Chapter 9

Multilinear Algebra and Determinants

We begin this chapter by investigating bilinear forms and quadratic forms on a vector space. Then we will move on to multilinear forms. We will show that the vector space of alternating n-linear forms has dimension one on a vector space of dimension n. This result will allow us to give a clean basis-free definition of the determinant of an operator.

This approach to the determinant via alternating multilinear forms leads to straightforward proofs of key properties of determinants. For example, we will see that the determinant is multiplicative, meaning that $\det(ST) = (\det S)(\det T)$ for all operators S and T on the same vector space. We will also see that T is invertible if and only if $\det T \neq 0$. Another important result states that the determinant of an operator on a complex vector space equals the product of the eigenvalues of the operator, with each eigenvalue included as many times as its multiplicity.

The chapter concludes with an introduction to tensor products.

---

standing assumptions for this chapter

- **F** denotes **R** or **C**.
- V and W denote finite-dimensional nonzero vector spaces over **F**.

The Mathematical Institute at the University of Göttingen. This building opened in 1930, when Emmy Noether (1882–1935) had already been a research mathematician and faculty member at the university for 15 years (the first eight years without salary). Noether was fired by the Nazi government in 1933. By then Noether and her collaborators had created many of the foundations of modern algebra, including an abstract algebra viewpoint that contributed to the development of linear algebra.

9A Bilinear Forms and Quadratic Forms

Bilinear Forms

A bilinear form on V is a function from $V \times V$ to $\mathbf{F}$ that is linear in each slot separately, meaning that if we hold either slot fixed then we have a linear function in the other slot. Here is the formal definition.

9.1 definition: *bilinear form*

A *bilinear form* on V is a function $\beta \colon V \times V \to \mathbf{F}$ such that

$$v \mapsto \beta(v, u) \quad \text{and} \quad v \mapsto \beta(u, v)$$

are both linear functionals on V for every $u \in V$.

For example, if V is a real inner product space, then the function that takes an ordered pair $(u, v) \in V \times V$ to $\langle u, v \rangle$ is a bilinear form on V. If V is a nonzero complex inner product space, then this function is not a bilinear form because the inner product is not linear in the second slot (complex scalars come out of the second slot as their complex conjugates).

*Recall that the term **linear functional**, used in the definition above, means a linear function that maps into the scalar field $\mathbf{F}$. Thus the term **bilinear functional** would be more consistent terminology than **bilinear form**, which unfortunately has become standard.*

If $\mathbf{F} = \mathbf{R}$, then a bilinear form differs from an inner product in that an inner product requires symmetry [meaning that $\beta(v, w) = \beta(w, v)$ for all $v, w \in V$] and positive definiteness [meaning that $\beta(v, v) > 0$ for all $v \in V \backslash \{0\}$], but these properties are not required for a bilinear form.

9.2 example: *bilinear forms*

- The function $\beta \colon \mathbf{F}^3 \times \mathbf{F}^3 \to \mathbf{F}$ defined by

$$\beta\big((x_1, x_2, x_3), (y_1, y_2, y_3)\big) = x_1 y_2 - 5 x_2 y_3 + 2 x_3 y_1$$

is a bilinear form on $\mathbf{F}^3$.

- Suppose A is an n-by-n matrix with $A_{j,k} \in \mathbf{F}$ in row j, column k. Define a bilinear form β_A on $\mathbf{F}^n$ by

$$\beta_A\big((x_1, ..., x_n), (y_1, ..., y_n)\big) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j y_k.$$

The first bullet point is a special case of this bullet point with $n = 3$ and

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -5 \\ 2 & 0 & 0 \end{pmatrix}.$$

- Suppose V is a real inner product space and $T \in \mathcal{L}(V)$. Then the function $\beta \colon V \times V \to \mathbf{R}$ defined by

$$\beta(u, v) = \langle u, Tv \rangle$$

   is a bilinear form on V.

- If n is a positive integer, then the function $\beta \colon \mathcal{P}_n(\mathbf{R}) \times \mathcal{P}_n(\mathbf{R}) \to \mathbf{R}$ defined by

$$\beta(p, q) = p(2) \cdot q'(3)$$

   is a bilinear form on $\mathcal{P}_n(\mathbf{R})$.

- Suppose $\varphi, \tau \in V'$. Then the function $\beta \colon V \times V \to \mathbf{F}$ defined by

$$\beta(u, v) = \varphi(u) \cdot \tau(v)$$

   is a bilinear form on V.

- More generally, suppose that $\varphi_1, ..., \varphi_n, \tau_1, ..., \tau_n \in V'$. Then the function $\beta \colon V \times V \to \mathbf{F}$ defined by

$$\beta(u, v) = \varphi_1(u) \cdot \tau_1(v) + \cdots + \varphi_n(u) \cdot \tau_n(v)$$

   is a bilinear form on V.

A bilinear form on V is a function from $V \times V$ to $\mathbf{F}$. Because $V \times V$ is a vector space, this raises the question of whether a bilinear form can also be a linear map from $V \times V$ to $\mathbf{F}$. Note that none of the bilinear forms in 9.2 are linear maps except in some special cases in which the bilinear form is the zero map. Exercise 3 shows that a bilinear form β on V is a linear map on $V \times V$ only if $\beta = 0$.

> **9.3** definition: $V^{(2)}$
>
> The set of bilinear forms on V is denoted by $V^{(2)}$.

With the usual operations of addition and scalar multiplication of functions, $V^{(2)}$ is a vector space.

For T an operator on an n-dimensional vector space V and a basis $e_1, ..., e_n$ of V, we used an n-by-n matrix to provide information about T. We now do the same thing for bilinear forms on V.

> **9.4** definition: *matrix of a bilinear form, $\mathcal{M}(\beta)$*
>
> Suppose β is a bilinear form on V and $e_1, ..., e_n$ is a basis of V. The *matrix* of β with respect to this basis is the n-by-n matrix $\mathcal{M}(\beta)$ whose entry $\mathcal{M}(\beta)_{j,k}$ in row j, column k is given by
>
> $$\mathcal{M}(\beta)_{j,k} = \beta(e_j, e_k).$$
>
> If the basis $e_1, ..., e_n$ is not clear from the context, then the notation $\mathcal{M}(\beta, (e_1, ..., e_n))$ is used.

Recall that $\mathbf{F}^{n,n}$ denotes the vector space of n-by-n matrices with entries in $\mathbf{F}$ and that $\dim \mathbf{F}^{n,n} = n^2$ (see 3.39 and 3.40).

9.5 $\dim V^{(2)} = (\dim V)^2$

Suppose $e_1, ..., e_n$ is a basis of V. Then the map $\beta \mapsto \mathcal{M}(\beta)$ is an isomorphism of $V^{(2)}$ onto $\mathbf{F}^{n,n}$. Furthermore, $\dim V^{(2)} = (\dim V)^2$.

Proof   The map $\beta \mapsto \mathcal{M}(\beta)$ is clearly a linear map of $V^{(2)}$ into $\mathbf{F}^{n,n}$.

For $A \in \mathbf{F}^{n,n}$, define a bilinear form β_A on V by

$$\beta_A(x_1 e_1 + \cdots + x_n e_n, y_1 e_1 + \cdots + y_n e_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j y_k$$

for $x_1, ..., x_n, y_1, ..., y_n \in \mathbf{F}$ (if $V = \mathbf{F}^n$ and $e_1, ..., e_n$ is the standard basis of $\mathbf{F}^n$, this β_A is the same as the bilinear form β_A in the second bullet point of Example 9.2).

The linear map $\beta \mapsto \mathcal{M}(\beta)$ from $V^{(2)}$ to $\mathbf{F}^{n,n}$ and the linear map $A \mapsto \beta_A$ from $\mathbf{F}^{n,n}$ to $V^{(2)}$ are inverses of each other because $\beta_{\mathcal{M}(\beta)} = \beta$ for all $\beta \in V^{(2)}$ and $\mathcal{M}(\beta_A) = A$ for all $A \in \mathbf{F}^{n,n}$, as you should verify.

Thus both maps are isomorphisms and the two spaces that they connect have the same dimension. Hence $\dim V^{(2)} = \dim \mathbf{F}^{n,n} = n^2 = (\dim V)^2$.   ∎

Recall that C^{t} denotes the transpose of a matrix C. The matrix C^{t} is obtained by interchanging the rows and the columns of C.

9.6   *composition of a bilinear form and an operator*

Suppose β is a bilinear form on V and $T \in \mathcal{L}(V)$. Define bilinear forms α and ρ on V by

$$\alpha(u, v) = \beta(u, Tv) \quad \text{and} \quad \rho(u, v) = \beta(Tu, v).$$

Let $e_1, ..., e_n$ be a basis of V. Then

$$\mathcal{M}(\alpha) = \mathcal{M}(\beta)\mathcal{M}(T) \quad \text{and} \quad \mathcal{M}(\rho) = \mathcal{M}(T)^{\mathrm{t}}\mathcal{M}(\beta).$$

Proof   If $j, k \in \{1, ..., n\}$, then

$$\begin{aligned}
\mathcal{M}(\alpha)_{j,k} &= \alpha(e_j, e_k) \\
&= \beta(e_j, Te_k) \\
&= \beta\left(e_j, \sum_{m=1}^{n} \mathcal{M}(T)_{m,k} e_m\right) \\
&= \sum_{m=1}^{n} \beta(e_j, e_m)\mathcal{M}(T)_{m,k} \\
&= (\mathcal{M}(\beta)\mathcal{M}(T))_{j,k}.
\end{aligned}$$

Thus $\mathcal{M}(\alpha) = \mathcal{M}(\beta)\mathcal{M}(T)$. The proof that $\mathcal{M}(\rho) = \mathcal{M}(T)^{\mathrm{t}}\mathcal{M}(\beta)$ is similar.   ∎

The result below shows how the matrix of a bilinear form changes if we change the basis. The formula in the result below should be compared to the change-of-basis formula for the matrix of an operator (see 3.84). The two formulas are similar, except that the transpose C^t appears in the formula below and the inverse C^{-1} appears in the change-of-basis formula for the matrix of an operator.

9.7 *change-of-basis formula*

Suppose $\beta \in V^{(2)}$. Suppose $e_1, ..., e_n$ and $f_1, ..., f_n$ are bases of V. Let

$$A = \mathcal{M}(\beta, (e_1, ..., e_n)) \quad \text{and} \quad B = \mathcal{M}(\beta, (f_1, ..., f_n))$$

and $C = \mathcal{M}(I, (e_1, ..., e_n), (f_1, ..., f_n))$. Then

$$A = C^t B C.$$

Proof The linear map lemma (3.4) tells us that there exists an operator $T \in \mathcal{L}(V)$ such that $T f_k = e_k$ for each $k = 1, ..., n$. The definition of the matrix of an operator with respect to a basis implies that

$$\mathcal{M}(T, (f_1, ..., f_n)) = C.$$

Define bilinear forms α, ρ on V by

$$\alpha(u, v) = \beta(u, Tv) \quad \text{and} \quad \rho(u, v) = \alpha(Tu, v) = \beta(Tu, Tv).$$

Then $\beta(e_j, e_k) = \beta(T f_j, T f_k) = \rho(f_j, f_k)$ for all $j, k \in \{1, ..., n\}$. Thus

$$\begin{aligned}
A &= \mathcal{M}(\rho, (f_1, ..., f_n)) \\
&= C^t \mathcal{M}(\alpha, f_1, ..., f_n)) \\
&= C^t B C,
\end{aligned}$$

where the second and third lines each follow from 9.6. ∎

9.8 example: *the matrix of a bilinear form on $\mathcal{P}_2(\mathbf{R})$*

Define a bilinear form β on $\mathcal{P}_2(\mathbf{R})$ by $\beta(p, q) = p(2) \cdot q'(3)$. Let

$$A = \mathcal{M}(\beta, (1, x - 2, (x - 3)^2)) \quad \text{and} \quad B = \mathcal{M}(\beta, (1, x, x^2))$$

and $C = \mathcal{M}(I, (1, x - 2, (x - 3)^2), (1, x, x^2))$. Then

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 & 6 \\ 0 & 2 & 12 \\ 0 & 4 & 24 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 1 & -2 & 9 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{pmatrix}.$$

Now the change-of-basis formula 9.7 asserts that $A = C^t B C$, which you can verify with matrix multiplication using the matrices above.

