Math S-21b – Lecture #11 Notes

Today's lecture is all about **determinants**. We'll discuss how to define them, how to calculate them, learn the all-important property known as **multilinearity**, and show that a square matrix **A** is invertible if and only if its determinant is nonzero. We'll also derive some useful geometric applications that will allow us to not only calculate length, area, and volume, but also to define geometric content (*k*-volume) in higher dimensions. We will also give an interpretation of the determinant as an "expansion factor" for geometric content. We'll wrap it up with a few minor results (Cramer's Rule and a not-too-practical formula for the inverse of a matrix).

Defining the determinant

You are probably already familiar with the determinant in the case of 2×2 and perhaps 3×3 matrices. Let's start with those and "reverse engineer" the general definition for any square matrix.

$\boxed{1\times 1 \text{ matrix}}$: Just for the sake of consistency, let's define $\det[a]=a$ for a 1×1 matrix.

$\boxed{2\times 2 \text{ matrix}}$: We define $\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad-bc$.

$\boxed{3\times 3 \text{ matrix}}$: We define

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

$$= a_{11}(a_{22}a_{33}-a_{23}a_{32}) - a_{12}(a_{21}a_{33}-a_{23}a_{31}) + a_{13}(a_{21}a_{32}-a_{22}a_{31})$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

This definition is based on a fact that we have not yet established called the ***Laplace expansion***, but let's take this as given and see what, if any, pattern it suggests. Note that there is just 1 term for the determinant of a 1×1 matrix, 2 terms for a 2×2 matrix (one positive, one negative), and $3!=6$ terms for a 3×3 matrix (half of them positive and half negative). Also note that the number of factors in each term grows with the size of the matrix. A more subtle observation is that, at least as written for the 3×3 case, all terms are of the form $a_{1x}a_{2y}a_{3z}$ and the choices of x, y, z correspond precisely with the different ways of permuting the characters in 123, i.e. $\{123, 132, 213, 231, 312, 321\}$. Finally, note that the sign of each term corresponds to whether this is an even permutation (positive if obtained by an even number of transpositions of the characters starting with 123) or an odd permutation (negative if obtained by an even number of transpositions).

Based on these observations, we might (correctly) speculate that for an $n\times n$ matrix we should define the determinant as follows:

Definition: Given an $n\times n$ matrix $\mathbf{A}=\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$, we define

$\det \mathbf{A} = \sum_{\sigma\in P(n)} \operatorname{sgn}(\sigma) a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}$ where $P(n)$ denotes the set of all permutations of the characters $12\cdots n$; σ denoting an individual permutation; $\sigma(i)$ denoting where the character i is mapped under that permutation; and $\operatorname{sgn}(\sigma)=+1$ if σ is an even permutation and $\operatorname{sgn}(\sigma)=-1$ if σ is an odd permutation. There will be $n!$ terms in the sum – corresponding to the number of permutations in $P(n)$.

There are other ways to define the determinant, but this is a practical definition at least in the case of relatively small matrices.

Two simple observations

1) If **A** is either **upper triangular** or **lower triangular**, all but one of the terms in the determinant will vanish and the determinant will be simply the **product of its diagonal entries**.

2) For any $n \times n$ matrix, $\boxed{\det \mathbf{A}^{\mathrm{T}} = \det \mathbf{A}}$. [The sum is the same, just rearranged and with the same signs.]

Multilinearity

Note that the determinant is, in fact, a function $\det : \mathbf{R}^{n \times n} \to \mathbf{R}$ that takes any $n \times n$ matrix **A** and yields the real number $\det \mathbf{A}$. As a function from one linear space to another, <u>the determinant is not linear</u>. For example, if we were to scale a 2×2 matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (with $\det \mathbf{A} = ad - bc$), we have $t\mathbf{A} = \begin{bmatrix} at & bt \\ ct & dt \end{bmatrix}$ and $\det(t\mathbf{A}) = t^2 ad - t^2 bc = t^2(ad - bc) = t^2 \det \mathbf{A}$.

