

WOMP 2001: LINEAR ALGEBRA

DAN GROSSMAN

REFERENCE

Roman, S. *Advanced Linear Algebra*, GTM #135. (Not very good.)

1. VECTOR SPACES

Let k be a field, e.g., \mathbb{R} , \mathbb{Q} , \mathbb{C} , \mathbf{F}_q , $K(t)$,...

Definition. A *vector space* over k is a set V with two operations

$$+ : V \times V \rightarrow V \quad \text{and} \quad \cdot : k \times V \rightarrow V$$

satisfying some familiar axioms. A *subspace* of V is a subset $W \subset V$ for which

- $0 \in W$,
- If $w_1, w_2 \in W$, $a \in k$, then $aw_1 + w_2 \in W$.

The *quotient* of V by the subspace $W \subset V$ is the vector space whose elements are subsets of the form (“affine translates”)

$$v + W \stackrel{\text{def}}{=} \{v + w : w \in W\}$$

(for which $v + W = v' + W$ iff $v - v' \in W$, also written $v \equiv v' \pmod{W}$), and whose operations $+, \cdot$ are those naturally induced from the operations on V .

Exercise 1. Verify that our definition of the vector space V/W makes sense.

Given a finite collection of elements (“vectors”) $v_1, \dots, v_m \in V$, their *span* is the subspace

$$\langle v_1, \dots, v_m \rangle \stackrel{\text{def}}{=} \{a_1v_1 + \dots + a_mv_m : a_1, \dots, a_m \in k\}.$$

Exercise 2. Verify that this is a subspace.

There may sometimes be redundancy in a spanning set; this is expressed by the notion of *linear dependence*. The collection $v_1, \dots, v_m \in V$ is said to be *linearly dependent* if there is a *linear combination*

$$a_1v_1 + \dots + a_mv_m = 0, \quad \text{some } a_i \neq 0.$$

This is equivalent to being able to express at least one of the v_i as a linear combination of the others.

Exercise 3. Verify this equivalence.

Theorem. Let V be a vector space over a field k . If there exists a finite set $\{v_1, \dots, v_m\} \subset V$ whose span equals V , then:

- (1) there is a minimal such set (i.e., one from which no element can be removed while still spanning V);
- (2) every minimal spanning set is also a maximal linearly independent set (i.e., one to which no vector can be appended while still being linearly independent), and vice versa;
- (3) All sets as in (2) have the same cardinality, called the *dimension* of V .

Exercise 4. Prove this theorem.

Exercise 5. Prove that if $U, W \subset V$ are two subspaces, then

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

(Here, $U + W$ is the set of vectors of the form $u + w \in V$, with $u \in U$ and $w \in W$, which is a subspace (as is $U \cap W$).)

We will discuss only finite-dimensional vector spaces, satisfying the hypothesis of the theorem.

2. LINEAR MAPS

Let V, W be vector spaces over the same field k .

Definition. A map $\varphi : V \rightarrow W$ is *linear* if for all $v, v' \in V$, $a \in k$,

$$(1) \quad \varphi(av + v') = a\varphi(v) + \varphi(v').$$

The set of all such linear maps forms the vector space $\text{Hom}_k(V, W)$.

Given ordered bases v_1, \dots, v_n for V and w_1, \dots, w_m for W , we can define a 1–1 correspondence

$$\{\text{linear maps } \varphi : V \rightarrow W\} \longleftrightarrow \{m \times n \text{ matrices over } k\},$$

which depends on the choice of bases, as follows:

First, write each $\varphi(v_i) \in W$ as

$$\varphi(v_i) = \sum_{\alpha=1}^m a_{\alpha i} w_{\alpha},$$

and associate to φ the matrix $A = (a_{\alpha i})$. In other words, the *columns* of A are the coordinates, with respect to the range's basis, of the images of the domain's basis.

Conversely, given an $m \times n$ matrix with entries in the field k , we can define a linear map $\varphi : V \rightarrow W$ by first defining $\varphi(v_i)$ for the basis vectors using (2), and then extending to arbitrary elements of V using linearity (1). Of course, these processes are inverse to one another.

If we replace our bases v_i and w_{α} by new bases

$$v'_i = \sum_j q_{ji} v_j, \quad w'_{\alpha} = \sum_{\beta} p_{\beta\alpha} w_{\beta},$$

then the matrix associated to φ in the new basis differs from that in the old by the relation

$$(2) \quad A' = P^{-1} \cdot A \cdot Q.$$

One can think of the general properties of linear maps as those properties of matrices that are invariant under this operation.

