Chapter 1

Spaces

1.1 Fields

Definition 1.1. A field is a 5-tuple $(F, +, \cdot, 0, 1)$ satisfying the following:

1. F is closed under the $+$ and $\cdot$ binary operations.
2. $0, 1 \in F$ and $1 \neq 0$.
3. $(F, +, 0)$ is an abelian group.
4.
   - $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in F$.
   - $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ for all $\alpha, \beta, \gamma \in F$.
   - $\alpha 1 = \alpha$ for every $\alpha \in F$.
   - For every $\alpha \neq 0$ in F, there exists $\alpha^{-1} \in F$ such that $\alpha\alpha^{-1} = 1$.
5. $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ for all $\alpha, \beta, \gamma \in F$.

Exercise 1.

1. Prove that α^{-1} (when $\alpha \neq 0$) and $-\alpha$ are unique for a given $\alpha \in F$.
2. Show that $0\alpha = \alpha 0 = 0$ for all $\alpha \in F$. (Hint: $0 + 0 = 0$). Show that $\alpha\beta = 0$ if and only if $\alpha = 0$ or $\beta = 0$.
3. Show that 0 and 1 are unique, in the sense that if $(F, +, \cdot, 0, 1)$ and $(F, +, \cdot, 0', 1')$ are both fields, then $0 = 0'$ and $1 = 1'$. The following can help us prove that an object is not a field: suppose that $(F, +, \cdot, 0, 1)$ is not a field, but 0 and 1 satisfy $\alpha + 0 = \alpha$ and $\alpha 1 = \alpha$ for every $\alpha \in F$, then $(F, +, \cdot, 0', 1')$ is not a field for any $0', 1' \in F$. Prove that $\mathbb{N}$, $\mathbb{Z}$ are not fields.
4. Suppose that F is a field. Define 1^c to be the *sum* of 1 with itself c times, where $c \in \mathbb{N}$. Prove that $1^c = 0$ for some c, and that the smallest such c must be prime. (Hint: $\alpha\beta = 0 \implies \alpha = 0$ or $\beta = 0$).

In what follows, we call the elements of a field scalars.

1.2 Vector Spaces

Definition 1.2. A vector space over field F is a 4-tuple $(V, +, \cdot, \mathbf{0})$ satisfying the following:

1. V is closed under $+ : V \times V \to V$ and $\cdot : F \times V \to V$.
2. $(V, +)$ is an abelian group.
3. $\alpha(\beta\vec{v}) = (\alpha\beta)\vec{v}$ for all $\alpha, \beta \in F$ and $\vec{v} \in V$.
4. $1\vec{v} = \vec{v}$ for all $\vec{v} \in V$.
5. $(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v}$ for all $\alpha, \beta \in F$ and $\vec{v} \in V$.
6. $\alpha(\vec{v} + \vec{w}) = \alpha\vec{v} + \alpha\vec{w}$ for all $\alpha \in F$ and $\vec{v}, \vec{w} \in V$.

In what follows, we call the elements of a vector space vectors. If the underlying field is $\mathbb{R}$, we shall call the space a *real vector space*. Similarly we have *complex vector spaces* and *rational vector spaces*.

Example 1.

1. $\mathbb{C}^n$ is a complex vector space under the natural $+$ and $\cdot$ operations. We shall refer to $\mathbb{C}^n$ as n-*dimensional complex coordinate space*. Similarly $\mathbb{F}^n$ is a vector space over $\mathbb{F}$. Note that $\mathbb{C}^n$ is a real vector space but $\mathbb{R}^n$ is not a complex vector space.
2. The set $\mathcal{P}_n$ of polynomials of degree at most $n-1$ is a vector space with $+$ and $\cdot$ being the addition and scalar multiplication operations on polynomials.
3. $\mathcal{O} = \{\mathbf{0}\}$ is a trivial vector space over any field F.

From now on, we shall use 0 to denote $\mathbf{0}$. This should not be confused with the scalar 0. To make matters worse, 0 will also be used to denote the trivial linear functional and linear transformation. Fortunately these relations among the various interpretation of 0 are such that, after this word of warning, there should be no confusion from this practice.

Exercise 2.

1. Prove that $-v$ is unique for a given $v \in V$.
2. Show that $0v = 0$ for all $v \in V$. Prove that $\alpha v = 0$ if and only if $\alpha = 0$ or $v = \mathbf{0}$.

1.3 Linear dependence

In what follows, when we speak of a set of vectors $\{x_i\}$, we admit the possibility of two different indices corresponding to the same vector. That is, what is important is not which vectors appear, but rather how they appear.

Definition 1.3. A finite set of vectors $\{x_i\}$ is linearly dependent if there exist a corresponding set $\{\alpha_i\}$ of scalars, not all zero, such that

$$\sum_i \alpha_i x_i = 0.$$

If on the other hand, $\sum_i \alpha_i x_i = 0$ implies that $\alpha_i = 0$ for each i, the set is linearly independent.

We shall adopt the abbreviation LD for linearly dependent and LI for linearly independent. Note also that the empty sum is 0, and so the empty set of vectors is LI.

The reason for the word *dependent* is as follows: suppose that $\{x_i\}$ is LD, and fixing i say $\alpha_i \neq 0$. Then we can write x_i as a (*linear* - only using $+$) combination of the (scaled versions of) other vectors in the set. In this sense, x_i is *dependent* on the other vectors in the set.

Remark. Note that the empty set of vectors being linearly independent is useful, because using this, any finite set of vectors is either linearly dependent or linearly independent.

Exercise 3. Suppose $\{x_i\}$ is LI. Then show that any subset of $\{x_i\}$ is also LI. Equivalently, if $\{x_i\}$ is LD, then any superset of $\{x_i\}$ is also LD.

We shall say, whenever $x = \sum_i \alpha_i x_i$ (i ranges over a finite set) that x is a linear combination (LC) of $\{x_i\}$.

Definition 1.4 (Span). The span of a set of vectors $\{x_i\}$ is the set of all finite linear combinations of $\{x_i\}$.

Exercise 4. Suppose $\{x_i\}$ is LI and $x \in V$. Then x is a LC of $\{x_i\}$ if and only if $\{x_i\} \cup \{x\}$ is LD.

Theorem 1.5. *The set of non-zero vectors $x_1, \ldots, x_n$ is linearly dependent if and only if some x_k, $2 \leq k \leq n$, is a linear combination of the preceding vectors $x_1, \ldots, x_{k-1}$. Moreover, the span of $x_1, \ldots, x_n$ is the same as the span of $x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$.*

Proof. Suppose that k is the smallest index ($k > 1$ since $x_1 \neq 0$) such that $x_1, \ldots, x_k$ is LD. Note that such a k exists because $x_1, \ldots, x_n$ is LD. Exercise 4 yields the first result. To finish, note that any finite linear combination of $x_1, \ldots, x_n$ can be changed to a linear combination of $x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$ by replacing x_k in the LC by its linear combination in terms of $x_1, \ldots, x_{k-1}$. Conversely, any LC of $x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$ is obviously one of $x_1, \ldots, x_n$. □

1.4 Bases

Definition 1.6. A (linear) basis or coordinate system in a space V is a set $\mathcal{X}$ of linearly independent vectors such that every vector in V is a linear combination of vectors in $\mathcal{X}$. A vector space V is finite-dimensional if it has a finite basis.

Example 2.

1. Consider the set of vectors $e_1, \ldots, e_n$ in $\mathbb{F}^n$ where e_i is the vector with 1 in the i-th coordinate and 0 elsewhere. This is called the canonical basis for $\mathbb{F}^n$.
2. The set of polynomials $\{1, x, x^2, \ldots, x^{n-1}\}$ is a basis for $\mathcal{P}_n$. The space of polynomials $\mathcal{P}$ is not finite-dimensional.

Remark. Since we take the empty sum is the zero vector, $\varnothing$ is a basis for $\mathcal{O}$.

