CHAPTER 1

Linear Vector Spaces

Definition 1.0.1. A linear vector space over a field F is a triple $(V, +, \cdot)$, where V is a set, $+ : V \times V \to V$ and $\cdot : F \times V \to V$ are maps with the properties :

(i) $(\forall x, y \in V)\,,\ x + y = y + x$.

(ii) $(\forall x, y, z \in V)\,,\ x + (y + z) = (x + y) + z$.

(iii) $\exists 0_V \in V\ :\ (\forall x \in V)\,,\ x + 0_V = 0_V + x = x$.

(iv) $\forall x \in V,\ \exists - x \in V\ :\ x + (-x) = (-x) + x = 0_V$.

(v) $(\forall \lambda \in F, x, y \in V)\ \ \lambda\,(x + y) = \lambda x + \lambda y$.

(vi) $(\forall \lambda, \mu \in F, x \in V)\ (\lambda + \mu)\,x = \lambda x + \mu x$.

(vii) $(\forall \lambda, \mu \in F, x \in V)\ (\lambda\mu)\,x = \lambda\,(\mu x)$.

(viii) $(\forall x \in V)\ 1_F x = x$.

Note that in the above and the following we will abbreviate $\lambda \cdot x$ by λx. We sometimes talk of the map $\cdot$ as the action of the field F on the vector space V. Sometimes we will refer to the set V as the vector space (where the $+$ and $\cdot$ is obvious from the context). Elements of the set V are called vectors, while those of F are called scalars. If the field F is either $\mathbb{R}$ or $\mathbb{C}$ (which are the only cases we will be interested in), we call V a real vector space or a complex vector space, respectively.

Example 1.0.2. (1) The set $\mathbb{R}^n$ of n-tuples $(x_1, x_2, \ldots, x_n)$ of real numbers with $+$ and $\cdot$ defined by

$$\begin{aligned} (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \\ \lambda \cdot (x_1, x_2, \ldots, x_n) &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n) \end{aligned}$$

is a linear vector space over the field $\mathbb{R}$.

(2) The set $\mathbb{C}^n$ of n-tuples of complex numbers (similar).

(3) Let $\mathrm{Mat}_{m\times n}(F)$ be the set of all $m \times n$ F-valued matrices. Then $\mathrm{Mat}_{m\times n}(F)$ is a vector space under usual addition of matrices and multiplication by scalars.

(4) Let $\mathbb{R}_{n+1}[X]$ be the set of all polynomials up to degree n, *i.e.* all expressions of the form $a_0 + a_1 X + a_2 X^2 + \ldots + a_n X^n$, where $a_i \in \mathbb{R}$, is a vector space over $\mathbb{R}$ with the usual definitions of addition of polynomials and multiplication of polynomials by numbers. (Note that we use the suffix $n+1$ and not n - because the number of parameters you need to specify a polynomial is $n+1$ - some authors refer to this set as $\mathbb{R}_n[X]$.

(5) $\mathbb{R}$ is a vector space over $\mathbb{R}$! Similarly $\mathbb{C}$ is one over $\mathbb{C}$. Note that $\mathbb{C}$ is also a vector space over $\mathbb{R}$ - though a different one from the previous example! Also note that $\mathbb{R}$ is *not* a vector space over $\mathbb{C}$.

Theorem 1.0.3. *If V is a vector space over F, then*

(1) $(\forall \lambda \in F)\ \lambda 0_V = 0_V$.

(2) $(\forall x \in V)\ 0_F x = 0_V$.

(3) *If* $\lambda x = 0_V$ *then either* $\lambda = 0_F$ *or* $x = 0_V$.

(4) $(\forall x \in V,\ \lambda \in F)\ \ (-\lambda)\,x = -\,(\lambda x) = \lambda\,(-x)$.

Proof. Exercise! ☐

Definition 1.0.4. Let $(V, +, \cdot)$ be a vector space over a field F. If a nonempty subset $W \subseteq V$ is a vector space over the same field F with the operations $+$ and $\cdot$ being restricted to $W \times W$ and $F \times W$, respectively, then $(W, +, \cdot)$ is called a subspace of $(V, +, \cdot)$. We usually say that W is a subspace of V.

THEOREM 1.0.5. *A nonempty $W \subseteq V$ is a subspace of V if*

$$(\forall x, y \in W,\ \lambda, \mu \in F)\ :\ \lambda x + \mu y \in W$$

EXAMPLE 1.0.6. (1) Every vector space is (trivially) a subspace of itself.

(2) Every subspace of V has the element 0_V. Thus the smallest subspace of V is the singleton set $\{0_V\}$.

(3) In $\mathbb{R}^3$ the subspaces are (i) the origin, (ii) lines passing through the origin, (iii) planes passing through the origin, and (iv) the whole space.

THEOREM 1.0.7. *The intersection of any collection of subspaces of V is a subspace of V.*

PROOF. Let C be a collection of subspaces of V and let $T = \cap_{W\in C} W$ be their intersection. Since $0_V \in W$, $\forall W \in C$, $0_V \in T$, so that $T \neq \emptyset$. Let $x, y \in T$. Then $\forall W \in C$, $x, y \in W$ and thus $\lambda x + \mu y \in W$ $\forall \lambda, \mu \in F$. So $\lambda x + \mu y \in T$, and thus T is a subspace. □

Note that this is not true for a union.

Given a subset $S \subseteq V$, we define $\langle S\rangle$ as the intersection of all subspaces of V that contain S. This is a subspace of V and is the smallest subspace containing S.

