Some notes on linear algebra

Throughout these notes, k denotes a field (often called the *scalars* in this context). Recall that this means that there are two binary operations on k, denoted $+$ and $\cdot$, that $(k, +)$ is an abelian group, $\cdot$ is commutative and associative and distributes over addition, there exists a multiplicative identity 1, and, for all $t \in k$, $t \neq 0$, there exists a multiplicative inverse for t, denoted t^{-1}. The main example in this course will be $k = \mathbb{C}$, but we shall also consider the cases $k = \mathbb{R}$ and $k = \mathbb{Q}$. Another important case is $k = \mathbb{F}_q$, a finite field with q elements, where $q = p^n$ is necessarily a prime power (and p is a prime number). For example, for $q = p$, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. If we drop the requirement that multiplication be commutative, then such a k is called a *division algebra* or *skew field*. It is hard to find noncommutative examples of division rings, though. For example, by a theorem of Wedderburn, a finite division ring is a field. One famous example is the *quaternions*

$$\mathbb{H} = \mathbb{R} + \mathbb{R} \cdot i + \mathbb{R} \cdot j + \mathbb{R} \cdot k.$$

A typical element of $\mathbb{H}$ is then of the form $a_0 + a_1 i + a_2 j + a_3 k$. Here $i^2 = j^2 = k^2 = -1$, and

$$ij = k = -ji; \qquad jk = i = -kj; \qquad ki = j = -ik.$$

These rules then force the rules for multiplication for $\mathbb{H}$, and a somewhat lengthy check shows that $\mathbb{H}$ is a (noncommutative) division algebra.

We shall occasionally relax the condition that every nonzero element has a multiplicative inverse. Thus a *ring* R is a set with two binary operations $+$ and $\cdot$, that $(R, +)$ is an abelian group, $\cdot$ is associative and distributes over addition, and there exists a multiplicative identity 1. If $\cdot$ is commutative as well, we call R a *commutative ring*. A standard example of a noncommutative ring is the ring $\mathbb{M}_n(\mathbb{R})$ of $n \times n$ matrices with coefficients in $\mathbb{R}$ with the usual operations of matrix addition and multiplication, for $n > 1$, and with multiplicative identity the $n \times n$ identity matrix I. (For $n = 1$, it is still a ring but in fact it is isomorphic to $\mathbb{R}$ and hence is commutative.) More generally, if k is any field, for example $k = \mathbb{C}$ and $n > 1$, then the set $\mathbb{M}_n(k)$ of $n \times n$ matrices with coefficients in k with the usual operations of matrix addition and multiplication is a noncommutative ring.

1 Definition of a vector space

Definition 1.1. A *k-vector space* or simply a *vector space* is a triple $(V, +, \cdot)$, where $(V, +)$ is an abelian group (the *vectors*), and there is a function $k \times V \to V$ (*scalar multiplication*), whose value at (t, v) is just denoted $t \cdot v$ or tv, such that

1. For all $s, t \in k$ and $v \in V$, $s(tv) = (st)v$.
2. For all $s, t \in k$ and $v \in V$, $(s + t)v = sv + tv$.
3. For all $t \in k$ and $v, w \in V$, $t(v + w) = tv + tw$.
4. For all $v \in V$, $1 \cdot v = v$.

A *vector subspace* W of a k-vector space V is a subgroup W of $(V, +)$ such that, for all $t \in k$ and $w \in W$, $tw \in W$, i.e. W is a subgroup closed under scalar multiplication. It is then straightforward to check that W is again a k-vector space. For example, $\{0\}$ and V are always vector subspaces of V.

It is a straightforward consequence of the axioms that $0v = 0$ for all $v \in V$, where the first 0 is the element $0 \in F$ and the second is $0 \in V$, that, for all $t \in F$, $t0 = 0$ (both 0's here are the zero vector), and that, for all $v \in V$, $(-1)v = -v$. Hence, W is a vector subspace of $V \iff W$ is nonempty and is closed under addition and scalar multiplication.

Example 1.2. (0) The set $\{0\}$ is a k-vector space.

(1) The n-fold Cartesian product k^n is a k-vector space, where k^n is a group under componentwise addition and scalar multiplication, i.e.

$$\begin{aligned}(a_1, \dots, a_n) + (b_1, \dots, b_n) &= (a_1 + b_1, \dots, a_n + b_n)\\ t(a_1, \dots, a_n) &= (ta_1, \dots, ta_n).\end{aligned}$$

(2) If X is a set and k^X denotes the group of functions from X to k, then k^X is a k-vector space. Here addition is defined pointwise: $(f + g)(x) = f(x) + g(x)$, and similarly for scalar multiplication: $(tf)(x) = tf(x)$.

(3) With X a set as in (2), we consider the subset $k[X] \subseteq k^X$ of functions $f\colon X \to k$ such that the set $\{x \in X : f(x) \neq 0\}$ is a finite subset of X, then $k[X]$ is easily seen to be a vector space under pointwise addition and scalar multiplication of functions, so that $k[X]$ is a vector subspace of k^X. Of course, $k[X] = k^X$ if and only if X is finite. In general, for a function $f\colon X \to k$, we define the *support* of f or $\operatorname{Supp} f$ by:

$$\operatorname{Supp} f = \{x \in X : f(x) \neq 0\}.$$

Then $k[X]$ is the subset of k^X consisting of all functions whose support is finite.

Given $x \in X$, let $\delta_x \in k[X]$ be the function defined by

$$\delta_x(y) = \begin{cases} 1, & \text{if } y = x; \\ 0, & \text{otherwise.} \end{cases}$$

The function δ_x is often called the *delta function* at x. Then clearly, for every $f \in k[X]$,

$$f(x) = \sum_{x \in X} f(x)\delta_x,$$

where in fact the above sum is finite (and therefore meaningful). We will often identify the function δ_x with the element $x \in X$. In this case, we can view X as a subset of $k[X]$, and every element of $k[X]$ can be uniquely written as $\sum_{x \in X} t_x \cdot x$, where the $t_x \in k$ and $t_x = 0$ for all but finitely many x. The case $X = \{1, \dots, n\}$ corresponds to k^n, and the element δ_i to the vector $e_i = (0, \dots, 1, \dots, 0)$ (the i^{th} component is 1 and all other components are 0).

For a general ring R, the analogue of a vector space is an R-module, with some care if R is not commutative:

Definition 1.3. A *left R-module* M is is a triple $(M, +, \cdot)$, where $(M, +)$ is an abelian group and there is a function $R \times M \to M$, whose value at (r, m) is just denoted $r \cdot m$ or rm, such that

1. For all $r, s \in R$ and $m \in M$, $r(sm) = (rs)m$.
2. For all $r, s \in R$ and $m \in M$, $(r + s)m = rm + sm$.
3. For all $r \in R$ and $m, n \in M$, $r(m + n) = rm + rn$.
4. For all $m \in M$, $1 \cdot m = m$.

Left submodules of an R-module M are defined in the obvious way. A *right R-module* is an abelian group M with a multiplication by R on the *right*, i.e. a function $R \times M \to M$ whose value at (r, m) is denoted mr, satisfying $(ms)r = m(sr)$ and the remaining analogues of (2), (3), (4) above. Thus, if R is commutative, there is no difference between left and right R-modules.

Example 1.4. (1) The n-fold Cartesian product R^n is a left R-module, where R^n is a group under componentwise addition and scalar multiplication, i.e.

$$\begin{aligned}(a_1, \dots, a_n) + (b_1, \dots, b_n) &= (a_1 + b_1, \dots, a_n + b_n)\\ r(a_1, \dots, a_n) &= (ra_1, \dots, ra_n).\end{aligned}$$

Of course, it can also be made into a right R-module.

(2) If X is a set and R^X denotes the group of functions from X to R, then R^X is a left R-module using pointwise addition and scalar multiplication. Similarly the subset $R[X]$, defined in the obvious way, is a submodule.

(3) Once we replace fields by rings, there are many more possibilities for R-modules. For example, if I is an ideal in R (or a left ideal if R is not commutative), then I is an R-submodule of R and the quotient ring R/I (in the commutative case, say) is also an R-module. For example, $\mathbb{Z}/n\mathbb{Z}$ is a $\mathbb{Z}$-module, but it looks very different from a $\mathbb{Z}$-module of the form $\mathbb{Z}^n$. For another example, which will be more relevant to us, if $R = \mathbb{M}_n(k)$ is the (noncommutative) ring of $n \times n$ matrices with coefficients in the field k, then k^n is a left R-module, by defining $A \cdot v$ to be the usual multiplication of the matrix A with the vector v (viewed as a column vector).

2 Linear maps

Linear maps are the analogue of homomorphisms and isomorphisms:

Definition 2.1. Let V_1 and V_2 be two k-vector spaces and let $F: V_1 \to V_2$ be a function (= map). Then F is *linear* or a *linear map* if it is a group homomorphism from $(V_1, +)$ to $(V_2, +)$, i.e. is additive, and satisfies: For all $t \in k$ and $v \in V_1$, $F(tv) = tF(v)$. The composition of two linear maps is again linear. For example, for all V_1 and V_2, the constant function $F = 0$ is linear. If $V_1 = V_2 = V$, then the identity function $\mathrm{Id}: V \to V$ is linear.

The function F is a *linear isomorphism* or briefly an *isomorphism* if it is both linear and a bijection; in this case, it is easy to check that F^{-1} is also linear. We say that V_1 and V_2 are *isomorphic*, written $V_1 \cong V_2$, if there exists an isomorphism F from V_1 to V_2.

