Linear Algebra II

Richard Earl

Hilary Term 2024

0. INTRODUCTION AND PRELIMINARY MATERIAL

0.1 Syllabus

Introduction to determinant of a square matrix: existence and uniqueness. Proof of existence by induction. Proof of uniqueness by deriving explicit formula from the properties of the determinant. Permutation matrices. (No general discussion of permutations). Basic properties of determinant, relation to volume. Multiplicativity of the determinant, computation by row operations. [2]

Determinants and linear transformations: definition of the determinant of a linear transformation, multiplicativity, invertibility and the determinant. [0.5]

Eigenvectors and eigenvalues, the characteristic polynomial, trace. Eigenvectors for distinct eigenvalues are linearly independent. Discussion of diagonalization. Examples. Eigenspaces, geometric and algebraic multiplicity of eigenvalues. Eigenspaces form a direct sum. [2.5]

Gram-Schmidt procedure. Spectral theorem for real symmetric matrices. Quadratic forms and real symmetric matrices. Application of the spectral theorem to putting quadrics into normal form by orthogonal transformations and translations. [3]

0.2 Reading list

(1) T. S. Blyth and E. F. Robertson, Basic Linear Algebra (Springer, London, 2nd edition 2002).
(2) C. W. Curtis, Linear Algebra – An Introductory Approach (Springer, New York, 4th edition, reprinted 1994).
(3) R. B. J. T. Allenby, Linear Algebra (Arnold, London, 1995).
(4) D. A. Towers, A Guide to Linear Algebra (Macmillan, Basingstoke 1988).
(5) S. Lang, Linear Algebra (Springer, London, Third Edition, 1987).
(6) R. Earl, Towards Higher Mathematics - A Companion (Cambridge University Press, Cambridge, 2017)

0.3 Introduction

Towards the end of the *Linear Algebra I* course, it was explained how a linear map $T: V \to V$ can be represented, with respect to a choice of basis, by a square $n \times n$ matrix where $n = \dim V$. When we make a choice of basis $\{e_1, \ldots, e_n\}$ for V, then a vector $v \in V$ becomes represented by a unique *co-ordinate vector* $(c_1, \ldots, c_n) \in \mathbb{R}^n$ such that

$$v = c_1 e_1 + \cdots + c_n e_n$$

and T becomes represented by the matrix $A = (a_{ij})$ where

$$Te_i = a_{1i} e_1 + \cdots + a_{ni} e_n.$$

Note that the co-ordinates of Te_i are the entries of the ith column of A.

Given the *same* linear map T can be represented by infinitely many different matrices, at least two questions arise:

- What do these different matrices have in common, given they represent the same linear map?
- Is there a best matrix representative – for purposes of computation or comprehension – amongst all these different matrices?

The second question will lead us to a discussion of *eigenvectors* and *diagonalizability*. Should we be able to find a matrix representative that is diagonal, then many calculations will be considerably simpler. *If* this is possible, and it is an 'if', then the linear map is said to be diagonalizable and the vectors in the basis are called eigenvectors. An eigenvector is a non-zero vector v such that $Tv = \lambda v$ for some scalar λ known as the *eigenvalue* of v.

Returning to the first question, we shall find that all the algebraic properties of T apply to each of its matrix representatives. If A and B are two matrices representing T then there is an invertible matrix P such that

$$A = P^{-1} B P.$$

We can then show that each matrix representative has the same determinant, trace, rank, nullity, eigenvalues and functional properties – e.g. T is self-inverse. Any calculation we make, pertaining to the algebra of T, reassuring yields the same answer. The matrix P is a *change of basis matrix* providing an *invertible change of variable*.

However, the same cannot be said of geometric properties of T. In general, an invertible change of variable will alter lengths, angles, areas, volumes etc.. If, say, we wished to change variables to show a curve that isn't in normal form – such as

$$x^2 + xy + y^2 = 1$$

– is in fact an ellipse, and determine its area, then we need to ensure that the area remains invariant under the change of co-ordinates. The matrices that preserve the scalar product – and so preserve angle, distance, area – are the *orthogonal* matrices. That is, $P^{-1} = P^T$. It is an easy check to see that the only matrices which *might* be diagonalized by an orthogonal change of variable are the symmetric matrices. At the end of the course we meet the important *spectral theorem* which shows the converse: symmetric matrices *can* be diagonalized by an orthogonal change of variable.

0.4 Notation

$(\mathbf{v}_1 \,|\, \mathbf{v}_2 \,|\, \cdots \,|\, \mathbf{v}_n)$ denotes the $m \times n$ matrix with columns $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^m_{\text{col}}$
$(\mathbf{r}_1 \,/\, \mathbf{r}_2 \,/\, \cdots \,/\, \mathbf{r}_m)$ denotes the $m \times n$ matrix with rows $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m \in \mathbb{R}^n$
$\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ denotes the canonical basis for $\mathbb{R}^n$.
L_A denotes, for an $m \times n$ matrix A, the map $\mathbb{R}^n_{\text{col}} \to \mathbb{R}^m_{\text{col}}$ given by $\mathbf{x} \mapsto A\mathbf{x}$.
$M_i(\lambda)$ denotes the ERO that multiplies the ith row by $\lambda \neq 0$.
S_{ij} denotes the ERO that swaps the ith and jth rows.
$A_{ij}(\lambda)$ denotes the ERO that adds $\lambda \times$ (row i) to row j.
$\text{diag}(\alpha_1, \ldots, \alpha_n)$ denotes the diagonal $n \times n$ matrix with entries $\alpha_1, \ldots, \alpha_n$.
$[A]_{ij}$ denotes the (i, j)th entry of a matrix A.

1. DETERMINANTS.

1.1 Definitions

A square matrix has a number associated with it called its *determinant*. There are various different ways of introducing determinants, each of which has its advantages but none of which is wholly ideal as will become clearer below. The definition we shall use is an inductive one, defining the determinant of an $n \times n$ matrix in terms of $(n-1) \times (n-1)$ determinants. Quite what the determinant of a matrix signifies will be discussed shortly in Remark 11.