If $(x, y) \in \mathbb{R}^2$ is not the origin and $S = \{(x, y)\}$, then $\langle S\rangle$ is the line joining (x, y) to the origin. If $S = \emptyset$, then $\langle S\rangle$ is $\{0_V\}$.

DEFINITION 1.0.8. Let V be a vector space over F and let S be a nonempty subset V. Then we say that $v \in V$ is a linear combination of elements of S if $\exists x_1, x_2, \ldots, x_r \in S$ and $\lambda_1, \lambda_2, \ldots, \lambda_r \in F$ such that

$$v = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_r x_r$$

This helps us to define

DEFINITION 1.0.9. Given a nonempty subset $S \subseteq V$, the span of S, denoted $\operatorname{Span} S$ is defined to be the set of all linear combinations of elements of S.

Since linear combinations of linear combinations are linear combinations, it is obvious that $\operatorname{Span} S$ is a subspace. If $S = \emptyset$, we define $\operatorname{Span} S$ to be the singleton vector space $\{0_V\}$.

THEOREM 1.0.10. $\langle S\rangle = \mathit{Span}\, S$

PROOF. The result follows trivially for $S = \emptyset$. Let us now consider $S \neq \emptyset$.

$\forall x \in S$, since $1_F x = x \in \operatorname{Span} S$, we have $S \subseteq \operatorname{Span} S$. Thus $\operatorname{Span} S$ is a subspace containing S. Since $\langle S\rangle$ is the smallest subspace containing S, we have $\langle S\rangle \subseteq \operatorname{Span} S$.

Let $v \in \operatorname{Span} S$. Then, $\exists x_1, \ldots, x_r \in S$, $\lambda_1, \ldots, \lambda_r \in F$ such that $v = \sum_{i=1}^{r} \lambda_i x_i$. If W is a subspace of V containing S, then $x_1, \ldots, x_r \in W$ and thus $v \in W$. Since $\langle S\rangle$ is a subspace containing S, it follows that $v \in \langle S\rangle$ and hence $\operatorname{Span} S \subseteq \langle S\rangle$. □

DEFINITION 1.0.11. A subset S of a vector space V is called a spanning set if $\operatorname{Span} S = V$.

EXAMPLE 1.0.12. In $\mathbb{R}^3$, the vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ form a spanning set, since any vector $(x, y, z) \in \mathbb{R}^3$ can be written as a linear combination of these three. Of course, this is not the only choice - for example, the set $\{(1, 1, 1), (1, -2, 1), (0, 1, 3), (0, 0, 1), (-1, 5, 7)\}$ is also a spanning set.

DEFINITION 1.0.13. A nonempty subset S of a vector space V is said to be linearly independent if the only way of expressing 0_V as a linear combination of elements of S is the trivial way, *i.e.* if $x_1, x_2, \ldots, x_n \in S$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in F$ then

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n = 0 \implies \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0.$$

A subset is called linearly dependent if it is not independent.

EXAMPLE 1.0.14. In $\mathbb{R}^3$, the vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ form a linearly independent set.

THEOREM 1.0.15. *No linearly independent subset of a vector space V can contain the vector 0_V.*

We can restate the theorem above as "every subset of a vector space containing the null vector is linearly dependent".

THEOREM 1.0.16. *A superset of a linearly dependent set is linearly dependent. Every subset of a linearly independent set is linearly independent.*

PROOF. Let S be linearly dependent and let $T \supseteq S$. Then, $\exists x_1, \dots, x_n \in S$ such that there is a nontrivial linear combination of these elements that yields 0_V. Since $x_1, \dots, x_n \in T$ as well, T is linearly dependent. The proof of the next part of the statement follows trivially. □

THEOREM 1.0.17. *Let V be a vector space over a field F. If $S \subseteq V$ has at least two elements then it is linearly dependent iff one of its elements can be written as a linear combination of the others.*

PROOF. (only if) Let S be linearly dependent. Then, $\exists x_1, \dots, x_n \in S$ and $\lambda_1, \dots, \lambda_n \in F$ such that $\lambda_1 x_1 + \ldots + \lambda_n x_n = 0$, with not all the $\lambda_i = 0$. Since $\exists \lambda_i \neq 0$, we can relabel if necessary so that $\lambda_1 \neq 0$. Then, we can write

$$x_1 = \left(-\frac{\lambda_2}{\lambda_1}\right) x_2 + \ldots + \left(-\frac{\lambda_n}{\lambda_1}\right) x_n.$$

(if) If it is possible to write one of the elements of S, say x_1, as a linear combination of the rest

$$x_1 = \mu_2 x_2 + \ldots + \mu_n x_n$$

Then we can write

$$(-1)\, x_1 + \mu_2 x_2 + \ldots + \mu_n x_n = 0_V$$

which is a non-trivial linear combination. □

DEFINITION 1.0.18. A linearly independent spanning set of a vector space is called its basis.

EXAMPLE 1.0.19. In $\mathbb{R}^n$ the n vectors $(1, 0, 0, \dots, 0)$, $(0, 1, 0, \dots, 0)$, …,$(0, 0, \dots, 0, 1)$ form a basis. This is called the natural basis of $\mathbb{R}^n$.

THEOREM 1.0.20. *A nonempty subset S of a vector space V is a basis of V iff every element of V can be expressed in a unique way as a linear combination of elements of S.*

PROOF. (only if) Let S be a basis. Then $\operatorname{Span} S = V$ and thus $\forall x \in V$ we can expand x as a linear combination $\sum_i \lambda_i x_i$ of elements of S. To see that this is unique, note that $\sum_i \lambda_i x_i = \sum_i \mu_i x_i$ implies $\sum_i (\lambda_i - \mu_i)\, x_i = 0_V$ and since the x_i's are independent, we have $\lambda_i - \mu_i = 0$ for all i, so that $\lambda_i = \mu_i$.