If $F\colon V_1 \to V_2$ is a linear map, then we define

$$\begin{aligned} \operatorname{Ker} F &= \{v \in V_1 : F(v) = 0\}; \\ \operatorname{Im} F &= \{w \in V_2 : \text{ there exists } v \in V_1 \text{ such that } F(v) = w. \end{aligned}$$

Lemma 2.2. *Let $F\colon V_1 \to V_2$ be a linear map. Then $\operatorname{Ker} F$ is a subspace of V_1 and $\operatorname{Im} F$ is a subspace of V_2. Moreover, F is injective $\iff$ $\operatorname{Ker} F = \{0\}$ and F is surjective $\iff$ $\operatorname{Im} F = V_2$.* □

To find examples of linear maps, we use the following:

Lemma 2.3. *Let W be a vector space and let $w_1, \dots, w_n \in W$. There is a unique linear map $F\colon k^n \to W$ defined by $F(t_1, \dots, t_n) = \sum_{i=1}^n t_i w_i$. It satisfies: For every i, $F(e_i) = w_i$.*

This lemma may be generalized as follows:

Lemma 2.4. *Let X be a set and let V be a vector space. For each function $f\colon X \to V$, there is a unique linear map $F\colon k[X] \to V$ defined by $F(\sum_{x \in X} t_x \cdot x) = \sum_{x \in X} t_x f(x)$ (a finite sum). In particular, F is specified by the requirement that $F(x) = f(x)$ for all $x \in X$. Finally, every linear function $k[X] \to V$ is of this form.* □

For this reason, $k[X]$ is sometimes called the *free vector space* on the set X.

3 Linear independence and span

Let us introduce some terminology:

Definition 3.1. Let V be a k-vector space and let $v_1, \dots, v_d \in V$ be a sequence of vectors. A *linear combination* of $v_1, \dots, v_d$ is a vector of the form $t_1 v_1 + \cdots + t_d v_d$, where the $t_i \in k$. The *span* of $\{v_1, \dots, v_d\}$ is the set of all linear combinations of $v_1, \dots, v_d$. Thus

$$\operatorname{span}\{v_1, \dots, v_d\} = \{t_1 v_1 + \cdots + t_d v_d : t_i \in k \text{ for all } i\}.$$

By definition (or logic), $\operatorname{span} \emptyset = \{0\}$.

With V and $v_1, \dots, v_d$ as above, we have the following properties of span:

(i) $\operatorname{span}\{v_1, \dots, v_d\}$ is a vector subspace of V containing v_i for every i. In fact, it is the image of the linear map $F\colon k^d \to V$ defined by $F(t_1, \dots, t_d) = \sum_{i=1}^d t_i v_i$ above.

(ii) If W is a vector subspace of V containing $v_1, \dots, v_d$, then

$$\operatorname{span}\{v_1, \dots, v_d\} \subseteq W.$$

In other words, $\operatorname{span}\{v_1, \dots, v_d\}$ is the smallest vector subspace of V containing $v_1, \dots, v_d$.

(iii) For every $v \in V$, $\operatorname{span}\{v_1, \dots, v_d\} \subseteq \operatorname{span}\{v_1, \dots, v_d, v\}$, and equality holds if and only if $v \in \operatorname{span}\{v_1, \dots, v_d\}$.

Definition 3.2. A sequence of vectors $v_1, \dots, v_d$ such that

$$V = \operatorname{span}\{v_1, \dots, v_d\}$$

will be said to *span* V. Thus, $v_1, \dots, v_d$ span $V \iff$ the linear map $F\colon k^d \to V$ defined by $F(t_1, \dots, t_d) = \sum_{i=1}^d t_i v_i$ is surjective.

Definition 3.3. A k-vector space V is a *finite dimensional vector space* if there exist $v_1, \dots, v_d \in V$ such that $V = \operatorname{span}\{v_1, \dots, v_d\}$. The vector space V is *infinite dimensional* if it is not finite dimensional.

For example, k^n is finite-dimensional. But, if X is an infinite set, then $k[X]$ and hence k^X are not finite-dimensional vector spaces.

Definition 3.4. A sequence $w_1, \dots, w_r \in V$ is *linearly independent* if the following holds: if there exist $t_i \in k$ such that

$$t_1 w_1 + \cdots + t_r w_r = 0,$$

then $t_i = 0$ for all i. The sequence $w_1, \dots, w_r$ is *linearly dependent* if it is not linearly independent. Note that $w_1, \dots, w_r \in V$ are linearly independent $\iff$ $\operatorname{Ker} F = \{0\}$, where $F\colon k^n \to V$ is the linear map defined by $F(t_1, \dots, t_r) = \sum_{i=1}^r t_i w_i \iff$ the map F defined above is injective. It then follows that the vectors $w_1, \dots, w_r$ are linearly independent if and only if, given $t_i, s_i \in k$, $1 \le i \le r$, such that

$$t_1 w_1 + \cdots + t_r w_r = s_1 w_1 + \cdots + s_r w_r,$$

then $t_i = s_i$ for all i.

Note that the definition of linear independence does **not** depend **only** on the set $\{w_1, \dots, w_r\}$—if there are any repeated vectors $w_i = w_j$, then we can express 0 as the nontrivial linear combination $w_i - w_j$. Likewise if one of the w_i is zero then the set is linearly dependent.

Definition 3.5. Let V be a k-vector space. The vectors $v_1, \dots, v_d$ are a *basis* of V if they are linearly independent and $V = \text{span}\{v_1, \dots, v_d\}$. Equivalently, the vectors $v_1, \dots, v_d$ are a basis of $V \iff$ the linear map $F\colon k^d \to V$ defined by $F(t_1, \dots, t_d) = \sum_{i=1}^d t_i v_i$ is an isomorphism.

Remark 3.6. There are analogous notions of linear independence, span, and basis for an infinite dimensional vector space V: If $\{v_i : i \in I\}$ is a collection of vectors in V indexed by a set I, then $\{v_i : i \in I\}$ is *linearly independent* if the following holds: if there exist $t_i \in k$ such that $t_i = 0$ for all but finitely many $i \in I$ and $\sum_{i \in I} t_i v_i = 0$, then $t_i = 0$ for all i. The *span* of $\{v_i : i \in I\}$ is the subset

$$\left\{ \sum_{i \in I} t_i v_i : t_i \in k,\ t_i = 0 \text{ for all but finitely many } i \in I \right\},$$

and $\{v_i : i \in I\}$ *spans* V if its span is equal to V. Finally, $\{v_i : i \in I\}$ is a *basis* of V if it is linearly independent and spans V, or equivalently if every $v \in V$ can be uniquely expressed as $\sum_{i \in I} t_i v_i$, where $t_i \in k$ and $t_i = 0$ for all but finitely many $i \in I$.

With this definition, the set X (or equivalently the set $\{\delta_x : x \in X\}$) is a basis for $k[X]$.

We summarize the salient facts about span, linear independence, and bases as follows:

1. Suppose that $V = \text{span}\{v_1, \dots, v_n\}$ and that $w_1, \dots, w_\ell$ are linearly independent vectors in V. Then $\ell \le n$.

2. If V is a finite dimensional k-vector space, then every two bases for V have the same number of elements—call this number the *dimension* of V which we write as $\dim V$ or $\dim_k V$ if we want to emphasize the field k. Thus for example $\dim k^n = n$, and, more generally, if X is a finite set, then $\dim k[X] = \#(X)$.

3. If $V = \text{span}\{v_1, \dots, v_d\}$ then some subsequence of $v_1, \dots, v_d$ is a basis for V. Hence $\dim V \le d$, and if $\dim V = d$ then $v_1, \dots, v_d$ is a basis for V.

4. If V is a finite-dimensional vector space and $w_1, \dots, w_\ell$ are linearly independent vectors in V, then there exist vectors

$$w_{\ell+1}, \dots, w_r \in V$$

such that $w_1, \dots, w_\ell, w_{\ell+1}, \dots, w_r$ is a basis for V. Hence $\dim V \geq \ell$, and if $\dim V = \ell$ then $w_1, \dots, w_\ell$ is a basis for V.

5. If V is a finite-dimensional vector space and W is a vector subspace of V, then $\dim W \leq \dim V$. Moreover $\dim W = \dim V$ if and only if $W = V$.

6. If V is a finite-dimensional vector space and $F\colon V \to V$ is linear, then F is injective $\iff$ F is surjective $\iff$ F is an isomorphism.

Remark 3.7. If V is a finite dimensional k-vector space and $W \subseteq V$, then W is a vector subspace of V if and only if it is of the form $\operatorname{span}\{v_1, \dots, v_d\}$ for some $v_1, \dots, v_d \in V$. This is no longer true if V is not finite dimensional.

4 The group algebra

We apply the construction of $k[X]$ to the special case where $X = G$ is a group (written multiplicatively). In this case, we can view $k[G]$ as functions $f\colon G \to k$ with finite support or as formal sums $\sum_{g\in G} t_g \cdot g$, where $t_g = 0$ for all but finitely many g. It is natural to try to extend the multiplication in G to a multiplication on $k[G]$, by defining the product of the formal symbols g and h to be gh (using the product in G and then expanding by using the obvious choice of rules. Explicitly, we define

$$\left(\sum_{g\in G} s_g \cdot g\right)\left(\sum_{g\in G} t_g \cdot g\right) = \sum_{g\in G}\left(\sum_{\substack{h_1,h_2\in G\\ h_1h_2=g}} s_{h_1}t_{h_2}\right)\cdot g.$$

It is straightforward to check that, with our finiteness assumptions, the inner sums in the formula above are all finite and the coefficients of the product as defined above are 0 for all but finitely many g.

If we view elements of $k[G]$ instead as functions, then the product of two functions f_1 and f_2 is called the *convolution* $f_1 * f_2$, and is defined as follows:

$$(f_1 * f_2)(g) = \sum_{\substack{h_1,h_2\in G\\ h_1h_2=g}} f_1(h_1)f_2(h_2) = \sum_{h\in G} f_1(h)f_2(h^{-1}g) = \sum_{h\in G} f_1(gh^{-1})f_2(h).$$

Again, it is straightforward to check that the sums above are finite and that $\operatorname{Supp} f_1 * f_2$ is finite as well.