Definition. The *kernel* of a linear map $\varphi : V \rightarrow W$ is the subspace

$$\text{Ker}(\varphi) = \{v \in V : \varphi(v) = 0 \in W\} \subset V,$$

and the *image* is the subspace (verify)

$$\text{Im}(\varphi) = \{\varphi(v) \in W : v \in V\} \cong V/\text{Ker}(\varphi).$$

The *cokernel* is the quotient space

$$\text{Coker}(\varphi) = W/\text{Im}(\varphi).$$

The *rank* of φ is the dimension of $\text{Im}(\varphi)$, and equals the dimension of the span of the columns of any matrix associated to φ (a subspace of \mathbb{R}^m , if $\dim(W) = m$).

Exercise 6. Prove that two $m \times n$ matrices A, A' have the same rank *iff* there exist P, Q satisfying (2).

A sequence of vector spaces and linear maps

$$\cdots \xrightarrow{\varphi_{i-2}} V_{i-1} \xrightarrow{\varphi_{i-1}} V_i \xrightarrow{\varphi_i} \cdots$$

is called an *exact sequence* if for each i ,

$$\text{Im}(\varphi_{i-1}) = \text{Ker}(\varphi_i).$$

Exercise 7. Show that for an exact sequence of finite-dim'l vector spaces

$$0 \rightarrow V_1 \rightarrow \cdots \rightarrow V_m \rightarrow 0,$$

we have the dimension formula

$$\dim(V_1) - \dim(V_2) + \cdots + (-1)^{m-1} \dim(V_m) = 0.$$

In particular, having a so-called *short exact sequence* of the form

$$0 \rightarrow U \rightarrow V \rightarrow W \rightarrow 0$$

means that U is isomorphic to a subspace of V , and that $W \cong V/U$.

2.1. Endomorphisms. Linear maps $\varphi : V \rightarrow V$ from a vector space V (over k) to itself are called *endomorphisms* of V , and can be deeply analyzed. This should mirror the study of square ($n \times n$) matrices, but note that we don't have to use the full (excessive) change-of-basis formula (2)—we instead apply the *same* change of basis $v'_i = \sum_j q_{ji} v_j$ in both the domain and the range, and obtain the formula

$$(3) \quad A' = P^{-1} \cdot A \cdot P.$$

So the study of endomorphisms is the same as the study of those properties of square matrices that are invariant under changes of the form (3). (By contrast, Exercise 6 shows that the only property invariant under (2) is the rank.) The space of endomorphisms is a vector space (in fact, an associative k -algebra) $\text{End}_k(V)$.

An endomorphism φ has a *determinant*, $\det(\varphi) \in k$, which one typically defines as the determinant of the matrix representing φ in any basis; this requires one to know what the determinant of a matrix is, and to know that all matrices related as in (3) have the same determinant. The definition of a matrix' determinant is wretched; we will see below that there are a few direct definitions of the determinant of the endomorphism itself, from which it is obvious *a posteriori* that different representative matrices have the same determinant. Of course, the most important facts about the determinant are that: $\det(\varphi \circ \psi) = \det(\varphi)\det(\psi)$; $\text{Ker}(\varphi) = \{0\} \Leftrightarrow \det(\varphi) \neq 0$; $\det(I) = 1$.

An endomorphism φ also has a *trace*, $\text{tr}(\varphi)$, for which exactly the same comments as in the preceding paragraph hold (except that the matrix version is not wretched, but arbitrary and seemingly worthless).

There are other ways to get at the determinant and trace of an endomorphism, which follow naturally from the following all-important considerations. Given an endomorphism $\varphi : V \rightarrow V$, it is natural to ask for vectors $v \in V$ on which φ acts only by scaling: $v \neq 0$ is called an *eigenvector* for φ if

$$\varphi(v) = \lambda \cdot v \quad \text{for some } \lambda \in k.$$

This λ is then the associated *eigenvalue*. Knowing the eigenvalues and eigenvectors of φ goes a long way toward completely understanding it.

