DETERMINANT

1. 2×2 Determinants

Fix a matrix $A \in \mathbb{R}^{2\times 2}$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Define the *determinant* of A, to be

$$\det(A) = ad - bc.$$

EXAMPLE: $\det \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} = 2(2) - (-1)(-4) = 0.$

$\det(A)$ encodes important geometric and algebraic information about A. For instance, the matrix A is invertible if and only if $\det(A) \neq 0$. To see this suppose $\det(A) \neq 0$, in this case the matrix

$$B = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

is well-defined. One computes,

$$BA = \frac{1}{ad - bc} \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = I_2.$$

and so A is invertible and $A^{-1} = B$. In the other direction, if A is not invertible, then the columns are linearly dependent and so either $a = kb$ and $c = kd$ or $b = ka$ and $d = kc$ for some $k \in \mathbb{R}$. Hence, $\det(A) = kbd - kbd = 0$ or $= kac - kac = 0$. More geometrically, if

$$S = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : 0 \leq x_1, x_2 \leq 1 \right\}$$

is the unit square and

$$A(S) = \{A\vec{x} : \vec{x} \in S\}$$

is the image under multiplication by A of S, then the area of $A(S)$, which we denote by $|A(S)|$ is

$$|A(S)| = |\det(A)|.$$

2. Determinants of $n \times n$ matrices

We seek a good notion for $\det(A)$ when $A \in \mathbb{R}^{n\times n}$. It is improtant to remember that while this is of historical and theoretical importance, it is not as significant for computational purposes. Begin by fixing a matrix $A \in \mathbb{R}^{n\times n}$ of the form

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

We define a *pattern*, P, of A to be a choice of n entries of A with the property that no two choices lie in the same row or column of A. This means there are

$n! = n(n-1)(n-1) * \cdots * 2 * 1$ different patterns. For example the *diagonal pattern* is $P = (a_{11}, a_{22}, \ldots, a_{nn})$.

Given a pattern, P, define $\text{prod}(P)$ to be the product of the entries picked by P.

EXAMPLE: If D is the diagonal pattern of A, then $\text{prod}(D) = a_{11}a_{22}\cdots a_{nn}$. Moreover, if $A = I_n$, then $prod(P) = 0$ for all patterns except the diagonal pattern (for which $prod(P) = 1$).

Two entries of a pattern are said to be *inverted* if one of them is located to the right and above the other in the matrix. As such, no entries in the diagonal pattern are inverted. Let $\#P$ be the number of inversions in P (i.e. number of pairs of inverted elements) and define the *signature* of P, to be

$$\text{sgn}P = (-1)^{\#P}.$$

We define the *determinant* of A to be the sum over all patterns of A of $\text{sgn}(P)\text{prod}(P)$. That is,

$$\det A = \sum_P \text{sgn}(P)\text{prod}(P)$$

The number of terms in the sum grows rapidly, for instance when $n = 4$ this is already 24 terms.

EXAMPLE: $\det I_n = 1$ as the only pattern with $\text{prod}P \neq 0$ has $\text{prod}P = 1$ and no inversions so $\text{sgn}(P) = 1$.

EXAMPLE: In the 2×2 case there are two patterns P_1 and P_2, the diagonal and off diagonal pattern. Have $\text{prod}(P_1) = ad$ $\text{sgn}(P_1) = 1$ while $\text{prod}(P_2) = bc$ and $\text{sgn}(P_2) = -1$. Summing this up recovers our original formula.

EXAMPLE: If

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

then $det(A) = -6$. This is because there is only one pattern, P, with $\text{prod}(P) = -6 \neq 0$ and it has $\#P = 2$ so $\text{sgn}(P) = 1$.

Theorem 2.1. *If A is upper triangular,*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

then $\det(A) = a_{11}a_{22}\cdots a_{nn}$.

Proof. There is exactly one pattern that doesn't contain a zero entry, the diagonal pattern, this pattern has no inversions. □

3. Basic Properties of Determinant

We collect here some basic properties of the determinant. One the one hand, these properties will allow us to justify some of the applications of the determinant, on the other they give efficient computational methods.

