LITERALLY EVERYTHING I KNOW ABOUT

Linear Algebra

Warren Alphonso

A very reductionist summary of LINEAR ALGEBRA AND ITS APPLICATIONS by Lay, Lay, and McDonald, as well as LINEAR ALGEBRA DONE WRONG by Treil.

June 1, 2019

Contents

Chapter 1

Determinants

1.1 Properties

In this chapter, we only consider determinants of $n \times n$ matrices. We will think of the determinant as the n-dimensional volume of the parallelepiped determined by our n vectors, $v_1, \cdots, v_n$. For dimensions 2 and 3, "volume" of the parallelepiped is determined with the *base times height* rule: we pick one vector and define height to be the distance between this vector and the subspace spanned by the $n-1$ remaining vectors. Then we define base to be the $(n-1)$-dimensional volume of the parallelepiped determined by the $n-1$ vectors.

This understanding allows determinants the following properties:

1. **Linearity in each argument:** Multiplying some vector v_k by α, then the height is multiplied by the same constant which means the determinant is multiplied by the same constant. This means the determinant is *linear in each argument*, which means if we fix $n-1$ vectors, the determinant is linear with respect to the final vector.

   Linearity means that for an $n \times n$ matrix A, $det(\alpha A) = \alpha^n det(A)$, because multiplying A by α is equivalent to multiplying n columns by α.

2. **Preservation under column replacement:**

$$det(v_1, \cdots, v_j + \alpha v_k, \cdots, v_k, \cdots, v_n) = det(v_1, \cdots, v_j, \cdots, v_k, \cdots, v_n)$$

   Realize that this is true because the "height" of $v_j + \alpha v_k$ is the same as the "height" of v_j, since "height" is defined in relation to the distance from the remaining subspace.

3. **Antisymmetry:** Swapping two vectors means the determinant changes signs.

$$det(v_1, \cdots, v_k, \cdots, v_j, \cdots, v_n) = -det(v_1, \cdots, v_j, \cdots, v_k, \cdots, v_n)$$

   This does not seem natural at first, but we can prove it by applying Preservation under column replacement thrice and then linearity.

$$det(v_1, \cdots, v_k, \cdots, v_j, \cdots, v_n)$$

$$= det(v_1, \cdots, v_k, \cdots, v_j - v_k, \cdots, v_n)$$

$$= det(v_1, \cdots, v_k + (v_j - v_k), \cdots, v_j - v_k, \cdots, v_n)$$

$$= det(v_1, \cdots, v_j, \cdots, v_j - v_k - (v_j), \cdots, v_n)$$

$$= det(v_1, \cdots, v_j, \cdots, -v_k, \cdots, v_n)$$

$$= -det(v_1, \cdots, v_j, \cdots, v_k, \cdots, v_n)$$

4. **Normalization:** For the standard basis, the corresponding parallelepiped is the n-dimensional unit cube so its volume is 1.

$$det(I) = 1$$

Using these, we can derive more basic properties of determinants for a square matrix A:

1. If A has a zero column, then $det(A) = 0$.

2. If A has two equal columns, then $det(A) = 0$.

3. If one column of A is a multiple of another, then $det(A) = 0$.

1.2 Computing the Determinant

The **determinant of diagonal matrices** is the product of the diagonal entries. Note that any diagonal matrix $\{a_1, \cdots, a_k\}$ can be obtained by multiplying column k by a_k.

The **determinant of triangular matrices** is also the product of the diagonal entries. This is because an upper or lower triangular matrix can be reduced to a diagonal matrix with the same diagonal entries through column operations.

Theorem 1.2.1. $det(A) = 0$ *if and only if A is not invertible.*

Proof. Recall that we can only use column operations when reducing a matrix to find the determinant, which is equivalent to doing row operations on A^T. If the echelon form of A^T does not have pivots in every column and row, then the product of diagonal entries will be 0. Not having pivots in every column and row also means the matrix is not invertible, so the two conclusions are equivalent. ∎

We will now prove some nontrivial properties of determinants, but to do so, we will need the following two theorems.

