Princeton University MAT 202 Spring 2008

Willie W. Wong

Mar 31 - Apr 4

MAR 31, 2008

- Using the 3×3 determinant, we want to construct a formula for determinants of $N \times N$ matrices. Recall that for a 3×3 matrix A, we have

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

- This leads us to the definition of sub-matrices and matrix minors:
  - Recall that for an $N \times M$ matrix A, we use the lower case letter a to denote its components, in particular, a_{ij} denote the entry on the ith row and jth column of A.
  - We shall, henceforth, write A_{ij} (with upper case A) to denote the sub-matrix given by removing the ith row and jth column of A. A_{ij} is an $(N-1) \times (M-1)$ matrix. So for a 3×3 matrix A, we can write

$$A_{31} = \begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix} \quad , \quad A_{32} = \begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix} \quad , \quad A_{33} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

    and thus the expression for the determinant as

$$\det(A) = a_{31}\det(A_{31}) - a_{32}\det(A_{32}) + a_{33}\det(A_{33})$$

    (mathematicians often call a sum like this, with one term positive, the next term negative, and the term afterwards positive, and so on, an *alternating sum*).
  - When A is square, we call the expression $\det(A_{ij})$ the ijth *minor* of the matrix A.

- This allows us to introduce the Laplace/cofactor expansion for the determinant: Given an $N \times N$ matrix A, we can fix a row i and define the determinant

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

we can also, equivalently, fix a column j and define the determinant

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

The determinant thus defined is independent of whether you chose a row expansion or a column expansion, and is independent of which row or column you fix in the beginning.

– The factor $(-1)^{i+j}$ makes the sum alternating in sign.

– The expression $(-1)^{i+j} \det(A_{ij})$ is also often called the ijth *cofactor* of A.

– Notice that this is a *recursive* definition: to calculate the determinant for a matrix of size $N \times N$, we need to first calculate the minors, which are determinants of matrices of size $(N-1) \times (N-1)$. Once we get down to 3×3, we can use Sarrus's rule from last Friday to finish.

– The calculation is usually very long and tedious: for the determinant of a 5×5 matrix, after fixing a column, say, we see that we need to calculate 5 matrix minors, i.e. determinants of 5 different 4×4 matrices. For each of those, we need to calculate determinants of 4 3×3 matrices. All in all, we need to invoke Sarrus's rule 20 times....

– A general trick: expand down a column (or across a row) where most of the terms are 0s. Suppose

$$A = \begin{bmatrix} 1 & 5 & 4 & 3 & 6 & 0 \\ 2 & 0 & 3 & 1 & 0 & 7 \\ 0 & 3 & 1 & 0 & 2 & 0 \\ 0 & 5 & 0 & 1 & 3 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 2 & 0 \end{bmatrix}$$

Noting that the last column has only 1 term that is non-zero, we see that doing Laplace's expansion down the last column gives us only 1 non-zero term

$$\det(A) = 0 + (-1)^{2+6} 7 \det(A_{26}) + 0 + 0 + 0 + 0$$

For $B = A_{26}$, we notice that the first column only has a non-zero term in the 11 position:

$$\det(B) = (-1)^{1+1} \cdot 1 \cdot \det(B_{11})$$

Now $C = B_{11}$ is

$$C = \begin{bmatrix} 3 & 1 & 0 & 2 \\ 5 & 0 & 1 & 3 \\ 2 & 0 & 0 & 1 \\ 3 & 0 & 0 & 2 \end{bmatrix}$$

so we now use Laplace expansion on the second column,

$$\det(C) = (-1)^{1+2} \cdot 1 \cdot \det(C_{12})$$

the resulting 3×3 matrix we can evaluate using Sarrus's rule:

$$\det(C_{12}) = \det(\begin{bmatrix} 5 & 1 & 3 \\ 2 & 0 & 1 \\ 3 & 0 & 2 \end{bmatrix}) = 3 - 4 = -1$$