Exercise 5. Suppose V has basis $\mathcal{X}$. Show that every x can be written uniquely as a linear combination of vectors in $\mathcal{X}$. That is, if $x = \sum_i \alpha_i x_i$, then each α_i is uniquely determined by x.

Theorem 1.7. *If V is a finite-dimensional vector space and if $\{y_1, \ldots, y_m\}$ is an independent set of vectors in V, then unless the y_is already form a basis, we can find vectors $y_{m+1}, \ldots, y_{m+p}$ such that the set $\{y_1, \ldots, y_m, y_{m+1}, \ldots, y_{m+p}\}$ is a basis. In other words, every linearly independent set can be extended to a basis.*

Proof. Since V is finite dimensional, it admits a basis $x_1, \ldots, x_n$. If $m = 0$, we set $y_{0+i} = x_i$ and win. Otherwise, we apply Theorem 1.5 to the linearly dependent set $y_1, \ldots, y_m, x_1, \ldots, x_n$ to get that some x_k is a linear combination of the preceding vectors (it cannot be a y_i that we throw out, since $\{y_i\}$ is LI). If the resulting set is independent, then we set $y_{m+1}, \cdots y_{m+p}$ to be the remaining x_is and win. Otherwise, we apply Theorem 1.5 repeatedly till this is the case. Notice that this will stop in a finite number of steps, since the number of x_is is finite. Also, by the span result of Theorem 1.5, the resulting set is a basis. □

Theorem 1.8. *Suppose V is a finite-dimensional vector space with basis $x_1, \ldots, x_n$. Then any linearly independent set of vectors in V has at most n vectors.*

Proof. This is essentially a refined version of the above proof. Apart from extending an LI set to a basis, it also bounds its size. Suppose $\mathcal{Y} = \{y_1, \ldots, y_m\}$ is LI. Consider the LD set $S = \{y_m, x_1, \ldots, x_n\}$. Apply Theorem 1.5 and throw an x_i out from S. Add y_{m-1} to the *front* of S; the set is again LD. Another x_i is thrown, add y_{m-2}. Repeat till all elements of $\mathcal{Y}$ have been added. At every stage, there *must have been* an x_i to throw out, since $\mathcal{Y}$ is LI. Thus $m \leq n$. □

Essentially, the previous proof tells us that every additional vector in an LI set must take the spot of some element of a basis - and thus there can only be so many elements in an LI set.

Corollary. Any two bases of a finite-dimensional vector space have the same number of elements.

Definition 1.9. The dimension of a finite-dimensional vector space V is the number of elements in any basis of V. We denote this by $\dim V$.

Exercise 6. Suppose V is a finite-dimensional vector space with $\dim V = n$, and $\mathcal{X}$ a set of vectors in V. Then any two of the following imply that $\mathcal{X}$ is a basis for V:

1. $\mathcal{X}$ is LI.
2. $\mathcal{X}$ spans V.
3. $\mathcal{X}$ has n elements.

1.5 Isomorphism

In this section, we show the intuitive result that every finite-dimensional vector space over field F is essentially the same as some F^n. The definition of isomorphism is identical to the definition from Abstract Algebra.

Definition 1.10 (Isomorphic vector spaces). Two vector spaces U and V over the same field F are isomorphic (denoted $U \cong V$) if there exists a bijection $T : U \to V$ such that

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

for all $u, v \in U$ and $\alpha, \beta \in F$. We call such a T an isomorphism.

Remark. Any isomorphism T must take 0_U to 0_V.

Exercise 7. The isomorphism relation $\sim$ defined by $U \sim V$ if U and V are isomorphic is an equivalence relation.

Exercise 8. Suppose U and V are isomorphic vector spaces over F. Then $\dim U = \dim V$. In fact, any isomorphism $T : U \to V$ maps each basis of U to a basis of V.

Which vector spaces are in the equivalence class under $\sim$ of $\mathbb{F}^n$? Of course, any such space has to have dimension n. Remarkably, *every* space with dimension n is isomorphic to $\mathbb{F}^n$.

Theorem 1.11. *[Isomorphism theorem] Suppose V is a finite-dimensional vector space over F with $\dim V = n$. Then V is isomorphic to $\mathbb{F}^n$.*

Proof. Let $v_1, \ldots, v_n$ be a basis of V. Verify that the map $T : V \to \mathbb{F}^n$ defined by $T(x = \alpha_1 v_1 + \cdots + \alpha_n v_n) = (\alpha_1, \ldots, \alpha_n)$ is an isomorphism. The same map may be described as follows: Consider the canonical basis $e_1, \ldots, e_n$ of $\mathbb{F}^n$, and let T map each v_i to e_i. Notice that linearity fixes the value of T on all of V, so T is uniquely determined. Indeed, it is the isomorphism we seek. □

Corollary. Suppose that vector spaces V and W have the same dimension n. Then $V \cong W$.

Exercise 9.

1. Show that the real vector spaces $\mathbb{C}^n$ and $\mathbb{R}^{2n}$ are isomorphic.
2. Suppose that V is a vector space over $\mathbb{Z}_p$ with dimension n. Find $|V|$.

1.6 Subspaces

Definition 1.12. A nonempty subset U of a vector space V is a subspace of V if U is a vector space under the same operations as V. We denote this by $U \leq V$.

Exercise 10. Let $U \subseteq V$ be nonempty. Show that $U \leq V$ if and only if U is closed under addition and scalar multiplication, that is, for every $x, y \in U$ and $\alpha, \beta \in F$, we have $\alpha x + \beta y \in U$.

This latter criterion is typically most convenient in establishing that a given subset is a subspace.

Remark. Note that $0 \subseteq U$ for any subspace U of V. We say that two subspaces $\mathfrak{M}, \mathfrak{N}$ of V are disjoint if $\mathfrak{M} \cap \mathfrak{N} = 0$.

Exercise 11.

1. Suppose $U \leq V$ with $\dim U = \dim V$. Show that $U = V$.
2. The following fact is quite useful in establishing isomorphism. Suppose the linear map $T : U \to V$ is injective, and $\dim U = \dim V$. Show that T is surjective and hence an isomorphism.

Exercise 12.

1. Consider a set $\mathfrak{M}_1, \ldots$ of subspaces of V. Show that $\mathfrak{M} = \bigcap_i \mathfrak{M}_i$ is a subspace of V.
2. Suppose U and W are subspaces of V. Show that $U \cup W$ is a subspace if and only if $U \subseteq W$ or

$W \subseteq U$. Unions are not "well-closed" essentially.

Definition 1.13. Consider a set $S \subseteq V$. The intersection of all subspaces of V containing S is called the subspace spanned by S, and is denoted by $\operatorname{span} S$.

Exercise 13. Consider a set $S \subseteq V$. Then the subspace spanned by S is precisely the span of S.

The above exercise establishes the fact that the span of a set of vectors is a subspace, and that it is actually the minimal (in terms of set inclusion) subspace containing the set.

Exercise 14. If $\mathcal{H}$ and $\mathcal{K}$ are two subspaces and $\mathfrak{M}$ the subspace spanned by $\mathcal{H} \cup \mathcal{K}$, then $\mathfrak{M}$ is the same as the set of all vectors of the form $x + y$, where $x \in \mathcal{H}$ and $y \in \mathcal{K}$.

Prompted by the above exercise, we denote $\mathfrak{M} = \mathcal{H} + \mathcal{K}$. This is called the sum of spaces $\mathcal{H}$ and $\mathcal{K}$. We shall say that a subspace $\mathcal{K}$ of a vector space V is a complement of a subspace $\mathcal{H}$ of V if $V = \mathcal{H} + \mathcal{K}$ and $\mathcal{H} \cap \mathcal{K} = \mathcal{O}$.