Theorem 3.1. *For all $A \in \mathbb{R}^{n\times n}$ we have* $\det(A) = \det(A^\top)$.

Proof. For any pattern P of A we there is a corresponding pattern $P^\top$ of $A^\top$ obtained in the obvious fashion. The numerical entries of P and $P^\top$ are the same and so $\mathrm{prod}(P) = \mathrm{prod}(P^\top)$. With a little more work one verifies that $\#P = \#P^\top$ so $\mathrm{sgn}(P) = \mathrm{sgn}(P^\top)$ and the result follows from the definition of determinant. $\square$

An important consequence of this result is that any property of the determinant that relates to the columns of A has an analogous property relating to the rows of A.

Theorem 3.2. *Fix $\vec{a}_2, \ldots, \vec{a}_n \in \mathbb{R}^n$ the function*

$$T(\vec{x}) = \det \begin{bmatrix} \vec{x} & | & \vec{a}_2 & | & \cdots & | & \vec{a}_n \end{bmatrix}$$

is a linear transform $T : \mathbb{R}^n \to \mathbb{R}$.

Proof. Lets show $T(k\vec{x}) = kT(\vec{x})$ the argument that $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$ is along similar lines. Pick $\vec{x} \in \mathbb{R}^n$ and $k \in \mathbb{R}$. Let

$$A = \begin{bmatrix} k\vec{x} & | & \vec{a}_2 & | & \cdots & | & \vec{a}_n \end{bmatrix} \text{ and } A' = \begin{bmatrix} \vec{x} & | & \vec{a}_2 & | & \cdots & | & \vec{a}_n \end{bmatrix}$$

If P is a pattern for A and P' is the same pattern for A', then we have

$$\mathrm{prod}(P) = k\mathrm{prod}(P') \text{ and } \mathrm{sgn}(P) = \mathrm{sgn}(P')$$

as exactly one entry of P lies in the first column. Hence,

$$\det(A) = \sum_P \mathrm{prod}(P)\mathrm{sgn}(P) = \sum_{P'} k\mathrm{prod}(P')\mathrm{sgn}(P') = k \sum_{P'} \mathrm{prod}(P')\mathrm{sgn}(P') = \det(A').$$

$\square$

Clearly, the same argument holds for any column in place of the first one and also for any row. We next note that swapping two columns changes the value of the determinant by a sign.

Theorem 3.3. *For fixed $\vec{a}_3, \ldots, \vec{a}_n \in \mathbb{R}^n$, the map*

$$B(\vec{x}, \vec{y}) = \det \begin{bmatrix} \vec{x} & | & \vec{y} & | & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}$$

satisfies $B(\vec{x}, \vec{y}) = -B(\vec{y}, \vec{x})$.

The function B is said to be anti-symmetric (a symmetric function, S is one for which $S(\vec{x}, \vec{y}) = S(\vec{y}, \vec{x})$) and and so this is called the anti-symmetry property of the determinant.

Proof. For any pattern P of A, let P' be the pattern of A' obtained by swapping the entries in the first and second column. Clearly, $\mathrm{prod}(P) = \mathrm{prod}(P')$, but $\#P' = \#P \pm 1$ (i.e., either add or destroy an inversion). This implies $\mathrm{sgn}(P) = -\mathrm{sgn}(P')$. Hence,

$$\det(A) = \sum_P \mathrm{prod}(P)\mathrm{sgn}(P) = \sum_{P'} \mathrm{prod}(P')(-\mathrm{sgn}(P')) = -\sum_{P'} \mathrm{prod}(P')\mathrm{sgn}(P') = -\det(A').$$

$\square$

The same proof works for any pair of adjacent columns (or rows). The conclusion is true for any pair of distinct columns (or rows) as one can swap two columns by making an odd number of swaps of adjacent columns. For example, to swap hte first and third, swap the first and second, then hte second and third and finally first and second.

Corollary. *If A has two columns (or two rows) the same, then* $\det(A) = 0$.