With either of the above descriptions, one can then check that $k[G]$ is a ring. The messiest calculation is associativity of the product. For instance, we can write

$$\begin{aligned} f_1 * (f_2 * f_3)(g) &= \sum_{x \in G} f_1(x)(f_2 * f_3)(x^{-1}g) = \sum_{x \in G}\sum_{y \in G} f_1(x)f_2(y)f_3(y^{-1}x^{-1}g) \\ &= \sum_{\substack{x,y,z \in G \\ xyz = g}} f_1(x)f_2(y)f_3(z). \end{aligned}$$

Similarly,

$$\begin{aligned} (f_1 * f_2) * f_3(g) &= \sum_{z \in G} (f_1 * f_2)(gz^{-1})f_3(z) = \sum_{x \in G}\sum_{z \in G} f_1(x)f_2(x^{-1}gz^{-1})f_3(z) \\ &= \sum_{\substack{x,y,z \in G \\ xyz = g}} f_1(x)f_2(y)f_3(z). \end{aligned}$$

Thus $f_1 * (f_2 * f_3) = (f_1 * f_2) * f_3$. Left and right distributivity can also be checked by a (somewhat easier) calculation. If $1 \in G$ is the identity element, let 1 denote the element of $k[G]$ which is $1 \cdot 1$ (the first 1 is $1 \in k$ and the second is $1 \in G$). Thus we identify 1 with the element $\delta_1 \in k[G]$. Then $1 * f = f * 1 = f$. In fact, we leave it as an exercise to show that, for all $h \in G$, and all $f \in k[G]$,

$$\begin{aligned} (\delta_h * f)(g) &= f(h^{-1}g); \\ (f * \delta_h)(g) &= f(gh^{-1}). \end{aligned}$$

In other words, convolution of the function f with δ_h has the effect of translating the function f, i.e. shifting the variable.

The ring $k[G]$ is commutative $\iff$ G is abelian. It is essentially never a division ring. For example, if $g \in G$ has order 2, so that $g^2 = 1$, then the element $1 + g \in k[G]$ is a zero divisor: let $-g$ denote the element $(-1) \cdot g$, where $-1 \in k$ is the additive inverse of 1. Then

$$(1 + g)(1 - g) = 1 + g - g - g^2 = 1 - 1 = 0.$$

5 Linear functions and matrices

Let $F\colon k^n \to k^m$ be a linear function. Then

$$F(t_1, \ldots, t_n) = F(\sum_i t_i e_i) = \sum_i t_i F(e_i).$$

We can write $F(e_i)$ in terms of the basis $e_1, \dots, e_m$ of k^m: suppose that $F(e_i) = \sum_{j=1}^{m} a_{ji}e_j = (a_{1i}, \dots, a_{mi})$. We can then associate to F an $m \times n$ matrix with coefficients in k as follows: Define

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Then the columns of A are the vectors $F(e_i)$. Thus every linear map $F\colon k^n \to k^m$ corresponds to an $m \times n$ matrix with coefficients in k. We denote by $\mathbb{M}_{m,n}(k)$ the set of all such. Clearly $\mathbb{M}_{m,n}(k)$ is a vector space of dimension mn, with basis

$$\{E_{ij}, 1 \le i \le m, 1 \le j \le n\},$$

where E_{ij} is the $n \times n$ matrix whose $(i,j)^{\text{th}}$ entry is 1 and such that all other entries are 0. Equivalently, $\mathbb{M}_{m,n}(k) \cong k[X]$, where $X = \{1, \dots, m\} \times \{1, \dots, n\}$ is the set of ordered pairs (i,j) with $1 \le i \le m$ and $1 \le j \le n$. If $m = n$, we abbreviate $\mathbb{M}_{m,n}(k)$ by $\mathbb{M}_n(k)$; it is the space of *square matrices* of size n.

Composition of linear maps corresponds to matrix multiplication: If $F\colon k^n \to k^m$ and $G\colon k^m \to k^\ell$ are linear maps corresponding to the matrices A and B respectively, then $G \circ F\colon k^n \to k^\ell$ corresponds to the $\ell \times n$ matrix BA. Here, the (i,j) entry of BA is given by

$$\sum_{k=1}^{m} b_{ik}a_{kj}.$$

Composition of matrices is associative and distributes over addition of matrices; moreover $(tB)A = B(tA) = t(BA)$ if this is defined. The *identity matrix* $I \in \mathbb{M}_n(k)$ is the matrix whose diagonal entries are 1 and such that all other entries are 0; it corresponds to the identity map $\mathrm{Id}\colon k^n \to k^n$. Thus as previously mentioned $\mathbb{M}_n(k)$ is a ring.

A matrix $A \in \mathbb{M}_n(k)$ is *invertible* if there exists a matrix $B \in \mathbb{M}_n(k)$ such that $BA = I$. Then necessarily $AB = I$, and we write $B = A^{-1}$. If $F\colon k^n \to k^n$ is the linear map corresponding to the matrix A, then A is invertible $\iff$ F is an isomorphism $\iff$ F is injective $\iff$ F is surjective. We have the following formula for two invertible matrices $A_1, A_2 \in \mathbb{M}_n(k)$: if A_1, A_2 are invertible, then so is A_1A_2, and

$$(A_1A_2)^{-1} = A_2^{-1}A_1^{-1}.$$

The subset of $\mathbb{M}_n(k)$ consisting of invertible matrices is then a group, denoted by $GL_n(k)$.

Suppose that V_1 and V_2 are two finite dimensional vector spaces, and choose bases $v_1, \dots, v_n$ of V_1 and $w_1, \dots, w_m$ of V_2, so that $n = \dim V_1$ and $m = \dim V_2$. If $F\colon V_1 \to V_2$ is linear, then, given the bases $v_1, \dots, v_n$ and $w_1, \dots, w_m$ we again get an $m \times n$ matrix A by the formula: $A = (a_{ij})$, where a_{ij} is defined by

$$F(v_i) = \sum_{j=1}^{m} a_{ji} w_j.$$

We say that A is the matrix *associated to the linear map f and the bases $v_1, \dots, v_n$ and $w_1, \dots, w_m$*. Given any collection of vectors $u_1, \dots, u_k$ in V_2, there is a unique linear map $F\colon V_1 \to V_2$ defined by $F(v_i) = u_i$ for all i. In fact, if we let $H\colon k^n \to V_1$ be the "change of basis" map defined by $H(t_1, \dots, t_n) = \sum_{i=1}^{n} t_i w_i$, then H is an isomorphism and so has an inverse H^{-1}, which is again linear. Let $G\colon k^n \to V_2$ be the map $G(t_1, \dots, t_n) = \sum_{i=1}^{n} t_i u_i$. Then the map F is equal to $G \circ H^{-1}$, proving the existence and uniqueness of F.

A related question is the following: Suppose that $v_1, \dots, v_n$ is a basis of k^n and that $w_1, \dots, w_m$ is a basis of k^m. Let F be a linear map $k^n \to k^m$. The we have associated a matrix A to F, using the standard bases of k^n and k^m, i.e. $F(e_i) = \sum_{j=1}^{m} a_{ji} e_j$. We can also associate a matrix B to F by using the bases $v_1, \dots, v_n$ and $w_1, \dots, w_m$, i.e. $F(v_i) = \sum_{j=1}^{m} b_{ji} w_j$. What is the relationship between A and B? This question can be answered most simply in terms of linear maps. Let $G\colon k^n \to k^m$ be the linear map corresponding to the matrix B in the usual way: $G(e_i) = \sum_{j=1}^{k} b_{ji} e_j$. As above, let $H_1\colon k^n \to k^n$ be the "change of basis" map in k^n defined by $H_1(e_i) = v_i$, and let $H_2\colon k^m \to k^m$ be the corresponding change of basis map in k^m, defined by $H_2(e_j) = w_j$. Then clearly (by applying both sides to v_i for every i)

$$F = H_2 \circ G \circ H_1^{-1}.$$

Thus, if H_1 corresponds to the $n \times n$ (square) matrix C_1 and H_2 corresponds to the $m \times m$ matrix C_2, then F, which corresponds to the matrix A, also corresponds to the matrix $C_2 \cdot B \cdot C_1^{-1}$, and thus we have the following equality of $m \times n$ matrices:

$$A = C_2 \cdot B \cdot C_1^{-1}.$$

A case that often arises is when $n = m$, so that A and B are square matrices, and the two bases $v_1, \dots, v_n$ and $w_1, \dots, w_n$ are equal. In this case $C_1 =$

$C_2 = C$, say, and the above formula reads

$$A = C \cdot B \cdot C^{-1}.$$

We say that A is obtained by *conjugating B by C*.

Similarly, suppose that V is a finite dimensional vector space and that $F\colon V \to V$ is a linear map. Choosing a basis $v_1, \dots, v_n$ of V identifies F with a square matrix A. Equivalently, if $H_1\colon k^n \to V$ is the isomorphism $H_1(t_1, \dots, t_n) = \sum_i t_i v_i$ defined by the basis $v_1, \dots, v_n$, then A corresponds to the linear map $G_1\colon k^n \to k^n$ defined by $G_1 = H_1^{-1} \circ F \circ H_1$. Choosing a different basis $w_1, \dots, w_n$ of V gives a different isomorphism $H_2\colon k^n \to V$ defined by $H_2(t_1, \dots, t_n) = \sum_i t_i w_i$, and hence a different matrix B corresponding to $G_2 = H_2^{-1} \circ F \circ H_2$. Then, using $(H_2^{-1} \circ H_1)^{-1} = H_1^{-1} \circ H_2$, we have

$$G_2 = (H_2^{-1} \circ H_1) \circ G_1 \circ (H_2^{-1} \circ H_1)^{-1} = H \circ G_1 \circ H^{-1},$$

say, where $H = H_2^{-1} \circ H_1\colon k^n \to k^n$ is invertible, and corresponds to the invertible matrix C. Thus

$$B = CAC^{-1}.$$

It follows that two different choices of bases in V lead to matrices $A, B \in \mathbb{M}_n(k)$ which are **conjugate**.