Symmetric Bilinear Forms

9.9   definition: *symmetric bilinear form, $V_{\text{sym}}^{(2)}$*

A bilinear form $\rho \in V^{(2)}$ is called *symmetric* if

$$\rho(u, w) = \rho(w, u)$$

for all $u, w \in V$. The set of symmetric bilinear forms on V is denoted by $V_{\text{sym}}^{(2)}$.

9.10   example: *symmetric bilinear forms*

- If V is a real inner product space and $\rho \in V^{(2)}$ is defined by

$$\rho(u, w) = \langle u, w \rangle,$$

  then ρ is a symmetric bilinear form on V.

- Suppose V is a real inner product space and $T \in \mathcal{L}(V)$. Define $\rho \in V^{(2)}$ by

$$\rho(u, w) = \langle u, Tw \rangle.$$

  Then ρ is a symmetric bilinear form on V if and only if T is a self-adjoint operator (the previous bullet point is the special case $T = I$).

- Suppose $\rho \colon \mathcal{L}(V) \times \mathcal{L}(V) \to \mathbf{F}$ is defined by

$$\rho(S, T) = \operatorname{tr}(ST).$$

  Then ρ is a symmetric bilinear form on $\mathcal{L}(V)$ because trace is a linear functional on $\mathcal{L}(V)$ and $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ for all $S, T \in \mathcal{L}(V)$; see 8.56.

9.11   definition: *symmetric matrix*

A square matrix A is called *symmetric* if it equals its transpose.

An operator on V may have a symmetric matrix with respect to some but not all bases of V. In contrast, the next result shows that a bilinear form on V has a symmetric matrix with respect to either all bases of V or with respect to no bases of V.

9.12   *symmetric bilinear forms are diagonalizable*

Suppose $\rho \in V^{(2)}$. Then the following are equivalent.
(a) ρ is a symmetric bilinear form on V.

(b) $\mathcal{M}(\rho, (e_1, ..., e_n))$ is a symmetric matrix for every basis $e_1, ..., e_n$ of V.

(c) $\mathcal{M}(\rho, (e_1, ..., e_n))$ is a symmetric matrix for some basis $e_1, ..., e_n$ of V.

(d) $\mathcal{M}(\rho, (e_1, ..., e_n))$ is a diagonal matrix for some basis $e_1, ..., e_n$ of V.

Proof   First suppose (a) holds, so ρ is a symmetric bilinear form. Suppose $e_1, ..., e_n$ is a basis of V and $j, k \in \{1, ..., n\}$. Then $\rho(e_j, e_k) = \rho(e_k, e_j)$ because ρ is symmetric. Thus $\mathcal{M}(\rho, (e_1, ..., e_n))$ is a symmetric matrix, showing that (a) implies (b).

Clearly (b) implies (c).

Now suppose (c) holds and $e_1, ..., e_n$ is a basis of V such that $\mathcal{M}(\rho, (e_1, ..., e_n))$ is a symmetric matrix. Suppose $u, w \in V$. There exist $a_1, ..., a_n, b_1, ..., b_n \in \mathbf{F}$ such that $u = a_1 e_1 + \cdots + a_n e_n$ and $w = b_1 e_1 + \cdots + b_n e_n$. Now

$$\begin{aligned}
\rho(u, w) &= \rho\left(\sum_{j=1}^{n} a_j e_j, \sum_{k=1}^{n} b_k e_k \right) \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k \rho(e_j, e_k) \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k \rho(e_k, e_j) \\
&= \rho\left(\sum_{k=1}^{n} b_k e_k, \sum_{j=1}^{n} a_j e_j \right) \\
&= \rho(w, u),
\end{aligned}$$

where the third line holds because $\mathcal{M}(\rho)$ is a symmetric matrix. The equation above shows that ρ is a symmetric bilinear form, proving that (c) implies (a).

At this point, we have proved that (a), (b), (c) are equivalent. Because every diagonal matrix is symmetric, (d) implies (c). To complete the proof, we will show that (a) implies (d) by induction on $n = \dim V$.

If $n = 1$, then (a) implies (d) because every 1-by-1 matrix is diagonal. Now suppose $n > 1$ and the implication (a) $\implies$ (d) holds for one less dimension. Suppose (a) holds, so ρ is a symmetric bilinear form. If $\rho = 0$, then the matrix of ρ with respect to every basis of V is the zero matrix, which is a diagonal matrix. Hence we can assume that $\rho \neq 0$, which means there exist $u, w \in V$ such that $\rho(u, w) \neq 0$. Now

$$2\rho(u, w) = \rho(u + w, u + w) - \rho(u, u) - \rho(w, w).$$

Because the left side of the equation above is nonzero, the three terms on the right cannot all equal 0. Hence there exists $v \in V$ such that $\rho(v, v) \neq 0$.

Let $U = \{u \in V : \rho(u, v) = 0\}$. Thus U is the null space of the linear functional $u \mapsto \rho(u, v)$ on V. This linear functional is not the zero linear functional because $v \notin U$. Thus $\dim U = n - 1$. By our induction hypothesis, there is a basis $e_1, ..., e_{n-1}$ of U such that the symmetric bilinear form $\rho|_{U \times U}$ has a diagonal matrix with respect to this basis.

Because $v \notin U$, the list $e_1, ..., e_{n-1}, v$ is a basis of V. Suppose $k \in \{1, ..., n-1\}$. Then $\rho(e_k, v) = 0$ by the construction of U. Because ρ is symmetric, we also have $\rho(v, e_k) = 0$. Thus the matrix of ρ with respect to $e_1, ..., e_{n-1}, v$ is a diagonal matrix, completing the proof that (a) implies (d). ∎

The previous result states that every symmetric bilinear form has a diagonal matrix with respect to some basis. If our vector space happens to be a real inner product space, then the next result shows that every symmetric bilinear form has a diagonal matrix with respect to some *orthonormal* basis. Note that the inner product here is unrelated to the bilinear form.

9.13   *diagonalization of a symmetric bilinear form by an orthonormal basis*

Suppose V is a real inner product space and ρ is a symmetric bilinear form on V. Then ρ has a diagonal matrix with respect to some orthonormal basis of V.

Proof   Let $f_1, ..., f_n$ be an orthonormal basis of V. Let $B = \mathcal{M}(\rho, (f_1, ..., f_n))$. Then B is a symmetric matrix (by 9.12). Let $T \in \mathcal{L}(V)$ be the operator such that $\mathcal{M}(T, (f_1, ..., f_n)) = B$. Thus T is self-adjoint.

The real spectral theorem (7.29) states that T has a diagonal matrix with respect to some orthonormal basis $e_1, ..., e_n$ of V. Let $C = \mathcal{M}(I, (e_1, ..., e_n), (f_1, ..., f_n))$. Thus $C^{-1}TC$ is the matrix of T with respect to the basis $e_1, ..., e_n$ (by 3.84). Hence $C^{-1}TC$ is a diagonal matrix. Now

$$M(\rho, (e_1, ..., e_n)) = C^{\mathrm{t}}TC = C^{-1}TC,$$

where the first equality holds by 9.7 and the second equality holds because C is a unitary matrix with real entries (which implies that $C^{-1} = C^{\mathrm{t}}$; see 7.57). ∎

Now we turn our attention to alternating bilinear forms. Alternating multilinear forms will play a major role in our approach to determinants later in this chapter.

9.14   definition: *alternating bilinear form, $V_{\mathrm{alt}}^{(2)}$*

A bilinear form $\alpha \in V^{(2)}$ is called *alternating* if

$$\alpha(v, v) = 0$$

for all $v \in V$. The set of alternating bilinear forms on V is denoted by $V_{\mathrm{alt}}^{(2)}$.

9.15   example: *alternating bilinear forms*

- Suppose $n \geq 3$ and $\alpha \colon \mathbf{F}^n \times \mathbf{F}^n \to \mathbf{F}$ is defined by

$$\alpha((x_1, ..., x_n), (y_1, ..., y_n)) = x_1 y_2 - x_2 y_1 + x_1 y_3 - x_3 y_1.$$

  Then α is an alternating bilinear form on $\mathbf{F}^n$.

- Suppose $\varphi, \tau \in V'$. Then the bilinear form α on V defined by

$$\alpha(u, w) = \varphi(u)\tau(w) - \varphi(w)\tau(u)$$

  is alternating.

The next result shows that a bilinear form is alternating if and only if switching the order of the two inputs multiplies the output by -1.

> **9.16** *characterization of alternating bilinear forms*
>
> A bilinear form α on V is alternating if and only if
>
> $$\alpha(u, w) = -\alpha(w, u)$$
>
> for all $u, w \in V$.

Proof   First suppose that α is alternating. If $u, w \in V$, then

$$
\begin{aligned}
0 &= \alpha(u + w, u + w) \\
&= \alpha(u, u) + \alpha(u, w) + \alpha(w, u) + \alpha(w, w) \\
&= \alpha(u, w) + \alpha(w, u).
\end{aligned}
$$

Thus $\alpha(u, w) = -\alpha(w, u)$, as desired.

To prove the implication in the other direction, suppose $\alpha(u, w) = -\alpha(w, u)$ for all $u, w \in V$. Then $\alpha(v, v) = -\alpha(v, v)$ for all $v \in V$, which implies that $\alpha(v, v) = 0$ for all $v \in V$. Thus α is alternating. ∎

Now we show that the vector space of bilinear forms on V is the direct sum of the symmetric bilinear forms on V and the alternating bilinear forms on V.

> **9.17** $V^{(2)} = V_{\text{sym}}^{(2)} \oplus V_{\text{alt}}^{(2)}$
>
> The sets $V_{\text{sym}}^{(2)}$ and $V_{\text{alt}}^{(2)}$ are subspaces of $V^{(2)}$. Furthermore,
>
> $$V^{(2)} = V_{\text{sym}}^{(2)} \oplus V_{\text{alt}}^{(2)}.$$

Proof   The definition of symmetric bilinear form implies that the sum of any two symmetric bilinear forms on V is a bilinear form on V, and any scalar multiple of any bilinear form on V is a bilinear form on V. Thus $V_{\text{sym}}^{(2)}$ is a subspace of $V^{(2)}$. Similarly, the verification that $V_{\text{alt}}^{(2)}$ is a subspace of $V^{(2)}$ is straightforward.

Next, we want to show that $V^{(2)} = V_{\text{sym}}^{(2)} + V_{\text{alt}}^{(2)}$. To do this, suppose $\beta \in V^{(2)}$. Define $\rho, \alpha \in V^{(2)}$ by

$$\rho(u, w) = \frac{\beta(u, w) + \beta(w, u)}{2} \quad \text{and} \quad \alpha(u, w) = \frac{\beta(u, w) - \beta(w, u)}{2}.$$

Then $\rho \in V_{\text{sym}}^{(2)}$ and $\alpha \in V_{\text{alt}}^{(2)}$, and $\beta = \rho + \alpha$. Thus $V^{(2)} = V_{\text{sym}}^{(2)} + V_{\text{alt}}^{(2)}$.

Finally, to show that the intersection of the two subspaces under consideration equals $\{0\}$, suppose $\beta \in V_{\text{sym}}^{(2)} \cap V_{\text{alt}}^{(2)}$. Then 9.16 implies that

$$\beta(u, w) = -\beta(w, u) = -\beta(u, w)$$

for all $u, w \in V$, which implies that $\beta = 0$. Thus $V^{(2)} = V_{\text{sym}}^{(2)} \oplus V_{\text{alt}}^{(2)}$, as implied by 1.46. ∎

Quadratic Forms

> **9.18** definition: *quadratic form associated with a bilinear form, q_β*
>
> For β a bilinear form on V, define a function $q_\beta \colon V \to \mathbf{F}$ by $q_\beta(v) = \beta(v, v)$. A function $q \colon V \to \mathbf{F}$ is called a *quadratic form* on V if there exists a bilinear form β on V such that $q = q_\beta$.