Theorem 1.2.2. *For a square matrix A and elementary matrix E,*

$$det(AE) = det(A)det(E)$$

Proof. Right multiplication of an elementary matrix is simply a column operation. Since a column operation is obtained from the identity matrix by the column operation, its determinant is 1 times the effect of the column operation. ∎

Theorem 1.2.3. *Any invertible matrix is a product of elementary matrices.*

Proof. We know that any invertible matrix is *row equivalent* to the identity matrix, which is its reduced echelon form. So

$$I = E_n E_{n-1} \cdots E_1 A$$

which means we can write A in terms of the identity and the inverses of some elementary matrices

$$A = E_1^{-1} \cdots E_{n-1}^{-1} E_n^{-1} I = E_1^{-1} \cdots E_{n-1}^{-1} E_n^{-1}$$

since the inverse of an elementary matrix is an elementary matrix, the proof is complete. ∎

Now for two important theorems:

Theorem 1.2.4. *For a square matrix A,*

$$det(A) = det(A^T)$$

Proof. A key observation is that $det(E) = det(E^T)$ for any elementary matrix E.
Notice also that it is sufficient to prove the theorem *only* for *invertible matrices A* since if A is not invertible then A^T is also not invertible and both determinants are 0, trivially proving the theorem.
Now, by the above theorem we can write

$$A = E_1 E_2 \cdots E_n$$

which means

$$det(A) = det(E_1)det(E_2) \cdots det(E_n)$$

We can also write

$$A^T = E_n^T \cdots E_2^T E_1^T = E_n \cdots E_2 E_1$$

which means

$$det(A^T) = det(E_n) \cdots det(E_2)det(E_1)$$

which is equivalent to $det(A)$. ∎

This theorem means that column operations have the same effect on determinants as row operations, so we can use either when reducing matrices to compute determinants.

Theorem 1.2.5. *For $n \times n$ matrices A, B,*

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

Proof. Two cases:
Case 1: *B* is invertible.
This means we can write

$$B = E_1 E_2 \cdots E_n$$

and so

$$det(AB) = det(A)[det(E_1)det(E_2)\cdots det(E_n)] = det(A)det(B)$$

Case 2: B is not invertible. If B is not invertible, we will prove that the product AB is also not invertible so $det(AB) = det(A)det(B) \to 0 = 0$.

We proceed by contradiction. Assume $AB = C$ is invertible. Then we left multiply both sides by C^{-1} to get $C^{-1}AB = I$, which means $C^{-1}A$ is the left inverse of B, but because B is square, that means $C^{-1}A$ is the inverse of B. Since we know B is not invertible, we have a contradiction. ∎

1.3 Cofactor Expansion

For an $n \times n$ matrix A, let $A_{j,k}$ denote the $(n-1) \times (n-1)$ matrix obtained from A by crossing out row j and column k.

Theorem 1.3.1. *Cofactor expansion of determinant For each $j, 1 \le j \le n$, determinant of A can be expanded in the row number j as*

$$det(A) = a_{j,1}(-1)^{j+1}det(A_{j,1}) + a_{j,2}(-1)^{j+2}det(A_{j,2}) + \cdots + a_{j,n}(-1)^{j+n}det(A_{j,n})$$

A similar expansion can be done for columns.

Proof. We will prove the expansion for row 1. This can be generalized by swapping row 1 with another row. Additionally, since $det(A) = det(A^T)$, column expansion follows automatically.

