Some proofs about determinants

Samuel R. Buss - Spring 2003

Revision 2.1
(Preliminary, corrections appreciated!)

These notes are written to supplement sections 2.1 and 2.2 of the textbook *Linear Algebra with Applications* by S. Leon for my Math 20F class at UCSD. In those sections, the definition of determinant is given in terms of the cofactor expansion along the first row, and then a theorem (Theorem 2.1.1) is stated that the determinant can also be computed by using the cofactor expansion along any row or along any column. This fact is true (of course), but its proof is certainly not obvious. Unfortunately, Leon's text does not give any proof of this theorem, and then uses it heavily in subsequent proofs.

Since most of the book is good about giving proofs, or at least proof sketches, it is galling that such a fundamental result is stated without proof. Accordingly, the following notes give a sketch of how to prove the theorems in sections 2.1 and 2.2 without depending on any unproved theorems. The intent is that the reader can read this in conjunction with Leon's textbook. By following the sequence of definitions and theorems given below, and by filling the details of the proofs, the reader can give a complete proof of all the results.

1 Definition of determinants

For our definition of determinants, we express the determinant of a square matrix A in terms of its cofactor expansion along the first column of the matrix. This is different than the definition in the textbook by Leon: Leon uses the cofactor expansion along the first *row*. It will take some work, but we shall later see that this is equivalent to our definition.

Formally, we define the determinant as follows:

Definition Let A be a $n \times n$ matrix. Then the *determinant* of A is defined by the following. If A is 1×1 so that $A = (a_{1,1})$, then $\det(A) = a_{1,1}$. Otherwise, if $n > 1$,

$$\det(A) \;=\; \sum_{i=1}^{n} a_{i,1} A_{i,1}, \tag{1}$$

where $A_{i,j}$ is the (i,j)-*cofactor* associated with A. In other words, it is the scalar value

$$A_{i,j} \;=\; (-1)^{i+j} \det(M_{i,j}),$$

where $M_{i,j}$ is the $(n-1) \times (n-1)$ matrix obtained from A by removing its i-th row and j-th column. The $M_{i,j}$'s are called the *minors* of A.

In this note, we assume that all matrices are square. We use the notations $A_{i,j}$ and $M_{i,j}$ to refer to the cofactors and minors of A. When working with multiple matrices, we use also use $M^A_{i,j}$ to denote the minor $M_{i,j}$ of A. Likewise, for B a matrix, we use $B_{i,j}$ and $M^B_{i,j}$ to denote the cofactors and minors of B.

2 How row operations affect determinants

We now present a series of theorems about determinants that should be proved in the order presented. These theorems are aimed at showing how row and column operations affect determinants. Indeed, as we shall see, row and column operations preserve the property of the determinant being non-zero. More generally, there are simple rules that tell how a determinant when a row or column operation is applied.

Theorem 1 (Multiplying a row by a scalar.) *Let A be a square matrix. Let B be obtained from A by multiplying the kth row of A by α. Then*

$$\det(B) \;=\; \alpha \cdot \det(A).$$

Proof: We prove the theorem by induction on n. The base case, where A is 1×1 is very simple, since $\det(B) = b_{1,1} = \alpha a_{1,1} = \alpha \det(A)$.

For the induction step, we assume the theorem holds for all $(n-1)\times(n-1)$ matrices and prove it for the $n \times n$ matrix A. Recall that the determinant of A is

$$\det(A) \;=\; \sum_{i=1}^{n} a_{i,1}A_{i,1}.$$

Likewise, the determinant of B is

$$\det(B) \;=\; \sum_{i=1}^{n} b_{i,1}B_{i,1}.$$

Consider the ith term in these two summations. First suppose $i = k$. Then $b_{i,1} = \alpha a_{i,1}$. Also, since A and B differ in only their kth rows, $M^B_{i,1} = M^A_{i.1}$, and thus $A_{i,1} = B_{i,1}$. Thus, for $i = k$, $b_{i,1}B_{i,1} = \alpha a_{i,1}A_{i,1}$. Second, suppose $i \neq k$. Then $b_{i,1} = a_{i,1}$. Also, $M^B_{i,1}$ is obtained from $M^A_{i,1}$ by multiplying one of its rows by α. Therefore, by the induction hypothesis, $B_{i,1} = \alpha A_{i,1}$. Thus, we again have $b_{i,1}B_{i,1} = \alpha a_{i,1}A_{i,1}$.