Note that if β is a bilinear form, then $q_\beta = 0$ if and only if β is alternating.

> **9.19** example: *quadratic form*
>
> Suppose β is the bilinear form on $\mathbf{R}^3$ defined by
>
> $$\beta((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1 y_1 - 4x_1 y_2 + 8x_1 y_3 - 3x_3 y_3.$$
>
> Then q_β is the quadratic form on $\mathbf{R}^3$ given by the formula
>
> $$q_\beta(x_1, x_2, x_3) = x_1^2 - 4x_1 x_2 + 8x_1 x_3 - 3x_3^2.$$

The quadratic form in the example above is typical of quadratic forms on $\mathbf{F}^n$, as shown in the next result.

> **9.20** *quadratic forms on $\mathbf{F}^n$*
>
> Suppose n is a positive integer and q is a function from $\mathbf{F}^n$ to $\mathbf{F}$. Then q is a quadratic form on $\mathbf{F}^n$ if and only if there exist numbers $A_{j,k} \in \mathbf{F}$ for $j, k \in \{1, \dots, n\}$ such that
>
> $$q(x_1, \dots, x_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k$$
>
> for all $(x_1, \dots, x_n) \in \mathbf{F}^n$.

Proof First suppose q is a quadratic form on $\mathbf{F}^n$. Thus there exists a bilinear form β on $\mathbf{F}^n$ such that $q = q_\beta$. Let A be the matrix of β with respect to the standard basis of $\mathbf{F}^n$. Then for all $(x_1, \dots, x_n) \in \mathbf{F}^n$, we have the desired equation

$$q(x_1, \dots, x_n) = \beta((x_1, \dots, x_n), (x_1, \dots, x_n)) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k.$$

Conversely, suppose there exist numbers $A_{j,k} \in \mathbf{F}$ for $j, k \in \{1, \dots, n\}$ such that

$$q(x_1, \dots, x_n) = \sum_{k=1}^{n} \sum_{j=1}^{k} A_{j,k} x_j x_k$$

for all $(x_1, \dots, x_n) \in \mathbf{F}^n$. Define a bilinear form β on $\mathbf{F}^n$ by

$$\beta((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sum_{k=1}^{n} \sum_{j=1}^{k} A_{j,k} x_j y_k.$$

Then $q = q_\beta$, as desired. ∎

Although quadratic forms are defined in terms of an arbitrary bilinear form, the equivalence of (a) and (b) in the result below shows that a *symmetric* bilinear form can always be used.

9.21 *characterization of quadratic forms*

Suppose $q \colon V \to \mathbf{F}$ is a function. The following are equivalent.

(a) q is a quadratic form.

(b) There exists a unique symmetric bilinear form ρ on V such that $q = q_\rho$.

(c) $q(\lambda v) = \lambda^2 q(v)$ for all $\lambda \in \mathbf{F}$ and all $v \in V$, and the function

$$(u, w) \mapsto q(u + w) - q(u) - q(w)$$

is a symmetric bilinear form on V.

(d) $q(2v) = 4q(v)$ for all $v \in V$, and the function

$$(u, w) \mapsto q(u + w) - q(u) - q(w)$$

is a symmetric bilinear form on V.

Proof First suppose (a) holds, so q is a quadratic form. Hence there exists a bilinear form β such that $q = q_\beta$. By 9.17, there exist a symmetric bilinear form ρ on V and an alternating bilinear form α on V such that $\beta = \rho + \alpha$. Now

$$q = q_\beta = q_\rho + q_\alpha = q_\rho.$$

If $\rho' \in V_{\text{sym}}^{(2)}$ also satisfies $q_{\rho'} = q$, then $q_{\rho' - \rho} = 0$; thus $\rho' - \rho \in V_{\text{sym}}^{(2)} \cap V_{\text{alt}}^{(2)}$, which implies that $\rho' = \rho$ (by 9.17). This completes the proof that (a) implies (b).

Now suppose (b) holds, so there exists a symmetric bilinear form ρ on V such that $q = q_\rho$. If $\lambda \in \mathbf{F}$ and $v \in V$ then

$$q(\lambda v) = \rho(\lambda v, \lambda v) = \lambda \rho(v, \lambda v) = \lambda^2 \rho(v, v) = \lambda^2 q(v),$$

showing that the first part of (c) holds.

If $u, w \in V$, then

$$q(u + w) - q(u) - q(w) = \rho(u + w, u + w) - \rho(u, u) - \rho(w, w) = 2\rho(u, w).$$

Thus the function $(u, w) \mapsto q(u+w) - q(u) - q(w)$ equals 2ρ, which is a symmetric bilinear form on V, completing the proof that (b) implies (c).

Clearly (c) implies (d).

Now suppose (d) holds. Let ρ be the symmetric bilinear form on V defined by

$$\rho(u, w) = \frac{q(u + w) - q(u) - q(w)}{2}.$$

If $v \in V$, then

$$\rho(v, v) = \frac{q(2v) - q(v) - q(v)}{2} = \frac{4q(v) - 2q(v)}{2} = q(v).$$

Thus $q = q_\rho$, completing the proof that (d) implies (a). ∎

9.22   example: *symmetric bilinear form associated with a quadratic form*

Suppose q is the quadratic form on $\mathbf{R}^3$ given by the formula

$$q(x_1, x_2, x_3) = x_1^2 - 4x_1x_2 + 8x_1x_3 - 3x_3^2.$$

A bilinear form β on $\mathbf{R}^3$ such that $q = q_\beta$ is given by Example 9.19, but this bilinear form is not symmetric, as promised by 9.21(b). However, the bilinear form ρ on $\mathbf{R}^3$ defined by

$$\rho((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 4x_1y_3 + 4x_3y_1 - 3x_3y_3$$

is symmetric and satisfies $q = q_\rho$.

The next result states that for each quadratic form we can choose a basis such that the quadratic form looks like a weighted sum of squares of the coordinates, meaning that there are no cross terms of the form x_jx_k with $j \neq k$.

> **9.23**   *diagonalization of quadratic form*
>
> Suppose q is a quadratic form on V.
>
> (a) There exist a basis $e_1, ..., e_n$ of V and $\lambda_1, ..., \lambda_n \in \mathbf{F}$ such that
>
> $$q(x_1e_1 + \cdots + x_ne_n) = \lambda_1x_1^2 + \cdots + \lambda_nx_n^2$$
>
> for all $x_1, ..., x_n \in \mathbf{F}$.
>
> (b) If $\mathbf{F} = \mathbf{R}$ and V is an inner product space, then the basis in (a) can be chosen to be an orthonormal basis of V.

Proof

(a) There exists a symmetric bilinear form ρ on V such that $q = q_\rho$ (by 9.21). Now there exists a basis $e_1, ..., e_n$ of V such that $\mathcal{M}(\rho, (e_1, ..., e_n))$ is a diagonal matrix (by 9.12). Let $\lambda_1, ..., \lambda_n$ denote the entries on the diagonal of this matrix. Thus

$$\rho(e_j, e_k) = \begin{cases} \lambda_j & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases}$$

for all $j, k \in \{1, ..., n\}$. If $x_1, ..., x_n \in \mathbf{F}$, then

$$\begin{aligned} q(x_1e_1 + \cdots + x_ne_n) &= \rho(x_1e_1 + \cdots + x_ne_n, x_1e_1 + \cdots + x_ne_n) \\ &= \sum_{k=1}^n \sum_{j=1}^n x_jx_k\rho(e_j, e_k) \\ &= \lambda_1x_1^2 + \cdots + \lambda_nx_n^2, \end{aligned}$$

as desired.

(b) Suppose $\mathbf{F} = \mathbf{R}$ and V is an inner product space. Then 9.13 tells us that the basis in (a) can be chosen to be an orthonormal basis of V. ∎

Exercises 9A

1 Prove that if β is a bilinear form on $\mathbf{F}$, then there exists $c \in \mathbf{F}$ such that

$$\beta(x, y) = cxy$$

for all $x, y \in \mathbf{F}$.

2 Let $n = \dim V$. Suppose β is a bilinear form on V. Prove that there exist $\varphi_1, ..., \varphi_n, \tau_1, ..., \tau_n \in V'$ such that

$$\beta(u, v) = \varphi_1(u) \cdot \tau_1(v) + \cdots + \varphi_n(u) \cdot \tau_n(v)$$

for all $u, v \in V$.

> *This exercise shows that if $n = \dim V$, then every bilinear form on V is of the form given by the last bullet point of Example 9.2.*

3 Suppose $\beta \colon V \times V \to \mathbf{F}$ is a bilinear form on V and also is a linear functional on $V \times V$. Prove that $\beta = 0$.

4 Suppose V is a real inner product space and β is a bilinear form on V. Show that there exists a unique operator $T \in \mathcal{L}(V)$ such that

$$\beta(u, v) = \langle u, Tv \rangle$$

for all $u, v \in V$.

> *This exercise states that if V is a real inner product space, then every bilinear form on V is of the form given by the third bullet point in 9.2.*

5 Suppose β is a bilinear form on a real inner product space V and T is the unique operator on V such that $\beta(u, v) = \langle u, Tv \rangle$ for all $u, v \in V$ (see Exercise 4). Show that β is an inner product on V if and only if T is an invertible positive operator on V.

6 Prove or give a counterexample: If ρ is a symmetric bilinear form on V, then

$$\{v \in V : \rho(v, v) = 0\}$$

is a subspace of V.

7 Explain why the proof of 9.13 (diagonalization of a symmetric bilinear form by an orthonormal basis on a real inner product space) fails if the hypothesis that $\mathbf{F} = \mathbf{R}$ is dropped.

8 Find formulas for $\dim V_{\text{sym}}^{(2)}$ and $\dim V_{\text{alt}}^{(2)}$ in terms of $\dim V$.

9 Suppose that n is a positive integer and $V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1)\}$. Define $\alpha \colon V \times V \to \mathbf{R}$ by

$$\alpha(p, q) = \int_0^1 pq'.$$

Show that α is an alternating bilinear form on V.
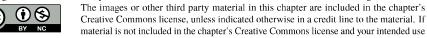
10   Suppose that n is a positive integer and

$$V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1) \text{ and } p'(0) = p'(1)\}.$$

Define $\rho \colon V \times V \to \mathbf{R}$ by

$$\rho(p, q) = \int_0^1 pq''.$$

Show that ρ is a symmetric bilinear form on V.

9B *Alternating Multilinear Forms*

Multilinear Forms

9.24 definition: V^m

For m a positive integer, define V^m by

$$V^m = \underbrace{V \times \cdots \times V}_{m \text{ times}}.$$

Now we can define m-linear forms as a generalization of the bilinear forms that we discussed in the previous section.

9.25 definition: *m-linear form, $V^{(m)}$, multilinear form*

- For m a positive integer, an *m-linear form* on V is a function $\beta \colon V^m \to \mathbf{F}$ that is linear in each slot when the other slots are held fixed. This means that for each $k \in \{1, ..., m\}$ and all $u_1, ..., u_m \in V$, the function

$$v \mapsto \beta(u_1, ..., u_{k-1}, v, u_{k+1}, ..., u_m)$$

  is a linear map from V to $\mathbf{F}$.

- The set of m-linear forms on V is denoted by $V^{(m)}$.

- A function β is called a *multilinear form* on V if it is an m-linear form on V for some positive integer m.

In the definition above, the expression $\beta(u_1, ..., u_{k-1}, v, u_{k+1}, ..., u_m)$ means $\beta(v, u_2, ..., u_m)$ if $k = 1$ and means $\beta(u_1, ..., u_{m-1}, v)$ if $k = m$.