To calculate these, we start with the eigenvalues, representing φ by the matrix $A = (a_{ij})$. Note that v can be an eigenvector for φ , with eigenvalue λ , if and only if

$$v \in \text{Ker}(\varphi - \lambda \cdot I),$$

where $I : V \rightarrow V$ is the identity map; equivalently, λ is an eigenvalue for φ if and only if

$$\det(A - \lambda \cdot I) = 0.$$

With A known, this is a polynomial of degree n in the variable λ —called the *characteristic polynomial* of A (or of φ , as it is independent of the choice of basis)—and therefore has at most n distinct roots; if k is algebraically closed, then it has exactly n roots, counted with multiplicity. The multiplicity of λ as a root of the characteristic polynomial is called the *algebraic multiplicity* of λ as an eigenvalue.

Once one knows a particular eigenvalue λ (obtained as a root of the characteristic polynomial), it is easy to compute its associated eigenvectors (which form a subspace of V , called the λ -*eigenspace*) by solving the homogeneous system of linear equations for unknown v :

$$(A - \lambda \cdot I)v = 0.$$

The dimension of the λ -eigenspace (the solution space of this system) is the *geometric multiplicity* of λ as an eigenvalue, and is less than or equal to the algebraic multiplicity. What *does* equal the algebraic multiplicity is the dimension of the *generalized λ -eigenspace*:

$$\{v \in V : (A - \lambda \cdot I)^n v = 0 \text{ for some } n = 1, 2, 3, \dots\}.$$

Returning for a moment to the characteristic polynomial: the product of the roots (which will lie in k even if some of the roots do not) is the determinant of φ , and their sum (which similarly lies in k) is the trace. This is not obvious.

Don't confuse this with the *minimal polynomial* of A (or of φ), which is the unique monic polynomial $p(X)$ of least degree in one variable that is satisfied by $\varphi \in \text{End}_k(V)$. An important fact is that the minimal polynomial of A *divides* the characteristic polynomial of A , so in particular, A satisfies its own characteristic equation.

2.2. Normal forms for endomorphisms. A *Jordan block* is a matrix of the form, for some $\lambda \in k$,

$$J_1(\lambda) = (\lambda), \quad J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad \dots \quad J_m(\lambda) = \begin{pmatrix} \lambda & 1 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \lambda & 1 \\ \dots & \dots & \dots & \lambda \end{pmatrix}.$$

Theorem (Jordan normal form). *If k is algebraically closed, and $\varphi \in \text{End}_k(V)$, then there is a basis for V for which the associated matrix is made up of Jordan blocks on the diagonal:*

$$\begin{pmatrix} J_{i_1}(\lambda_1) & 0 & \dots & \dots \\ 0 & J_{i_2}(\lambda_2) & \dots & \dots \\ \vdots & \vdots & \ddots & \dots \\ 0 & 0 & \dots & J_{i_r}(\lambda_r) \end{pmatrix}.$$

This basis is not quite unique, but the matrix in so-called “Jordan normal form” is uniquely determined (up to re-ordering of the blocks). Here, each of $\lambda_1, \dots, \lambda_r$ is an eigenvalue, but they are not necessarily distinct; however, every eigenvalue appears at least once.

The optimal situation is that in which all of the Jordan blocks are of size 1. This is equivalent to having a basis of eigenvectors, and we call the endomorphism φ (or any matrix representing it) *diagonalizable*. If the eigenvalues are distinct and lie in k , then φ is diagonalizable; and if φ is diagonalizable, then the eigenvalues all lie in k , but they may coincide. So it's hard to tell if a matrix (or endomorphism) is diagonalizable, without going through and calculating the Jordan normal form. Later (in the context of inner-product spaces) we'll see a form of diagonalizability with a much cleaner theory.

The reason that we require that k be algebraically closed in this theorem is that we need all of the eigenvalues—roots of the characteristic polynomial—to lie in the field k . So even if k is not algebraically closed, but the characteristic polynomial factors completely (into linear factors) over k , then the Jordan normal form is still available.

However, this does not always occur, and for other cases we have a less useful alternative, called *rational normal form*. Given a monic polynomial

$$p(X) = a_0 + a_1X + \cdots + a_{m-1}X^{m-1} + X^m,$$

we associate the *rational block matrix*

$$R_p = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{m-1} \end{pmatrix}.$$

Now, given $\varphi \in \text{End}_k(V)$, we factor the characteristic polynomial as far as possible over k :

$$p_\varphi(X) = p_1(X) \cdots p_s(X).$$

Each of these factors has an associated rational block, and it is a theorem that there is a basis for V for which the matrix corresponding to φ is of the block form

$$\begin{pmatrix} R_{p_1} & 0 & \cdots & 0 \\ 0 & R_{p_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R_{p_s} \end{pmatrix}.$$