Notation 1 *Given a square matrix A and $1 \leqslant I, J \leqslant n$, we write A_{IJ} for the $(n-1) \times (n-1)$ matrix formed by removing the Ith row and the Jth column from A.*

Example 2 *Let*

$$A = \begin{pmatrix} 1 & -3 & 2 \\ 0 & 7 & 1 \\ -5 & 1 & 3 \end{pmatrix}.$$

Then (a) removing the 2nd row and 3rd column or (b) removing the 3rd row and 1st column, we get

$$\text{(a)} \quad A_{23} = \begin{pmatrix} 1 & -3 \\ -5 & 1 \end{pmatrix}; \qquad \text{(b)} \quad A_{31} = \begin{pmatrix} -3 & 2 \\ 7 & 1 \end{pmatrix}.$$

Our inductive definition of a determinant is then:

Definition 3 *The determinant of a 1×1 matrix (a_{11}) is simply a_{11} itself. The* ***determinant*** $\det A$ *of an $n \times n$ matrix $A = (a_{ij})$ is then given by*

$$\det A = a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31} - \cdots + (-1)^{n+1} a_{n1} \det A_{n1}.$$

Notation 4 *The determinant of a square matrix A is denoted as $\det A$ and also sometimes as $|A|$. So we may also write the determinant of the matrix A in Example 2 as*

$$\begin{vmatrix} 1 & -3 & 2 \\ 0 & 7 & 1 \\ -5 & 1 & 3 \end{vmatrix}.$$

Proposition 5 *The determinants of 2×2 and 3×3 matrices are given by the following formulae.*
(a) For 2×2 matrices

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

(b) For 3×3 *matrices*

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}. \tag{1.1}$$

Proof. (a) Applying the above inductive definition, we have $\det A_{11} = \det(a_{22}) = a_{22}$ and $\det A_{21} = \det(a_{12}) = a_{12}$, so that

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} \det A_{11} - a_{21} \det A_{21} = a_{11}a_{22} - a_{12}a_{21}.$$

(b) For the 3×3 case

$$\det A_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \qquad \det A_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, \qquad \det A_{31} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix},$$

so that

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}) \end{aligned}$$

using the formula for 2×2 determinants. This rearranges to (1.1). ■

Example 6 *Let* R_θ *and* S_θ *be the rotation and reflection matrices*

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad S_\theta = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}.$$

R_θ *represents rotation by* θ *anti-clockwise about the origin and* S_θ *represents reflection in* $y = \tan\theta$. *Note, for any* θ, *that*

$$\det R_\theta = \cos^2\theta + \sin^2\theta = 1, \qquad \det S_\theta = -\cos^2 2\theta - \sin^2 2\theta = -1.$$

Example 7 *Returning to the matrix from Example 2, we have*

$$\begin{aligned} \begin{vmatrix} 1 & -3 & 2 \\ 0 & 7 & 1 \\ -5 & 1 & 3 \end{vmatrix} &= \underbrace{1 \times 7 \times 3}_{21} + \underbrace{(-3) \times 1 \times (-5)}_{15} + \underbrace{2 \times 0 \times 1}_{0} \\ &\quad -\underbrace{1 \times 1 \times 1}_{1} - \underbrace{(-3) \times 0 \times 3}_{0} - \underbrace{2 \times 7 \times (-5)}_{-70} \\ &= 105. \end{aligned}$$

Remark 8 *In the* 2×2 *and* 3×3 *cases,* ***but only in these cases,*** *there is a simple way to remember the determinant formula. The* 2×2 *formula is the product of entries on the left-to-right diagonal minus the product of those on the right-to-left diagonals. If, in the* 3×3 *case, we allow diagonals to 'wrap around' the vertical sides of the matrix – for example as below*

$$\begin{pmatrix} & \searrow & \\ & & \searrow \\ \searrow & & \end{pmatrix}, \qquad \begin{pmatrix} \swarrow & & \\ & & \swarrow \\ & \swarrow & \end{pmatrix},$$

– then from this point of view a 3×3 *matrix has three left-to-right diagonals and three right-to-left. A* 3×3 *determinant then equals the sum of the products of entries on the three left-to-right diagonals minus the products from the three right-to-left diagonals. This method of calculation does* ***not*** *apply to* $n \times n$ *determinants when* $n \geqslant 4$.

Definition 9 *Let* A *be an* $n \times n$ *matrix. Given* $1 \leqslant I, J \leqslant n$ *the* (I, J)*th* ***cofactor*** *of* A*, denoted* $C_{IJ}(A)$ *or just* C_{IJ}*, is defined as* $C_{IJ} = (-1)^{I+J} \det A_{IJ}$ *and so the determinant* $\det A$ *can be rewritten as*

$$\det A = a_{11}C_{11} + a_{21}C_{21} + \cdots + a_{n1}C_{n1}.$$

Proposition 10 *Let* A *be a triangular matrix. Then* $\det A$ *equals the product of the diagonal entries of* A*. In particular it follows that* $\det I_n = 1$ *for any* n*.*

Proof. This is left to Sheet 1, S3. ■

Remark 11 *(**Summary of Determinant's Properties**) As commented earlier, there are different ways to introduce determinants, each with their own particular advantages and disadvantages.*

- *With Definition 3, the determinant of an* $n \times n$ *matrix is at least unambiguously and relatively straightforwardly given. There are other (arguably more natural) definitions which require some initial work to show that they're well-defined. For example, we shall see that* $\det$ *has the following algebraic properties*

  (i) $\det$ *is linear in the rows (or columns) of a matrix (see Theorem 13(A)).*

  (ii) if a matrix has two equal rows then its determinant is zero (see Theorem 13(B)).

  (iii) $\det I_n = 1$.

  In fact, these three algebraic properties uniquely characterize a function $\det$ *which assigns a number to each* $n \times n$ *matrix (Proposition 26). As a consequence of this uniqueness it also follows that*

  (∗) $\det A^T = \det A$ *for any square matrix* A *(see Corollary 20).*

  The problem with the above approach is that the existence and uniqueness of such a function are still moot.

- *Using Definition 3 we avoid these issues, but unfortunately we currently have no real sense of what the determinant might convey about a matrix. The determinant of a* 2×2 *matrix is uniquely characterized by the two following geometric properties. Given a* 2×2 *matrix* A*, with associated map* L_A*, it is then the case that*

  (a) for any region S *of the xy-plane, we have*

$$\text{area of } L_A(S) = |\det A| \times (\text{area of } S). \tag{1.2}$$

  (b) The sense of any angle under L_A *is reversed when* $\det A < 0$ *but remains the same when* $\det A > 0$.

  We can demonstrate (a) and (b) by noting that the Jacobian of

$$L_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$$

  equals

$$\frac{\partial(f_1, f_2)}{\partial(x, y)} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

  These two properties best show the significance of determinants. Thinking along these lines, the following properties should seem natural enough:

  (α) $\det AB = \det A \det B$ *(Corollary 19).*

  (β) a square matrix is singular if and only it has zero determinant (Corollary 18).

  However, whilst these geometric properties might better motivate the importance of determinants, they would be less useful in calculating determinants. Their meaning would also be less clear if we were working in more than three dimensions (at least until we had defined volume and sense/orientation in higher dimensions) or if we were dealing with matrices with complex numbers as entries.

