers, define $X$ to be the subspace of all complex triples of the form $(\xi_1, \xi_2, 0)$, $Y$ to be the subspace of all complex triples of the form $(0, \xi_2, 0)$ and $Z$ to be the subspace of all complex triples of the form $(0, 0, \xi_3)$. Then,

1. $Y \cap Z = X \cap Z = \{0\}$.
2. $X \cap Y$ is the subspace of $V$ consisting of all triples of the form $(0, \xi_2, 0)$.
3. $Y + Z$ is the subspace of $V$ consisting of all complex triples of the form $(0, \xi_2, \xi_3)$.
4. $X + Y$ is the subspace of $V$ consisting of all triples of the form $(\xi_1, \xi_2, 0)$.
5. $X + Z = X + Y + Z = V$.
6. $X \cap (Y + Z)$ is the subspace of $V$ consisting of all triples of the form $(0, \xi_2, 0)$.
7. $(X \cap Y) + (X \cap Z)$ is the subspace of $V$ consisting of all triples of the form $(0, \xi_2, 0)$.

**Theorem 6.2. (Dimension Formula)** If $X$ and $Y$ are subspaces of the finite-dimensional vector space $V$ then

$$\dim X + \dim Y = \dim(X + Y) + \dim(X \cap Y).$$
\textbf{Proof.} Assume the dimensions of $X$ and $Y$ are $n_1$ and $n_2$ and the dimension of $X \cap Y$ is $m$. Choose a basis 
\[ \{\alpha_1, \alpha_2, \cdots, \alpha_m\} \]
of $X \cap Y$. By Theorem \?, it can be augmented to form a basis 
\[ \{\alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \beta_2, \cdots, \beta_{n_1-m}\} \]
of $X$. It can also be augmented to form a basis 
\[ \{\alpha_1, \alpha_2, \cdots, \alpha_m, \gamma_1, \gamma_2, \cdots, \gamma_{n_2-m}\} \]
of $Y$.

We claim that 
\[ \{\alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \beta_2, \cdots, \beta_{n_1-m}, \gamma_1, \gamma_2, \cdots, \gamma_{n_2-m}\} \]
is a basis of $X + Y$. If the claim is true then the dimension of $X + Y$ is $n_1 + n_2 - m$ and the theorem is proved.

We prove the claim. Since 
\[ X = \text{span} \{\alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \beta_2, \cdots, \beta_{n_1-m}\} \]
and 
\[ Y = \text{span} \{\alpha_1, \alpha_2, \cdots, \alpha_m, \gamma_1, \gamma_2, \cdots, \gamma_{n_2-m}\} \],
it is obvious that 
\[ X + Y = \text{span} \{\alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \beta_2, \cdots, \beta_{n_1-m}, \gamma_1, \gamma_2, \cdots, \gamma_{n_2-m}\} \].
We only need to prove that 
\[ \{\alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \beta_2, \cdots, \beta_{n_1-m}, \gamma_1, \gamma_2, \cdots, \gamma_{n_2-m}\} \]
is linearly independent. Assume that 
\[ k_1\alpha_1 + \cdots + k_m\alpha_m + p_1\beta_1 + \cdots + p_{n_1-m}\beta_{n_1-m} + q_1\gamma_1 + \cdots + q_{n_1-m}\gamma_{n_2-m} = 0 \]
Let 
\[ \alpha = k_1\alpha_1 + \cdots + k_m\alpha_m + p_1\beta_1 + \cdots + p_{n_1-m}\beta_{n_1-m} = -q_1\gamma_1 - \cdots - q_{n_1-m}\gamma_{n_2-m} \]
From the first equality we know that $\alpha \in X$ and from the second equality we deduce that $\alpha \in Y$. Therefore, $\alpha \in X \cap Y$. Therefore, $\alpha$ can be expressed as a linear combination of $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$, that is

$$\alpha = l_1 \alpha_1 + l_2 \alpha_2 + \cdots + l_m \alpha_m.$$  

Then,

$$l_1 \alpha_1 + l_2 \alpha_2 + \cdots + l_m \alpha_m + q_1 \gamma_1 + \cdots + q_{n_1-m} \gamma_{n_2-m} = 0.$$  

Since $\{\alpha_1, \alpha_2, \cdots, \alpha_m, \gamma_1, \gamma_2, \cdots, \gamma_{n_2-m}\}$ is linearly independent, $l_1 = l_2 = \cdots = l_m = 0$ and $q_1 = q_2 = \cdots + q_{n_2-m} = 0$. Therefore, $\alpha = 0$. The first equality above becomes

$$k_1 \alpha_1 + \cdots + k_m \alpha_m + p_1 \beta_1 + \cdots + \phi_{n_1-m} \beta_{n_1-m} = \alpha = 0.$$  

It conclude that

$$k_1 = k_2 = \cdots = k_m = p_1 = \cdots = p_{n_1-m} = 0,$$  

since $\{\alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \beta_2, \cdots, \beta_{n_1-m}\}$ is linearly independent. Therefore

$$\{\alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \beta_2, \cdots, \beta_{n_1-m}, \gamma_1, \gamma_2, \cdots, \gamma_{n_2-m}\}$$

is linearly independent. This completes the proof.  

**Corollary 6.3.** Let $V_1$ and $V_2$ be subspaces of a n-dimensional vector space. If the sum of the dimensions of $V_1$ and $V_2$ is greater than the n, then $V_1$ and $V_2$ contains a common nonzero vector.

**Proof.** It leaves to you as a homework.  

**Exercises**

1. Find a vector space $V$ and subspaces $X$, $Y$ and $Z$ such that

$$X \cap (Y + Z) \neq (X \cap Y) + (X \cap Z).$$

2. Show that if $X$ and $Y$ are two-dimensional subspaces of a three-dimensional space then $X$ and $Y$ are not disjoint.

3. Prove that if $X$ is an m-dimensional subspace of an n-dimensional vector space then every complement of $X$ has dimension $n - m$.

4. A function $f$ in $R^R$ is called even if $f(-t) = f(t)$ for all real $t$ and is called odd if $f(-t) = -f(t)$ for all real $t$. Prove that the even functions and the odd functions are disjoint subspaces of $R^R$.  

12
7 Direct Sums

An important case of sum of subspaces is the direct sum.

**Definition 7.1.** Let $V_1$ and $V_2$ be subspaces of a vector space $V$. The sum $V_1 + V_2$ is called a direct sum if, for any vector $\alpha$ in $V_1 + V_2$, there is a unique decomposition

$$\alpha = \alpha_1 + \alpha_2,$$

where $\alpha_1 \in V_1$ and $\alpha_2 \in V_2$. The direct sum is denoted by $V_1 \bigoplus V_2$.

**Theorem 7.1.** The sum $V_1 + V_2$ is the direct sum if and only if that

$$\alpha_1 + \alpha_2 = 0, \alpha_i \in V_i \ (i = 1, 2)$$

implies that $\alpha_1 = \alpha_2 = 0$.

**Proof.** The condition of the theorem states that the decomposition of the zero vector is unique. It is necessary obviously. Let us prove the sufficiency.

Suppose that $\alpha \in V_1 + V_2$ has two decompositions:

$$\alpha = \alpha_1 + \alpha_2 = \beta_1 + \beta_2,$$

where $\alpha_i, \beta_i \in V_i$, for $i = 1, 2$. Then $(\alpha_1 - \beta_1) + (\alpha_2 - \beta_2) = 0$. Note that $\alpha_i - \beta_i \in V_i$ for $i = 1, 2$. By the assumption of the theorem, we deduce that $\alpha_i - \beta_i = 0$, i.e., $\alpha_i = \beta_i$ for $i = 1, 2$. That is the decomposition is unique.

**Corollary 7.2.** Sum $V_1 + V_2$ is the direct sum if and only if $V_1 \cap V_2 = \{0\}$.

**Proof.** Sufficiency. Suppose that $\alpha_1 + \alpha_2 = 0$ and $\alpha_i \in V_i$ for $i = 1, 2$. Then $\alpha_1 = -\alpha_2 \in V_1 \cap V_2$. By the assumption, we deduce that $\alpha_1 = \alpha_2 = 0$ That is the sum is the direct sum.

Necessity. For any vector $\alpha \in V_1 \cap V_2$, we write $0 = \alpha + (-\alpha)$ where $\alpha \in V_1$ and $-\alpha \in V_2$. Since the sum $V_1 + V_2$ is the direct sum, we deduce that $\alpha = 0$, i.e., $V_1 \cap V_2 = \{0\}$.

**Theorem 7.3.** Let $V_1$ and $V_2$ be subspaces of a vector space $V$ and $W = V_1 + V_2$. Then $W = V_1 \bigoplus V_2$ if and only if

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2.$$ 

**Proof.** It is a direct consequence of Dimension Formula.
Theorem 7.4. Let $U$ be a subspace of a $n$-dimensional space $W$. There exists a subspace of $W$ such that $W = U \oplus V$. In other words, every subspace of a finite-dimensional vector space has a complement.