Consider the special case when the first row has *only one* nonzero term, $a_{1,1}$. Performing column operations on columns $2, \cdots, n$, we transform A to lower triangular form. Now

$$det(A) = (\text{product of diagonal entries}) \times (\text{correcting factor from column operations})$$

but since the *product of diagonal entries* except $a_{1,1}$ times the *correcting factor* is exactly $det(A_{1,1})$, we can write

$$det(A) = a_{1,1}det(A_{1,1})$$

Now consider the case when all entries in the first row except $a_{1,2}$ are zeros. We can reduce this to the previous case by swapping columns 1 and 2, so $det(A) = (-1)a_{1,2}det(A_{1,2})$.

If $a_{1,3}$ is the only nonzero term in the first row, we can reduce this to the previous case by swapping columns 2 and 3, so $det(A) = a_{1,3}det(A_{1,3})$.

These special cases are important because we have linearity of the determinant. If the matrix $A^{(k)}$ is obtained by replacing all A's entries in the first row with 0 except for $a_{1,k}$, then linearity of the determinant implies

$$det(A) = det(A^{(1)}) + \cdots + det(A^{(n)}) = \sum_{k=1}^{n} det(A^{(k)})$$

Based on our analysis of special cases, we know

$$det(A^{(k)}) = (-1)^{1+k} a_{1,k} det(A_{1,k})$$

so

$$det(A) = \sum_{k=1}^{n} (-1)^{1+k} a_{1,k} det(A_{1,k})$$

To get the expansion for the second row, we swap rows so multiply by -1. For the third row, multiply by -1 again to get the original equation, and so on. ∎

Cofactor expansion is not practical for anything larger than a 3×3 matrix, but it has great theoretical importance.

Definition 1.3.1. Formally, the numbers

$$C_{j,k} = (-1)^{j+k} det(A_{j,k})$$

are called **cofactors**.
The matrix $C = \{C_{j,k}\}_{j,k=1}^{n}$ whose entries are *cofactors* of a given matrix A is called the **cofactor matrix** of A.

Theorem 1.3.2. *Cofactor formula for inverse Let A be an invertible matrix and let C be its cofactor matrix. Then*

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

Proof. Let us find the product AC^T.
The jth diagonal entry is obtained by multiplying the jth row of A by the jth row of C,

$$(AC^T)_{j,j} = a_{j,1} C_{j,1} + \cdots + a_{j,n} C_{j,n} = det(A)$$

by cofactor expansion.
To get the off-diagonal terms, we multiply the kth row of A with the jth row of C, $j \neq k$,

$$a_{k,1} C_{j,1} + \cdots + a_{k,n} C_{j,n}$$

If we look at this as a cofactor expansion of the jth row, this is the determinant of the matrix A except that we replace row j with row k. Since two rows of our matrix coincide, the determinant will be 0, which means all off-diagonal terms will be 0, thus

$$AC^T = det(A)I$$

Dividing by $det(A)$ gives us

$$A \frac{1}{det(A)} C^T = I$$

so $\frac{1}{det(A)} C^T$ is the right inverse of A, which means it is the inverse of A since A is square. ∎

Since for invertible matrices, $Ax = b$ has a unique solution, we have

$$x = A^{-1}b = \frac{1}{det(A)}C^T b$$

Theorem 1.3.3. *Cramer's Rule For invertible matrix A, entry k of the solution to $Ax = b$ is given by*

$$x_k = \frac{det(B_k)}{det(A)}$$

where B_k is obtained from A by replacing column k with b.

Proof. After our above theorem, we need only prove that entry k of $C^T b = det(B_k)$.

We know entry k of $C^T b$ is equivalent to the product of the kth row of C^T and b, which is equivalent to the product of the kth column of C and b.

Element j,k of C is obtained by crossing out the jth row and kth column of A and computing the determinant of the remaining matrix. Multiplying the kth column of C with b is equivalent to

$$b_1 C_{1,k} + \cdots + b_n C_{n,k}$$

which is the same as the cofactor expansion of B_k. ∎

One application of the cofactor formula is a shortcut to inverting 2×2 matrices. For the matrix

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

The cofactor matrix is made up of 4 1×1 matrices,

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

which means

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