Since $b_{i,1}B_{i,1} = \alpha a_{i,1}A_{i,1}$ holds for all i, we conclude that $\det(B) = \alpha \det(A)$, and the theorem is proved. □

Corollary 2 *Let A be a square matrix. If any row of A is all zero, then* $\det(A) = 0$.

Proof: This is an immediate corollary of Theorem 1 using $\alpha = 0$. □

Our next theorems use matrices A, B and C. These are always assumed to be square and have the same dimensions. Furthermore, our proofs will use the notations $A_{i,j}$, $B_{i,j}$ and $C_{i,j}$ for the cofactors of A, B and C. We also use the notations $M^A_{i,j}$, $M^B_{i,j}$ and $M^C_{i,j}$ for the minors of the three matrices. Recall that $\mathbf{a}_{:,i}$, $\mathbf{b}_{:,i}$, and $\mathbf{c}_{:,i}$ denote the ith rows of the matrices A, B, and C.

Theorem 3 *Suppose i_0 is a fixed number such that $1 \leq k \leq n$. Also suppose A, B and C satisfy*

$$\mathbf{c}_{:,k} \;=\; \mathbf{a}_{:,k} + \mathbf{b}_{:,k}$$

and that, for all $i \neq k$,

$$\mathbf{a}_{:,i} \;=\; \mathbf{b}_{:,i} \;=\; \mathbf{c}_{:,i}.$$

Then

$$\det(C) \;=\; \det(A) \;+\; \det(B).$$

The hypotheses of the theorem say that A, B, and C are the same, except that the k row of C is the sum of the corresponding rows of A and B.

> **Proof:** The proof uses induction on n. The base case $n = 1$ is trivially true. For the induction step, we assume that the theorem holds for all $(n-1)\times(n-1)$ matrices and prove it for the $n \times n$ matrices A, B, C. From the definitions $\det(A)$, $\det(B)$, and $\det(C)$, it will suffice to prove that
>
> $$c_{i,1}C_{i,1} \;=\; a_{i,1}A_{i,1} + b_{i,1}B_{i,1} \tag{2}$$
>
> holds for all $i = 1, \ldots, n$. First, suppose $i = k$. Then $c_{i,1} = a_{i,1} + b_{i,1}$. Also, since the matrices differ only in their kth rows, $C_{i,1} = A_{i,1} = B_{i,1}$. Thus, equation (2) holds for $i = k$. Second, suppose $i \neq k$. Then, $c_{i,1} = a_{i,1} = b_{i,1}$. Also, by the induction hypothesis, we have that $C_{i,1} = A_{i,1} + B_{i,1}$. This is because $M^A_{i,1}$, $M^B_{i,1}$, and $M^C_{i,1}$ in equal all but one of their rows; the remaining row in $M^C_{i,1}$ is the sum of the corresponding rows in $M^A_{i,1}$ and $M^B_{i,1}$. So again, (2) holds. □

Theorem 4 *Suppose that B is obtained from A by swapping two of the rows of A. Then $\det(B) = -\det(A)$.*

> **Proof:** We shall first prove the theorem under the assumption that row 1 is swapped with row k, for $k > 1$. This will be sufficient to prove the theorem for swapping any two rows, since swapping rows k and k' is equivalent to performing three swaps: first swapping rows 1 and k, then swapping rows 1 and k', and finally swapping rows 1 and k.
>
> The proof is by induction on n. The base case $n = 1$ is completely trivial. (Or, if you prefer, you may take $n = 2$ to be the base case, and the theorem is easily proved using the formula for the determinant of a 2×2 matrix.)
>
> The definitions of the determinants of A and B are:
>
> $$\det(A) = \sum_{i=1}^{n} a_{i,1}A_{i,1} \qquad \text{and} \qquad \det(B) = \sum_{i=1}^{n} b_{i,1}B_{i,1}.$$
>
> First suppose $i \notin \{1, k\}$. In this case, it is clear that $M^A_{i,1}$ and $M^B_{i,1}$ are the same except for two rows being swapped. Therefore, $A_{i,1} = -B_{i,1}$. Since also $a_{i,1} = b_{i,1}$, we have that $b_{i,1}B_{i,1} = -a_{i,1}A_{i,1}$.
>
> It remains to consider the $i = 1$ and $i = k$ terms. We claim that
>
> $$a_{k,1}A_{k,1} = -b_{1,1}B_{1,1} \qquad \text{and} \qquad a_{1,1}A_{1,1} = -b_{k,1}B_{k,1}.$$