More generally, for any $n \times n$ matrix, we have $\det(t\mathbf{A}) = t^n \det \mathbf{A}$.

However, **the determinant is linear in any single row or column**. This is known as **multilinearity**.

$\boxed{2 \times 2 \text{ example}}$: $\det \begin{bmatrix} 3 & x_1 \\ 2 & x_2 \end{bmatrix} = 3x_2 - 2x_1 = \begin{bmatrix} -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

3×3 example:

$$\det\begin{bmatrix} 1 & 2 & 1 \\ 3 & 4 & -2 \\ x_1 & x_2 & x_3 \end{bmatrix} = 1(4x_3+2x_2) - 2(3x_3+2x_1) + 1(3x_2-4x_1) = -8x_1+5x_2-2x_3$$

$$= \begin{bmatrix} -8 & 5 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

The multilinearity property gives several immediate corollaries.

In terms of the *k*th column of a matrix:

$$\det\begin{bmatrix} \uparrow & & \uparrow & & \uparrow \\ \mathbf{v}_1 & \ldots & \mathbf{x}+\mathbf{y} & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow & & \downarrow \end{bmatrix} = \det\begin{bmatrix} \uparrow & & \uparrow & & \uparrow \\ \mathbf{v}_1 & \ldots & \mathbf{x} & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow & & \downarrow \end{bmatrix} + \det\begin{bmatrix} \uparrow & & \uparrow & & \uparrow \\ \mathbf{v}_1 & \ldots & \mathbf{y} & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow & & \downarrow \end{bmatrix}$$

$$\text{and } \det\begin{bmatrix} \uparrow & & \uparrow & & \uparrow \\ \mathbf{v}_1 & \ldots & r\mathbf{x} & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow & & \downarrow \end{bmatrix} = r\det\begin{bmatrix} \uparrow & & \uparrow & & \uparrow \\ \mathbf{v}_1 & \ldots & \mathbf{x} & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow & & \downarrow \end{bmatrix}$$

In terms of the *k*th row of a matrix:

$$\det\begin{bmatrix} \leftarrow & \mathbf{v}_1 & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{x}+\mathbf{y} & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_n & \rightarrow \end{bmatrix} = \det\begin{bmatrix} \leftarrow & \mathbf{v}_1 & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{x} & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_n & \rightarrow \end{bmatrix} + \det\begin{bmatrix} \leftarrow & \mathbf{v}_1 & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{y} & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_n & \rightarrow \end{bmatrix}$$

and $\det\begin{bmatrix} \leftarrow & \mathbf{v}_1 & \rightarrow \\ & \vdots & \\ \leftarrow & r\mathbf{x} & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_n & \rightarrow \end{bmatrix} = r\det\begin{bmatrix} \leftarrow & \mathbf{v}_1 & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{x} & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_n & \rightarrow \end{bmatrix}$.

This actually explains the **Laplace expansion**. Choose any row or column of the $n\times n$ matrix $\mathbf{A}$ and for each entry a_{ij} of that row or column, let $\mathbf{A}_{ij}$ be its minor – the $(n-1)\times(n-1)$ matrix obtained by deleting the *i*th row and *j*th column of the matrix **A**.

Then, in terms of the *i*th row, $\boxed{\det\mathbf{A} = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det\mathbf{A}_{ij}}$; and in terms of the *j*th column, $\boxed{\det\mathbf{A} = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det\mathbf{A}_{ij}}$.