(if) Suppose that every element of V can be expressed in a unique way as a linear combination of elements of S. Then $\operatorname{Span} S = V$. Again, since $\sum_i 0_F x_i = 0_V$ and the expansion is unique for all $x \in V$, the only linear combination that gives the null vector is the trivial one - so that S is also linearly independent. □

EXAMPLE 1.0.21. Every element in $\mathbb{R}_{n+1}[X]$ can be expanded as a unique linear combination $a_0 + a_1 X + a_2 X^2 + \ldots + a_n X^n$ of elements of the subset $\{1, X, X^2, \dots, X^n\}$, and hence this is a basis.

EXAMPLE 1.0.22. For $i = 1, \dots, n$ let $a_i = \{a_{i1}, a_{i2}, \dots, a_{in}\}$. Then $\{a_1, a_2, \dots, a_n\}$ is a basis of $\mathbb{R}^n$ iff the matrix $A = [a_{ij}]_{n \times n}$ is invertible. To see this, let $x = (x_1, x_2, \dots, x_n)$ and consider the expansion

$$x = \lambda_1 a_1 + \ldots + \lambda_n a_n.$$

This corresponds to the system of equations

$$A^t \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

So, an unique linear combination is possible if and only if A^t is invertible, *i.e.* A is invertible.

THEOREM 1.0.23. *Let V be a vector space that is spanned by a finite set $\{v_1, \dots, v_n\}$. If $I = \{w_1, \dots, w_m\}$ is a linearly independent subset of V then necessarily $m \leq n$.*

PROOF. Consider $w_1 \in I$. Since $\{v_1, \ldots, v_n\}$ is a spanning set, $\exists \lambda_1, \ldots, \lambda_n \in F$ such that $w_1 = \lambda_1 v_1 + \ldots + \lambda_n v_n$. Since $w_1 \neq 0_V$ (because I is linearly independent), not all $\lambda_i = 0_F$. By a change of indices if necessary, we can take $\lambda_1 \neq 0_F$. Then $v_1 = \lambda_1^{-1} w_1 - \lambda_1^{-1}\lambda_2 v_2 - \ldots - \lambda_1^{-1}\lambda_n v_n$, which shows that

$$V = \text{Span}\ \{v_1, v_2, \ldots, v_n\} \subseteq \text{Span}\ \{w_1, v_2, \ldots, v_n\} = V.$$

Again, $w_2 = \mu_1 w_1 + \mu_2 v_2 + \ldots + \mu_n v_n$ and not all of $\mu_2, \ldots, \mu_n = 0_F$ (otherwise we would have $w_2 = \mu_1 w_1$ - contradicing the independence of I). Again, we can relabel if necessary to get $\mu_2 \neq 0_F$, and show that $v_2 = \mu_2^{-1} w_2 - \mu_2^{-1}\mu_1 w_1 - \mu_2^{-1}\mu_3 v_3 - \ldots - \mu_2^{-1}\mu_n v_n$. This helps us to show that Span $\{w_1, w_2, v_3, \ldots, v_n\} = V$.

If $n < m$, we could proceed in the same way and finally get Span $\{w_1, w_2, \ldots, w_n\} = V$. Then we will have a contradiction, because we would be able to express the rest of the w's - $w_{n+1}, \ldots, w_m$ as linear combinations of $w_1, \ldots, w_n$. Thus $m \leq n$. □

In other words, you can not have more linearly independent vectors than there are in a spanning set.

COROLLARY 1.0.24. *If V has a finite basis B then every basis of V is finite and has the same number of elements as B.*

PROOF. Let B^* be another basis for V. If B^* is an infinite set then, since any subset of a linearly independent set is linearly independent, we can get finite linearly independent subsets of B^* with arbitrarily many elements, thus contradicting the last theorem. So, B^* must be finite.

Since B is a spanning set and B^* is linearly independent, we have $|B| \geq |B^*|$. By reversing the two bases, we have $|B^*| \geq |B|$. Thus $|B| = |B^*|$. □

DEFINITION 1.0.25. A finite dimensional vector space V is one which has a finite basis. The number of elements in any basis of V is called the dimension of V, and is denoted by $\dim V$.

EXAMPLE 1.0.26. A basis for the vector space $\text{Mat}_{m\times n}(F)$ is the set $\{e_{pq} : 1 \leq p \leq m,\ 1 \leq q \leq n\}$ where e_{pq} is a $m \times n$ matrix with the elements $(e_{pq})_{ij} = \delta_{ip}\delta_{jq}$. Thus $\dim \text{Mat}_{m\times n}(F) = mn$.

THEOREM 1.0.27. *Let V be a finite dimensional vector space. If G is a finite spanning set of V and if $I \subseteq G$ is linearly independent, then there is a basis B of V such that $I \subseteq B \subseteq G$.*

PROOF. If $\text{Span}\, I = V$ then $B = I$ and we have nothing left to prove. Suppose $\text{Span}\, I \neq V$. Then $I \neq G$ and $G \setminus \text{Span}\, I \neq \emptyset$ (If $G \subseteq \text{Span}\, I$, we would have $\text{Span}\, I \supseteq \text{Span}\, G = V$). Thus, $\exists g_1 \in G$ such that $g_1 \notin \text{Span}\, I$. Then $I' = I \cup \{g_1\} \subseteq G$ is linearly independent. If $\text{Span}\, I' = V$,we are done. If not, $\exists g_2 \in G$ such that $g_2 \notin \text{Span}\, I'$ and we get a linearly independent set $I'' = I \cup \{g_1, g_2\}$. If $\text{Span}\, I'' = V$ we are done. If not, we continue. Since G is a finite set, this process will come to an end. Finally we will get a linearly independent set $I \cup \{g_1, g_2, \ldots, g_r\} \subseteq G$ which spans V. This is the basis B and obviously $I \subseteq B \subseteq G$. □