- *The current definition appears to lend some importance to the first column; Definition 3 is sometimes referred to as expansion along the first column. From Sheet 1, P1 one might (rightly) surmise that determinants can be calculated by expanding along any row or column (Theorem 28).*

- *Finally, calculation is difficult and inefficient using Definition 3. (For example, the formula for an* $n \times n$ *determinant involves the sum of* $n!$ *separate products (Propositions 26 and 27(b)). We shall, in due course, see that a much better way to calculate determinants is via EROs. This method works well with specific examples but less well in general as too many special cases arise; if we chose to define determinants this way, even determining the general formulae for* 2×2 *and* 3×3 *determinants would become something of a chore.*

In the following we rigorously develop the theory of determinants. These proofs are often technical and not particularly illuminating and only a selection of the proofs will be covered in lectures. I'd suggest the significant properties of determinants are (i), (ii), (iii), (∗), (a), (b), (α), (β) above and these should be committed to memory. The next significant result (or method) appears in Remark 21 where we begin the discussion of calculating determinants efficiently.

Notation 12 *(a) We shall write* $(\mathbf{r}_1/\cdots/\mathbf{r}_n)$ *for the* $n\times n$ *matrix with rows* $\mathbf{r}_1,\cdots,\mathbf{r}_n\in\mathbb{R}^n$.
(b) We shall write $(\mathbf{v}_1|\cdots|\mathbf{v}_n)$ *for the* $n\times n$ *matrix with columns* $\mathbf{v}_1,\cdots,\mathbf{v}_n\in\mathbb{R}^n_{\text{col}}$.
(c) We shall write $\mathbf{e}_1,\ldots,\mathbf{e}_n$ *for the standard basis of* $\mathbb{R}^n$.

Theorem 13 *The map* det *defined in Definition 3 has the following properties.*

(A) det *is linear in each row. That is,* $\det C=\lambda\det A+\mu\det B$ *where*

$$\begin{aligned} A &= (\mathbf{r}_1/\cdots/\mathbf{r}_{i-1}/\mathbf{r}_i/\mathbf{r}_{i+1}/\cdots/\mathbf{r}_n),\\ B &= (\mathbf{r}_1/\cdots/\mathbf{r}_{i-1}/\mathbf{v}/\mathbf{r}_{i+1}/\cdots/\mathbf{r}_n),\\ C &= (\mathbf{r}_1/\ldots/\mathbf{r}_{i-1}/\lambda\mathbf{r}_i+\mu\mathbf{v}/\mathbf{r}_{i+1}/\ldots/\mathbf{r}_n).\end{aligned}$$

(B) If $A=(\mathbf{r}_1/\cdots/\mathbf{r}_n)$ *with* $\mathbf{r}_i=\mathbf{r}_j$ *for some* $i\neq j$, *then* $\det A=0$.
(B') If the matrix B *is produced by swapping two different rows of* A *then* $\det B=-\det A$.

Before proceeding to the main proof we will first prove the following.

Lemma 14 *Together, properties (A) and (B) are equivalent to properties (A) and (B').*

Proof. Suppose that det has properties (A), (B). Let $A=(\mathbf{r}_1/\cdots/\mathbf{r}_n)$ and B be produced by swapping rows i and j where $i<j$. Then

$$\begin{aligned} 0 &= \det(\mathbf{r}_1/\cdots/\mathbf{r}_i+\mathbf{r}_j/\cdots/\mathbf{r}_i+\mathbf{r}_j/\cdots/\mathbf{r}_n) \qquad \text{[by (B)]}\\ &= \det(\mathbf{r}_1/\cdots/\mathbf{r}_i/\cdots/\mathbf{r}_i+\mathbf{r}_j/\cdots/\mathbf{r}_n)+\det(\mathbf{r}_1/\cdots/\mathbf{r}_j/\cdots/\mathbf{r}_i+\mathbf{r}_j/\cdots/\mathbf{r}_n) \qquad \text{[by (A)]}\\ &= \{\det(\mathbf{r}_1/\cdots/\mathbf{r}_i/\cdots/\mathbf{r}_i/\cdots/\mathbf{r}_n)+\det(\mathbf{r}_1/\cdots/\mathbf{r}_i/\cdots/\mathbf{r}_j/\cdots/\mathbf{r}_n)\}\\ &\quad +\{\det(\mathbf{r}_1/\cdots/\mathbf{r}_j/\cdots/\mathbf{r}_i/\cdots/\mathbf{r}_n)+\det(\mathbf{r}_1/\cdots/\mathbf{r}_j/\cdots/\mathbf{r}_j/\cdots/\mathbf{r}_n)\} \qquad \text{[by (A)]}\\ &= \{0+\det A\}+\{\det B+0\} \qquad \text{[by (B)]}\\ &= \det A+\det B\end{aligned}$$

and so property (B') follows. Conversely, if det has properties (A), (B') and $\mathbf{r}_i=\mathbf{r}_j$ for $i\neq j$ then

$$\det(\mathbf{r}_1/\cdots/\mathbf{r}_i/\cdots/\mathbf{r}_j/\cdots/\mathbf{r}_n)=\det(\mathbf{r}_1/\cdots/\mathbf{r}_j/\cdots/\mathbf{r}_i/\cdots/\mathbf{r}_n),$$

as the two matrices are equal, but by property (B')

$$\det(\mathbf{r}_1/\cdots/\mathbf{r}_i/\cdots/\mathbf{r}_j/\cdots/\mathbf{r}_n)=-\det(\mathbf{r}_1/\cdots/\mathbf{r}_j/\cdots/\mathbf{r}_i/\cdots/\mathbf{r}_n),$$

so that both determinants are in fact zero. ■

We continue now with the proof of Theorem 13.

Proof. (A) If $n=1$ then (A) equates to the identity $(\lambda a_{11}+\mu v_1)=\lambda(a_{11})+\mu(v_1)$. As an inductive hypothesis, suppose that (A) is true for $(n-1)\times(n-1)$ matrices. We are looking to show that the $n\times n$ determinant function is linear in the ith row. Note, for $j\neq i$, that $C_{j1}(C)=\lambda C_{j1}(A)+\mu C_{j1}(B)$ by our inductive hypothesis as these cofactors relate to

$(n-1)\times(n-1)$ determinants. Also $C_{i1}(C)=C_{i1}(A)=C_{i1}(B)$ as C_{i1} is independent of the ith row. Hence

$$\begin{aligned}\det C &= a_{11}C_{11}(C)+\cdots+(\lambda a_{i1}+\mu v_1)C_{i1}(C)+\cdots+a_{n1}C_{n1}(C)\\ &= a_{11}(\lambda C_{11}(A)+\mu C_{11}(B))+\cdots+\lambda a_{i1}C_{i1}(A)+\mu v_1 C_{i1}(B)+\cdots+a_{n1}(\lambda C_{n1}(A)+\mu C_{n1}(B))\\ &= \lambda\{a_{11}C_{11}(A)+\cdots+a_{n1}C_{n1}(A)\}+\mu\{a_{11}C_{11}(B)+\cdots+v_1C_{i1}(B)+\cdots+a_{n1}C_{n1}(B)\}\\ &= \lambda\det A+\mu\det B.\end{aligned}$$

We have therefore proved (A) for all square matrices. In what follows, note that if (B) is true of certain matrices then so is (B') as we have shown that (A) and (B) are equivalent to (A) and (B').

