Chapter Four

Determinants

In the first chapter we highlighted the special case of linear systems with the same number of equations as unknowns, those of the form $T\vec{x} = \vec{b}$ where T is a square matrix. We noted that there are only two kinds of T's. If T is associated with a unique solution for any $\vec{b}$, such as for the homogeneous system $T\vec{x} = \vec{0}$, then T is associated with a unique solution for every such $\vec{b}$. We call such a matrix nonsingular. The other kind of T, where every linear system for which it is the matrix of coefficients has either no solution or infinitely many solutions, we call singular.

In our work since then this distinction has been a theme. For instance, we now know that an $n \times n$ matrix T is nonsingular if and only if each of these holds:

- any system $T\vec{x} = \vec{b}$ has a solution and that solution is unique;
- Gauss-Jordan reduction of T yields an identity matrix;
- the rows of T form a linearly independent set;
- the columns of T form a linearly independent set, a basis for $\mathbb{R}^n$;
- any map that T represents is an isomorphism;
- an inverse matrix T^{-1} exists.

So when we look at a square matrix, one of the first things that we ask is whether it is nonsingular.

This chapter develops a formula that determines whether T is nonsingular. More precisely, we will develop a formula for 1×1 matrices, one for 2×2 matrices, etc. These are naturally related; that is, we will develop a family of formulas, a scheme that describes the formula for each size.

Since we will restrict the discussion to square matrices, in this chapter we will often simply say 'matrix' in place of 'square matrix'.

I Definition

Determining nonsingularity is trivial for 1×1 matrices.

$$\begin{pmatrix} a \end{pmatrix} \quad \text{is nonsingular iff} \quad a \neq 0$$

Corollary Three.IV.4.11 gives the 2×2 formula.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{is nonsingular iff} \quad ad - bc \neq 0$$

We can produce the 3×3 formula as we did the prior one, although the computation is intricate (see Exercise 9).

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \quad \text{is nonsingular iff} \quad aei + bfg + cdh - hfa - idb - gec \neq 0$$

With these cases in mind, we posit a family of formulas: a, $ad-bc$, etc. For each n the formula defines a *determinant* function $\det_{n\times n}\colon \mathcal{M}_{n\times n} \to \mathbb{R}$ such that an $n\times n$ matrix T is nonsingular if and only if $\det_{n\times n}(T) \neq 0$. (We usually omit the subscript $n\times n$ because the size of T describes which determinant function we mean.)

I.1 Exploration

This subsection is an optional motivation and development of the general definition. The definition is in the next subsection.

Above, in each case the matrix is nonsingular if and only if some formula is nonzero. But the three formulas don't show an obvious pattern. We may spot that the 1×1 term a has one letter, that the 2×2 terms ad and bc have two letters, and that the 3×3 terms each have three letters. We may even spot that in those terms there is a letter from each row and column of the matrix, e.g., in the cdh term one letter comes from each row and from each column.

$$\begin{pmatrix} & & c \\ d & & \\ & h & \end{pmatrix}$$

But these observations are perhaps more puzzling than enlightening. For instance, we might wonder why some terms are added but some are subtracted.

A good strategy for solving problems is to explore which properties the solution must have, and then search for something with those properties. So we shall start by asking what properties we'd like the determinant formulas to have.

At this point, our main way to decide whether a matrix is singular or not is to do Gaussian reduction and then check whether the diagonal of the echelon form matrix has any zeroes, that is, whether the product down the diagonal is zero. So we could guess that whatever determinant formula we find, the proof that it is right may involve applying Gauss's Method to the matrix to show that in the end the product down the diagonal is zero if and only if our formula gives zero.

This suggests a plan: we will look for a family of determinant formulas that are unaffected by row operations and such that the determinant of an echelon form matrix is the product of its diagonal entries. In the rest of this subsection we will test this plan against the 2×2 and 3×3 formulas. In the end we will have to modify the "unaffected by row operations" part, but not by much.

First we check whether the 2×2 and 3×3 formulas are unaffected by the row operation of combining: if

$$T \xrightarrow{k\rho_i+\rho_j} \hat{T}$$

then is $\det(\hat{T}) = \det(T)$? This check of the 2×2 determinant after the $k\rho_1 + \rho_2$ operation

$$\det(\begin{pmatrix} a & b \\ ka+c & kb+d \end{pmatrix}) = a(kb+d) - (ka+c)b = ad - bc$$

shows that it is indeed unchanged, and the other 2×2 combination $k\rho_2+\rho_1$ gives the same result. Likewise, the 3×3 combination $k\rho_3 + \rho_2$ leaves the determinant unchanged

$$\begin{aligned} \det(\begin{pmatrix} a & b & c \\ kg+d & kh+e & ki+f \\ g & h & i \end{pmatrix}) &= a(kh+e)i + b(ki+f)g + c(kg+d)h \\ &\quad - h(ki+f)a - i(kg+d)b - g(kh+e)c \\ &= aei + bfg + cdh - hfa - idb - gec \end{aligned}$$

as do the other 3×3 row combination operations.

So there seems to be promise in the plan. Of course, perhaps if we had worked out the 4×4 determinant formula and tested it then we might have found that it is affected by row combinations. This is an exploration and we do not yet have all the facts. Nonetheless, so far, so good.

Next we compare $\det(\hat{T})$ with $\det(T)$ for row swaps. Here we hit a snag: the

$2{\times}2$ row swap $\rho_1 \leftrightarrow \rho_2$ does not yield $ad - bc$.

$$\det(\begin{pmatrix} c & d \\ a & b \end{pmatrix}) = bc - ad$$

And this $\rho_1 \leftrightarrow \rho_3$ swap inside of a $3{\times}3$ matrix

$$\det(\begin{pmatrix} g & h & i \\ d & e & f \\ a & b & c \end{pmatrix}) = gec + hfa + idb - bfg - cdh - aei$$

also does not give the same determinant as before the swap since again there is a sign change. Trying a different $3{\times}3$ swap $\rho_1 \leftrightarrow \rho_2$

$$\det(\begin{pmatrix} d & e & f \\ a & b & c \\ g & h & i \end{pmatrix}) = dbi + ecg + fah - hcd - iae - gbf$$

also gives a change of sign.

So row swaps appear in this experiment to change the sign of a determinant. This does not wreck our plan entirely. We hope to decide nonsingularity by considering only whether the formula gives zero, not by considering its sign. Therefore, instead of expecting determinant formulas to be entirely unaffected by row operations we modify our plan so that on a swap they will change sign.

Obviously we finish by comparing $\det(\hat{T})$ with $\det(T)$ for the operation of multiplying a row by a scalar. This

$$\det(\begin{pmatrix} a & b \\ kc & kd \end{pmatrix}) = a(kd) - (kc)b = k \cdot (ad - bc)$$

ends with the entire determinant multiplied by k, and the other $2{\times}2$ case has the same result. This $3{\times}3$ case ends the same way

$$\begin{aligned} \det(\begin{pmatrix} a & b & c \\ d & e & f \\ kg & kh & ki \end{pmatrix}) &= ae(ki) + bf(kg) + cd(kh) \\ &\qquad -(kh)fa - (ki)db - (kg)ec \\ &= k \cdot (aei + bfg + cdh - hfa - idb - gec) \end{aligned}$$

as do the other two $3{\times}3$ cases. These make us suspect that multiplying a row by k multiplies the determinant by k. As before, this modifies our plan but does not wreck it. We are asking only that the zero-ness of the determinant formula be unchanged, not focusing on the its sign or magnitude.

So in this exploration out plan got modified in some inessential ways and is now: we will look for $n{\times}n$ determinant functions that remain unchanged under the operation of row combination, that change sign on a row swap, that rescale on the rescaling of a row, and such that the determinant of an echelon form matrix is the product down the diagonal. In the next two subsections we will see that for each n there is one and only one such function.

Finally, for the next subsection note that factoring out scalars is a row-wise operation: here

$$\det(\begin{pmatrix} 3 & 3 & 9 \\ 2 & 1 & 1 \\ 5 & 11 & -5 \end{pmatrix}) = 3 \cdot \det(\begin{pmatrix} 1 & 1 & 3 \\ 2 & 1 & 1 \\ 5 & 11 & -5 \end{pmatrix})$$

the 3 comes only out of the top row only, leaving the other rows unchanged. Consequently in the definition of determinant we will write it as a function of the rows $\det(\vec{\rho}_1, \vec{\rho}_2, \ldots \vec{\rho}_n)$, rather than as $\det(T)$ or as a function of the entries $\det(t_{1,1}, \ldots, t_{n,n})$.

Exercises

✓ **1.1** Evaluate the determinant of each.

(a) $\begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 4 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 3 & -1 \end{pmatrix}$

1.2 Evaluate the determinant of each.

(a) $\begin{pmatrix} 2 & 0 \\ -1 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 1 & 1 \\ 0 & 5 & -2 \\ 1 & -3 & 4 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{pmatrix}$

✓ **1.3** Verify that the determinant of an upper-triangular 3×3 matrix is the product down the diagonal.

$$\det(\begin{pmatrix} a & b & c \\ 0 & e & f \\ 0 & 0 & i \end{pmatrix}) = aei$$

Do lower-triangular matrices work the same way?

✓ **1.4** Use the determinant to decide if each is singular or nonsingular.

(a) $\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$ (c) $\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$

1.5 Singular or nonsingular? Use the determinant to decide.

(a) $\begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 2 \\ 0 & 1 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 4 & 1 & 3 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 1 & 0 \\ 3 & -2 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

✓ **1.6** Each pair of matrices differ by one row operation. Use this operation to compare $\det(A)$ with $\det(B)$.

(a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ $B = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$

(b) $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ $B = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$

(c) $A=\begin{pmatrix}1&-1&3\\2&2&-6\\1&0&4\end{pmatrix}$ $B=\begin{pmatrix}1&-1&3\\1&1&-3\\1&0&4\end{pmatrix}$

1.7 Show this.

$$\det(\begin{pmatrix}1&1&1\\a&b&c\\a^2&b^2&c^2\end{pmatrix})=(b-a)(c-a)(c-b)$$

✓ 1.8 Which real numbers x make this matrix singular?

$$\begin{pmatrix}12-x&4\\8&8-x\end{pmatrix}$$

1.9 Do the Gaussian reduction to check the formula for 3×3 matrices stated in the preamble to this section.

$$\begin{pmatrix}a&b&c\\d&e&f\\g&h&i\end{pmatrix}\text{ is nonsingular iff } aei+bfg+cdh-hfa-idb-gec\neq0$$

1.10 Show that the equation of a line in $\mathbb{R}^2$ thru (x_1,y_1) and (x_2,y_2) is given by this determinant.

$$\det(\begin{pmatrix}x&y&1\\x_1&y_1&1\\x_2&y_2&1\end{pmatrix})=0\qquad x_1\neq x_2$$

✓ 1.11 Many people know this mnemonic for the determinant of a 3×3 matrix: first repeat the first two columns and then sum the products on the forward diagonals and subtract the products on the backward diagonals. That is, first write

$$\left(\begin{array}{ccc|cc}h_{1,1}&h_{1,2}&h_{1,3}&h_{1,1}&h_{1,2}\\h_{2,1}&h_{2,2}&h_{2,3}&h_{2,1}&h_{2,2}\\h_{3,1}&h_{3,2}&h_{3,3}&h_{3,1}&h_{3,2}\end{array}\right)$$

and then calculate this.

$$\begin{aligned}&h_{1,1}h_{2,2}h_{3,3}+h_{1,2}h_{2,3}h_{3,1}+h_{1,3}h_{2,1}h_{3,2}\\&-h_{3,1}h_{2,2}h_{1,3}-h_{3,2}h_{2,3}h_{1,1}-h_{3,3}h_{2,1}h_{1,2}\end{aligned}$$

(a) Check that this agrees with the formula given in the preamble to this section.
(b) Does it extend to other-sized determinants?

1.12 The *cross product* of the vectors

$$\vec{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\qquad\vec{y}=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$$

is the vector computed as this determinant.

$$\vec{x}\times\vec{y}=\det(\begin{pmatrix}\vec{e}_1&\vec{e}_2&\vec{e}_3\\x_1&x_2&x_3\\y_1&y_2&y_3\end{pmatrix})$$

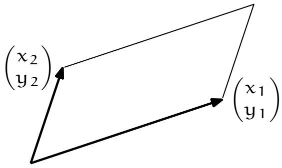
Note that the first row's entries are vectors, the vectors from the standard basis for $\mathbb{R}^3$. Show that the cross product of two vectors is perpendicular to each vector.

1.13 Prove that each statement holds for 2×2 matrices.
(a) The determinant of a product is the product of the determinants $\det(ST)=\det(S)\cdot\det(T)$.

(b) If T is invertible then the determinant of the inverse is the inverse of the determinant $\det(T^{-1}) = (\det(T))^{-1}$.

Matrices T and T' are *similar* if there is a nonsingular matrix P such that $T' = PTP^{-1}$. (We shall look at this relationship in Chapter Five.) Show that similar 2×2 matrices have the same determinant.

✓ 1.14 Prove that the area of this region in the plane



is equal to the value of this determinant.

$$\det(\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix})$$

Compare with this.

$$\det(\begin{pmatrix} x_2 & x_1 \\ y_2 & y_1 \end{pmatrix})$$

1.15 Prove that for 2×2 matrices, the determinant of a matrix equals the determinant of its transpose. Does that also hold for 3×3 matrices?

✓ 1.16 Is the determinant function linear — is $\det(x\cdot T + y\cdot S) = x\cdot\det(T) + y\cdot\det(S)$?

1.17 Show that if A is 3×3 then $\det(c\cdot A) = c^3\cdot\det(A)$ for any scalar c.

1.18 Which real numbers θ make

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

singular? Explain geometrically.

? 1.19 [Am. Math. Mon., Apr. 1955] If a third order determinant has elements 1, 2, ..., 9, what is the maximum value it may have?

I.2 Properties of Determinants

We want a formula to determine whether an $n\times n$ matrix is nonsingular. We will not begin by stating such a formula. Instead we will begin by considering the function that such a formula calculates. We will define this function by its properties, then prove that the function with these properties exists and is unique, and also describe how to compute it. (Because we will eventually show that the function exists and is unique, from the start we will just say '$\det(T)$' instead of 'if there is a unique determinant function then $\det(T)$'.)

2.1 Definition A $n \times n$ *determinant* is a function $\det\colon \mathcal{M}_{n\times n} \to \mathbb{R}$ such that

(1) $\det(\vec{\rho}_1, \ldots, k\cdot\vec{\rho}_i + \vec{\rho}_j, \ldots, \vec{\rho}_n) = \det(\vec{\rho}_1, \ldots, \vec{\rho}_j, \ldots, \vec{\rho}_n)$ for $i \neq j$

(2) $\det(\vec{\rho}_1, \ldots, \vec{\rho}_j, \ldots, \vec{\rho}_i, \ldots, \vec{\rho}_n) = -\det(\vec{\rho}_1, \ldots, \vec{\rho}_i, \ldots, \vec{\rho}_j, \ldots, \vec{\rho}_n)$ for $i \neq j$

(3) $\det(\vec{\rho}_1, \ldots, k\vec{\rho}_i, \ldots, \vec{\rho}_n) = k \cdot \det(\vec{\rho}_1, \ldots, \vec{\rho}_i, \ldots, \vec{\rho}_n)$ for any scalar k

(4) $\det(I) = 1$ where I is an identity matrix

(the $\vec{\rho}$'s are the rows of the matrix). We often write $|T|$ for $\det(T)$.

2.2 Remark Condition (2) is redundant since

$$T \xrightarrow{\rho_i+\rho_j} \xrightarrow{-\rho_j+\rho_i} \xrightarrow{\rho_i+\rho_j} \xrightarrow{-\rho_i} \hat{T}$$

swaps rows i and j. We have listed it just for consistency with the Gauss's Method presentation in earlier chapters.

2.3 Remark Condition (3) does not have a $k \neq 0$ restriction, although the Gauss's Method operation of multiplying a row by k does have it. The next result shows that we do not need that restriction here.

2.4 Lemma A matrix with two identical rows has a determinant of zero. A matrix with a zero row has a determinant of zero. A matrix is nonsingular if and only if its determinant is nonzero. The determinant of an echelon form matrix is the product down its diagonal.

Proof To verify the first sentence swap the two equal rows. The sign of the determinant changes but the matrix is the same and so its determinant is the same. Thus the determinant is zero.

For the second sentence multiply the zero row by two. That doubles the determinant but it also leaves the row unchanged, and hence leaves the determinant unchanged. Thus the determinant must be zero.

Do Gauss-Jordan reduction for the third sentence, $T \to \cdots \to \hat{T}$. By the first three properties the determinant of T is zero if and only if the determinant of $\hat{T}$ is zero (although the two could differ in sign or magnitude). A nonsingular matrix T Gauss-Jordan reduces to an identity matrix and so has a nonzero determinant. A singular T reduces to a $\hat{T}$ with a zero row; by the second sentence of this lemma its determinant is zero.

The fourth sentence has two cases. If the echelon form matrix is singular then it has a zero row. Thus it has a zero on its diagonal and the product down its diagonal is zero. By the third sentence of this result the determinant is zero and therefore this matrix's determinant equals the product down its diagonal.