Theorem 1.14. *A subspace U of an n-dimensional vector space V is a vector space of dimension at most n.*

Proof. If $U = \mathcal{O}$, then $\dim U = 0$, and we are done. Otherwise, $\exists$ nonzero $u_1 \in U$. Let U_1 be the subspace spanned by u_1. If $U_1 = U$, we are done. Otherwise, there is $u_2 \in U \setminus U_1$. Let U_2 be the subspace spanned by u_1, u_2. If $U_2 = U$, we are done. Otherwise, there is $u_3 \in U \setminus U_2$. Continuing in this fashion, we get a sequence of subspaces $U_1 \subset U_2 \subset \dots$ of U. Since V is finite-dimensional, this sequence must terminate, since the set $u_1, \dots$ is LI, say at U_k. Then $U_k = U$, and so $\dim U = k \leq n$. □

A similar argument shows the following:

Exercise 15 (Basis extension for subspaces). Given any m-dimensional subspace U of n-dimensional vector space V and a basis $x_1, \dots, x_m$ of U, we can find a basis $x_1, \dots, x_m, x_{m+1}, \dots, x_n$ of V.

The above exercise also implies Theorem 1.7.

Exercise 16.

1. Show that every subspace of a vector space V has a complement.
2. Show that every non-trivial subspace of a vector space V (i.e. one that is not $\mathcal{O}$ or V) does not have a unique complement.
3. If U is a m-dimensional subspace of n-dimensional vector space V, then show that every complement of U in V has dimension $n - m$.

Exercise 17. An easy consequence of the basis extension theorem is the following inclusion-exclusion style result:

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

(Hint: Extend a basis of $U \cap W$ to bases of U and W.) The general theorem is also true. Use the above

hint and the combinatorial inclusion-exclusion principle to show that

$$\dim(U_1 + \cdots + U_m) = \sum_{i=1}^{m} \dim U_i - \sum_{i<j} \dim(U_i \cap U_j) + \cdots + (-1)^{m-1} \dim(U_1 \cap \cdots \cap U_m).$$

1.7 Dual Spaces

Definition 1.15 (Linear functional). A linear functional on a vector space V over field F is a function $y : V \to F$ such that $y(\alpha x + \beta x') = \alpha y(x) + \beta y(x')$ for all $x, x' \in V$ and $\alpha, \beta \in F$. The set of all linear functionals on V is denoted by V^*.

As with isomorphisms, the values of a linear functional on a basis of V determine its values on all of V. Also, every linear functional y sends 0_V to 0_F.

Example 3.

1. For $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$, the functions $y(x) = x_1$ or $y(x) = x_1 + x_2$ or more generally $y(x) = \sum_i \alpha_i x_i$ are linear functionals on $\mathbb{C}^n$.
2. The zero functional, $y(x) = 0$ for all $x \in V$, is a linear functional on V.

Definition 1.16 (Dual space). The dual space of a vector space V over field F is the vector space V^* over F of all linear functionals on V. The operation of addition and scalar multiplication on V^* are defined pointwise.

Exercise 18. Show that the dual space of a vector space V is a vector space.

Definition 1.17 (Bracket notation). The notation $[x, y]$ is a substitute for the ordinary function symbol $y(x)$.

The notation $[x, y]$ is a symbolic way of writing down the recipe for actual operations performed; for example, if $y(x) = x^2$, it corresponds to the sentence [take a number, and square it].

The defining property of a linear functional then becomes

$$[\alpha x + \beta x', y] = \alpha[x, y] + \beta[x', y],$$

and the definition of the linear operations for linear functionals becomes

$$[x, \alpha y + \beta y'] = \alpha[x, y] + \beta[x, y'].$$

The two relations together say that $[x, y]$ is a bilinear form of $x \in V$ and $y \in V^*$.

Exercise 19. Fix a basis $x_1, \ldots, x_n$ of V. Show that every function y of the form

$$y(x = \sum_i \alpha_i x_i) = \sum_i \alpha_i \beta_i$$

is a linear functional. Conversely, show that every linear functional on V is of this form for some β_i.

The previous exercise yields the useful Riesz representation theorem, which we shall state here though the notion of a matrix and matrix multiplication has not been defined yet, since it is a useful way to think about the dual space.

Theorem 1.18 (Riesz). *Let V be a finite-dimensional vector space with some basis $\mathcal{X} = \{x_1, \dots, x_n\}$. Associate each $y \in V^*$ with the vector $\beta = \sum_i \beta_i x_i$ (represented $[\beta_1, \dots, \beta_n]^T$) where $y(x_i) = \beta_i$ for each i. This correspondence is an isomorphism between V^* and V, and further for each $x = \sum_i \alpha_i x_i$ (represented $[\alpha_1, \dots, \alpha_n]^T$) we have*

$$[x, y] = x^T \beta.$$

Exercise 20.

1. Show that the value of $[x_i, y]$ for a basis $x_1, \dots, x_n$ of V determines y uniquely.
2. Show that the value of $[x, y_i]$ for a basis $y_1, \dots, y_n$ of V^* determines x uniquely.

Exercise 21. Prove that if y and z are linear functionals on V such that

$$[x, z] = 0 \implies [x, y] = 0,$$

then there exists scalar α such that $y = \alpha z$.

It is clear that given a basis $x_1, \dots, x_n$ of V and a set of scalars $\alpha_1, \dots, \alpha_n$, there is one and only one functional y satisfying $[x_i, y] = \alpha_i$ for each i. Consider the functionals $y_1, \dots, y_n$ defined by $[x_i, y_j] = \delta_{ij}$, where

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

is the Kronecker delta.

Exercise 22. Show that $y_1, \dots, y_n$ is a basis of V^*.

Definition 1.19 (Dual basis). Fix a basis $\mathcal{X} = \{x_1, \dots, x_n\}$ of V. The basis $\mathcal{X}^* = \{y_1, \dots, y_n\}$ of V^* defined by $[x_i, y_j] = \delta_{ij}$ is called the dual basis of $x_1, \dots, x_n$.

The reason for dual is due to the following theorem:

Theorem 1.20. *Let V be a finite-dimensional vector space. The dual space V^* is isomorphic to V.*

Proof. V and V^* have the same dimension by the existence of a dual basis. The result then follows from Theorem 1.11. □

If we look at $[x, y]$ as a function of x for fixed $y = y_0$, we see a linear functional acting on V. If, however, we fix $x = x_0$, we then see a linear functional on V^*, that is, an element of V^{**}. By this method we have exhibited *some* linear functionals on V^*; have we exhibited them all? For the finite-dimensional case, the answer is, remarkably, yes.

Theorem 1.21 (Double dual). *Let V be a finite-dimensional vector space. Corresponding to each linear functional z_0 on V^* there is a vector x_0 in V such that*

$$z_0(y) = [x_0, y] = y(x_0)$$

for every $y \in V^*$. The correspondence $z_0 \leftrightarrow x_0$ between V^{**} and V is an isomorphism, called the natural isomorphism.

Proof. Consider the map T from V to V^{**} defined by $T(x) = [x, \cdot]$ viewed as an element of V^{**}. T is linear, since $T(\alpha x + \beta x') = \alpha[x, \cdot] + \beta[x', \cdot] = \alpha T(x) + \beta T(x')$. T is injective, since $T(x) = 0$ implies $[x, y] = 0$ for all y, and so $x = 0$. $T(V)$ is an n-dimensional subspace of V^{**}, since for a basis $x_1, \ldots, x_n$ of V, $T(x_1), \ldots, T(x_n)$ is one of $T(V)$: $\sum_i \alpha_i T(x_i) = 0 \implies T(\sum_i \alpha_i x_i) = 0 \implies \sum_i \alpha_i x_i = 0 \implies \alpha_i = 0$ for all i. So $T(V) = V^{**}$, since V^{**} is n-dimensional. Thus T is surjective (establishing the existence of an x_0 for every z_0), and hence, also an isomorphism. □

In view of the above theorem, it is frequently convenient to be mildly sloppy and identify V^{**} with V, and we shall say that the element z_0 of V^{**} is the same as the element x_0 of V whenever $z_0(y) = [x_0, y]$ for each y.