COROLLARY 1.0.28. *Every linearly independent subset I of a finite dimensional vector space V can be extended to form a basis.*

PROOF. Since V is finite dimensional, it has a finite basis B. The subset $I \cup B \supseteq I$ is a finite set that spans V. Now we can use the theorem above. □

COROLLARY 1.0.29. *Every finite spanning set of a finite dimensional vector space contains a basis.*

PROOF. Let G be a finite spanning set of V. Since $\emptyset \subseteq G$ is linearly independent, according to the theorem we have a basis B of V such that $\emptyset \subseteq B \subseteq G$. □

COROLLARY 1.0.30. *If V is of dimension n then every linearly independent set consisting of $n$elements is a basis of V.*

PROOF. If I is a linearly independent set containing n elements, then $\exists B$, a basis of V containing I. But $|B| = n$ and thus $B = I$. □

COROLLARY 1.0.31. *If S is a subset of a finite dimensional vector space V, then the following statements are equivalent.*

(1) *S is a basis.*

(2) *S is a maximal independent set (*i.e. *if I is an independent set with $S \subseteq I$, then $S = I$).*

(3) *S is a minimal spanning set (*i.e. *if $G \subseteq S$ spans V then $G = S$).*

PROOF. (1) $\implies$ (2) : If I is independent with $S \subseteq I$, there is a basis B with $I \subseteq B$. Then $I \subseteq B$ - but $|I| = |B|$. Thus $I = B$.

(2) $\implies$ (1): By the first Corollary above, there exists a basis B with $S \subseteq B$. But B is independent and thus, by hypothesis $B = S$.

(1) $\implies$ (3) : If $\operatorname{Span} G = V$ then $\exists B \subseteq G$ where B is a basis of V. Thus $B \subseteq G \subseteq S$. But since both B and S are bases, we have $|B| = |S|$, and thus $B = S$.

(3) $\implies$ (1) : Since S spans V, there is a basis B with $\emptyset \subseteq B \subseteq S$. But B then also spans V and according to the hypothesis, $B = S$. ☐

COROLLARY 1.0.32. *If V is of dimension n then every subset containing more than n elements is linearly dependent. Again, no subset with less than n elements can span V.*

THEOREM 1.0.33. *Let V be afinite dimensional vector space. If $W$is a subspace of V, then W is also finite dimensional, and*

$$\dim W \leq \dim V.$$

Moreover, we have $\dim W = \dim V$ *iff* $W = V$.

PROOF. Let $\dim V = n$. If I is a linearly independent subset of W then it is also a linearly independent subset of V and thus $|I| \leq n$. Thus, a maximal such subset B exists and is, then, a basis of W. Hence W is also of finite dimension and $\dim W \leq \dim V$.

Again, if $\dim W = \dim V$, we have $|B| = n$. Thus B is a linearly independent subset of V with $n$elements. So, $B$is also a basis of V and thus $W = \operatorname{Span} B = V$. ☐

CHAPTER 2

Linear mappings

DEFINITION 2.0.34. If V and W are vector spaces over the same field F, then a map $f : V \to W$ is called a linear map if

(1) $(\forall x, y \in V) \qquad f(x+y) = f(x) + f(y)$, and

(2) $(\forall x \in V)(\forall \lambda \in F) \qquad f(\lambda x) = \lambda f(x)$.

EXAMPLE 2.0.35. The map $f : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$(x, y) \mapsto (x+y, x-y, y)$$

is linear.

EXAMPLE 2.0.36. The n maps $\mathrm{pr}_i : \mathbb{R}^n \to \mathbb{R}$ defined by

$$(x_1, \ldots, x_i, \ldots, x_n) \mapsto x_i$$

are linear. pr_i is called the i-th projection map.

EXAMPLE 2.0.37. The map $D : \mathbb{R}_{n+1}[X] \to \mathbb{R}_{n+1}[X]$ defined by

$$a_0 + a_1 X + a_2 X^2 + \ldots + a_n X^n \mapsto a_1 + 2a_2 X + \ldots + n a_n X^{n-1}$$

is linear.

THEOREM 2.0.38. *If the map $f : V \to W$ is linear, then*

(1) $\quad f(0_V) = 0_W$, *and*

(2) $\quad f(-x) = -f(x)$

PROOF. (1) $f(0_V) = f(0_F x) = 0_F f(x) = 0_W$.
(2) $0_W = f(0_V) = f(x + (-x)) = f(x) + f(-x)$ - which gives the result if we add $-f(x)$ to both sides. □

DEFINITION 2.0.39. Given a linear map $f : V \to W$ and $\lambda \in F$ we define the map $\lambda f : V \to W$ by

$$(\forall v \in V) \quad (\lambda f)(v) = \lambda f(v)$$

DEFINITION 2.0.40. Given two linear maps $f, g : V \to W$ we define the sum map $f + g : V \to W$ by

$$(\forall v \in V) \quad (f+g)(v) = f(v) + g(v)$$

It is easy to check that the maps λf and $f + g$ defined above are linear. If we denote the set of all linear maps from V to W by $\mathrm{Lin}(V, W)$, this shows that $\mathrm{Lin}(V, W)$ is a vector space over F under the addition and multiplication by scalar operation defined above.