Definition 5.1. Let V be a vector space. We let $GL(V)$ be the group of isomorphisms from V to itself. If V is finite dimensional, with $\dim V = n$, then $GL(V) \cong GL_n(k)$ by choosing a basis of V. More precisely, fixing a basis $\{v_1, \dots, v_n\}$, we define $\Phi\colon GL(V) \to GL_n(k)$ by setting $\Phi(F) = A$, where A is the matrix of F with respect to the basis $\{v_1, \dots, v_n\}$. Choosing a different basis $\{w_1, \dots, w_n\}$ replaces Φ by $C\Phi C^{-1}$, where C is the appropriate change of basis matrix.

6 Determinant and trace

We begin by recalling the standard properties of the determinant. For every n, we have a function $\det\colon \mathbb{M}_n(k) \to k$, which is a linear function of the columns of A (i.e. is a linear function of the i^{th} column if we hold all of the other columns fixed) and is *alternating*: given $i \neq j$, $1 \leq i, j \leq n$, if A' is the matrix obtained by switching the i^{th} and j^{th} columns of A, then $\det A' = -\det A$. It is unique up the normalizing condition that

$$\det I = 1.$$

One can show that $\det A$ can be evaluated by expanding about the i^{th} row for any i:

$$\det A = \sum_k (-1)^{i+k} a_{ik} \det A_{ik}.$$

Here $(a_{i1}, \ldots, a_{in})$ is the i^{th} row of A and A_{ik} is the $(n-1) \times (n-1)$ matrix obtained by deleting the i^{th} row and k^{th} column of A. $\det A$ can also be evaluated by expanding about the j^{th} column of A:

$$\det A = \sum_k (-1)^{k+j} a_{kj} \det A_{kj}.$$

In terms of the symmetric group S_n,

$$\det A = \sum_{\sigma \in S_n} (\operatorname{sign} \sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n},$$

where $\operatorname{sign}(\sigma)$ (sometimes denoted $\varepsilon(\sigma)$) is $+1$ if σ is even and -1 if σ is odd.

For example, if A is an upper triangular matrix ($a_{ij} = 0$ if $i > j$), and in particular if A is a diagonal matrix ($a_{ij} = 0$ for $i \neq j$) then $\det A$ is the product of the diagonal entries of A.

Another useful fact is the following:

$$\det A = \det({}^t A),$$

where, if $A = (a_{ij})$, then ${}^t A$ is the matrix with $(i, j)^{\text{th}}$ entry a_{ji}.

We have the important result:

Proposition 6.1. *For all $n \times n$ matrices A and B, $\det AB = \det A \det B$.*

Proposition 6.2. *The matrix A is invertible $\iff$ $\det A \neq 0$, and in this case*

$$\det(A^{-1}) = \frac{1}{\det A}.$$

Hence $\det \colon GL_n(k) \to k^*$ is a homomorphism, where k^* is the multiplicative group of nonzero elements of k.

Corollary 6.3. *If C is an invertible $n \times n$ matrix and A is any $n \times n$ matrix, then $\det A = \det(CAC^{-1})$. Thus, the determinant of a linear map can be read off from its matrix with respect to any basis.*

Proof. Immediate from $\det(CAC^{-1}) = \det C \cdot \det A \cdot \det(C^{-1}) = \det C \cdot \det A \cdot (\det C)^{-1} = \det A$. $\square$

An important consequence of the corollary is the following: let V be a finite dimensional vector space and let $F: V \to V$ be a linear map. Choosing a basis $v_1, \dots, v_n$ for V gives a matrix A associated to F and the basis $v_1, \dots, v_n$. Changing the basis $v_1, \dots, v_n$ replaces A by CAC^{-1} for some invertible matrix C. Since $\det A = \det(CAC^{-1})$, in fact the determinant $\det F$ is **well-defined**, i.e. independent of the choice of basis.

Thus for example if $v_1, \dots, v_n$ is a basis of k^n and, for every i, there exists a $d_i \in k$ such that $Av_i = d_i v_i$, then $\det A = d_1 \cdots d_n$.

One important application of determinants is to finding eigenvalues and eigenvectors. Recall that, if $F: k^n \to k^n$ (or $F: V \to V$, where V is a finite dimensional vector space) is a linear map, a nonzero vector $v \in k^n$ (or V) is an *eigenvector* of F with *eigenvalue* $\lambda \in k$ if $F(v) = \lambda v$. Since $F(v) = v \iff \lambda v - F(v) = 0 \iff v \in \operatorname{Ker}(\lambda \operatorname{Id} - F)$, we see that v is an eigenvector of F with eigenvalue $\lambda \iff \det(\lambda I - A) = 0$, where A is the matrix associated to F (and some choice of basis in the case of a finite dimensional vector space V). Defining the *characteristic polynomial* $p_A(t) = \det(tI - A)$, it is easy to see that p_A is a polynomial of degree n whose roots are the eigenvalues of F.

A linear map $F: V \to V$ is *diagonalizable* if there exists a basis $v_1, \dots, v_n$ of V consisting of eigenvectors; similarly for a matrix $A \in \mathbb{M}_n(k)$. This says that the matrix A for F in the basis $v_1, \dots, v_n$ is a diagonal matrix (the only nonzero entries are along the diagonal), and in fact $a_{ii} = \lambda_i$, where λ_i is the eigenvalue corresponding to the eigenvector v_i. Hence, for all positive integers d (and all integers if F is invertible), $v_1, \dots, v_n$ are also eigenvectors for F^d or A^d, and the eigenvalue corresponding to v_i is λ_i^d. If the characteristic polynomial p_A has n distinct roots, then A is diagonalizable. In general, not every matrix is diagonalizable, even for an algebraically closed field, but we do have the following:

Proposition 6.4. *If k is algebraically closed, for example if $k = \mathbb{C}$, and if V is a finite dimensional vector space and $F: V \to V$ is a linear map, then there exists at least one eigenvector for F.*

Remark 6.5. Let $A \in \mathbb{M}_n(k)$. If $f(t) = \sum_{i=0}^d a_i t^i \in k[t]$ is a polynomial in t, then we can apply f to the matrix A: $f(A) = \sum_{i=0}^d a_i A^i \in \mathbb{M}_n(k)$. Moreover, evaluation at A defines a homomorphism $\mathrm{ev}_A: k[t] \to \mathbb{M}_n(k)$, viewing $\mathbb{M}_n(k)$ as a (non-commutative) ring. There is a unique monic polynomial $m_A(t)$ such that (i) $m_A(A) = 0$ and (ii) if $f(t) \in k[t]$ is any polynomial such that $f(A) = 0$, then $m_A(t)$ divides $f(t)$ in the ring $k[t]$. In fact, we can take $m_A(t)$ to be the monic generator of the principal ideal $\operatorname{Ker} \mathrm{ev}_A$.

Then one has the *Cayley-Hamilton theorem*: If $p_A(t)$ is the characteristic polynomial of A, then $p_A(A) = 0$ and hence $m_A(t)$ divides $p_A(t)$. For example, if A is diagonalizable with eigenvalues $\lambda_1, \dots, \lambda_n$, then

$$p_A(t) = (t - \lambda_1) \cdots (t - \lambda_n),$$

and it is easy to see that $p_A(A) = 0$ in this case. On the other hand, in this case the minimal polynomial is the product of the factors $(t - \lambda_i)$ for **distinct** eigenvalues λ_i. Hence we see that, in the case where A is diagonalizble, $m_A(t)$ divides $p_A(t)$, but they are equal $\iff$ there are no repeated eigenvalues.

Finally, we define the *trace* of a matrix. If $A = (a_{ij}) \in \mathbb{M}_n(k)$, we define

$$\operatorname{Tr} A = \sum_{i=1}^{n} a_{ii}.$$

Thus the trace of a matrix is the sum of its diagonal entries. Clearly, $\operatorname{Tr}\colon \mathbb{M}_n(k) \to k$ is a linear map. As such, it is much simpler that the very nonlinear function det. However, they are related in several ways. For example, expanding out the characteristic polynomial $p_a(t) = \det(t \operatorname{Id} - A)$ as a function of t, one can check that

$$p_a(t) = t^n - (\operatorname{Tr} A)t^{n-1} + \cdots + (-1)^n \det A.$$

In a related interpretation, Tr is the derivative of det at the identity. One similarity between Tr and det is the following:

Proposition 6.6. *For all $A, B \in \mathbb{M}_n(k)$,*

$$\operatorname{Tr}(AB) = \operatorname{Tr}(BA).$$

Proof. By definition of matrix multiplication, AB is the matrix whose (i, j) entry is $\sum_{r=1}^{n} a_{ir}b_{rj}$, and hence

$$\operatorname{Tr}(AB) = \sum_{i=1}^{n} \sum_{r=1}^{n} a_{ir}b_{ri}.$$

Since $\operatorname{Tr}(BA)$ is obtained by reversing the roles of A and B,

$$\operatorname{Tr}(BA) = \sum_{i=1}^{n} \sum_{r=1}^{n} b_{ir}a_{ri} = \sum_{i=1}^{n} \sum_{r=1}^{n} a_{ri}b_{ir} = \operatorname{Tr}(AB),$$

after switching the indices r and i. $\square$

Corollary 6.7. *For $A \in \mathbb{M}_n(k)$ and $C \in GL_n(k)$,*

$$\operatorname{Tr}(CAC^{-1}) = \operatorname{Tr} A.$$

Hence, if V is a finite dimensional vector space and $F\colon V \to V$ is a linear map, then $\operatorname{Tr} F$ is well-defined. □

7 Direct sums

Definition 7.1. Let V_1 and V_2 be two vector spaces. We define the *direct sum* or *external direct sum* $V_1 \oplus V_2$ to be the product $V_1 \times V_2$, with + and scalar multiplication defined componentwise:

$$(v_1, v_2) + (w_1, w_2) = (v_1 + w_1, v_2 + w_2) \text{ for all } v_1, w_1 \in V_1 \text{ and } v_2, w_2 \in V_2;$$
$$t(v_1, v_2) = (tv_1, tv_2) \text{ for all } v_1 \in V_1,\ v_2 \in V_2, \text{ and } t \in k.$$

It is easy to see that, with this definition, $V_1 \oplus V_2$ is a vector space. It contains subspaces $V_1 \oplus \{0\} \cong V_1$ and $\{0\} \oplus V_2 \cong V_2$, and every element of $V_1 \oplus V_2$ can be uniquely written as a sum of an element of $V_1 \oplus \{0\}$ and and element of $\{0\} \oplus V_2$. Define $i_1\colon V_1 \to V_1 \oplus V_2$ by $i_1(v) = (v, 0)$, and similarly define $i_2\colon V_2 \to V_1 \oplus V_2$ by $i_1(w) = (0, w)$. Then i_1 and i_2 are linear and i_1 is an isomorphism from V_1 to $V_1 \oplus \{0\}$ and likewise for i_2.