Proof. Let $W$ be a vector space of dimension $n$ and let $U$ be a subspace of dimension $m$. Then, there is a basis $X = \{x_1, \ldots, x_m\}$ in $U$. By a previous proposition, there is a basis $Z = \{x_1, \ldots, x_m, y_1, \ldots, y_{n-m}\}$ in $W$. Define

$$V = \text{span} \{y_1, \ldots, y_{n-m}\}.$$ 

Then, the dimension of $V$ is $n - m$ since $\{y_1, \ldots, y_n\}$ is linearly independent.

We need to show that $W = U \oplus V$. We show that for every $z$ in $W$ there is exactly one pair of vectors $x$ in $U$ and $y$ in $V$ such that $z = x + y$.

Indeed, since $Z$ is a basis in $W$, for every $z$ in $W$ there are unique scalars $a_1, \ldots, a_m$ and $b_1, \ldots, b_{n-m}$ such that $z = a_1 x_1 + \cdots + a_m x_m + b_1 y_1 + \cdots + b_{n-m} y_{n-m}$ Therefore, since $U \cap V = \{0\}$, there are uniquely defined vectors $x = a_1 x_1 + \cdots + a_m x_m$ and $y = b_1 y_1 + \cdots + b_{n-m} y_{n-m}$ such that $z = x + y$. \qed

The concept of direct sums can be extended to many subspaces.

Definition 7.2. Let $V_1, V_2, \cdots, V_s$ be subspaces of a vector space $V$. The sum $V_1 + V_2 + \cdots + V_s$ is called the direct sum of $V_1, V_2, \cdots, V_s$ if, for any vector $\alpha$ in $V_1, V_2, \cdots, V_s$, there is a unique decomposition

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_s,$$

where $\alpha_i \in V_i$ for $i = 1, 2, \cdots, s$. The direct sum is denoted by $V_1 \oplus V_2 \oplus \cdots \oplus V_s$.

Similarly, we have,

Theorem 7.5. Let $V_1, V_2, \cdots, V_s$ be subspaces of a vector space $V$. The following statements are equivalent

(i) $W = \sum V_i$ is the direct sum.

(ii) Decomposition of the zero vector is unique.

(iii) $V_i \cap \sum_{j \neq i} V_j = \{0\}$ for $i = 1, 2, \cdots, s$.

(iv) $\dim W = \sum \dim V_i$. 

14
8 Direct Sums

We begin our presentation of direct sums by considering two examples.

**Example.** Let $F$ be a field. Define $U = F^m$, $V = F^n$ and $W = F^{m+n}$. If $x = (x_1, \ldots, x_n)$ is a vector in $U$ and $y = (y_1, \ldots, y_m)$ is a vector in $V$, the vector $z = (x_1, \ldots, x_n, y_1, \ldots, y_m)$ is a vector in $W$. We might write $z = (x, y)$ and think of this construction as a method for obtaining longer vectors from shorter ones. Let us temporarily give an unofficial definition of the symbol $\bigoplus$, by proposing to define this method of obtaining $W$ from $U$ and $V$ as the result of the operation $W = U \bigoplus V$.

Each vector $x$ in $U$ can be identified with $(x, 0)$ in $W = U \bigoplus V$ and each vector $y$ in $V$ can be identified with $(0, y)$ in $W = U \bigoplus V$. Can this construction be extended to a more general scheme that can be applied to every pair of vector spaces? We answer this question below. Before we do so, it is worthwhile to consider another example.

**5.2 Example** Suppose we start with $W = F^{m+n}$. Define $U$ to be the subspace of $F^{m+n}$ consisting of all vectors $x$ of the form $x = (x_1, \ldots, x_m, 0, \ldots, 0)$. Define $V$ to be the subspace of $F^{m+n}$ consisting of all vectors $y$ of the form $y = (0, \ldots, 0, y_1, \ldots, y_n)$. Then, $U$ and $V$ are disjoint subspaces of $W$ and $W = U + V$, i.e.: $U$ and $V$ are complements of each other. Since $U$ is isomorphic to $F^m$ and $V$ is isomorphic to $F^n$, we can identify $U$ with $F^m$ and $V$ with $F^n$. But, if we do this, we can no longer write $W = U + V$, since neither the vectors in $U$ nor the vectors in $V$ are necessarily also vectors in $W$. (We can, however, write $W = U \bigoplus V$, consistently with the previous example.)

If we agree that this distinction is only technical and not substantive, we can simply ignore it. Indeed, this is the course of action we shall take in what follows.

**Remark 8.1.** The operation $U \bigoplus V$ described in the first example above is an example of an “external direct sum” of two vector spaces (over the same field), while the operation $U \bigoplus V$ described in the second example above is an example of an “interior direct sum” of two disjoint subspaces of a vector space. Strictly speaking, they are different operations.

**Definition 8.1.** Definition Let $U$ and $V$ be vector spaces over the same field. We define their direct sum $U \bigoplus V$ to be the collection of all ordered pairs $(x, y)$ with $x$ in $U$ and $y$ in $V$. We make $U \bigoplus V$ into a vector space by defining scalar multiplication by $a(x, y) = (ax, ay)$ and vector addition by $(x, y) + (x', y') = (x + x', y + y')$.

The zero vector in $U \bigoplus V$ is $(0, 0)$. Note that

- $U$ is isomorphic to $\{(x, y) \in U \bigoplus V \mid y = 0\}$, a subspace of $U \bigoplus V$, and
- $V$ is isomorphic to $\{(x, y) \in U \bigoplus V \mid x = 0\}$, a subspace of $U \bigoplus V$.

So, although neither $U$ nor $V$ is a subspace of $U \bigoplus V$, each is isomorphic to a subspace of $U \bigoplus V$. With the proper identifications, by abuse of language, we speak of $U$ and
V as being subspaces of $U \oplus V$. Note that $U \cap V = \{0\}$ (the singleton of the zero vector in $U \oplus V$).

Now, let’s turn things around. If $U$ and $V$ are subspaces of a vector space $W$, when can we say that $W = U \oplus V$? We answer this question below.

**Theorem 8.1.** Let $U$ and $V$ be subspaces of a vector space $W$. The following three conditions are equivalent.

1. $W = U \oplus V$.
2. $U$ and $V$ are complements of each other. That is, $W = U + V$ and $U \cap V = \{0\}$.
3. For every vector $z$ in $W$ there is exactly one pair of vectors $x$ in $U$ and $y$ in $V$ such that $z = x + y$.

**Proof.** We prove that (1) implies (2). Every element $z$ of $W$ can be written $z = (x, y) = (x, 0) + (0, y)$. Hence, $Z = U + V$. If $W = U \oplus V$ then $z = (x, y) \in U$ and $z = (x, y) \in V$ imply $y = 0$ and $x = 0$, i.e., $z = 0$. Hence, $U \cap V = \{0\}$.

We prove that (2) implies (3). If $W = U + V$ then every element $z$ of $W$ is of the form $z = x + y$, with $x$ in $U$ and $y$ in $V$. If $z = x' + y'$ also, with $x'$ in $U$ and $y'$ in $V$, then $x - x' = y - y'$. Since $x - x'$ is in $U$, $y - y'$ is in $V$, and $U \cap V = \{0\}$, it follows that $x - x' = 0$ and $y - y' = 0$, showing that the representation $z = x + y$ is unique.

We prove that (3) implies (1). Suppose that for every $z$ in $W$ the representation $z = x + y$, with $x$ in $U$ and $y$ in $V$ is unique. Since we have agreed to identify every $x$ in $U$ with $(x, 0)$ in $U \oplus V$ and to identify every $y$ in $V$ with $(0, y)$ in $U \oplus V$, $z$ can be identified uniquely with $(x, y)$, since $(x, y) = (x, 0) + (0, y) = x + y = z$. With this identification, the elements $z$ of $W$ are precisely the elements $(x, y)$ of $U \oplus V$. 

**Theorem 8.2.** If $U$ and $V$ are vector spaces over the same field then

$$\dim U \oplus V = \dim U + \dim V.$$ 

**Proof.** It is a direct consequence of Dimension Formula.

**Theorem 8.3.** Let $U$ be a subspace of a $n$-dimensional space $W$. There exists a subspace of $W$ such that $W = U \oplus V$. In other words, every subspace of a finite-dimensional vector space has a complement.

**Proof.** Let $W$ be a vector space of dimension $n$ and let $U$ be a subspace of dimension $m$. Then, there is a basis $X = \{x_1, \cdots, x_m\}$ in $U$. By a previous proposition, there is a basis $Z = \{x_1, \cdots, x_m, y_1, \cdots, y_{n-m}\}$ in $W$. Define

$$V = \text{span} \ \{y_1, \cdots, y_{n-m}\}.$$
Then, the dimension of $V$ is $n - m$ since $\{y_1, \ldots, y_n\}$ is linearly independent.

We need to show that $W = U \oplus V$. We show that for every $z$ in $W$ there is exactly one pair of vectors $x$ in $U$ and $y$ in $V$ such that $z = x + y$.