If the echelon form matrix is nonsingular then none of its diagonal entries is zero so we can use condition (3) to get 1's on the diagonal.

$$\begin{vmatrix} t_{1,1} & t_{1,2} & & t_{1,n} \\ 0 & t_{2,2} & & t_{2,n} \\ & & \ddots & \\ 0 & & & t_{n,n} \end{vmatrix} = t_{1,1} \cdot t_{2,2} \cdots t_{n,n} \cdot \begin{vmatrix} 1 & t_{1,2}/t_{1,1} & & t_{1,n}/t_{1,1} \\ 0 & 1 & & t_{2,n}/t_{2,2} \\ & & \ddots & \\ 0 & & & 1 \end{vmatrix}$$

(We need that diagonal entries are nonzero to write, e.g., $t_{1,2}/t_{1,1}$.) Then the Jordan half of Gauss-Jordan elimination leaves the identity matrix.

$$= t_{1,1} \cdot t_{2,2} \cdots t_{n,n} \cdot \begin{vmatrix} 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ & & \ddots & \\ 0 & & & 1 \end{vmatrix} = t_{1,1} \cdot t_{2,2} \cdots t_{n,n} \cdot 1$$

So in this case also, the determinant is the product down the diagonal. QED

That gives us a way to compute the value of a determinant function on a matrix: do Gaussian reduction, keeping track of any changes of sign caused by row swaps and any scalars that we factor out, and finish by multiplying down the diagonal of the echelon form result. This algorithm is as fast as Gauss's Method and so is practical on all of the matrices that we will see.

2.5 Example Doing 2×2 determinants with Gauss's Method

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

doesn't give a big time savings because the 2×2 determinant formula is easy. However, a 3×3 determinant is often easier to calculate with Gauss's Method than with its formula.

$$\begin{vmatrix} 2 & 2 & 6 \\ 4 & 4 & 3 \\ 0 & -3 & 5 \end{vmatrix} = \begin{vmatrix} 2 & 2 & 6 \\ 0 & 0 & -9 \\ 0 & -3 & 5 \end{vmatrix} = -\begin{vmatrix} 2 & 2 & 6 \\ 0 & -3 & 5 \\ 0 & 0 & -9 \end{vmatrix} = -54$$

2.6 Example Determinants bigger than 3×3 go quickly with the Gauss's Method procedure.

$$\begin{vmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 5 \\ 0 & 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & -1 & -3 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & -1 & -3 \\ 0 & 0 & 0 & 5 \end{vmatrix} = -(-5) = 5$$

The prior example illustrates an important point. Although we have not yet found a 4×4 determinant formula, if one exists then we know what value it gives to the matrix — if there is a function with properties (1)-(4) then on the above matrix the function must return 5.

2.7 Lemma For each n, if there is an $n\times n$ determinant function then it is unique.

Proof. Perform Gauss's Method on the matrix, keeping track of how the sign alternates on row swaps and any row-scaling factors, and then multiply down the diagonal of the echelon form result. By the definition and the lemma, all $n\times n$ determinant functions must return this value on the matrix. QED

The 'if there is an $n\times n$ determinant function' emphasizes that, although we can use Gauss's Method to compute the only value that a determinant function could possibly return, we haven't yet shown that such a function exists for all n. The rest of this section does that.

Exercises

For these, assume that an $n\times n$ determinant function exists for all n.

✓ 2.8 Use Gauss's Method to find each determinant.

(a) $\begin{vmatrix} 3 & 1 & 2 \\ 3 & 1 & 0 \\ 0 & 1 & 4 \end{vmatrix}$ (b) $\begin{vmatrix} 1 & 0 & 0 & 1 \\ 2 & 1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{vmatrix}$

2.9 Use Gauss's Method to find each.

(a) $\begin{vmatrix} 2 & -1 \\ -1 & -1 \end{vmatrix}$ (b) $\begin{vmatrix} 1 & 1 & 0 \\ 3 & 0 & 2 \\ 5 & 2 & 2 \end{vmatrix}$

2.10 For which values of k does this system have a unique solution?

$$\begin{aligned} x \quad + z - w &= 2 \\ y - 2z \quad &= 3 \\ x \quad + kz \quad &= 4 \\ z - w &= 2 \end{aligned}$$

✓ 2.11 Express each of these in terms of $|H|$.

(a) $\begin{vmatrix} h_{3,1} & h_{3,2} & h_{3,3} \\ h_{2,1} & h_{2,2} & h_{2,3} \\ h_{1,1} & h_{1,2} & h_{1,3} \end{vmatrix}$

(b) $\begin{vmatrix} -h_{1,1} & -h_{1,2} & -h_{1,3} \\ -2h_{2,1} & -2h_{2,2} & -2h_{2,3} \\ -3h_{3,1} & -3h_{3,2} & -3h_{3,3} \end{vmatrix}$

(c) $\begin{vmatrix} h_{1,1}+h_{3,1} & h_{1,2}+h_{3,2} & h_{1,3}+h_{3,3} \\ h_{2,1} & h_{2,2} & h_{2,3} \\ 5h_{3,1} & 5h_{3,2} & 5h_{3,3} \end{vmatrix}$

✓ 2.12 Find the determinant of a diagonal matrix.

2.13 Describe the solution set of a homogeneous linear system if the determinant of the matrix of coefficients is nonzero.

✓ 2.14 Show that this determinant is zero.

$$\begin{vmatrix} y+z & x+z & x+y \\ x & y & z \\ 1 & 1 & 1 \end{vmatrix}$$

2.15 (a) Find the $1{\times}1$, $2{\times}2$, and $3{\times}3$ matrices with i,j entry given by $(-1)^{i+j}$.
(b) Find the determinant of the square matrix with i,j entry $(-1)^{i+j}$.

2.16 (a) Find the $1{\times}1$, $2{\times}2$, and $3{\times}3$ matrices with i,j entry given by $i+j$.
(b) Find the determinant of the square matrix with i,j entry $i+j$.

✓ 2.17 Show that determinant functions are not linear by giving a case where $|A+B| \neq |A|+|B|$.

2.18 The second condition in the definition, that row swaps change the sign of a determinant, is somewhat annoying. It means we have to keep track of the number of swaps, to compute how the sign alternates. Can we get rid of it? Can we replace it with the condition that row swaps leave the determinant unchanged? (If so then we would need new $1{\times}1$, $2{\times}2$, and $3{\times}3$ formulas, but that would be a minor matter.)

2.19 Prove that the determinant of any triangular matrix, upper or lower, is the product down its diagonal.

2.20 Refer to the definition of elementary matrices in the Mechanics of Matrix Multiplication subsection.
(a) What is the determinant of each kind of elementary matrix?
(b) Prove that if E is any elementary matrix then $|ES| = |E||S|$ for any appropriately sized S.
(c) *(This question doesn't involve determinants.)* Prove that if T is singular then a product TS is also singular.
(d) Show that $|TS| = |T||S|$.
(e) Show that if T is nonsingular then $|T^{-1}| = |T|^{-1}$.

2.21 Prove that the determinant of a product is the product of the determinants $|TS| = |T|\,|S|$ in this way. Fix the $n \times n$ matrix S and consider the function $d\colon \mathcal{M}_{n\times n} \to \mathbb{R}$ given by $T \mapsto |TS|/|S|$.
(a) Check that d satisfies condition (1) in the definition of a determinant function.
(b) Check condition (2).
(c) Check condition (3).
(d) Check condition (4).
(e) Conclude the determinant of a product is the product of the determinants.

2.22 A *submatrix* of a given matrix A is one that we get by deleting some of the rows and columns of A. Thus, the first matrix here is a submatrix of the second.

$$\begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} \qquad \begin{pmatrix} 3 & 4 & 1 \\ 0 & 9 & -2 \\ 2 & -1 & 5 \end{pmatrix}$$

Prove that for any square matrix, the rank of the matrix is r if and only if r is the largest integer such that there is an $r{\times}r$ submatrix with a nonzero determinant.

✓ 2.23 Prove that a matrix with rational entries has a rational determinant.

? 2.24 [Am. Math. Mon., Feb. 1953] Find the element of likeness in (a) simplifying a fraction, (b) powdering the nose, (c) building new steps on the church, (d) keeping emeritus professors on campus, (e) putting B, C, D in the determinant

$$\begin{vmatrix} 1 & a & a^2 & a^3 \\ a^3 & 1 & a & a^2 \\ B & a^3 & 1 & a \\ C & D & a^3 & 1 \end{vmatrix}.$$

I.3 The Permutation Expansion

The prior subsection defines a function to be a determinant if it satisfies four conditions and shows that there is at most one $n \times n$ determinant function for each n. What is left is to show that for each n such a function exists.

But, we easily compute determinants: we use Gauss's Method, keeping track of the sign changes from row swaps, and end by multiplying down the diagonal. How could they not exist?

The difficulty is to show that the computation gives a well-defined — that is, unique — result. Consider these two Gauss's Method reductions of the same matrix, the first without any row swap

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \xrightarrow{-3\rho_1+\rho_2} \begin{pmatrix} 1 & 2 \\ 0 & -2 \end{pmatrix}$$

and the second with one.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \xrightarrow{\rho_1\leftrightarrow\rho_2} \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} \xrightarrow{-(1/3)\rho_1+\rho_2} \begin{pmatrix} 3 & 4 \\ 0 & 2/3 \end{pmatrix}$$

Both yield the determinant -2 since in the second one we note that the row swap changes the sign of the result we get by multiplying down the diagonal. The fact that we are able to proceed in two ways opens the possibility that the two give different answers. That is, the way that we have given to compute determinant values does not plainly eliminate the possibility that there might be, say, two reductions of some 7×7 matrix that lead to different determinant values. In that case we would not have a function, since the definition of a function is that for each input there must be exactly associated one output. The rest of this section shows that the definition Definition 2.1 never leads to a conflict.

To do this we will define an alternative way to find the value of a determinant. (This alternative is less useful in practice because it is slow. But it is very useful for theory.) The key idea is that condition (3) of Definition 2.1 shows that the determinant function is not linear.

3.1 Example With condition (3) scalars come out of each row separately,

$$\begin{vmatrix} 4 & 2 \\ -2 & 6 \end{vmatrix} = 2 \cdot \begin{vmatrix} 2 & 1 \\ -2 & 6 \end{vmatrix} = 4 \cdot \begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix}$$

not from the entire matrix at once. So, where

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}$$

then $\det(2A) \neq 2 \cdot \det(A)$ (instead, $\det(2A) = 4 \cdot \det(A)$).

Since scalars come out a row at a time we might guess that determinants are linear a row at a time.

3.2 Definition Let V be a vector space. A map $f\colon V^n \to \mathbb{R}$ is *multilinear* if

(1) $f(\vec{\rho}_1, \ldots, \vec{v} + \vec{w}, \ldots, \vec{\rho}_n) = f(\vec{\rho}_1, \ldots, \vec{v}, \ldots, \vec{\rho}_n) + f(\vec{\rho}_1, \ldots, \vec{w}, \ldots, \vec{\rho}_n)$

(2) $f(\vec{\rho}_1, \ldots, k\vec{v}, \ldots, \vec{\rho}_n) = k \cdot f(\vec{\rho}_1, \ldots, \vec{v}, \ldots, \vec{\rho}_n)$

for $\vec{v}, \vec{w} \in V$ and $k \in \mathbb{R}$.

3.3 Lemma Determinants are multilinear.

Proof Property (2) here is just Definition 2.1's condition (3) so we need only verify property (1).

There are two cases. If the set of other rows $\{\vec{\rho}_1, \ldots, \vec{\rho}_{i-1}, \vec{\rho}_{i+1}, \ldots, \vec{\rho}_n\}$ is linearly dependent then all three matrices are singular and so all three determinants are zero and the equality is trivial.

Therefore assume that the set of other rows is linearly independent. We can make a basis by adding one more vector $\langle \vec{\rho}_1, \ldots, \vec{\rho}_{i-1}, \vec{\beta}, \vec{\rho}_{i+1}, \ldots, \vec{\rho}_n \rangle$. Express $\vec{v}$ and $\vec{w}$ with respect to this basis

$$\begin{aligned} \vec{v} &= v_1\vec{\rho}_1 + \cdots + v_{i-1}\vec{\rho}_{i-1} + v_i\vec{\beta} + v_{i+1}\vec{\rho}_{i+1} + \cdots + v_n\vec{\rho}_n \\ \vec{w} &= w_1\vec{\rho}_1 + \cdots + w_{i-1}\vec{\rho}_{i-1} + w_i\vec{\beta} + w_{i+1}\vec{\rho}_{i+1} + \cdots + w_n\vec{\rho}_n \end{aligned}$$

and add.

$$\vec{v} + \vec{w} = (v_1 + w_1)\vec{\rho}_1 + \cdots + (v_i + w_i)\vec{\beta} + \cdots + (v_n + w_n)\vec{\rho}_n$$

Consider the left side of (1) and expand $\vec{v} + \vec{w}$.

$$\det(\vec{\rho}_1, \ldots,\ (v_1 + w_1)\vec{\rho}_1 + \cdots + (v_i + w_i)\vec{\beta} + \cdots + (v_n + w_n)\vec{\rho}_n,\ \ldots, \vec{\rho}_n) \quad (*)$$

By the definition of determinant's condition (1), the value of $(*)$ is unchanged by the operation of adding $-(v_1+w_1)\vec{\rho}_1$ to the i-th row $\vec{v}+\vec{w}$. The i-th row becomes this.

$$\vec{v}+\vec{w}-(v_1+w_1)\vec{\rho}_1=(v_2+w_2)\vec{\rho}_2+\cdots+(v_i+w_i)\vec{\beta}+\cdots+(v_n+w_n)\vec{\rho}_n$$

Next add $-(v_2+w_2)\vec{\rho}_2$, etc., to eliminate all of the terms from the other rows. Apply condition (3) from the definition of determinant.

$$\begin{aligned}\det(\vec{\rho}_1,\ldots,\vec{v}+\vec{w},\ldots,\vec{\rho}_n)\\ &=\det(\vec{\rho}_1,\ldots,(v_i+w_i)\cdot\vec{\beta},\ldots,\vec{\rho}_n)\\ &=(v_i+w_i)\cdot\det(\vec{\rho}_1,\ldots,\vec{\beta},\ldots,\vec{\rho}_n)\\ &=v_i\cdot\det(\vec{\rho}_1,\ldots,\vec{\beta},\ldots,\vec{\rho}_n)+w_i\cdot\det(\vec{\rho}_1,\ldots,\vec{\beta},\ldots,\vec{\rho}_n)\end{aligned}$$

Now this is a sum of two determinants. To finish, bring v_i and w_i back inside in front of the $\vec{\beta}$'s and use row combinations again, this time to reconstruct the expressions of $\vec{v}$ and $\vec{w}$ in terms of the basis. That is, start with the operations of adding $v_1\vec{\rho}_1$ to $v_i\vec{\beta}$ and $w_1\vec{\rho}_1$ to $w_i\vec{\rho}_1$, etc., to get the expansions of $\vec{v}$ and $\vec{w}$.
QED

Multilinearity allows us to expand a determinant into a sum of determinants, each of which involves a simple matrix.

3.4 Example Use property (1) of multilinearity to break up the first row

$$\begin{vmatrix}2&1\\4&3\end{vmatrix}=\begin{vmatrix}2&0\\4&3\end{vmatrix}+\begin{vmatrix}0&1\\4&3\end{vmatrix}$$

and then use (1) again to break each along the second row.

$$=\begin{vmatrix}2&0\\4&0\end{vmatrix}+\begin{vmatrix}2&0\\0&3\end{vmatrix}+\begin{vmatrix}0&1\\4&0\end{vmatrix}+\begin{vmatrix}0&1\\0&3\end{vmatrix}$$

The result is four determinants. In each row of each of the four there is a single entry from the original matrix.

3.5 Example In the same way, a 3×3 determinant separates into a sum of many simpler determinants. Splitting along the first row produces three determinants (we have highlighted the zero in the $1,3$ position to set it off visually from the zeroes that appear as part of the splitting).

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

In turn, each of the above splits in three along the second row. Then each of the nine splits in three along the third row. The result is twenty seven determinants, such that each row contains a single entry from the starting matrix.

$$= \begin{vmatrix} 2 & 0 & 0 \\ 4 & 0 & 0 \\ 2 & 0 & 0 \end{vmatrix} + \begin{vmatrix} 2 & 0 & 0 \\ 4 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} + \begin{vmatrix} 2 & 0 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 5 \end{vmatrix} + \begin{vmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 2 & 0 & 0 \end{vmatrix} + \cdots + \begin{vmatrix} 0 & 0 & -1 \\ 0 & 0 & \boxed{0} \\ 0 & 0 & 5 \end{vmatrix}$$

So multilinearity will expand an $n \times n$ determinant into a sum of n^n-many determinants, where each row of each determinant contains a single entry from the starting matrix.

In this expansion, although there are lots of terms, most of them have a determinant of zero.

3.6 Example In each of these examples from the prior expansion, two of the entries from the original matrix are in the same column.

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

For instance, in the first matrix the 2 and the 4 both come from the first column of the original matrix. In the second matrix the -1 and 5 both come from the third column. And in the third matrix the 0 and 5 both come from the third column. Any such matrix is singular because one row is a multiple of the other. Thus any such determinant is zero, by Lemma 2.4.

With that observation the above expansion of the 3×3 determinant into the sum of the twenty seven determinants simplifies to the sum of these six where the entries from the original matrix come one per row, and also one per column.

$$\begin{aligned} \begin{vmatrix} 2 & 1 & -1 \\ 4 & 3 & \boxed{0} \\ 2 & 1 & 5 \end{vmatrix} &= \begin{vmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{vmatrix} + \begin{vmatrix} 2 & 0 & 0 \\ 0 & 0 & \boxed{0} \\ 0 & 1 & 0 \end{vmatrix} \\ &\quad + \begin{vmatrix} 0 & 1 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 5 \end{vmatrix} + \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & \boxed{0} \\ 2 & 0 & 0 \end{vmatrix} \\ &\quad + \begin{vmatrix} 0 & 0 & -1 \\ 4 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & -1 \\ 0 & 3 & 0 \\ 2 & 0 & 0 \end{vmatrix} \end{aligned}$$

In that expansion we can bring out the scalars.

$$
\begin{aligned}
&= (2)(3)(5)\begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix} + (2)(\boxed{0})(1)\begin{vmatrix}1&0&0\\0&0&1\\0&1&0\end{vmatrix}\\
&\quad+ (1)(4)(5)\begin{vmatrix}0&1&0\\1&0&0\\0&0&1\end{vmatrix} + (1)(\boxed{0})(2)\begin{vmatrix}0&1&0\\0&0&1\\1&0&0\end{vmatrix}\\
&\quad+ (-1)(4)(1)\begin{vmatrix}0&0&1\\1&0&0\\0&1&0\end{vmatrix} + (-1)(3)(2)\begin{vmatrix}0&0&1\\0&1&0\\1&0&0\end{vmatrix}
\end{aligned}
$$

To finish, evaluate those six determinants by row-swapping them to the identity matrix, keeping track of the sign changes.

$$
\begin{aligned}
&= 30\cdot(+1) + 0\cdot(-1)\\
&\quad + 20\cdot(-1) + 0\cdot(+1)\\
&\quad - 4\cdot(+1) - 6\cdot(-1) = 12
\end{aligned}
$$

That example captures this subsection's new calculation scheme. Multilinearity expands a determinant into many separate determinants, each with one entry from the original matrix per row. Most of these have one row that is a multiple of another so we omit them. We are left with the determinants that have one entry per row and column from the original matrix. Factoring out the scalars further reduces the determinants that we must compute to the one-entry-per-row-and-column matrices where all entries are 1's.