The direct sum $V_1 \oplus V_2$ has the following "universal" property:

Lemma 7.2. *If V_1, V_2, W are vector spaces and $F_1\colon V_1 \to W$, $F_2\colon V_2 \to W$ are linear maps, then the function $F_1 \oplus F_2\colon V_1 \oplus V_2 \to W$ defined by*

$$(F_1 \oplus F_2)(v_1, v_2) = F_1(v_1) + F_2(v_2)$$

is linear, and satisfies $F(v_1, 0) = F_1(v_1)$, $F(0, v_2) = F_2(v_2)$, i.e. $F_1 = F \circ i_1$, $F_2 = F \circ i_2$. Conversely, given a linear map $G\colon V_1 \oplus V_2 \to W$, if we define $G_1\colon V_1 \to W$ by $G_1(v_1) = G(v_1, 0)$ and $G_2\colon V_2 \to W$ by $G_2(v_2) = G(0, v_2)$, so that $G_1 = G \circ i_1$ and $G_2 = G \circ i_2$, then G_1 and G_2 are linear and $G = G_1 \oplus G_2$. □

The direct sum of V_1 and V_2 also has a related property that makes it into the *product* of V_1 and V_2. Let $\pi_1\colon V_1 \oplus V_2 \to V_1$ be the projection onto the firsy factor: $\pi_1(v_1, v_2) = v_1$, and similarly for $\pi_2\colon V_1 \oplus V_2 \to V_2$. It is easy to check from the definitions that the π_i are linear and surjective, with $\pi_1 \circ i_1 = \mathrm{Id}$ on V_1 and $\pi_2 \circ i_2 = \mathrm{Id}$ on V_2. Note that $\pi_1 \circ i_2 = 0$ and $\pi_2 \circ i_1 = 0$.

Lemma 7.3. *If V_1, V_2, W are vector spaces and $G_1 \colon W \to V_1$, $G_2 \colon W \to V_2$ are linear maps, then the function $(G_1, G_2) \colon W \to V_1 \oplus V_2$ defined by*

$$(G_1, G_2)(w) = (G_1(w), G_2(w))$$

is linear, and satisfies $\pi_1 \circ (G_1, G_2) = G_1$, $\pi_2 \circ (G_1, G_2) = G_2$. Conversely, given a linear map $G \colon W \to V_1 \oplus V_2$, if we define $G_1 \colon W \to V_1$ by $G_1 = \pi_1 \circ G$, $G_2 \colon W \to V_2$ by $\pi_2 \circ G$, then G_1 and G_2 are linear and $G = (G_1, G_2)$. $\square$

Remark 7.4. It may seem strange to introduce the notation $\oplus$ for the Cartesian product. One reason for doing so is that we can define the direct sum and direct product for an arbitrary collection $V_i, i \in I$ of vector spaces V_i. In this case, when I is infinite, the direct sum and direct product have very different properties.

Finally, if we have vector spaces V, W, U, X and linear maps $F \colon V \to U$ and $G \colon W \to X$, we can define the linear map $F \oplus G \colon V \oplus W \to U \oplus X$ (same notation as Lemma 7.2) by the formula

$$(F \oplus G)(v, w) = (F(v), G(w)).$$

The direct sum $V_1 \oplus \cdots \oplus V_d = \bigoplus_{i=1}^{d} V_i$ of finitely many vector spaces is similarly defined, and has similar properties. It is also easy to check for example that

$$V_1 \oplus V_2 \oplus V_3 \cong (V_1 \oplus V_2) \oplus V_3 \cong V_1 \oplus (V_2 \oplus V_3).$$

It is straightforward to check the following (proof as exercise):

Lemma 7.5. *Suppose that $v_1, \dots, v_n$ is a basis for V_1 and that $w_1, \dots, w_m$ is a basis for V_2. Then $(v_1, 0), \dots, (v_n, 0), (0, w_1), \dots, (0, w_m)$ is a basis for $V_1 \oplus V_2$. Hence V_1 and V_2 are finite dimensional $\iff$ $V_1 \oplus V_2$ is finite dimensional, and in this case*

$$\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2. \quad \square$$

In particular, we can identify $k^n \oplus k^m$ with k^{n+m}. Of course, in some sense we originally defined k^n as $\underbrace{k \oplus \cdots \oplus k}_{n \text{ times}}$.

A linear map $F \colon k^n \to k^r$ is the same thing as an $r \times n$ matrix A, and a linear map $G \colon k^m \to k^r$ is the same thing as an $r \times m$ matrix B. It is then easy to check that the linear map $F \oplus G \colon k^n \oplus k^m \cong k^{n+m} \to k^r$ is the $r \times (n+m)$ matrix $\begin{pmatrix} A & B \end{pmatrix}$. Likewise, given linear maps $F \colon k^r \to k^n$ and

$G\colon k^r \to k^m$ corresponding to matrices A, B, the linear map $(F, G)\colon k^r \to k^{m+n}$ corresponds to the $(n+m) \times r$ matrix $\begin{pmatrix} A \\ B \end{pmatrix}$. Finally, if given linear maps $F\colon k^n \to k^r$ and $G\colon k^m \to k^s$ corresponding to matrices A, B, the linear map $F \oplus G\colon k^{n+m} \to k^{r+s}$ corresponds to the $(r+s) \times (n+m)$ matrix $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$. In particular, in the case $n = r$ and $m = s$, A, B, and $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$ are square matrices (of sizes n, m, and $n+m$ respectively). We thus have:

Proposition 7.6. *In the above notation,*

$$\operatorname{Tr}(F \oplus G) = \operatorname{Tr} F + \operatorname{Tr} G. \quad \square$$

8 Internal direct sums

The term "external direct sum" suggests that there should also be a notion of an "internal direct sum." Suppose that W_1 and W_2 are subspaces of a vector space V. Define

$$W_1 + W_2 = \{w_1 + w_2 : w_1 \in W_1, w_2 \in W_2\}.$$

It is easy to see directly that $W_1 + W_2$ is a subspace of V. To see this another way, let $i_1\colon W_1 \to V$ and $i_2\colon W_2 \to V$ denote the inclusions, which are linear. Then as we have seen above there is an induced linear map $i_1 \oplus i_2\colon W_1 \oplus W_2 \to V$, defined by

$$(i_1 \oplus i_2)(w_1, w_2) = w_1 + w_2,$$

viewed as an element of V. Thus $W_1 + W_2 = \operatorname{Im}(i_1 \oplus i_2)$, and hence is a subspace of V by Lemma 2.2. The sum $W_1 + \cdots + W_d$ of an arbitrary number of subspaces is defined in a similar way.

Definition 8.1. The vector space V is an *internal direct sum* of the two subspaces W_1 and W_2 if the linear map $(i_1 \oplus i_2)\colon W_1 \oplus W_2 \to V$ defined above is an isomorphism.

We shall usually omit the word "internal" and understand the statement that V is a direct sum of the subspaces W_1 and W_2 to always mean that it is the internal direct sum.

Proposition 8.2. *Let W_1 and W_2 be subspaces of a vector space V. Then V is the direct sum of W_1 and W_2 $\iff$ the following two conditions hold:*

(i) $W_1 \cap W_2 = \{0\}$.

(ii) $V = W_1 + W_2$.

Proof. The map $(i_1 \oplus i_2)\colon W_1 \oplus W_2 \to V$ is an isomorphism $\iff$ it is injective and surjective. Now $\operatorname{Ker}(i_1 \oplus i_2) = \{(w_1, w_2) \in W_1 \oplus W_2 : w_1 + w_2 = 0\}$. But $w_1 + w_2 = 0 \iff w_2 = -w_1$, and this is only possible if $w_1 \in W_1 \cap W_2$. Hence

$$\operatorname{Ker}(i_1 \oplus i_2) = \{(w, -w) \in W_1 \oplus W_2 : w \in W_1 \cap W_2\}.$$

In particular, $W_1 \cap W_2 = \{0\} \iff \operatorname{Ker}(i_1 \oplus i_2) = \{0\} \iff i_1 \oplus i_2$ is injective. Finally, since $\operatorname{Im}(i_1 \oplus i_2) = W_1 + W_2$, we see that $W_1 + W_2 = V$ $\iff i_1 \oplus i_2$ is surjective. Putting this together, V is the direct sum of W_1 and $W_2 \iff i_1 \oplus i_2$ is both injective and surjective $\iff$ both (i) and (ii) above hold. □

More generally, it is easy to check the following:

Proposition 8.3. *Let $W_1, \dots, W_d$ be subspaces of a vector space V. Then V is the direct sum of $W_1, \dots, W_d$, i.e. the induced linear map $W_1 \oplus \cdots \oplus W_d \to V$ is an isomorphism, $\iff$ the following two conditions hold:*

(i) *Given $w_i \in W_i$, if $w_1 + \cdots + w_d = 0$, then $w_i = 0$ for every i.*

(ii) *$V = W_1 + \cdots + W_d$.* □

In general, to write a finite dimensional vector space as a direct sum $W_1 \oplus \cdots \oplus W_d$, where none of the W_i is 0, is to decompose the vector space into hopefully simpler pieces. For example, every finite dimensional vector space V is a direct sum $V = L_1 \oplus \cdots \oplus L_n$, where the L_i are one dimensional subspaces. Note however that, if V is one dimensional, then it is not possible to write $V \cong W_1 \oplus W_2$, where both of W_1, W_2 are nonzero. So, in this sense, the one dimensional vector spaces are the building blocks for all finite dimensional vector spaces.

