Determinants via Exterior Algebra

In the following, V will always be a finite-dimensional vector space over $\mathbb{R}$.

1 Review of Exterior Algebras

We can define the exterior/wedge product on V as the universal alternating bilinear map on $V \times V$. This will let us analyze general alternating bilinear (and later multilinear) maps h through linear maps instead, denoted t below. Then, we'll focus on perhaps the most important alternating multilinear maps: the determinant.

Definition 1. *The* exterior product $\wedge : V \times V \to V \wedge V$ *is a universal alternating bilinear map, where* $(v_1, v_2) \mapsto v_1 \wedge v_2$*. That is, if* $h : V \times V \to W$ *is any alternating bilinear map, then there is a unique linear map* $t : V \wedge V \to W$ *such that the following diagram commutes:*

$$\begin{array}{ccc} V \times V & \xrightarrow{\ \wedge\ } & V \wedge V \\ & \searrow^{h} & \vdots\, t \\ & & W \end{array}$$

Remark 2 (On universal properties; may be skipped)**.** *It may seem strange that we never specified what* $V \wedge V$ *was nor what* $\wedge$ *really did. Indeed, what we've really defined is the* universal property *for alternating bilinear maps rather than the exterior product (then we set the exterior product as the canonical universal alternating bilinear map).*

Let's give the previous definition more formally: we say that a map $u : V \times V \to U$ *is a universal alternating bilinear map (where* U *is a vector space) if:*

(i) u *itself is an alternating bilinear map, and*

(ii) for any alternating bilinear maps $h : V \times V \to W$*, there is a unique linear maps* $t : U \to W$ *such that* $h = t \circ u$*.*

This second condition states that the following diagram commutes:

$$\begin{array}{ccc} V \times V & \xrightarrow{\ u\ } & U \\ & \searrow^{h} & \vdots\, t \\ & & W \end{array}$$

Furthermore, we call the pair (u, U) *a* universal element *with respect to this property. Notice that we say* (u, U) *is* a *universal element because multiple maps could satisfy this universal property. However, it turns out that all universal elements are* uniquely isomorphic*. Here's why:*

Take (u, U) *and* (u', U') *two universal elements. As* u' *is alternating bilinear, by universality of* u*, there is a unique linear map* $t : U \to U'$ *such that* $u' = t \circ u$*. By the same argument, there is also* *a unique linear map $t' : U' \to U$ such that $u = t' \circ u'$. In particular, $t : U \to U'$ is an isomorphism (and unique with respect to this property).*

Now, because all universal elements are uniquely isomorphic, we can define $V \wedge V$ to be the canonical universal element (so rename the vector space U by $V \wedge V$), and instead of writing $u(v_1, v_2)$, we'll write $v_1 \wedge v_2$. Finally, note that it might very well be the case that no universal elements exist (in which case all that we've stated so far is trivially true). Rest assured that a construction of a universal element is possible (briefly, quotient $V \otimes V$ with the subspace generated by $\{v \otimes v : v \in V\}$); however, this construction is perhaps more formal than illuminating.

We can define the universal alternating trilinear map in similar fashion. Or, more generally, the universal alternating p-linear map, which gives rise to the vector space $\Lambda^p(V) = V \wedge \cdots \wedge V$ (just V wedged with itself p times). Thus, $\Lambda^2(V) = V \wedge V$. We'll identify $\Lambda^1(V) := V$ and $\Lambda^0(V) := \mathbb{R}$. We call the sequence:

$$\Lambda(V) := (\Lambda^0(V), \Lambda^1(V), \Lambda^2(V), \cdots)$$

the *exterior algebra* of V. Here are some properties of $\Lambda(V)$:

- $v_1 \wedge v_2 = -v_2 \wedge v_1$, which implies $v \wedge v = 0$. More generally, if $\sigma \in \Sigma_p$ is a permutation, then:

  $$v_1 \wedge \cdots \wedge v_p = (-1)^{\text{sgn}(\sigma)} v_{\sigma_1} \wedge \cdots \wedge v_{\sigma_p}, \tag{1}$$

  and if $v_i = v_j$ for any $i \neq j$, then the product is equal to 0.

- elements of $\Lambda^p(V)$ are called *p-vectors*, spanned by elements of the form $v_1 \wedge \cdots \wedge v_p$ where $v_i \in V$. Such elements $v_1 \wedge \cdots \wedge v_n$ are called *decomposable*. Thus, general elements of $\Lambda^p(V)$ are finite sums of decomposable p-vectors. Not all p-vectors are decomposable.

- if $\{e_1, \ldots, e_n\}$ is a basis of V, then $\Lambda^p(V)$ is a $\binom{n}{p}$-dimensional vector space with basis $e_{k_1} \wedge \cdots \wedge e_{k_p}$ where $k_1 < \cdots < k_p$. Abusing notation, we denote this basis element by $e_{\mathbf{k}}$, where $\mathbf{k} \in \binom{n}{p}$ is a choice of p indexes from $[n]$. If $p > n$, then $\Lambda^p(V) = 0$.