For example, in terms of the 1st row of a matrix $\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$, we can express

$\begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix} = a_{11}\begin{bmatrix} 1 & \cdots & 0 \end{bmatrix} + \cdots + a_{1n}\begin{bmatrix} 0 & \cdots & 1 \end{bmatrix}$. Applying linearity in the 1st row, this gives;

$\det\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = a_{11}\det\begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} + \cdots + a_{1n}\det\begin{bmatrix} 0 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$ and because of all the 0's in the first row of each, and some observations about even vs. odd permutations to determine the signs, this becomes:

$$\det \mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = +a_{11} \det \begin{bmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{bmatrix} - \cdots + (-1)^{1+n} a_{1n} \det \begin{bmatrix} a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

$$= +a_{11} \det \mathbf{A}_{11} - \cdots + (-1)^{1+n} a_{1n} \det \mathbf{A}_{1n}$$

The same idea applies to any choice of row or column with appropriate signs.

Example: If $\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 3 \end{bmatrix}$, we can choose to expand along any row or column. We often choose a row with one or more 0's in order to minimize the number of nonzero terms in the sum, but not necessarily.

Expanding along the <u>1st row</u> gives $\det \mathbf{A} = 3\begin{vmatrix} 1 & 0 \\ 1 & 3 \end{vmatrix} - 2\begin{vmatrix} 1 & 0 \\ -1 & 3 \end{vmatrix} + 1\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = 3(3) - 2(3) + 1(2) = 5$.

Expanding along the <u>2nd row</u> gives

$$\det \mathbf{A} = -1\begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} + 1\begin{vmatrix} 3 & 1 \\ -1 & 3 \end{vmatrix} - 0\begin{vmatrix} 3 & 2 \\ -1 & 1 \end{vmatrix} = -1(5) + 1(10) - 0(\cdots) = 5.$$

Expanding along the <u>3rd column</u> gives

$$\det \mathbf{A} = 1\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} - 0\begin{vmatrix} 3 & 2 \\ -1 & 1 \end{vmatrix} + 3\begin{vmatrix} 3 & 2 \\ 1 & 1 \end{vmatrix} = 1(2) - 0(\cdots) + 3(1) = 5.$$

Effect of elementary row operations on the determinant

For any $n \times n$ matrix $\mathbf{A}$, we have the following properties:

(a) $\mathbf{A} \xrightarrow[\text{by } k]{\text{scale row}} \mathbf{B} \quad \Rightarrow \quad \det \mathbf{B} = k \det \mathbf{A}$

$\mathbf{A} \xrightarrow[\text{by } 1/k,\ k \neq 0]{\text{scale row}} \mathbf{B} \quad \Rightarrow \quad \det \mathbf{B} = \frac{1}{k} \det \mathbf{A}$

(b) $\mathbf{A} \xrightarrow[\text{two rows}]{\text{interchange}} \mathbf{B} \quad \Rightarrow \quad \det \mathbf{B} = -\det \mathbf{A}$

(c) $\mathbf{A} \xrightarrow[\text{one row to another}]{\text{add a multiple of}} \mathbf{B} \quad \Rightarrow \quad \det \mathbf{B} = \det \mathbf{A}$

Property (a) follows directly from linearity in any one row. Property (b) follows by observing that all the terms in the determinant will be the same except that even permutations will become odd and vice-versa. This causes all the signs to be reversed. Property (b) also implies that if a matrix has two identical rows, then its determinant must be zero. Property (c) requires a small argument for justification:

$$\det \begin{bmatrix} & \vdots & \\ \leftarrow & \mathbf{v}_i & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_j + k\mathbf{v}_i & \rightarrow \\ & \vdots & \end{bmatrix} = \det \begin{bmatrix} & \vdots & \\ \leftarrow & \mathbf{v}_i & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_j & \rightarrow \\ & \vdots & \end{bmatrix} + k \det \begin{bmatrix} & \vdots & \\ \leftarrow & \mathbf{v}_i & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_i & \rightarrow \\ & \vdots & \end{bmatrix}$$

$$= \det \begin{bmatrix} & \vdots & \\ \leftarrow & \mathbf{v}_i & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_j & \rightarrow \\ & \vdots & \end{bmatrix} + k \cdot 0 = \det \begin{bmatrix} & \vdots & \\ \leftarrow & \mathbf{v}_i & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_j & \rightarrow \\ & \vdots & \end{bmatrix}$$

There are at least two significant results that flow from these observations. The first has to do with simplification of the calculation of a determinant by first doing some row reduction. The second will give a new criterion for invertibility of a matrix.