A 1-linear form on V is a linear functional on V. A 2-linear form on V is a bilinear form on V. You can verify that with the usual addition and scalar multiplication of functions, $V^{(m)}$ is a vector space.

9.26 example: *m-linear forms*

- Suppose $\alpha, \rho \in V^{(2)}$. Define a function $\beta \colon V^4 \to \mathbf{F}$ by

$$\beta(v_1, v_2, v_3, v_4) = \alpha(v_1, v_2)\, \rho(v_3, v_4).$$

  Then $\beta \in V^{(4)}$.

- Define $\beta \colon (\mathcal{L}(V))^m \to \mathbf{F}$ by

$$\beta(T_1, ..., T_m) = \operatorname{tr}(T_1 \cdots T_m).$$

  Then β is an m-linear form on $\mathcal{L}(V)$.

Alternating multilinear forms, which we now define, play an important role as we head toward defining determinants.

9.27   definition: *alternating forms,* $V_{\text{alt}}^{(m)}$

Suppose m is a positive integer.

- An m-linear form α on V is called *alternating* if $\alpha(v_1, ..., v_m) = 0$ whenever $v_1, ..., v_m$ is a list of vectors in V with $v_j = v_k$ for some two distinct values of j and k in $\{1, ..., m\}$.
- $V_{\text{alt}}^{(m)} = \{\alpha \in V^{(m)} : \alpha \text{ is an alternating } m\text{-linear form on } V\}$.

You should verify that $V_{\text{alt}}^{(m)}$ is a subspace of $V^{(m)}$. See Example 9.15 for examples of alternating 2-linear forms. See Exercise 2 for an example of an alternating 3-linear form.

The next result tells us that if a linearly dependent list is input to an alternating multilinear form, then the output equals 0.

9.28   *alternating multilinear forms and linear dependence*

Suppose m is a positive integer and α is an alternating m-linear form on V. If $v_1, ..., v_m$ is a linearly dependent list in V, then

$$\alpha(v_1, ..., v_m) = 0.$$

Proof   Suppose $v_1, ..., v_m$ is a linearly dependent list in V. By the linear dependence lemma (2.19), some v_k is a linear combination of $v_1, ..., v_{k-1}$. Thus there exist $b_1, ..., b_{k-1}$ such that $v_k = b_1 v_1 + \cdots + b_{k-1} v_{k-1}$. Now

$$\alpha(v_1, ..., v_m) = \alpha\left(v_1, ..., v_{k-1}, \sum_{j=1}^{k-1} b_j v_j, v_{k+1}, ..., v_m\right)$$

$$= \sum_{j=1}^{k-1} b_j\, \alpha(v_1, ..., v_{k-1}, v_j, v_{k+1}, ..., v_m)$$

$$= 0. \qquad \blacksquare$$

The next result states that if $m > \dim V$, then there are no alternating m-linear forms on V other than the function on V^m that is identically 0.

9.29   *no nonzero alternating m-linear forms for $m > \dim V$*

Suppose $m > \dim V$. Then 0 is the only alternating m-linear form on V.

Proof   Suppose that α is an alternating m-linear form on V and $v_1, ..., v_m \in V$. Because $m > \dim V$, this list is not linearly independent (by 2.22). Thus 9.28 implies that $\alpha(v_1, ..., v_m) = 0$. Hence α is the zero function from V^m to $\mathbf{F}$. $\qquad \blacksquare$

Alternating Multilinear Forms and Permutations

> **9.30** *swapping input vectors in an alternating multilinear form*
>
> Suppose m is a positive integer, α is an alternating m-linear form on V, and $v_1, ..., v_m$ is a list of vectors in V. Then swapping the vectors in any two slots of $\alpha(v_1, ..., v_m)$ changes the value of α by a factor of -1.

Proof   Put $v_1 + v_2$ in both the first two slots, getting

$$0 = \alpha(v_1 + v_2, v_1 + v_2, v_3, ..., v_m).$$

Use the multilinear properties of α to expand the right side of the equation above (as in the proof of 9.16) to get

$$\alpha(v_2, v_1, v_3, ..., v_m) = -\alpha(v_1, v_2, v_3, ..., v_m).$$

Similarly, swapping the vectors in any two slots of $\alpha(v_1, ..., v_m)$ changes the value of α by a factor of -1. ∎

To see what can happen with multiple swaps, suppose α is an alternating 3-linear form on V and $v_1, v_2, v_3 \in V$. To evaluate $\alpha(v_3, v_1, v_2)$ in terms of $\alpha(v_1, v_2, v_3)$, start with $\alpha(v_3, v_1, v_2)$ and swap the entries in the first and third slots, getting $\alpha(v_3, v_1, v_2) = -\alpha(v_2, v_1, v_3)$. Now in the last expression, swap the entries in the first and second slots, getting

$$\alpha(v_3, v_1, v_2) = -\alpha(v_2, v_1, v_3) = \alpha(v_1, v_2, v_3).$$

More generally, we see that if we do an odd number of swaps, then the value of α changes by a factor of -1, and if we do an even number of swaps, then the value of α does not change.

To deal with arbitrary multiple swaps, we need a bit of information about permutations.

> **9.31** definition: *permutation,* perm m
>
> Suppose m is a positive integer.
>
> - A *permutation* of $(1, ..., m)$ is a list $(j_1, ..., j_m)$ that contains each of the numbers $1, ..., m$ exactly once.
> - The set of all permutations of $(1, ..., m)$ is denoted by perm m.

For example, $(2, 3, 4, 5, 1) \in$ perm 5. You should think of an element of perm m as a rearrangement of the first m positive integers.

The number of swaps used to change a permutation $(j_1, ..., j_m)$ to the standard order $(1, ..., m)$ can depend on the specific swaps selected. The following definition has the advantage of assigning a well-defined sign to every permutation.

<div style="border:1px solid; background:#f5f0a0; padding:8px;">

9.32   definition: *sign of a permutation*

The *sign* of a permutation $(j_1, ..., j_m)$ is defined by

$$\text{sign}(j_1, ..., j_m) = (-1)^N,$$

where N is the number of pairs of integers (k, ℓ) with $1 \leq k < \ell \leq m$ such that k appears after ℓ in the list $(j_1, ..., j_m)$.

</div>

Hence the sign of a permutation equals 1 if the natural order has been changed an even number of times and equals -1 if the natural order has been changed an odd number of times.

9.33   example: *signs*

- The permutation $(1, ..., m)$ [no changes in the natural order] has sign 1.

- The only pair of integers (k, ℓ) with $k < \ell$ such that k appears after ℓ in the list $(2, 1, 3, 4)$ is $(1, 2)$. Thus the permutation $(2, 1, 3, 4)$ has sign -1.

- In the permutation $(2, 3, ..., m, 1)$, the only pairs (k, ℓ) with $k < \ell$ that appear with changed order are $(1, 2), (1, 3), ..., (1, m)$. Because we have $m - 1$ such pairs, the sign of this permutation equals $(-1)^{m-1}$.

<div style="border:1px solid; background:#b8cce4; padding:8px;">

9.34   *swapping two entries in a permutation*

Swapping two entries in a permutation multiplies the sign of the permutation by -1.

</div>

Proof   Suppose we have two permutations, where the second permutation is obtained from the first by swapping two entries. The two swapped entries were in their natural order in the first permutation if and only if they are not in their natural order in the second permutation. Thus we have a net change (so far) of 1 or -1 (both odd numbers) in the number of pairs not in their natural order.

Consider each entry between the two swapped entries. If an intermediate entry was originally in the natural order with respect to both swapped entries, then it is now in the natural order with respect to neither swapped entry. Similarly, if an intermediate entry was originally in the natural order with respect to neither of the swapped entries, then it is now in the natural order with respect to both swapped entries. If an intermediate entry was originally in the natural order with respect to exactly one of the swapped entries, then that is still true. Thus the net change (for each pair containing an entry between the two swapped entries) in the number of pairs not in their natural order is 2, -2, or 0 (all even numbers).

For all other pairs of entries, there is no change in whether or not they are in their natural order. Thus the total net change in the number of pairs not in their natural order is an odd number. Hence the sign of the second permutation equals -1 times the sign of the first permutation. ∎

9.35 *permutations and alternating multilinear forms*

Suppose m is a positive integer and $\alpha \in V_{\text{alt}}^{(m)}$. Then

$$\alpha(v_{j_1}, ..., v_{j_m}) = (\text{sign}(j_1, ..., j_m))\alpha(v_1, ..., v_m)$$

for every list $v_1, ..., v_m$ of vectors in V and all $(j_1, ..., j_m) \in \text{perm } m$.

Proof Suppose $v_1, ..., v_m \in V$ and $(j_1, ..., j_m) \in \text{perm } m$. We can get from $(j_1, ..., j_m)$ to $(1, ..., m)$ by a series of swaps of entries in different slots. Each such swap changes the value of α by a factor of -1 (by 9.30) and also changes the sign of the remaining permutation by a factor of -1 (by 9.34). After an appropriate number of swaps, we reach the permutation $1, ..., m$, which has sign 1. Thus the value of α changed signs an even number of times if $\text{sign}(j_1, ..., j_m) = 1$ and an odd number of times if $\text{sign}(j_1, ..., j_m) = -1$, which gives the desired result. ∎

Our use of permutations now leads in a natural way to the following beautiful formula for alternating n-linear forms on an n-dimensional vector space.

9.36 *formula for* $(\dim V)$*-linear alternating forms on* V

Let $n = \dim V$. Suppose $e_1, ..., e_n$ is a basis of V and $v_1, ..., v_n \in V$. For each $k \in \{1, ..., n\}$, let $b_{1,k}, ..., b_{n,k} \in \mathbf{F}$ be such that

$$v_k = \sum_{j=1}^{n} b_{j,k} e_j.$$

Then

$$\alpha(v_1, ..., v_n) = \alpha(e_1, ..., e_n) \sum_{(j_1, ..., j_n) \in \text{perm } n} (\text{sign}(j_1, ..., j_n))b_{j_1,1} \cdots b_{j_n,n}$$

for every alternating n-linear form α on V.

Proof Suppose α is an alternating n-linear form α on V. Then

$$\alpha(v_1, ..., v_n) = \alpha\left(\sum_{j_1=1}^{n} b_{j_1,1} e_{j_1}, ..., \sum_{j_n=1}^{n} b_{j_n,n} e_{j_n}\right)$$

$$= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} b_{j_1,1} \cdots b_{j_n,n}\, \alpha(e_{j_1}, ..., e_{j_n})$$

$$= \sum_{(j_1, ..., j_n) \in \text{perm } n} b_{j_1,1} \cdots b_{j_n,n}\, \alpha(e_{j_1}, ..., e_{j_n})$$

$$= \alpha(e_1, ..., e_n) \sum_{(j_1, ..., j_n) \in \text{perm } n} (\text{sign}(j_1, ..., j_n))b_{j_1,1} \cdots b_{j_n,n},$$

where the third line holds because $\alpha(e_{j_1}, ..., e_{j_n}) = 0$ if $j_1, ..., j_n$ are not distinct integers, and the last line holds by 9.35. ∎

The following result will be the key to our definition of the determinant in the next section.

9.37 $\dim V_{\text{alt}}^{(\dim V)} = 1$

The vector space $V_{\text{alt}}^{(\dim V)}$ has dimension one.

Proof   Let $n = \dim V$. Suppose α and α' are alternating n-linear forms on V with $\alpha \neq 0$. Let $e_1, ..., e_n$ be such that $\alpha(e_1, ..., e_n) \neq 0$. There exists $c \in \mathbf{F}$ such that

$$\alpha'(e_1, ..., e_n) = c\alpha(e_1, ..., e_n).$$

Furthermore, 9.28 implies that $e_1, ..., e_n$ is linearly independent and thus is a basis of V.