Indeed, since $Z$ is a basis in $W$, for every $z$ in $W$ there are unique scalars $a_1, \ldots, a_m$ and $b_1, \ldots, b_{n-m}$ such that $z = a_1x_1 + \cdots + a_mx_m + b_1y_1 + \cdots + b_{n-m}y_{n-m}$. Therefore, since $U \cap V = \{0\}$, there are uniquely defined vectors $x = a_1x_1 + \cdots + a_mx_m$ and $y = b_1y_1 + \cdots + b_{n-m}y_{n-m}$ such that $z = x + y$.

\[ \square \]

### 9 Isomorphisms

A bijection between two sets is an isomorphism of sets, i.e.: two isomorphic sets can at most differ in the naming of their elements. A bijection between vector spaces is not enough to consider the two vector spaces as differing at most by a renaming of their elements, the bijection must preserve the vector space structure. We elaborate

**Definition 9.1.** If $V$ and $W$ are vector spaces over the same field $F$, a mapping $T : V \rightarrow W$ is called a linear transformation if

1. $T(x + y) = T(x) + T(y)$ for all $x$ and $y$ in $V$; and
2. $T(ax) = aT(x)$ for all $x$ in $V$ all $a$ in $F$.

Condition (1) and (2) is equivalent to the following

- $T(ax + by) = aT(x) + bT(y)$ for all $x$ and $y$ in $V$ and all $a$ and $b$ in $F$.

Note that in the left-hand side of (1) the expression $ax + by$ refers to scalar multiplication and vector addition in $V$, while in the right-hand side of (1) the expression $aT(x) + bT(y)$ refers to scalar multiplication and vector addition in $W$. If, in addition to being linear, $T$ is also bijective, it is called a linear isomorphism, or a vector space isomorphism between $V$ and $W$, in which case the two spaces are said to be isomorphic.

Note that the inverse $T^{-1}$ of a linear isomorphism $T$ is linear. Indeed, the bijection $T$ has a bijective inverse $T^{-1}$. To see that it is linear, evaluate $T^{-1}(ax' + by')$. If $T(x) = x'$ and $T(y) = y'$ then $T(ax + by) = aT(x) + bT(y) = ax' + by'$, which implies

$$T^{-1}(ax' + by') = T^{-1}(T(ax + by)) = ax + by = aT^{-1}(x') + bT^{-1}(y').$$

**Remark 9.1.** While an isomorphism of sets is any bijection between the two sets, an isomorphism of vector spaces is a linear bijection between the two vector spaces.

**Remark 9.2.** Suppose $V$, $W$ and $Z$ are vector spaces over the same field. If $T_1 : V \rightarrow W$ and $T_2 : W \rightarrow Z$ are linear isomorphisms then $T_2 \circ T_1 : V \rightarrow Z$ is also a linear isomorphism.
Indeed, for all \( x \) and \( y \) in \( V \) and all scalars \( a \) and \( b \) we have
\[
(T_2 \circ T_1)(ax + by) = T_2(T_1(ax + by)) \\
= T_2(aT_1(x) + bT_1(y)) \\
= aT_2(T_1(x)) + bT_2(T_1(y)) \\
= a(T_2T_1)(x) + b(T_2T_1)(y)
\]

Let us, but just temporarily, write \( V \sim W \) to indicate that the vector spaces \( V \) and \( W \) (over the same field) are isomorphic. Then,

(i) \( V \sim V \) (take \( T : V \to V : x \to x \)),

(ii) \( V \sim W \Rightarrow W \sim V \) (given \( T : V \to W \), we have \( T^{-1} : W \to V \)), and

(iii) \( (V \sim W) \) and \( (W \sim Z) \Rightarrow V \sim Z \) (given \( T_1 : V \to W \) and \( T_2 : W \to Z \), we have \( T_2 \circ T_1 : V \to Z \)).

It is an equivalence relation over the set of all vector spaces.

**Theorem 9.1.** *Every n-dimensional vector space \( V \) over the field \( F \) is isomorphic to \( F^n \).*

**Proof.** Let \( V \) be an \( n \)-dimensional vector space over the field \( F \) and let
\[
X = \{x_1, \ldots, x_n\}
\]
be a basis in \( V \). Define \( T : V \to F^n \) as follows. Every \( x \) in \( V \) has a unique representation of the form
\[
x = a_1x_1 + \cdots + a_nx_n.
\]

Define
\[
T(x) = (a_1, \ldots, a_n).
\]

It is called the coordinate of \( x \) under the basis \( X \).

The mapping \( T \) is surjective, since for every \( (a_1, \ldots, a_n) \) in \( F^n \) there is some \( x \) in \( V \) such that \( T(x) = (a_1, \ldots, a_n) \). (In fact, \( x = a_1x_1 + \cdots + a_nx_n \).

The mapping \( T \) is also injective, since the representation \( x = a_1x_1 + \cdots + a_nx_n \) is unique. (Hence, \( T(x) \neq T(y) \) implies \( x \neq y \).)

Further, \( T \) is also linear. Indeed, if \( x = a_1x_1 + \cdots + a_nx_n \) and \( y = b_1x_1 + \cdots + b_nx_n \), then
\[
T(ax + by) = T((a_1x_1 + \cdots + a_nx_n) + (b_1x_1 + \cdots + b_nx_n)) \\
= T((a_1 + b_1)x_1 + \cdots + (a_n + b_n)x_n) \\
= ((a_1 + b_1), \ldots, (a_n + b_n)) \\
= (aa_1, \ldots, aa_n) + (bb_1, \ldots, bb_n) \\
= a(a_1, \ldots, a_n) + b(b_1, \ldots, b_n) \\
= aT(x) + bT(y).
\]
Theorem 9.2. Two finite dimensional spaces over a field $F$ are isomorphic if and only if the two spaces have the same dimension.

Exercises
1. Prove that $P_n(R)$ is isomorphic to $R^{n+1}$.
2. Consider $C$ as a vector space over $R$ by defining scalar multiplication as the usual multiplication of a real number by a complex number and by defining vector addition as the usual addition of complex numbers. Prove that (the vector space) $C$ is isomorphic to $R^2$.
3. Prove that if $X$ and $Y$ are subspaces of a vector space $V$ then $X \cup Y$ is a subspace of $V$ iff $X \subset Y$ or $Y \subset X$.
4. For any set $V$ the power set $P(V)$ of $V$ is the set of all subsets of $V$. If $V$ is a vector space, let $S(V)$ denote the set (a subsetof $P(V)$) of all subspaces of $V$. For a finite-dimensional vector space $V$ define

$$\sigma : P(V) \rightarrow S(V) : S \rightarrow \text{span } S.$$ 

(a) Prove that $\sigma(\sigma(S)) = \sigma(S)$ for every $S$ in $P(V)$.
(b) Prove that $(S_1 \subset S_2) \Rightarrow \sigma(S_1) \subset \sigma(S_2)$ for all $S_1$ and $S_2$ in $P(V)$. (c) Prove that $\sigma(S_1 \cup S_2) = \sigma(S_1) + \sigma(S_2)$ for all $S_1$ and $S_2$ in $P(V)$.

10 Linear Functionals

**Definition 10.1.** Let $V$ be a vector space over the field $F$. A linear functional on $V$ is a function $\lambda : V \rightarrow F$ such that for all vectors $x$ and $y$ and for all scalars $a$ and $b$

$$\lambda(ax + by) = a\lambda(x) + b\lambda(y).$$

The kernel of the linear functional $\lambda$ is the set $\ker \lambda = \{x \in V| \lambda(x) = 0\}$.

A hyperplane is a subspace that has dimension one less than the whole space.

**Examples**

**Ex. 1** On $C^n$, the vector space of all $n$-tuples $x = (\xi_1, \cdots, \xi_n)$ of complex numbers, the following are examples of linear functionals

(a) $\lambda(x) = 0$,
(b) $\lambda(x) = \xi_1$,
1. $\lambda(x) = a_1\xi_1 + \cdots + a_n\xi_n$, for scalars $a_1, \cdots, a_n$. 

Ex. 2 On $C(R)$, the vector space of all continuous functions $f : R \rightarrow R$, the following are examples of linear functionals

(a) $\lambda(f) = 0$,
(b) $\lambda(f) = f(0)$,
(c) $\lambda(f) = \int_a^b \alpha(t)f(t) dt$, for any bounded interval $(a, b)$ and any continuous a function $\alpha : [a, b] \rightarrow R$.

Remark 10.1. If $\lambda_1$ and $\lambda_2$ are linear functionals on a vector space $V$ over a field $F$, for any pair $\alpha$ and $\beta$ of scalars, define $\lambda : V \rightarrow F$ by

$$\lambda(x) = \alpha \lambda_1(x) + \beta \lambda_2(x)$$

for all $x$ in $V$. Then $\lambda$ is a linear functional, denoted by $\lambda = \alpha \lambda_1 + \beta \lambda_2$. Indeed,

$$\lambda(ax + by) = \alpha \lambda_1(ax + by) + \beta \lambda_2(ax + by)$$

$$= \alpha(a \lambda_1(x) + a \lambda_1(y)) + \beta(a \lambda_2(x) + b \lambda_2(y))$$

$$= a(\alpha \lambda_1(x) + \beta \lambda_2(x)) + b(\alpha \lambda_1(y) + \beta \lambda_2(y))$$

$$= a \lambda(x) + b \lambda(y)$$

for all vectors $x$ and $y$ and for all scalars $a$ and $b$. This suggests the definitions given below.