Proof. Swapping the two repeated columns yields A back, so $\det(A) = -\det(A) \Rightarrow \det(A) = 0$. $\square$

4. Determinant and Gauss-Jordan Elimination

Recall, the following three elementary row operations one can perform on a matrix A:

(1) (scale) Multiply one row of A by $k \in \mathbb{R}$, $k \neq 0$.
(2) (swap) Swap two different rows of A.
(3) (add) Add a multiple of one row of A to a different row.

Using the properties of the determinant, we obtain the following result describing how elementary row operations affect the determinant.

Theorem 4.1. *Fix a matrix $A \in \mathbb{R}^{n\times n}$, if*

(1) *A' is obtained from A by scaling a row by k, then* $\det(A') = k\det(A)$
(2) *A' is obtained from A by swapping two rows of A, then* $\det(A') = -\det(A)$
(3) *A' is obtained from A by adding a multiple of one row to another, then* $\det(A') = \det(A)$.

Proof. Claim (1) follows immediately from linearity property of the determinant. Claim (2) follows from the anti-symmetry property of the determinant. Claim (3) follows from the linearity property and the fact that repeated identical rows yields zero determinant. $\square$

Corollary. *If B is obtained from A by applying a sequence of elementary row operations, then*

$$\det(B) = (-1)^s k_1 k_2 \cdots k_r \det(A)$$

where s is the number of swaps and $k_1, \ldots, k_r$ scaling factors (so there are r total scalings).

As one obtains $\operatorname{rref}(A)$ by repeated applications of elementary row operations one has

Corollary. *For any $A \in \mathbb{R}^{n\times n}$,* $\det(A) = (-1)^s k_1^{-1} \cdots k_r^{-1} \det(\operatorname{rref}(A))$.

This last fact allows us to prove one of our main theoretical applications of the determinant.

Theorem 4.2. *A is invertible if and only if* $\det(A) \neq 0$.

Proof. We first note that, by definition, $\operatorname{rref}(A)$ is upper triangular. Moreover, at least one entry on the diagonal of $\operatorname{rref}(A)$ is zero unless $\operatorname{rref}(A) = I_n$. That is either $\det(\operatorname{rref}(A)) = 1$ and $A = I_n$ or $\det(\operatorname{rref}(A)) = 0$ and $A \neq I_n$. Applying the above corollary, $\det(A) = (-1)^s k_1^{-1} \cdots k_r^{-1} \det(\operatorname{rref}(A))$ and so $\det(A) = 0$ if and only if $\det(A) = 0$ that is if and only if $\operatorname{rref}(A) \neq I_n$. As A is invertible only when $\operatorname{rref}(A) = I_n$ this proves the claim. $\square$

EXAMPLE: Compute $\det(A)$ for

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

In this case, applying elementary row operations yields

$$A \to \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 7 & 5 & -3 \\ 1 & 1 & 2 & -1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 7 & 5 & -3 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 7 & 5 & -3 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix} = B$$

Got from A to B with two swaps and no scales and B is upper triangular so

$$\det(A) = (-1)^2 \det(B) = 1 * 7 * (-1) * (-2) = 14.$$

5. Determinant of a Product

An important property of the determinant is that the determinant of a product of two matrices is the product of their determinants.

Theorem 5.1. *Let A and B be $n \times n$ matrices and m a positive integer*

(1) $\det(AB) = \det(A)\det(B)$
(2) $\det(A^m) = (\det(A))^m$
(3) *If A is invertible, then* $\det(A^{-1}) = \frac{1}{\det(A)}$.

Proof. To see claim (1) first observe that if A is not invertible, then $\operatorname{rank}(A) < n$ and $\det(A) = 0$. In this case, $\operatorname{rank}(AB) \leq \operatorname{rank}(A) < n$ and so AB is also not invertible. Hence, $\det(AB) = 0$ verifying the result. If A is invertible, then (this is an exercise in the textbook):

$$\operatorname{rref}[A|AB] = [I_n|B]$$

Hence, as we are applying the same elementary row operations to both A and AB

$$\det(A) = (-1)^s k_1 \cdots k_r$$

and

$$\det(AB) = (-1)k_1 \cdots k_r \det(B) = \det(A)\det(B).$$

This proves the claim. Claim (2) follows easily from Claim (1). Finally, Claim (3) follows by noting that $I_n = A^{-1}A$ and so Claim (1) means that

$$1 = \det(I_n) = \det(A^{-1}A) = \det(A^{-1})\det(A)$$

and the result follows by dividing. □

An immediate consequence of this rule is that similar matrices have the same determinant.