We can calculate the determinant of a matrix by "double tracking" the steps in row reduction and the effect of each step on the value of the determinant. This is especially useful for larger matrices.

Example: Calculate $\det \mathbf{A}$ for the matrix $\mathbf{A} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 1 & 4 \\ 0 & 4 & 5 \end{bmatrix}$.

Solution:

$$\underset{\det \mathbf{A}}{\begin{bmatrix} 2 & 3 & 1 \\ 1 & 1 & 4 \\ 0 & 4 & 5 \end{bmatrix}} \to \underset{-\det \mathbf{A}}{\begin{bmatrix} 1 & 1 & 4 \\ 2 & 3 & 1 \\ 0 & 4 & 5 \end{bmatrix}} \to \underset{-\det \mathbf{A}}{\begin{bmatrix} 1 & 1 & 4 \\ 0 & 1 & -7 \\ 0 & 4 & 5 \end{bmatrix}} \to \underset{-\det \mathbf{A}}{\begin{bmatrix} 1 & 0 & 11 \\ 0 & 1 & -7 \\ 0 & 0 & 33 \end{bmatrix}} \to \underset{-\frac{1}{33}\det \mathbf{A}}{\begin{bmatrix} 1 & 0 & 11 \\ 0 & 1 & -7 \\ 0 & 0 & 1 \end{bmatrix}} \to \underset{-\frac{1}{33}\det \mathbf{A}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}$$

We could conclude from the 4th entry when we obtained an upper triangular matrix that $-\det \mathbf{A} = 33$, so $\det \mathbf{A} = -33$. We could also have completed the row reduction to get to reduced row-echelon form. This would give that $-\frac{1}{33}\det \mathbf{A} = 1$, so $\det \mathbf{A} = -33$.

Invertibility and the determinant

Suppose we began with a matrix **A** and carried out a sequence of steps to obtain $\mathrm{rref}(\mathbf{A})$. This sequence of steps would involve s row swaps which would affect the determinant by multiplying by $(-1)^s$, r row scalings by factors $\frac{1}{k_1}, \frac{1}{k_2}, \cdots, \frac{1}{k_r}$ (where $k_1, k_2, \cdots, k_r \neq 0$), and some number of steps where a multiple of a pivot row is added to another row. The effect of these row operations on the determinant then gives that $\det[\mathrm{rref}(\mathbf{A})] = (-1)^s \frac{1}{k_1}\frac{1}{k_2}\cdots\frac{1}{k_r}\det(\mathbf{A})$.

From this we conclude that $\boxed{\det(\mathbf{A}) = (-1)^s k_1 k_2 \cdots k_r \det[\text{rref}(\mathbf{A})]}$.

There are only two possible values for $\det[\text{rref}(\mathbf{A})]$. If the matrix **A** is invertible with rank *n*, then $\text{rref}(\mathbf{A}) = \mathbf{I}_n$ and $\det[\text{rref}(\mathbf{A})] = 1$. If the matrix **A** is not invertible with rank $k < n$, then $\text{rref}(\mathbf{A})$ will have at least one all-zero row and $\det[\text{rref}(\mathbf{A})] = 0$. From the result above, this gives the following important theorem:

> **Theorem**: An $n \times n$ matrix **A** is invertible if and only if $\det \mathbf{A} \neq 0$.

There are a number of other facts about determinants of both practical and theoretical value.

Proposition: If **A** and **B** are $n \times n$ matrices, then $\boxed{\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})}$.