Definition 10.2. Given a vector space $V$ over the field $F$, denote by $V'$ the collection of all linear functionals on $V$. We make $V'$ into a vector space over $F$ by defining the following.

(1) The zero element of $V'$ is the linear functional 0, defined by $0(x) = 0$ for all $x$ in $V$.

(2) Scalar multiplication: For $\lambda$ in $V'$ and scalar $a$ define $a \lambda$ in $V'$ by

$$(a \lambda)(x) = a \lambda(x)$$

for all $x$ in $V$.

(3) Vector addition: For $\lambda_1$ and $\lambda_2$ in $V'$ define $\lambda_1 + \lambda_2$ in $V'$ by

$$(\lambda_1 + \lambda_2)(x) = \lambda_1(x) + \lambda_2(x)$$

for all $x$ in $V$. 

20
The vector space $V'$ is called the dual space of $V$. We write $\langle x, \lambda \rangle$ for $\lambda(x)$.

Remark 10.2. The reason for calling $V'$ the dual space of $V$ and for introducing the notation $\langle x, \lambda \rangle$ becomes transparent when we observe that

$$\langle ax_1 + bx_2, \lambda \rangle = a \langle x_1, \lambda \rangle + b \langle x_2, \lambda \rangle \quad (10.1)$$

and

$$\langle x, a\lambda_1 + b\lambda_2 \rangle = a \langle x, \lambda_1 \rangle + b \langle x, \lambda_2 \rangle. \quad (10.2)$$

In other words, we have a function $\langle \cdot, \cdot \rangle : V \times V' \rightarrow F$ that is linear in the first argument, i.e.: (10.1) holds, and is also linear in the second argument, i.e.: (10.2) holds. The function $\langle \cdot, \cdot \rangle$ is called a bilinear functional. Later we will see an obvious reason to adopt this notation.

11 Dual Bases

Theorem 11.1. Let $V$ be a vector space with basis $\{x_1, \cdots, x_n\}$. To every set $\{a_1, \cdots, a_n\}$ of scalars there corresponds a unique $\lambda$ in $V'$ such that

$$\langle x_i, \lambda \rangle = a_i \quad (11.1)$$

for $i = 1, \cdots, n$.

Proof. Suppose that $V$ is a vector space over the field $F$ and that $\{x_1, \cdots, x_n\}$ is a basis in $V$. Let $\{a_1, \cdots, a_n\}$ be a given set of scalars.

We show existence. Define $\lambda : V \rightarrow F$ as follows. Every vector $x$ in $V$ has the unique representation $x = b_1 x_1 + \cdots + b_n x_n$. Define

$$\langle x, \lambda \rangle = b_1 a_1 + \cdots + b_n a_n.$$

Then $\langle x_i, \lambda \rangle = a_i$ for $i = 1, \cdots, n$, i.e.: $\lambda$ satisfies (11.1).

We show that $\lambda$ is linear. Suppose that $y$ is also a vector in $V$. Then, $y$ has the unique representation $y = c_1 x_1 + \cdots + c_n x_n$. Hence, for scalars $a$ and $b$, the vector $ax + by$ has the unique representation

$$ax + by = a(b_1 x_1 + \cdots + b_n x_n) + b(c_1 x_1 + \cdots + c_n x_n) = (ab_1 + bc_1)x_1 + \cdots + (ab_n + bc_n)x_n.$$

Therefore,

$$\langle ax + by, \lambda \rangle = (ab_1 + bc_1)a_1 + \cdots + (ab_n + bc_n)a_n$$

$$= a(b_1 a_1 + \cdots + b_n a_n) + b(c_1 a_1 + \cdots + c_n a_n) = a \langle x, \lambda \rangle + b \langle y, \lambda \rangle,$$
establishing the linearity of $\lambda$.

We show uniqueness. Let $\lambda_1$ and $\lambda_2$ be linear functionals on $V$ that satisfy (11.1). Every $x$ in $V$ has the unique representation of the form $x = b_1x_1 + \cdots + b_nx_n$. Hence, thanks to (11.1),

$$\langle x, \lambda_1 \rangle - \langle x, \lambda_2 \rangle = (b_1\langle x_1, \lambda_1 \rangle + \cdots + b_n\langle x_n, \lambda_1 \rangle) - (b_1\langle x_1, \lambda_2 \rangle + \cdots + b_n\langle x_n, \lambda_2 \rangle)$$

$$= (b_1a_1 + \cdots + b_na_n) - (b_1a_1 + \cdots + b_na_n)$$

$$= 0$$

Therefore, (11.1) implies $\langle x, \lambda_1 \rangle = \langle x, \lambda_2 \rangle$ for every vector $x$ in $V$, i.e.: $\lambda_1 = \lambda_2$.

\textbf{Theorem 11.2. (Dual Basis)} Let $V$ be an $n$-dimensional vector space. To every basis $X = \{x_1, \cdots, x_n\}$ in $V$ there corresponds exactly one basis $X' = \{\lambda_1, \cdots, \lambda_n\}$ in $V'$ such that $\langle x_i, \lambda_j \rangle = 0$ for $i \neq j$ and $\langle x_i, \lambda_i \rangle = 1$.

\textbf{Proof.} Thanks to the previous proposition, for each $j$, $1 \leq j \leq n$, there is a unique $\lambda_j$ in $V'$ such that $\langle x_i, \lambda_j \rangle = 0$ for $i \neq j$ and $\langle x_i, \lambda_i \rangle = 1$. Define $X' = \{\lambda_1, \cdots, \lambda_n\}$. We prove that $X'$ spans $V'$. Given any $\lambda$ in $V'$, define $a_i = \langle x_i, \lambda \rangle$ for $i = 1, \cdots, n$. Then, since every $x$ in $V$ is of the form $x = b_1x_1 + \cdots + b_nx_n$, we have

$$\langle x, \lambda \rangle = \langle b_1x_1 + \cdots + b_nx_n, \lambda \rangle$$

$$= b_1\langle x_1, \lambda \rangle + \cdots + b_n\langle x_n, \lambda \rangle$$

$$= b_1a_1 + \cdots + b_na_n$$

and, for $i = 1, \cdots, n$,

$$\langle x, \lambda_i \rangle = \langle b_1x_1 + \cdots + b_nx_n, \lambda_i \rangle$$

$$= b_1\langle x_1, \lambda_i \rangle + \cdots + b_n\langle x_n, \lambda_i \rangle$$

$$= b_i.$$

Hence,

$$\langle x, \lambda \rangle = a_1\langle x, \lambda_1 \rangle + \cdots + a_n\langle x, \lambda_n \rangle$$

$$= \langle x, a_1\lambda_1 + \cdots + a_n\lambda_n \rangle,$$

which holds for every $x$ in $V$. Therefore, $\lambda = a_1\lambda_1 + \cdots + a_n\lambda_n$. This shows that $\{\lambda_1, \cdots, \lambda_n\}$ spans $V'$.

We prove that $X'$ is linearly independent. Suppose that

$$a_1\lambda_1 + \cdots + a_n\lambda_n = 0. \quad (11.2)$$

22
Then, for every $x$ in $V$
\[
\langle x, a_1 \lambda_1 + \cdots + a_n \lambda_n \rangle = a_1 \langle x, \lambda_1 \rangle + \cdots + a_n \langle x, \lambda_n \rangle.
\]
Hence, for $x = x_i$ the above yields
\[
0 = a_1 \langle x_i, \lambda_1 \rangle + \cdots + a_n \langle x_i, \lambda_n \rangle = a_i \langle x_i, \lambda_i \rangle = a_i.
\]
The above holds for $i = 1, \ldots, n$. Hence, (11.2) implies $a_1 = \cdots = a_n = 0$. This shows that $\{\lambda_1, \ldots, \lambda_n\}$ is linearly independent. \hfill \qed

**Corollary 11.3.** The dimension of a finite-dimensional vector space and the dimension of its dual space are the same.

**Proof.** $X$ and $X'$ have the same number of elements. \hfill \qed

**Theorem 11.4.** Let $V$ be a finite-dimensional vector space. For every nonzero $x$ in $V$ there exists a linear functional $\lambda$ on $V$ such that $\langle x, \lambda \rangle \neq 0$.