DEFINITION 2.0.41. If $f : V \to W$ is linear we define $f^{\rightarrow} : 2^V \to 2^W$ by

$$(\forall X \subseteq V) \quad f^{\rightarrow}(X) = \{f(x) : x \in X\}$$

and $f^{\leftarrow} : 2^W \to 2^V$ by

$$(\forall Y \subseteq W) \quad f^{\leftarrow}(Y) = \{x \in V : f(x) \in Y\}.$$

We call $f^{\rightarrow}(X)$ the direct image of X under f, while $f^{\leftarrow}(Y)$ is called the inverse image of Y under f.

Note that most authors denote $f^{\rightarrow}(X)$ and $f^{\leftarrow}(Y)$ by $f(X)$ and $f^{-1}(Y)$, respectively.

It is easy to see that the maps $f^{\rightarrow}$ and $f^{\leftarrow}$ are inclusion preserving, *i.e.* if $X_1 \subseteq X_2 \subseteq V$ then $f^{\rightarrow}(X_1) \subseteq f^{\rightarrow}(X_1)$ and if $Y_1 \subseteq Y_2 \subseteq W$ then $f^{\leftarrow}(Y_1) \subseteq f^{\leftarrow}(Y_2)$.

Theorem 2.0.42. *Both $f^{\rightarrow}$ and $f^{\leftarrow}$ carry subspaces into subspace,* i.e. *if X is a subspace of V, then $f^{\rightarrow}(X)$ is a subspace of W and if Y is a subspace of W then $f^{\leftarrow}(Y)$ is a subspace of V.*

Proof. Let X be a subspace of V. Let $w_1, w_2 \in f^{\rightarrow}(X)$ and $\lambda_1, \lambda_2 \in F$. Then $\exists x_1, x_2 \in X$ such that $f(x_1) = w_1$ and $f(x_2) = w_2$. Then

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 f(x_1) + \lambda_2 f(x_2) = f(\lambda_1 x_1 + \lambda_2 x_2) \in f^{\rightarrow}(X)$$

Thus, $f^{\rightarrow}(X)$ is a subspace.

Again, let Y be a subspace of W. Let $v_1, v_2 \in f^{\leftarrow}(Y)$ and $\lambda_1, \lambda_2 \in F$. Then $\exists y_1, y_2 \in Y$ such that $f(v_1) = y_1$ and $f(v_2) = y_2$. Then

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = \lambda_1 y_1 + \lambda_2 y_2 \in Y$$

and thus $\lambda_1 v_1 + \lambda_2 v_2 \in f^{\leftarrow}(Y)$, which proves that $f^{\leftarrow}(Y)$ is a subspace. □

Definition 2.0.43. The subspace $f^{\rightarrow}(V)$ of W is called the image of f and is denoted by $\operatorname{Im} f$.

Definition 2.0.44. The supspace $f^{\leftarrow}\{0_W\}$ is called the kernel of f and is denoted by $\operatorname{Ker} f$.

Example 2.0.45. For the map in example 2.0.35, the image is

$$\operatorname{Im} f = \{(x+y, x-y, y) \,:\, x, y \in \mathbb{R}\} = \operatorname{Span}\{(1,1,0), (1,-1,1)\}.$$

and the Kernel is

$$\operatorname{Ker} f = \{(0,0)\}.$$

Example 2.0.46. For the i-th projection map pr_i(example 2.0.36), we have

$$\operatorname{Im} \mathrm{pr}_i = \mathbb{R}, \quad \operatorname{Ker} \mathrm{pr}_i = \{x \in \mathbb{R}^n \,:\, x_i = 0\}.$$

Example 2.0.47. For the differentiation map D on $\mathbb{R}_{n+1}[X]$ (example 2.0.37) we have

$$\operatorname{Im} D = \mathbb{R}_n[X], \quad \operatorname{Ker} D = \mathbb{R}.$$

Theorem 2.0.48. *If $f : V \to W$ is linear then f is injective iff $Ker\, f = \{0_V\}$.*

Proof. (only if) Let f be injective. Since $f(0_V) = 0_W$ for all linear maps, $0_V \in \operatorname{Ker} f$. On the other hand, if $x \in \operatorname{Ker} f$, then $f(x) = 0_W = f(0_V)$. Since f is injective, we have $x = 0_V$.

(if) Let $\operatorname{Ker} f = \{0_V\}$. Then if $f(x) = f(y)$ we have $f(x-y) = f(x) - f(y) = 0_W$ so that $x - y \in \operatorname{Ker} f \implies x - y = 0_V \implies x = y$. Thus f is injective. □

Theorem 2.0.49. *If $f : V \to W$ is linear and injective and if $\{v_1, \ldots, v_n\}$ is a linearly independent subset of V, then $\{f(v_1), \ldots, f(v_n)\}$ is linearly independent.*

Proof. Let $\lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) = 0_W$. Then $f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = 0_W$ and hence $\lambda_1 v_1 + \ldots + \lambda_n v_n \in \operatorname{Ker} f$. Since f is injective, $\operatorname{Ker} f = \{0_V\}$ so that $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0_V$ and since $\{v_1, v_2, \ldots, v_n\}$ is linearly independent, we have $\lambda_1 = \ldots = \lambda_n = 0$. Thus $\{f(v_1), \ldots, f(v_n)\}$ is linearly independent. □

Theorem 2.0.50. *Let V and W be finite dimensional vector spaces over the same field F. If $f : V \to W$ is linear we have*

$$\dim V = \dim Im\, f + \dim Ker\, f$$

Proof. Let $\{v_1, \ldots, v_n\}$ be a basis of $\operatorname{Ker} f \subseteq V$ and $\{w_1, \ldots, w_m\}$ be a basis of $\operatorname{Im} f \subseteq W$. Since $w_i \in \operatorname{Im} f$, $\exists v_i^* \in V$ such that $f(v_i^*) = w_i$. Consider the set $J = \{v_1^*, \ldots, v_m^*, v_1, \ldots, v_n\} \subseteq V$.