Exercise 23. Show that the dual basis $\mathcal{X}^{**} \in V^{**}$ identifies with $\mathcal{X} \in V$.

1.8 Annihilators

Definition 1.22. The annihilator S^0 of a subset S of a vector space V is the set of vectors $y \in V^*$ such that $[x, y] = 0$ for all $x \in S$.

Notice that the annihilator of a subset is a subspace of the dual space. We have the following nice result:

Theorem 1.23. *If U is an m-dimensional subspace of V, then U^0 is an $(n - m)$-dimensional subspace of V^*.*

Proof. Suppose that a basis for U is $u_1, \ldots, u_m$ and that extended to V is $u_1, \ldots, u_n$. The esential constraint on a vector $y = \sum_i \alpha_i u_i^*$ in U^0 is that $y(u_i) = 0$ for $i = 1, \ldots, m$. $y(u_{m+1}) \ldots y(u_n)$ can be chosen arbitrarily. We leave it to the reader to verify that $y_{m+1}, \ldots, y_n$ indeed form a basis for U^0. □

Exercise 24. Suppose that $U \subseteq W$ are subspaces of vector space V. Show that $\dim U = \dim W$ if and only if $U = W$.

Theorem 1.24 (Involution). *If U is a subspace of V, then $U^{00} = U$.*

Proof. U^{00} is the set of vectors $z \in V^{**}$ (so $x \in V$) such that $[x, y] = 0$ for all $y \in U^0$. For each $x \in U$, $[x, y] = 0$ for all $y \in U^0$ so $U \subseteq U^{00} \subseteq V$. But $\dim U^{00} = n - (n - m) = m = \dim U$, so $U = U^{00}$. □

Exercise 25.

1. Let $y \neq 0 \in V^*$. Prove that the set of vectors $x \in V$ with $[x, y] = 0$ is an $(n - 1)$ dimensional subspace of V.

2. Let $y(x) = \zeta_1 + \zeta_2 + \zeta_3$ whenever $x = (\zeta_1, \zeta_2, \zeta_3)$ is a vector in $\mathbb{C}^3$. Find a basis for the subspace of $\mathbb{C}^3$ consisting of all vectors x such that $y(x) = 0$.

Exercise 26 (System of linear equations).

1. Consider a system of m equations (with right-hand side 0) in n variables represented as a set of m linear functionals over $\mathbb{C}^{n*}$: that is, the system of linear equations is rephrased as $[x, y_i] = 0$ for $i = 1, \ldots, m$. Prove that for $m < n$, the kernel of this system (the set of solutions x) contains a nonzero vector. Find the dimension of the kernel.
2. Suppose now that the equations are of the form $[x, y_i] = \alpha_i$. Let $m < n$ again. What are the conditions on the α_i's for the system to have a solution?

Exercise 27. Let M, N be subspaces of a finite-dimensional vector space V. Prove that $(M + N)^0 = M^0 \cap N^0$ and $(M \cap N)^0 = M^0 + N^0$. The involution property of the annihilator operation can sometimes come in handy.

1.9 Direct Sums

We shall study several important general methods of making new vector spaces out of old ones. The sum construction we have already seen to make new subspaces out of existing ones.

Definition 1.25 (Direct Sum). if U and V are vector spaces over the same field F, their direct sum $W = U \oplus V$ whose elements are all the ordered pairs $\langle u, v \rangle$ with $u \in U$ and $v \in V$. The operations of addition and scalar multiplication are defined by

$$\langle u_1, v_1 \rangle + \langle u_2, v_2 \rangle = \langle u_1 + u_2, v_1 + v_2 \rangle$$
$$\alpha \langle u, v \rangle = \langle \alpha u, \alpha v \rangle$$

Exercise 28. Prove that the direct sum of two vector spaces is a vector space.

Exercise 29. Prove that the direct sum of two finite-dimensional vector spaces is finite-dimensional and that $\dim(U \oplus V) = \dim U + \dim V$.

The set of vectors $\langle x, 0 \rangle$ is a subspace of $U \oplus V$ isomorphic to U. It is convenient to identify U with this subspace. Similarly, we identify V with the subspace $\langle 0, v \rangle$ of $U \oplus V$. With this identification, we can think of U and V as subspaces of $U \oplus V$.

The distinction between $\langle u, v \rangle$ and $u + v$ with U, V disjoint is more pedantic than conceptual - it is much the same distinction as $((1, 2), 3)$ and $(1, 2, 3)$ - and we shall usually ignore it.

The question then arises: what is the relation between U and V when we consider these spaces as subspaces of the big space W?

Theorem 1.26. *if U and V are subspaces of a vector space W, then the following three conditions are equivalent.*

1. $W = U \oplus V$.
2. $U \cap V = 0$ *and* $U + V = W$ *(i.e. U and V are complements of each other).*

3. *Every vector $w \in W$ can be written in one and only one way as $w = u + v$ with $u \in U$ and $v \in V$.*

Proof. The proofs (1) $\implies$ (2) $\implies$ (3) are left as exercises. The implication (3) $\implies$ (1) is almost by definition. We have that every vector $z \in W$ can be written $z = \langle u, 0\rangle + \langle 0, v\rangle = \langle u, v\rangle$ for some $u \in U$ and $v \in V$. This correspondence is injective; thus $W \subseteq U \oplus V$. but every vector in $U \oplus V$ can be written in the form $\langle u, v\rangle = u + v \in W$ so $U \oplus V \subseteq W$. □

The condition (2) provides an easy-to-check criterion for determining if $W = U \oplus V$. Picking bases for U and V, the condition (2) is equivalent to the statement that the concatenation of the two bases is a basis for W (that is, the concatenation must be independent - the intersection condition - and also span W - the sum condition).

If two subspaces U and V of W are disjoint and they span W, it is usual to say that W is the internal direct sum of U and V.

Exercise 30. If W is n-dimensional and U is an m-dimensional subspace of W, prove that there exists an $(n - m)$-dimensional subspace V of W such that $W = U \oplus V$.

We next characterize the linear functionals on a direct sum. For simplicity (but without loss of generality), we assume that M and N are disjoint subspaces of V.

Exercise 31. Suppose that M and N are subspaces of V, and that $V = M \oplus N$. Show that M^* is isomorphic to N^0 and N^* to M^0. Show also that $V^* = M^0 \oplus N^0$.

To conclude, note that one can generalize each result proved here to the direct sum of any finite number of vector spaces; there is nothing special about the number two.

Exercise 32. Show that the $\oplus$ operator is commutative and associative upto isomorphism.

1.10 Quotient Spaces

If M is a subspace of vector space V, there are usually many complements of M. There is no natural way of choosing one from among the wealth of complements. There is, however, a natural construction that associates with M and V a new vector space that, for all practical purposes, plays the role of a complement of M. The theoretical advantage is that it does not depend upon a basis, or for that matter, on choosing anything at all.

Definition 1.27 (coset). We define the coset of M in V containing x to be the set $x+M = \{x+m : m \in M\}$.

We do not distinguish cosets by their representatives x.

Exercise 33.

1. The cosets of M in V form a partition of V.
2. We define the sum of two cosets like subspace addition: we write $H + K$ for the set of all sums $u + v$ with $u \in H$ and $v \in K$. Prove that $H + K$ is a coset of M in V.
3. We define the scalar multiplication operator in the natural way: $\alpha H = \{\alpha h : h \in H\}$. Prove that αH is also a coset of M in V.
4. Verify that the set of cosets forms a vector space over F with the operations as defined above. (For the reader familiar with group theory, the normality of M in V is trivial from the commutativity

in the vector space, so the set of cosets is always a vector space).

The vector space of the set of cosets is called the quotient space V/M.