**Proof.** Let $V$ be an $n$-dimensional vector space and let $x$ be any vector in $V$. There is a basis $X = \{x_1, \ldots, x_n\}$ in $V$ and a dual basis $X' = \{\lambda_1, \ldots, \lambda_n\}$ in $V'$. The vector $x$ has the unique representation $x = a_1 x_1 + \cdots + a_n x_n$ and we have $\langle x, \lambda_i \rangle = a_i$ for $i = 1, \ldots, n$. Therefore, with $N = \{1, \ldots, n\}$,
\[
(\forall \lambda \in V')(\langle x, \lambda \rangle = 0) \Rightarrow (\forall i \in N)(\langle x, \lambda_i \rangle = 0) \Rightarrow (\forall i \in N)(a_i = 0) \Rightarrow x = 0.
\]
Then, by contraposition,
\[
x \neq 0 \Rightarrow (\exists i \in N)(a_i \neq 0) \Rightarrow (\exists i \in N)(\langle x, \lambda_i \rangle \neq 0) \Rightarrow (\exists \lambda \in V')(\langle x, \lambda \rangle \neq 0).
\]

**Corollary 11.5.** If $V$ is a finite-dimensional vector space then for any two distinct vectors $x$ and $y$ in $V$ there exists a linear functional $\lambda$ such that $\langle x, \lambda \rangle \neq \langle y, \lambda \rangle$.

**Proof.** Let $V$ be a finite-dimensional vector space and let $x$ and $y$ be elements of $V$. If $x \neq y$ then $x - y \neq 0$ and there exists a $\lambda$ in $V'$ such that $\langle x - y, \lambda \rangle \neq 0$). Therefore, $\langle x, \lambda \rangle \neq \langle y, \lambda \rangle$. \hfill \qed

**Exercises**

1. Define a nonzero $\lambda \in (R^3)'$ such that $\langle x_1, \lambda \rangle = \langle x_2, \lambda \rangle = 0$ for the two vectors $x_1 = (1, -1, 1)$ and $x_2 = (-1, 1, 1)$.
2. Define $\lambda$ in $(C^n)'$ by $\langle x, \lambda \rangle = \xi_0 + \cdots + \xi_n$ for all vectors $x = (\xi_0, \cdots, \xi_n)$ in $C^n$. Find a basis for the subspace $X = \{x \in C^n | \langle x, \lambda \rangle = 0\}$.
3. Consider the basis $X = \{(1, 0, 0), (1, 1, 0), (0, 1, 1)\}$ in $C^3$. What is the dual basis $X'$?
12 The Second Dual

The dual $V'$ of a vector space $V$ is also a vector space. Therefore, $V'$ has a dual space $(V')'$, which we shall simply denote by $V''$. The vector space $V''$ is called the second dual of $V$. So far, the only thing we know about the second dual is that, since a finite-dimensional vector space and its dual have the same dimension, is that if $V$ is finite-dimensional then $\dim V'' = \dim V' = \dim V$.

Now, $V$, $V'$ and $V''$ are vector spaces over the same field and, if $V$ is finite dimensional, all three spaces have the same dimension. Hence, they are isomorphic. However, this does not reveal a much deeper connection between the elements of $V$ and those of $V''$, a correspondence we shall call “natural isomorphism” between a finite-dimensional vector space and its second dual.

Let $V$ be a finite-dimensional vector space over the field $F$. For each $x$ in $V$ define $\mu_x = \langle x, \cdot \rangle$, i.e.: define

$$\mu_x : V' \rightarrow F : \lambda \rightarrow \langle x, \lambda \rangle.$$

Then, for any two elements $\lambda_1$ and $\lambda_2$ of $V'$ and scalars $a$ and $b$,

$$\mu_x(a\lambda_1 + b\lambda_2) = \langle x, a\lambda_1 + b\lambda_2 \rangle = a\langle x, \lambda_1 \rangle + b\langle x, \lambda_2 \rangle = a\mu_x(\lambda_1) + b\mu_x(\lambda_2).$$

We have defined a linear mapping from $V'$ to $F$. Therefore, $\langle x, \cdot \rangle$ is an element of $V''$. In other words, $\langle x, \cdot \rangle$ is a linear functional on $V'$ for every $x$ in $V$. Is every functional $\mu$ on $V'$ of the form $\mu = \langle x, \cdot \rangle$ for some $x$ in $V$? We answer this question below.

**Theorem 12.1.** If $V$ is a finite-dimensional vector space then $\Phi : V \rightarrow V'' : x \rightarrow \langle x, \cdot \rangle$ is a vector space isomorphism (the natural isomorphism) between $V$ and its second dual $V''$. In other words, to every $\mu$ in $V''$ there corresponds exactly one $x$ in $V$ such that $\mu(\lambda) = \langle x, \lambda \rangle$ for all $\lambda$ in $V'$.

**Proof.** Suppose $V$ is a finite-dimensional vector space. Then, for every $x$ and $y$ in $V$ and for all scalars $a$ and $b$,

$$\Phi(ax + by)(\lambda) = \langle ax + by, \lambda \rangle = a\langle x, \lambda \rangle + b\langle y, \lambda \rangle = a\Phi(x)(\lambda) + b\Phi(y)(\lambda)$$

for all $\lambda$ in $V'$, i.e.: $\Phi(ax + by) = a\Phi(x) + b\Phi(y)$. This shows that $\Phi$ is linear.

We show that $\Phi$ is injective. Suppose that $x$ and $y$ are in $V$. Then $\Phi(x)(\lambda) = \langle x, \lambda \rangle$ and $\Phi(y)(\lambda) = \langle y, \lambda \rangle$ for all $\lambda$ in $V'$. Hence,

$$\Phi(x) = \Phi(y) \Rightarrow (\forall \lambda \in V')(\Phi(x)(\lambda) = \Phi(y)(\lambda))$$

$$\Rightarrow (\forall \lambda \in V')(\Phi(x - y)(\lambda) = 0))$$

$$\Rightarrow x = y.$$
We show that $\Phi$ is surjective. Let $W$ denote the range of $\Phi$, i.e.: the subset of $V''$ consisting of all elements of the form $\langle x, \cdot \rangle$ for some $x$ in $V$. Since $\langle 0, \cdot \rangle$ is in $W$ and, for all scalars $a$ and $b$, $\langle ax + by, \cdot \rangle$ is also in $W$ whenever $\langle x, \cdot \rangle$ and $\langle y, \cdot \rangle$ are in $W$, it follows that $W$ is a subspace of $V''$. By the injectivity of $\Phi$, this function is an isomorphism between $V$ and $W$. Hence, $W$ is a subspace of $V''$ with the same dimension as $V''$. It follows that $W = V''$, i.e.: that every $\mu$ in $V''$ is of the form $\mu = \langle x, \cdot \rangle$ for some $x$ in $V$.

**Remark 12.1.** Remark A finite-dimensional vector space that is isomorphic to its second dual is called reflexive. We have shown that every finite-dimensional vector space is reflexive. Since every finite-dimensional vector space and its dual have the same dimension, the two are isomorphic. Why has this isomorphism not been talked about and, instead, we elaborated on the isomorphism between a finite-dimensional vector space and its second dual? The reason is that there is always a natural isomorphism between a finite-dimensional vector space and its second dual, while a corresponding “natural isomorphism” does not always exist between a finite-dimensional vector space and its (first) dual. Later we shall consider the proper setting where a “natural isomorphism” can be constructed between some vector spaces and their dual spaces. It is commonplace to blur the difference between mathematical structures that are different, but isomorphic in some natural way. Mathematics is full of such abuses of language. Accordingly, we identify any finite-dimensional vector space $V$ with its second dual space. In other words, we blur the difference between $x$ in $V$ and $\langle x, \cdot \rangle$ in $V''$. To put it even more specifically, every $x$ in $V$ “is a linear functional” on $V'$ via $\lambda \mapsto \langle x, \lambda \rangle$ and every functional $\mu$ on $V''$ (which corresponds a unique $x$ in $V$ such that: $\mu(\lambda) = \langle x, \lambda \rangle$ for all $\lambda$ in $V'$) “is a vector” $x$ in $V$. An obvious consequence of this identification is that any basis in a finite-dimensional vector space $V$ is also a basis in $V''$.

If the last statement above seems to be too abstract, consider this. Every $\mu$ in $V''$ is a linear functional on $V'$, i.e.: 

$$\mu(\lambda) = \langle x, \lambda \rangle,$$ 

for all $\lambda$ in $V'$, where $x$ in $V$ is uniquely determined by $\mu$. If $X = \{x_1, \cdots, x_n\}$ is a basis in $V$ and $x = a_1x_1 + \cdots + a_nx_n$, then, thanks to (12.1), by linearity, 

$$\mu(\lambda) = \langle a_1x_1 + \cdots + a_nx_n, \lambda \rangle = a_1\langle x_1, \lambda \rangle + \cdots + a_n\langle x_n, \lambda \rangle,$$

for all $\lambda$ in $V'$. Therefore, for the given $\mu$ 

$$\mu = a_1\langle x_1, \cdot \rangle + \cdots + a_n\langle x_n, \cdot \rangle,$$

showing that $\{\langle x_1, \cdot \rangle, \cdots, \langle x_n, \cdot \rangle\}$, which we identify with $\{x_1, \cdots, x_n\}$, is a basis in $V''$. 