Recall Definition Three.IV.3.14, that a *permutation matrix* is square, with entries 0's except for a single 1 in each row and column. We now introduce a notation for permutation matrices.

3.7 Definition An n-*permutation* is a function on the first n positive integers $\phi\colon \{1,\ldots,n\}\to\{1,\ldots,n\}$ that is one-to-one and onto.

In a permutation each number $1,\ldots,n$ appears as output for one and only one input. We can denote a permutation as a sequence $\phi = \langle \phi(1), \phi(2), \ldots, \phi(n)\rangle$.

3.8 Example The 2-permutations are the functions $\phi_1\colon\{1,2\}\to\{1,2\}$ given by $\phi_1(1)=1$, $\phi_1(2)=2$, and $\phi_2\colon\{1,2\}\to\{1,2\}$ given by $\phi_2(1)=2$, $\phi_2(2)=1$. The sequence notation is shorter: $\phi_1=\langle 1,2\rangle$ and $\phi_2=\langle 2,1\rangle$.

3.9 Example In the sequence notation the 3-permutations are $\phi_1=\langle 1,2,3\rangle$, $\phi_2=\langle 1,3,2\rangle$, $\phi_3=\langle 2,1,3\rangle$, $\phi_4=\langle 2,3,1\rangle$, $\phi_5=\langle 3,1,2\rangle$, and $\phi_6=\langle 3,2,1\rangle$.

Let ι_j be the row vector that is all 0's except for a 1 in entry j, so that the four-wide ι_2 is $(0\ \ 1\ \ 0\ \ 0)$. Then our notation for permutation matrices is: with any $\phi = \langle \phi(1), \ldots, \phi(n) \rangle$ associate the matrix whose rows are $\iota_{\phi(1)}, \ldots, \iota_{\phi(n)}$. For instance, associated with the 4-permutation $\phi = \langle 3, 2, 1, 4 \rangle$ is the matrix whose rows are the corresponding ι's.

$$P_\phi = \begin{pmatrix} \iota_3 \\ \iota_2 \\ \iota_1 \\ \iota_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

3.10 Example These are the permutation matrices for the 2-permutations listed in Example 3.8.

$$P_{\phi_1} = \begin{pmatrix} \iota_1 \\ \iota_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad P_{\phi_2} = \begin{pmatrix} \iota_2 \\ \iota_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

For instance, P_{ϕ_2}'s first row is $\iota_{\phi_2(1)} = \iota_2$ and its second is $\iota_{\phi_2(2)} = \iota_1$.

3.11 Example Consider the 3-permutation $\phi_5 = \langle 3, 1, 2 \rangle$. The permutation matrix P_{ϕ_5} has rows $\iota_{\phi_5(1)} = \iota_3$, $\iota_{\phi_5(2)} = \iota_1$, and $\iota_{\phi_5(3)} = \iota_2$.

$$P_{\phi_5} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

3.12 Definition The *permutation expansion* for determinants is

$$\begin{vmatrix} t_{1,1} & t_{1,2} & \ldots & t_{1,n} \\ t_{2,1} & t_{2,2} & \ldots & t_{2,n} \\ & \vdots & & \\ t_{n,1} & t_{n,2} & \ldots & t_{n,n} \end{vmatrix} = \begin{aligned} & t_{1,\phi_1(1)} t_{2,\phi_1(2)} \cdots t_{n,\phi_1(n)} |P_{\phi_1}| \\ & + t_{1,\phi_2(1)} t_{2,\phi_2(2)} \cdots t_{n,\phi_2(n)} |P_{\phi_2}| \\ & \ \vdots \\ & + t_{1,\phi_k(1)} t_{2,\phi_k(2)} \cdots t_{n,\phi_k(n)} |P_{\phi_k}| \end{aligned}$$

where $\phi_1, \ldots, \phi_k$ are all of the n-permutations.

We can restate the formula in *summation notation*

$$|T| = \sum_{\text{permutations } \phi} t_{1,\phi(1)} t_{2,\phi(2)} \cdots t_{n,\phi(n)} \, |P_\phi|$$

read aloud as, "the sum, over all permutations ϕ, of terms having the form $t_{1,\phi(1)} t_{2,\phi(2)} \cdots t_{n,\phi(n)} |P_\phi|$."

3.13 Example The familiar 2×2 determinant formula follows from the above

$$\begin{vmatrix} t_{1,1} & t_{1,2} \\ t_{2,1} & t_{2,2} \end{vmatrix} = t_{1,1}t_{2,2}\cdot|P_{\phi_1}| + t_{1,2}t_{2,1}\cdot|P_{\phi_2}|$$
$$= t_{1,1}t_{2,2}\cdot\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + t_{1,2}t_{2,1}\cdot\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$$
$$= t_{1,1}t_{2,2} - t_{1,2}t_{2,1}$$

as does the 3×3 formula.

$$\begin{aligned}\begin{vmatrix} t_{1,1} & t_{1,2} & t_{1,3} \\ t_{2,1} & t_{2,2} & t_{2,3} \\ t_{3,1} & t_{3,2} & t_{3,3} \end{vmatrix} &= t_{1,1}t_{2,2}t_{3,3}\,|P_{\phi_1}| + t_{1,1}t_{2,3}t_{3,2}\,|P_{\phi_2}| + t_{1,2}t_{2,1}t_{3,3}\,|P_{\phi_3}| \\ &\quad + t_{1,2}t_{2,3}t_{3,1}\,|P_{\phi_4}| + t_{1,3}t_{2,1}t_{3,2}\,|P_{\phi_5}| + t_{1,3}t_{2,2}t_{3,1}\,|P_{\phi_6}| \\ &= t_{1,1}t_{2,2}t_{3,3} - t_{1,1}t_{2,3}t_{3,2} - t_{1,2}t_{2,1}t_{3,3} \\ &\quad + t_{1,2}t_{2,3}t_{3,1} + t_{1,3}t_{2,1}t_{3,2} - t_{1,3}t_{2,2}t_{3,1}\end{aligned}$$

Computing a determinant with the permutation expansion typically takes longer than with Gauss's Method. However, we will use it to prove that the determinant function is well-defined. We will just state the result here and defer its proof to the following subsection.

3.14 Theorem For each n there is an $n\times n$ determinant function.

Also in the next subsection is the proof of the result below (they are together because the two proofs overlap).

3.15 Theorem The determinant of a matrix equals the determinant of its transpose.

Because of this theorem, while we have so far stated determinant results in terms of rows, all of the results also hold in terms of columns.

3.16 Corollary A matrix with two equal columns is singular. Column swaps change the sign of a determinant. Determinants are multilinear in their columns.

Proof For the first statement, transposing the matrix results in a matrix with the same determinant, and with two equal rows, and hence a determinant of zero. Prove the other two in the same way. QED

We finish this subsection with a summary: determinant functions exist, are unique, and we know how to compute them. As for what determinants are about, perhaps these lines [Kemp] help make it memorable.

Determinant none,
 Solution: lots or none.
Determinant some,
 Solution: just one.

Exercises

This summarizes our notation for the 2- and 3- permutations.

i	1	2
$\phi_1(i)$	1	2
$\phi_2(i)$	2	1

i	1	2	3
$\phi_1(i)$	1	2	3
$\phi_2(i)$	1	3	2
$\phi_3(i)$	2	1	3
$\phi_4(i)$	2	3	1
$\phi_5(i)$	3	1	2
$\phi_6(i)$	3	2	1

✓ **3.17** Compute the determinant by using the permutation expansion.

(a) $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$ **(b)** $\begin{vmatrix} 2 & 2 & 1 \\ 3 & -1 & 0 \\ -2 & 0 & 5 \end{vmatrix}$

✓ **3.18** Compute these both with Gauss's Method and the permutation expansion formula.

(a) $\begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix}$ **(b)** $\begin{vmatrix} 0 & 1 & 4 \\ 0 & 2 & 3 \\ 1 & 5 & 1 \end{vmatrix}$

✓ **3.19** Use the permutation expansion formula to derive the formula for 3×3 determinants.

3.20 List all of the 4-permutations.

3.21 A permutation, regarded as a function from the set $\{1, .., n\}$ to itself, is one-to-one and onto. Therefore, each permutation has an inverse.
 (a) Find the inverse of each 2-permutation.
 (b) Find the inverse of each 3-permutation.

3.22 Prove that f is multilinear if and only if for all $\vec{v}, \vec{w} \in V$ and $k_1, k_2 \in \mathbb{R}$, this holds.

$$f(\vec{\rho}_1, \ldots, k_1\vec{v}_1 + k_2\vec{v}_2, \ldots, \vec{\rho}_n) = k_1 f(\vec{\rho}_1, \ldots, \vec{v}_1, \ldots, \vec{\rho}_n) + k_2 f(\vec{\rho}_1, \ldots, \vec{v}_2, \ldots, \vec{\rho}_n)$$

3.23 How would determinants change if we changed property (4) of the definition to read that $|I| = 2$?

3.24 Verify the second and third statements in Corollary 3.16.

✓ **3.25** Show that if an $n\times n$ matrix has a nonzero determinant then we can express any column vector $\vec{v} \in \mathbb{R}^n$ as a linear combination of the columns of the matrix.

3.26 [Strang 80] True or false: a matrix whose entries are only zeros or ones has a determinant equal to zero, one, or negative one.

3.27 **(a)** Show that there are 120 terms in the permutation expansion formula of a 5×5 matrix.
 (b) How many are sure to be zero if the $1, 2$ entry is zero?

3.28 How many n-permutations are there?

3.29 Show that the inverse of a permutation matrix is its transpose.

3.30 A matrix A is *skew-symmetric* if $A^{T} = -A$, as in this matrix.

$$A = \begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix}$$

Show that $n \times n$ skew-symmetric matrices with nonzero determinants exist only for even n.

✓ 3.31 What is the smallest number of zeros, and the placement of those zeros, needed to ensure that a 4×4 matrix has a determinant of zero?

✓ 3.32 If we have n data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ and want to find a polynomial $p(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1 x + a_0$ passing through those points then we can plug in the points to get an n equation/n unknown linear system. The matrix of coefficients for that system is the *Vandermonde matrix*. Prove that the determinant of the transpose of that matrix of coefficients

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ {x_1}^2 & {x_2}^2 & \ldots & {x_n}^2 \\ & \vdots & & \\ {x_1}^{n-1} & {x_2}^{n-1} & \ldots & {x_n}^{n-1} \end{vmatrix}$$

equals the product, over all indices $i, j \in \{1, \ldots, n\}$ with $i < j$, of terms of the form $x_j - x_i$. (This shows that the determinant is zero, and the linear system has no solution, if and only if the x_i's in the data are not distinct.)

3.33 We can divide a matrix into *blocks*, as here,

$$\left(\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 0 \\ \hline 0 & 0 & -2 \end{array}\right)$$

which shows four blocks, the square 2×2 and 1×1 ones in the upper left and lower right, and the zero blocks in the upper right and lower left. Show that if a matrix is such that we can partition it as

$$T = \left(\begin{array}{c|c} J & Z_2 \\ \hline Z_1 & K \end{array}\right)$$

where J and K are square, and Z_1 and Z_2 are all zeroes, then $|T| = |J| \cdot |K|$.

✓ 3.34 Prove that for any $n \times n$ matrix T there are at most n distinct reals r such that the matrix $T - rI$ has determinant zero (we shall use this result in Chapter Five).

? 3.35 [Math. Mag., Jan. 1963, Q307] The nine positive digits can be arranged into 3×3 arrays in $9!$ ways. Find the sum of the determinants of these arrays.

3.36 [Math. Mag., Jan. 1963, Q237] Show that

$$\begin{vmatrix} x-2 & x-3 & x-4 \\ x+1 & x-1 & x-3 \\ x-4 & x-7 & x-10 \end{vmatrix} = 0.$$

? 3.37 [Am. Math. Mon., Jan. 1949] Let S be the sum of the integer elements of a magic square of order three and let D be the value of the square considered as a determinant. Show that D/S is an integer.

? **3.38** [Am. Math. Mon., Jun. 1931] Show that the determinant of the n^2 elements in the upper left corner of the Pascal triangle

$$\begin{array}{cccccc} 1 & 1 & 1 & 1 & . & . \\ 1 & 2 & 3 & . & . & \\ 1 & 3 & . & . & & \\ 1 & . & . & & & \\ . & & & & & \\ . & & & & & \end{array}$$

has the value unity.

I.4 Determinants Exist

This subsection contains proofs of two results from the prior subsection. It is optional. We will use the material developed here only in the Jordan Canonical Form subsection, which is also optional.

We wish to show that for any size n, the determinant function on $n{\times}n$ matrices is well-defined. The prior subsection develops the permutation expansion formula.

$$\begin{aligned}\begin{vmatrix} t_{1,1} & t_{1,2} & \ldots & t_{1,n} \\ t_{2,1} & t_{2,2} & \ldots & t_{2,n} \\ & \vdots & & \\ t_{n,1} & t_{n,2} & \ldots & t_{n,n} \end{vmatrix} &= t_{1,\phi_1(1)}t_{2,\phi_1(2)}\cdots t_{n,\phi_1(n)}|P_{\phi_1}| \\ &\quad + t_{1,\phi_2(1)}t_{2,\phi_2(2)}\cdots t_{n,\phi_2(n)}|P_{\phi_2}| \\ &\quad \vdots \\ &\quad + t_{1,\phi_k(1)}t_{2,\phi_k(2)}\cdots t_{n,\phi_k(n)}|P_{\phi_k}| \\ &= \sum_{\text{permutations }\phi} t_{1,\phi(1)}t_{2,\phi(2)}\cdots t_{n,\phi(n)}\,|P_\phi|\end{aligned}$$

This reduces the problem of showing that the determinant is well-defined to only showing that the determinant is well-defined on the set of permutation matrices.

A permutation matrix can be row-swapped to the identity matrix. So one way that we can calculate its determinant is by keeping track of the number of swaps. However, we still must show that the result is well-defined. Recall what the difficulty is: the determinant of

$$P_\phi = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

could be computed with one swap

$$P_\phi \quad \xrightarrow{\rho_1\leftrightarrow\rho_2} \quad \begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}$$

or with three.

$$P_\phi \quad \xrightarrow{\rho_3\leftrightarrow\rho_1} \quad \xrightarrow{\rho_2\leftrightarrow\rho_3} \quad \xrightarrow{\rho_1\leftrightarrow\rho_3} \quad \begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}$$

Both reductions have an odd number of swaps so in this case we figure that $|P_\phi| = -1$ but if there were some way to do it with an even number of swaps then we would have the determinant giving two different outputs from a single input. Below, Corollary 4.5 proves that this cannot happen — there is no permutation matrix that can be row-swapped to an identity matrix in two ways, one with an even number of swaps and the other with an odd number of swaps.

4.1 Definition In a permutation $\phi = \langle \ldots, k, \ldots, j, \ldots \rangle$, elements such that $k > j$ are in an *inversion* of their natural order. Similarly, in a permutation matrix two rows

$$P_\phi = \begin{pmatrix} \vdots \\ \iota_k \\ \vdots \\ \iota_j \\ \vdots \end{pmatrix}$$

such that $k > j$ are in an *inversion*.

4.2 Example This permutation matrix

$$\begin{pmatrix} 1&0&0&0\\ 0&0&1&0\\ 0&1&0&0\\ 0&0&0&1 \end{pmatrix} = \begin{pmatrix} \iota_1 \\ \iota_3 \\ \iota_2 \\ \iota_4 \end{pmatrix}$$

has a single inversion, that ι_3 precedes ι_2.

4.3 Example There are three inversions here:

$$\begin{pmatrix} 0&0&1\\ 0&1&0\\ 1&0&0 \end{pmatrix} = \begin{pmatrix} \iota_3 \\ \iota_2 \\ \iota_1 \end{pmatrix}$$

ι_3 precedes ι_1, ι_3 precedes ι_2, and ι_2 precedes ι_1.

4.4 Lemma A row-swap in a permutation matrix changes the number of inversions from even to odd, or from odd to even.

Proof Consider a swap of rows j and k, where $k > j$.

If the two rows are adjacent

$$P_\phi = \begin{pmatrix} \vdots \\ \iota_{\phi(j)} \\ \iota_{\phi(k)} \\ \vdots \end{pmatrix} \xrightarrow{\rho_k \leftrightarrow \rho_j} \begin{pmatrix} \vdots \\ \iota_{\phi(k)} \\ \iota_{\phi(j)} \\ \vdots \end{pmatrix}$$

then since inversions involving rows not in this pair are not affected, the swap changes the total number of inversions by one, either removing or producing one inversion depending on whether $\phi(j) > \phi(k)$ or not. Consequently, the total number of inversions changes from odd to even or from even to odd.