- If $T : V \to V$ is a linear map, then consider the alternating p-linear map $\tau : V^p \to \Lambda^p(V)$, by

  $$(v_1, \ldots, v_p) \overset{\tau}{\mapsto} Tv_1 \wedge \cdots \wedge Tv_p.$$

  By universality, there is a unique linear map $t : \Lambda^p(V) \to \Lambda^p(V)$ such that $\tau = t \circ \wedge^p$. Because this map is uniquely induced by T, let's denote it $\Lambda^p(T)$.

2 Determinants

Let's consider this last point for $p = n$. Notice that $\Lambda^n(V)$ is a 1-dimensional vector space, with basis $e_1 \wedge \cdots \wedge e_n$. Given a map $T : V \to V$ such that $T(e_i) = v_i$, the induced map $\Lambda^n(T)$ is:

$$e_1 \wedge \cdots \wedge e_n \overset{\Lambda^n(T)}{\longmapsto} v_1 \wedge \cdots \wedge v_n,$$

where $v_1 \wedge \cdots \wedge v_n$ expanded in the basis is just $\alpha e_1 \wedge \cdots \wedge e_n$ for some $\alpha \in \mathbb{R}$. So we can identify $\Lambda^n(T)$ with the constant α.

We should interpret the n-vector $e_1 \wedge \cdots \wedge e_n$ as a unit of n-volume element in V. Then, $\Lambda^n(T)$ tells us how much this volume element is scaled by the map T; the value α is the familiar quantity $\det(T)$. We can verify this explicitly:

Proposition 3. *If $T : V \to V$ is a linear transformation of an n-dimensional vector space, then*

$$Tu_1 \wedge \cdots \wedge Tu_n = \det(T) u_1 \wedge \cdots \wedge u_n$$

for all $u_1, \ldots, u_n \in V$.

Proof. It suffices to show this for a basis $\mathbf{e} = \{e_1, \ldots, e_n\}$, in which case, we can consider the matrix of T rel $\mathbf{e}$. Namely, $T(e_j) = \sum e_i T_{ij}$. Then,

$$Te_1 \wedge \cdots \wedge Te_n = \left(\sum_i e_i T_{i1} \right) \wedge \cdots \wedge \left(\sum_i e_i T_{in} \right).$$

Expanding out by multilinearity yields n^n terms, but any terms with a repeated index is 0, so we are left with a sum over all permutations $\sigma \in \Sigma_n$. From Equation 1, we obtain:

$$Te_1 \wedge \cdots \wedge Te_n = \left(\sum_{\sigma \in \Sigma_n} (-1)^{\mathrm{sgn}(\sigma)} T_{\sigma_1 1} \cdots T_{\sigma_n n} \right) e_1 \wedge \cdots \wedge e_n.$$

The term in the parenthesis is the usual expansion of the determinant of T. □

Let's write $v_i = Te_i$. We can give an interpretation of $v_1 \wedge \cdots \wedge v_n$ that is essentially agnostic to T: the *magnitude* of $v_1 \wedge \cdots \wedge v_n$ is the volume of the parallelpiped spanned by the vectors $v_1, \ldots, v_n$ in V. Of course, this statement is technically not well-formed, because $\Lambda^n(V)$ as a vector space has no intrinsic measure of length (and neither does V). But once V is an inner product space (i.e. once we've fixed a basis), then $\Lambda^n(V)$ is also made into an inner product space. Naturally, the inner product is defined so that:

$$|e_1 \wedge \cdots \wedge e_n|^2 = \langle e_1 \wedge \cdots \wedge e_n, e_1 \wedge \cdots \wedge e_n \rangle = 1.$$

This analysis extends to $\Lambda^p(V)$ for all p. That is, $\Lambda^p(V)$ is also made into an inner product space. Decomposable p-vectors $v_1 \wedge \cdots \wedge v_p$ encoding information about (i) the subspace spanned by the vectors $v_1, \ldots, v_p$—encapsulated in the direction of the p-vector itself, and (ii) the p-dimensional volume of the parallelpiped spanned by those vectors, where:

$$\mathrm{vol}(v_1, \ldots, v_p) = |v_1 \wedge \cdots \wedge v_p|.$$

With this interpretation, Cauchy-Binet's formula becomes very simple. Recall that Cauchy-Binet states: given the p vectors $v_1, \ldots, v_p \in V$. Let X be the matrix where the ith column is v_i, so:

$$X = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_p \\ | & & | \end{bmatrix}.$$

Cauchy-Binet states that:

$$\det(X^T X) = \sum_{\mathbf{k} \in \binom{n}{p}} \det(X_{\mathbf{k}}^T X_{\mathbf{k}}),$$

where $X_{\mathbf{k}}$ is the matrix minor by taking all columns and the $k_1, \ldots, k_p$ rows of X.