So tracing back we have

$$\det(A) = 7 \cdot 1 \cdot (-1) \cdot 1 \cdot (-1) = 7$$

- When ever we have a recursive definition, facts about the definition can most easily be proven using the method of mathematical induction.
  - When using induction, the *goal* is always to prove some statement X for all natural numbers.
  - If we try to prove the statement directly, it will take infinite amount of time: there are infinitely many natural numbers after all, and if we examine each case individually, it will take forever.
  - The principle of mathematical induction allows us to show that X must be true for all natural numbers by showing that " we know how to prove it for an arbitrary natural number."
  - The basic idea of mathematical induction is a lot like climbing a ladder: suppose you have a ladder going straight up from the ground floor to the 12th floor of Fine Hall. I want to convince you that I can, by climbing the ladder, get to the 12th floor of Fine Hall. One way to convince you is by actually climbing all the way to the 12th floor and waving to you thence. This is the brute force approach. A different way is like this:
    - I first hop onto the ladder at the ground floor, convincing you that I know how to get *on* the first rung.
    - I next climb a couple steps higher, convincing you that once I am on the ladder, I know how to climb up.
    - Then I can just hop off the ladder to your satisfaction since the only skills I need to climb a ladder to the 12th floor of Fine Hall have been demonstrated.
  - Mathematical induction works formally exactly the same way. By showing the following two things you can show that a statement X must be true for all natural numbers:
    - First you need to show that X is true for the number 1. This is getting on the ladder.
    - Next you first suppose that X is true for the number N (this is called the *induction hypothesis*; this is saying "hey look! I am already on the Nth rung of the ladder"), then you prove, using the induction hypothesis, that X is true for the number $N+1$ (this is saying "hey look again! Now I've climbed one rung higher from where I was before").
- (Whew! Real life finally caught up to the notes.)

- We begin with an application of the principle of mathematical induction: we show that if a matrix is upper triangular, i.e., if it looks like

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1N} \\ 0 & a_{22} & a_{23} & \dots & a_{2N} \\ 0 & 0 & a_{33} & \dots & a_{3N} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & & a_{NN} \end{bmatrix}$$

  then $\det(A)$ is the product of the diagonal entries:

$$\det(A) = a_{11}a_{22}a_{33}\cdots a_{NN}$$

  - First we need to check the fact is true for a simple case. For the case $N = 1$, it is trivial: $\det(A) = a_{11}$. For the case $N = 2$, we have $\det(A) = ad - bc$, but in this case, $c = 0$, so $\det(A) = ad = a_{11}a_{22}$. For the case $N = 3$, using Sarrus's rule, we can calculate simply that the statement is true.
  - Next we assume the fact is true for N and show that it is true for $N + 1$. Assume now that A is an $(N + 1) \times (N + 1)$ matrix that is upper triangular, this means that $a_{ij} = 0$ whenever $i > j$. We evaluate the determinant of A using Laplace expansion on the first column:

$$\det(A) = (-1)^{1+1}a_{11}\det(A_{11})$$

    since all other terms $a_{i1} = 0$ for $i > 1$. Let $B = A_{11}$ be the $N \times N$ minor matrix. There is a simple relation between the entries of B and the entries of A:

$$b_{ij} = a_{i+1,j+1}$$

    so B is necessarily upper triangular: if $i > j$, then $i + 1 > j + 1$, so $b_{ij} = a_{i+1,j+1} = 0$ by assumption that A is upper triangular. Now we use the induction hypothesis on B: B is an $N \times N$ upper triangular matrix, so its determinant should be the product of its diagonal terms.

$$\det(B) = b_{11}b_{22}\cdots b_{NN} = a_{22}a_{33}\cdots a_{N+1,N+1}$$

    Therefore

$$\det(A) = a_{11}\det(B) = a_{11}a_{22}\cdots a_{N+1,N+1}$$

    the product of the diagonal terms.
  - Remark: it is important to justify the reduction to the induction hypothesis. In our case, we need to take care to make sure that the sub-matrix $A_{11} = B$ is in fact an upper-triangular matrix with dimension $N \times N$.