In fact, once we prove these two identities, the theorem will be proved. By symmetry, it will suffice to prove the first identity. For this, first note that $a_{k,1} = b_{1,1}$. Second, note that $M^B_{1,1}$ is obtained from $M^A_{k,1}$ by reordering the rows $1, 2, \ldots, k-1$ of $M^A_{k,1}$ into the order $2, 3, \ldots, k-1, 1$. This reordering can be done by swapping row 1 with row 2, then swapping that row with row 3, etc., ending with swap with row $k-1$. This is a total of $k-2$ row swaps. So, by the induction hypothesis,

$$\det(M^B_{1,1}) \;=\; (-1)^{k-2}\det(M^A_{k,1}) \;=\; (-1)^k\det(M^A_{k,1}).$$

Since $B_{k,1} = (-1)^{k+1}\det(M^B_{k,1})$ and $A_{1,1} = \det M^A_{1,1}$, we have established that $A_{k,1} = -B_{1,1}$. Thus, $a_{k,1}A_{k,1} = -b_{1,1}B_{1,1}$.

This completes the proof of the theorem. □

Corollary 5 *If two rows of A are equal, then $\det(A) = 0$.*

Proof: This is an immediate consequence of Theorem 4 since if the two equal rows are switched, the matrix is unchanged, but the determinant is negated. □

Corollary 6 *If B is obtained from A by adding α times row i to row j (where $i \neq j$), then $\det(B) = \det(A)$. (This is a row operation of type 3.)*

Proof: Let C be the matrix obtained from A by replacing row j with row i. Then, by Theorem 5, $\det(C) = 0$. Now, modify C by multiplying row j by α to obtain D. By Theorem 1, $\det(D) = \alpha\det(C) = 0$. Now, by Theorem 3,

$$\det(B) \;=\; \det(A) \;+\; \det(D) \;=\; \det(A) + 0 \;=\; \det(A).$$

□

Summary of section: Among other things, we have shown how the determinant matrix changes under row operations and column operations. For row operations, this can be summarized as follows:

R1 If two rows are swapped, the determinant of the matrix is negated. (Theorem 4.)

R2 If one row is multiplied by α, then the determinant is multiplied by α. (Theorem 1.)

R3 If a multiple of a row is added to another row, the determinant is unchanged. (Corollary 6.)

R4 If there is a row of all zeros, or if two rows are equal, then the determinant is zero. (Corollary 2 and Corollary 5.)

For column operations, we have similar facts, which we list here for convenience. To prove them, we must first prove that $\det(A) = \det(A^T)$, which will be done later as Theorem 15.

C1 If two columns are swapped, the determinant of the matrix is negated. (Theorem 22.)

C2 If one column is multiplied by α, then the determinant is multiplied by α. (Theorem 19.)

C3 If a multiple of a column is added to another column, the determinant is unchanged. (Corollary 24.)

C4 If there is a column of all zeros, or if two columns are equal, then the determinant is zero. (Corollary 20 and Corollary 23.)

3 Diagonal and tridiagonal matrices

The next theorem states that the determinants of upper and lower triangular matrices are obtained by multiplying the entries on the diagonal of the matrix.

Theorem 7 *Let A be an upper triangular matrix (or, a lower triangular matrix). Then, $\det(A)$ is the product of the diagonal elements of A, namely*

$$\det(A) \;=\; \prod_{i=1}^{n} a_{i,i}.$$

Proof: The proof is by induction on n. For the base case, $n = 1$, the theorem is obviously true. Now consider the induction case, $n > 1$, with A upper triangular or lower triangular. By the induction hypothesis, $M^A_{1,1}$ is the product of all the entries on the diagonal of A except $a_{1,1}$. Thus, $a_{1,1}A_{1,1}$ is the product of the diagonal entries of A. Therefore, from the formula (1) for the determinant of A, it will suffice to prove that

$$a_{i,1}A_{i,1} \;=\; 0,$$

for $i > 1$. Now, if A is upper triangular, then $a_{i,1} = 0$ when $i > 0$. On the other hand, if A is lower triangular and $i > 1$, then the first row of $M^A_{i,1}$ contains all zeros, so $A_{i,1} = 0$ by Theorem 2.

That completes the proof by induction. □.

Since a diagonal matrix is both upper triangular and lower triangular, Theorem 7 applies also to diagonal matrices.

Corollary 8 *Let I be an identity matrix. Then* $\det(I) = 1$.

4 Determinants of elementary matrices

Theorem 9 *Let E be an elementary matrix of type I. Then* $\det(E) = -1$.