Theorem 1.28. *If* M *and* N *are complements in* V*, then the correspondence that assigns to each vector* $y \in N$ *the coset* $y + M$ *is an isomorphism between* N *and* V/M*.*

Proof. The map is clearly linear. It is injective because $M \cap N = 0$. It is surjective because $M + N = V$. □

Corollary. If M is a subspace of finite-dimensional vector space V, then V/M is finite-dimensional and $\dim V/M = \dim V - \dim M$.

Exercise 34. Let M be a subspace of V.

1. Assign to each $y \in (V/M)^*$ the functional $z \in V^*$ defined by $z(x) = y(x + M)$ for all $x \in V$. Show that every such z is an element of M^0. Moreover, prove that T is an isomorphism $(V/M)^* \to M^0$. (the key point here is that this particular map is an isomorphism; the isomorphism between the spaces is trivial by dimensionality).
2. Show that $V^* \cong M^* \oplus M^0$. Corresponding to every coset $y + M^0$ in V^*/M^0 there is a linear functional z on V such that $z(x) = y(x)$ for all $x \in M$. Show that the map $T : y + M^0 \mapsto z$ is well-defined. That is, for every $y, y' \in V^*$, if $y + M^0 = y' + M^0$, then $y = y'$ on M. Show also that T is an isomorphism $V^*/M^0 \to M^*$.

1.11 Bilinear Forms

Consider a direct sum $W = U \oplus V$. It will be convenient to use the $\langle x, y \rangle$ representation for elements of W here. We write $w(\langle x, y \rangle) = w(x, y)$ for a function w on W. The linear functionals w are not of much more interest; we have already described them in Exercise 31. We turn our attention to other functions on W; in particular, the bilinear forms.

Definition 1.29. A scalar-valued function w on W is called a bilinear form (or bilinear functional) if it is linear in each variable separately: For all $x, x_1, x_2 \in U$ and $y, y_1, y_2 \in V$ and all $\alpha, \beta \in F$, we have

$$w(\alpha x_1 + \beta x', y) = \alpha w(x, y) + \beta w(x', y)$$
$$w(x, \alpha y + \beta y') = \alpha w(x, y) + \beta w(x, y')$$

Remark. Note that every bilinear form is 0_F on each $\langle x, 0 \rangle$ and $\langle 0, y \rangle$.

We have already seen one such bilinear functional: suppose that $W = V \oplus V^*$. Set $w(x, y) = [x, y]$. Then w is a bilinear form on W.

Exercise 35. Let $W = U \oplus V$ and let $u \in U^*, v \in V^*$. Show that the function $w(x, y) = u(x)v(y)$ is a bilinear form on W. (note however, that the function $w(x, y) = u(x) + v(y)$ is not a bilinear form on W).

Indeed, supposing that $u_1^*, \ldots, u_m^*$ is a basis for U^* and $v_1^*, \ldots, v_n^*$ is a basis for V^*, then

$$w = \sum_{i,j} \alpha_{ij} u_i^*(x) v_j^*(y)$$

is a bilinear form on W. It is reasonable to expect that every bilinear form on W can be written in this form, much like the case for linear functionals.

Exercise 36. Define the addition and scalar multiplication operations on bilinear forms to be pointwise. Under these operations, the set of bilinear forms $W^{(2)}$ on W is a vector space over F.

We define the dual basis for the space of bilinear forms just like we did for the dual space: letting $\mathcal{X} = \{x_1, \ldots, x_m\}$ be a basis for U and $\mathcal{Y} = \{y_1, \ldots, y_n\}$ be a basis for V, define the set of bilinear forms w_{ij} by

$$w_{ij}(x_k, y_l) = \delta_{ik}\delta_{jl}$$

with extension to all of W by bilinearity.

Exercise 37. Show that the set $\{w_{ij}\}$ is a basis for $W^{(2)}$. Hence conclude that $\dim W^{(2)} = \dim U \cdot \dim V$.

Exercise 38. Given a bilinear form w, we define the transpose w^t of w by $w^t(x, y) = w(y, x)$. w is said to be symmetric if $w = w^t$ and skew-symmetric if $w = -w^t$. Show that every bilinear form can be written as the sum of a symmetric and a skew-symmetric bilinear form.

Exercise 39. Suppose that w is a bilinear form on $W = V \oplus V$. A quadratic form on V is a function q on V defined by $q(x) = w(x, x)$.

1. If w is symmetric and 4 is invertible in the underlying field, show that q determines w uniquely.
2. If w is skew-symmetric, show that q is the zero function.

Finally, a Riesz-like (see Theorem 1.18) result for bilinear forms:

Exercise 40. Let w be a linear form on $V \oplus W$. Fix bases $\mathcal{X} = \{x_1, \ldots, x_n\}$ of V and $\mathcal{Y} = \{y_1, \ldots, y_n\}$ of W, and consider matrix A with entries $a_{ij} = w(x_i, y_j)$. Show that for each x, y (represented as column vectors in bases $\mathcal{X}$, $\mathcal{Y}$ respectively) we have

$$w(x, y) = x^T A y.$$

Further, show that w is symmetric iff A is symmetric, and similarly for skew-symmetric.

1.12 Multilinear Forms

We foray briefly into the world of multilinear forms because of their importance in the study of determinants later on.

Definition 1.30. Suppose $V_1, \ldots, V_k$ are vector spaces over the same field F. A k-linear form is a scalar-valued function w from $V_1 \oplus \cdots \oplus V_k$ to F that is linear in each variable for any fixed value of the other variables.

That is, for each i, scalars $\alpha_1, \alpha_2 \in F$, vectors $y_j \in V_j$ for $j \neq i$ and $x_1, x_2 \in V_i$ we have

$$w(y_1, \ldots, y_{i-1}, \alpha_1 x_1 + \alpha_2 x_2, y_{i+1}, \ldots, y_k) = \alpha_1 w(\ldots, y_{i-1}, x_1, y_{i+1}, \ldots) + \alpha_2 w(\ldots, y_{i-1}, x_2, y_{i+1}, \ldots).$$

The linear functionals are precisely the 1-linear forms. As with these, pointwise addition and multiplication make the set of k-linear forms into a vector space over F.

Exercise 41. Let $W^{(k)}$ denote the vector space of all k-linear forms on $W = V_1 \oplus \cdots \oplus V_k$. Show that $\dim W^{(k)} = \prod_{i=1}^{k} \dim V_i$.

Example 4. Consider $W = \mathbb{R}^2 \oplus \mathbb{R}^3$. Let $\hat{\imath}, \hat{\jmath}$ be the standard basis for $\mathbb{R}^2$ and e_1, e_2, e_3 be the standard basis for $\mathbb{R}^3$. Consider $x = x_1\hat{\imath} + x_2\hat{\jmath}$ and $y = y_1e_1 + y_2e_2 + y_3e_3$.

- If w is a linear functional on W, then

$$w(x, y) = x_1w(\hat{\imath}, 0) + x_2w(\hat{\jmath}, 0) + y_1w(0, e_1) + y_2w(0, e_2) + y_3w(0, e_3).$$

- If w is instead a bilinear functional on W, then

$$w(x, y) = x_1y_1w(\hat{\imath}, e_1) + x_1y_2w(\hat{\imath}, e_2) + x_1y_3w(\hat{\imath}, e_3) + x_2y_1w(\hat{\jmath}, e_1) + x_2y_2w(\hat{\jmath}, e_2) + x_2y_3w(\hat{\jmath}, e_3).$$

Let us denote the k-fold direct sum $V \oplus \dots V$ by $V^{\oplus k}$. A k-linear form on $V^{\oplus k}$ is symmetric if for each $\pi in S_k$ (the set of permutations of $[k]$ aka the symmetric group of order k), one has $\pi w = w$, where

$$\pi w(x_1, \dots, x_k) = w(x_{\pi(1)}, \dots, x_{\pi(k)}).$$

A k-linear form is skew-symmetric if $\pi w = \operatorname{sgn}(\pi) w$ for each $\pi \in S_k$ ($\operatorname{sgn}(\pi)$ is the sign of the permutation π).