(B) For a 2×2 matrix, if $\mathbf{r}_1=\mathbf{r}_2$ then

$$\det A=\begin{pmatrix} a_{11} & a_{12}\\ a_{11} & a_{12}\end{pmatrix}=a_{11}a_{12}-a_{12}a_{11}=0.$$

So (B) (and hence (B')) hold for 2×2 matrices. Assume now that (B) (or equivalently (B')) is true for $(n-1)\times(n-1)$ matrices. Let $A=(\mathbf{r}_1/\cdots/\mathbf{r}_n)$ with $\mathbf{r}_i=\mathbf{r}_j$ where $i<j$. Then

$$\det A=a_{11}C_{11}(A)+\cdots+a_{n1}C_{n1}(A)=a_{i1}C_{i1}(A)+a_{j1}C_{j1}(A)$$

by the inductive hypothesis as A_{k1} has two equal rows when $k\neq i,j$. Note that as $\mathbf{r}_i=\mathbf{r}_j$, with one copy of each being removed from A_{i1} and A_{j1}, then the rows of A_{i1} are the same as the rows of A_{j1} but come in a different order. The rows of A_{i1} and A_{ji} can be reordered to be the same as follows: what remains of $\mathbf{r}_j$ in A_{i1} can be moved up to the position of $\mathbf{r}_i$'s remainder in A_{j1} by swapping it $j-i-1$ times, each time with the next row above. (Note that we cannot simply swap the rows $\mathbf{r}_i$ and $\mathbf{r}_j$ in A to show $\det A=0$ as this would be assuming (B') for $n\times n$ matrices which is equivalent to what we're trying to prove.) By our inductive hypothesis

$$\begin{aligned}\det A &= a_{i1}C_{i1}(A)+a_{j1}C_{j1}(A)\\ &= (-1)^{1+i}a_{i1}\det A_{i1}+(-1)^{1+j}a_{j1}\det A_{j1} \qquad \text{[by definition of cofactors]}\\ &= (-1)^{1+i}a_{i1}(-1)^{j-i-1}\det A_{j1}+(-1)^{1+j}a_{j1}\det A_{j1} \qquad \text{[by } j-i-1 \text{ uses of (B')]}\\ &= (-1)^{j}(a_{i1}-a_{j1})\det A_{j1}=0 \qquad [\text{as } a_{i1}=a_{j1} \text{ because } \mathbf{r}_i=\mathbf{r}_j].\end{aligned}$$

Hence (B) is true for $n\times n$ determinants and the result follows by induction. ∎

Corollary 15 *Let A be an $n\times n$ matrix and λ a real number.*
(a) If the matrix B is formed by multiplying a row of A by λ then $\det B=\lambda\det A$.
(b) $\det(\lambda A)=\lambda^n\det A$.
(c) If any row of A is zero then $\det A=0$.

Proof. (a) This follows from the fact that det is linear in its rows, and then if (a) is applied consecutively to each of the n rows part (b) follows. Finally if $\mathbf{r}_i=\mathbf{0}$ for some i, then $\mathbf{r}_i=0\mathbf{r}_i$ and so (c) follows from part (a). ∎

Notation 16 *We will denote the three EROs as:*

(a) $M_i(\lambda)$ *denotes multiplication of the* i*th row by* $\lambda \neq 0$.

(b) S_{ij} *denotes swapping the* i*th and* j*th rows.*

(c) $A_{ij}(\lambda)$ *denotes adding* $\lambda \times$ *(row* i*) to row* j.

Lemma 17 *(a) The determinants of the elementary matrices are*

$$\det M_i(\lambda) = \lambda; \qquad \det S_{ij} = -1; \qquad \det A_{ij}(\lambda) = 1.$$

In particular, elementary matrices have non-zero determinants.

(b) If E, A *are* $n \times n$ *matrices and* E *is elementary then* $\det EA = \det E \det A$.

(c) If E *is an elementary matrix then* $\det E^T = \det E$.

Proof. We shall prove (a) and (b) together. If $E = M_i(\lambda)$ and then $\det EA = \lambda \det A$ by Corollary 15(a). If we choose $A = I_n$ then we find $\det M_i(\lambda) = \lambda$ and so we also have $\det EA = \det E \det A$ when $E = M_i(\lambda)$.

If $E = S_{ij}$ then by Theorem 13(B') $\det EA = -\det A$. If we take $A = I_n$ then we see $\det S_{ij} = -1$ and then we also have $\det EA = \det E \det A$ when $E = S_{ij}$.

If $E = A_{ij}(\lambda)$ and $A = (\mathbf{r}_1 / \cdots / \mathbf{r}_n)$ then

$$\begin{aligned}
\det(EA) &= \det(\mathbf{r}_1 / \cdots / \mathbf{r}_i + \lambda \mathbf{r}_j / \cdots / \mathbf{r}_j / \cdots / \mathbf{r}_n) \\
&= \det(\mathbf{r}_1 / \cdots / \mathbf{r}_i / \cdots / \mathbf{r}_j / \cdots / \mathbf{r}_n) + \lambda \det(\mathbf{r}_1 / \cdots / \mathbf{r}_j / \cdots / \mathbf{r}_j / \cdots / \mathbf{r}_n) \qquad \text{[by Theorem 13 (A)]} \\
&= \det A + 0 \qquad \text{[by Theorem 13 (B)]} \\
&= \det A.
\end{aligned}$$

If we take $A = I_n$ then $\det A_{ij}(\lambda) = 1$ and so $\det EA = \det E \det A$ also follows when $E = A_{ij}(\lambda)$.

(c) Note that $M_i(\lambda)$ and S_{ij} are symmetric and so there is nothing to prove in these cases. Finally

$$\det(A_{ij}(\lambda)^T) = \det A_{ji}(\lambda) = 1 = \det A_{ij}(\lambda).$$

■

Corollary 18 ***(Criterion for Invertibility)*** *A square matrix* A *is invertible if and only if* $\det A \neq 0$, *in which case* $\det(A^{-1}) = (\det A)^{-1}$.