25
13 Annihilators

Definition 13.1. Let $V$ be a vector space. Given any subset $S$ of $V$ the annihilator $S^0$ of $S$ is defined by

$$S^0 = \{ \lambda \in V' \mid \langle \forall x \in S \rangle(\langle x, \lambda \rangle = 0) \}.$$ 

In other words, the annihilator $S^0$ of $S$ is the collection of all linear functionals on $V$ that vanish at every $x$ in $S$.

Remark 13.1. If $V$ is a vector space, then $\{0\}^0 = V'$ and $V^0 = \{0\}$, the singleton of the zero functional on $V$. If $V$ is finite-dimensional and $S$ contains a nonzero vector $x$ then there is a linear functional $\lambda$ on $V$ such that $\langle x, \lambda \rangle \neq 0$. Hence, $S^0$ cannot contain all linear functionals on $V$, showing that $S^0$ must be a proper subset of $V'$.

Remark 13.2. $S^0$ is always a vector space (a subspace of $V'$, of course), even when $S$ is not a subspace of $V$. Indeed, if $\langle x, \lambda_1 \rangle = 0$ and $\langle x, \lambda_2 \rangle = 0$ for all $x$ in $S$, then $0 = a\langle x, \lambda_1 \rangle + b\langle x, \lambda_2 \rangle = \langle x, a\lambda_1 + b\lambda_2 \rangle$ for all $x$ in $S$, showing that $\lambda_1 \in S^0$ and $\lambda_2 \in S^0 \Rightarrow a\lambda_1 + b\lambda_2 \in S^0$ for all scalars $a$ and $b$.

Theorem 13.1. Suppose $V$ is a finite-dimensional vector space. If $X$ is a subspace of $V$ then $\dim X^0 = \dim V - \dim X$.

Proof Suppose $V$ is an $n$-dimensional vector space and $X$ is a subspace of $V$. Let $Y = \{y_1, \ldots, y_n\}$ be a basis in $V$ such that $\{y_1, \ldots, y_m\}$ spans $X$. Consider the dual basis $Y' = \{\lambda_1, \ldots, \lambda_n\}$ in $V'$ and let $Z$ denote the subspace of $V'$ spanned by $\{\lambda_{m+1}, \ldots, \lambda_n\}$. We have $\dim V = n$, $\dim X = m$ and $\dim Z = n - m$. We prove that $Z = X^0$ to establish the proposition.

We prove that $Z \subset X^0$. Every $x$ in $X$ is of the form $x = a_1y_1 + \cdots + a_my_m$. Hence, for every $j$, $m + 1 \leq j \leq n$, and for all $x$ in $X$ we have

$$\langle x, \lambda_j \rangle = \langle a_1y_1 + \cdots + a_my_m, \lambda_j \rangle = a_1\langle y_1, \lambda_j \rangle + \cdots + a_m\langle y_m, \lambda_j \rangle = 0,$$

since $\langle y_i, \lambda_j \rangle = 0$ for $1 \leq i \leq m$ and $m + 1 \leq j \leq n$ (i.e.: $i \neq j$). We have shown that $\lambda_j \in X^0$ for $j = m + 1, \ldots, n$. Hence, any linear combination of $\lambda_{m+1}, \ldots, \lambda_n$ also belongs to $X^0$, i.e.: $Z = \text{span} \{\lambda_{m+1}, \ldots, \lambda_n\} \subset X^0$.

We prove that $X^0 \subset Z$. Every $\lambda$ in $X^0$ is of the form $\lambda = b_1\lambda_1 + \cdots + b_n\lambda_n$ (as is every element of $V'$). Further, $\lambda$ in $X^0$ implies that $\lambda$ vanishes on every element of
and, in particular, on the elements \( y_1, \ldots, y_m \). Hence, for \( i = 1, \ldots, m \),

\[
0 = \langle y_i, \lambda \rangle \\
= \langle y_i, b_1 \lambda_1 + \cdots + b_n \lambda_n \rangle \\
= b_1 \langle y_i, \lambda_1 \rangle + \cdots + b_n \langle y_i, \lambda_n \rangle \\
= b_i,
\]

showing that

\[
\lambda = b_1 \lambda_1 + \cdots + b_m \lambda_m + b_{m+1} \lambda_{m+1} + \cdots + b_n \lambda_n,
\]
i.e.: \( \lambda \in \text{span} \{ \lambda_{m+1}, \ldots, \lambda_n \} = Z \). We proved that \( \lambda \in X^0 \) implies \( \lambda \in Z \), i.e.: \( X^0 \subset Z \).

**Remark 13.3.** Since we identify every finite-dimensional vector space with its second dual, does it follow that the annihilator of an annihilator of a subspace can be identified with that subspace? Below, we show that the answer is affirmative. First, let us agree to write \( S^{00} \) for \((S^0)^0\).

**Theorem 13.2.** Let \( V \) be a finite-dimensional vector space. If \( X \) is any subspace of \( V \) then \( X^{00} = X \).

**Proof.** Let \( V \) be an n-dimensional vector space and let \( X \) be a subspace of \( V \). Then, by identifying \( V \) with its second dual,

\[
X^{00} = \{ x \in V | (\forall \lambda \in X^0)(\langle x, \lambda \rangle = 0) \}.
\]

Recall the definition of

\[
X^0 : = \{ \lambda \in V' | (\forall x \in X)(\langle x, \lambda \rangle = 0) \}.
\]

Hence, \( (\forall x \in X)(\forall \lambda \in X^0)(\langle x, \lambda \rangle = 0) \), which implies

\[
x \in X \Rightarrow (\forall \lambda \in X^0)(\langle x, \lambda \rangle = 0),
\]

which says \( x \in X \Rightarrow x \in X^{00} \), i.e.: \( X \subset X^{00} \).

Can \( X \) be a proper subset of \( X^{00} \)? Let’s count dimensions. We have

\[
\dim X^0 = \dim V - \dim X,
\]

\[
\dim X^{00} = \dim V' - \dim X^0,
\]

\[
\dim V = \dim V'.
\]

Hence, \( \dim X = \dim V - \dim X^0 = \dim V' - \dim X^0 = \dim X^{00} \). The only subspace of \( X^{00} \) with the same dimension as \( X^{00} \) is \( X^{00} \). Hence, \( X^{00} = X \).
Theorem 13.3. Let $U$ and $V$ be subspaces of a vector space $W$. If $W = U \oplus V$ then the following hold.

1) The annihilator $U^0$ of $U$ and the annihilator $V^0$ of $V$ are disjoint, i.e.: $U^0 \cap V^0 = \{0\}$.

2) $W' = U^0 \oplus V^0$.

3) The dual space $V'$ of $V$ is isomorphic to the annihilator $U^0$ of $U$.

4) The dual space $U'$ of $U$ is isomorphic to the annihilator $V^0$ of $V$.

Proof. We prove (1). Recall that 
\[ \lambda \in U^0 \iff (\forall x \in U)(\langle x, \lambda \rangle = 0) \]
and 
\[ \lambda \in V' \iff (\forall y \in V)(\langle y, \lambda \rangle = 0) \].
Suppose that $\lambda \in U^0 \cap V^0$. Then, $\langle x, \lambda \rangle = 0$ for all $x$ in $U$ and $\langle y, \lambda \rangle = 0$ for all $y$ in $V$. Hence, since every $z$ in $W$ has the unique representation $z = x + y$, with $x$ in $U$ and $y$ in $V$, we have
\[ \langle z, \lambda \rangle = \langle x + y, \lambda \rangle = \langle x, \lambda \rangle + \langle y, \lambda \rangle = 0. \]
We have shown that $\lambda \in U^0 \cap V^0 \Rightarrow \lambda = 0$, i.e.: $U^0 \cap V^0 = \{0\}$.

We prove (2). Corresponding to every $\lambda$ in $W'$ we define the linear functionals $\lambda_1$ and $\lambda_2$ on $W$ as follows. Every $z$ in $W$ has the unique representation $z = x + y$, with $x$ in $U$ and $y$ in $V$. Define $\lambda_1$ by $\langle z, \lambda_1 \rangle = \langle y, \lambda \rangle$ and define $\lambda_2$ by $\langle z, \lambda_2 \rangle = \langle x, \lambda \rangle$. We claim that $\lambda_1 \in U^0$ and $\lambda_2 \in V^0$. Indeed (keeping in mind the unique representation $z = x + y$),
\[ z \in U \Rightarrow \langle z, \lambda_1 \rangle = \langle 0, \lambda \rangle = 0 \Rightarrow \lambda_1 \in U^0, \]
and
\[ z \in V \Rightarrow \langle z, \lambda_2 \rangle = \langle 0, \lambda \rangle = 0 \Rightarrow \lambda_2 \in V^0. \]
On the other hand, for every $z$ in $W$
\[ \langle z, \lambda \rangle = \langle x + y, \lambda \rangle = \langle x, \lambda \rangle + \langle y, \lambda \rangle = \langle z, \lambda_1 \rangle + \langle z, \lambda_2 \rangle = \langle z, \lambda_1 + \lambda_2 \rangle, \]
i.e.: $\lambda = \lambda_1 + \lambda_2$, showing that $W' = U^0 + V^0$. Hence, since $U^0 \cap V^0 = \{0\}$, $W' = U^0 \oplus V^0$.