Exercise 42. Show that w is skew-symmetric iff $\pi w = -w$ for every odd permutation w.

Finally, a form w is alternating if $w(x_1, \dots, x_k)$ is 0 whenever two of the x_i's are equal.

Theorem 1.31. *Every alternating multilinear form is skew-symmetric.*

Proof. Exercise. Use $0 = w(x_1 + x_2, x_1 + x_2, \dots, x_k)$. □

Note that the converse is not true in full generality.

Exercise 43.

1. Suppose w is a skew-symmetric multilinear form. Suppose that 2 is invertible in the underlying field. Show that w is alternating.
2. Consider the field $F = \mathbb{Z}_2$ and the vector space $V = F^2$. In this field, $w = -w$ so symmetry is identical to skew-symmetry. Show that there exists a symmetric multilinear form on V that is not alternating.

It turns out that alternating forms have a lot to do with linear dependence.

Exercise 44. Let $x_1, \dots, x_k$ be linearly dependent, and let w be an alternating multilinear form on $V^{\oplus k}$. Then $w(x_1, \dots, x_k) = 0$. Hint: Use Theorem 1.5.

Theorem 1.32. *Let $\dim V = n$, and suppose that $w \neq 0$ is an alternating multilinear form on $V^{\oplus n}$. Then for every basis $\mathcal{X} = \{x_1, \dots, x_n\}$ of V, $w(x_1, \dots, x_n) \neq 0$.*

Proof. Consider $w(y_1, \dots, y_n)$ for general $y_i \in V$. Suppose that for each i, $y_i = \sum_j \alpha_{ij} x_j$. Then expanding by multilinearity and using alternation,

$$\begin{aligned} w(y_1, \dots, y_n) &= \sum_{j_1, \dots, j_n \in [n]^n} (\alpha_{1 j_1} \dots \alpha_{n j_n}) \, w(x_{j_1}, \dots, x_{j_n}) \\ &= \sum_{\pi \in S_n} \left(\prod_{i=1}^{n} \alpha_{i \pi(i)} \right) (\pi w)(x_1, \dots, x_n) \\ &= \left(\sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} \alpha_{i \pi(i)} \right) w(x_1, \dots, x_n). \end{aligned}$$

Thus $w(x_1, \dots, x_n) = 0 \iff w = 0$ as needed. □

Using the same proof idea yields the following useful result:

Exercise 45. Any two alternating n-linear forms on a vector space of dimension n are linearly dependent, i.e. scalar multiples of each other.

Exercise 46. What is the dimension of the space of all symmetric k-linear forms? What about the skew-symmetric forms? The alternating forms? For simplicity, you may assume the vector space involved is $V^{\oplus n}$, and that $k \leq n$. (Hint: Pick a basis for V, and then expand out $w(y_1, \dots, y_k)$ for general $y_1, \dots, y_k \in V$. Which $w(\dots)$ values do you need, and which ones are then determined?)

To finish, we make one final point about the connections of multilinear forms to algebra: we have seen that 1-linear forms - the functionals - are dual to vectors, and similarly that 2-linear forms are dual to matrices. The general object here is a tensor, and makes up the study of tensor algebra.

1.13 Tensor Products - Rewrite!

We look at another way to put two vector spaces together to create a third. Consider the following motivating example. Let U be the set of polynomials in variable s and V that in t. Let W be the set of two-variable polynomials in variables s, t. W is not a sum of U and W, of course, but in a sense it *is* related to the vectors in U and V; for example, the polynomial product of $u \in U$ and $v \in V$ defines an element $z(s,t) = u(s)v(t)$ of W. And indeed, the cartesian product of the bases of U and V - the set $\{s^i : i \geq 0\} \times \{t^j : j \geq 0\}$ - is a basis for W. We shall now formalize this construction.

Definition 1.33. The tensor product $U \otimes V$ of two finite-dimensional vector spaces U and V (over the same field F) is the dual of the vector space of bilinear forms on $U \oplus V$.

For each $x, y \in U \oplus V$, we also define the element $z := x \otimes y$ in $U \otimes V$ by $z(w) = w(x, y)$ for all $w \in (U \oplus V)^{(2)}$. We call z the tensor product of x and y.

Notice that we get the product rule for dimensions: $\dim(U \otimes V) = \dim U \cdot \dim V$ for free from the definition.

Theorem 1.34. *If $\mathcal{X} = \{x_1, \dots, x_m\}$ is a basis for U and $\mathcal{Y} = \{y_1, \dots, y_n\}$ is a basis for V, then the set $\{z_{ij} = x_i \otimes y_j\}$ is a basis for $U \otimes V$.*

Proof. Let $\{w_{ij}\}$ be the (basis of) bilinear forms defined by $w_{ij}(x_k, y_l) = \delta_{ik}\delta_{jl}$. Notice that $z_{kl}(w_{ij}) = \delta_{(i,j)=(k,l)}$; since $\{w_{ij}\}$ is a basis, $\{z_{ij}\}$ is a (the dual) basis for $U \otimes V$. □

Exercise 47. Let U, V, W be finite-dimensional vector spaces over F. Prove that $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$ and $U \otimes V \cong V \otimes U$. Show that

$$U \otimes (V \oplus W) = (U \otimes V) \oplus (U \otimes W).$$

Chapter 2

Transformations

2.1 Linear Transformations

We come now to the objects that really make vector spaces interesting.

Definition 2.1. A linear transformation or operator A on a vector space V is a correspondence that assigns to each vector $x \in V$ a vector Ax in V in such a way that

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay$$

for every choice of scalars α and β and vectors x and y in V.

Example 5.

1. The zero transformation $0 : x \mapsto 0$ and the identity $1 : x \mapsto x$.
2. The projection: Let x_0 be a fixed vector in V, and y_0 any functional in V. Then the map $x \mapsto y_0(x)x_0$ is a linear transformation. Readers familiar with matrices will recognize the matrix $x_0 y_0^T$ associated to this transformation.

It is clear that the pointwise sum and scalar multiplication of linear transformations yield linear transformations. Thus the space of linear transformations also forms a vector space, which we denote by $\mathcal{L}(V)$ or $\mathcal{L}(V, V)$.

In particular, for arbitrary set $\{x_1, \ldots, x_n\}$ of vectors and linear functionals $\{y_1, \ldots, y_n\}$ the following function is a linear transformation:

$$x \mapsto \sum_{i=1}^{n} y_i(x)x_i.$$

Exercise 48. Suppose that $\mathcal{X} = \{x_1, \ldots, x_n\}$ is a basis for V. Show that every linear transformation $A \in \mathcal{L}(V, V)$ can be written uniquely as

$$Ax = \sum_{i=1}^{n} y_i(x)x_i$$

for some set $\{y_1, \ldots, y_n\}$ of linear functionals. Conclude that $\dim \mathcal{L}(V, V) = (\dim V)^2$.

Again, for those familiar with matrices, one can think of each functional (in vector form) as a row of the matrix A.

A few more examples of linear transformations:

Example 6.

1. Let V be the space of all polynomials in x of degree at most n. Then the derivative operator D is a linear transformation on V. Wrt the basis set $\{1, x, \ldots, x^n\}$, find the associated linear functionals as defined in the above exercise. Similarly the function $x(t) \mapsto x(t+1) - x(t)$ as well.
2. Let m be a polynomial in t. Then the map $p(t) \mapsto m(t)Dp(t)$ is a linear transformation on P.
3. Let $V = \mathbb{C}$; then the map $z \mapsto \bar{z}$ is a linear transformation on V.
4. Let V be the space of k-linear forms on a vector space; then the map $Aw = \sum_{\pi \in S_k} \operatorname{sgn}(\pi)\pi w$ is a linear transformation on V. This is the *alternating* operator.

A more general definition of a linear transformation allows for the domain and codomain to be different vector spaces.

Definition 2.2. Let V and W be vector spaces. A linear transformation A from V to W is a function $V \to W$ such that

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay$$

for every choice of scalars α and β and vectors x and y in V.