Proof. If A is invertible then it row-reduces to the identity; that is, there are elementary matrices $E_1, \ldots, E_k$ such that $E_k \cdots E_1 A = I$. Hence, by repeated use of Lemma 17(b),

$$1 = \det I = \det E_k \times \cdots \times \det E_1 \times \det A.$$

In particular $\det A \neq 0$. Further as $E_k \cdots E_1 = A^{-1}$ then $\det(A^{-1}) = (\det A)^{-1}$. If, however, A is singular then A reduces to a matrix R with at least one zero row so that $\det R = 0$. So as before

$$\det E_k \times \cdots \times \det E_1 \times \det A = \det R = 0$$

for some elementary matrices E_i. As $\det E_i \neq 0$ for each i, it follows that $\det A = 0$. ■

Corollary 19 ***(Product Rule for Determinants)*** *Let A, B be $n \times n$ matrices. Then* $\det AB = \det A \det B$.

Proof. If A and B are invertible then they can be written as products of elementary matrices; say $A = E_1 \ldots E_k$ and $B = F_1 \ldots F_l$. Then

$$\det AB = \det E_1 \times \cdots \times \det E_k \times \det F_1 \times \cdots \times \det F_l = \det A \det B$$

by Lemma 17(b). Otherwise one (or both) of A or B is singular. Then AB is singular and so $\det AB = 0$. But, also $\det A \times \det B = 0$ as one or both of A, B is singular. ■

Corollary 20 ***(Transpose Rule for Determinants)*** *Let A be a square matrix. Then*

$$\det A^T = \det A.$$

Proof. A is invertible if and only if A^T is invertible. If A is invertible then $A = E_1 \ldots E_k$ for some elementary matrices E_i. Now $A^T = E_k^T \cdots E_1^T$ by the product rule for transposes and so, by Lemma 17(c) and the product rule above,

$$\det A^T = \det E_k^T \times \cdots \times \det E_1^T = \det E_k \times \cdots \times \det E_1 = \det A.$$

If A is singular then so is A^T and so $\det A = 0 = \det A^T$. ■

Remark 21 *Corollaries 18, 19, 20, represent the most important algebraic properties of determinants. However we are still lumbered with a very inefficient way of calculating determinants in Definition 3. That definition is practicable up to 3×3 matrices but rapidly becomes laborious after that. A much more efficient way to calculate determinants is using EROs and ECOs, and we have been in a position to do this since showing* $\det EA = \det E \times \det A$ *for elementary E. An ECO involves postmultiplication by an elementary matrix but the product rule shows they will have the same effects on the determinant. Spelling this out:*

- *Adding a multiple of a row (resp. column) to another row (resp. column) has no effect on a determinant.*
- *Multiplying a row or column of the determinant by a scalar λ multiplies the determinant by λ.*
- *Swapping two rows or two columns of a determinant multiplies the determinant by -1.*

The following examples will hopefully make clear how to efficiently calculate determinants using EROs and ECOs.

Example 22 *Use EROs and ECOs to calculate the following 4×4 determinants.*

$$\begin{vmatrix} 1 & 2 & 0 & 3 \\ 4 & -3 & 1 & 0 \\ 0 & 2 & 5 & -1 \\ 2 & 3 & 1 & 2 \end{vmatrix}, \qquad \begin{vmatrix} 2 & 2 & 1 & -3 \\ 0 & 6 & -2 & 1 \\ 3 & 2 & 1 & 1 \\ 4 & 2 & -1 & 2 \end{vmatrix}.$$

Solution.

$$\begin{vmatrix} 1 & 2 & 0 & 3 \\ 4 & -3 & 1 & 0 \\ 0 & 2 & 5 & -1 \\ 2 & 3 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 0 & 3 \\ 0 & -11 & 1 & -12 \\ 0 & 2 & 5 & -1 \\ 0 & -1 & 1 & -4 \end{vmatrix} = \begin{vmatrix} -11 & 1 & -12 \\ 2 & 5 & -1 \\ -1 & 1 & -4 \end{vmatrix}$$

$$= \begin{vmatrix} 0 & -10 & 32 \\ 0 & 7 & -9 \\ -1 & 1 & -4 \end{vmatrix} = -1 \times \begin{vmatrix} -10 & 32 \\ 7 & -9 \end{vmatrix} = 134,$$

where, in order, we (i) add appropriate multiples of the row 1 to lower rows to clear the rest of column 1, (ii) expand along column 1, (iii) add appropriate multiples of row 3 to rows 1 and 2 to clear the rest of column 1, (iv) expand along column 1 and (v) employ the 2×2 determinant formula.

$$\begin{vmatrix} 2 & 2 & 1 & -3 \\ 0 & 6 & -2 & 1 \\ 3 & 2 & 1 & 1 \\ 4 & 2 & -1 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 2 & 1 & -3 \\ 0 & 6 & -2 & 1 \\ 0 & -1 & -\frac{1}{2} & \frac{11}{2} \\ 0 & -2 & -3 & 8 \end{vmatrix} = 2 \begin{vmatrix} 6 & -2 & 1 \\ -1 & -\frac{1}{2} & \frac{11}{2} \\ -2 & -3 & 8 \end{vmatrix}$$

$$= 2 \begin{vmatrix} 0 & -5 & 34 \\ -1 & -\frac{1}{2} & \frac{11}{2} \\ 0 & -2 & -3 \end{vmatrix} = 2 \begin{vmatrix} -5 & 34 \\ -2 & -3 \end{vmatrix} = 166,$$

where, in order, we (i) add appropriate multiples of row 1 to lower rows to clear the rest of column 1, (ii) expand along column 1, (iii) add appropriate multiples of row 2 to rows 1 and 3 to clear the rest of column 1, (iv) expand along column 1 and (v) employ the 2×2 determinant formula.

Alternatively, for this second determinant, it may have made more sense to column-reduce as the third column has a helpful leading 1 and we could have instead calculated the determinant as follows.

$$\begin{vmatrix} 2 & 2 & 1 & -3 \\ 0 & 6 & -2 & 1 \\ 3 & 2 & 1 & 1 \\ 4 & 2 & -1 & 2 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 4 & 10 & -2 & -5 \\ 1 & 0 & 1 & 4 \\ 6 & 4 & -1 & -1 \end{vmatrix} = \begin{vmatrix} 4 & 10 & -5 \\ 1 & 0 & 4 \\ 6 & 4 & -1 \end{vmatrix}$$

$$= \begin{vmatrix} 0 & 10 & -21 \\ 1 & 0 & 4 \\ 0 & 4 & -25 \end{vmatrix} = - \begin{vmatrix} 10 & -21 \\ 4 & -25 \end{vmatrix} = 166.$$

where, in order, we (i) add appropriate multiples of column 3 to other columns to clear the rest of row 1, (ii) expand along row 1 (iii) add appropriate multiples of row 2 to rows 1 and 3 to clear the rest of column 1, (iv) expand along column 1 and (v) employ the 2×2 determinant formula. ■

- We will demonstrate in Theorem 28 the as-yet-unproven equivalence of expanding along *any* row or column.