We prove (3). We need a linear isomorphism between $V'$ and $U^0$. In the mapping $W' \to U^0 : \lambda \to \lambda_1$ defined above, $\lambda_1$ is defined by
\[ \langle z, \lambda_1 \rangle = \langle y, \lambda \rangle \quad (13.1) \]
(Keep in mind the unique representation \( z = x + y \).) If we consider all \( z \) of the form \( z = y \) (all \( z \) in \( V \)), then the rule given by (13.1) reduces to \( \langle y, \lambda_1 \rangle = \langle y, \lambda \rangle \) and defines a bijection between \( V' \) and \( U^0 \).

To establish that this bijection is linear, suppose that \( \mu \) is a functional on \( V \) (so that \( \mu_1 \) in \( U^0 \) is defined by \( \langle y, \mu_1 \rangle = \langle y, \mu \rangle \)). Then, for all scalars \( a \) and \( b \),

\[
\langle y, a\lambda + b\mu \rangle = a\langle y, \lambda \rangle + b\langle y, \mu \rangle = a\langle y, \lambda_1 \rangle + b\langle y, \mu_1 \rangle,
\]

showing that \( a\lambda + b\mu \rightarrow a\lambda_1 + b\mu_1 \). We have obtained the desired linear isomorphism between \( U' \) and \( V^0 \).

The proof of (4) is carried out by making obvious modifications to the proof of (3). \( \square \)

**Remark 13.4.** The above proof uses neither coordinates nor finite-dimensional arguments. The vector spaces need not be finite-dimensional.

**Exercises**

1. Prove that \( S^{00} = \text{span} \ S \) for any subset \( S \) of a finite-dimensional vector space.

2. Suppose that \( S_1 \) and \( S_2 \) are subsets of a vector space \( V \). Prove that if \( S_1 \subset S_2 \) then \( S_2^0 \subset S_1^0 \).

3. Suppose \( X \) and \( Y \) are subspaces of a finite-dimensional vector space.
   (a) Prove that \( (X \cap Y)^0 = X^0 + Y^0 \).
   (b) Prove that \( (X + Y)^0 = X^0 \cap Y^0 \).

4. Show that the cancellation law \( X \bigoplus Y = X \bigoplus Z \Rightarrow Y = Z \) does not always apply to the direct sum of subspaces. In other words, give an example of a vector space \( V \) with subspaces \( X, Y \) and \( Z \) such that \( V = X \bigoplus Y = X \bigoplus Z \) and \( Y \neq Z \).

5. Suppose \( U, V \) and \( W \) are vector spaces.
   (a) Does \( U \bigoplus (V \bigoplus W) = (U \bigoplus V) \bigoplus W \) hold? If not, is there a linear isomorphism between the two sides of the equation?
   (b) Does \( U \bigoplus V = V \bigoplus U \) hold? If not, is there a linear isomorphism between the two sides of the equation?

6. The subspaces \( X, Y \) and \( Z \) of the vector space \( V \) are said to be independent if the sum of any two is disjoint from the third. Further, define \( X + Y + Z \) in the obvious way:

\[
X + Y + Z = \{x + y + z \in V \mid x \in X, y \in Y \text{ and } z \in Z\}.
\]
(a) Suppose that \( x, y \) and \( z \) are vectors in \( V \). Further, suppose that \( X = \text{span}\{x\}, Y = \text{span}\{y\} \) and \( Z = \text{span}\{z\} \). Prove that \( \{x, y, z\} \) is linearly independent iff \( X, Y \) and \( Z \) are independent.

(b) Prove that if \( V = X \bigoplus (Y \bigoplus Z) \) then \( V = X + Y + Z \) and the subspaces are independent. Do the same for \( V = (X \bigoplus Y) \bigoplus Z \).

(c) If \( V = X + Y + Z \) and the subspaces \( X, Y, Z \), are independent, does it follow that i. \( V = X \bigoplus (Y \bigoplus Z) \)? ii. \( V = (X \bigoplus Y) \bigoplus Z \)?

14 Quotient Sets

**Definition 14.1.** A binary relation \( \simeq \) on a set \( S \) is a subset of \( S \times S \). We write \( x \simeq y \) for \((x, y) \in \simeq \). The binary relation \( \simeq \) is called an equivalence relation if it is

1. reflexive: \( x \simeq x \) for all \( x \) in \( S \),
2. symmetric: \( x \simeq y \Rightarrow y \simeq x \) for all \( x \) and \( y \) in \( S \),
3. transitive: \( x \simeq y \) and \( y \simeq z \Rightarrow x \simeq z \) for all \( x \), \( y \), and \( z \) in \( S \).

If \( \simeq \) is an equivalence relation on \( S \) we define the function \( \phi_{\simeq} : S \rightarrow \mathcal{P}(S) \) by \( \phi_{\simeq}(x) = \{y \in S| x \simeq y\} \), where \( \mathcal{P}(S) \) denotes the power set of \( S \), the set of all subsets of \( S \). The range of \( \phi_{\simeq} \) is called the quotient set of \( S \) modulo \( \simeq \) and is denoted by \( S/\simeq \). We define the projection \( p_{\simeq} : S \rightarrow S/\simeq : x \rightarrow \phi_{\simeq}(x) \). Clearly, \( p_{\simeq} \) is surjective and \( S/\simeq \) does not contain the empty set, since \( x \in p_{\simeq}(x) \) for every \( x \) in \( S \) (\( x \simeq x \) for every \( x \) in \( S \)). The set \( p_{\simeq}(x) \) is called the equivalence class of \( x \). In other words, \( S/\simeq = \{p_{\simeq}(x) \in \mathcal{P}(S)| x \in S\} \) is the set of all equivalence classes defined by \( \simeq \) on \( S \).

A partition \( P \) of a set \( S \) is any collection of disjoint nonempty subsets of \( S \) whose union is \( S \).

**Theorem 14.1.** If \( \simeq \) is an equivalence relation on the set \( S \) then the elements of \( S/\simeq \) are disjoint.

Proof If \( z \in p_{\simeq}(x) \) and \( z \in p_{\simeq}(y) \) then \( x \simeq z \) and \( y \simeq z \). Hence, thanks to symmetry, \( x \simeq z \) and \( z \simeq y \). Therefore, thanks to transitivity, \( x \simeq y \), which implies \( p_{\simeq}(x) = p_{\simeq}(y) \). Hence, \( p_{\simeq}(x) \neq p_{\simeq}(y) \) implies \( p_{\simeq}(x) \cap p_{\simeq}(y) = \emptyset \).

**Corollary 14.2.** If \( \simeq \) is an equivalence relation on the set \( S \) then \( S/\simeq \) is a partition of \( S \).

Proof We already know that \( S/\simeq \) is a collection of nonempty disjoint subsets of \( S \). It remains to be shown that \( S \) is the union of the elements in \( S/\simeq \). Indeed, every \( x \) in \( S \) is an element of \( p_{\simeq}(x) \). Hence, \( S \subset \cup_{x \in S}p_{\simeq}(x) \). On the other hand, \( p_{\simeq}(x) \) is a subset of \( S \). Hence, \( \cup_{x \in S}p_{\simeq}(x) \subset S \).
Theorem 14.3. Every partition of a set defines an equivalence relation on that set.

Proof Let $P$ be a partition of $S$. Then, since the union of the elements of $P$ equals $S$, every $x$ in $S$ belongs to some $X$ in $P$. Moreover, since the elements of $P$ are disjoint, this $X$ is unique.

Define the binary relation $\simeq$ on $S$ by

$$x \simeq y \iff (\exists X \in P)(y \in X \land x \in X).$$

Then, $x \simeq x$ for every $x$ in $S$. Further, the symmetry of the definition implies that if $x \simeq y$ then $y \simeq x$. Finally, if $x \simeq y$ and $y \simeq z$ then $x$ and $y$ belong to some $X$ in $P$, and $y$ and $z$ belong to some $Y$ in $P$. Since $z$ belongs to exactly one set in $P$, it follows that $X = Y$ and, hence, that $x \simeq z$.

Theorem 14.4. If $\simeq$ is an equivalence relation on the set $S$ then to every function $f : S \to X$ such that $x \simeq y \Rightarrow f(x) = f(y)$ there corresponds exactly one function $g : S/\simeq \to X$ such that $f = g \circ p\simeq$, i.e.: such that the diagram below commutes.

$$
\begin{array}{ccc}
S & \xrightarrow{p\simeq} & S/\simeq \\
\downarrow f & & \downarrow g \\
& X \\
\end{array}
$$

Moreover, if $f$ is surjective and $f(x) = f(y) \Rightarrow x \simeq y$ then $g$ is bijective.