- The above calculations suggest that it is vastly simpler to compute the determinant for an upper-triangular matrix than for an arbitrary matrix. Now... if only we can make a matrix upper-triangular... A-ha! We CAN!
- Remember Gauss-Jordan elimination? We can use just part of it to get a matrix in upper-triangular form: we don't need use the full force to reduce to RREF. So all that's left is to figure out what the elementary row operations that we employ in G-J affect the determinant.
- We claim:
  - if matrix B is obtained from matrix A by multiplying a fixed row by the number k, then
$$\det(B) = k \det(A)$$
  - if matrix B is obtained from matrix A by swapping two rows, then
$$\det(B) = -\det(A)$$
  - if matrix B is obtained from matrix A by adding a multiple of one row to another, then
$$\det(B) = \det(A)$$
- Proof of the claims (if we have time in class):
  - Suppose B is obtained from multiplying the ith row of A by the number k, then we do Laplace expansion on that row:
$$\det(B) = \sum_j (-1)^{i+j} b_{ij} \det(B_{ij})$$
Now, since all the other rows are unchanged, we have that $B_{ij} = A_{ij}$. And since the ith row is changed by a factor of k, we have $b_{ij} = ka_{ij}$, so
$$\det(B) = \sum_j k(-1)^{i+j} a_{ij} \det(A_{ij}) = k \det(A)$$
  - By induction. First, check that this is true for $N = 2$ (there are not rows to swap in a 1×1 matrix). Let
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$
then
$$\det(A) = ad - bc = -(cb - da) = -\det(B)$$
Now, assuming that the fact is true for $N \times N$ matrices, we check for $(N+1) \times (N+1)$ matrices. Suppose B is obtained from A by switching two rows. Suppose further that

the number of rows of A is at least 3 (since we've just dealt with the case $N = 2$). Then we can do Laplace expansion on a row that is *not* swapped:

$$\det(B) = \sum_j (-1)^{i+j} b_{ij} \det(B_{ij})$$

Since the ith row is not swapped: $a_{ij} = b_{ij}$. Furthermore, it is clear that B_{ij} is obtained from A_{ij} by swapping two rows, and thus we can use the induction hypothesis: $\det(B_{ij}) = -\det(A_{ij})$. Therefore

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

* Notice that an immediate corollary of this fact is that if a matrix A has two rows that are identical, its determinant is 0. This follows because, supposing rows p and q of A are identical. Then we can swap the two rows and keep the matrix the same. By the above argument
$$\det(A) = -\det(A)$$
where A on the left hand side is the "swapped" version and A on the right hand side is the "original" version. The only number that is equal to minus itself is 0, so $\det(A) = 0$.

– Suppose we get B from adding k times the pth row to the qth row of A. Then Laplace expanding on the qth row gets

$$\begin{aligned}\det(B) = \sum_j (-1)^{q+j} b_{qj} \det(B_{qj}) &= \sum_j (-1)^{q+j}(a_{qj} + k a_{pj})(\det(A_{qj}) \\ &= \det(A) + k\sum_j (-1)^{q+j} a_{pj} \det(A_{qj})\end{aligned}$$

the second term on the far right is k times the determinant of a matrix that we obtained from A by replacing its qth row by its pth row: the matrix has identical p and q rows. Therefore by the remark above, its determinant is 0. And so $\det(B) = \det(A)$

• So: how do we use this fact to compute the determinant of some matrix? Let's do an example: let

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 & 1 & 2 \\ 1 & -1 & 1 & 2 & 1 & 1 \\ 0 & 3 & 1 & 2 & 1 & 3 \\ 1 & 1 & 2 & 2 & 4 & 3 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & 5 & 1 & 3 & 4 \end{bmatrix}$$

– First, we eliminate the first column by adding -1 times the first row to rows 2, 4, and 5. This does not change the determinant, so we get

$$\begin{bmatrix} 1 & 1 & 2 & 2 & 1 & 2 \\ 0 & -2 & -1 & 0 & 0 & -1 \\ 0 & 3 & 1 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & -1 & -1 & -2 & 0 & -2 \\ 0 & 3 & 5 & 1 & 3 & 4 \end{bmatrix} ; m = 1$$

where m is the multiplier from the G-J process.