Proof: If the matrix **A** is not invertible, then **AB** will also not be invertible and $\det \mathbf{A} = 0$ and $\det(\mathbf{AB}) = 0$, so the result holds in this case. <u>A homework exercise</u> shows that in the case where **A** is invertible and **B** is an arbitrary $n \times n$ matrix, then $\text{rref}[\mathbf{A} \,|\, \mathbf{AB}] = [\mathbf{I}_n \,|\, \mathbf{B}]$. If the row reduction from **A** to $\mathbf{I}_n$ involves the same row operations as outlined previously, then these same row operations would be applied in reducing **AB** to **B**, so $\det(\mathbf{AB}) = (-1)^s k_1 k_2 \cdots k_r \det(\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B})$.

Proposition: If **A** is invertible, then $\boxed{\det(\mathbf{A}^{-1}) = \dfrac{1}{\det(\mathbf{A})}}$.

Proof: If **A** is invertible, then $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$, so $\det(\mathbf{A}^{-1}\mathbf{A}) = \det(\mathbf{A}^{-1})\det(\mathbf{A}) = \det(\mathbf{I}_n) = 1$, so $\det(\mathbf{A}^{-1})$ and $\det(\mathbf{A})$ are reciprocals.

Proposition: If two $n \times n$ matrices **A** and **B** are <u>similar</u>, then $\det \mathbf{A} = \det \mathbf{B}$.

Proof: Two $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ are similar if an only if $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$ for some invertible (change of basis) matrix $\mathbf{S}$. Therefore $\det \mathbf{B} = \det(\mathbf{S}^{-1}\mathbf{AS}) = \det(\mathbf{S}^{-1})\det(\mathbf{A})\det(\mathbf{S}) = \det(\mathbf{A})$.

This last proposition yields an important corollary:

Corollary: Suppose V is a finite-dimensional vector space and $T : V \to V$ is a linear transformation. Then the determinant $\det(T)$ is well-defined. That is, if $\mathcal{B}$ is any basis for V and if $\mathbf{A} = [T]_{\mathcal{B}}$ is the matrix of T relative to this basis, and if we define $\det(T) = \det(\mathbf{A})$, then this value will be the same no matter what basis we choose.

Proof: If we choose any other basis then the matrix of T relative to this other basis will be $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$ for some invertible (change of basis) matrix $\mathbf{S}$. Therefore $\det(T) = \det \mathbf{A} = \det \mathbf{B}$ from the previous proposition.

Geometry and the determinant

If we merge some of the previous information about Gram-Schmidt orthogonalization and QR factorization with the current facts about determinants, we can derive some important and useful results. Recall that if $\{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$ are linearly independent and if we write

$$\mathbf{A} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_k \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

, then the Gram-Schmidt process gave:

$$\mathbf{A} = \underbrace{\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_k \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}}_{\substack{n\times k \text{ matrix w/linearly} \\ \text{independent columns}}} = \underbrace{\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}}_{\substack{n\times k \text{ matrix} \\ \text{w/orthonormal columns}}} \underbrace{\begin{bmatrix} r_{11} & \mathbf{v}_2 \cdot \mathbf{u}_1 & \cdots & \mathbf{v}_k \cdot \mathbf{u}_1 \\ 0 & r_{22} & \cdots & \mathbf{v}_k \cdot \mathbf{u}_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{kk} \end{bmatrix}}_{\substack{k\times k \text{ upper triangular matrix} \\ \text{with nonzero diagonal entries}}} = \mathbf{QR}.$$

The columns of the matrix **A** are the original vectors; the columns of the matrix **Q** are those of the Gram-Schmidt basis; and the entries of the matrix **R** capture all of the geometric aspects of the original basis, i.e. lengths, areas, etc. and the non-orthogonality of the original vectors. The <u>*k*-volume</u> of the parallelepiped determined by $\{\mathbf{v}_1,\cdots,\mathbf{v}_k\}$ is just the product of the diagonal entries of **R**, i.e. $r_{11}r_{22}\cdots r_{kk}=\det\mathbf{R}$.