Example 23 *Let a, x be real numbers. Determine the following 3×3 and $n \times n$ determinants.*

$$(a) \quad \begin{vmatrix} x & a & a \\ x & x & a \\ x & x & x \end{vmatrix}, \qquad (b) \quad \begin{vmatrix} x & 1 & \cdots & 1 \\ 1 & x & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & x \end{vmatrix}.$$

Solution. (a) Subtracting row 3 from the other rows, and expanding along column 1, we obtain

$$\begin{vmatrix} x & a & a \\ x & x & a \\ x & x & x \end{vmatrix} = \begin{vmatrix} 0 & a-x & a-x \\ 0 & 0 & a-x \\ x & x & x \end{vmatrix} = x \begin{vmatrix} a-x & a-x \\ 0 & a-x \end{vmatrix} = x(a-x)^2.$$

Similarly for (b) if we note that the sum of each column is the same and then add the bottom $n-1$ rows to the first row (which won't affect the determinant), we see it equals

$$\begin{aligned} & \begin{vmatrix} x+n-1 & x+n-1 & \cdots & x+n-1 \\ 1 & x & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & x \end{vmatrix} \\ = \quad & (x+n-1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & x & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & x \end{vmatrix} \\ = \quad & (x+n-1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & x-1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x-1 \end{vmatrix} \end{aligned}$$

where, in order, we (i) take the common factor of $x+n-1$ out of the first row, (ii) subtract the first row from each of the other rows, (iii) note the determinant is upper triangular to finally obtain a result of $(x+n-1)(x-1)^{n-1}$. ■

We conclude this section by defining the Vandermonde[1] determinant useful in *interpolation*.

Example 24 ***(Vandermonde Matrix)*** *For $n \geqslant 2$ and real numbers $x_1, \ldots, x_n$ we define*

$$V_n = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} \qquad \text{and then} \qquad \det V_n = \prod_{i>j} (x_i - x_j).$$

In particular, V_n is invertible if and only if the x_i are distinct.

Solution. This is left to Sheet 1, Exercise 5. ■

[1] After the French mathematician Alexandre-Théophile Vandermonde (1735-1796).

1.2 Permutation Matrices

It was claimed in Remark 11 that the determinant function for $n \times n$ matrices is entirely determined by certain algebraic properties. In light of Lemma 14, these properties are equivalent to

(i) det is linear in the rows of a matrix.

(ii) if a matrix has two equal rows then its determinant is zero.

(ii)' if the matrix B is produced by swapping two of the rows of A then $\det B = -\det A$.

(iii) $\det I_n = 1$.

To see why these properties determine det, we first consider the $n = 2$ case. Given a 2×2 matrix $A = (a_{ij})$, we can calculate its determinant as follows. As det is linear in row 1 then

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

which, as det is linear in row 2, equals

$$\left\{ \begin{vmatrix} a_{11} & 0 \\ 0 & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 \\ a_{21} & 0 \end{vmatrix} \right\} + \left\{ \begin{vmatrix} 0 & a_{12} \\ a_{21} & 0 \end{vmatrix} + \begin{vmatrix} 0 & a_{12} \\ 0 & a_{22} \end{vmatrix} \right\}.$$

Again as det is linear in rows the above equals

$$a_{11}a_{22} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + a_{11}a_{21} \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} + a_{12}a_{21} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} + a_{12}a_{22} \begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix}.$$

Then, using (ii), this equals

$$a_{11}a_{22} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + a_{12}a_{21} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$$

which, using (ii)', equals

$$a_{11}a_{22} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} - a_{12}a_{21} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}.$$

Finally, using (iii), we've shown

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} == a_{11}a_{22} - a_{12}a_{21}$$

If we were to argue similarly for a 3×3 matrix $A = (a_{ij})$, we could first use linearity to expand the determinant into a linear combination of $3^3 = 27$ determinants, with entries 1 and 0, each multiplied by a *monomial* $a_{1i}a_{2j}a_{3k}$. But we can ignore those cases where i, j, k involves some repetition as the corresponding determinant is zero. There would, in fact, be only $3! = 6$ non-zero contributions giving us the formula

$$a_{11}a_{22}a_{23}\begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix}+a_{12}a_{23}a_{31}\begin{vmatrix}0&1&0\\0&0&1\\1&0&0\end{vmatrix}+a_{13}a_{21}a_{32}\begin{vmatrix}0&0&1\\1&0&0\\0&1&0\end{vmatrix}$$
$$+a_{12}a_{21}a_{33}\begin{vmatrix}0&1&0\\1&0&0\\0&0&1\end{vmatrix}+a_{13}a_{22}a_{31}\begin{vmatrix}0&0&1\\0&1&0\\1&0&0\end{vmatrix}+a_{11}a_{23}a_{32}\begin{vmatrix}1&0&0\\0&0&1\\0&1&0\end{vmatrix}.$$

The first determinant here is $\det I_3$ which we know to be 1. The other determinants all have the same rows $(1,0,0),(0,1,0),(0,0,1)$ as I_3 but appearing in some other order. In each case, it is possible (if necessary) to swap $(1,0,0)$ – which appears as some row of the determinant – with the first row, so that it is now in the correct place. Likewise the second row can be moved (if necessary) so it is in the right place. By a process of elimination the third row is now in the right place and we have transformed the determinant into $\det I_3$. We know what the effect of each such swap is, namely multiplying by -1, and so the six determinants above have values 1 or -1. For example,

$$\begin{vmatrix}0&1&0\\0&0&1\\1&0&0\end{vmatrix}=-\begin{vmatrix}1&0&0\\0&0&1\\0&1&0\end{vmatrix}=\begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix}=1,\qquad \begin{vmatrix}0&0&1\\0&1&0\\1&0&0\end{vmatrix}=-\begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix}=-1.$$

So finally we have, as we found in Proposition 5(b), that

$$\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{23}&a_{33}\end{vmatrix}=a_{11}a_{22}a_{23}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}-a_{11}a_{23}a_{32}.$$