Suppose $v_1, ..., v_n \in V$. Let $b_{j,k}$ be as in 9.36 for $j, k = 1, ..., n$. Then

$$\alpha'(v_1, ..., v_n) = \alpha'(e_1, ..., e_n) \sum_{(j_1, ..., j_n) \in \text{perm } n} (\text{sign}(j_1, ..., j_n)) b_{j_1, 1} \cdots b_{j_n, n}$$

$$= c\alpha(e_1, ..., e_n) \sum_{(j_1, ..., j_n) \in \text{perm } n} (\text{sign}(j_1, ..., j_n)) b_{j_1, 1} \cdots b_{j_n, n}$$

$$= c\alpha(v_1, ..., v_n),$$

where the first and last lines above come from 9.36. The equation above implies that $\alpha' = c\alpha$. Thus α', α is not a linearly independent list, which implies that $\dim V_{\text{alt}}^{(n)} \leq 1$.

To complete the proof, we only need to show that there exists a nonzero alternating n-linear form α on V (thus eliminating the possibility that $\dim V_{\text{alt}}^{(n)}$ equals 0). To do this, let $e_1, ..., e_n$ be any basis of V, and let $\varphi_1, ..., \varphi_n \in V'$ be the linear functionals on V that allow us to express each element of V as a linear combination of $e_1, ..., e_n$. In other words,

$$v = \sum_{j=1}^{n} \varphi_j(v) e_j$$

for every $v \in V$ (see 3.114). Now for $v_1, ..., v_n \in V$, define

9.38 $$\alpha(v_1, ..., v_n) = \sum_{(j_1, ..., j_n) \in \text{perm } n} (\text{sign}(j_1, ..., j_n)) \varphi_{j_1}(v_1) \cdots \varphi_{j_n}(v_n).$$

The verification that α is an n-linear form on V is straightforward.

To see that α is alternating, suppose $v_1, ..., v_n \in V$ with $v_1 = v_2$. For each $(j_1, ..., j_n) \in \text{perm } n$, the permutation $(j_2, j_1, j_3, ..., j_n)$ has the opposite sign. Because $v_1 = v_2$, the contributions from these two permutations to the sum in 9.38 cancel either other. Hence $\alpha(v_1, v_1, v_3, ..., v_n) = 0$. Similarly, $\alpha(v_1, ..., v_n) = 0$ if any two vectors in the list $v_1, ..., v_n$ are equal. Thus α is alternating.

Finally, consider 9.38 with each $v_k = e_k$. Because $\varphi_j(e_k)$ equals 0 if $j \neq k$ and equals 1 if $j = k$, only the permutation $(1, ..., n)$ makes a nonzero contribution to the right side of 9.38 in this case, giving the equation $\alpha(e_1, ..., e_n) = 1$. Thus we have produced a nonzero alternating n-linear form α on V, as desired.   ∎

Earlier we showed that the value of an alternating multilinear form applied to a linearly dependent list is 0; see 9.28. The next result provides a converse of 9.28 for n-linear multilinear forms when $n = \dim V$. In the following result, the statement that α is nonzero means (as usual for a function) that α is not the function on V^n that is identically 0.

> *The formula 9.38 used in the last proof to construct a nonzero alternating n-linear form came from the formula in 9.36, and that formula arose naturally from the properties of an alternating multilinear form.*

9.39 *alternating* $(\dim V)$-*linear forms and linear independence*

Let $n = \dim V$. Suppose α is a nonzero alternating n-linear form on V and $e_1, ..., e_n$ is a list of vectors in V. Then

$$\alpha(e_1, ..., e_n) \neq 0$$

if and only if $e_1, ..., e_n$ is linearly independent.

Proof   First suppose $\alpha(e_1, ..., e_n) \neq 0$. Then 9.28 implies that $e_1, ..., e_n$ is linearly independent.

To prove the implication in the other direction, now suppose $e_1, ..., e_n$ is linearly independent. Because $n = \dim V$, this implies that $e_1, ..., e_n$ is a basis of V (see 2.38).

Because α is not the zero n-linear form, there exist $v_1, ..., v_n \in V$ such that $\alpha(v_1, ..., v_n) \neq 0$. Now 9.36 implies that $\alpha(e_1, ..., e_n) \neq 0$.   ∎

Exercises 9B

1   Suppose m is a positive integer. Show that $\dim V^{(m)} = (\dim V)^m$.

2   Suppose $n \geq 3$ and $\alpha \colon \mathbf{F}^n \times \mathbf{F}^n \times \mathbf{F}^n \to \mathbf{F}$ is defined by

$$\alpha\big((x_1, ..., x_n), (y_1, ..., y_n), (z_1, ..., z_n)\big)$$
$$= x_1 y_2 z_3 - x_2 y_1 z_3 - x_3 y_2 z_1 - x_1 y_3 z_2 + x_3 y_1 z_2 + x_2 y_3 z_1.$$

Show that α is an alternating 3-linear form on $\mathbf{F}^n$.

3   Suppose m is a positive integer and α is an m-linear form on V such that $\alpha(v_1, ..., v_m) = 0$ whenever $v_1, ..., v_m$ is a list of vectors in V with $v_j = v_{j+1}$ for some $j \in \{1, ..., m-1\}$. Prove that α is an alternating m-linear form on V.

4   Prove or give a counterexample: If $\alpha \in V_{\text{alt}}^{(4)}$, then

$$\{(v_1, v_2, v_3, v_4) \in V^4 : \alpha(v_1, v_2, v_3, v_4) = 0\}$$

is a subspace of V^4.

5   Suppose m is a positive integer and β is an m-linear form on V. Define an m-linear form α on V by

$$\alpha(v_1, ..., v_m) = \sum_{(j_1, ..., j_m) \in \text{perm } m} (\text{sign}(j_1, ..., j_m)) \beta(v_{j_1}, ..., v_{j_m})$$

for $v_1, ..., v_m \in V$. Explain why $\alpha \in V_{\text{alt}}^{(m)}$.

6   Suppose m is a positive integer and β is an m-linear form on V. Define an m-linear form α on V by

$$\alpha(v_1, ..., v_m) = \sum_{(j_1, ..., j_m) \in \text{perm } m} \beta(v_{j_1}, ..., v_{j_m})$$

for $v_1, ..., v_m \in V$. Explain why

$$\alpha(v_{k_1}, ..., v_{k_m}) = \alpha(v_1, ..., v_m)$$

for all $v_1, ..., v_m \in V$ and all $(k_1, ..., k_m) \in \text{perm } m$.

7   Give an example of a nonzero alternating 2-linear form α on $\mathbf{R}^3$ and a linearly independent list v_1, v_2 in $\mathbf{R}^3$ such that $\alpha(v_1, v_2) = 0$.

This exercise shows that 9.39 can fail if the hypothesis that $n = \dim V$ is deleted.

9C   *Determinants*

Defining the Determinant

The next definition will lead us to a clean, beautiful, basis-free definition of the determinant of an operator.

9.40   definition: α_T

Suppose that m is a positive integer and $T \in \mathcal{L}(V)$. For $\alpha \in V_{\mathrm{alt}}^{(m)}$, define $\alpha_T \in V_{\mathrm{alt}}^{(m)}$ by

$$\alpha_T(v_1, ..., v_m) = \alpha(Tv_1, ..., Tv_m)$$

for each list $v_1, ..., v_m$ of vectors in V.

Suppose $T \in \mathcal{L}(V)$. If $\alpha \in V_{\mathrm{alt}}^{(m)}$ and $v_1, ..., v_m$ is a list of vectors in V with $v_j = v_k$ for some $j \neq k$, then $Tv_j = Tv_k$, which implies that $\alpha_T(v_1, ..., v_m) = \alpha(Tv_1, ..., Tv_m) = 0$. Thus the function $\alpha \mapsto \alpha_T$ is a linear map of $V_{\mathrm{alt}}^{(m)}$ to itself.

We know that $\dim V_{\mathrm{alt}}^{(\dim V)} = 1$ (see 9.37). Every linear map from a one-dimensional vector space to itself is multiplication by some unique scalar. For the linear map $\alpha \mapsto \alpha_T$, we now define $\det T$ to be that scalar.

9.41   definition: *determinant of an operator,* $\det T$

Suppose $T \in \mathcal{L}(V)$. The *determinant* of T, denoted by $\det T$, is defined to be the unique number in $\mathbf{F}$ such that

$$\alpha_T = (\det T) \, \alpha$$

for all $\alpha \in V_{\mathrm{alt}}^{(\dim V)}$.

9.42   example: *determinants of operators*

Let $n = \dim V$.

- If I is the identity operator on V, then $\alpha_I = \alpha$ for all $\alpha \in V_{\mathrm{alt}}^{(n)}$. Thus $\det I = 1$.

- More generally, if $\lambda \in \mathbf{F}$, then $\alpha_{\lambda I} = \lambda^n \alpha$ for all $\alpha \in V_{\mathrm{alt}}^{(n)}$. Thus $\det(\lambda I) = \lambda^n$.

- Still more generally, if $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$, then $\alpha_{\lambda T} = \lambda^n \alpha_T = \lambda^n (\det T) \alpha$ for all $\alpha \in V_{\mathrm{alt}}^{(n)}$. Thus $\det(\lambda T) = \lambda^n \det T$.

- Suppose $T \in \mathcal{L}(V)$ and there is a basis $e_1, ..., e_n$ of V consisting of eigenvectors of T, with corresponding eigenvalues $\lambda_1, ..., \lambda_n$. If $\alpha \in V_{\mathrm{alt}}^{(n)}$, then

$$\alpha_T(e_1, ..., e_n) = \alpha(\lambda_1 e_1, ..., \lambda_n e_n) = (\lambda_1 \cdots \lambda_n)\alpha(e_1, ..., e_n).$$

  If $\alpha \neq 0$, then 9.39 implies $\alpha(e_1, ..., e_n) \neq 0$. Thus the equation above implies

$$\det T = \lambda_1 \cdots \lambda_n.$$

Our next task is to define and give a formula for the determinant of a square matrix. To do this, we associate with each square matrix an operator and then define the determinant of the matrix to be the determinant of the associated operator.

9.43   definition: *determinant of a matrix,* $\det A$

Suppose that n is a positive integer and A is an n-by-n square matrix with entries in $\mathbf{F}$. Let $T \in \mathcal{L}(\mathbf{F}^n)$ be the operator whose matrix with respect to the standard basis of $\mathbf{F}^n$ equals A. The *determinant* of A, denoted by $\det A$, is defined by $\det A = \det T$.

9.44   example: *determinants of matrices*

- If I is the n-by-n identity matrix, then the corresponding operator on $\mathbf{F}^n$ is the identity operator I on $\mathbf{F}^n$. Thus the first bullet point of 9.42 implies that the determinant of the identity matrix is 1.

- Suppose A is a diagonal matrix with $\lambda_1, \dots, \lambda_n$ on the diagonal. Then the corresponding operator on $\mathbf{F}^n$ has the standard basis of $\mathbf{F}^n$ as eigenvectors, with eigenvalues $\lambda_1, \dots, \lambda_n$. Thus the last bullet point of 9.42 implies that $\det A = \lambda_1 \cdots \lambda_n$.

For the next result, think of each list $v_1, \dots, v_n$ of n vectors in $\mathbf{F}^n$ as a list of n-by-1 column vectors. The notation $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ then denotes the n-by-n square matrix whose k^{th} column is v_k for each $k = 1, \dots, n$.

9.45   *determinant is an alternating multilinear form*

Suppose that n is a positive integer. The map that takes a list $v_1, \dots, v_n$ of vectors in $\mathbf{F}^n$ to $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is an alternating n-linear form on $\mathbf{F}^n$.

Proof   Let $e_1, \dots, e_n$ be the standard basis of $\mathbf{F}^n$ and suppose $v_1, \dots, v_n$ is a list of vectors in $\mathbf{F}^n$. Let $T \in \mathcal{L}(\mathbf{F}^n)$ be the operator such that $Te_k = v_k$ for $k = 1, \dots, n$. Thus T is the operator whose matrix with respect to $e_1, \dots, e_n$ is $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$. Hence $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} = \det T$, by definition of the determinant of a matrix.