Note that with $\mathbf{A}=\mathbf{QR}$ we have $\mathbf{A}^{\mathrm{T}}\mathbf{A}=(\mathbf{QR})^{\mathrm{T}}\mathbf{QR}=\mathbf{R}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{QR}=\mathbf{R}^{\mathrm{T}}\mathbf{I}_k\mathbf{R}=\mathbf{R}^{\mathrm{T}}\mathbf{R}$. Therefore $\det(\mathbf{A}^{\mathrm{T}}\mathbf{A})=\det(\mathbf{R}^{\mathrm{T}}\mathbf{R})=\det(\mathbf{R}^{\mathrm{T}})\det(\mathbf{R})=\det(\mathbf{R})\det(\mathbf{R})=(\det\mathbf{R})^2=(k\text{-volume})^2$, so $\boxed{k\text{-volume}=\sqrt{\det(\mathbf{A}^{\mathrm{T}}\mathbf{A})}}$. This is a very handy way to calculate areas, volumes, and their higher-dimensional analogues.

Example: In $\mathbf{R}^3$, find the area of the parallelogram determined by the vectors

$$\mathbf{v}_1=\begin{bmatrix}1\\2\\3\end{bmatrix} \text{ and } \mathbf{v}_2=\begin{bmatrix}-1\\0\\2\end{bmatrix}.$$

Solution: In <u>multivariable calculus</u>, we would likely find the area of this parallelogram using the cross product. We would calculate that $\mathbf{v}_1\times\mathbf{v}_2=\begin{bmatrix}4\\-5\\2\end{bmatrix}$ and find its magnitude:

$\text{Area}=\|\mathbf{v}_1\times\mathbf{v}_2\|=\sqrt{16+25+4}=\sqrt{45}=3\sqrt{5}$. Using our <u>determinant method</u>, we write

$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 2 \end{bmatrix}$ and calculate $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 14 & 5 \\ 5 & 5 \end{bmatrix}$. So

$\det(\mathbf{A}^{\mathrm{T}}\mathbf{A}) = \det \begin{bmatrix} 14 & 5 \\ 5 & 5 \end{bmatrix} = 70 - 25 = 45$ and Area $=$ 2-volume $= \sqrt{\det(\mathbf{A}^{\mathrm{T}}\mathbf{A})} = \sqrt{45} = 3\sqrt{5}$.

It is important to note that the cross product is only defined in $\mathbf{R}^3$, so any method involving cross products has very limited applicability.

Special Case: Determinant of an *n*×*n* matrix as an expansion factor

If $\mathbf{A} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow \end{bmatrix}$ is an $n \times n$ matrix, then

$\det(\mathbf{A}^{\mathrm{T}}\mathbf{A}) = \det(\mathbf{A}^{\mathrm{T}})\det(\mathbf{A}) = (\det \mathbf{A})^2 = (n\text{-volume})^2$ and the *n*-volume determined by the vectors $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is given by $\sqrt{\det(\mathbf{A}^{\mathrm{T}}\mathbf{A})} = \sqrt{(\det \mathbf{A})^2} = |\det \mathbf{A}|$. If we further note that

$\begin{Bmatrix} \mathbf{v}_1 = \mathbf{A}\mathbf{e}_1 \\ \vdots \\ \mathbf{v}_n = \mathbf{A}\mathbf{e}_n \end{Bmatrix}$, we can observe that the unit *n*-cube determined by $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ is mapped to the parallelepiped determined by $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$, so the volume is expanded from 1 to $|\det \mathbf{A}|$. This result extends to any region in the domain and enables us to think of $|\det \mathbf{A}|$ as a "volume expansion factor". This provides a simple geometric interpretation of the fact that $\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$ (and therefore $|\det(\mathbf{AB})| = |\det \mathbf{A}||\det \mathbf{B}|$).