The general situation is hopefully now clear for an $n\times n$ matrix $A=(a_{ij})$. Using linearity to expand along each row in turn, $\det A$ can be written as the sum of n^n terms

$$\sum \det P_{i_1\cdots i_n}a_{1i_1}\cdots a_{ni_n}$$

where $P_{i_1\cdots i_n}$ is the matrix whose rows are $\mathbf{e}_{i_1},\ldots,\mathbf{e}_{i_n}$ – that is the entries of $P_{i_1\cdots i_n}$ are all zero except entries $(1,i_1),\ldots,(n,i_n)$ which are all 1. At the moment each of $i_1,\ldots,i_n$ can independently take a value between 1 and n, but most such choices lead to the determinant $\det P_{i_1\cdots i_n}$ being zero as some of the rows $\mathbf{e}_{i_1},\ldots,\mathbf{e}_{i_n}$ are repeated. In fact, $\det P_{i_1\cdots i_n}$ can only be non-zero when

$$\{i_1,\ldots,i_n\}=\{1,\ldots,n\}.$$

That is $i_1,\ldots,i_n$ are $1,\ldots n$ *in some order* or equivalently the rows of $P_{i_1\cdots i_n}$ are $\mathbf{e}_1,\ldots,\mathbf{e}_n$ *in some order.*

Definition 25 *An $n\times n$ matrix P is said to be a* ***permutation matrix*** *if its rows are $\mathbf{e}_1,\ldots,\mathbf{e}_n$ in some order. This is equivalent to saying that each row and column contains a single entry 1 with all other entries being zero.*

Thus we have shown:

Proposition 26 *The function* det *is entirely determined by the three algebraic properties (i), (ii) and (iii). Further, the determinant* $\det A$ *of an* $n \times n$ *matrix* $A = (a_{ij})$ *equals*

$$\det A = \sum \det P_{i_1 \cdots i_n}\, a_{1i_1} \cdots\, a_{ni_n} \tag{1.3}$$

where the sum is taken over all permutation matrices $P_{i_1 \cdots i_n} = (\mathbf{e}_{i_1}/\cdots/\mathbf{e}_{i_n})$.

We further note:

Proposition 27 *(a) The columns of a permutation matrix are* $\mathbf{e}_1^T, \ldots, \mathbf{e}_n^T$ *in some order.*

(b) The number of $n \times n$ *permutation matrices is* $n!$*.*

(c) A permutation matrix has determinant 1 *or* -1*.*

(d) When $n \geqslant 2$*, half the permutation matrices have determinant* 1 *and half have determinant* -1*.*

Proof. (a) The entries in the first column of a permutation matrix P are the first entries of $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n$ in some order and so are $1, 0, \ldots 0$ in some order – that is the first column is $\mathbf{e}_i^T$ for some i. Likewise each column of P is $\mathbf{e}_i^T$ for some i. If any of the columns of P were the same then this would mean that a row of P had two non-zero entries which cannot occur. So the columns are all distinct. As there are n columns then each of $\mathbf{e}_1^T, \ldots, \mathbf{e}_n^T$ appears exactly once.

(b) This is equal to the number of bijections from the set $\{1, 2, \ldots, n\}$ to itself.

(c) The rows of a permutation matrix P are $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in some order. We know that swapping two rows of a matrix has the effect of multiplying the determinant by -1. We can create a (possibly new) matrix P_1 by swapping the first row of P with the row $\mathbf{e}_1$ (which appears somewhere); of course no swap may be needed. The matrix P_1 has $\mathbf{e}_1$ as its first row and $\det P_1 = \pm \det P$ depending on whether a swap was necessary or not. We can continue in this fashion producing matrices $P_1, \ldots, P_n$ such that the first k rows of P_k are $\mathbf{e}_1, \ldots, \mathbf{e}_k$ in *that* order and $\det P_k = \pm \det P_{k-1}$ in each case, depending on whether or not we needed to make any swap to get $\mathbf{e}_k$ to the kth row. Eventually then $P_n = I_n$ and $\det P = \pm \det P_n = 1$ or -1 depending on whether an even or odd number of swaps had to be made to turn P into I_n.

(d) Let $n \geqslant 2$ and let S_{12} be the elementary $n \times n$ matrix associated with swapping the first and second rows of a matrix. If P is a permutation matrix then $S_{12}P$ is also a permutation matrix as its rows are still $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in some order; further

$$\det S_{12}P = \det S_{12} \times \det P = -\det P.$$

For each permutation matrix P with $\det P = 1$, we have $S_{12}P$ being a permutation matrix with $\det(S_{12}P) = -1$; conversely for every permutation matrix $\tilde{P}$ with $\det \tilde{P} = -1$ we have $S_{12}\tilde{P}$ being a permutation matrix with $\det(S_{12}\tilde{P}) = 1$ As these processes are inverses of one another, because $S_{12}(S_{12}P) = P$, there are equal numbers of determinant 1 and determinant -1 permutation matrices, each separately numbering $\frac{1}{2}n!$. ■

We now prove a result already mentioned in Remark 11. Our inductive definition of the determinant began by expanding down the first column. In fact it is the case that we will arrive at the same answer, the determinant, whichever column or row we expand along.

Theorem 28 ***(Equality of determinant expansions[2])*** *Let $A = (a_{ij})$ be an $n \times n$ matrix and let C_{ij} denote the (i,j)th cofactor of A. Then the determinant* $\det A$ *may be calculated by expanding along any column or row of A. So, for any $1 \leqslant i \leqslant n$, we have*

$$\begin{aligned} \det A &= a_{1i}C_{1i} + a_{2i}C_{2i} + \cdots + a_{ni}C_{ni} \qquad \textit{[this is expansion along the } i\textit{th column]} \quad (1.4) \\ &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \qquad \textit{[this is expansion along the } i\textit{th row].} \quad (1.5) \end{aligned}$$

Proof. We showed in Theorem 13 and Proposition 10 that det has properties (i), (ii), (iii), and have just shown in Proposition 26 that these properties uniquely determine the function det. Making the obvious changes to Theorem 13 and Proposition 10 it can similarly be shown, for any i, that the function which assigns

$$a_{1i}C_{1i} + a_{2i}C_{2i} + \cdots + a_{ni}C_{ni} \tag{1.6}$$

to the matrix $A = (a_{ij})$ also has properties (i), (ii), (iii). By uniqueness it follows that (1.6) also equals $\det A$. That is, expanding down any column also leads to the same answer of $\det A$. Then

$$\begin{aligned} \det A = \det A^T &= [A^T]_{1i}C_{1i}(A^T) + [A^T]_{2i}C_{2i}(A^T) + \cdots + [A^T]_{ni}C_{ni}(A^T) \\ &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \end{aligned}$$

by expanding down the ith column of A^T, but this is the same sum found when expanding along the ith row of A. ∎

In practical terms, however, Laplace's result isn't that helpful. We have already discounted repeated expansion along rows and columns of hard-to-calculate cofactors as a hugely inefficient means to find determinants (see Remarks 11 and 21). However, it does lead us to the following theorem of interest.