Let α be an alternating n-linear form on $\mathbf{F}^n$ such that $\alpha(e_1, \dots, e_n) = 1$. Then

$$
\begin{aligned}
\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} &= \det T \\
&= (\det T)\, \alpha(e_1, \dots, e_n) \\
&= \alpha(Te_1, \dots, Te_n) \\
&= \alpha(v_1, \dots, v_n),
\end{aligned}
$$

where the third line follows from the definition of the determinant of an operator. The equation above shows that the map that takes a list of vectors $v_1, \dots, v_n$ in $\mathbf{F}^n$ to $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is the alternating n-linear form α on $\mathbf{F}^n$.   ∎

The previous result has several important consequences. For example, it immediately implies that a matrix with two identical columns has determinant 0. We will come back to other consequences later, but for now we want to give a formula for the determinant of a square matrix. Recall that if A is a matrix, then $A_{j,k}$ denotes the entry in row j, column k of A.

9.46 *formula for determinant of a matrix*

Suppose that n is a positive integer and A is an n-by-n square matrix. Then

$$\det A = \sum_{(j_1,...,j_n) \in \text{perm } n} (\text{sign}(j_1,...,j_n)) A_{j_1,1} \cdots A_{j_n,n}.$$

Proof Apply 9.36 with $V = \mathbf{F}^n$ and $e_1, ..., e_n$ the standard basis of $\mathbf{F}^n$ and α the alternating n-linear form on $\mathbf{F}^n$ that takes $v_1, ..., v_n$ to $\det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ [see 9.45]. If each v_k is the kth column of A, then each $b_{j,k}$ in 9.36 equals $A_{j,k}$. Finally,

$$\alpha(e_1, ..., e_n) = \det \begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix} = \det I = 1.$$

Thus the formula in 9.36 becomes the formula stated in this result. ∎

9.47 example: *explicit formula for determinant*

- If A is a 2-by-2 matrix, then the formula in 9.46 becomes

$$\det A = A_{1,1} A_{2,2} - A_{2,1} A_{1,2}.$$

- If A is a 3-by-3 matrix, then the formula in 9.46 becomes

$$\det A = A_{1,1} A_{2,2} A_{3,3} - A_{2,1} A_{1,2} A_{3,3} - A_{3,1} A_{2,2} A_{1,3}$$
$$- A_{1,1} A_{3,2} A_{2,3} + A_{3,1} A_{1,2} A_{2,3} + A_{2,1} A_{3,2} A_{1,3}.$$

The sum in the formula in 9.46 contains $n!$ terms. Because $n!$ grows rapidly as n increases, the formula in 9.46 is not a viable method to evaluate determinants even for moderately sized n. For example, $10!$ is over three million, and $100!$ is approximately 10^{158}, leading to a sum that the fastest computer cannot evaluate. We will soon see some results that lead to faster evaluations of determinants than direct use of the sum in 9.46.

9.48 *determinant of upper-triangular matrix*

Suppose that A is an upper-triangular matrix with $\lambda_1, ..., \lambda_n$ on the diagonal. Then $\det A = \lambda_1 \cdots \lambda_n$.

Proof If $(j_1, ..., j_n) \in \text{perm } n$ with $(j_1, ..., j_n) \neq (1, ..., n)$, then $j_k > k$ for some $k \in \{1, ..., n\}$, which implies that $A_{j_k,k} = 0$. Thus the only permutation that can make a nonzero contribution to the sum in 9.46 is the permutation $(1, ..., n)$. Because $A_{k,k} = \lambda_k$ for each $k = 1, ..., n$, this implies that $\det A = \lambda_1 \cdots \lambda_n$. ∎

Properties of Determinants

Our definition of the determinant leads to the following magical proof that the determinant is multiplicative.

9.49 *determinant is multiplicative*

(a) Suppose $S, T \in \mathcal{L}(V)$. Then $\det(ST) = (\det S)(\det T)$.

(b) Suppose A and B are square matrices of the same size. Then

$$\det(AB) = (\det A)(\det B)$$

Proof

(a) Let $n = \dim V$. Suppose $\alpha \in V_{\text{alt}}^{(n)}$ and $v_1, ..., v_n \in V$. Then

$$\begin{aligned}
\alpha_{ST}(v_1, ..., v_n) &= \alpha(STv_1, ..., STv_n) \\
&= (\det S)\alpha(Tv_1, ..., Tv_n) \\
&= (\det S)(\det T)\alpha(v_1, ..., v_n),
\end{aligned}$$

where the first equation follows from the definition of α_{ST}, the second equation follows from the definition of $\det S$, and the third equation follows from the definition of $\det T$. The equation above implies that $\det(ST) = (\det S)(\det T)$.

(b) Let $S, T \in \mathcal{L}(\mathbf{F}^n)$ be such that $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = B$, where all matrices of operators in this proof are with respect to the standard basis of $\mathbf{F}^n$. Then $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T) = AB$ (see 3.43). Thus

$$\det(AB) = \det(ST) = (\det S)(\det T) = (\det A)(\det B),$$

where the second equality comes from the result in (a). ∎

The determinant of an operator determines whether the operator is invertible.

9.50 *invertible* $\iff$ *nonzero determinant*

An operator $T \in \mathcal{L}(V)$ is invertible if and only if $\det T \neq 0$. Furthermore, if T is invertible, then $\det(T^{-1}) = \frac{1}{\det T}$.

Proof   First suppose T is invertible. Thus $TT^{-1} = I$. Now 9.49 implies that

$$1 = \det I = \det(TT^{-1}) = (\det T)\left(\det(T^{-1})\right).$$

Hence $\det T \neq 0$ and $\det(T^{-1})$ is the multiplicative inverse of $\det T$.

To prove the other direction, now suppose $\det T \neq 0$. Suppose $v \in V$ and $v \neq 0$. Let $v, e_2, ..., e_n$ be a basis of V and let $\alpha \in V_{\text{alt}}^{(n)}$ be such that $\alpha \neq 0$. Then $\alpha(v, e_2, ..., e_n) \neq 0$ (by 9.39). Now

$$\alpha(Tv, Te_2, ..., Te_n) = (\det T)\alpha(v, e_2, ..., e_n) \neq 0,$$

Thus $Tv \neq 0$. Hence T is invertible. ∎

An n-by-n matrix A is invertible (see 3.80 for the definition of an invertible matrix) if and only if the operator on $\mathbf{F}^n$ associated with A (via the standard basis of $\mathbf{F}^n$) is invertible. Thus the previous result shows that a square matrix A is invertible if and only if $\det A \neq 0$.

9.51   *eigenvalues and determinants*

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Then λ is an eigenvalue of T if and only if $\det(\lambda I - T) = 0$.

Proof   The number λ is an eigenvalue of T if and only if $T - \lambda I$ is not invertible (see 5.7), which happens if and only if $\lambda I - T$ is not invertible, which happens if and only if $\det(\lambda I - T) = 0$ (by 9.50). ■

Suppose $T \in \mathcal{L}(V)$ and $S \colon W \to V$ is an invertible linear map. To prove that $\det(S^{-1}TS) = \det T$, we could try to use 9.49 and 9.50, writing

$$\det(S^{-1}TS) = (\det S^{-1})(\det T)(\det S)$$
$$= \det T.$$

That proof works if $W = V$, but if $W \neq V$ then it makes no sense because the determinant is defined only for linear maps from a vector space to itself, and S maps W to V, making $\det S$ undefined. The proof given below works around this issue and is valid when $W \neq V$.

9.52   *determinant is a similarity invariant*

Suppose $T \in \mathcal{L}(V)$ and $S \colon W \to V$ is an invertible linear map. Then

$$\det(S^{-1}TS) = \det T.$$

Proof   Let $n = \dim W = \dim V$. Suppose $\tau \in W_{\text{alt}}^{(n)}$. Define $\alpha \in V_{\text{alt}}^{(n)}$ by

$$\alpha(v_1, \ldots, v_n) = \tau(S^{-1}v_1, \ldots, S^{-1}v_n)$$

for $v_1, \ldots, v_n \in V$. Suppose $w_1, \ldots, w_n \in W$. Then

$$\tau_{S^{-1}TS}(w_1, \ldots, w_n) = \tau(S^{-1}TSw_1, \ldots, S^{-1}TSw_n)$$
$$= \alpha(TSw_1, \ldots, TSw_n)$$
$$= \alpha_T(Sw_1, \ldots, Sw_n)$$
$$= (\det T)\alpha(Sw_1, \ldots, Sw_n)$$
$$= (\det T)\tau(w_1, \ldots, w_n).$$

The equation above and the definition of the determinant of the operator $S^{-1}TS$ imply that $\det(S^{-1}TS) = \det T$. ■

For the special case in which $V = \mathbf{F}^n$ and $e_1, ..., e_n$ is the standard basis of $\mathbf{F}^n$, the next result is true by the definition of the determinant of a matrix. The left side of the equation in the next result does not depend on a choice of basis, which means that the right side is independent of the choice of basis.

9.53   *determinant of operator equals determinant of its matrix*

Suppose $T \in \mathcal{L}(V)$ and $e_1, ..., e_n$ is a basis of V. Then

$$\det T = \det \mathcal{M}(T, (e_1, ..., e_n)).$$

Proof   Let $f_1, ..., f_n$ be the standard basis of $\mathbf{F}^n$. Let $S\colon \mathbf{F}^n \to V$ be the linear map such that $Sf_k = e_k$ for each $k = 1, ..., n$. Thus $\mathcal{M}(S, (f_1, ..., f_n), (e_1, ..., e_n))$ and $\mathcal{M}(S^{-1}, (e_1, ..., e_n), (f_1, ..., f_n))$ both equal the n-by-n identity matrix. Hence

9.54 $$\mathcal{M}(S^{-1}TS, (f_1, ..., f_n)) = \mathcal{M}(T, (e_1, ..., e_n)),$$

as follows from two applications of 3.43. Thus

$$\det T = \det(S^{-1}TS)$$

$$= \det \mathcal{M}(S^{-1}TS, (f_1, ..., f_n))$$

$$= \det \mathcal{M}(T, (e_1, ..., e_n)),$$

where the first line comes from 9.52, the second line comes from the definition of the determinant of a matrix, and the third line follows from 9.54.   ∎

The next result gives a more intuitive way to think about determinants than the definition or the formula in 9.46. We could make the characterization in the result below the definition of the determinant of an operator on a finite-dimensional complex vector space, with the current definition then becoming a consequence of that definition.

9.55   *if $\mathbf{F} = \mathbf{C}$, then determinant equals product of eigenvalues*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then $\det T$ equals the product of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.

Proof   There is a basis of V with respect to which T has an upper-triangular matrix with the diagonal entries of the matrix consisting of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity—see 8.37. Thus 9.53 and 9.48 imply that $\det T$ equals the product of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.   ∎

As the next result shows, the determinant interacts nicely with the transpose of a square matrix, with the dual of an operator, and with the adjoint of an operator on an inner product space.

9.56 *determinant of transpose, dual, or adjoint*

(a) Suppose A is a square matrix. Then $\det A^t = \det A$.

(b) Suppose $T \in \mathcal{L}(V)$. Then $\det T' = \det T$.

(c) Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Then
$$\det(T^*) = \overline{\det T}.$$

Proof

(a) Let n be a positive integer. Define $\alpha \colon (\mathbf{F}^n)^n \to \mathbf{F}$ by
$$\alpha\big(\big(\; v_1 \;\;\cdots\;\; v_n \;\big)\big) = \det\big(\big(\; v_1 \;\;\cdots\;\; v_n \;\big)^t\big)$$
for all $v_1, \ldots, v_n \in \mathbf{F}^n$. The formula in 9.46 for the determinant of a matrix shows that α is an n-linear form on $\mathbf{F}^n$.

