DETERMINANTS

S.R. DOTY

The *determinant* function det is a function from $n \times n$ matrices to scalars, defined recursively by the rules:

(D1) $\det A = a$ if $A = [a]$ is a 1×1 matrix.
(D2) If $A = (a_{i,j})_{1 \le i,j \le n}$ then $\det A = \sum_j (-1)^{1+j} a_{1,j} \det A_{1,j}$,

where $A_{1,j}$ stands for a certain *submatrix* of A. In general, for any i, j the submatrix $A_{i,j}$ is defined to be the $(n-1) \times (n-1)$ submatrix obtained from A by deleting the ith row and jth column. The formula in (D2) is called a first row *Laplace expansion* of the determinant, and the determinants of the submatrices A_{ij} are known as *subminors*.

A common alternative notation for $\det A$ is $|A|$. It should be noted that determinants only make sense for a square matrix A. If a matrix A is not square, then its determinant $|A|$ is *undefined*.

The above recursive definition is actually not a very good way to define determinants, but the better ways require more abstraction, so we will be content with the above definition.

Example 1. It follows immediately from the definition that the determinant of an arbitrary 2×2 matrix is given by the formula

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

For 2×2 matrices, remembering this formula is probably the best course of action.

Example 2. We now compute the determinant of an arbitrary 3×3 matrix using the recursive definition (i.e., doing a Laplace expansion along the first row):

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

Date: Fall 2013.

which simplifies to

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg)$$

There is a standard trick for easily computing 3×3 determinants without remembering a formula. See Chapter Four, Section I, Exercise 1.11 in Hefferon's book for details.

Once determinants have been defined one can use the definition to inductively prove certain basic properties of determinants of matrices, which are as follows:

(1) It is possible to Laplace expand along *any* row or column. The formula for expanding along the ith row is the same as the above formula for expanding along the first, except that you must replace 1 by i in the above:

$$\det A = \sum_j (-1)^{i+j} a_{i,j} \det A_{i,j}$$

where we sum over j and hold i fixed. To expand on the jth column, the formula is the same, except that j is fixed and we sum over i:

$$\det A = \sum_i (-1)^{i+j} a_{i,j} \det A_{i,j}.$$

(2) If A is an $n \times n$ matrix with real entries then the absolute value $|\det A|$ of the determinant of A gives the volume of the n-dimensional parallelepiped in $\mathbb{R}^n$ generated by the row (or column) vectors of A. This comes up in multivariable calculus, in the theory of multiple integrals. In particular, in case $n = 2$ the magnitude of the determinant gives the area of the parallelogram generated by the row (or column) vectors of the matrix.

(3) $\det A \neq 0$ if and only if A is nonsingular. Equivalently, $\det A = 0$ if and only if A is singular.

(4) $\det A^{\mathsf{T}} = \det A$ where A^{T} is the transpose of A.

(5) $\det A^{-1} = 1/(\det A)$, provided A is nonsingular.

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

(7) (Triangular Property) If A is upper or lower triangular then $\det A$ is the product of the entries on the main diagonal.

(8) (Swapping Property) If two rows (or columns) of A are interchanged, then the determinant changes in sign but not magnitude.

(9) (Scaling Property) If one row (or column) of A is multiplied by a scalar c to form a new matrix B then $\det B = c(\det A)$.

(10) (Pivoting Property) If a scalar multiple of one row is added to another row then the determinant remains unchanged. (The same holds for columns.)

The best way to compute a larger determinant is not to use the recursive definition, but rather to use properties (8)–(10) above to row reduce the given matrix to an echelon form, which is always an upper triangular matrix, keeping track of all scaling operations along the way, and then use property (7) on the resulting echelon form. For a 10 by 10 matrix this reduces the number of multiplications needed from about forty million to a few hundred, which is a reduction of roughly 99.999% of the work involved. You usually *don't* want to compute a determinant by the recursive definition, unless it's 4 by 4 or smaller!

One consequence of the recursive definition is that the determinant $\det A$ depends polynomially on the entries of the matrix A, since only additions, subtractions, and multiplications of the entries are used to compute $\det A$.

Explicit formula for $\det A$

The following explicit formula for the determinant $\det A$ of any $n \times n$ matrix A may be proved by induction on n, starting from the recursive definition:

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}.$$

In the formula, S_n is the *symmetric group*, consisting of all permutations σ of the set $\{1, 2, \dots, n\}$. A permutation is just a bijection of the set onto itself. For example,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 6 & 4 & 1 & 2 & 7 & 5 \end{pmatrix} \in S_7$$

defines the permutation sending $1 \to 3$, $2 \to 6$, $3 \to 4$, $\dots$, $7 \to 5$. The *sign* of a permutation σ may be defined by

$$\operatorname{sgn}(\sigma) = (-1)^N$$

where N counts the number of *descents* (or inversions) in σ, a descent being an $i < j$ for which $\sigma(i) > \sigma(j)$. For instance, in the example $\sigma \in S_7$ given above, $\operatorname{sgn}(\sigma) = 1$ since σ has eight descents (in the second row of σ observe that 3, 6, and 4 are larger than both 1 and

2, which gives six descents, and then 6 and 7 are larger than 5, which gives two more, for a total of eight).

Cramer's rule for A^{-1}

Finally, it should be mentioned that there is also an explicit formula for the inverse of an invertible matrix A, in terms of determinants, known as *Cramer's rule.* The formula is

$$A^{-1} = \frac{1}{\det A}(\widetilde{A})^{\mathsf{T}}$$

where the (i, j) entry of $\widetilde{A}$ equals $(-1)^{i+j} \det A_{i,j}$. Some people call the matrix $(\widetilde{A})^{\mathsf{T}}$ in the above formula the *adjoint* matrix of A. You do not need this formula, since for large matrices the technique of Gaussian row reduction $[A \mid I] \to [I \mid A^{-1}]$ discussed earlier in the course is much simpler. However, in the 2×2 case Cramer's rule does give the useful formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

which is sometimes a convenient time-saver.

Cramer's rule for solving $A\vec{x} = \vec{b}$

Similarly, there is another explicit formula, also called *Cramer's rule*, for the entries in the (unique) solution vector $\vec{x}$ to a nonsingular linear system $A\vec{x} = \vec{b}$. This formula expresses the entries of $\vec{x}$ entirely in terms of determinants. It is a pretty formula, but also quite a useless one, in the sense that for large systems, computing the solution vector $\vec{x}$ this way would be prohibitively time-consuming, while Gaussian elimination would solve the system relatively quickly. Moreover, the formula only works in the nonsingular case. So I don't give the formula here, and refer you to Wikipedia or some other reference.