Since the product of two matrices corresponds to the composition of linear transformations, and if applying the matrix **B** scales volume by $|\det \mathbf{B}|$, and this is followed by applying the matrix **A** which scales volume by $|\det \mathbf{A}|$, then the composition should scale volume by the product $|\det \mathbf{A}||\det \mathbf{B}|$.

It's not hard to reason that the <u>sign of the determinant</u> will be positive if the linear transformation is "orientation preserving" and negative if the transformation is "orientation reversing." Indeed, we can define these terms by the sign of the determinant.

In the special case when a system of *n* linear equations in *n* variables has a unique solution, determinants provide a formula for this unique solution. This is known as Cramer's Rule.

Cramer's Rule: Suppose a linear system is represented as $\mathbf{Ax} = \mathbf{b}$ where **A** is an $n \times n$ matrix with rank *n*. Let $\mathbf{A}_k$ be the $n \times n$ matrix obtained by replacing the *k*th column of **A** with the column vector **b**. If the solution to the system is $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, then $x_k = \dfrac{\det \mathbf{A}_k}{\det \mathbf{A}}$ for all *k*.

Proof: Suppose **x** solves $\mathbf{Ax} = \mathbf{b}$, then

$$\det \mathbf{A}_k = \det \begin{bmatrix} \uparrow & & \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{b} & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow & & \downarrow \end{bmatrix} = \det \begin{bmatrix} \uparrow & & \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{Ax} & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow & & \downarrow \end{bmatrix}$$

$$= \det \begin{bmatrix} \uparrow & & \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & (x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n) & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow & & \downarrow \end{bmatrix} = x_k \det \begin{bmatrix} \uparrow & & \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_k & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow & & \downarrow \end{bmatrix} = x_k \det \mathbf{A}$$

where we have liberally applied several previous results. So $x_k = \dfrac{\det \mathbf{A}_k}{\det \mathbf{A}}$.

Example: Solve the linear system $\begin{Bmatrix} 2x + y - 3z = 1 \\ -3x + 4y + 5z = 3 \\ x - y + 6z = -4 \end{Bmatrix}$ using Cramer's Rule.

Solution: We have $\mathbf{A} = \begin{bmatrix} 2 & 1 & -3 \\ -3 & 4 & 5 \\ 1 & -1 & 6 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$. We first calculate

$\det \mathbf{A} = 2(29) - 1(-23) - 3(-1) = 58 + 23 + 3 = 84 \neq 0$, so the system will yield a unique solution.

We next write $\mathbf{A}_1 = \begin{bmatrix} 1 & 1 & -3 \\ 3 & 4 & 5 \\ -4 & -1 & 6 \end{bmatrix}$, $\mathbf{A}_2 = \begin{bmatrix} 2 & 1 & -3 \\ -3 & 3 & 5 \\ 1 & -4 & 6 \end{bmatrix}$, and $\mathbf{A}_3 = \begin{bmatrix} 2 & 1 & 1 \\ -3 & 4 & 3 \\ 1 & -1 & -4 \end{bmatrix}$ and calculate

$\det \mathbf{A}_1 = 1(29) - 1(38) - 3(13) = 29 - 38 - 39 = -48$ and

$\det \mathbf{A}_2 = 2(38) - 1(-23) - 3(9) = 76 + 23 - 27 = 72$ and

$\det \mathbf{A}_3 = 2(-13) - 1(9) + 1(-1) = -26 - 9 - 1 = -36$. So $x_1 = \frac{-48}{84} = -\frac{4}{7}$, $x_2 = \frac{72}{84} = \frac{6}{7}$, and

$x_3 = \frac{-36}{84} = -\frac{3}{7}$.