Theorem 5.2. *IF A and B are similar matrices, then* $\det(A) = \det(B)$.

Proof. Remember A and B are similar if and only if $AS = SB$ for some invertible S. Hence,

$$\det(A)\det(S) = \det(AS) = \det(SB) = \det(S)\det(B) = \det(B)\det(S)$$

. As S isinvertible, $\det(S) \neq 0$, and so we divide it out of both sides to obtain $\det(A) = \det(B)$. □

Interestingly, this fact allows one to define a notion of determinant for any linear transformation between linear spaces. To see this, first fix a linear transformation $T : V \to V$ where V is some finite dimensional linear space. If $\mathcal{B}$ and $\mathcal{U}$ are (different) bases of V, then $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{U}}$ are similar matrices (recall, these are the matrices of T with respect to the two bases). Indeed, if $S_{\mathcal{B}\to\mathcal{U}} \in \mathbb{R}^{n\times n}$ is the change of basis matrix, then

$$S_{\mathcal{B}\to\mathcal{U}}[T]_{\mathcal{B}} = [T]_{\mathcal{U}} S_{\mathcal{B}\to\mathcal{U}}$$

As such,

$$\det[T]_{\mathcal{B}} = \det[T]_{\mathcal{U}}$$

For this reason, we may define

$$\det T := \det[T]_{\mathcal{B}}$$

where $\mathcal{B}$ is any choice of basis.

6. Laplace Expansion

The most computationally efficient method of computing the determinant is usually to use Gaussian elimination as discussed above. However, in some cases (especially when one or more entries in the matrix are variables) it is convenient to have a closed form expression. In this case, the method known as the Laplace (or cofactor) expansion is often the most efficient method. This method is particularly suited to matrices with lots of zero entries.

Fix a matrix

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

Define $A_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ to be the matrix obtained by deleting the ith row and jth column of A. EXAMPLE: For $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, $A_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}$. We call the value $D_{ij} = \det A_{ij}$ a *minor* of A. The number $C_{ij} = (-1)^{i+j} D_{ij}$ is a *cofactor* of A.

We can use the cofactors to give a formula for the determinant called the *Laplace* or *Cofactor* expansion

Theorem 6.1. *Expanding along the jth column:*

$$\det A = \sum_{i=1}^{n} a_{ij} C_{ij} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

Expanding along the ith row:

$$\det A = \sum_{j=1}^{n} a_{ij} C_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}.$$

To remember the signs useful to remember the 3×3 case

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

EXAMPLE: By expanding along middle column, we compute:

$$\det \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = -0 \cdot \det \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} + 1 \cdot \det \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} - 0 \cdot \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} = 1.$$

7. Determinant of orthogonal matrices

For any orthogonal matrix $Q \in \mathbb{R}^{n\times n}$, $\det(Q) = \pm 1$. To see this, recall Q orthogonal $\iff I_n = Q^\top Q$. Hence,

$$1 = \det(I_n) = \det(Q^\top Q) = \det(Q^\top)\det(Q) = \det(Q)^2.$$

This means $\det(Q) = \pm 1$ as desired.

EXAMPLE: For rotations

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

we have

$$\det R_\theta = \cos^2\theta + \sin^2\theta = 1$$

Reflections (across line spanned by $(\cos\phi/2)\vec{e}_1 + (\sin\phi/2)\vec{e}_2$) have matrix

$$T_\phi = \begin{bmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{bmatrix}$$

and in this case

$$\det T_\phi = -\cos^2\phi - \sin^2\phi = -1.$$

8. Determinant as Area, Volume, etc.

Fix vectors $\vec{v}_1, \cdots, \vec{v}_m \in \mathbb{R}^n$ for $m \leq n$. The *parallelepiped defined* or *spanned* by $\vec{v}_1, \ldots, \vec{v}_m$ is the set

$$P = P(\vec{v}_1, \ldots, \vec{v}_m) = \{t_1\vec{v}_1 + \ldots + t_n\vec{v}_n : 0 \leq t_1, \ldots, t_n \leq 1\}$$