This formula is more clearly seen as a particular Pythagorean theorem, recalling that $\det(X^T X)$ gives the squared p-dimensional volume of the parallelpiped spanned by its columns. That is, $\det(X^T X) = |v_1 \wedge \cdots \wedge v_p|^2$.

We derive Cauchy-Binet from the following lemma, expanding $v_1 \wedge \cdots \wedge v_p$ in the basis of $\Lambda^p(V)$:

Lemma 4 (Cauchy-Binet)**.** *Let V be an n-dimensional inner product space with orthonormal basis $\mathbf{e}$, while $\mathbf{v}$ is a list of p elements of V, and X the matrix rel $\mathbf{e}$ of this list, in the sense that $v_i = \sum_j X_{ij} e_j$ for $i \in [p]$. Then,*

$$v_1 \wedge \cdots \wedge v_p = \sum_{\mathbf{k} \in \binom{n}{p}} \det(X_{\mathbf{k}}) e_{\mathbf{k}},$$

where $X_{\mathbf{k}}$ is the matrix minor of X by taking all columns and the $k_1, \ldots, k_p$ rows.

Then, the Cauchy-Binet formula falls out from the Pythagorean theorem, and the fact that $\det(X_{\mathbf{k}})^2 = \det(X_{\mathbf{k}}^T X_{\mathbf{k}})$.

Proof. We write the left-hand side out as:

$$v_1 \wedge \cdots \wedge v_p = \left(\sum_{i \in [n]} e_i X_{i1} \right) \wedge \cdots \wedge \left(\sum_{i \in [n]} e_i X_{ip} \right),$$

where we may expand out via multilinearity; for each basis element $e_{\mathbf{k}}$ for $\mathbf{k} \in \binom{n}{p}$, we have the contribution:

$$\left(\sum_{i \in \mathbf{k}} e_i X_{i1} \right) \wedge \cdots \wedge \left(\sum_{i \in \mathbf{k}} e_i X_{in} \right) = \det(X_{\mathbf{k}}) e_{\mathbf{k}},$$

from Proposition 3. Summing over all $\mathbf{k} \in \binom{n}{p}$ yields the desired formula. □

Corollary 5. *Let V, $\mathbf{e}$, $\mathbf{v}$ and X as before. Let $T : V \to V$ be any linear transformation and identify T with its matrix rel $\mathbf{e}$. Then,*

$$Tv_1 \wedge \cdots \wedge Tv_p = \sum_{\mathbf{h} \in \binom{n}{p}} \sum_{\mathbf{k} \in \binom{n}{p}} e_{\mathbf{h}} \det(T_{\mathbf{h} \times \mathbf{k}}) \det(X_{\mathbf{k}}),$$

where $T_{\mathbf{h} \times \mathbf{k}}$ is the matrix minor of T with rows $h_1, \ldots, h_p$ and columns $k_1, \ldots, k_p$.

Proof. We just apply $\Lambda^p(T)$ to $v_1 \wedge \cdots \wedge v_p$. Because $\Lambda^p(T)$ is linear, Lemma 4 gives:

$$\Lambda^p(T)(v_1 \wedge \cdots \wedge v_p) = Tv_1 \wedge \cdots \wedge Tv_p = \sum_{\mathbf{k} \in \binom{n}{p}} \det(X_{\mathbf{k}}) \cdot \Lambda^p(T) e_{\mathbf{k}},$$

where $\Lambda^p(T) e_{\mathbf{k}}$ is, by Lemma 4 again:

$$Te_{k_1} \wedge \cdots \wedge Te_{k_p} = \sum_{\mathbf{h} \in \binom{n}{p}} \det(T_{\mathbf{h} \times \mathbf{k}}) e_{\mathbf{h}}.$$

Substituting this into the first equation gives the desired formula. □

In particular, if T is diagonal, then $\det(T_{\mathbf{h} \times \mathbf{k}}) = 0$ if $\mathbf{h} \neq \mathbf{k}$. Denote the diagonal matrix by $Q = \operatorname{diag}(q_1, \ldots, q_n)$. Then, the the formula simplifies to:

$$Qv_1 \wedge \cdots \wedge Qv_p = \sum_{\mathbf{k} \in \binom{n}{p}} \det(X_{\mathbf{k}}) \det(Q_{\mathbf{k} \times \mathbf{k}}) e_{\mathbf{k}} = \sum_{\mathbf{k} \in \binom{n}{p}} e_{\mathbf{k}} \det(X_{\mathbf{k}}) \prod_{i \in \mathbf{k}} q_i.$$

It follows that:

$$\det(X^T Q^2 X) = \sum_{\mathbf{k} \in \binom{n}{p}} \det(X_{\mathbf{k}}^T X_{\mathbf{k}}) \prod_{i \in \mathbf{k}} q_i^2. \tag{2}$$