Cookbook recipe for finding the inverse of an invertible matrix

If you look carefully at Cramer's Rule, you may notice that it actually provides a formula for the inverse of any invertible matrix. The fact that $\det \mathbf{A}$ should appear in the denominators is clear enough , and we omit most of the remaining details, but with a little effort we can arrive at the following (not particularly useful) result:

Recipe for $\mathbf{A}^{-1}$: Given an $n\times n$ matrix, we first calculate $\det\mathbf{A}$. If $\det\mathbf{A}=0$, stop – the matrix is not invertible. If $\det\mathbf{A}\neq 0$, we continue. For each entry a_{ij} of the matrix, let $\mathbf{A}_{ij}$ be its minor – the $(n-1)\times(n-1)$ matrix obtained by deleting the *i*th row and *j*th column of the matrix **A**. We define the **cofactors** by $\boxed{\operatorname{cof}(a_{ij})=(-1)^{i+j}\det\mathbf{A}_{ij}}$. If we assemble all of these cofactors into a matrix, we call this $\operatorname{cof}(\mathbf{A})$. We then transpose this matrix to get the adjoint matrix $\boxed{\operatorname{adj}(\mathbf{A})=[\operatorname{cof}(\mathbf{A})]^{\mathrm{T}}}$. Then $\boxed{\mathbf{A}^{-1}=\frac{1}{\det(\mathbf{A})}\operatorname{adj}(\mathbf{A})}$.

A simple procedure for carrying this out is to:

(a) Calculate the determinant of the given matrix. If it's nonzero, continue.

(b) Calculate the matrix consisting of the determinants of the respective minors for every entry of the given matrix.

(c) Adjust all the signs using the checkerboard pattern: $\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$.

(d) Transpose the resulting matrix to get the adjoint.

(e) Multiply by the reciprocal of the determinant to get the inverse matrix.

Example: Find the inverse of the matrix $\mathbf{A} = \begin{bmatrix} 2 & 1 & -3 \\ -3 & 4 & 5 \\ 1 & -1 & 6 \end{bmatrix}$.

Solution: (a) $\det \mathbf{A} = 2(29) - 1(-23) - 3(-1) = 58 + 23 + 3 = 84 \neq 0$.

(b) The determinant of the minors gives: $\begin{bmatrix} 29 & -23 & -1 \\ 3 & 15 & -3 \\ 17 & 1 & 11 \end{bmatrix}$

(c) Adjust the signs to get the matrix of cofactors: $\begin{bmatrix} 29 & 23 & -1 \\ -3 & 15 & 3 \\ 17 & -1 & 11 \end{bmatrix}$

(d) Transpose to get the adjoint: $\begin{bmatrix} 29 & -3 & 17 \\ 23 & 15 & -1 \\ -1 & 3 & 11 \end{bmatrix}$

(e) Multiply by the reciprocal of the determinant to get $\mathbf{A}^{-1} = \frac{1}{84}\begin{bmatrix} 29 & -3 & 17 \\ 23 & 15 & -1 \\ -1 & 3 & 11 \end{bmatrix}$.

Had we proceeded this way, we would have solved the system in the previous example as

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \frac{1}{84}\begin{bmatrix} 29 & -3 & 17 \\ 23 & 15 & -1 \\ -1 & 3 & 11 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} -48 \\ 72 \\ -36 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} -4 \\ 6 \\ -3 \end{bmatrix}.$$

Note: The impracticality of this method starts to become clear when we look at 4×4 matrices which would involve the calculation 16 determinants of 3×3 matrices in addition to the original 4×4 determinant which requires the calculation of other 3×3 determinants to bring the total to 20 such determinants (in addition to the other calculations).

For a 5×5 matrix, we would have to calculate $25+5=30$ determinants of 4×4 matrices each of which would require the calculation of smaller determinants. In general, it is far quicker to solve using row reduction methods, and row reduction has the additional advantage of yielding solutions in the case of consistent systems with rank less than *n*.

Notes by Robert Winters