Let $x \in V$. Since $f(x) \in \operatorname{Im} f$, we can write

$$f(x) = \sum_{i=1}^{m} \lambda_i w_i = \sum_{i=1}^{m} \lambda_i f(v_i^*) = f\left(\sum_{i=1}^{m} \lambda_i v_i^*\right)$$

so that $f(x - \sum_{i=1}^{m} \lambda_i v_i^*) = 0_W$. Thus $x - \sum_{i=1}^{m} \lambda_i v_i^* \in \operatorname{Ker} f$ so that

$$x - \sum_{i=1}^{m} \lambda_i v_i^* = \sum_{j=1}^{n} \mu_j v_j \implies x = \sum_{i=1}^{m} \lambda_i v_i^* + \sum_{j=1}^{n} \mu_j v_j.$$

Thus J is a spanning set.

Again, if $\sum_{i=1}^{m}\lambda_i v_i^* + \sum_{j=1}^{n}\mu_j v_j = 0_V$, we have $f\left(\sum_{i=1}^{m}\lambda_i v_i^* + \sum_{j=1}^{n}\mu_j v_j\right) = 0_W$ so that $\sum_{i=1}^{m}\lambda_i w_i = 0_W$. Since $\{w_1, \ldots, w_m\}$ is linearly independent we have $\lambda_i = 0$. But then $\sum_{j=1}^{n}\mu_j v_j = 0$ so that $\mu_j = 0$. Thus J is linearly independent and is hence a basis.

Thus

$$\dim V = |J| = m + n = \dim \operatorname{Im} f + \dim \operatorname{Ker} f$$

□

DEFINITION 2.0.51. If f is a linear map $\dim \operatorname{Im} f$ is called the rank of f and $\dim \operatorname{Ker} f$ is called the nullity of f. Thus theorem 2.0.50 can be written as

$$rank + nullity = dimension\ of\ departure\ space.$$

THEOREM 2.0.52. *Let V and W be vector spaces of dimension n over the same field F. If $f : V \to W$ is linear then the following statements are equivalent*

(1) f *is injective.*

(2) f *is surjective.*

(3) f *is bijective.*

(4) f *carries bases to bases,* i.e. *if $\{v_1, \ldots, v_n\}$ is a basis for V then $\{f(v_1), \ldots, f(v_n)\}$ is a basis of W.*

PROOF. (1) $\implies$ (2) $\implies$ (3) :

Let $f : V \to W$ be linear and injective. Then $\operatorname{Ker} f = \{0_V\}$ so that $\dim \operatorname{Ker} f = 0$. According to theorem 2.0.50, we have $\dim \operatorname{Im} f = n$ and thus $\operatorname{Im} f = W$, so that f is surjective. Thus f is both injective and surjective, so that it is bijective.

(2) $\implies$ (1) $\implies$ (3) :

Let $f : V \to W$ be linear and surjective. Then $\operatorname{Im} f = W$ so that $\dim \operatorname{Im} f = \dim W = n$ and thus $\dim \operatorname{Ker} f = 0$ from theorem 2.0.50. Thus $\operatorname{Ker} f = \{0_V\}$ and so f is injective and hence bijective.

That (3) $\implies$ (1) and (3) $\implies$ (2) is obvious.

(1) $\implies$ (4) :

Let $f : V \to W$ be linear and injective and let $\{v_1, \ldots, v_n\}$ be a basis for V. Then $\{f(v_1), \ldots, f(v_n)\}$ are n distinct (because f is injective) liearly independent vectors (from theorem 2.0.49) in W. Thus this provides a basis of W (corollary 1.0.30).

(4) $\implies$ (1) :

Let $f : V \to W$ be linear and carry bases of V to bases of W, *i.e.* if $\{v_1, \ldots, v_n\}$ is a basis for V then $\{f(v_1), \ldots, f(v_n)\}$is a basis of W. Let $x \in \operatorname{Ker} f$ and $x = \sum_{i=1}^{n}\lambda_i v_i$. Then $0_V = f(x) = \sum_{i=1}^{n}\lambda_{1i} f(v_i)$ and thus $\lambda_i = 0$ for $i = 1, 2, \ldots, n$ so that $x = 0_V$. So $\operatorname{Ker} f = \{0_V\}$ and so, according to theorem 2.0.48, f is injective. □

DEFINITION 2.0.53. Let V and W be linear vector spaces over the same field F. A linear map $f : V \to W$ is called a linear isomorphism if it is a bijection. Two vector spaces V and W are called isomorphic, denoted by $V \cong W$ if there is a isomorphism from one to the other.

REMARK 2.0.54. The relation "is isomorphic to" between vector spaces is an equivalence relation.

THEOREM 2.0.55. *Let V be a vector space of dimension $n \geq 1$ over a field F. Then $V \cong F^n$.*

PROOF. Let $\{v_1, v_2, \ldots v_n\}$ be a basis of V. Consider the map $f : V \to F^n$ defined by

$$f\left(\sum_i \lambda_i v_i\right) = (\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Since $\{v_1, v_2, \ldots v_n\}$ is a basis, every vector $x \in V$ has a unique expansion of the form $\sum_i \lambda_i x_i$. Thus the map f is an isomorphism. It is also obviously linear. □

THEOREM 2.0.56. *If V and W are vector spaces of dimension n over the same field F, then $V \cong W$.*

PROOF. By the previous theorem, we have $V \cong F^n$ and $W \cong F^n$. Then $\exists$ linear isomorphisms $f : V \to F^n$ and $g : W \to F^n$. Then, the map $g^{-1} \circ f : V \to W$ is a linear bijection, and hence an isomorphism. □

For a general map f between two spaces V and W, we have to specify what f does to each element of V. We now show that for a linear map, we have to only worry about what the map does to each element of a basis of V.