Definition 8.4. If W is a subspace of V, then a *complement to W* is a subspace W' of V such that V is the direct sum of W and W'. Given W, it is easy to check that a complement to W always exists, and in fact there are many of them.

One important way to realize V as a direct sum is as follows: suppose that $V = W_1 \oplus W_2$. On the vector space $W_1 \oplus W_2$, we have the projection map $p_1\colon W_1 \oplus W_2 \to W_1 \oplus W_2$ defined by $p_1(w_1, w_2) = (w_1, 0)$. Let $p\colon V \to W_1$

denote the corresponding linear map via the isomorphism $V \cong W_1 \oplus W_2$. Concretely, given $v \in V$, $p(v)$ is defined as follows: write v (uniquely) as $w_1 + w_2$, where $w_1 \in W_1$ and $w_2 \in W_2$. Then $p(v) = w_1$. The linear map p has the following properties:

1. $\operatorname{Im} p = W_1$.
2. For all $w \in W_1$, $p(w) = w$.
3. $\operatorname{Ker} p = W_2$.

The next proposition says that we can reverse this process, and will be a basic tool in decomposing V as a direct sum.

Proposition 8.5. *Let V be a vector space and let $p\colon V \to V$ be a linear map. Let W be the subspace* $\operatorname{Im} p$*, and suppose that, for all $w \in W$, $p(w) = w$. Then, for $W' = \operatorname{Ker} p$, V is the direct sum of W and W'.*

Proof. We must show that W and W' satisfy (i) and (ii) of Proposition 8.3. First suppose that $w \in W \cap W' = W \cap \operatorname{Ker} p$. Then $p(w) = w$ since $w \in W$, and $p(w) = 0$ since $w \in W' = \operatorname{Ker} p$. Hence $w = 0$, i.e. $W \cap W' = \{0\}$. Now let $v \in V$. Then $w = p(v) \in W$, and $v = p(v) + (v - p(v)) = w + w'$, say, where we have set $w = p(v)$ and $w' = v - p(v) = v - w$. Since $w \in W$,

$$p(v - w) = p(v) - p(w) = w - w = 0,$$

where we have used $p(w) = w$ since $w \in W$. Thus $w' \in \operatorname{Ker} p$ and hence $v = w + w'$, where $w \in W$ and $w' \in W'$. Hence $W + W' = V$. It follows that both (i) and (ii) of Proposition 8.3 are satisfied, and so V is the direct sum of W and W'. $\square$

Definition 8.6. Let V be a vector space and W a subspace. A linear map $p\colon V \to V$ is a *projection onto* W if $\operatorname{Im} = W$ and $p(w) = w$ for all $w \in W$. Note that p is not in general uniquely determined by W.

Proposition 8.7. *Let V be a finite dimensional vector space and let W be a subspace. If $p\colon V \to V$ is a projection onto W, then*

$$\operatorname{Tr} p = \dim W.$$

Proof. There exists a basis $w_1, \dots, w_a, w_{a+1}, \dots, w_n$ of V for which $w_1, \dots, w_a$ is a basis of W and $w_{a+1}, \dots, w_n$ is a basis of $\operatorname{Ker} p$. Here $a = \dim W$. In this basis, the matrix for p is a diagonal matrix A with diagonal entries $a_{ii} = 1$ for $i \leq a$ and $a_{ii} = 0$ for $i > a$. Thus $\operatorname{Tr} p = \operatorname{Tr} A = a = \dim W$. $\square$

9 New vector spaces from old

There are many ways to construct new vector spaces out of one or more given spaces; the direct sum $V_1 \oplus V_2$ is one such example. As we saw there, we also want to be able to construct new linear maps from old ones. We give here some other examples:

Quotient spaces: Let V be a vector space, and let W be a subspace of V. Then, in particular, W is a subgroup of V (under addition), and since V is abelian W is automatically normal. Thus we can consider the group of cosets

$$V/W = \{v + W : v \in V\}.$$

It is a group under coset addition. Also, given a coset $v + W$, we can define scalar multiplication by the rule

$$t(v + W) = tv + W.$$

We must check that this is well-defined (independent of the choice of representative of the coset $v + W$): if v and $v + w$ are two elements of the same coset, then $t(v + w) = tv + tw$, and this is in the same coset as tv since $tw \in W$. It is then easy to check that the vector space axioms hold for V/W under coset addition and scalar multiplication as defined above.

If V is finite dimensional, then so is W. Let $w_1, \ldots, w_a$ be a basis for W. As we have see, we can complete the linearly independent vectors $w_1, \ldots, w_a$ to a basis $w_1, \ldots, w_a, w_{a+1}, \ldots, w_n$ for V. It is then easy to check that $w_{a+1} + W, \ldots, w_n + W$ form a basis for V/W. In particular,

$$\dim(V/W) = \dim V - \dim W.$$

We have the natural surjective homomorphism $\pi \colon V \to V/W$, and it is easy to check that it is linear. Given any linear map $F \colon V/W \to U$, where U is a vector space, $G = F \circ \pi \colon V \to U$ is a linear map satisfying: $G(w) = 0$ for all $w \in W$. We can reverse this process: If $G \colon V \to U$ is a linear map such that $G(w) = 0$ for all $w \in W$, then define $F \colon V/W \to U$ by the formula: $F(v + W) = G(v)$. If we replace v by some other representative in the coset $v + W$, necessarily of the form $v + w$ with $w \in W$, then $G(v + w) = G(v) + G(w) = G(v)$ is unchanged. Hence the definition of F is independent of the choice of representative, so there is an induced homomorphism $F \colon V/W \to U$. It is easy to check that F is linear. Summarizing:

Proposition 9.1. *Let V be a vector space and let W be a subspace. For a vector space U, there is a bijection (described above) from the set of linear* *maps $F\colon V/W \to U$ and the set of linear maps $G\colon V \to U$ such that $G(w) = 0$ for all $w \in W$.* □

Dual spaces and spaces of linear maps: Let V be a vector space and define the *dual vector space* V^* (sometimes written $V^\vee$) by:

$$V^* = \{F \in k^V : F \text{ is linear}\}.$$

(Recall that k^V denotes the set of functions from V to k.) Thus the elements of V^* are linear maps $V \to k$. Similarly, if V and W are two vector spaces, define

$$\mathrm{Hom}(V, W) = \{F \in W^V : F \text{ is linear}\}.$$

(Here, again, W^V denotes the set of functions from V to W.) Then the elements of $\mathrm{Hom}(V, W)$ are linear maps $F\colon V \to W$, and $V^* = \mathrm{Hom}(V, k)$.

Proposition 9.2. $\mathrm{Hom}(V, W)$ *is a vector subspace of W^V, and hence a vector space in its own right under pointwise addition and scalar multiplication. In particular, V^* is a vector space under pointwise addition and scalar multiplication.*

Proof. By definition, $\mathrm{Hom}(V, W)$ is a subset of W^V. It is closed under pointwise addition since, if $F_1\colon V \to W$ and $F_2\colon V \to W$ are both linear, then so is $F_1 + F_2$ (proof as exercise). Similarly, $\mathrm{Hom}(V, W)$ is closed under scalar multiplication. Since $0 \in \mathrm{Hom}(V, W)$, and $\mathrm{Hom}(V, W)$ is a subgroup of W^V closed under scalar multiplication, $\mathrm{Hom}(V, W)$ is a vector subspace of W^V. □

Suppose that V is finite dimensional with basis $v_1, \dots, v_n$. We define the *dual basis* $v_1^*, \dots, v_n^*$ of V^* as follows: $v_i^*\colon V \to k$ is the unique linear map such that $v_i^*(v_j) = 0$ if $i \neq j$ and $v_i^*(v_i) = 1$. (Recall that every linear map $V \to k$ is uniquely specified by its values on a basis, and every possible set of values arises in this way.) We summarize the above requirement by writing $v_i^*(v_j) = \delta_{ij}$, where δ_{ij}, the *Kronecker delta function*, is defined to be 1 if $i = j$ and 0 otherwise. It is easy to check that $v_1^*, \dots, v_n^*$ is in fact a basis of V^*, hence $\dim V^* = \dim V$ if V is finite dimensional. In particular, for $V = k^n$ with the standard basis $e_1, \dots, e_n$, $V^* \cong k^n$ with basis $e_1^*, \dots, e_n^*$. But thus isomorphism changes in a complicated way with respect to linear maps, and cannot be defined intrinsically (i.e. without choosing a basis).

More generally, suppose that V and W are both finite dimensional, with bases $v_1, \dots, v_n$ and $w_1, \dots, w_m$ respectively. Define the linear map

$v_i^* w_j \colon V \to W$ as follows:

$$v_i^* w_j(v_\ell) = \begin{cases} 0, & \text{if } \ell \neq i; \\ w_j, & \text{if } \ell = i. \end{cases}$$

In other words, for all $v \in V$, $(v_i^* w_j)(v) = v_i^*(v) w_j$. Suppose that $F \colon V \to W$ corresponds to the matrix $A = (a_{ij})$ with respect to the bases $v_1, \dots, v_n$ and $w_1, \dots, w_m$. In other words, $F(v_i) = \sum_{j=1}^m a_{ji} w_j$. Then by definition $F = \sum_{i,j} a_{ji} v_i^* w_j$. It then follows that the $v_i^* w_j$, $1 \leq i \leq n$, $1 \leq j \leq m$, are a basis for $\mathrm{Hom}(V, W)$ (we saw a special case of this earlier for $V \cong k^n$, $W \cong k^m$, where we denoted $e_j^* e_i$ by E_{ij}). Thus

$$\dim \mathrm{Hom}(V, W) = (\dim V)(\dim W).$$

The "universal" properties of direct sums can be expressed as follows:

Proposition 9.3. *Let V, W, U be vector spaces. Then*

$$\mathrm{Hom}(V \oplus W, U) \cong \mathrm{Hom}(V, U) \oplus \mathrm{Hom}(W, U);$$
$$\mathrm{Hom}(U, V \oplus W) \cong \mathrm{Hom}(U, V) \oplus \mathrm{Hom}(U, W).$$

In fact, the statements of Lemma 7.2 and Lemma 7.3 give an explicit construction of the isomorphisms.