Again, the space of linear transformations from V to W forms a vector space denoted $\mathcal{L}(V, W)$.

Example 7.

1. Every linear functional on V is a linear transformation $V \to F$.
2. The projection operator $\langle x, t\rangle \mapsto x$ is a linear transformation $V \oplus W \to V$.
3. Let $N \leq V$ be a subspace. The map $x \mapsto x + N$ is a linear transformation $V \to V/N$.

The key thing about linear transformations are that they are functions - and so can be composed (if compatible).

Definition 2.3 (Product). Let $A : V \to W$ and $B : U \to V$ be linear transformations. Then the product $AB : U \to W$ is defined by

$$(AB)u = A(Bu).$$

Exercise 49. Show that the product of two linear transformation is also a linear transformation.

Notice that the product of linear transformations is not commutative in general. However, it is associative.

Given $A \in \mathcal{L}(V, V)$, we define the iterated multiplications in the natural way: $A^0 := 1$, $A^{n+1} = A^n A$. Note that A^m and A^n commute for all m and n. Given polynomial $p(t)$, we define $p(A)$ to be the linear transformation $p(A) = \sum_{i=0}^{n} a_i A^i \in \mathcal{L}(V, V)$.

One disconcerting thing that we shall get out of the way now:

Definition 2.4 (Divisors of Zero). Suppose $A \neq 0 \in \mathcal{L}(V, W)$, $B \neq 0 \in \mathcal{L}(U, V)$ are transformations satisfying $AB = 0$. Then A is a left divisor of zero and B is a right divisor of zero.

For example, denoting the iterated differentiation operator on P_n by D^k, we have $D^k D^{n-k} = 0$ for $1 < k < n$.

2.2 Inverses

Suppose that $A \in \mathcal{L}(V, V)$ is a bijection. We then say that A is invertible and denote its inverse by the inverse function A^{-1}.

Exercise 50. Show that A^{-1} is also a linear transformation $V \to V$.

Theorem 2.5 (Uniqueness of the inverse). *If A, B and C are linear transformations on V such that $AB = 1$ and $CA = 1$, then A is invertible and $B = C = A^{-1}$.*

Proof. We have $C = C1 = CAB = 1B = B$. A more illuminating proof follows. If $Ax_1 = Ax_2$, then $CAx_1 = CAx_2$, so $x_1 = x_2$. Thus A is injective. Similarly, suppose that $y \in V$. Then $y = A(By)$, so A is surjective. Thus A is invertible. Multiply by A^{-1} on the left to get $B = A^{-1}$ and on the right to get $C = A^{-1}$. □

Exercise 51. If A is invertible, show that A^{-1} is invertible and that $(A^{-1})^{-1} = A$.

Exercise 52. Suppose A and B are linear transformations on the same vector space V such that $AB = 1$. We say A is a left inverse for B and B is a right inverse for A. Suppose that A has a unique right inverse B. Show that A is invertible and $B = A^{-1}$. (Hint: $A(BA - 1) = 0$).

A more remarkable result holds for linear transformations over finite-dimensional vector spaces: one of injectivity or surjectivity is enough to guarantee invertibility:

Theorem 2.6 (Invertible linear transformations). *Let A be a linear transformation on a finite-dimensional vector space V. Then the following are equivalent:*

1. *A is invertible.*
2. *A is injective. Equivalently, $Ax = 0$ implies $x = 0$.*
3. *A is surjective. For every $y \in V$, $\exists x \in V$ such that $Ax = y$.*

Proof. (1) $\implies$ (2), (3) are trivial. We establish the reverse implications.

(2) $\implies$ (1): Suppose A is injective. Let $\mathcal{X} = \{x_1, \ldots, x_n\}$ be a basis for V. By injectivity and linearity, it is easy to verify that $A\mathcal{X} = \{Ax_1, \ldots, Ax_n\}$ must be a basis for V. But the range of A is precisely the span of $A\mathcal{X}$ and so is V. Thus A is surjective - and hence invertible.

(3) $\implies$ (1): Suppose A is surjective. Let $\mathcal{Y} = \{y_1, \ldots, y_n\}$ be a basis for V. Choose x_i such that $Ax_i = y_i$ for each i. Then $\mathcal{X} = \{x_i\}$ is a basis for V, because $\sum_i \alpha_i x_i = 0 \implies \sum_i \alpha_i y_i = 0$. So every x can be written $\sum_i \alpha_i x_i$ and so $Ax = 0 \iff \sum_i \alpha_i y_i = 0 \iff \alpha_i = 0$ for all i i.e. $x = 0$. Thus A is injective - and hence invertible. □

Exercise 53. Suppose V is finite-dimensional. Then $AB = 1$ implies that both A and B are invertible (and inverses of each other).

Exercise 54. If A and B are invertible, then show that AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$. Let $\alpha \neq 0$, then prove that αA is invertible and $(\alpha A)^{-1} = \alpha^{-1}A^{-1}$.

The above exercises establish that A^n is invertible for every invertible A and integer n.

Exercise 55. Show that

$$A, B \text{ invertible} \iff AB, BA \text{ invertible.}$$

2.3 Matrices

Definition 2.7 (Matrix). An $m \times n$ matrix A on a set S is a function $A : [m] \times [n] \to S$. A matrix is typically denoted by a rectangular array:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

where we use A_{ij} to denote the value of A at (i, j). A matrix is often also denoted by the numbers A_{ij} written as a doubly-indexed set - for example, $\{A_{ij}\}_{i,j}$ where the bounds of i (rows) and j (columns) are clear from context.

A matrix is said to be square if $m = n$.

Under pointwise addition and scalar mulitplication, the set of all $m \times n$ matrices on field F forms a vector space, denoted $M_{m,n}(F)$ or $F^{m \times n}$.

The matrix is a most useful tool for finite-dimensional vector spaces.

Definition 2.8. Let V be an n-dimensional vector space and $\mathcal{X}$ be a basis for V. Let A be a linear transformation on V. Since every vector is a linear combination of the x_is, we have in particular

$$Ax_j = \sum_i \alpha_{ij} x_i$$

for each $j = 1, \dots, n$. The set of scalars $\{\alpha_{ij}\}_{i,j}$ define the matrix of transformation A wrt basis $\mathcal{X}$ and is denoted $[A; \mathcal{X}]$ or $[A]_{\mathcal{X}}^{\mathcal{X}}$ or $[A]_{\mathcal{X}}$ or simply $[A]$.

Remark. Note that the appearance of the square array associated with $[A]$ varies with the ordering of $\mathcal{X}$ - the matrix is thus ideally associated with a transformation, a basis, and an ordering of that basis.

Remark. Notice the unusual indexing:

$$Ax_j = \sum_i \alpha_{ij} x_i$$

instead of the more "natural"

$$Ax_i = \sum_j \alpha_{ij} x_j.$$

This stems from the fact that vectors are represented as column vectors - that is, matrices with a single column. Following this convention, the vector Ax_j is more ideally represented as a column of $[A]$, and not a row (this is not just a choice of aesthetics; matrix multiplication is also simplified by this convention). The jth column being Ax_j necessitates the "unusual" indexing (the first index of the matrix is always the row, and the second the column).

If transformations A, B have matrices $[A]$, $[B]$, what about $[\alpha A + \beta B]$? We want the matrix to continue to satisfy its defining point of correspondence to the transformation:

$$Ax_j = \sum_i [A]_{ij} x_i.$$

Exercise 56. Show that $[\alpha A + \beta B] = \alpha[A] + \beta[B]$. This equation is nice thanks to our definition of matrix addition and scalar multiplication.

What about $[AB]$?

Exercise 57. We would like to define a multiplication operation for matrices that has the pleasing property

$$[AB] = [A][B].$$

What is the notion of matrix multiplication that we must use?

The following definition of matrix multiplication looks strange at first, but is naturally implied by the above exercise.