Proof: Any such E is obtained from the identity matrix by interchanging two rows. Thus, $\det(E) = -1$ follows from the facts that the identity has determinant 1 (Corollary 8) and that swapping two rows negates the determinant (Theorem 4). □

Theorem 10 *Let E be an elementary matrix of Type II, with $e_{i,i} = \alpha$. Then* $\det(E) = \alpha$.

> **Proof:** This is an immediate consequence of Theorem 7. □

Theorem 11 *Let E be an elementary matrix of Type III. Then* $\det(E) = 1$.

> **Proof:** This is immediate from Theorem 7. □

The next two corollaries will come in handy.

Corollary 12 *Let E be an elementary matrix. Then* $\det(E^T) = \det(E)$.

Corollary 13 *Let E be an elementary matrix. Then* $\det(E^{-1}) = 1/\det(E)$.

As we shall see, Corollary 12 actually holds for any square matrix A, not just for elementary matrices. And, Corollary 13 holds for any invertible matrix. Both corollaries are easily proved from the previous three theorems.

The next theorem is an important technical tool. (It will be superseded by Theorem 17 below.)

Theorem 14 *Let A be a square matrix, and let E be an elementary matrix. Then*

$$\det(EA) \;=\; \det(E)\det(A).$$

> **Proof:** This is an immediate consequence of Theorems 9-11 and Theorems 1, 4, and 6. □

5 How to compute a determinant efficiently

We know that any matrix can be put in row echelon form by elementary operations. That is to say, any matrix A can be transformed into a row echelon form matrix B by elementary row operations. This gives us

$$B \;=\; E_k E_{k-1} \cdots E_2 E_1 A$$

where B is in row echelon form and hence upper triangular. Since B is upper triangular, we can easily compute its determinant. By Theorem 14,

$$\det(B) \;=\; \det(E_k)\det(E_{k-1})\cdots\det(E_2)\det(E_1)\det(A).$$

Then, by Corollary 13,

$$\det(A) \;=\; \det(E_k)^{-1}\det(E_{k-1})^{-1}\cdots\det(E_2)^{-1}\det(E_1)^{-1}\det(B).$$

This gives an algorithm for computing the determinant, $\det(A)$, of A.

This algorithm is quite efficient; however, for hand calculation it is sometimes easier to not put A in row echelon form, but instead to only get B upper triangular. In this case, the determinant of B is still easily computable; in addition, the upper triangular B can be obtained using only row operations of types I and III, each of the elementary matrices E_i has determinant ± 1, and thus $\det(E_i) = \det(E_i)^{-1} = \pm 1$, so we need only keep track of the sign changes. This latter algorithm is the one we advocate in class lecture as being the best one to use.

6 Determinants and invertibility

Theorem 15 *A square matrix A is invertible if and only if its reduced row echelon form is the identity matrix. Furthermore, it is invertible if and only if its row echelon form does not have any free variables.*

> **Proof:** To prove this theorem, note that the conditions are satisfied if and only if there is no row of zeros in the (reduced) row echelon form of A. This is equivalent to the condition that the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. $\square$.

Corollary 16 *A is invertible if and only if* $\det(A) \neq 0$.

> **Proof:** By the discussion at the beginning of this section, A has determinant zero if and only if its reduced row echelon form B has determinant zero. Now, if A is invertible, B is the identity and hence has determinant equal to one; i.e., if A is invertible, A has nonzero determinant. Otherwise, B has a row of all zeros and thus has determinant zero, so A has determinant equal to zero. $\square$

7 Determinants of products of matrices

A very important fact about matrices is that $\det(AB) = \det(A) \cdot \det(B)$.

Theorem 17 *Let A and B be $n \times n$ matrices. Then*

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

> **Proof:** First suppose $\det(B) = 0$. Then $\det(B)$ is not invertible, so there is a non-trivial solution to $B\mathbf{x} = \mathbf{0}$. This is also a non-trivial solution to $AB\mathbf{x} = \mathbf{0}$, so AB is not invertible and thus has determinant 0. Then $\det(AB) = 0 = \det(A)\det(B)$ in this case.
>
> Second suppose $\det(A) = 0$ and $\det(B) \neq 0$. Since A is not invertible, there is a non-trivial solution $\mathbf{y} \neq \mathbf{0}$ to $A\mathbf{y} = 0$. But then, $\mathbf{x} = B^{-1}(\mathbf{y})$ is a non-trivial solution to $AB\mathbf{x} = \mathbf{0}$. Therefore, AB is not invertible, so $\det(AB) = 0$. So again, $\det(AB) = 0 = \det(A)\det(B)$.
>
> Now suppose that $\det(A) \neq 0$. Then the rref form of A is just the identity I. This means there are elementary matrices E_i so that
>
> $$A = E_k E_{k-1} \cdots E_2 E_1.$$
>
> Then,
>
> $$\det(A) = \det(E_k)\det(E_{k-1}) \cdots \det(E_2)\det(E_1).$$