THEOREM 2.0.57. *Let V and W be finite dimensional vector spaces over the same field F. If $\{v_1, v_2, \ldots, v_n\}$ be a basis of V and let $w_1, w_2, \ldots, w_n$ be vectors (not necessarily distinct) in W then there is a unique linear map $f : V \to W$ such that $f(v_i) = w_i$ for $i = 1, 2, \ldots, n$.*

PROOF. Since every element of V can be uniquely expressed in the form $\sum_i \lambda_i v_i$, we can define a map $f : V \to W$ by

$$f\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i w_i$$

It is easy to check that this map is linear. Moreover, for each i, we have

$$f(v_i) = f\left(\sum_{j=1}^{n} \delta_{ij} v_j\right) = \sum_{j=1}^{n} \delta_{ij} w_j = w_i.$$

To show that this map f is unque, let $g : V \to W$ be a linear map with the property that $g(v_i) = w_i$. Then for $\forall v \in V$, $v = \sum_i \lambda_i v_i$, we have

$$g(v) = g\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i g(v_i) = \sum_{i=1}^{n} \lambda_i w_i = f(v).$$

□

COROLLARY 2.0.58. *A linear map is completely and uniquely determined by its action on the basis of its departure space.*

COROLLARY 2.0.59. *Two linear mappings $f, g : V \to W$ are equal iff they agree on a basis of V.*

CHAPTER 3

Linear maps and matrices

DEFINITION 3.0.60. Let V be a finite dimensional vector space. An ordered basis of V is a finite sequence $(v_i)_{1\leq i\leq n}$ of elements of V such that the set $\{v_1, v_2, \ldots, v_n\}$ is a basis of V.

We have seen that a linear map f between two spaces V and W is uniquely determined by its action on a basis of V. Given ordered bases $(v_i)_m$ and $(w_j)_n$ on V and W, respectively we define the matrix form of f with respect to the bases $(v_i)_m$ and $(w_j)_n$, denoted by Mat f, as a matrix with the elements a_{ij} given by

$$f(v_i) = \sum_{j=1}^{n} a_{ji} w_j, \qquad i = 1, 2, \ldots, m$$

Note the (slightly counterintuitive) order of the indices on a. Also note that if V and W are m and n dimensional, respectively, then for a linear map $f : V \to W$ we have $\mathrm{Mat} f \in \mathrm{Mat}_{n\times m}(F)$.

For the special map $\mathrm{id}_V : V \to V$ which maps every $v \in V$ to itself, it is easy to see that $\mathrm{Matid}_V = I_n$ if V is a n dimensional vector space and any given ordered basis $(v_i)_n$ of V is used.

THEOREM 3.0.61. *Let V and W be linear vector spaces over the same field F. Then for all linear maps $f, g : V \to W$ and $\lambda \in F$, and with respect to fixed ordered bases*

$$\begin{aligned} Mat(f+g) &= Matf + Matg \\ Mat(\lambda f) &= \lambda Matf \end{aligned}$$

PROOF. Let $\mathrm{Mat} f = [x_{ij}]_{n\times m}$ and $\mathrm{Mat} g = [y_{ij}]_{n\times m}$ relative to fixed ordered bases $(v_i)_m$ of V and $(w_j)_n$ of W. Then for $i = 1, 2, \ldots, m$

$$f(v_i) = \sum_{j=1}^{n} x_{ji} w_j, \qquad g(v_i) = \sum_{j=1}^{n} y_{ji} w_j$$

so that

$$(f+g)(v_i) = f(v_i) + g(v_i) = \sum_{j=1}^{n} (x_{ji} + y_{ji}) w_j,$$

and thus $\mathrm{Mat}(f+g) = [x_{ij} + y_{ij}]_{n\times m} = \mathrm{Mat} f + \mathrm{Mat} g$. Similarly

$$(\lambda f)(v_i) = \lambda f(v_i) = \lambda \sum_{j=1}^{n} x_{ji} w_j,$$

so that $\mathrm{Mat}(\lambda f) = [\lambda x_{ij}]_{n\times m} = \lambda \mathrm{Mat} f$. □

This theorem implies that the map $\vartheta : \mathrm{Lin}(V, W) \to \mathrm{Mat}_{n\times m}(F)$ defined by $f \mapsto \mathrm{Mat} f$ is a linear map. It is also clearly a bijection (why?). Thus we have

THEOREM 3.0.62. *If V, W are vector spaces of dimension m, n over F then*

$$Lin(V, W) \cong Mat_{n\times m}(F).$$

THEOREM 3.0.63. *Consider linear maps $f : U \to V$ and $g : V \to W$ with fixed ordered bases $(u_i)_m$, $(v_j)_n$ and $(w_k)_p$ of U, V and W respectively. We have*

$$Mat(g \circ f) = Matg\, Matf$$

Proof. Let $\mathrm{Mat} f = [x_{ij}]_{n\times m}$ and $\mathrm{Mat} g = [y_{ij}]_{p\times n}$. Then

$$\begin{aligned} g\circ f(u_i) &= g(f(u_i)) = g\left(\sum_{j=1}^{n} x_{ji}v_j\right) \\ &= \sum_{j=1}^{n} x_{ji} g(v_j) = \sum_{j=1}^{n} x_{ji}\left(\sum_{k=1}^{p} y_{kj}w_k\right) \\ &= \sum_{k=1}^{p}\left(\sum_{j=1}^{n} y_{kj}x_{ji}\right) w_k \end{aligned}$$