If the rows are not adjacent then we can swap them via a sequence of adjacent swaps, first bringing row k up

$$\begin{pmatrix} \vdots \\ \iota_{\phi(j)} \\ \iota_{\phi(j+1)} \\ \iota_{\phi(j+2)} \\ \vdots \\ \iota_{\phi(k)} \\ \vdots \end{pmatrix} \xrightarrow{\rho_k \leftrightarrow \rho_{k-1}} \xrightarrow{\rho_{k-1} \leftrightarrow \rho_{k-2}} \dots \xrightarrow{\rho_{j+1} \leftrightarrow \rho_j} \begin{pmatrix} \vdots \\ \iota_{\phi(k)} \\ \iota_{\phi(j)} \\ \iota_{\phi(j+1)} \\ \vdots \\ \iota_{\phi(k-1)} \\ \vdots \end{pmatrix}$$

and then bringing row j down.

$$\xrightarrow{\rho_{j+1} \leftrightarrow \rho_{j+2}} \xrightarrow{\rho_{j+2} \leftrightarrow \rho_{j+3}} \dots \xrightarrow{\rho_{k-1} \leftrightarrow \rho_k} \begin{pmatrix} \vdots \\ \iota_{\phi(k)} \\ \iota_{\phi(j+1)} \\ \iota_{\phi(j+2)} \\ \vdots \\ \iota_{\phi(j)} \\ \vdots \end{pmatrix}$$

Each of these adjacent swaps changes the number of inversions from odd to even or from even to odd. The total number of swaps $(k-j)+(k-j-1)$ is odd.

Thus, in aggregate, the number of inversions changes from even to odd, or from odd to even. QED

4.5 Corollary If a permutation matrix has an odd number of inversions then swapping it to the identity takes an odd number of swaps. If it has an even number of inversions then swapping to the identity takes an even number.

Proof The identity matrix has zero inversions. To change an odd number to zero requires an odd number of swaps, and to change an even number to zero requires an even number of swaps. QED

4.6 Example The matrix in Example 4.3 can be brought to the identity with one swap $\rho_1 \leftrightarrow \rho_3$. (So the number of swaps needn't be the same as the number of inversions, but the oddness or evenness of the two numbers is the same.)

4.7 Definition The *signum* of a permutation $\operatorname{sgn}(\phi)$ is -1 if the number of inversions in ϕ is odd and is $+1$ if the number of inversions is even.

4.8 Example Using the notation for the 3-permutations from Example 3.8 we have

$$P_{\phi_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad P_{\phi_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

so $\operatorname{sgn}(\phi_1) = 1$ because there are no inversions, while $\operatorname{sgn}(\phi_2) = -1$ because there is one.

We still have not shown that the determinant function is well-defined because we have not considered row operations on permutation matrices other than row swaps. We will finesse this issue. We will define a function $d\colon \mathcal{M}_{n\times n} \to \mathbb{R}$ by altering the permutation expansion formula, replacing $|P_\phi|$ with $\operatorname{sgn}(\phi)$.

$$d(T) = \sum_{\text{permutations } \phi} t_{1,\phi(1)} t_{2,\phi(2)} \cdots t_{n,\phi(n)} \operatorname{sgn}(\phi)$$

The advantage of this formula is that the number of inversions is clearly well-defined — just count them. Therefore, we will be finished showing that an $n\times n$ determinant function exists when we show that this d satisfies the conditions required of a determinant.

4.9 Lemma The function d above is a determinant. Hence determinants exist for every n.

Proof We must check that it has the four conditions from the definition of determinant, Definition 2.1.

Condition (4) is easy: where I is the $n \times n$ identity, in

$$d(I) = \sum_{\text{perm } \phi} \iota_{1,\phi(1)}\iota_{2,\phi(2)} \cdots \iota_{n,\phi(n)} \operatorname{sgn}(\phi)$$

all of the terms in the summation are zero except for the one where the permutation ϕ is the identity, which gives the product down the diagonal, which is one.

For condition (3) suppose that $T \xrightarrow{k\rho_i} \hat{T}$ and consider $d(\hat{T})$.

$$\begin{aligned}\sum_{\text{perm } \phi} \hat{t}_{1,\phi(1)} \cdots \hat{t}_{i,\phi(i)} \cdots \hat{t}_{n,\phi(n)} \operatorname{sgn}(\phi)& \\ &= \sum_{\phi} t_{1,\phi(1)} \cdots k t_{i,\phi(i)} \cdots t_{n,\phi(n)} \operatorname{sgn}(\phi)\end{aligned}$$

Factor out k to get the desired equality.

$$= k \cdot \sum_{\phi} t_{1,\phi(1)} \cdots t_{i,\phi(i)} \cdots t_{n,\phi(n)} \operatorname{sgn}(\phi) = k \cdot d(T)$$

For (2) suppose that $T \xrightarrow{\rho_i \leftrightarrow \rho_j} \hat{T}$. We must show that $d(\hat{T})$ is the negative of $d(T)$.

$$d(\hat{T}) = \sum_{\text{perm } \phi} \hat{t}_{1,\phi(1)} \cdots \hat{t}_{i,\phi(i)} \cdots \hat{t}_{j,\phi(j)} \cdots \hat{t}_{n,\phi(n)} \operatorname{sgn}(\phi) \qquad (*)$$

We will show that each term in $(*)$ is associated with a term in $d(T)$, and that the two terms are negatives of each other. Consider the matrix from the multilinear expansion of $d(\hat{T})$ giving the term $\hat{t}_{1,\phi(1)} \cdots \hat{t}_{i,\phi(i)} \cdots \hat{t}_{j,\phi(j)} \cdots \hat{t}_{n,\phi(n)} \operatorname{sgn}(\phi)$.

$$\begin{pmatrix} & \vdots & \\ \hat{t}_{i,\phi(i)} & & \\ & \vdots & \\ & & \hat{t}_{j,\phi(j)} \\ & \vdots & \end{pmatrix}$$

It is the result of the $\rho_i \leftrightarrow \rho_j$ operation performed on this matrix.

$$\begin{pmatrix} & \vdots & \\ & & t_{i,\phi(j)} \\ & \vdots & \\ t_{j,\phi(i)} & & \\ & \vdots & \end{pmatrix}$$

That is, the term with hatted t's is associated with this term from the $d(T)$ expansion: $t_{1,\sigma(1)} \cdots t_{j,\sigma(j)} \cdots t_{i,\sigma(i)} \cdots t_{n,\sigma(n)} \operatorname{sgn}(\sigma)$, where the permutation σ equals ϕ but with the i-th and j-th numbers interchanged, $\sigma(i) = \phi(j)$ and $\sigma(j) = \phi(i)$. The two terms have the same multiplicands $\hat{t}_{1,\phi(1)} = t_{1,\sigma(1)}$, ..., including the entries from the swapped rows $\hat{t}_{i,\phi(i)} = t_{j,\phi(i)} = t_{j,\sigma(j)}$ and $\hat{t}_{j,\phi(j)} = t_{i,\phi(j)} = t_{i,\sigma(i)}$. But the two terms are negatives of each other since $\operatorname{sgn}(\phi) = -\operatorname{sgn}(\sigma)$ by Lemma 4.4.

Now, any permutation ϕ can be derived from some other permutation σ by such a swap, in one and only one way. Therefore the summation in $(*)$ is in fact a sum over all permutations, taken once and only once.

$$\begin{aligned}
d(\hat{T}) &= \sum_{\text{perm } \phi} \hat{t}_{1,\phi(1)} \cdots \hat{t}_{i,\phi(i)} \cdots \hat{t}_{j,\phi(j)} \cdots \hat{t}_{n,\phi(n)} \operatorname{sgn}(\phi) \\
&= \sum_{\text{perm } \sigma} t_{1,\sigma(1)} \cdots t_{j,\sigma(j)} \cdots t_{i,\sigma(i)} \cdots t_{n,\sigma(n)} \cdot \big(-\operatorname{sgn}(\sigma)\big)
\end{aligned}$$

Thus $d(\hat{T}) = -d(T)$.

Finally, for condition (1) suppose that $T \xrightarrow{k\rho_i+\rho_j} \hat{T}$.

$$\begin{aligned}
d(\hat{T}) &= \sum_{\text{perm } \phi} \hat{t}_{1,\phi(1)} \cdots \hat{t}_{i,\phi(i)} \cdots \hat{t}_{j,\phi(j)} \cdots \hat{t}_{n,\phi(n)} \operatorname{sgn}(\phi) \\
&= \sum_{\phi} t_{1,\phi(1)} \cdots t_{i,\phi(i)} \cdots (kt_{i,\phi(j)} + t_{j,\phi(j)}) \cdots t_{n,\phi(n)} \operatorname{sgn}(\phi)
\end{aligned}$$

Distribute over the addition in $kt_{i,\phi(j)} + t_{j,\phi(j)}$.

$$\begin{aligned}
= \sum_{\phi} \big[& t_{1,\phi(1)} \cdots t_{i,\phi(i)} \cdots kt_{i,\phi(j)} \cdots t_{n,\phi(n)} \operatorname{sgn}(\phi) \\
& + t_{1,\phi(1)} \cdots t_{i,\phi(i)} \cdots t_{j,\phi(j)} \cdots t_{n,\phi(n)} \operatorname{sgn}(\phi) \big]
\end{aligned}$$

Break it into two summations.

$$\begin{aligned}
= \sum_{\phi} & t_{1,\phi(1)} \cdots t_{i,\phi(i)} \cdots kt_{i,\phi(j)} \cdots t_{n,\phi(n)} \operatorname{sgn}(\phi) \\
& + \sum_{\phi} t_{1,\phi(1)} \cdots t_{i,\phi(i)} \cdots t_{j,\phi(j)} \cdots t_{n,\phi(n)} \operatorname{sgn}(\phi)
\end{aligned}$$

Recognize the second one.

$$\begin{aligned}
= k \cdot \sum_{\phi} & t_{1,\phi(1)} \cdots t_{i,\phi(i)} \cdots t_{i,\phi(j)} \cdots t_{n,\phi(n)} \operatorname{sgn}(\phi) \\
& + d(T)
\end{aligned}$$

Consider the terms $t_{1,\phi(1)} \cdots t_{i,\phi(i)} \cdots t_{i,\phi(j)} \cdots t_{n,\phi(n)} \operatorname{sgn}(\phi)$. Notice the subscripts; the entry is $t_{i,\phi(j)}$, not $t_{j,\phi(j)}$. The sum of these terms is the determinant of a matrix S that is equal to T except that row j of S is a copy of row i of T, that is, S has two equal rows. In the same way that we proved Lemma 2.4 we can see that $d(S) = 0$: a swap of S's equal rows will change the sign of $d(S)$ but since the matrix is unchanged by that swap the value of $d(S)$ must also be unchanged, and so that value must be zero. QED

We have now proved that determinant functions exist for each size $n \times n$. We already know that for each size there is at most one determinant. Therefore, for each size there is one and only one determinant function.

We end this subsection by proving the other result remaining from the prior subsection.

4.10 Theorem The determinant of a matrix equals the determinant of its transpose.

PROOF The proof is best understood by doing the general 3×3 case. That the argument applies to the $n\times n$ case will be clear.

Compare the permutation expansion of the matrix T

$$\begin{aligned}\begin{vmatrix} t_{1,1} & t_{1,2} & t_{1,3} \\ t_{2,1} & t_{2,2} & t_{2,3} \\ t_{3,1} & t_{3,2} & t_{3,3} \end{vmatrix} &= t_{1,1}t_{2,2}t_{3,3}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} + t_{1,1}t_{2,3}t_{3,2}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} \\ &\quad + t_{1,2}t_{2,1}t_{3,3}\begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} + t_{1,2}t_{2,3}t_{3,1}\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} \\ &\quad + t_{1,3}t_{2,1}t_{3,2}\begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} + t_{1,3}t_{2,2}t_{3,1}\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}\end{aligned}$$

with the permutation expansion of its transpose.

$$\begin{aligned}\begin{vmatrix} t_{1,1} & t_{2,1} & t_{3,1} \\ t_{1,2} & t_{2,2} & t_{3,2} \\ t_{1,3} & t_{2,3} & t_{3,3} \end{vmatrix} &= t_{1,1}t_{2,2}t_{3,3}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} + t_{1,1}t_{3,2}t_{2,3}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} \\ &\quad + t_{2,1}t_{1,2}t_{3,3}\begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} + t_{2,1}t_{3,2}t_{1,3}\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} \\ &\quad + t_{3,1}t_{1,2}t_{2,3}\begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} + t_{3,1}t_{2,2}t_{1,3}\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}\end{aligned}$$

Compare first the six products of t's. The ones in the expansion of T are the same as the ones in the expansion of the transpose; for instance, $t_{1,2}t_{2,3}t_{3,1}$ is in the top and $t_{3,1}t_{1,2}t_{2,3}$ is in the bottom. That's perfectly sensible — the six in the top arise from all of the ways of picking one entry of T from each row and column while the six in the bottom are all of the ways of picking one entry of T from each column and row, so of course they are the same set.

Next observe that in the two expansions, each t-product expression is not necessarily associated with the same permutation matrix. For instance, on the top $t_{1,2}t_{2,3}t_{3,1}$ is associated with the matrix for the map $1 \mapsto 2$, $2 \mapsto 3$, $3 \mapsto 1$. On the bottom $t_{3,1}t_{1,2}t_{2,3}$ is associated with the matrix for the map $1 \mapsto 3$, $2 \mapsto 1$, $3 \mapsto 2$. The second map is inverse to the first. This is also perfectly sensible — both the matrix transpose and the map inverse flip the $1, 2$ to $2, 1$, flip the $2, 3$ to $3, 2$, and flip $3, 1$ to $1, 3$.

We finish by noting that the determinant of P_ϕ equals the determinant of $P_{\phi^{-1}}$, as shown in Exercise 16. QED

Exercises

These summarize the notation used in this book for the 2- and 3- permutations.

i	1	2
$\phi_1(i)$	1	2
$\phi_2(i)$	2	1

i	1	2	3
$\phi_1(i)$	1	2	3
$\phi_2(i)$	1	3	2
$\phi_3(i)$	2	1	3
$\phi_4(i)$	2	3	1
$\phi_5(i)$	3	1	2
$\phi_6(i)$	3	2	1

4.11 Give the permutation expansion of a general 2×2 matrix and its transpose.

✓ 4.12 *This problem appears also in the prior subsection.*

(a) Find the inverse of each 2-permutation.

(b) Find the inverse of each 3-permutation.

✓ 4.13 (a) Find the signum of each 2-permutation.

(b) Find the signum of each 3-permutation.

4.14 Find the only nonzero term in the permutation expansion of this matrix.

$$\begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

Compute that determinant by finding the signum of the associated permutation.

4.15 [Strang 80] What is the signum of the n-permutation $\phi = \langle n, n-1, \ldots, 2, 1\rangle$?

4.16 Prove these.

(a) Every permutation has an inverse.

(b) $\operatorname{sgn}(\phi^{-1}) = \operatorname{sgn}(\phi)$

(c) Every permutation is the inverse of another.

4.17 Prove that the matrix of the permutation inverse is the transpose of the matrix of the permutation $P_{\phi^{-1}} = {P_\phi}^{\mathsf{T}}$, for any permutation ϕ.

✓ **4.18** Show that a permutation matrix with m inversions can be row swapped to the identity in m steps. Contrast this with Corollary 4.5.

✓ **4.19** For any permutation ϕ let $g(\phi)$ be the integer defined in this way.

$$g(\phi) = \prod_{i<j} [\phi(j) - \phi(i)]$$

(This is the product, over all indices i and j with $i < j$, of terms of the given form.)

(a) Compute the value of g on all 2-permutations.

(b) Compute the value of g on all 3-permutations.

(c) Prove that $g(\phi)$ is not 0.

(d) Prove this.

$$\operatorname{sgn}(\phi) = \frac{g(\phi)}{|g(\phi)|}$$

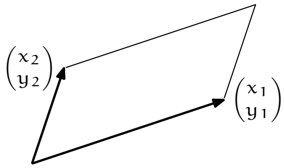
Many authors give this formula as the definition of the signum function.

II Geometry of Determinants

The prior section develops the determinant algebraically, by considering formulas satisfying certain conditions. This section complements that with a geometric approach. Beyond its intuitive appeal, an advantage of this approach is that while we have so far only considered whether or not a determinant is zero, here we shall give a meaning to the value of the determinant. (The prior section treats the determinant as a function of the rows but this section focuses on columns.)

II.1 Determinants as Size Functions

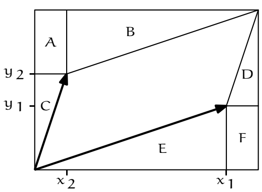
This parallelogram picture is familiar from the construction of the sum of the two vectors.



1.1 Definition In $\mathbb{R}^n$ the *box* (or *parallelepiped*) formed by $\langle\vec{v}_1,\ldots,\vec{v}_n\rangle$ is the set $\{t_1\vec{v}_1+\cdots+t_n\vec{v}_n \mid t_1,\ldots,t_n\in[0\,\ldots\,1]\}$.

Thus the parallelogram above is the box formed by $\langle\binom{x_1}{y_1},\binom{x_2}{y_2}\rangle$. A three-space box is shown in Example 1.4.

We can find the area of the above box by drawing an enclosing rectangle and subtracting away areas not in the box.



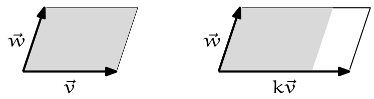
$$\begin{aligned}
&\text{area of parallelogram}\\
&\quad=\text{area of rectangle}-\text{area of }A-\text{area of }B\\
&\qquad-\cdots-\text{area of }F\\
&\quad=(x_1+x_2)(y_1+y_2)-x_2y_1-x_1y_1/2\\
&\qquad-x_2y_2/2-x_2y_2/2-x_1y_1/2-x_2y_1\\
&\quad=x_1y_2-x_2y_1
\end{aligned}$$

That the area equals the value of the determinant

$$\begin{vmatrix}x_1 & x_2\\ y_1 & y_2\end{vmatrix}=x_1y_2-x_2y_1$$

is no coincidence. The definition of determinants contains four properties that we know lead to a unique function for each dimension n. We shall argue that these properties make good postulates for a function that measure the size of boxes in n-space.