3. BILINEAR FORMS

Definition. A *bilinear form* on a vector space V (over a field k) is a map

$$B : V \times V \rightarrow k$$

satisfying

$$B(av + v', w) = aB(v, w) + B(v', w), \quad B(v, aw + w') = aB(v, w) + B(v, w').$$

Given an ordered basis v_1, \dots, v_n for V , we can define a 1-1 correspondence

$$\{\text{bilinear forms on } V\} \longleftrightarrow \{n \times n \text{ matrices over } k\},$$

which depends on the choice of basis, as follows:

Given a bilinear form B , associate the matrix whose (i, j) -entry is $B(v_i, v_j)$. Conversely, given a matrix (b_{ij}) , associate the bilinear form B whose values on the basis vectors are

$$B(v_i, v_j) = b_{ij},$$

and whose values on all other vectors are determined by bilinearity. These processes are inverse to one another. Given a bilinear form B , the effect on a change of basis on the associated matrix is as follows: if $v_i = \sum_j a_{ji}v'_j$, then $b_{ij} = \sum_{k,l} a_{ki}b'_{kl}a_{lj}$, or in matrix notation,

$$(4) \quad B = A^t \cdot B' \cdot A.$$

Note that there is no such thing as the “determinant” (or “trace”) of a bilinear form, because different matrices representing the same bilinear form have (in general) different determinants (or traces).

3.1. Symmetric bilinear forms. The most important bilinear forms are *symmetric*, $B(v, w) = B(w, v)$. The associated matrix, with respect to *any* basis, is then symmetric; note that the operation (4) will always convert a symmetric B' into a symmetric B . In case the field $k = \mathbb{R}$, the most important symmetric bilinear forms are those that are *positive-definite*, meaning

$$B(v, v) \geq 0 \quad \forall v, \quad \text{equality iff } v = 0.$$

A positive-definite symmetric bilinear form (over \mathbb{R}) is called an *inner-product*, and when only one inner-product is under consideration, we often denote it by $\langle v, v' \rangle$. In this case, we have the important *Cauchy-Schwarz inequality*, derived as follows: fix $v, w \in V$, and then note that for all $t \in \mathbb{R}$,

$$0 \leq \langle v - tw, v - tw \rangle = \langle w, w \rangle t^2 - 2\langle v, w \rangle t + \langle v, v \rangle.$$

This quadratic polynomial in t , always being positive, can have at most one root, so its discriminant is non-positive, i.e.,

$$\boxed{\langle v, w \rangle^2 \leq \langle v, v \rangle \langle w, w \rangle}$$

This works in infinite dimensions (for instance, when the inner-product on a function space is defined by integration), and over \mathbb{C} (with a slightly more careful proof, reflecting the fact that in that context, an *inner-product* is defined as a *Hermitian symmetric* (as opposed to symmetric), *sesquilinear* (as opposed to bilinear) form that is positive definite).

One immediate feature of a symmetric bilinear form is its *kernel*

$$B^\perp = \{v \in V : B(v, w) = 0 \text{ for all } w \in V\}.$$

B is said to be *non-degenerate* if $B^\perp = \{0\}$; and even if B is degenerate, it induces a non-degenerate symmetric bilinear form on $\bar{V} \stackrel{\text{def}}{=} V/B^\perp$.

Exercise 8. Prove the last claim.

One fixed non-degenerate bilinear form B on a vector space V induces an identification

$$\{\text{bilinear forms on } V\} \longleftrightarrow \{\text{endomorphisms of } V\},$$

as follows. Given $\varphi \in \text{End}_k(V)$, we set

$$C(v, v') \stackrel{\text{def}}{=} B(v, \varphi(v')).$$

The inverse is a little trickier; given $C(\cdot, \cdot)$, we let $\varphi(v)$ be the unique vector $v' \in V$ such that (3.1) holds, and one has to prove that such a unique v' exists (which is only true in finite dim'ns). Equivalently, if we choose an ordered basis for V (w/r/t B), then bilinear forms and endomorphisms each correspond to $n \times n$ matrices, and *if* that basis is *orthonormal*, then all these identifications become compatible (and in particular, that of (3.1) is independent of basis). Note that because orthogonal matrices satisfy $P^{-1} = P^t$, in this setting the change-of-basis (4) for bilinear forms and the change-of-basis (3) for endomorphisms are the same.