– Next we multiply the second row by 1/2, this means we need to multiply m by 2 (whatever we put on m will the inverse of the multiplier in the above claim). And we use this new second row to eliminate the leading numbers in rows 3, 5, and 6. Notice we don't worry about row 1, since we are only interested in making it upper triangular, and not RREF. Adding the second row to other rows does not change the determinant, so in the end, we get

$$\begin{bmatrix} 1 & 1 & 2 & 2 & 1 & 2 \\ 0 & -1 & -1/2 & 0 & 0 & -1/2 \\ 0 & 0 & -1/2 & 2 & 1 & 3/2 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & -1/2 & -2 & 0 & -3/2 \\ 0 & 0 & 7/2 & 1 & 3 & 5/2 \end{bmatrix} ; m = 2$$

– Next we use the third row to eliminate the leading numbers in rows 5 and 6 (not changing m).

$$\begin{bmatrix} 1 & 1 & 2 & 2 & 1 & 2 \\ 0 & -1 & -1/2 & 0 & 0 & -1/2 \\ 0 & 0 & -1/2 & 2 & 1 & 3/2 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & -4 & -1 & -3 \\ 0 & 0 & 0 & 15 & 10 & 13 \end{bmatrix} ; m = 2$$

– We need now to swap rows 4 and 5 (multiplying m by -1) and multiply row 4 by $1/4$ (multiplying m by 4) and eliminating the leading number in the 6th row (not changing m)

$$\begin{bmatrix} 1 & 1 & 2 & 2 & 1 & 2 \\ 0 & -1 & -1/2 & 0 & 0 & -1/2 \\ 0 & 0 & -1/2 & 2 & 1 & 3/2 \\ 0 & 0 & 0 & -1 & -1/4 & -3/4 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 25/4 & 7/4 \end{bmatrix} ; m = 2 \cdot (-1) \cdot 4 = -8$$

- Lastly we eliminate the leading number in the final row (not changing m)

$$\begin{bmatrix} 1 & 1 & 2 & 2 & 1 & 2 \\ 0 & -1 & -1/2 & 0 & 0 & -1/2 \\ 0 & 0 & -1/2 & 2 & 1 & 3/2 \\ 0 & 0 & 0 & -1 & -1/4 & -3/4 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1/3 \end{bmatrix} ; m = -8$$

- So we have that, in the end

$$\det(A) = m \cdot \det(\begin{bmatrix} 1 & 1 & 2 & 2 & 1 & 2 \\ 0 & -1 & -1/2 & 0 & 0 & -1/2 \\ 0 & 0 & -1/2 & 2 & 1 & 3/2 \\ 0 & 0 & 0 & -1 & -1/4 & -3/4 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1/3 \end{bmatrix})$$

so

$$\det(A) = -8 \cdot 1 \cdot -1 \cdot -\frac{1}{2} \cdot -1 \cdot 3 \cdot -\frac{1}{3} = -4$$

- Notice that since everything used in deriving the algorithm above is based on Laplace expansion, and Laplace expansion works for both columns and rows, we can also perform column operations with similar results: swapping columns gives a factor of (-1), and multiplying a column with a number k gives a factor of k, while adding a multiple of one column to another changes not the determinant.

- Here we collect some raw facts about determinants:

  - The determinant of a product is the product of the determinant: let A and B be both $N \times N$ matrices, then
  $$\det(AB) = \det(A)\det(B)$$
  - The determinants of similar matrices are the same: let $A = S^{-1}BS$, then
  $$\det(A) = \det(B)$$
  - The determinant of the inverse is the inverse of the determinant: if A is invertible (i.e. $\det(A) \neq 0$), then
  $$\det(A^{-1}) = \frac{1}{\det(A)} = (\det(A))^{-1}$$
  - The determinant of the transpose is the same: let A be a square matrix
  $$\det(A^T) = \det(A)$$

- One last useful fact about the determinant is that it extends to a linear transformation of vectors in the following way: fix $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{M-1}$ vectors in $\mathbb{R}^M$. Then the transformation $T : \mathbb{R}^M \to \mathbb{R}^M$ given by

$$T(\boldsymbol{x}) = \det(\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \ldots & \boldsymbol{v}_{M-1} & \boldsymbol{x}\end{bmatrix})$$

  is a linear transformation. Notice that since we can swap columns at the expense of multiplying by an over all (-1), we don't have to insert the $\boldsymbol{x}$ vector in the end: it could be in any position in the matrix.