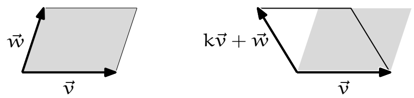
For instance, a function that measures the size of the box should have the property that multiplying one of the box-defining vectors by a scalar will multiply the size by that scalar.



Shown here is $k = 1.4$. On the right the rescaled region is in solid lines with the original region shaded for comparison.

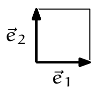
That is, we can reasonably expect that $\text{size}(\ldots, k\vec{v}, \ldots) = k \cdot \text{size}(\ldots, \vec{v}, \ldots)$. Of course, this condition is one of those in the definition of determinants.

Another property of determinants that should apply to any function measuring the size of a box is that it is unaffected by row combinations. Here are before-combining and after-combining boxes (the scalar shown is $k = -0.35$).



The box formed by v and $k\vec{v} + \vec{w}$ slants differently than the original one but the two have the same base and the same height, and hence the same area. So we expect that size is not affected by a shear operation $\text{size}(\ldots, \vec{v}, \ldots, \vec{w}, \ldots) = \text{size}(\ldots, \vec{v}, \ldots, k\vec{v} + \vec{w}, \ldots)$. Again, this is a determinant condition.

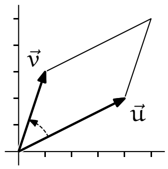
We expect that the box formed by unit vectors has unit size



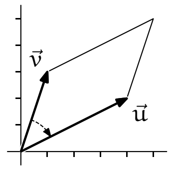
and we naturally extend that to any n-space $\text{size}(\vec{e}_1, \ldots, \vec{e}_n) = 1$.

Condition (2) of the definition of determinant is redundant, as remarked following the definition. We know from the prior section that for each n the determinant exists and is unique so we know that these postulates for size functions are consistent and that we do not need any more postulates. Therefore, we are justified in interpreting $\det(\vec{v}_1, \ldots, \vec{v}_n)$ as giving the size of the box formed by the vectors.

1.2 Remark Although condition (2) is redundant it raises an important point. Consider these two.



$$\begin{vmatrix} 4 & 1 \\ 2 & 3 \end{vmatrix} = 10$$

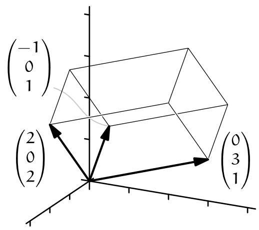


$$\begin{vmatrix} 1 & 4 \\ 3 & 2 \end{vmatrix} = -10$$

Swapping the columns changes the sign. On the left, starting with $\vec{u}$ and following the arc inside the angle to $\vec{v}$ (that is, going counterclockwise), we get a positive size. On the right, starting at $\vec{v}$ and going to $\vec{u}$, and so following the clockwise arc, gives a negative size. The sign returned by the size function reflects the *orientation* or *sense* of the box. (We see the same thing if we picture the effect of scalar multiplication by a negative scalar.)

1.3 Definition The *volume* of a box is the absolute value of the determinant of a matrix with those vectors as columns.

1.4 Example By the formula that takes the area of the base times the height, the volume of this parallelepiped is 12. That agrees with the determinant.



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

We can also compute the volume as the absolute value of this determinant.

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

1.5 Theorem A transformation $t\colon \mathbb{R}^n \to \mathbb{R}^n$ changes the size of all boxes by the same factor, namely, the size of the image of a box $|t(S)|$ is $|T|$ times the size of the box $|S|$, where T is the matrix representing t with respect to the standard basis.

That is, the determinant of a product is the product of the determinants $|TS| = |T| \cdot |S|$.

The two sentences say the same thing, first in map terms and then in matrix terms. This is because $|t(S)| = |TS|$, as both give the size of the box that is the image of the unit box $\mathcal{E}_n$ under the composition $t \circ s$, where the maps are represented with respect to the standard basis. We will prove the second sentence.

Proof First consider the $|T| = 0$ case, the case that T is singular and does not have an inverse. Observe that if TS is invertible then there is an M such that $(TS)M = I$, so $T(SM) = I$, and so T is invertible. The contrapositive of that observation is that if T is not invertible then neither is TS — if $|T| = 0$ then $|TS| = 0$.

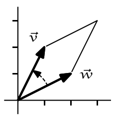
Now consider the case that T is nonsingular. Any nonsingular matrix factors into a product of elementary matrices $T = E_1 E_2 \cdots E_r$. To finish this argument we will verify that $|ES| = |E| \cdot |S|$ for all matrices S and elementary matrices E. The result will then follow because $|TS| = |E_1 \cdots E_r S| = |E_1| \cdots |E_r| \cdot |S| = |E_1 \cdots E_r| \cdot |S| = |T| \cdot |S|$.

There are three types of elementary matrix. We will cover the $M_i(k)$ case; the $P_{i,j}$ and $C_{i,j}(k)$ checks are similar. The matrix $M_i(k)S$ equals S except that row i is multiplied by k. The third condition of determinant functions then gives that $|M_i(k)S| = k \cdot |S|$. But $|M_i(k)| = k$, again by the third condition because $M_i(k)$ is derived from the identity by multiplication of row i by k. Thus $|ES| = |E| \cdot |S|$ holds for $E = M_i(k)$. QED

1.6 Example Application of the map t represented with respect to the standard bases by

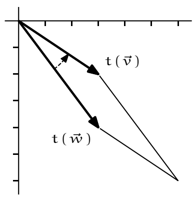
$$\begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix}$$

will double sizes of boxes, e.g., from this



$$\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3$$

to this



$$\begin{vmatrix} 3 & 3 \\ -4 & -2 \end{vmatrix} = 6$$

1.7 Corollary If a matrix is invertible then the determinant of its inverse is the inverse of its determinant $|T^{-1}| = 1/|T|$.

Proof $1 = |I| = |TT^{-1}| = |T| \cdot |T^{-1}|$ QED

Exercises

1.8 Find the volume of the region defined by the vectors.

(a) $\langle \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 4 \end{pmatrix} \rangle$

(b) $\langle \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}, \begin{pmatrix} 8 \\ -3 \\ 8 \end{pmatrix} \rangle$

(c) $\langle \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 0 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 7 \end{pmatrix} \rangle$

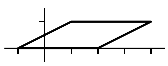
✓ 1.9 Is

$$\begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}$$

inside of the box formed by these three?

$$\begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 2 \\ 6 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix}$$

✓ 1.10 Find the volume of this region.



✓ 1.11 Suppose that $|A| = 3$. By what factor do these change volumes?
(a) A **(b)** A^2 **(c)** A^{-2}

✓ 1.12 By what factor does each transformation change the size of boxes?

(a) $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2x \\ 3y \end{pmatrix}$ **(b)** $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 3x - y \\ -2x + y \end{pmatrix}$ **(c)** $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x - y \\ x + y + z \\ y - 2z \end{pmatrix}$

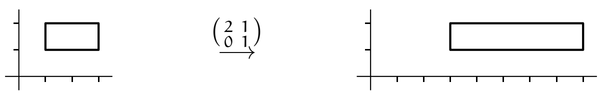
1.13 What is the area of the image of the rectangle $[2..4] \times [2..5]$ under the action of this matrix?

$$\begin{pmatrix} 2 & 3 \\ 4 & -1 \end{pmatrix}$$

1.14 If $t\colon \mathbb{R}^3 \to \mathbb{R}^3$ changes volumes by a factor of 7 and $s\colon \mathbb{R}^3 \to \mathbb{R}^3$ changes volumes by a factor of $3/2$ then by what factor will their composition changes volumes?

1.15 In what way does the definition of a box differ from the definition of a span?

✓ 1.16 Why doesn't this picture contradict Theorem 1.5?



area is 2 determinant is 2 area is 5

✓ 1.17 Does $|TS| = |ST|$? $|T(SP)| = |(TS)P|$?

1.18 **(a)** Suppose that $|A| = 3$ and that $|B| = 2$. Find $|A^2 \cdot B^{\mathsf{T}} \cdot B^{-2} \cdot A^{\mathsf{T}}|$.
(b) Assume that $|A| = 0$. Prove that $|6A^3 + 5A^2 + 2A| = 0$.

✓ **1.19** Let T be the matrix representing (with respect to the standard bases) the map that rotates plane vectors counterclockwise thru θ radians. By what factor does T change sizes?

✓ **1.20** Must a transformation $t\colon \mathbb{R}^2 \to \mathbb{R}^2$ that preserves areas also preserve lengths?

✓ **1.21** What is the volume of a parallelepiped in $\mathbb{R}^3$ bounded by a linearly dependent set?

✓ **1.22** Find the area of the triangle in $\mathbb{R}^3$ with endpoints $(1,2,1)$, $(3,-1,4)$, and $(2,2,2)$. (Area, not volume. The triangle defines a plane — what is the area of the triangle in that plane?)

✓ **1.23** An alternate proof of Theorem 1.5 uses the definition of determinant functions.

(a) Note that the vectors forming S make a linearly dependent set if and only if $|S| = 0$, and check that the result holds in this case.

(b) For the $|S| \neq 0$ case, to show that $|TS|/|S| = |T|$ for all transformations, consider the function $d\colon \mathcal{M}_{n\times n} \to \mathbb{R}$ given by $T \mapsto |TS|/|S|$. Show that d has the first property of a determinant.

(c) Show that d has the remaining three properties of a determinant function.

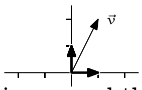
(d) Conclude that $|TS| = |T| \cdot |S|$.

1.24 Give a non-identity matrix with the property that $A^{\mathsf{T}} = A^{-1}$. Show that if $A^{\mathsf{T}} = A^{-1}$ then $|A| = \pm 1$. Does the converse hold?

1.25 The algebraic property of determinants that factoring a scalar out of a single row will multiply the determinant by that scalar shows that where H is 3×3, the determinant of cH is c^3 times the determinant of H. Explain this geometrically, that is, using Theorem 1.5. (The observation that increasing the linear size of a three-dimensional object by a factor of c will increase its volume by a factor of c^3 while only increasing its surface area by an amount proportional to a factor of c^2 is the *Square-cube law* [Wikipedia, Square-cube Law].)

✓ **1.26** We say that matrices H and G are *similar* if there is a nonsingular matrix P such that $H = P^{-1}GP$ (we will study this relation in Chapter Five). Show that similar matrices have the same determinant.

1.27 We usually represent vectors in $\mathbb{R}^2$ with respect to the standard basis so vectors in the first quadrant have both coordinates positive.



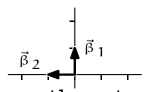
$$\mathrm{Rep}_{\mathcal{E}_2}(\vec{v}) = \begin{pmatrix} +3 \\ +2 \end{pmatrix}$$

Moving counterclockwise around the origin, we cycle thru four regions:

$$\cdots \longrightarrow \begin{pmatrix} + \\ + \end{pmatrix} \longrightarrow \begin{pmatrix} - \\ + \end{pmatrix} \longrightarrow \begin{pmatrix} - \\ - \end{pmatrix} \longrightarrow \begin{pmatrix} + \\ - \end{pmatrix} \longrightarrow \cdots.$$

Using this basis

$$B = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \rangle$$



gives the same counterclockwise cycle. We say these two bases have the same *orientation*.

(a) Why do they give the same cycle?

(b) What other configurations of unit vectors on the axes give the same cycle?

(c) Find the determinants of the matrices formed from those (ordered) bases.
(d) What other counterclockwise cycles are possible, and what are the associated determinants?
(e) What happens in $\mathbb{R}^1$?
(f) What happens in $\mathbb{R}^3$?

A fascinating general-audience discussion of orientations is in [Gardner].

1.28 *This question uses material from the optional Determinant Functions Exist subsection.* Prove Theorem 1.5 by using the permutation expansion formula for the determinant.

✓ 1.29 **(a)** Show that this gives the equation of a line in $\mathbb{R}^2$ thru (x_2, y_2) and (x_3, y_3).

$$\begin{vmatrix} x & x_2 & x_3 \\ y & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix} = 0$$

(b) [Petersen] Prove that the area of a triangle with vertices (x_1, y_1), (x_2, y_2), and (x_3, y_3) is

$$\frac{1}{2}\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix}.$$

(c) [Math. Mag., Jan. 1973] Prove that the area of a triangle with vertices at (x_1, y_1), (x_2, y_2), and (x_3, y_3) whose coordinates are integers has an area of N or $N/2$ for some positive integer N.

III Laplace's Formula

Determinants are a font of interesting and amusing formulas. Here is one that is often used to compute determinants by hand.

III.1 Laplace's Expansion

The example shows a 3×3 case but the approach works for any size $n > 1$.

1.1 Example Consider the permutation expansion.

$$\begin{aligned}
\begin{vmatrix} t_{1,1} & t_{1,2} & t_{1,3} \\ t_{2,1} & t_{2,2} & t_{2,3} \\ t_{3,1} & t_{3,2} & t_{3,3} \end{vmatrix} &= t_{1,1}t_{2,2}t_{3,3}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} + t_{1,1}t_{2,3}t_{3,2}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} \\
&\quad + t_{1,2}t_{2,1}t_{3,3}\begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} + t_{1,2}t_{2,3}t_{3,1}\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} \\
&\quad + t_{1,3}t_{2,1}t_{3,2}\begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} + t_{1,3}t_{2,2}t_{3,1}\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}
\end{aligned}$$

Pick a row or column and factor out its entries; here we do the entries in the first row.

$$\begin{aligned}
&= t_{1,1}\cdot\left[t_{2,2}t_{3,3}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} + t_{2,3}t_{3,2}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix}\right] \\
&\quad + t_{1,2}\cdot\left[t_{2,1}t_{3,3}\begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} + t_{2,3}t_{3,1}\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix}\right] \\
&\quad + t_{1,3}\cdot\left[t_{2,1}t_{3,2}\begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} + t_{2,2}t_{3,1}\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}\right]
\end{aligned}$$

In those permutation matrices, swap to get the first rows into place. This requires one swap to each of the permutation matrices on the second line, and two swaps to each on the third line. (Recall that row swaps change the sign of the determinant.)

$$
\begin{aligned}
&= t_{1,1}\cdot\left[t_{2,2}t_{3,3}\begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix}+t_{2,3}t_{3,2}\begin{vmatrix}1&0&0\\0&0&1\\0&1&0\end{vmatrix}\right]\\
&\quad - t_{1,2}\cdot\left[t_{2,1}t_{3,3}\begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix}+t_{2,3}t_{3,1}\begin{vmatrix}1&0&0\\0&0&1\\0&1&0\end{vmatrix}\right]\\
&\quad + t_{1,3}\cdot\left[t_{2,1}t_{3,2}\begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix}+t_{2,2}t_{3,1}\begin{vmatrix}1&0&0\\0&0&1\\0&1&0\end{vmatrix}\right]
\end{aligned}
$$

On each line the terms in square brackets involve only the second and third row and column, and simplify to a 2×2 determinant.

$$= t_{1,1}\cdot\begin{vmatrix}t_{2,2}&t_{2,3}\\t_{3,2}&t_{3,3}\end{vmatrix} - t_{1,2}\cdot\begin{vmatrix}t_{2,1}&t_{2,3}\\t_{3,1}&t_{3,3}\end{vmatrix} + t_{1,3}\cdot\begin{vmatrix}t_{2,1}&t_{2,2}\\t_{3,1}&t_{3,2}\end{vmatrix}$$

The formula given in Theorem 1.5, which generalizes this example, is a *recurrence* — the determinant is expressed as a combination of determinants. This formula isn't circular because it gives the $n\times n$ case in terms of smaller ones.

1.2 Definition For any $n\times n$ matrix T, the $(n-1)\times(n-1)$ matrix formed by deleting row i and column j of T is the i,j *minor* of T. The i,j *cofactor* $T_{i,j}$ of T is $(-1)^{i+j}$ times the determinant of the i,j minor of T.