3.2. Normalizing symmetric bilinear forms. There are normal forms for symmetric bilinear forms that are much simpler than those for endomorphisms. The first important normal-form result applies when k is a field for which -1 is a perfect square (e.g., \mathbb{C} ...but again, this has nothing to do with the notion of inner-product over \mathbb{C}):

Theorem. *If B is a symmetric bilinear form on a vector space V , over a field k for which -1 is a perfect square, then there is a basis for V with respect to which B corresponds to the matrix*

$$\begin{pmatrix} I_p & 0 \\ 0 & 0_r \end{pmatrix},$$

where $r = \dim(B^\perp)$, 0_r is the $r \times r$ zero-matrix, $p = \dim(V) - r$, and I_p is the $p \times p$ identity matrix.

Exercise 9. Prove this theorem. (Hint: think about the Gramm-Schmidt process used to find orthonormal bases.)

Now we consider the case where $k = \mathbb{R}$: Given a symmetric bilinear form B , one can use a Gramm-Schmidt-like procedure to produce a basis for which the associated matrix is of the form, for some integers (p, q, r) with $p + q + r = n$,

$$\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0_r \end{pmatrix}.$$

In such a basis v_1, \dots, v_n , we'll have that $\langle v_1, \dots, v_p \rangle$ span a maximal subspace restricted to which B is positive definite, and *all* such maximal subspaces have the same dimension; likewise for v_{p+1}, \dots, v_{p+q} and negativity. Of course, $v_{p+q+1}, \dots, v_{p+q+r}$ form a basis for the kernel of B . These three dimensions (p, q, r) are thus uniquely determined by B , and are called its *signature*. This collection of facts is often called *Sylvester's theorem*, and there is a similar version of it for Hermitian forms over \mathbb{C} .

Be careful: the set of positive vectors (i.e., for which $B(v, v) \geq 0$) is not a subspace.

Exercise 10. Become familiar with the geometry of \mathbb{R}^3 , with the bilinear form

$$B((x, y, z), (x', y', z')) = xx' + yy' - zz'.$$

The case where $k = \mathbb{Q}$ involves severe difficulties, I think.

3.3. Adjoints of linear maps. If V, W are inner-product spaces, then for any linear map $\varphi : V \rightarrow W$, there is a linear map

$$\varphi^* : W \rightarrow V,$$

defined (after a little proof) by the requirement that

$$\langle \varphi(v), w \rangle_W = \langle v, \varphi^*(w) \rangle_V \quad \text{for all } v \in V, w \in W.$$

Note that it is nothing like an inverse. With respect to *orthonormal* ordered bases, the matrix of φ^* is the transpose of the matrix of φ . We'll get a different perspective on the adjoint later, in discussing dual vector spaces.

In the context of endomorphisms $\varphi : V \rightarrow V$ (which was already more interesting than the study of linear maps), when V has an inner-product, the notion of adjoint exists and produces a very rich theory, whose generalizations are everywhere in mathematics. One has the notions of

- self-adjoint operators, for which $\varphi^* = \varphi$, so that $\langle \varphi(v), v' \rangle = \langle v, \varphi(v') \rangle$, and in an orthonormal basis have symmetric matrices;
- skew-adjoint operators, for which $\varphi^* = -\varphi$, so that $\langle \varphi(v), v' \rangle = -\langle v, \varphi(v') \rangle$, and in an orthonormal basis have skew-symmetric matrices;
- over \mathbb{C} , unitary operators (and over \mathbb{R} , orthogonal operators) for which $\varphi^* = \varphi^{-1}$, so that $\langle v, v' \rangle = \langle \varphi(v), \varphi(v') \rangle$, and in an orthonormal basis have unitary (resp. orthogonal) matrices.
- normal operators, for which $\varphi \circ \varphi^* = \varphi^* \circ \varphi$.

These natural classes of operators are important in trying to understand diagonalizability of φ . Note that over \mathbb{C} or \mathbb{R} , self-adjoint operators have all real eigenvalues, skew-adjoint operators have all imaginary eigenvalues, unitary operators have all eigenvalues of absolute value 1. Note also that all three of these classes fall into the class of normal operators.

Exercise 11. Prove that eigenspaces of a normal operator (for different eigenvalues) are pairwise orthogonal.

Previously, we considered the question of when an endomorphism of a vector space has a basis of eigenvectors (i.e., when it is diagonalizable), and got a not-so-great answer. The preceding exercise suggests that we ask when an operator on an inner-product space has an *orthonormal* basis of eigenvectors. This is answered by the spectral theorem.