Proof Suppose $\simeq$ is an equivalence relation on the set $S$ and $f : S \to X$ is such that $x \simeq y \Rightarrow f(x) = f(y)$. Then $f$ is constant on the elements of each equivalence class $p\simeq(x)$, i.e.: if $f(x) = c$ then $f(y) = c$ for all $y \in p\simeq(x)$. Define $g : S/\simeq \to X$ by $g(p\simeq(x)) = f(x)$. Then,

$$(g \circ p\simeq)(x) = g(p\simeq(x)) = f(x).$$

If $f$ is surjective then so is $g$. Suppose $f(x) = f(y) \Rightarrow x \simeq y$. Then $g(p\simeq(x)) = g(p\simeq(y))$ implies $f(x) = f(y)$, which implies $x \simeq y$, i.e.: $p\simeq(x) = p\simeq(y)$, showing that $g$ is injective.

15 Quotient Spaces

Definition 15.1. (Cosets) Let $U$ be a subspace of a vector space $V$ and let $v$ be a vector in $V$. Define the coset (also called affine subspace) $v + U$ of $U$ by

$$v + U = \{v + x \in V \mid x \in U\}.$$ 

We say that $S$ is a coset of $U$ if $S = v + U$ for some vector $v$. 

31
Remark 15.1. Note that $U$ is a coset of $U$, since $U = 0 + U$. Note that $v + U = u + U$ does not imply $v = u$, i.e.: it is possible that $v + U = u + U$ and $v \neq u$, as shown in Figure 2 for the case $V = R^2$. In $R^2$, if $U$ is a line containing the origin and $v$ is any vector then $v + U$ is a line parallel to $U$. The collection of all cosets of $U$ is the collection of all lines parallel to $U$ (including, of course, $U$).

Definition 15.2. Define $X + Y = \{x + y \in V \| x \in X \land y \in Y\}$ for any two cosets $X$ and $Y$ of any subspace $U$ of the vector space $V$.

Theorem 15.1. Let $U$ be a subspace of a vector space $V$. If $X$ and $Y$ are cosets of $U$ so is $X + Y$.

Proof Suppose $U$ is a subspace of a vector space $V$. If $X$ and $Y$ are cosets of $U$ then $X = u + U$ and $Y = v + U$ for some vectors $u$ and $v$ in $V$.

We show that $X + Y \subset (u + v) + U$. Each vector $x_1$ in $X$ is of the form $u + z_1$ for some $z_1$ in $U$ and each vector $x_2$ in $Y$ is of the form $x_2 = v + z_2$ for some $z_2$ in $U$. Therefore, each vector $x$ in $X + Y$ is of the form $x = (u + v) + (z_1 + z_2)$. Since $U$ is a vector space, $z_1 + z_2$ is in $U$. Hence, $x = (u + v) + z$ for some $z$ in $U$, i.e.: $x$ belongs to the coset $(u + v) + U$ of $U$.

We show that $(u + v) + U \subset X + Y$. Every vector $x$ in $(u + v) + U$ is of the form $x = (u + v) + z$ for some $z$ in $U$ and, therefore, of the form $x = (u + z) + (v + 0)$, with $z$ and $0$ both in $U$, i.e.: $u + z$ is in $u + U$ (which is $X$)and $v + 0$ is in $v + U$ (which is $Y$). Hence, $x$ is in $X + Y$.

Remark 15.2. The above proposition establishes that

$$(u + U) + (v + U) = (u + v) + U = (u + v) + (U + U)$$

for any two cosets $u + U$ and $v + U$ of $U$. (Note that $U + U = U$, since $U$ is a subspace.

Corollary 15.2. Coset addition is associative and commutative.

Proof Commutativity is an immediate consequence of the definition of coset addition. We show associativity. Consider any three cosets $X$, $Y$ and $Z$ of $U$. Then $X = u + U$, $Y = v + U$ and $Z = w + U$ for some vectors $u$, $v$ and $w$ in $V$. Then, since vector addition is associative and $U = U + U = (U + U) + U = U + (U + U)$,

$$(X + Y) + Z = ((u + U) + (v + U)) + (w + U)$$
$$= (u + v + (U + U)) + (w + U)$$
$$= ((u + v) + w) + (U + U)$$
$$= (u + (v + w)) + (U + U)$$
$$= (u + U) + ((v + U) + (w + U))$$
$$= X + (Y + Z)$$
**Definition 15.3.** Let $U$ be a subspace of a vector space $V$. We make the set of all cosets of $U$ into a vector space as follows.

- Coset addition is already defined and is associative and commutative.
- Recall that $U = 0 + U$ is a coset of $U$. For any coset $X$ of $U$ we have $X + U = X$, i.e.: $U$ is the origin.
- Scalar multiplication is defined as follows. For a scalar $a$, $a \neq 0$, define $aX$ as the set of all vectors $ax$ with $x$ in $X$. Define $0X = U$.
- Define $-X$ as set of all vectors $-x$ with $x$ in $X$. Then $X + (-X) = U$.

With above definitions, the set of all cosets of $U$ is a vector space, the quotient space of $V$ modulo $U$, and is denoted by $V/U$ (read $V$ modulo $U$).

**Theorem 15.3.** Suppose $U$ and $V$ are complementary subspaces of a vector space $W$, i.e.: $U + V = W$ and $U \cap V = \{0\}$. The mapping

$$\Phi : V \to W/U : v \to v + U$$

is a vector space isomorphism.

**Proof** We prove that $\Phi$ is injective. If $\Phi(v_1) = \Phi(v_2)$ then $v_1 + U = v_2 + U$. Then, $v_1$ is an element of $v_2 + U$. Hence, $v_1 = v_2 + x$ for some $x$ in $U$. Therefore, $v_1 - v_2 = x$. It follows that $v_1 - v_2 = 0$, since $v_1 - v_2$ belongs to $V$, $x$ belongs to $U$ and $U \cap V = \{0\}$.

We prove that $\Phi$ is surjective. Choose any coset $v + U$ of $U$. Recall that $v = x + y$ with $x$ in $U$ and $y$ in $V$, since $U + V = W$. Hence, since $U + U = U$ and $x + U = U$,

$$v + U = (x + U) + U = (x + U) + (y + U) = U + (y + U) = y + U = \Phi(y).$$

Therefore, for every coset $v + U$ of $U$ there is some $y$ in $V$ such that $\Phi(y) = v + U$.

We prove that $\Phi$ is linear. For all scalars $a$ and $b$ and all $v_1$ and $v_2$ in $V$

$$\Phi(av_1 + bv_2) = (av_1 + bv_2) + U = (av_1 + U) + (bv_2 + U) = (av_1 + aU) + (bv_2 + bU) = a(v_1 + U) + b(v_2 + U) = a\Phi(v_1) + b\Phi(v_2)$$

since $U + U = U = aU = bU$.

**Corollary 15.4.** If $U$ is a subspace of the finite-dimensional vector space $W$ then $\dim(W/U) = \dim W - \dim U$.
Proof Suppose $W$ is a finite-dimensional vector space and $U$ is a subspace of $W$. There is a subspace $V$ of $W$ such that $W = U \oplus V$ and $\dim W = \dim U + \dim V$. Thanks to the theorem, $V$ is isomorphic to $W/U$. Hence, $\dim(W/U) = \dim V = \dim W - \dim U$.

**Exercises**

1. Suppose $U$ is a subspace of the vector space $V$. For any two vectors $u$ and $v$ in $V$ write $u \simeq v$ (read “$u$ and $v$ are congruent modulo $U$”) if $u - v$ is a vector in $U$.

   (a) Prove that $\simeq$ is an equivalence relation on $V$.

   (b) Prove that if $u_1 \simeq v_1$ and $u_2 \simeq v_2$ then $au_1 + bu_2 \simeq av_2 + bv_2$ for all scalars $a$ and $b$.

   (c) Prove that $V/\simeq = V/U$.

   Hint: To visualize the geometric meaning of $\simeq$ for the case $V = \mathbb{R}^2$, consider the situation shown in Figure 3. Given a subspace $U$, the vector $u$ defines the coset $\mathcal{X}$. The vector $v$ also defines the same coset. Therefore, the vector $u - v$ is parallel to $U$, i.e.: is an element of $U$. More generally, when $u$ and $v$ are not necessarily in the coset $\mathcal{X}$ shown in the figure, $u - v \in U$ is another way of saying that $u + U$ and $v + U$ are the same coset ($u + U = v + U$).

2. Suppose $U$ is a subspace of a vector space $V$ over the field $F$. Define $f : (V/U)' \to F^V : \lambda \to \mu$, where $\mu(x) = \langle x + U, \lambda \rangle$ for all $x$ in $V$.

   (a) Prove that $f(\lambda)$ is a linear functional on $V$ for every $\lambda$ in $(V/U)'$.

   (b) Prove that $f$ is a linear isomorphism between $(V/U)'$ and $U^0$.

3. Suppose $U$ is a subspace of the vector space $V$. Define

   $$f : V' \to U' : \lambda \to \mu,$$

   where $\langle x, \mu \rangle = \langle x, \lambda \rangle$.

   (a) Prove that if $\lambda_1 + U^0 = \lambda_2 + U^0$ then $f(\lambda_1) = f(\lambda_2)$.

   (b) Find a linear isomorphism $g$ between $V'/U^0$ and $U'$. Hint: Fill in the question mark in the diagram below, then obtain $g$.  

34