APR 4, 2008

- Today we talk about geometric interpretation of the determinant.
- We begin by looking at the 2×2 case.
  - The scaling matrix $A = kI_2$, its determinant is $\det(A) = k^2$
  - The pure rotational matrix $A = \begin{bmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{bmatrix}$ has determinant $\det(A) = \cos^2\theta + \sin^2\theta = 1$.
  - The reflection matrix $A = \begin{bmatrix}2a^2-1 & 2ab \\ 2ab & 2b^2-1\end{bmatrix}$ with $a^2 + b^2 = 1$ has determinant

$$\det(A) = 4a^2b^2 - 2a^2 - 2b^2 + 1 - 4a^2b^2 = -1$$

- For arbitrary $M \times M$ scaling matrix $A = kI_M$, we have that $\det(A) = k^M$. The power M suggest that the determinant should have something to do with an M-dimensional volume.
- Let's look at A an orthogonal matrix: We know that $A^T A = I_M$, or, $A^T = A^{-1}$. Using the properties of determinants that we learned on Wednesday, we know that

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

  for an orthogonal matrix, so

$$\det(A)^2 = 1$$

  In other words, $\det(A) = \pm 1$. For the case $\det(A) = 1$ we will say that A is a *rotational matrix*, and its corresponding linear transformation a *rotation*.
- All this is to motivate 1 point: the determinant is a volume, geometrically speaking.
- The picture is much clearer if we look at a three dimensional case using Gram-Schmidt.
- How does the determinant of a matrix behave under Gram-Schmidt?

– Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_M$ be vectors in $\mathbb{R}^M$. And let $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_M$ be the unit vectors obtained from them using the G-S process. We want to related the determinants of the matrices A and B where

$$A = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 & \ldots & \boldsymbol{v}_M \end{bmatrix}, \quad B = \begin{bmatrix} \boldsymbol{u}_1 & \ldots & \boldsymbol{u}_M \end{bmatrix}$$

– (Remember that row operations and column operations behave the same with respect to the determinants. So we now express the G-S process using elementary row/column operations and see how the determinants relate.)

– We get $\boldsymbol{u}_1$ from $\boldsymbol{v}_1$ by dividing by its norm, so we have that

$$\det(A) = |\boldsymbol{v}_1| \det(\begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{v}_2 & \ldots & \boldsymbol{v}_M \end{bmatrix}$$

– Next, we obtain $\boldsymbol{u}_2$ by first projecting, and then dividing. Remember that

$$\boldsymbol{u}_2 = \boldsymbol{v}_2^{\perp} / |\boldsymbol{v}_2^{\perp}|$$

and

$$\boldsymbol{v}_2^{\perp} = \boldsymbol{v}_2 - (\boldsymbol{v}_2 \cdot \boldsymbol{u}_1)\boldsymbol{u}_1$$

The important part is that $\boldsymbol{v}_2^{\perp}$ is $\boldsymbol{v}_2$ *minus a scalar multiple of* $\boldsymbol{u}_1$, so

$$\det(\begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{v}_2 & \ldots & \boldsymbol{v}_M \end{bmatrix} = \det(\begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{v}_2^{\perp} & \boldsymbol{v}_3 & \ldots & \boldsymbol{v}_M \end{bmatrix}$$

And furthermore

$$\det(\begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{v}_2^{\perp} & \boldsymbol{v}_3 & \ldots & \boldsymbol{v}_M \end{bmatrix} = |\boldsymbol{v}_2^{\perp}| \det(\begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{v}_3 & \ldots & \boldsymbol{v}_M \end{bmatrix}$$

– For $\boldsymbol{u}_3$, the process is similar: first we build $\boldsymbol{v}_3^{\perp}$ by subtracting from $\boldsymbol{v}_3$ scalar multiples of $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. Next we divide $\boldsymbol{v}_3^{\perp}$ by its length. The first step does not change the determinant. The second step gives you the length of $\boldsymbol{v}_3^{\perp}$.

– So putting it all together:

$$\det(A) = |\boldsymbol{v}_1||\boldsymbol{v}_2^{\perp}||\boldsymbol{v}_3^{\perp}| \cdots |\boldsymbol{v}_M^{\perp}| \det(B)$$

– Now, if $\boldsymbol{v}$s were not linearly independent, A is not invertible, and $\det(A) = 0$; but when $\boldsymbol{v}$s are linearly independent, B is now a matrix with column vectors being orthonormal basis! This means that B is an orthogonal matrix. And so its determinant is either 1 or -1. So if we take the absolute value:

$$|\det(A)| = |\boldsymbol{v}_1||\boldsymbol{v}_2^{\perp}||\boldsymbol{v}_3^{\perp}| \cdots |\boldsymbol{v}_M^{\perp}|$$

- Let us now go back to the 2×2 case. (Draw picture with $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ and form the parallelogram associated to them.) The area of a parallelogram is base times height. Letting $\boldsymbol{v}_1$ be the base, we see that $\boldsymbol{v}_2^{\perp}$ gives you precisely the height. So $|\det(A)|$ is the area of the parallelogram.