1.3 Example The $1,2$ cofactor of the matrix from Example 1.1 is the negative of the second 2×2 determinant.

$$T_{1,2} = -1\cdot\begin{vmatrix}t_{2,1}&t_{2,3}\\t_{3,1}&t_{3,3}\end{vmatrix}$$

1.4 Example Where

$$T=\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}$$

these are the $1,2$ and $2,2$ cofactors.

$$T_{1,2}=(-1)^{1+2}\cdot\begin{vmatrix}4&6\\7&9\end{vmatrix}=6\qquad T_{2,2}=(-1)^{2+2}\cdot\begin{vmatrix}1&3\\7&9\end{vmatrix}=-12$$

1.5 Theorem (Laplace Expansion of Determinants) Where T is an $n \times n$ matrix, we can find the determinant by expanding by cofactors on any row i or column j.

$$\begin{aligned}|T| &= t_{i,1} \cdot T_{i,1} + t_{i,2} \cdot T_{i,2} + \cdots + t_{i,n} \cdot T_{i,n} \\ &= t_{1,j} \cdot T_{1,j} + t_{2,j} \cdot T_{2,j} + \cdots + t_{n,j} \cdot T_{n,j}\end{aligned}$$

Proof Exercise 25. QED

1.6 Example We can compute the determinant

$$|T| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

by expanding along the first row, as in Example 1.1.

$$|T| = 1 \cdot (+1) \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} + 2 \cdot (-1) \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \cdot (+1) \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = -3 + 12 - 9 = 0$$

Or, we could expand down the second column.

$$|T| = 2 \cdot (-1) \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 5 \cdot (+1) \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} + 8 \cdot (-1) \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = 12 - 60 + 48 = 0$$

1.7 Example A row or column with many zeroes suggests a Laplace expansion.

$$\begin{vmatrix} 1 & 5 & 0 \\ 2 & 1 & 1 \\ 3 & -1 & 0 \end{vmatrix} = 0 \cdot (+1) \begin{vmatrix} 2 & 1 \\ 3 & -1 \end{vmatrix} + 1 \cdot (-1) \begin{vmatrix} 1 & 5 \\ 3 & -1 \end{vmatrix} + 0 \cdot (+1) \begin{vmatrix} 1 & 5 \\ 2 & 1 \end{vmatrix} = 16$$

We finish by applying Laplace's expansion to derive a new formula for the inverse of a matrix. With Theorem 1.5, we can calculate the determinant of a matrix by taking linear combinations of entries from a row with their associated cofactors.

$$t_{i,1} \cdot T_{i,1} + t_{i,2} \cdot T_{i,2} + \cdots + t_{i,n} \cdot T_{i,n} = |T| \qquad (*)$$

Recall that a matrix with two identical rows has a zero determinant. Thus, weighing the cofactors by entries from row k with $k \neq i$ gives zero

$$t_{i,1} \cdot T_{k,1} + t_{i,2} \cdot T_{k,2} + \cdots + t_{i,n} \cdot T_{k,n} = 0 \qquad (**)$$

because it represents the expansion along the row k of a matrix with row i equal to row k. This summarizes $(*)$ and $(**)$.

$$\begin{pmatrix} t_{1,1} & t_{1,2} & \ldots & t_{1,n} \\ t_{2,1} & t_{2,2} & \ldots & t_{2,n} \\ & \vdots & & \\ t_{n,1} & t_{n,2} & \ldots & t_{n,n} \end{pmatrix} \begin{pmatrix} T_{1,1} & T_{2,1} & \ldots & T_{n,1} \\ T_{1,2} & T_{2,2} & \ldots & T_{n,2} \\ & \vdots & & \\ T_{1,n} & T_{2,n} & \ldots & T_{n,n} \end{pmatrix} = \begin{pmatrix} |T| & 0 & \ldots & 0 \\ 0 & |T| & \ldots & 0 \\ & \vdots & & \\ 0 & 0 & \ldots & |T| \end{pmatrix}$$

Note that the order of the subscripts in the matrix of cofactors is opposite to the order of subscripts in the other matrix; e.g., along the first row of the matrix of cofactors the subscripts are $1, 1$ then $2, 1$, etc.

1.8 Definition The matrix *adjoint* to the square matrix T is

$$\mathrm{adj}(T) = \begin{pmatrix} T_{1,1} & T_{2,1} & \ldots & T_{n,1} \\ T_{1,2} & T_{2,2} & \ldots & T_{n,2} \\ & \vdots & & \\ T_{1,n} & T_{2,n} & \ldots & T_{n,n} \end{pmatrix}$$

where $T_{j,i}$ is the j, i cofactor.

1.9 Theorem Where T is a square matrix, $T \cdot \mathrm{adj}(T) = \mathrm{adj}(T) \cdot T = |T| \cdot I$. Thus if T has an inverse, if $|T| \neq 0$, then $T^{-1} = (1/|T|) \cdot \mathrm{adj}(T)$.

Proof Equations $(*)$ and $(**)$. QED

1.10 Example If

$$T = \begin{pmatrix} 1 & 0 & 4 \\ 2 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}$$

then $\mathrm{adj}(T)$ is

$$\begin{pmatrix} T_{1,1} & T_{2,1} & T_{3,1} \\ T_{1,2} & T_{2,2} & T_{3,2} \\ T_{1,3} & T_{2,3} & T_{3,3} \end{pmatrix} = \begin{pmatrix} \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} & -\begin{vmatrix} 0 & 4 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 0 & 4 \\ 1 & -1 \end{vmatrix} \\ -\begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 4 \\ 1 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 4 \\ 2 & -1 \end{vmatrix} \\ \begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -4 \\ -3 & -3 & 9 \\ -1 & 0 & 1 \end{pmatrix}$$

and taking the product with T gives the diagonal matrix $|T| \cdot I$.

$$\begin{pmatrix} 1 & 0 & 4 \\ 2 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -4 \\ -3 & -3 & 9 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

The inverse of T is $(1/-3) \cdot \mathrm{adj}(T)$.

$$T^{-1} = \begin{pmatrix} 1/-3 & 0/-3 & -4/-3 \\ -3/-3 & -3/-3 & 9/-3 \\ -1/-3 & 0/-3 & 1/-3 \end{pmatrix} = \begin{pmatrix} -1/3 & 0 & 4/3 \\ 1 & 1 & -3 \\ 1/3 & 0 & -1/3 \end{pmatrix}$$

The formulas from this subsection are often used for by-hand calculation and are sometimes useful with special types of matrices. However, for generic matrices they are not the best choice because they require more arithmetic than, for instance, the Gauss-Jordan method.

Exercises

✓ 1.11 Find the cofactor.

$$T = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 1 & 3 \\ 0 & 2 & -1 \end{pmatrix}$$

(a) $T_{2,3}$ **(b)** $T_{3,2}$ **(c)** $T_{1,3}$

✓ 1.12 Find the determinant by expanding

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

(a) on the first row **(b)** on the second row **(c)** on the third column.

1.13 Find the adjoint of the matrix in Example 1.6.

✓ 1.14 Find the matrix adjoint to each.

(a) $\begin{pmatrix} 2 & 1 & 4 \\ -1 & 0 & 2 \\ 1 & 0 & 1 \end{pmatrix}$ **(b)** $\begin{pmatrix} 3 & -1 \\ 2 & 4 \end{pmatrix}$ **(c)** $\begin{pmatrix} 1 & 1 \\ 5 & 0 \end{pmatrix}$ **(d)** $\begin{pmatrix} 1 & 4 & 3 \\ -1 & 0 & 3 \\ 1 & 8 & 9 \end{pmatrix}$

✓ 1.15 Find the inverse of each matrix in the prior question with Theorem 1.9.

1.16 Find the matrix adjoint to this one.

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

✓ 1.17 Expand across the first row to derive the formula for the determinant of a 2×2 matrix.

✓ 1.18 Expand across the first row to derive the formula for the determinant of a 3×3 matrix.

✓ 1.19 **(a)** Give a formula for the adjoint of a 2×2 matrix.
(b) Use it to derive the formula for the inverse.

✓ 1.20 Can we compute a determinant by expanding down the diagonal?

1.21 Give a formula for the adjoint of a diagonal matrix.

✓ 1.22 Prove that the transpose of the adjoint is the adjoint of the transpose.

1.23 Prove or disprove: $\text{adj}(\text{adj}(T)) = T$.

1.24 A square matrix is *upper triangular* if each i, j entry is zero in the part above the diagonal, that is, when $i > j$.

(a) Must the adjoint of an upper triangular matrix be upper triangular? Lower triangular?

(b) Prove that the inverse of a upper triangular matrix is upper triangular, if an inverse exists.

1.25 *This question requires material from the optional Determinants Exist subsection.* Prove Theorem 1.5 by using the permutation expansion.

1.26 Prove that the determinant of a matrix equals the determinant of its transpose using Laplace's expansion and induction on the size of the matrix.

? 1.27 Show that

$$F_n = \begin{vmatrix} 1 & -1 & 1 & -1 & 1 & -1 & \ldots \\ 1 & 1 & 0 & 1 & 0 & 1 & \ldots \\ 0 & 1 & 1 & 0 & 1 & 0 & \ldots \\ 0 & 0 & 1 & 1 & 0 & 1 & \ldots \\ . & . & . & . & . & . & \ldots \end{vmatrix}$$

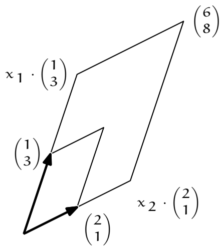
where F_n is the n-th term of $1, 1, 2, 3, 5, \ldots, x, y, x + y, \ldots$, the Fibonacci sequence, and the determinant is of order $n - 1$. [Am. Math. Mon., Jun. 1949]

Cramer's Rule

A linear system is equivalent to a linear relationship among vectors.

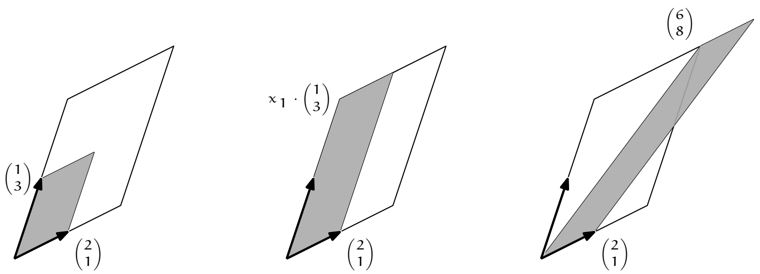
$$\begin{aligned} x_1 + 2x_2 &= 6 \\ 3x_1 + x_2 &= 8 \end{aligned} \qquad \Longleftrightarrow \qquad x_1 \cdot \begin{pmatrix}1\\3\end{pmatrix} + x_2 \cdot \begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}6\\8\end{pmatrix}$$

In the picture below the small parallelogram is formed from sides that are the vectors $\binom{1}{3}$ and $\binom{2}{1}$. It is nested inside a parallelogram with sides $x_1\binom{1}{3}$ and $x_2\binom{2}{1}$. By the vector equation, the far corner of the larger parallelogram is $\binom{6}{8}$.



This drawing restates the algebraic question of finding the solution of a linear system into geometric terms: by what factors x_1 and x_2 must we dilate the sides of the starting parallelogram so that it will fill the other one?

We can use this picture, and our geometric understanding of determinants, to get a new formula for solving linear systems. Compare the sizes of these shaded boxes.



The second is defined by the vectors $x_1\binom{1}{3}$ and $\binom{2}{1}$ and one of the properties of the size function — the determinant — is that therefore the size of the second box is x_1 times the size of the first. The third box is derived from the second by shearing, adding $x_2\binom{2}{1}$ to $x_1\binom{1}{3}$ to get $x_1\binom{1}{3}+x_2\binom{2}{1}=\binom{6}{8}$, along with $\binom{2}{1}$. The determinant is not affected by shearing so the size of the third box equals that of the second.

Taken together, we have this.

$$\begin{vmatrix} 6 & 2 \\ 8 & 1 \end{vmatrix} = \begin{vmatrix} x_1\cdot 1 & 2 \\ x_1\cdot 3 & 1 \end{vmatrix} = x_1\cdot\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix}$$

Solving gives the value of one of the variables.

$$x_1 = \frac{\begin{vmatrix} 6 & 2 \\ 8 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix}} = \frac{-10}{-5} = 2$$

The generalization of this example is *Cramer's Rule*: if $|A|\neq 0$ then the system $A\vec{x}=\vec{b}$ has the unique solution $x_i=|B_i|/|A|$ where the matrix B_i is formed from A by replacing column i with the vector $\vec{b}$. The proof is Exercise 3.

For instance, to solve this system for x_2

$$\begin{pmatrix} 1 & 0 & 4 \\ 2 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

we do this computation.

$$x_2 = \frac{\begin{vmatrix} 1 & 2 & 4 \\ 2 & 1 & -1 \\ 1 & -1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & 4 \\ 2 & 1 & -1 \\ 1 & 0 & 1 \end{vmatrix}} = \frac{-18}{-3}$$

Cramer's Rule lets us by-eye solve systems that are small and simple. For example, we can solve systems with two equations and two unknowns, or three equations and three unknowns, where the numbers are small integers. Such cases appear often enough that many people find this formula handy.

But using it to solving large or complex systems is not practical, either by hand or by a computer. A Gauss's Method-based approach is faster.

Exercises

1 Use Cramer's Rule to solve each for each of the variables.

(a) $\begin{aligned} x - y &= 4 \\ -x + 2y &= -7 \end{aligned}$ (b) $\begin{aligned} -2x + y &= -2 \\ x - 2y &= -2 \end{aligned}$

2 Use Cramer's Rule to solve this system for z.

$$\begin{aligned} 2x + y + z &= 1 \\ 3x \phantom{{}+y} + z &= 4 \\ x - y - z &= 2 \end{aligned}$$

3 Prove Cramer's Rule.

4 Here is an alternative proof of Cramer's Rule that doesn't overtly contain any geometry. Write X_i for the identity matrix with column i replaced by the vector $\vec{x}$ of unknowns $x_1, \ldots, x_n$.

(a) Observe that $AX_i = B_i$.

(b) Take the determinant of both sides.

5 Suppose that a linear system has as many equations as unknowns, that all of its coefficients and constants are integers, and that its matrix of coefficients has determinant 1. Prove that the entries in the solution are all integers. (*Remark.* This is often used to invent linear systems for exercises.)

6 Use Cramer's Rule to give a formula for the solution of a two equations/two unknowns linear system.

7 Can Cramer's Rule tell the difference between a system with no solutions and one with infinitely many?

8 The first picture in this Topic (the one that doesn't use determinants) shows a unique solution case. Produce a similar picture for the case of infinitely many solutions, and the case of no solutions.

Topic

Speed of Calculating Determinants

The permutation expansion formula for computing determinants is useful for proving theorems but the method of using row operations is faster for finding the determinants of a large matrix. We can make this statement precise by finding how many operations each method performs.

We compare the speed of two algorithms by finding for each one how the time taken grows as the size of its input data set grows. For instance, if we increase the size of the input by a factor of ten does the time taken grow by a factor of ten, or by a factor of a hundred, or by a factor of a thousand? That is, is the time proportional to the size of the data set, or to the square of that size, or to the cube of that size, etc.? An algorithm whose time taken is proportional to the square is faster than one whose time taken is proportional to the cube.

In the permutation expansion formula

$$\begin{vmatrix} t_{1,1} & t_{1,2} & \dots & t_{1,n} \\ t_{2,1} & t_{2,2} & \dots & t_{2,n} \\ & \vdots & & \\ t_{n,1} & t_{n,2} & \dots & t_{n,n} \end{vmatrix} = \sum_{\text{permutations } \phi} t_{1,\phi(1)} t_{2,\phi(2)} \cdots t_{n,\phi(n)} \left|P_\phi\right|$$

there are $n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$ different n-permutations. This factorial function grows quickly; when n is only 10 then the expansion above has $10! = 3,628,800$ terms, each with n multiplications and a permutation matrix to do. Observe that performing $n!$ many operations is performing more than n^2 many operations because multiplying the first two factors in $n!$ gives $n \cdot (n-1)$, which for large n is approximately n^2 and then multiplying in more factors will make the factorial even larger. Similarly, the factorial function grows faster than the cube, or the fourth power, or any polynomial function. So a computer program that uses the permutation expansion formula, and thus performs a number of operations that is greater than or equal to the factorial of the number of rows, would be very slow.

In contrast, the time taken by the row reduction method does not grow so fast. Below is a fragment of row-reduction code. It is in the computer language FORTRAN, which is widely used for numeric code.

```
DO 10 ROW=1, N-1
  PIVINV=1.0/A(ROW,ROW)
  DO 20 I=ROW+1, N
    DO 30 J=I, N
      A(I,J)=A(I,J)-PIVINV*A(ROW,J)
    30 CONTINUE
  20 CONTINUE
10 CONTINUE
```

The matrix is in the $\mathtt{N}\times\mathtt{N}$ array `A`. The program's outer loop runs through each `ROW` between `1` and `N-1` and does the entry-by-entry combination $-\mathtt{PIVINV}\cdot\rho_{\mathtt{ROW}}+\rho_{\mathtt{I}}$ with the lower rows.

For each `ROW`, the nested `I` and `J` loops perform the combination with the lower rows by doing arithmetic on the entries in `A` that are below and to the right of `A(ROW,ROW)`. There are $(\mathtt{N}-\mathtt{ROW})^2$ such entries. On average, `ROW` will be $\mathtt{N}/2$. So this program will perform the arithmetic about $(\mathtt{N}/2)^2$ times. Taking into account the outer loop gives the estimate that the running time of the algorithm is proportional to the cube of the number of equations. (The code above is naive; for example it does not handle the case that the entry `A(ROW,ROW)` is zero. Analysis of a finished version that includes all of the tests and subcases is messier but gives the same conclusion.)

Finding the fastest algorithm to compute the determinant is a topic of current research. So far, people have found algorithms that run in time between the square and cube of `N`.

The contrast between the methods of permutation expansion and row operations for computing determinants makes the point that although in principle they give the same answer, in practice we want the one with the best performance.

Exercises

1 Computer systems generate random numbers (of course, these are only pseudo-random in that they come from an algorithm but they pass a number of reasonable statistical tests for randomness).

(a) Fill a 5×5 array with random numbers say, in the range $[0\,..\,1))$. See if it is singular. Repeat that experiment a few times. Are singular matrices frequent or rare (in this sense)?

(b) Time your computer algebra system at finding the determinant of ten 5×5 arrays of random numbers. Find the average time per array. Repeat the prior item for 15×15 arrays, 25×25 arrays, 35×35 arrays, etc. You may find that you need to get above a certain size to get a timing that you can use. (Notice that, when an array is singular, we can sometimes decide that quickly, for instance if the first row equals the second. In the light of your answer to the first part, do you expect that singular systems play a large role in your average?)

(c) Graph the input size versus the average time.

2 Compute the determinant of each of these by hand using the two methods discussed above.

(a) $\begin{vmatrix} 2 & 1 \\ 5 & -3 \end{vmatrix}$ **(b)** $\begin{vmatrix} 3 & 1 & 1 \\ -1 & 0 & 5 \\ -1 & 2 & -2 \end{vmatrix}$ **(c)** $\begin{vmatrix} 2 & 1 & 0 & 0 \\ 1 & 3 & 2 & 0 \\ 0 & -1 & -2 & 1 \\ 0 & 0 & -2 & 1 \end{vmatrix}$

Count the number of multiplications and divisions used in each case, for each of the methods. (Computer multiplications and divisions take longer than additions and subtractions, so algorithm designers worry about them more.)