Theorem. (1) An operator on a complex inner-product space V is orthonormally diagonalizable if and only if it is normal.

(2) An operator on a real inner-product space V is orthonormally diagonalizable if and only if it is self-adjoint.

(2) follows from (1) by the observation that among normal operators, those with real eigenvalues are precisely those that are self-adjoint.

Sometimes people will say: a real square matrix A is conjugate, by an orthogonal matrix, to a (real) diagonal matrix if and only if A is symmetric; and a complex square matrix A is conjugate, by a unitary matrix, to a diagonal matrix if and only if $AA^* = A^*A$.

Much more useful, and also much more difficult, is the version for Hilbert spaces (infinite dim'l).

One common application is the construction of the “square-root” of certain operators; this is ubiquitous in analysis. A self-adjoint operator φ (on a real inner-product space) is said to be positive if

$$\langle \varphi(v), v \rangle \geq 0 \quad \text{for all } v.$$

This is the same as saying that the symmetric bilinear form corresponding to φ (using the inner-product) is positive semi-definite. Now that we can diagonalize such an operator, we can define a new operator with the same eigenvectors, but whose eigenvalues are the square roots of those of φ . This will be a positive operator whose “square” equals φ .

3.4. Skew-symmetric bilinear forms. Also of some importance are *skew-symmetric* bilinear forms on V , satisfying

$$B(v, w) = -B(w, v).$$

Often one finds the definition to be instead

$$B(v, v) = 0 \quad \text{for all } v,$$

which is equivalent *unless* the field k has characteristic 2.

Exercise 12. Prove that these are equivalent.

There is also a normal form for skew-symmetric bilinear forms. We first define again

$$B^\perp = \{v \in V : B(v, w) = 0 \text{ for all } w \in V\},$$

and the following remarkable fact holds:

Theorem. $\dim(V) - \dim(B^\perp)$ is even, say $= 2k$, and there is a basis for V in which the associated matrix equals (in blocks of size k , k , and $n - 2k$)

$$\begin{pmatrix} 0 & I_k & 0 \\ -I_k & 0 & 0 \\ 0 & 0 & 0_{n-2k} \end{pmatrix}.$$

Exercise 13. Prove this theorem.

4. DUAL VECTOR SPACES

The *dual space* of a vector space V over k is the vector space

$$V^* = \{\text{Linear maps (“functionals”) } \varphi : V \rightarrow k\},$$

with the obvious operations $+$, \cdot . Associated to a basis v_1, \dots, v_n of V is the *dual basis* $\varphi_1, \dots, \varphi_n$ of V^* , defined by

$$(5) \quad \varphi_i(a_1v_1 + \dots + a_nv_n) = a_i.$$

In particular, if V is finite-dim'l, then so is V^* , with the same dimension. In this case, there is an isomorphism $V \rightarrow V^*$ defined on basis vectors by sending $v_i \mapsto \varphi_i$, and extending by linearity. Two warnings:

- Choosing a different basis for V (and consequently a different dual basis for V^*) will give a different isomorphism $V \rightarrow V^*$, so we say that V and V^* are not *canonically* isomorphic; the isomorphism is somewhat artificial.
- For infinite-dim'l vector spaces, it may not even be the case that $V \cong V^*$.

One reason that this confusion is particularly tempting is the following: an inner-product B on V induces an isomorphism $V \xrightarrow{\sim} V^*$, by

$$v \mapsto B(v, \cdot).$$

(In fact, any non-degenerate bilinear form induces such an isomorphism.) Now, when one works with $V = \mathbb{R}^n$, it is difficult to avoid being aware of the standard inner-product and thereby identifying V with V^* ...but be careful.

If $\psi : V \rightarrow W$ is a linear map between vector spaces (over the same field), then there is a *dual map* (sometimes called the *transpose*), defined by “pre-composition” of a functional (on W) with ψ :

$$\psi^* : W^* \rightarrow V^*, \quad \psi^*(\mu) = \mu \circ \psi.$$

If the matrix of ψ (with respect to some bases for V and W) is known, then the matrix for ψ^* , with respect to the dual bases for V^* and W^* , is the transpose of the original matrix.

Exercise 14. Show that for a finite-dim'l vector space V , there is an isomorphism $(V^*)^* \cong V$, which is *independent* of the choice of basis.

Exercise 15. Explain how the dual map ψ^* defined here can be identified with the adjoint of a linear map discussed previously, in the context of inner-product spaces.