by Theorem 14. Using Theorem 14 again gives

$$\det(AB) \;=\; \det(E_k)\det(E_{k-1})\cdots\det(E_2)\det(E_1)\det(B) \;=\; \det(A)\det(B).$$

So the Theorem is proved. □

8 Determinants of transposes

Theorem 18 $\det(A^T) = \det(A)$.

Proof: Express A in row echelon form B, i.e.,

$$A \;=\; E_k E_{k-1}\cdots E_2 E_1 B.$$

So, by Theorem 17,

$$\det(A) = \det(E_k)\det(E_{k-1})\cdots\det(E_2)\det(E_1)\det(B).$$

The matrix B is upper triangular, so B^T is lower triangular. In addition, B^T and B have the same diagonal entries and thus the same determinant. We also have

$$A^T \;=\; B^T E_1^T E_2^T \cdots E_{k-1}^T E_k^T.$$

Using Theorem 17 again,

$$\det(A^T) = \det(E_k^T)\det(E_{k-1}^T)\cdots\det(E_2^T)\det(E_1^T)\det(B^T).$$

The theorem now follows from Corollary 12. □

9 How column operations effect determinants

Now that we have proved Theorem 18 that determinants are preserved under taking transposes, we automatically know that all the facts established in section 2 for row operations also hold for column operations:

Theorem 19 (Multiplying a column by a scalar.) *Let A be a square matrix. Let B be obtained from A by multiplying the kth column of A by α. Then*

$$\det(B) \;=\; \alpha\cdot\det(A).$$

Recall that $\mathbf{a}_i$, $\mathbf{b}_i$, and $\mathbf{c}_i$ denote the ith columns of the matrices A, B, and C.

Corollary 20 *Let A be a square matrix. If any column of A is all zero, then* $\det(A) = 0$.

Theorem 21 *Suppose j_0 is a fixed number such that $1 \le j_0 \le n$. Also suppose A, B and C satisfy*

$$\mathbf{c}_{j_0} \;=\; \mathbf{a}_{j_0} + \mathbf{b}_{j_0}$$

and that, for all $j \neq j_0$,

$$\mathbf{a}_j \;=\; \mathbf{b}_j \;=\; \mathbf{c}_j.$$

Then

$$\det(C) \;=\; \det(A) \;+\; \det(B).$$

The hypotheses of the theorem say that A and B and C are the same, except that the j_0 column of C is the sum of the corresponding columns of A and B.

Theorem 22 *Suppose that B is obtained from A by swapping two of the columns of A. Then* $\det(B) = -\det(A)$.

Corollary 23 *If two columns of A are equal, then* $\det(A) = 0$.

Corollary 24 *If B is obtained from A by adding α times column i to column j ($i \neq j$), then* $\det(B) = \det(A)$.

10 Cofactor expansion along any column or row

We are now in a position to prove that the determinant can be calculated in terms of its cofactor expansion along any column, or along any row.

The definition of the cofactor expansion along column j is:

$$\sum_{i=1}^{n} a_{i,j} A_{i,j}. \tag{3}$$

The definition of the cofactor expansion along row i is:

$$\sum_{j=1}^{n} a_{i,j} A_{i,j}. \tag{4}$$

Theorem 25 *For any $j = 1, \ldots, n$,* $\det(A)$ *is equal to the quantity (3).*

Theorem 26 *For any $i = 1, \ldots, n$,* $\det(A)$ *is equal to the quantity (4).*

Theorem 25 may be proved as follows: Let B be obtained from A by swapping columns 1 and j. By Theorem 4, $\det(B) = -\det(A)$. Then, prove that the terms in the cofactor expansion (3) along column j are equal to the negations of the terms in the definition of the determinant of B (using the cofactor expansion along column 1 to compute the determinant of B); in other words, prove that $a_{i,j}A_{i,j} = -b_{i,1}B_{i,1}$ for $i = 1, 2, \ldots, n$. We leave the details to the reader.

Theorem 26 is an immediate consequence of Theorem 25 and the fact that determinants are preserved under transposes (Theorem 18).