3 What 10×10 array can you invent that takes your computer the longest time to reduce? The shortest?

4 The FORTRAN language specification requires that arrays be stored "by column," that is, the entire first column is stored contiguously, then the second column, etc. Does the code fragment given take advantage of this, or can it be rewritten to make it faster, by taking advantage of the fact that computer fetches are faster from contiguous locations?

Topic

Chiò's Method

When doing Gauss's Method on a matrix that contains only integers people often like to keep it that way. To avoid fractions in the reduction of this matrix

$$A=\begin{pmatrix}2&1&1\\3&4&-1\\1&5&1\end{pmatrix}$$

they may start by multiplying the lower rows by 2

$$\xrightarrow[2\rho_3]{2\rho_2}\begin{pmatrix}2&1&1\\6&8&-2\\2&10&2\end{pmatrix}\qquad(*)$$

so that elimination in the first column goes like this.

$$\xrightarrow[-\rho_1+\rho_3]{-3\rho_1+\rho_2}\begin{pmatrix}2&1&1\\0&5&-5\\0&8&0\end{pmatrix}\qquad(**)$$

This all-integer approach is easier for mental calculations. And, using integer arithmetic on a computer avoids some sticky issues involving floating point calculations [Kahan]. So there are sound reasons for this approach.

Another advantage of this approach is that we can easily apply Laplace's expansion to the first column of $(**)$ and then get the determinant by remembering to divide by 4 because of $(*)$.

Here is the general 3×3 case of this approach to finding the determinant. First, assuming $a_{1,1}\neq0$, we can rescale the lower rows.

$$A=\begin{pmatrix}a_{1,1}&a_{1,2}&a_{1,3}\\a_{2,1}&a_{2,2}&a_{2,3}\\a_{3,1}&a_{3,2}&a_{3,3}\end{pmatrix}\xrightarrow[a_{1,1}\rho_3]{a_{1,1}\rho_2}\begin{pmatrix}a_{1,1}&a_{1,2}&a_{1,3}\\a_{2,1}a_{1,1}&a_{2,2}a_{1,1}&a_{2,3}a_{1,1}\\a_{3,1}a_{1,1}&a_{3,2}a_{1,1}&a_{3,3}a_{1,1}\end{pmatrix}$$

This rescales the determinant by $a_{1,1}^2$. Now eliminate down the first column.

$$\xrightarrow[-a_{3,1}\rho_1+\rho_3]{-a_{2,1}\rho_1+\rho_2} \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ 0 & a_{2,2}a_{1,1}-a_{2,1}a_{1,2} & a_{2,3}a_{1,1}-a_{2,1}a_{1,3} \\ 0 & a_{3,2}a_{1,1}-a_{3,1}a_{1,2} & a_{3,3}a_{1,1}-a_{3,1}a_{1,3} \end{pmatrix}$$

Let C be the $1,1$ minor. By Laplace the determinant of the above matrix is $a_{1,1}\det(C)$. We thus have $a_{1,1}^2\det(A) = a_{1,1}\det(C)$ and since $a_{1,1} \neq 0$ this gives $\det(A) = \det(C)/a_{1,1}$.

To do larger matrices we must see how to compute the minor's entries. The pattern above is that each element of the minor is a 2×2 determinant. For instance, the entry in the minor's upper left $a_{2,2}a_{1,1} - a_{2,1}a_{1,2}$, which is the $2,2$ entry in the above matrix, is the determinant of the matrix of these four elements of A.

$$\begin{pmatrix} \boxed{a_{1,1}} & \boxed{a_{1,2}} & a_{1,3} \\ \boxed{a_{2,1}} & \boxed{a_{2,2}} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

And the minor's lower left, the $3,2$ entry from above, is the determinant of the matrix of these four.

$$\begin{pmatrix} \boxed{a_{1,1}} & \boxed{a_{1,2}} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ \boxed{a_{3,1}} & \boxed{a_{3,2}} & a_{3,3} \end{pmatrix}$$

So, where A is $n\times n$ for $n \geqslant 3$, we let Chiò's matrix C be the $(n-1)\times(n-1)$ matrix whose i,j entry is the determinant

$$\begin{vmatrix} a_{1,1} & a_{1,j+1} \\ a_{i+1,1} & a_{i+1,j+1} \end{vmatrix}$$

where $1 < i,j \leqslant n$. *Chiò's method* for finding the determinant of A is that if $a_{1,1} \neq 0$ then $\det(A) = \det(C)/a_{1,1}^{n-2}$. (By the way, nothing in Chiò's formula requires that the numbers be integers; it applies to reals as well.)

To illustrate we find the determinant of this 3×3 matrix.

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 4 & -1 \\ 1 & 5 & 1 \end{pmatrix}$$

This is Chiò's matrix.

$$C = \begin{pmatrix} \begin{vmatrix} 2 & 1 \\ 3 & 4 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ 3 & -1 \end{vmatrix} \\ \begin{vmatrix} 2 & 1 \\ 1 & 5 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} 5 & -5 \\ 9 & 1 \end{pmatrix}$$

The formula for $3{\times}3$ matrices $\det(A) = \det(C)/a_{1,1}$ gives $\det(A) = (50/2) = 25$.

For a larger determinant we must do multiple steps but each involves only $2{\times}2$ determinants. So we can often calculate the determinant just by writing down a bit of intermediate information. For instance, with this $4{\times}4$ matrix

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

we can mentally doing each of the $2{\times}2$ calculations and only write down the $3{\times}3$ result.

$$C_3 = \begin{pmatrix} \begin{vmatrix} 3 & 0 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 1 & 0 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} \\ \begin{vmatrix} 3 & 0 \\ 2 & -1 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 2 & 0 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 2 & 3 \end{vmatrix} \\ \begin{vmatrix} 3 & 0 \\ 1 & 0 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 1 & 0 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} 6 & -1 & 2 \\ -3 & -2 & 7 \\ 0 & -1 & 2 \end{pmatrix}$$

Note that the determinant of this is $a_{1,1}^{4-2} = 3^2$ times the determinant of A.

To finish, iterate. Here is Chiò's matrix of C_3.

$$C_2 = \begin{pmatrix} \begin{vmatrix} 6 & -1 \\ -3 & -2 \end{vmatrix} & \begin{vmatrix} 6 & 2 \\ -3 & 7 \end{vmatrix} \\ \begin{vmatrix} 6 & -1 \\ 0 & -1 \end{vmatrix} & \begin{vmatrix} 6 & 2 \\ 0 & 2 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -15 & 48 \\ -6 & 12 \end{pmatrix}$$

The determinant of this matrix is 6 times the determinant of C_3. The determinant of C_2 is 108. So $\det(A) = 108/(3^2 \cdot 6) = 2$.

Laplace's expansion formula reduces the calculation of an $n{\times}n$ determinant to the evaluation of a number of $(n-1){\times}(n-1)$ ones. Chiò's formula is also recursive but it reduces an $n{\times}n$ determinant to a single $(n-1){\times}(n-1)$ determinant, calculated from a number of $2{\times}2$ determinants. However, for large matrices Gauss's Method is better than either of these; for instance, it takes roughly half as many operations as Chiò's Method [Fuller & Logan].

Exercises

1 Use Chiò's Method to find each determinant.

(a) $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$ **(b)** $\begin{vmatrix} 2 & 1 & 4 & 0 \\ 0 & 1 & 4 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \end{vmatrix}$

2 What if $a_{1,1}$ is zero?

3 The *Rule of Sarrus* is a mnemonic that many people learn for the $3{\times}3$ determinant formula. To the right of the matrix, copy the first two columns.

$$\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}$$

Then the determinant is the sum of the three upper-left to lower-right diagonals minus the three lower-left to upper-right diagonals $aei+bfg+cdh-gec-hfa-idb$. Count the operations involved in Sarrus's formula and Chiò's formula for the 3×3 case and see which uses fewer.

4 Prove Chiò's formula.

Computer Code

This implements Chiò's Method. It is in the computer language Python (to make it more readable the code avoids some Python facilities). Note the recursive call in the final line of `chio_det`.

```
#!/usr/bin/python
# chio.py
#  Calculate a determinant using Chio's method.
# Jim Hefferon; Public Domain
# For demonstration only; for instance, does not handle the M[0][0]=0 case

def det_two(a,b,c,d):
    """Return the determinant of the 2x2 matrix [[a,b], [c,d]]"""
    return a*d-b*c

def chio_mat(M):
    """Return the Chio matrix as a list of the rows
        M  nxn matrix, list of rows"""
    dim=len(M)
    C=[]
    for row in range(1,dim):
        C.append([])
        for col in range(1,dim):
            C[-1].append(det_two(M[0][0], M[0][col], M[row][0], M[row][col]))
    return C

def chio_det(M,show=None):
    """Find the determinant of M by Chio's method
        M  mxm matrix, list of rows"""
    dim=len(M)
    key_elet=M[0][0]
    if dim==1:
        return key_elet
    return chio_det(chio_mat(M))/(key_elet**(dim-2))


if __name__=='__main__':
    M=[[2,1,1], [3,4,-1], [1,5,1]]
    print "M=",M
    print "Det is", chio_det(M)
```

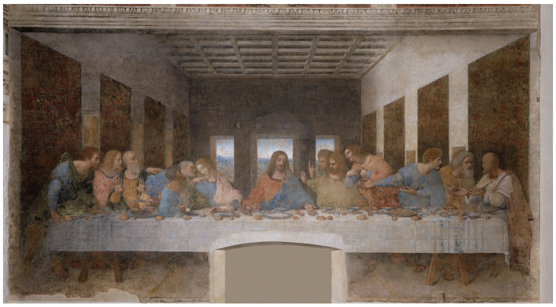
This is the result of calling the program from my command line.

```
$ python chio.py
M=[[2, 1, 1], [3, 4, -1], [1, 5, 1]]
Det is 25
```

Topic

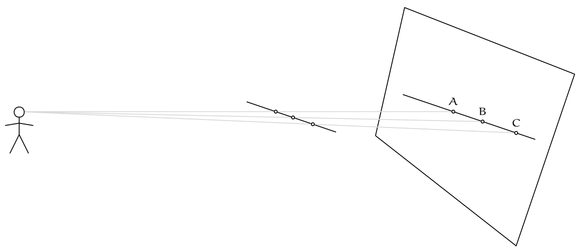
Projective Geometry

There are geometries other than the familiar Euclidean one. One such geometry arose when artists observed that what a viewer sees is not necessarily what is there. As an example, here is Leonardo da Vinci's *The Last Supper*.



Look at where the ceiling meets the left and right walls. In the room those lines are parallel but da Vinci has painted lines that, if extended, would intersect. The intersection is the *vanishing point*. This aspect of perspective is familiar as an image of railroad tracks that appear to converge at the horizon.

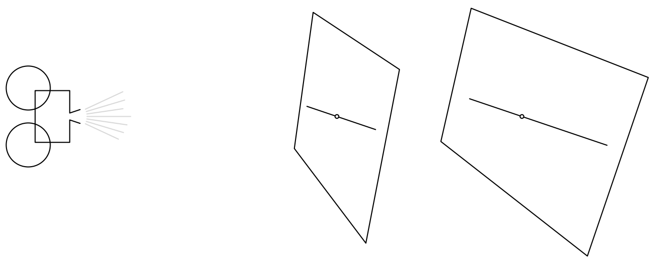
Da Vinci has adopted a model of how we see. Imagine a person viewing a room. From the person's eye, in every direction, carry a ray outward until it intersects something, such as a point on the line where the wall meets the ceiling. This first intersection point is what the person sees in that direction. Overall what the person sees is the collection of three-dimensional intersection points projected to a common two dimensional image.



This is a *central projection* from a single point. As the sketch shows, this projection is not orthogonal like the ones we have seen earlier because the line from the viewer to C is not orthogonal to the image plane. (This model is only an approximation — it does not take into account such factors as that we have binocular vision or that our brain's processing greatly affects what we perceive. Nonetheless the model is interesting, both artistically and mathematically.)

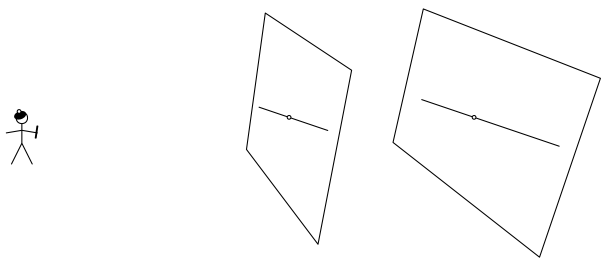
The operation of central projection preserves some geometric properties, for instance lines project to lines. However, it fails to preserve some others. One example is that equal length segments can project to segments of unequal length (above, AB is longer than BC because the segment projected to AB is closer to the viewer and closer things look bigger). The study of the effects of central projections is projective geometry.

There are three cases of central projection. The first is the projection done by a movie projector.



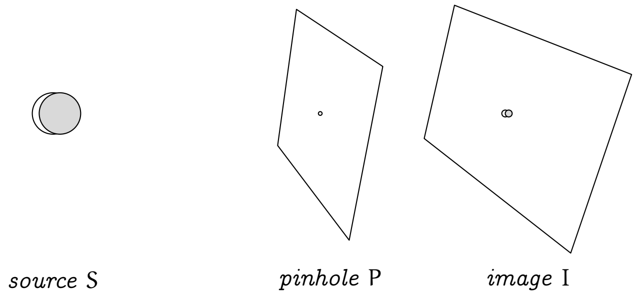
projector P *source* S *image* I

We can think that each source point is pushed from the domain plane S outward to the image plane I. The second case of projection is that of the artist pulling the source back to a canvas.



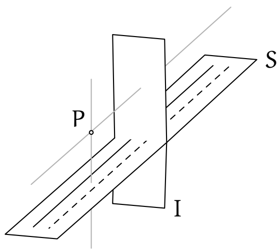
painter P *image* I *source* S

The two are different because first S is in the middle and then I. One more configuration can happen, with P in the middle. An example of this is when we use a pinhole to shine the image of a solar eclipse onto a paper.

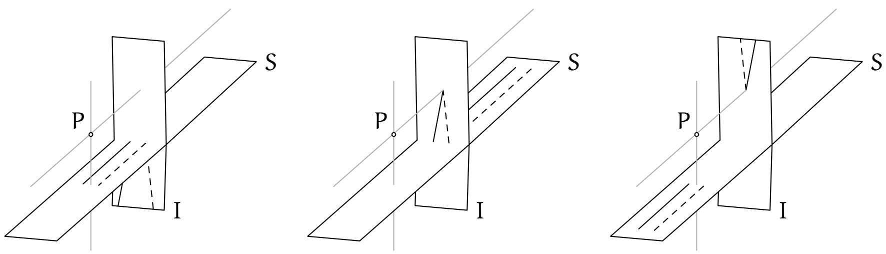


Although the three are not exactly the same, they are similar. We shall say that each is a central projection by P of S to I. We next look at three models of central projection, of increasing abstractness but also of increasing uniformity. The last model will bring out the linear algebra.

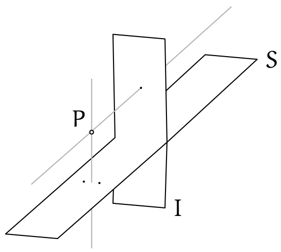
Consider again the effect of railroad tracks that appear to converge to a point. Model this with parallel lines in a domain plane S and a projection via a P to a codomain plane I. (The gray lines shown are parallel to the S plane and to the I plane.)



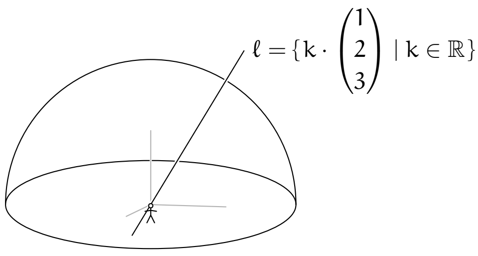
This single setting shows all three projection cases. The first picture below shows P acting as a movie projector by pushing points from part of S out to image points on the lower half of I. The middle picture shows P acting as the artist by pulling points from another part of S back to image points in the middle of I. In the third picture P acts as the pinhole, projecting points from S to the upper part of I. This third picture is the trickiest — the points that are projected near to the vanishing point are the ones that are far out on the lower left of S. Points in S that are near to the vertical gray line are sent high up on I.



There are two awkward things here. First, neither of the two points in the domain nearest to the vertical gray line (see below) has an image because a projection from those two is along the gray line that is parallel to the codomain plane (we say that these two are projected to infinity). The second is that the vanishing point in I isn't the image of any point from S because a projection to this point would be along the gray line that is parallel to the domain plane (we say that the vanishing point is the image of a projection from infinity).



For a model that eliminates this awkwardness, cover the projector P with a hemispheric dome. In any direction, defined by a line through the origin, project anything in that direction to the single spot on the dome where the line intersects. This includes projecting things on the line between P and the dome, as with the movie projector. It includes projecting things on the line further from P than the dome, as with the painter. More subtly, it also includes things on the line that lie behind P, as with the pinhole case.



More formally, for any nonzero vector $\vec{v} \in \mathbb{R}^3$, let the associated *point v in the projective plane* be the set $\{k\vec{v} \mid k \in \mathbb{R} \text{ and } k \neq 0\}$ of nonzero vectors lying on the same line through the origin as $\vec{v}$. To describe a projective point we can give any representative member of the line, so that the projective point shown above can be represented in any of these three ways.

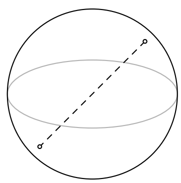
$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \begin{pmatrix} 1/3 \\ 2/3 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} -2 \\ -4 \\ -6 \end{pmatrix}$$

Each of these is a *homogeneous coordinate vector* for the point ℓ.