- (When the two vectors are parallel, the "parallelogram" has zero area: $\det(A) = 0$).
- For the 3×3 case, (draw picture with 3D parallelepiped), the volume of such a figure is the area of the base times its height. Let the base be the face formed by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, its area is $|\boldsymbol{v}_1||\boldsymbol{v}_2^\perp|$. The height it the perpendicular distance to the plane containing the base, that is precisely $|\boldsymbol{v}_3^\perp|$.
- Definition: parallelepipeds in $\mathbb{R}^M$. A M-parallelepiped corresponding to M vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_M$ is the figure in $\mathbb{R}^M$ formed by the collection of all the vectors

$$v = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_M\boldsymbol{v}_M$$

  with the coefficients between 0 and 1.
- The M-volume of a M-parallelepiped is also given by base times height: the base being the $M-1$-volume of the $M-1$-parallelepiped formed by the vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_{M-1}$.
- So it is now easy to see that for M vectors in $\mathbb{R}^M$, the volume of its parallelepiped is $|\det(A)|$, where A is the matrix of those vectors.
- What about if we want to find the M-volume of an M-parallelepiped sitting in $\mathbb{R}^N$? We can't just go and take the determinant, since the matrix is no longer square. Motivated by our discussion of least squares, we know that the matrix A^TA is a square matrix that is very closely related to A: in fact, if the vectors of A are linearly independent, A^TA is invertible. Claim: the M-volume of an M-parallelepiped in $\mathbb{R}^N$ is

$$\sqrt{\det(A^TA)}$$

- I will not prove this fact in detail: but let me just remark that in light of the Gram Schmidt process, it is only really necessary to check that the formula is correct if the vectors of A are orthonormal. Geometrically all parallelepipeds can be obtained from a cube via two transformations: stretching the length of one side and shearing the figure. Geometrically, the former stretches the volume by the same factor and the latter does not change the volume. In terms of determinants, the former gives an application of the 1st elementary transformation and the latter the third elementary operation. For orthonormal vectors, A^TA is the identity.
- The determinant can also be understood as an expansion factor in a linear transformation. Let $T : \mathbb{R}^M \to \mathbb{R}^M$ be represented by the matrix A. First, let's consider the volume of the cube formed by the standard basis vectors $\boldsymbol{e}_1, \dots, \boldsymbol{e}_M$. Simple geometrical considerations shows that its volume is 1. Now, let us look at the volume of the parallelepiped formed by the images of the standard basis vectors $T(\boldsymbol{e}_1), \dots, T(\boldsymbol{e}_M)$. Those vectors are exactly the column vectors of A, and so the volume of the parallelepiped thereof is $|\det(A)|$.
- This example can be reworked into one comparing the "before" and "after" of arbitrary parallelepipeds. Let $\boldsymbol{w}_1, \dots, \boldsymbol{w}_M$ be vectors in $\mathbb{R}^M$ and let T be a linear transformation

given by the square matrix A. We can write B for the square matrix with column vectors $\boldsymbol{w}$. Then the "before" volume of the parallelepiped formed by the vectors $\boldsymbol{w}$ is $|\det(B)|$. What about the after? The volume is

$$|\det(\begin{bmatrix} T(\boldsymbol{w}_1) & \dots & T(\boldsymbol{w}_M) \end{bmatrix})|$$

But the matrix inside the determinant is simply the matrix AB! So the "after" volume is

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

Therefore we arrived at the fact that $\det(A)$ is a generic expansion factor for the volume of any parallelepiped (in fact, any geometrical figure) under the linear transformation given by A.

Homework for this week

6.1: 8, 16, 36, 44, 56

6.2: 4, 10, 12, 16, 30, 50

6.3: 1, 4, 13, 18