Thus $\mathrm{Mat}(g\circ f) = [z_{ij}]_{p\times m}$ where

$$z_{ki} = \sum_{j=1}^{n} y_{kj}x_{ji}.$$

□

Corollary 3.0.64. *A square matrix is invertible iff it represents an isomorphism.*

Proof. Let A be an invertible $n\times n$ square matrix. Thus, there exists a $n\times n$ square matrix B such that $BA = I_n$. Let V be a n dimensional vector space and let $(v_i)_n$ be an ordered basis of V. Let $f, g : V \to V$ be linear maps represented by the matrices A, B with respect to the ordered basis $(v_i)_n$. Then $g\circ f$ is represented by the matrix $BA = I_n$ and thus $g\circ f = \mathrm{id}_V$ and so f is invertible and is thus an isomorphism.

Conversely, let the matrix A represent the linear isomorphism $f : V \to V$. Since f is a bijection, it has an inverse f^{-1} such that $f^{-1}\circ f = \mathrm{id}_V$. Then the matrix B representing f^{-1} obeys $BA = \mathrm{Mat} f^{-1}\mathrm{Mat} f = \mathrm{Mat}\left(f^{-1}\circ f\right) = \mathrm{Mat}\,\mathrm{id}_V = I_n$. Thus $B = A^{-1}$ and the matrix A is invertible. □

We next consider the question of change of basis and its effect on the matrix representative of a linear map.

Definition 3.0.65. Given two ordered bases $(v_i)_n$ and $(v_i')_n$ of a given vector space V the transition matrix from $(v_i)_n$ to $(v_i')_n$ is defined as the matrix representative of the identity map id_V with respect to the bases $(v_i)_n$ and $(v_i')_n$.

Remark 3.0.66. If $V = \mathbb{R}^n$ and $(v_i)_n$ is the natural ordered basis on V then it is easier to determine the transition matrix from $(v_i')_n$ to $(v_i)_n$, than the other way around.

From corollary 3.0.64 it is obvious that

Theorem 3.0.67. *Transition matrices are invertible.*

Theorem 3.0.68. *If a linear mapping $f : V \to W$ is represented relative to ordered bases $(v_i)_m, (w_i)_n$ by the $n\times m$ matrix A then relative to new ordered bases $(v_i')_m, (w_i')_n$ the matrix representing f is the $n\times m$ matrix $Q^{-1}AP$ where Q is the transition matrix from $(w_i')_n$ to $(w_i)_n$ and P is the transition matrix from $(v_i')_m$ to $(v_i)_m$.*

Proof. From the following diagram

$$\begin{array}{ccc} V;\ (v_i)_m & \xrightarrow{\ f;\ A\ } & W;\ (w_i)_n \\ {\scriptstyle \mathrm{id}_V;\ P}\uparrow & & \uparrow{\scriptstyle \mathrm{id}_W;\ Q} \\ V;\ (v_i')_m & \xrightarrow[\ f;\ B\ ]{} & W;\ (w_i')_n \end{array}$$

using

$$f\circ \mathrm{id}_V = \mathrm{id}_W \circ f = f$$

we get

$$AP = QB$$

so that $B = Q^{-1}AP$. □

The converse of this theorem is also true.

THEOREM 3.0.69. *Let $(v_i)_m, (w_i)_n$ be ordered bases of V, W, respectively. Suppose that A, B are $n \times m$ matrices such that there are invertible matrices P, Q such that $B = Q^{-1}AP$. Then there are ordered bases $(v_i')_m, (w_i')_n$ of V, W and a linear map $f : V \to W$ such that A is the matrix of f relative to $(v_i)_m, (w_i)_n$ and B is the matrix of f relative to $(v_i')_m, (w_i')_n$.*

PROOF. If $P = [p_{ij}]_{m\times m}$ and $Q = [q_{ij}]_{n\times n}$, define

$$(i = 1, \ldots, m) \quad v_i' = \sum_{j=1}^{m} p_{ji}v_j; \qquad (i = 1, \ldots, n) \quad w_i' = \sum_{j=1}^{n} q_{ji}w_j.$$

Since P is invertible, there is (corollary 3.0.64) an isomorphism $f_P : V \to V$ that is represented by P relative to the ordered basis $(v_i)_m$. Since, by definition $v_i' = f_P(v_i)$, it follows that $(v_i')_m$ is an ordered basis of V (by theorem 2.0.52). P is the transition matrix from $(v_i')_m$ to $(v_i)_m$. Similarly $(w_i')_n$ is an ordered basis of and Q is the transition matrix from $(w_i')_n$ to $(w_i)_n$.

Now let $f : V \to$W be the linear map whose matrix, relative to the bases $(v_i)_m$ and $(w_i)_n$, is A. Then by theorem 3.0.68 the matrix of f relative to the bases $(v_i')_m$ and $(w_i')_n$ is $Q^{-1}AP = B$. □

DEFINITION 3.0.70. If A, B are $n \times n$ matrices then $B$is called similar to A if there is an invertible matrix P such that $B = P^{-1}AP$.

The next two theorems are very simple to prove.

THEOREM 3.0.71. *For $n \times n$ matrices, the relation "is similar to" is an equivalence relation.*

THEOREM 3.0.72. *Two $n \times n$ matrices A, B are similar iff they represent the same linear map relative to possibly different bases.*