This picture and definition clarifies central projection but there is still something ungainly about the dome model: what happens when P looks down? Consider, in the sketch above, the part of P's line of sight that comes up towards us, out of the page. Imagine that this part of the line falls, to the equator and below. Now the part of the line ℓ that intersects the dome lies behind the page.

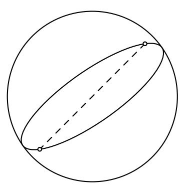
That is, as the line of sight continues down past the equator, the projective point suddenly shifts from the front of the dome to the back of the dome. (This brings out that the dome does not include the entire equator or else when the viewer is looking exactly along the equator then there would be two points in the line that are both on the dome. Instead we define the dome so that it includes the points on the equator with a positive y coordinate, as well as the point where $y = 0$ and x is positive.) This discontinuity means that we often have to treat equatorial points as a separate case. So while the railroad track model of central projection has three cases, the dome has two.

We can do better, we can reduce to a model having a single case. Consider a sphere centered at the origin. Any line through the origin intersects the sphere in two spots, said to be antipodal. Because we associate each line through the origin with a point in the projective plane, we can draw such a point as a pair of antipodal spots on the sphere. Below, we show the two antipodal spots connected by a dashed line to emphasize that they are not two different points, the pair of spots together make one projective point.



While drawing a point as a pair of antipodal spots on the sphere is not as intuitive as the one-spot-per-point dome mode, on the other hand the awkwardness of the dome model is gone in that as a line of view slides from north to south, no sudden changes happen. This central projection model is uniform.

So far we have described points in projective geometry. What about lines? What a viewer P at the origin sees as a line is shown below as a great circle, the intersection of the model sphere with a plane through the origin.



(We've included one of the projective points on this line to bring out a subtlety. Because two antipodal spots together make up a single projective point, the great circle's behind-the-paper part is the same set of projective points as its in-front-of-the-paper part.) Just as we did with each projective point, we can also describe a projective line with a triple of reals. For instance, the members of this plane through the origin in $\mathbb{R}^3$

$$\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x + y - z = 0\}$$

project to a line that we can describe with $(1 \;\; 1 \;\; -1)$ (using a row vector for this typographically distinguishes lines from points). In general, for any nonzero three-wide row vector $\vec{L}$ we define the associated *line in the projective plane*, to be the set $L = \{k\vec{L} \mid k \in \mathbb{R} \text{ and } k \neq 0\}$.

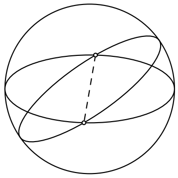
The reason this description of a line as a triple is convenient is that in the projective plane a point v and a line L are *incident* — the point lies on the line, the line passes through the point — if and only if a dot product of their representatives $v_1L_1 + v_2L_2 + v_3L_3$ is zero (Exercise 4 shows that this is independent of the choice of representatives $\vec{v}$ and $\vec{L}$). For instance, the projective point described above by the column vector with components 1, 2, and 3 lies in the projective line described by $(1 \;\; 1 \;\; -1)$, simply because any vector in $\mathbb{R}^3$ whose components are in ratio $1:2:3$ lies in the plane through the origin whose equation is of the form $k \cdot x + k \cdot y - k \cdot z = 0$ for any nonzero k. That is, the incidence formula is inherited from the three-space lines and planes of which v and L are projections.

With this, we can do analytic projective geometry. For instance, the projective line $L = (1 \;\; 1 \;\; -1)$ has the equation $1v_1 + 1v_2 - 1v_3 = 0$, meaning that for any projective point v incident with the line, any of v's representative homogeneous coordinate vectors will satisfy the equation. This is true simply because those vectors lie on the three space plane. One difference from Euclidean analytic geometry is that in projective geometry besides talking about the equation of a line, we also talk about the equation of a point. For the fixed point

$$v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

the property that characterizes lines incident on this point is that the components of any representatives satisfy $1L_1 + 2L_2 + 3L_3 = 0$ and so this is the equation of v.

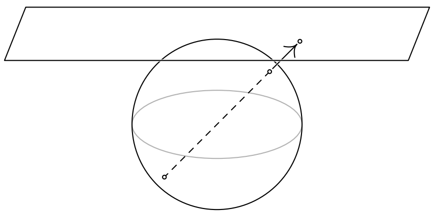
This symmetry of the statements about lines and points is the *Duality Principle* of projective geometry: in any true statement, interchanging 'point' with 'line' results in another true statement. For example, just as two distinct points determine one and only one line, in the projective plane two distinct lines determine one and only one point. Here is a picture showing two projective lines that cross in antipodal spots and thus cross at one projective point.



(∗)

Contrast this with Euclidean geometry, where two unequal lines may have a unique intersection or may be parallel. In this way, projective geometry is simpler, more uniform, than Euclidean geometry.

That simplicity is relevant because there is a relationship between the two spaces: we can view the projective plane as an extension of the Euclidean plane. Draw the sphere model of the projective plane as the unit sphere in $\mathbb{R}^3$. Take Euclidean 2-space to be the plane $z = 1$. As shown below, all of the points on the Euclidean plane are projections of antipodal spots from the sphere. Conversely, we can view some points in the projective plane as corresponding to points in Euclidean space. (Note that projective points on the equator don't correspond to points on the plane; instead we say these project out to infinity.)



(∗∗)

Thus we can think of projective space as consisting of the Euclidean plane with some extra points adjoined — the Euclidean plane is embedded in the projective plane. The extra points in projective space, the equatorial points, are called *ideal points* or *points at infinity* and the equator is called the *ideal line* or *line at infinity* (it is not a Euclidean line, it is a projective line).

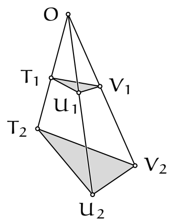
The advantage of this extension from the Euclidean plane to the projective plane is that some of the nonuniformity of Euclidean geometry disappears. For instance, the projective lines shown above in (∗) cross at antipodal spots, a single projective point, on the sphere's equator. If we put those lines into (∗∗) then they correspond to Euclidean lines that are parallel. That is, in moving from the Euclidean plane to the projective plane, we move from having two cases, that distinct lines either intersect or are parallel, to having only one case, that distinct lines intersect (possibly at a point at infinity).

A disadvantage of the projective plane is that we don't have the same familiarity with it as we have with the Euclidean plane. Doing analytic geometry in the projective plane helps because the equations lead us to the right conclusions. Analytic projective geometry uses linear algebra. For instance, for three points of the projective plane t, u, and v, setting up the equations for those points by fixing vectors representing each shows that the three are collinear if and only if the resulting three-equation system has infinitely many row vector solutions representing their line. That in turn holds if and only if this determinant is zero.

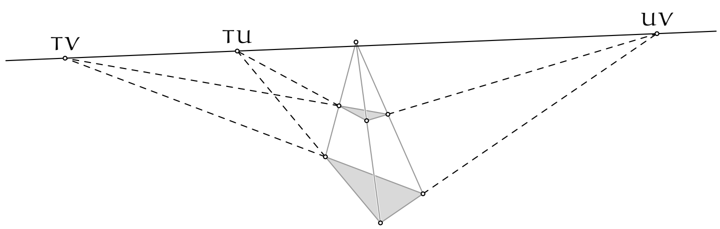
$$\begin{vmatrix} t_1 & u_1 & v_1 \\ t_2 & u_2 & v_2 \\ t_3 & u_3 & v_3 \end{vmatrix}$$

Thus, three points in the projective plane are collinear if and only if any three representative column vectors are linearly dependent. Similarly, by duality, three lines in the projective plane are incident on a single point if and only if any three row vectors representing them are linearly dependent.

The following result is more evidence of the niceness of the geometry of the projective plane. These two triangles are *in perspective* from the point O because their corresponding vertices are collinear.



Consider the pairs of corresponding sides: the sides T_1U_1 and T_2U_2, the sides T_1V_1 and T_2V_2, and the sides U_1V_1 and U_2V_2. *Desargue's Theorem* is that when we extend the three pairs of corresponding sides, they intersect (shown here as the points TU, TV, and UV). What's more, those three intersection points are collinear.



We will prove this using projective geometry. (We've drawn Euclidean figures because that is the more familiar image. To consider them as projective figures we can imagine that, although the line segments shown are parts of great circles and so are curved, the model has such a large radius compared to the size of the figures that the sides appear in our sketch to be straight.)

For the proof we need a preliminary lemma [Coxeter]: if W, X, Y, Z are four points in the projective plane, no three of which are collinear, then there are homogeneous coordinate vectors $\vec{w}$, $\vec{x}$, $\vec{y}$, and $\vec{z}$ for the projective points, and a basis B for $\mathbb{R}^3$, satisfying this.

$$\mathrm{Rep}_B(\vec{w}) = \begin{pmatrix}1\\0\\0\end{pmatrix} \quad \mathrm{Rep}_B(\vec{x}) = \begin{pmatrix}0\\1\\0\end{pmatrix} \quad \mathrm{Rep}_B(\vec{y}) = \begin{pmatrix}0\\0\\1\end{pmatrix} \quad \mathrm{Rep}_B(\vec{z}) = \begin{pmatrix}1\\1\\1\end{pmatrix}$$

To prove the lemma, because W, X, and Y are not on the same projective line, any homogeneous coordinate vectors $\vec{w}_0$, $\vec{x}_0$, and $\vec{y}_0$ do not line on the same plane through the origin in $\mathbb{R}^3$ and so form a spanning set for $\mathbb{R}^3$. Thus any homogeneous coordinate vector for Z is a combination $\vec{z}_0 = a\cdot\vec{w}_0 + b\cdot\vec{x}_0 + c\cdot\vec{y}_0$. Then let the basis be $B = \langle\vec{w}, \vec{x}, \vec{y}\rangle$ and take $\vec{w} = a\cdot\vec{w}_0$, $\vec{x} = b\cdot\vec{x}_0$, $\vec{y} = c\cdot\vec{y}_0$, and $\vec{z} = \vec{z}_0$.

To prove Desargue's Theorem use the lemma to fix homogeneous coordinate vectors and a basis.

$$\mathrm{Rep}_B(\vec{t}_1) = \begin{pmatrix}1\\0\\0\end{pmatrix} \quad \mathrm{Rep}_B(\vec{u}_1) = \begin{pmatrix}0\\1\\0\end{pmatrix} \quad \mathrm{Rep}_B(\vec{v}_1) = \begin{pmatrix}0\\0\\1\end{pmatrix} \quad \mathrm{Rep}_B(\vec{o}) = \begin{pmatrix}1\\1\\1\end{pmatrix}$$

The projective point T_2 is incident on the projective line OT_1 so any homogeneous coordinate vector for T_2 lies in the plane through the origin in $\mathbb{R}^3$ that is spanned by homogeneous coordinate vectors of O and T_1:

$$\mathrm{Rep}_B(\vec{t}_2) = a\begin{pmatrix}1\\1\\1\end{pmatrix} + b\begin{pmatrix}1\\0\\0\end{pmatrix}$$

for some scalars a and b. Hence the homogeneous coordinate vectors of members T_2 of the line OT_1 are of the form on the left below. The forms for U_2 and V_2 are similar.

$$\mathrm{Rep}_B(\vec{t}_2) = \begin{pmatrix}t_2\\1\\1\end{pmatrix} \qquad \mathrm{Rep}_B(\vec{u}_2) = \begin{pmatrix}1\\u_2\\1\end{pmatrix} \qquad \mathrm{Rep}_B(\vec{v}_2) = \begin{pmatrix}1\\1\\v_2\end{pmatrix}$$

The projective line T_1U_1 is the projection of a plane through the origin in $\mathbb{R}^3$. One way to get its equation is to note that any vector in it is linearly dependent on the vectors for T_1 and U_1 and so this determinant is zero.

$$\begin{vmatrix} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & z \end{vmatrix} = 0 \qquad \Longrightarrow \qquad z = 0$$

The equation of the plane in $\mathbb{R}^3$ whose image is the projective line T_2U_2 is this.

$$\begin{vmatrix} t_2 & 1 & x \\ 1 & u_2 & y \\ 1 & 1 & z \end{vmatrix} = 0 \qquad \Longrightarrow \qquad (1-u_2)\cdot x + (1-t_2)\cdot y + (t_2u_2-1)\cdot z = 0$$

Finding the intersection of the two is routine.

$$T_1U_1 \,\cap\, T_2U_2 = \begin{pmatrix} t_2-1 \\ 1-u_2 \\ 0 \end{pmatrix}$$

(This is, of course, a homogeneous coordinate vector of a projective point.) The other two intersections are similar.

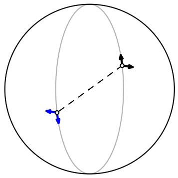
$$T_1V_1 \,\cap\, T_2V_2 = \begin{pmatrix} 1-t_2 \\ 0 \\ v_2-1 \end{pmatrix} \qquad U_1V_1 \,\cap\, U_2V_2 = \begin{pmatrix} 0 \\ u_2-1 \\ 1-v_2 \end{pmatrix}$$

Finish the proof by noting that these projective points are on one projective line because the sum of the three homogeneous coordinate vectors is zero.

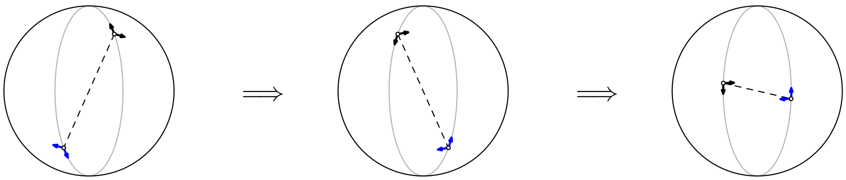
Every projective theorem has a translation to a Euclidean version, although the Euclidean result may be messier to state and prove. Desargue's theorem illustrates this. In the translation to Euclidean space, we must treat separately the case where O lies on the ideal line, for then the lines T_1T_2, U_1U_2, and V_1V_2 are parallel.

The remark following the statement of Desargue's Theorem suggests thinking of the Euclidean pictures as figures from projective geometry for a sphere model with very large radius. That is, just as a small area of the world seems to people living there to be flat, the projective plane is locally Euclidean.

We finish by pointing out one more thing about the projective plane. Although its local properties are familiar, the projective plane has a perhaps unfamiliar global property. The picture below shows a projective point. At that point we have drawn Cartesian axes, xy-axes. Of course, the axes appear in the picture at both antipodal spots, one in the northern hemisphere (that is, shown on the right) and the other in the south. Observe that in the northern hemisphere a person who puts their right hand on the sphere, palm down, with their thumb on the y axis will have their fingers pointing along the x-axis in the positive direction.



The sequence of pictures below show a trip around this space: the antipodal spots rotate around the sphere with the spot in the northern hemisphere moving up and over the north pole, ending on the far side of the sphere, and its companion coming to the front. (Be careful: the trip shown is not halfway around the projective plane. It is a full circuit. The spots at either end of the dashed line are the same projective point. So by the third sphere below the trip has pretty much returned to the same projective point where we drew it starting above.)



At the end of the circuit, the x part of the xy-axes sticks out in the other direction. That is, for a person to put their thumb on the y-axis and have their fingers point positively on the x-axis, they must use their left hand. The projective plane is not orientable — in this geometry, left and right handedness are not fixed properties of figures (said another way, we cannot describe a spiral as clockwise or counterclockwise).

This exhibition of the existence of a non-orientable space raises the question of whether our universe orientable. Could an astronaut leave earth right-handed and return left-handed? [Gardner] is a nontechnical reference. [Clarke] is a classic science fiction story about orientation reversal.

For an overview of projective geometry see [Courant & Robbins]. The approach we've taken here, the analytic approach, leads to quick theorems and illustrates the power of linear algebra; see [Hanes], [Ryan], and [Eggar]. But another approach, the synthetic approach of deriving the results from an axiom system, is both extraordinarily beautiful and is also the historical route of development. Two fine sources for this approach are [Coxeter] or [Seidenberg]. An easy and interesting application is in [Davies].

Exercises

1 What is the equation of this point?

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

2 **(a)** Find the line incident on these points in the projective plane.

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$$

(b) Find the point incident on both of these projective lines.

$$(1 \ \ 2 \ \ 3), (4 \ \ 5 \ \ 6)$$

3 Find the formula for the line incident on two projective points. Find the formula for the point incident on two projective lines.

4 Prove that the definition of incidence is independent of the choice of the representatives of p and L. That is, if p_1, p_2, p_3, and q_1, q_2, q_3 are two triples of homogeneous coordinates for p, and L_1, L_2, L_3, and M_1, M_2, M_3 are two triples of homogeneous coordinates for L, prove that $p_1L_1 + p_2L_2 + p_3L_3 = 0$ if and only if $q_1M_1 + q_2M_2 + q_3M_3 = 0$.

5 Give a drawing to show that central projection does not preserve circles, that a circle may project to an ellipse. Can a (non-circular) ellipse project to a circle?

6 Give the formula for the correspondence between the non-equatorial part of the antipodal modal of the projective plane, and the plane $z = 1$.

7 (Pappus's Theorem) Assume that T_0, U_0, and V_0 are collinear and that T_1, U_1, and V_1 are collinear. Consider these three points: (i) the intersection V_2 of the lines T_0U_1 and T_1U_0, (ii) the intersection U_2 of the lines T_0V_1 and T_1V_0, and (iii) the intersection T_2 of U_0V_1 and U_1V_0.

(a) Draw a (Euclidean) picture.

(b) Apply the lemma used in Desargue's Theorem to get simple homogeneous coordinate vectors for the T's and V_0.

(c) Find the resulting homogeneous coordinate vectors for U's (these must each involve a parameter as, e.g., U_0 could be anywhere on the T_0V_0 line).

(d) Find the resulting homogeneous coordinate vectors for V_1. (*Hint:* it involves two parameters.)

(e) Find the resulting homogeneous coordinate vectors for V_2. (It also involves two parameters.)

(f) Show that the product of the three parameters is 1.

(g) Verify that V_2 is on the T_2U_2 line.