Theorem 29 ***(Existence of the Adjugate)*** *Let A be an $n \times n$ matrix. Let C_{ij} denote the (i,j)th cofactor of A and let $C = (C_{ij})$ be the matrix of cofactors. Then*

$$C^T A = AC^T = \det A \times I_n.$$

In particular, if A is invertible, then

$$A^{-1} = \frac{C^T}{\det A} \tag{1.7}$$

Proof. Note

$$[C^T A]_{ij} = \sum_{k=1}^{n} [C^T]_{ik}[A]_{kj} = \sum_{k=1}^{n} C_{ki}a_{kj}.$$

When $i = j$ then

$$[C^T A]_{ii} = \sum_{k=1}^{n} a_{ki}C_{ki} = \det A$$

[2]This was proved by Pierre-Simon Laplace (1749-1827) in 1772, though Leibniz had been aware of this result a century earlier.

by Theorem 28 as this is the determinant calculated by expanding along the ith column. On the other hand, if $i \neq j$, then consider the matrix B which has the same columns as A except for the ith column of B which is a copy of A's jth column. As the ith and jth columns of B are equal then $\det B$ is zero. Note that the (k,i)th cofactor of B equals C_{ki} as A and B agree except in the ith column; so if expanding $\det B$ along its ith column we see

$$0 = \det B = \sum_{k=1}^{n} b_{ki} C_{ki} = \sum_{k=1}^{n} a_{kj} C_{ki} = [C^T A]_{ij}.$$

So $C^T A = \det A \times I_n$. That $AC^T = \det A \times I_n$ similarly follows. Finally if A is invertible, then $\det A \neq 0$, and (1.7) follows. ■

Definition 30 *With notation as in Theorem 29 the matrix C^T is called the **adjugate** of A (or sometimes the adjoint of A) and is written* adj$A$.

Corollary 31 ***(Cramer's Rule**[3]**)*** *Let A be an $n \times n$ matrix, $\mathbf{b}$ in $\mathbb{R}^n_{\text{col}}$ and consider the linear system $(A|\mathbf{b})$. The system has a unique solution if and only if $\det A \neq 0$, which is given by*

$$\mathbf{x} = \frac{C^T \mathbf{b}}{\det A}.$$

Proof. There is a solution if and only if L_A is onto which is then unique if and only if the kernel is trivial. That is A is invertible and the result follows from the previous theorem and Corollary 18. ■

Remark 32 *Writing $A = (a_{ij})$ and $\mathbf{b} = (b_1, \ldots, b_n)^T$ then Cramer's Rule with $n = 2$ expressly reads as*

$$x_1 = \frac{b_1 a_{22} - b_2 a_{12}}{\det A}, \qquad x_2 = \frac{b_2 a_{11} - b_1 a_{21}}{\det A},$$

where $\det A = a_{11}a_{22} - a_{12}a_{21}$*. When $n = 3$ Cramer's rule reads as*

$$x_1 = \frac{b_1 \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - b_2 \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + b_3 \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}}{\det A},$$

$$x_2 = \frac{-b_1 \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + b_2 \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - b_3 \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}}{\det A},$$

$$x_3 = \frac{b_1 \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} - b_2 \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} + b_3 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}{\det A},$$

where $\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}$.

Cramer's rule though is a seriously limited and impractical means of solving linear systems. The rule only applies when the matrix A is square and invertible, and the computational power required to calculate so many cofactors and $\det A$ make it substantially more onerous than row-reduction.

[3]Named after the Swiss mathematician, Gabriel Cramer (1704-1752), who discovered this result in 1750.

1.3 Determinants of Linear Maps

Definition 33 *Let $T\colon V \to V$ be a linear map of a finite dimensional vector space V. Then the determinant of T is defined by*

$$\det T = \det A$$

where A is a matrix representing T with respect to some basis for V.

Proposition 34 *(a) The determinant of a linear map is well-defined.*

(b) If $S\colon V \to V$ is a second linear map then

$$\det(ST) = \det S \times \det T.$$

(c) $T\colon V \to V$ is invertible if and only if $\det T \neq 0$. If T is invertible then

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

Proof. (a) As T may have many different matrix representatives, it is possible that different representatives might have different determinants. However any two representatives, A and B, of T are similar matrices so that $A = P^{-1}BP$ for some invertible matrix P. Specifically if

$$A = {}_{\mathcal{E}}T_{\mathcal{E}} \qquad \text{and} \qquad B = {}_{\mathcal{F}}T_{\mathcal{F}}$$

then $A = P^{-1}BP$ where $P = {}_{\mathcal{F}}I_{\mathcal{E}}$. By the product rule for determinants we have

$$\det A = \det\left(P^{-1}BP\right) = \frac{1}{\det P} \times \det B \times \det P = \det B,$$

and hence each matrix representative has the same determinant.

(b) Say that S and T are represented by A and B wrt the same basis for V. Then ST is represented by AB wrt the same basis. So, by the product rule,

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

(c) Say that T is invertible. Then there is a linear map $S\colon V \to V$ such that $ST = I = TS$. So

$$1 = \det I = \det S \times \det T,$$

showing $\det T \neq 0$. Conversely say that $\det T \neq 0$. Let A be a matrix representating T wrt some basis. So $\det A \neq 0$ and A is an invertible matrix. Let $S\colon V \to V$ be the linear map represented by A^{-1} wrt the same basis. Then ST is represented by $A^{-1}A = I$ wrt this basis, and TS is represented by $AA^{-1} = I$ wrt this basis. But the identity matrix represents the identity map wrt all bases and so

$$ST = I = TS.$$

Thus $S = T^{-1}$ and T is invertible. Finally, when T is invertible, we have

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

and the result follows. ■

Example 35 *Let* $V = \langle 1, x, x^2 \rangle$ *be the space of real polynomials in* x *of degree at most* 2. *Define* $D, T \colon V \to V$ *by*

$$(Df)(x) = f'(x), \qquad (Tf)(x) = f(x+1).$$

Evaluate $\det D$ *and* $\det T$.

Solution. As $D(1) = 0$ then D is not invertible and so $\det D = 0$. Alternatively the matrix for D wrt $\{1, x, x^2\}$ is

$$D = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

and $\det D = 0^3 = 0$ as this is an upper triangular matrix.

Now the matrix for T wrt the same basis is

$$T = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

and so $\det T = 1^3 = 1$. ■