Next we see how linear maps behave for these constructions. Suppose that V_1 and V_2 are two vector spaces and that $G \colon V_1 \to V_2$ is a linear map. Given an element F of V_2^*, in other words a linear map $F \colon V_2 \to k$, we can consider the composition on the right with G: define

$$G^*(F) = F \circ G \colon V_1 \to k.$$

Since the composition of two linear maps is linear, $G^*(F) \in V_1^*$, and a straightforward calculation shows that $G^*(F_1 + F_2) = G^*(F_1) + G^*(F_2)$ and that $G^*(tF) = tG^*(F)$, so that G^* is linear. In fact, one can show the following (proof omitted):

Proposition 9.4. *The function $G \mapsto G^*$ is a linear function from the vector space $\mathrm{Hom}(V_1, V_2)$ to the vector space $\mathrm{Hom}(V_2^*, V_1^*)$, and it is an isomorphism if V_1 and V_2 are finite dimensional.*

Remark 9.5. In case $V_1 = k^n$ and $V_2 = k^m$, a linear map from V_1 to V_2 is the same thing as an $m \times n$ matrix $A = (a_{ij})$. Using the isomorphisms $(k^n)^* \cong k^n$ and $(k^m)^* \cong k^m$ via the dual bases e_i^*, there is a natural isomorphism from

$\mathrm{Hom}(k^n, k^m) = \mathbb{M}_{m,n}(k)$ to $\mathrm{Hom}(k^m, k^n) = \mathbb{M}_{n,m}(k)$. In other words, we have a natural way to associate an $n \times m$ matrix to the $m \times n$ matrix A. It is easy to check that this is the *transpose matrix*

$$^tA = (a_{ji}).$$

Clearly, $A \mapsto {}^tA$ is an isomorphism from $\mathbb{M}_{m,n}(k)$ to $\mathbb{M}_{n,m}(k)$. In fact, its inverse is again transpose, and more precisely we have the obvious formula: for all $A \in \mathbb{M}_{m,n}(k)$,

$$^{tt}A = {}^t({}^tA) = A.$$

Note the reversal of order in Proposition 9.4. In fact, this extends to compositions as follows:

Proposition 9.6. *Suppose that V_1, V_2, V_3 are vector spaces and that $G_1 \colon V_1 \to V_2$ and $G_2 \colon V_2 \to V_3$ are linear maps. Then*

$$(G_2 \circ G_1)^* = G_1^* \circ G_2^*.$$

Proof. By definition, given $F \in V_3^*$,

$$(G_2 \circ G_1)^*(F) = F \circ (G_2 \circ G_1) = (F \circ G_2) \circ G_1,$$

since function composition is associative. On the other hand,

$$G_1^* \circ G_2^*(F) = G_1^*(G_2^*(F)) = G_1^*(F \circ G_2) = (F \circ G_2) \circ G_1,$$

proving the desired equality. $\square$

Remark 9.7. For the case of matrices, i.e. $V_1 = k^n$, $V_2 = k^m$, $V_3 = k^p$, the reversal of order in Proposition 9.6 is the familiar formula

$$^t(AB) = {}^tB \cdot {}^tA.$$

The change in the order above is just a fact of life, and it is the only possible formula along these lines if we want to line up domains and ranges in the right way. However, there is one case where we can keep the original order of composition. Suppose that $V_1 = V_2 = V$, say, and that G is invertible, i.e. is a linear isomorphism. Then we have the operation $G \mapsto G^{-1}$, and it also reverses the order of operations: $(G_1 \circ G_2)^{-1} = G_2^{-1} \circ G_1^{-1}$. So if we combine this with $G \mapsto G^*$, we get a function $G \mapsto (G^{-1})^*$, from invertible linear maps on V to invertible linear maps on V^*. given as follows: if $F \in V^*$, i.e. $F \colon V \to k$ is a linear function, define

$$G \cdot F = F \circ (G^{-1}).$$

(Note: a calculation shows that, if G is invertible, then G^* is invertible, and in fact we have the formula $(G^*)^{-1} = (G^{-1})^*$, so that these two operations commute.) In this case, the order is preserved, since, if G_1, G_2 are two isomorphisms, then

$$((G_1 \circ G_2)^{-1})^* = (G_2^{-1} \circ G_1^{-1})^* = (G_1^{-1})^* \circ (G_2^{-1})^*,$$

and so the order is unchanged. For example, taking $V_1 = V_2 = k^n$ in the above notation, for all $G_1, G_2 \in GL_n(k)$ and $F \in \mathrm{Hom}(k^n, W)$,

$$(G_1 \circ G_2) \cdot F = G_1 \cdot (G_2 \cdot F).$$

This says that the above defines a group action of $GL_n(k)$ on $\mathrm{Hom}(k^n, W)$.

In the case of $\mathrm{Hom}(V, W)$, we have to consider linear maps for both the domain and range. Given two vector spaces V_1 and V_2 and a linear map $G\colon V_1 \to V_2$, we can define G^* as before, by **right** composition: if $F \in \mathrm{Hom}(V_2, W)$, in other words if $F\colon V_2 \to W$ is a linear map, we can consider the linear map $G^*F = F \circ G\colon V_1 \to W$. Thus G^* is a map from $\mathrm{Hom}(V_2, W)$ to $\mathrm{Hom}(V_1, W)$, and it is linear. By the same kind of arguments as above for the dual space, if V_1, V_2, V_3 are vector spaces and $G_1\colon V_1 \to V_2$ and $G_2\colon V_2 \to V_3$ are linear maps, then

$$(G_2 \circ G_1)^* = G_1^* \circ G_2^*.$$

On the other hand, given two vector spaces W_1 and W_2 and a linear map $H\colon W_1 \to W_2$, we can define a function $H_*\colon \mathrm{Hom}(V, W_1) \to \mathrm{Hom}(V, W_2)$ by using **left** composition: for $F \in \mathrm{Hom}(V, W_1)$, we set

$$H_*(F) = H \circ F.$$

Note that the composition is linear, since both F and H were assumed linear, and hence $H^*(F) \in \mathrm{Hom}(V, W_2)$. It is again easy to check that H_* is a linear map from $\mathrm{Hom}(V, W_1)$ to $\mathrm{Hom}(V, W_2)$. Left composition preserves the order, though, since given $H_1\colon W_1 \to W_2$ and $H_2\colon W_2 \to W_3$, we have

$$\begin{aligned}(H_2 \circ H_1)_*(F) &= (H_2 \circ H_1) \circ F = H_2 \circ (H_1 \circ F) \\ &= (H_2)_*((H_1)_*(F)) = ((H_2)_* \circ (H_1)_*)(F).\end{aligned}$$

Finally, suppose given $G\colon V_1 \to V_2$ and $H\colon W_1 \to W_2$. We use the same symbol to denote $G^*\colon \mathrm{Hom}(V_2, W_1) \to \mathrm{Hom}(V_1, W_1)$ and $G^*\colon \mathrm{Hom}(V_2, W_2) \to \mathrm{Hom}(V_1, W_2)$, and also use H_* to denote both $H_*\colon \mathrm{Hom}(V_1, W_1) \to \mathrm{Hom}(V_1, W_2)$ and $H_*\colon \mathrm{Hom}(V_2, W_1) \to \mathrm{Hom}(V_2, W_2)$. With this understanding, given $F\colon V_2 \to W_1$,

$$G^* \circ H_*(F) = H \circ F \circ G = H_* \circ G^*(F).$$

Hence:

Proposition 9.8. *The functions* $G^* \circ H_*\colon \operatorname{Hom}(V_2, W_1) \to \operatorname{Hom}(V_1, W_2)$ *and* $H_* \circ G^*\colon \operatorname{Hom}(V_2, W_1) \to \operatorname{Hom}(V_1, W_2)$ *agree.* □

We paraphrase this by saying that *right and left composition commute*. Equivalently, there is a commutative diagram

$$\begin{array}{ccc} \operatorname{Hom}(V_2, W_1) & \xrightarrow{H_*} & \operatorname{Hom}(V_2, W_2) \\ {\scriptstyle G^*}\downarrow & & \downarrow{\scriptstyle G^*} \\ \operatorname{Hom}(V_1, W_1) & \xrightarrow[H_*]{} & \operatorname{Hom}(V_1, W_2). \end{array}$$

More on duality: Let V be a vector space and V^* its dual. If $V = k^n$, then $V^* \cong k^n$, and in fact a reasonable choice of basis is the dual basis $e_1^*, \dots, e_n^*$. Thus, if V is a finite dimensional vector space, then $V \cong V^*$ since both are isomorphic to k^n. However, the isomorphism depends on the choice of a basis of V.

However, let $V^{**} = (V^*)^*$ be the dual of the dual of V. We call V^{**} the *double dual* of V. There is a "natural" linear map $\mathrm{ev}\colon V \to V^{**}$ defined as follows: First, define the function $E\colon V^* \times V \to k$ by

$$E(f, v) = f(v),$$

and then set

$$\mathrm{ev}(v)(f) = E(f, v) = f(v).$$

In other words, $\mathrm{ev}(v)$ is the function on V^* defined by evaluation at v. For example, if $V = k^n$ with basis $e_1, \dots, e_n$, and $e_1^*, \dots, e_n^*$ is the dual basis of $(k^n)^*$, then it is easy to see that

$$\mathrm{ev}(e_i)(e_j^*) = \delta_{ij}.$$

This says that $\mathrm{ev}(e_i) = e_i^{**}$, where $e_1^{**}, \dots, e_n^{**}$ is the dual basis to $e_1^*, \dots, e_n^*$. Thus:

Proposition 9.9. *If* V *is finite dimensional, then* $\mathrm{ev}\colon V \to V^{**}$ *is an isomorphism.* □

Remark 9.10. If V is not finite dimensional, then V^* and hence V^{**} tend to be *much larger* than V, and so the map $V \to V^{**}$, which is always injective, is not surjective.