Suppose $v_1, \ldots, v_n \in \mathbf{F}^n$ and $v_j = v_k$ for some $j \neq k$. If B is an n-by-n matrix, then $\big(\; v_1 \;\;\cdots\;\; v_n \;\big)^t B$ cannot equal the identity matrix because row j and row k of $\big(\; v_1 \;\;\cdots\;\; v_n \;\big)^t B$ are equal. Thus $\big(\; v_1 \;\;\cdots\;\; v_n \;\big)^t$ is not invertible, which implies that $\alpha\big(\big(\; v_1 \;\;\cdots\;\; v_n \;\big)\big) = 0$. Hence α is an alternating n-linear form on $\mathbf{F}^n$.

Note that α applied to the standard basis of $\mathbf{F}^n$ equals 1. Because the vector space of alternating n-linear forms on $\mathbf{F}^n$ has dimension one (by 9.37), this implies that α is the determinant function. Thus (a) holds.

(b) The equation $\det T' = \det T$ follows from (a) and 9.53 and 3.132.

(c) Pick an orthonormal basis of V. The matrix of T^* with respect to that basis is the conjugate transpose of the matrix of T with respect to that basis (by 7.9). Thus 9.53, 9.46, and (a) imply that $\det(T^*) = \overline{\det T}$. ∎

9.57 *helpful results in evaluating determinants*

(a) If either two columns or two rows of a square matrix are equal, then the determinant of the matrix equals 0.

(b) Suppose A is a square matrix and B is the matrix obtained from A by swapping either two columns or two rows. Then $\det A = -\det B$.

(c) If one column or one row of a square matrix is multiplied by a scalar, then the value of the determinant is multiplied by the same scalar.

(d) If a scalar multiple of one column of a square matrix to added to another column, then the value of the determinant is unchanged.

(e) If a scalar multiple of one row of a square matrix to added to another row, then the value of the determinant is unchanged.

Proof All the assertions in this result follow from the result that the maps $v_1, ..., v_n \mapsto \det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ and $v_1, ..., v_n \mapsto \det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}^{\mathrm{t}}$ are both alternating n-linear forms on $\mathbf{F}^n$ [see 9.45 and 9.56(a)].

For example, to prove (d) suppose $v_1, ..., v_n \in \mathbf{F}^n$ and $c \in \mathbf{F}$. Then

$$\det \begin{pmatrix} v_1 + cv_2 & v_2 & \cdots & v_n \end{pmatrix}$$

$$= \det \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} + c \det \begin{pmatrix} v_2 & v_2 & v_3 & \cdots & v_n \end{pmatrix}$$

$$= \det \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix},$$

where the first equation follows from the multilinearity property and the second equation follows from the alternating property. The equation above shows that adding a multiple of the second column to the first column does not change the value of the determinant. The same conclusion holds for any two columns. Thus (d) holds.

The proof of (e) follows from (d) and from 9.56(a). The proofs of (a), (b), and (c) use similar tools and are left to the reader. ∎

For matrices whose entries are concrete numbers, the result above leads to a much faster way to evaluate the determinant than direct application of the formula in 9.46. Specifically, apply the Gaussian elimination procedure of swapping rows [by 9.48(b), this changes the determinant by a factor of -1], multiplying a row by a nonzero constant [by 9.48(c), this changes the determinant by the same constant], and adding a multiple of one row to another row [by 9.48(e), this does not change the determinant] to produce an upper-triangular matrix, whose determinant is the product of the diagonal entries (by 9.48). If your software keeps track of the number of row swaps and of the constants used when multiplying a row by a constant, then the determinant of the original matrix can be computed.

Because a number $\lambda \in \mathbf{F}$ is an eigenvalue of an operator $T \in \mathcal{L}(V)$ if and only if $\det(\lambda I - T) = 0$ (by 9.51), you may be tempted to think that one way to find eigenvalues quickly is to choose a basis of V, let $A = \mathcal{M}(T)$, evaluate $\det(\lambda I - A)$, and then solve the equation $\det(\lambda I - A) = 0$ for λ. However, that procedure is rarely efficient, except when $\dim V = 2$ (or when $\dim V$ equals 3 or 4 if you are willing to use the cubic or quartic formulas). One problem is that the procedure described in the paragraph above for evaluating a determinant does not work when the matrix includes a symbol (such as the λ in $\lambda I - A$). This problem arises because decisions need to be made in the Gaussian elimination procedure about whether certain quantities equal 0, and those decisions become complicated in expressions involving a symbol λ.

Recall that an operator on a finite-dimensional inner product space is unitary if it preserves norms (see 7.51 and the paragraph following it). Every eigenvalue of a unitary operator has absolute value 1 (by 7.54). Thus the product of the eigenvalues of a unitary operator has absolute value 1. Hence (at least in the case $\mathbf{F} = \mathbf{C}$) the determinant of a unitary operator has absolute value 1 (by 9.55). The next result gives a proof that works without the assumption that $\mathbf{F} = \mathbf{C}$.

> **9.58** *every unitary operator has determinant with absolute value* 1
>
> Suppose V is an inner product space and $S \in \mathcal{L}(V)$ is a unitary operator. Then $|\det S| = 1$.

Proof Because S is unitary, $I = S^*S$ (see 7.53). Thus

$$1 = \det(S^*S) = (\det S^*)(\det S) = \overline{(\det S)}(\det S) = |\det S|^2,$$

where the second equality comes from 9.49(a) and the third equality comes from 9.56(c). The equation above implies that $|\det S| = 1$. ∎

The determinant of a positive operator on an inner product space meshes well with the analogy that such operators correspond to the nonnegative real numbers.

> **9.59** *every positive operator has nonnegative determinant*
>
> Suppose V is an inner product space and $T \in \mathcal{L}(V)$ is a positive operator. Then $\det T \geq 0$.

Proof By the spectral theorem (7.29 or 7.31), V has an orthonormal basis consisting of eigenvectors of T. Thus by the last bullet point of 9.42, $\det T$ equals a product of the eigenvalues of T, possibly with repetitions. Each eigenvalue of T is a nonnegative number (by 7.38). Thus we conclude that $\det T \geq 0$. ∎

Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Recall that the list of nonnegative square roots of the eigenvalues of T^*T (each included as many times as its multiplicity) is called the list of singular values of T (see Section 7E).

> **9.60** $|\det T| = $ *product of singular values of* T
>
> Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Then
>
> $$|\det T| = \sqrt{\det(T^*T)} = \text{product of singular values of } T.$$

Proof We have

$$|\det T|^2 = \overline{(\det T)}(\det T) = \big(\det(T^*)\big)(\det T) = \det(T^*T),$$

where the middle equality comes from 9.56(c) and the last equality comes from 9.49(a). Taking square roots of both sides of the equation above shows that $|\det T| = \sqrt{\det(T^*T)}$.

Let $s_1, \ldots, s_n$ denote the list of singular values of T. Thus $s_1^2, \ldots, s_n^2$ is the list of eigenvalues of T^*T (with appropriate repetitions), corresponding to an orthonormal basis of V consisting of eigenvectors of T^*T. Hence the last bullet point of 9.42 implies that

$$\det(T^*T) = s_1^2 \cdots s_n^2.$$

Thus $|\det T| = s_1 \cdots s_n$, as desired. ∎

An operator T on a real inner product space changes volume by a factor of the product of the singular values (by 7.111). Thus the next result follows immediately from 7.111 and 9.60. This result explains why the absolute value of a determinant appears in the change of variables formula in multivariable calculus.

9.61   *T changes volume by factor of* $|\det T|$

Suppose $T \in \mathcal{L}(\mathbf{R}^n)$ and $\Omega \subseteq \mathbf{R}^n$. Then

$$\text{volume } T(\Omega) = |\det T|(\text{volume } \Omega).$$

For operators on finite-dimensional complex vector spaces, we now connect the determinant to a polynomial that we have previously seen.

9.62   *if* $\mathbf{F} = \mathbf{C}$, *then characteristic polynomial of* T *equals* $\det(zI - T)$

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, ..., \lambda_m$ denote the distinct eigenvalues of T, and let $d_1, ..., d_m$ denote their multiplicities. Then

$$\det(zI - T) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}.$$

Proof   There exists a basis of V with respect to which T has an upper-triangular matrix with each λ_k appearing on the diagonal exactly d_k times (by 8.37). With respect to this basis, $zI - T$ has an upper-triangular matrix with $z - \lambda_k$ appearing on the diagonal exactly d_k times for each k. Thus 9.48 gives the desired equation. ∎

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. The characteristic polynomial of T was defined in 8.26 as the polynomial on the right side of the equation in 9.62. We did not previously define the characteristic polynomial of an operator on a finite-dimensional real vector space because such operators may have no eigenvalues, making a definition using the right side of the equation in 9.62 inappropriate.

We now present a new definition of the characteristic polynomial, motivated by 9.62. This new definition is valid for both real and complex vector spaces. The equation in 9.62 shows that this new definition is equivalent to our previous definition when $\mathbf{F} = \mathbf{C}$ (8.26).

9.63   definition: *characteristic polynomial*

Suppose $T \in \mathcal{L}(V)$. The polynomial defined by

$$z \mapsto \det(zI - T)$$

is called the *characteristic polynomial* of T.

The formula in 9.46 shows that the characteristic polynomial of an operator $T \in \mathcal{L}(V)$ is a monic polynomial of degree dim V. The zeros in $\mathbf{F}$ of the characteristic polynomial of T are exactly the eigenvalues of T (by 9.51).

Previously we proved the Cayley–Hamilton theorem (8.29) in the complex case. Now we can extend that result to operators on real vector spaces.

9.64 *Cayley–Hamilton theorem*

Suppose $T \in \mathcal{L}(V)$ and q is the characteristic polynomial of T. Then $q(T) = 0$.

Proof If $\mathbf{F} = \mathbf{C}$, then the equation $q(T) = 0$ follows from 9.62 and 8.29.

Now suppose $\mathbf{F} = \mathbf{R}$. Fix a basis of V, and let A be the matrix of T with respect to this basis. Let S be the operator on $\mathbf{C}^{\dim V}$ such that the matrix of S (with respect to the standard basis of $\mathbf{C}^{\dim V}$) is A. For all $z \in \mathbf{R}$ we have

$$q(z) = \det(zI - T) = \det(zI - A) = \det(zI - S).$$

Thus q is the characteristic polynomial of S. The case $\mathbf{F} = \mathbf{C}$ (first sentence of this proof) now implies that $0 = q(S) = q(A) = q(T)$. ∎

The Cayley–Hamilton theorem (9.64) implies that the characteristic polynomial of an operator $T \in \mathcal{L}(V)$ is a polynomial multiple of the minimal polynomial of T (by 5.29). Thus if the degree of the minimal polynomial of T equals $\dim V$, then the characteristic polynomial of T equals the minimal polynomial of T. This happens for a very large percentage of operators, including over 99.999% of 4-by-4 matrices with integer entries in $[-100, 100]$ (see the paragraph following 5.25).

The last sentence in our next result was previously proved in the complex case (see 8.54). Now we can give a proof that works on both real and complex vector spaces.

9.65 *characteristic polynomial, trace, and determinant*

Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then the characteristic polynomial of T can be written as

$$z^n - (\operatorname{tr} T)z^{n-1} + \cdots + (-1)^n (\det T).$$

Proof The constant term of a polynomial function of z is the value of the polynomial when $z = 0$. Thus the constant term of the characteristic polynomial of T equals $\det(-T)$, which equals $(-1)^n \det T$ (by the third bullet point of 9.42).