When $n = m = 2$ this is a parallelogram. When $n = m = 3$ this is a solid. For example, the unit cube is $P(\vec{e}_1, \vec{e}_2, \vec{e}_3)$. Let $|P| = V(\vec{v}_1, \ldots, \vec{v}_m)$ be the m-dimensional volume of P.

EXAMPLE: When $m = 2$ then this is the area of a parallelogram. As such we have the formula

$$|P| = (base)(height) = ||\vec{v}_1||||\vec{v}_2^\perp||.$$

More generally, we inductively have

$$|P| = (base)(height) = |P(\vec{v}_1, \ldots, \vec{v}_{m-1})|||\vec{v}_m^\perp|| = ||\vec{v}_1||||\vec{v}_2^\perp||\cdots||\vec{v}_m^\perp||.$$

Let

$$M = \begin{bmatrix} \vec{v}_1 & | & \cdots & | & \vec{v}_m \end{bmatrix} \in \mathbb{R}^{n\times m}$$

Wen $n = m$, we can use the QR factorization to show

$$|\det(M)| = ||\vec{v}_1||||\vec{v}_2^\perp||\cdots||\vec{v}_n^\perp|| = |P|.$$

We do this for $m = n = 2$ as the general case follows from this.

First observe that we may assume the columns are linearly independent as otherwise the area is obviously zero and $\det(M) = 0$. When the columns are linearly independent we have

$$M = QR = \begin{bmatrix} \vec{u}_1 & | & \vec{u}_2 \end{bmatrix} \begin{bmatrix} ||\vec{v}_1|| & \vec{u}_1 \cdot \vec{v}_2 \\ 0 & ||\vec{v}_2^\perp|| \end{bmatrix}.$$

Hence,

$$\det(M) = \det(QR) = \det(Q)\det(R) = \pm||\vec{v}_1||||\vec{v}_2^\perp||$$

where we used that Q was orthogonal and R was upper triangular

When $m < n$, M is not square so the its determinant doesn't make sense. However, $M^\top M \in \mathbb{R}^{m\times m}$ is square. Moreover, if $M = QR$ is the QR-factorization, then

$$M^\top M = R^\top Q^\top QR = R^\top R$$

Hence, (if the columns of M are linearly independent)

$$\det(M^\top M) = \det(R^\top R) = \det(R)^2 = ||\vec{v}_1||^2||\vec{v}_2^\perp||^2 \cdots ||\vec{v}_m||^2 = |P(\vec{v}_1, \dots, \vec{v}_m)|^2.$$

In other words,

$$|P(\vec{v}_1, \dots, \vec{v}_m)| = \sqrt{\det(M^\top M)}.$$

EXAMPLE: Compute the volume and surface area area of the parallelpiped, P, with sides

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \text{ and } \vec{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

We have

$$\det \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = -\det \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = -1$$

so $|P| = 1$. Let $P_1 = P(\vec{v}_2, \vec{v}_3), P_2 = P(\vec{v}_1, \vec{v}_3), P_3 = P(\vec{v}_1, \vec{v}_2)$. The surface area of P is

$$A(P) = 2|P_1| + 2|P_2| + 2|P_3|$$

Now,

$$M_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ so } M_1^\top M_1 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

Hence,

$$|P_1| = \left(\det \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\right)^{1/2} = 1$$

$$M_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \text{ so } M_2^\top M_2 = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$$

Hence,

$$|P_2| = \left(\det \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}\right)^{1/2} = \sqrt{2}$$

Finally,

$$M_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \text{ so } M^\top M = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$$

Hence,

$$|P_3| = (\det \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix})^{1/2} = \sqrt{2}.$$

Thus,

$$A(P) = 1 + 2\sqrt{2} + 2\sqrt{2} = 1 + 4\sqrt{2}.$$