Definition 2.9 (Matrix multiplication). Let A be an $m \times n$ matrix and B an $n \times p$ matrix. Then the product AB is the $m \times p$ matrix defined by

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}.$$

Exercise 58 (The reason for the strange indexing). Suppose that we defined $[A]$ instead by the more natural choice,

$$Ax_i = \sum_j [A]_{ij} x_j.$$

Show that the product of matrices AB would then be defined by

$$[AB]_{ij} = \sum_k [B]_{ik} [A]_{kj}$$

instead - which "looks" more like a reasonable definition for $[B][A]$ that $[A][B]$! The difficulty rose precisely because B is applied first in the transformation AB - the inverted indices serve exactly to counter this when doing matrix mulitplication.

Definition 2.10. We associate to every vector $x = (\zeta_1, \ldots, \zeta_n) \in F^n$ the column vector

$$[x] = \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_n \end{bmatrix} \in F^{n \times 1}.$$

Let V be a finite-dimensional vector space, and fix basis $\mathcal{X}$ for V. As in the proof of Theorem 1.11, we associate with each $x \in V$ the solumn vector $[x]_{\mathcal{X}} \in F^{n \times 1}$ defined by

$$x = \sum_i \zeta_i x_i \iff [x]_{\mathcal{X}} = \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_n \end{bmatrix}.$$

We call $[x]_{\mathcal{X}}$ the coordinate vector of x wrt $\mathcal{X}$.

Exercise 59 (Multiplying matrices with vectors). Fix a finite-dimensional vector space V, a basis $\mathfrak{X}$ and a transformation A. Show that for each $x \in V$,

$$[Ax]_{\mathfrak{X}} = [A]_{\mathfrak{X}}[x]_{\mathfrak{X}}.$$

At first sight, this seems a pleasant coincidence. However, one can make it more natural: Treat each vector x as being the linear transformation $\alpha \mapsto \alpha x$ from F to V. One then sees that $Ax : F \to V$ is the composition of $x : F \to V$ and $A : V \to V$, since for each $\alpha \in F$

$$(Ax)(\alpha) = \alpha Ax = A(\alpha x) = A(x(\alpha)).$$

Thus $[Ax]_{\mathfrak{X}} = [A]_{\mathfrak{X}}[x]_{\mathfrak{X}}$ is just the usual equation one expects of a composition of transformations. (This is a bit of a stretch, but it is a nice way to think about it).

Note that we have not defined the matrix of a linear transformation from V to a different space W, but we have used it in the above exercise; we shall get to it shortly. Note also that $[x]_{\mathfrak{X}}$ is the matrix of the linear transformation x when the basis of F used is $\{1\}$.

Exercise 60. Show that matrix multiplication is associative and distributive over addition.

We next prove a result that establishes much the same thing as Theorem 1.11 did.

Theorem 2.11 (Isomorphism between linear transformations and matrices). *Fix finite-dimensional vector space V and basis $\mathfrak{X}$. Let $n = \dim V$. Then the map $A \mapsto [A]_{\mathfrak{X}}$ is an isomorphism between vector spaces $\mathcal{L}(V)$ and $F^{n\times n}$.*

Proof. It is clear that the map is linear. As in the case of linear functionals, it is clear that the value of A on $\mathfrak{X}$ determines A. Since the matrix $[A]_{\mathfrak{X}}$ precisely specifies this information, the map is injective. Further, from each matrix M, one can read off a linear transformation A by defining for each $x = \sum_i \zeta_i x_i$

$$Ax = \sum_j \zeta_j \left(\sum_i M_{ij} x_i \right).$$

This is clearly linear, and so the map is surjective. Thus the map is an isomorphism. □

Exercise 61. Extend the notion of the matrix associated to a linear transformation to the case where the domain and codomain are different vector spaces. Verify that the matrices of addition and composition of linear transformations still have the same symbolic form as before.

Thus in the general case (transformations $V \to W$), not every matrix representation is square, the matrix products AB and BA may have altogether different dimensions, and finally that matrices A and B many not even be multipliable.

Exercise 62. Suppose that A and B are multipliable matrices. Partition A into four rectangular blocks like so:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{ij} are now matrices. Partition B similarly so that the number of columns in the top left part of A is the same as the number of rows of the top left part of B. B now looks like

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Show that

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

Next, use subspaces and complements to express the above result in terms of linear transformations. Finally, can you generalize the result and interpretation to a larger number of blocks?

The form of writing a matrix in terms of matrices of linear transformations on subspaces is extremely useful at times and is called a block form of the matrix. Matrices written in terms of blocks are typically called block matrices.

2.4 Invariance and Projections

Definition 2.12 (Invariance). Let A be a linear transformation on V and let $M \leq V$ be a subspace. We say that M is invariant under V if

$$AM = \{Ax : x \in M\} \leq M,$$

that is, for every $x \in M$, $Ax \in M$.

Exercise 63. Suppose M is invariant under A. Pick a basis $\mathcal{X} = \{x_1, \ldots, x_k\}$ for M and extend it to a basis $\mathcal{Y} = \{x_1, \ldots, x_k, y_{k+1}, \ldots, y_n\}$ for V. Show that the matrix of A wrt $\mathcal{Y}$ has the form

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

where $A_{11} \in F^{k \times k}$ (and of course, 0 is the zero matrix in $F^{(n-k)\times k}$).

Definition 2.13 (Reducible linear transformations). Let A be a linear transformation and V be a vector space. We say that A is reducible to (M, N) if M, N are invariant under A and $V = M \oplus N$.

The above definition can also be turned around: Let M, N be two vector spaces, and let $V = M \oplus N$. Suppose A and B are linear transformations on M, N respectively. Then we can define linear transformation C on V, denoted $A \oplus B$, by

$$C\langle x, y\rangle = \langle Ax, By\rangle.$$

Identifying M, N with subspaces of V, C is reducible to (M, N) with the restrictions of C to M and N being A and B respectively.

Exercise 64. Fix A, M, N; suppose that $\mathcal{X}$ is a basis for M and $\mathcal{Y}$ is a basis for N. Show that A is reducible to (M, N) if and only if the matrix of A wrt basis $\mathcal{X} \cup \mathcal{Y}$ has the form

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

where A_{11} is the matrix of the restriction of A to M wrt $\mathcal{X}$ and A_{22} is the matrix of the restriction of A to N wrt $\mathcal{Y}$.

It is also easy to see that if $C = A \oplus B$, then $C^n = A^n \oplus B^n$, and in general for polynomial $p \in P$ we have $p(C) = p(A) \oplus p(B)$.

Another connection between direct sums and linear transformations is that of projections. These are a very powerful algebraic tool in studying the concept of direct sum.

Definition 2.14. If $V = M \oplus N$ - so that every $x \in V$ may be written uniquely as a sum $x + y$ with $x \in M$ and $y \in N$ - then the projection on M along N is the linear transformation P on V defined by

$$Pz = x.$$

Theorem 2.15. *A linear transformation E is a projection on some subspace if and only if it is idempotent, that is, $E^2 = E$ or $E(1 - E) = 0$.*

Proof. $\implies$ is trivial. For $\impliedby$, let

$$M = \{z \in V : Ez = z\}$$

and

$$N = \{z \in V : Ez = 0\}.$$

By linearity of E, these are subspaces. We claim that $V = M \oplus N$. Indeed, for each z, we may write $z = Ez + (1 - E)z$ with the first term in M and the second in N. Further, if $z \in M \cap N$, then $z = Ez = 0$. Thus $V = M \oplus N$. To finish, note that for each $z \in V$, $Ez = E(Ez)$ and so E is the projection on M along N. □

Exercise 65. Suppose that E is a projection on M along N.

1. Show that E is reducible to (M, N).
2. Show that $M = \{z : Ez = z\}$ and $N = \{z : Ez = 0\}$.
3. Show that $1 - E$ is also a projection; in fact, it is the projection on N along M.