Fix a basis of V, and let A be the matrix of T with respect to this basis. The matrix of $zI - T$ with respect to this basis is $zI - A$. The term coming from the identity permutation $\{1, \ldots, n\}$ in the formula 9.46 for $\det(zI - A)$ is

$$(z - A_{1,1}) \cdots (z - A_{n,n}).$$

The coefficient of z^{n-1} in the expression above is $-(A_{1,1} + \cdots + A_{n,n})$, which equals $-\operatorname{tr} T$. The terms in the formula for $\det(zI - A)$ coming from other elements of perm n contain at most $n - 2$ factors of the form $z - A_{k,k}$ and thus do not contribute to the coefficient of z^{n-1} in the characteristic polynomial of T. ∎

In the result below, think of the columns of the n-by-n matrix A as elements of $\mathbf{F}^n$. The norms appearing below

The next result was proved by Jacques Hadamard (1865–1963) in 1893.

then arise from the standard inner product on $\mathbf{F}^n$. Recall that the notation $R_{\cdot,k}$ in the proof below means the kth column of the matrix R (as was defined in 3.44).

9.66   *Hadamard's inequality*

Suppose A is an n-by-n matrix. Let $v_1, ..., v_n$ denote the columns of A. Then

$$|\det A| \leq \prod_{k=1}^{n} \|v_k\|.$$

Proof   If A is not invertible, then $\det A = 0$ and hence the desired inequality holds in this case.

Thus assume that A is invertible. The QR factorization (7.58) tells us that there exist a unitary matrix Q and an upper-triangular matrix R whose diagonal contains only positive numbers such that $A = QR$. We have

$$|\det A| = |\det Q|\, |\det R|$$

$$= |\det R|$$

$$= \prod_{k=1}^{n} R_{k,k}$$

$$\leq \prod_{k=1}^{n} \|R_{\cdot,k}\|$$

$$= \prod_{k=1}^{n} \|QR_{\cdot,k}\|$$

$$= \prod_{k=1}^{n} \|v_k\|,$$

where the first line comes from 9.49(b), the second line comes from 9.58, the third line comes from 9.48, and the fifth line holds because Q is an isometry.   ∎

To give a geometric interpretation to Hadamard's inequality, suppose $\mathbf{F} = \mathbf{R}$. Let $T \in \mathcal{L}(\mathbf{R}^n)$ be the operator such that $Te_k = v_k$ for each $k = 1, ..., n$, where $e_1, ..., e_n$ is the standard basis of $\mathbf{R}^n$. Then T maps the box $P(e_1, ..., e_n)$ onto the parallelepiped $P(v_1, ..., v_n)$ [see 7.102 and 7.105 for a review of this notation and terminology]. Because the box $P(e_1, ..., e_n)$ has volume 1, this implies (by 9.61) that the parallelepiped $P(v_1, ..., v_n)$ has volume $|\det T|$, which equals $|\det A|$. Thus Hadamard's inequality above can be interpreted to say that among all parallelepipeds whose edges have lengths $\|v_1\|, ..., \|v_n\|$, the ones with largest volume have orthogonal edges $\left(\text{and thus have volume } \prod_{k=1}^{n} \|v_k\|\right)$.

For a necessary and sufficient condition for Hadamard's inequality to be an equality, see Exercise 18.

The matrix in the next result is called the *Vandermonde matrix*. Vandermonde matrices have important applications in polynomial interpolation, the discrete Fourier transform, and other areas of mathematics. The proof of the next result is a nice illustration of the power of switching between matrices and linear maps.

> **9.67**   *determinant of Vandermonde matrix*
>
> Suppose $n > 1$ and $\beta_1, ..., \beta_n \in \mathbf{F}$. Then
> $$\det \begin{pmatrix} 1 & \beta_1 & \beta_1^2 & \cdots & \beta_1^{n-1} \\ 1 & \beta_2 & \beta_2^2 & \cdots & \beta_2^{n-1} \\ & & \ddots & & \\ 1 & \beta_n & \beta_n^2 & \cdots & \beta_n^{n-1} \end{pmatrix} = \prod_{1 \le j < k \le n} (\beta_k - \beta_j).$$

Proof   Let $1, z, ..., z^{n-1}$ be the standard basis of $\mathcal{P}_{n-1}(\mathbf{F})$ and let $e_1, ..., e_n$ denote the standard basis of $\mathbf{F}^n$. Define a linear map $S \colon \mathcal{P}_{n-1}(\mathbf{F}) \to \mathbf{F}^n$ by

$$Sp = (p(\beta_1), ..., p(\beta_n)).$$

Let A denote the Vandermonde matrix shown in the statement of this result. Note that
$$A = \mathcal{M}(S, (1, z, ..., z^{n-1}), (e_1, ..., e_n)).$$

Let $T \colon \mathcal{P}_{n-1}(\mathbf{F}) \to \mathcal{P}_{n-1}(\mathbf{F})$ be the operator on $\mathcal{P}_{n-1}(\mathbf{F})$ such that $T1 = 1$ and
$$Tz^k = (z - \beta_1)(z - \beta_2)\cdots(z - \beta_k)$$
for $k = 1, ..., n - 1$. Let $B = \mathcal{M}(T, (1, z, ..., z^{n-1}), (1, z, ..., z^{n-1}))$. Then B is an upper-triangular matrix all of whose diagonal entries equal 1. Thus $\det B = 1$ (by 9.48).

Let $C = \mathcal{M}(ST, (1, z, ..., z^{n-1}), (e_1, ..., e_n))$. Thus $C = AB$ (by 3.81), which implies that
$$\det A = (\det A)(\det B) = \det C.$$
The definitions of C, S, and T show that C equals

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & \beta_2 - \beta_1 & 0 & \cdots & 0 \\ 1 & \beta_2 - \beta_1 & (\beta_3 - \beta_1)(\beta_3 - \beta_2) & \cdots & 0 \\ & & & \ddots & \\ 1 & \beta_2 - \beta_1 & (\beta_3 - \beta_1)(\beta_3 - \beta_2) & \cdots & (\beta_n - \beta_1)(\beta_n - \beta_2)\cdots(\beta_n - \beta_{n-1}) \end{pmatrix}.$$

Now $\det A = \det C = \displaystyle\prod_{1 \le j < k \le n} (\beta_k - \beta_j)$, where we have used 9.56(a) and 9.48. ∎

Exercises 9C

1   Prove or give a counterexample: $S, T \in \mathcal{L}(V) \implies \det(S+T) = \det S + \det T$.

2   Suppose the first column of a square matrix A consists of all zeros except possibly the first entry $A_{1,1}$. Let B be the matrix obtained from A by deleting the first row and the first column of A. Show that $\det A = A_{1,1} \det B$.

3   Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that $\det(I + T) = 1$.

4   Suppose $S \in \mathcal{L}(V)$. Prove that S is unitary if and only if $|\det S| = \|S\| = 1$.

5   Suppose A is a block upper-triangular matrix

$$A = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_k along the diagonal is a square matrix. Prove that

$$\det A = (\det A_1)\cdots(\det A_m).$$

6   Suppose $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is an n-by-n matrix, with v_k denoting the kth column of A. Show that if $(m_1, ..., m_n) \in \text{perm } n$, then

$$\det \begin{pmatrix} v_{m_1} & \cdots & v_{m_n} \end{pmatrix} = (\text{sign}(m_1, ..., m_n)) \det A.$$

7   Suppose $T \in \mathcal{L}(V)$ is invertible. Let p denote the characteristic polynomial of T and let q denote the characteristic polynomial of T^{-1}. Prove that

$$q(z) = \frac{1}{p(0)} z^{\dim V} p\left(\frac{1}{z}\right)$$

for all nonzero $z \in \mathbf{F}$.

8   Suppose $T \in \mathcal{L}(V)$ is an operator with no eigenvalues (which implies that $\mathbf{F} = \mathbf{R}$). Prove that $\det T > 0$.

9   Suppose that V is a real vector space of even dimension, $T \in \mathcal{L}(V)$, and $\det T < 0$. Prove that T has at least two distinct eigenvalues.

10   Suppose V is a real vector space of odd dimension and $T \in \mathcal{L}(V)$. Without using the minimal polynomial, prove that T has an eigenvalue.

   This result was previously proved without using determinants or the characteristic polynomial—see 5.34.

11   Prove or give a counterexample: If $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\det T > 0$, then T has a square root.

   If $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, and $\det V \neq 0$, then V has a square root (see 8.41).

12   Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Define $p \colon \mathbf{F} \to \mathbf{F}$ by

$$p(z) = \det(zS - T).$$

Prove that p is a polynomial of degree $\dim V$ and that the coefficient of $z^{\dim V}$ in this polynomial is $\det S$.

13   Suppose $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, and $n = \dim V > 2$. Let $\lambda_1, ..., \lambda_n$ denote the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.

(a) Find a formula for the coefficient of z^{n-2} in the characteristic polynomial of T in terms of $\lambda_1, ..., \lambda_n$.

(b) Find a formula for the coefficient of z in the characteristic polynomial of T in terms of $\lambda_1, ..., \lambda_n$.

14   Suppose V is an inner product space and T is a positive operator on V. Prove that

$$\det \sqrt{T} = \sqrt{\det T}.$$

15   Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Use the polar decomposition to give a proof that

$$|\det T| = \sqrt{\det(T^*T)}$$

that is different from the proof given earlier (see 9.60).

16   Suppose $T \in \mathcal{L}(V)$. Define $g \colon \mathbf{F} \to \mathbf{F}$ by $g(x) = \det(I + xT)$. Show that $g'(0) = \operatorname{tr} T$.

> *Look for a clean solution to this exercise, without using the explicit but complicated formula for the determinant of a matrix.*

17   Suppose a, b, c are positive numbers. Find the volume of the ellipsoid

$$\left\{ (x, y, z) \in \mathbf{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1 \right\}$$

by finding a set $\Omega \subseteq \mathbf{R}^3$ whose volume you know and an operator T on $\mathbf{R}^3$ such that $T(\Omega)$ equals the ellipsoid above.

18   Suppose that A is an invertible square matrix. Prove that Hadamard's inequality (9.66) is an equality if and only if each column of A is orthogonal to the other columns.

19   Suppose V is an inner product space, $e_1, ..., e_n$ is an orthonormal basis of V, and $T \in \mathcal{L}(V)$ is a positive operator.

(a) Prove that $\det T \leq \prod_{k=1}^{n} \langle Te_k, e_k \rangle$.

(b) Prove that if T is invertible, then the inequality in (a) is an equality if and only if e_k is an eigenvector of T for each $k = 1, ..., n$.

20   Suppose A is an n-by-n matrix, and suppose c is such that $|A_{j,k}| \leq c$ for all $j, k \in \{1, ..., n\}$. Prove that

$$|\det A| \leq c^n n^{n/2}.$$

The formula for the determinant of a matrix (9.46) shows that $|\det A| \leq c^n n!$. However, the estimate given by this exercise is much better. For example, if $c = 1$ and $n = 100$, then $c^n n! \approx 10^{158}$, but the estimate given by this exercise is the much smaller number 10^{100}. If n is an integer power of 2, then the inequality above is sharp and cannot be improved.

21   Suppose n is a positive integer and $\delta \colon \mathbf{C}^{n,n} \to \mathbf{C}$ is a function such that

$$\delta(AB) = \delta(A) \cdot \delta(B)$$

for all $A, B \in \mathbf{C}^{n,n}$ and $\delta(A)$ equals the product of the diagonal entries of A for each diagonal matrix $A \in \mathbf{C}^{n,n}$. Prove that

$$\delta(A) = \det A$$

for all $A \in \mathbf{C}^{n,n}$.

Recall that $\mathbf{C}^{n,n}$ denotes set of n-by-n matrices with entries in $\mathbf{C}$. This exercise shows that the determinant is the unique function defined on square matrices that is multiplicative and has the desired behavior on diagonal matrices. This result is analogous to Exercise 10 in Section 8D, which shows that the trace is uniquely determined by its algebraic properties.

I find that in my own elementary lectures, I have, for pedagogical reasons, pushed determinants more and more into the background. Too often I have had the experience that, while the students acquired facility with the formulas, which are so useful in abbreviating long expressions, they often failed to gain familiarity with their *meaning*, and skill in manipulation prevented the student from going into all the details of the subject and so gaining a mastery.

—*Elementary Mathematics from an Advanced Standpoint: Geometry*, Felix Klein