5. TENSOR PRODUCT OF VECTOR SPACES

Theorem. *Given vector spaces V, W over k , there exists a vector space $V \otimes_k W$ and bilinear map $\pi : V \times W \rightarrow V \otimes_k W$ with the following “universal” property:*

for every vector space U (over k) and bilinear map $V \times W \rightarrow U$, there is a unique linear map $\varphi : V \otimes_k W \rightarrow U$ through which π factors:

$$V \times W \xrightarrow{\pi} V \otimes_k W \xrightarrow{\varphi} U.$$

Moreover, any two vector space/map pairs having the preceding property can be canonically identified.

This provides a good definition of the tensor product. The upshot is that one can replace the study of bilinear (and then tri-linear, etc.) maps by that of linear maps alone.

To actually understand $V \otimes_k W$, it is good to construct it. The idea is that its elements are (equivalence classes of) *formal expressions* of the form (with obvious operations $+, \cdot$)

$$a_1(v_1 \otimes w_1) + \cdots + a_s(v_s \otimes w_s),$$

where each $v_i \in V$, $w_i \in W$, $a_i \in k$, and the equivalence just mentioned is generated by the relations

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w, \quad v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2,$$

and

$$a(v \otimes w) = (av) \otimes w = v \otimes (aw).$$

If v_1, \dots, v_n is a basis for V , and w_1, \dots, w_m is a basis for W , then the elements $v_i \otimes w_\alpha$ form a basis for $V \otimes_k W$; in particular, the dimension of the tensor product is the product of the dimensions (unlike, e.g., direct product, for which dimensions add). Note also that there is a canonical identification $V \otimes W \cong W \otimes V$, but one should be careful about using it in cases where $W = V$.

A fundamental and widely-used fact is that there is a canonical identification

$$\mathrm{Hom}_k(V, W) \cong V^* \otimes_k W.$$

Exercise 16. Prove that there is such an identification, and if you use a basis to do so, prove that it the identification does not depend on the choice of basis (unlike, e.g., the identification between V and V^*).

Exercise 17. For the case of endomorphisms, we have $\mathrm{End}_k(V) \cong V^* \otimes_k V$. Describe the *trace* of an endomorphism not in terms of matrices (or eigenvalues) but more naturally in terms of $V^* \otimes_k V$.

Also note that with the identifications $V^* \otimes_k W \cong W \otimes_k V^*$ and $(V^*)^* \cong V$, this gives an identification

$$\mathrm{Hom}_k(V, W) \cong \mathrm{Hom}_k(W^*, V^*).$$

Exercise 18. Where have we seen this identification previously?

It is worth pointing out that one can take the various “tensor powers” $V^{\otimes m}$ of a vector space V , for all $m = 1, 2, \dots$; and there is a (non-commutative) graded k -algebra

$$\bigotimes V = k \oplus V \oplus (V \otimes V) \oplus (V^{\otimes 3}) \oplus \dots$$

5.1. Exterior powers of a vector space. Given a vector space V , for each m there is a vector space $\bigwedge^m V$ and a skew-symmetric, multi-linear map

$$\pi : \underbrace{V \times \dots \times V}_m \rightarrow \bigwedge^m V$$

having the following universal property:

for every vector space U (over k) and skew-symmetric multi-linear map $V^m \rightarrow U$, there is a unique linear map $\varphi : \bigwedge^m V \rightarrow U$ through which π factors:

$$V^m \xrightarrow{\pi} \bigwedge^m V \xrightarrow{\varphi} U.$$

Moreover, *any* vector space and map having the preceding property can be canonically identified, so this is a good definition of the *m th exterior power* of V (once these statements are proved). The upshot is that one can replace the study of skew-symmetric multi-linear maps by that of linear maps alone.

Again, this definition is incomprehensible. To really work with this, one can proceed in one of two ways.