Another example of duality arises as follows: Given two vector spaces V and W and $F \in \mathrm{Hom}(V, W)$, we have defined $F^*\colon \mathrm{Hom}(W, k) = W^* \to \mathrm{Hom}(V, k) = V^*$, and the linear map $\mathrm{Hom}(V, W) \to \mathrm{Hom}(W^*, V^*)$ defined by $F \mapsto F^*$ is an isomorphism. Repeating this construction, given $F \in \mathrm{Hom}(V, W)$ we have defined $F^* \in \mathrm{Hom}(W^*, V^*)$, and thus $F^{**} \in \mathrm{Hom}(V^{**}, W^{**})$. If V and W are finite dimensional, then $V^{**} \cong V$, by an explicit isomorphism which is the inverse of ev, and similarly $W^{**} \cong W$. Thus we can view F^{**} as an element of $\mathrm{Hom}(V, W)$, and it is straightforward to check that, via this identification,

$$F^{**} = F.$$

In fact, translating this statement into the case $V = k^n$, $W = k^m$, this statement reduces to the identity ${}^{tt}A = A$ for all $A \in \mathbb{M}_{m,n}(k)$ of Remark 9.5.

10 Tensor products

Tensor products are another very useful way of producing new vector spaces from old. We begin by recalling a definition from linear algebra:

Definition 10.1. Let V, W, U be vector spaces. A function $F\colon V \times W \to U$ is *bilinear* if it is linear in each variable when the other is held fixed: for all $w \in W$, the function $f(v) = F(v, w)$ is a linear function from V to U, and all $v \in V$, the function $g(w) = F(v, w)$ is a linear function from W to U. Multilinear maps $F\colon V_1 \times V_2 \times \cdots \times V_n \to U$ are defined in a similar way.

Bilinear maps occur throughout linear algebra. For example, for any field k we have the *standard inner product*, often denoted by $\langle \cdot, \cdot \rangle\colon k^n \times k^n \to k$, and it is defined by: if $v = (v_1, \dots, v_n)$ and $w = (w_1, \dots, w_n)$, then

$$\langle v, w \rangle = \sum_{i=1}^{n} v_i w_i.$$

More generally, if $B\colon k^n \times k^m \to k$ is any bilinear function, then it is easy to see that there exist unique $a_{ij} \in k$ such that, for all $v = (v_1, \dots, v_n) \in k^n$, $w = (w_1, \dots, w_m) \in k^m$,

$$B(v, w) = \sum_{i,j} a_{ij} v_i w_j.$$

Here, $a_{ij} = B(e_i, e_j)$ and the formula is obtained by expanding out

$$B(\sum_i v_i e_i, \sum_j w_j e_j)$$

using the defining properties of bilinearity.

Another important example of a bilinear function is composition of linear maps: given three vector spaces V, W, U, we have the composition function

$$C\colon \operatorname{Hom}(V,W) \times \operatorname{Hom}(W,U) \to \operatorname{Hom}(V,U)$$

defined by $C(F,G) = G \circ F$, and it is easily checked to be bilinear. Correspondingly, we have matrix multiplication

$$\mathbb{M}_{n,m}(k) \times \mathbb{M}_{m,\ell}(k) \to \mathbb{M}_{n,\ell}(k),$$

and the statement that matrix multiplication distributes over matrix addition (on both sides) and commutes with scalar multiplication of matrices (in the sense that $(tA)B = A(tB) = t(AB)$) is just the statement that matrix multiplication is bilinear.

The *tensor product* $V \otimes W$ of two vector spaces is a new vector space which is often most usefully described by a "universal property" with respect to bilinear maps: First, for all $v \in V$, $w \in W$, there is a symbol $v \otimes w \in V \otimes W$, which is bilinear, i.e. for all $v, v_i \in V$, $w, w_i \in W$ and $t \in k$,

$$\begin{aligned}
(v_1 + v_2) \otimes w &= (v_1 \otimes w) + (v_2 \otimes w);\\
v \otimes (w_1 + w_2) &= (v \otimes w_1) + (v \otimes w_2);\\
(tv) \otimes w &= v \otimes (tw) = t(v \otimes w).
\end{aligned}$$

Second, for every vector space U and bilinear function $F\colon V \times W \to U$, there is a unique **linear** function $\widehat{F}\colon V \otimes W \to U$ such that, for all $v \in V$ and $w \in W$, $F(v,w) = \widehat{F}(v \otimes w)$.

One can show that, for two vector spaces V and W, the tensor product $V \otimes W$ exists and is uniquely characterized, up to a unique isomorphism, by the above universal property. The construction is not very illuminating: even if V and W are finite dimensional, the construction of $V \otimes W$ starts with the very large (infinite dimensional if k is infinite) vector space $k[V \times W]$, the free vector space corresponding to the set $V \times W$ and then taking the quotient by the subspace generated by elements of the form $(v_1 + v_2, w) - (v_1, w) - (v_2, w)$, $(v, w_1 + w_2) - (v, w_1) - (v, w_2)$, $(tv, w) - t(v, w)$, $(v, tw) - t(v, w)$. The important features of the tensor product are as follows:

1. If V and W are finite dimensional, $v_1, \dots, v_n$ is a basis of V_1 and $w_1, \dots, w_m$ is a basis of W, then $v_i \otimes w_j$, $1 \le i \le n$, $1 \le j \le m$, is a basis of $V \otimes W$. Thus, in this case $V \otimes W$ is finite dimensional as well and
$$\dim(V \otimes W) = (\dim V) \cdot (\dim W).$$

2. Given vector spaces V_1, V_2, W_1, W_2 and linear maps $G\colon V_1 \to V_2$, $H\colon W_1 \to W_2$, then there is an induced linear map, denoted $G \otimes H$, from $V_1 \otimes W_1$ to $V_2 \otimes W_2$. It satisfies: for all $v \in V_1$ and $w \in W_1$,

$$(G \otimes H)(v \otimes w) = G(v) \otimes H(w).$$

   Moreover, given another pair of vector spaces V_3, W_3, and linear maps $G'\colon V_2 \to V_3$ and $H'\colon W_2 \to W_3$, then

$$(G' \otimes H') \circ (G \otimes H) = (G' \circ G) \otimes (H' \circ H).$$

3. There is a "natural" isomorphism $S\colon V \otimes W \cong W \otimes V$, which satisfies

$$S(v \otimes w) = w \otimes v.$$

   Thus, as with direct sum (but unlike Hom), tensor product is symmetric.

4. There is a "natural" isomorphism

$$(V_1 \oplus V_2) \otimes W \cong (V_1 \otimes W) \oplus (V_2 \otimes W),$$

   and similarly for the second factor, so that tensor product distributes over direct sum.

5. There are "natural" isomorphisms

$$V_1 \otimes (V_2 \otimes V_3) \cong (V_1 \otimes V_2) \otimes V_3 \cong V_1 \otimes V_2 \otimes V_3,$$

   where by definition $V_1 \otimes V_2 \otimes V_3$ has a universal property with respect to *trilinear maps* $V_1 \times V_2 \times V_3 \to U$.

6. There is a "natural" linear map L from $V^* \otimes W$ to $\mathrm{Hom}(V, W)$, which is an isomorphism if V is finite dimensional. (Note that this is consistent with the statement that, if both V and W are finite dimensional, then $\dim(V \otimes W) = \dim(\mathrm{Hom}(V, W)) = (\dim V)(\dim W)$.) Here natural means in particular the following: Suppose that V_1, V_2, W_2, W_2 are vector spaces, and that $G\colon V_1 \to V_2$ and $H\colon W_1 \to W_2$ are linear maps. We have defined the linear map $H_* \circ G^* = G^* \circ H_*$ from $\mathrm{Hom}(V_2, W_1)$ to $\mathrm{Hom}(V_1, W_2)$. On the other hand, we have the linear map $G^*\colon V_2^* \to V_1^*$ and hence there is a linear map $G^* \otimes H\colon V_2^* \otimes W_1 \to V_1^* \otimes W_2$, and these commute with the isomorphisms $L\colon V_2^* \otimes W_1 \to$

$\mathrm{Hom}(V_2, W_1)$ and $L'\colon V_1^* \otimes W_2 \to \mathrm{Hom}(V_1, W_2)$ in the sense that the following diagram is commutative:

$$\begin{array}{ccc} V_2^* \otimes W_1 & \xrightarrow{L} & \mathrm{Hom}(V_2, W_1) \\ {\scriptstyle G^* \otimes H}\Big\downarrow & & \Big\downarrow{\scriptstyle G^* \circ H_*} \\ V_1^* \otimes W_2 & \xrightarrow[L']{} & \mathrm{Hom}(V_1, W_2). \end{array}$$

In particular, if V is finite dimensional then there is a "natural" isomorphism from $\mathrm{Hom}(V, V)$ to $V^* \otimes V$. Using this, we can give an intrinsic definition of the trace as follows:

Proposition 10.2. *The function $E\colon V^* \times V \to k$ defined by*

$$E(f, v) = f(v)$$

is bilinear. Using the isomorphism $V^* \otimes V \cong \mathrm{Hom}(V, V)$, the corresponding linear map $V^* \otimes V \to k$ is identified with the trace. $\square$

Warning: In general, it can be somewhat tricky to work with tensor products. For example, every element of a tensor product $V \otimes W$ can be written as a **finite sum** of the form $\sum_i v_i \otimes w_i$, where $v_i \in V$ and $w_i \in W$. But it is not in general possible to write **every** element of $V \otimes W$ in the form $v \otimes w$, for a **single** choice of $v \in V$, $w \in W$. Also, the expression of an element as $\sum_i v_i \otimes w_i$ is far from unique. What this means in practice, for example, if that we **cannot** try to define a linear map $G\colon V \otimes W \to U$ by simply defining G on elements of the form $v \otimes w$, and expect that G is always well defined. Instead, it is best to use the universal property of the tensor product when attempting to define linear maps from $V \otimes W$ to some other vector space.