First, one can take $\bigwedge^m V$ to be a certain subspace of $V^{\otimes m}$, as follows. For each pair of indices $1 \leq i < j \leq m$, there is an endomorphism of $V^{\otimes m}$ defined by swapping the i th and j th factors in each monomial $v_1 \otimes \dots \otimes v_m$. Then the subspace $\bigwedge^m V$ consists of those tensors which are negated by these swaps, for each pair i, j . For instance, $\bigwedge^2 V$ consists of all linear combinations of tensors (equivalent to one) of the form

$$v \otimes v' - v' \otimes v.$$

A general such element may be written as a sum over permutations, for some vectors v_1, \dots, v_m :

$$\frac{1}{m!} \sum_{\sigma \in S_m} \mathrm{sgn}(\sigma) v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(m)}.$$

This element is denoted

$$v_1 \wedge \dots \wedge v_m.$$

Note that, for instance,

$$(6) \quad v_1 \wedge v_2 \wedge \dots \wedge v_m = -v_2 \wedge v_1 \wedge \dots \wedge v_m.$$

It is more fashionable to regard $\bigwedge^m V$ as a *quotient* space of $V^{\otimes m}$. Here, the subspace of $V^{\otimes m}$ by which we take the quotient consists of linear combinations of tensor multiples of 2-tensors of the form

$$v \otimes v.$$

Note that (if $\text{char}(k) \neq 2$) this subspace equals everything of the form

$$v \otimes v' + v' \otimes v.$$

(Another way to describe this is as starting with the tensor algebra $\otimes V$, considering the two-sided ideal generated by 2-tensors of the preceding form, taking the quotient by this ideal, and taking the degree- m part of what remains.) The equivalence class of $v_1 \otimes \cdots \otimes v_m$ in this quotient is denoted

$$v_1 \wedge \cdots \wedge v_m.$$

The relation (6) holds in this setting as well. These two constructions of $\wedge^m V$ are canonically equivalent, in a sense that can be made precise with a little classical representation theory.

There is no such thing as $V \wedge W$ for different vector spaces V, W ...because what would that mean?

Exercise 19. Prove that if $\dim(V) = n$, then for $m > n$, we have $\wedge^m V = \{0\}$. More generally, compute the dimension of $\wedge^m V$ for all m .

Exercise 20. Prove that $\wedge^m(V^*) \cong (\wedge^m V)^*$, under an isomorphism that does not depend on the choice of basis.

For instance, a skew-symmetric bilinear form as discussed previously is equivalent to an element $B \in \wedge^2(V^*)$.

We can now give the “right” definition of the determinant of an endomorphism $\varphi : V \rightarrow V$, assuming $\dim(V) = n$: the map φ induces an endomorphism $(\wedge^n \varphi) : \wedge^n V \rightarrow \wedge^n V$, and the space $\wedge^n V$ is 1-dimensional...any endomorphism of a 1-dimensional space is just scaling by a constant in k , and that constant is $\det(\varphi)$. (Write this out in a basis to really understand.)

5.2. Symmetric powers of a vector space. Given a vector space V , for each m there is a vector space $S^m V$ (sometimes denotes $\text{Sym}^m V$, or even $\odot^m V$) and a symmetric, multi-linear map

$$\pi : \underbrace{V \times \cdots \times V}_m \rightarrow S^m V$$

having the following universal property:

for every vector space U (over k) and symmetric multi-linear map $V^m \rightarrow U$, there is a unique linear map $\varphi : S^m V \rightarrow U$ through which π factors:

$$V^m \xrightarrow{\pi} S^m V \xrightarrow{\varphi} U.$$

Moreover, *any* vector space and map having the preceding property can be canonically identified, so this is a good definition of the *m th symmetric power* of V (once these statements are proved). The upshot is that one can replace the study of symmetric multi-linear maps by that of linear maps alone.

Again, this definition is incomprehensible. To really work with this, one can proceed in one of two ways, in analogy with those used for exterior powers. The whole thing is left to the reader.

Note that $S^2(V^*)$ may be identified with the vector space of symmetric bilinear forms on V , discussed previously. Over \mathbb{R} , there is an open convex cone in $S^2(V^*)$ corresponding to those symmetric bilinear forms which are also positive-definite. The standard action of $GL(V)$ on V induces actions on all tensor, exterior, and symmetric powers of V , and in particular, it induces an action on $S^2(V^*)$, for which this cone is a single orbit. The stabilizer of a point in this orbit is a subgroup isomorphic to the orthogonal group $O(V) \subset GL(V)$, and therefore the collection of all inner-products on V is isomorphic to the quotient

$$GL(V)/O(V).$$

This identification of one of the most important of all *symmetric spaces* is a reason why this is useful.

Another is that the elements of $S^m(V^*)$ may be thought of as degree- m homogeneous polynomials on V . The full symmetric algebra $SV^* = \oplus_m S^m(V^*)$ is therefore the homogeneous coordinate ring of the projective space $\mathbf{P}V^*$, which is a starting point for algebraic geometry.