DETERMINANTS

MAX STEINBERG FOR OLGA RADKO MATH CIRCLE
ADVANCED 2

1. Warm Up

Before we can learn about determinants, we should go over some basics of linear algebra to ensure we are all on the same page. If you are already familiar with matrices, row reduction, and matrix multiplication, feel free to go on to the next section. We will only be working with *square matrices*, that is, matrices that have the same number of rows and columns.

Definition 1. A (real) **matrix** of size n-by-n is a collection of n^2 real numbers arranged into n rows/columns. For example, if $n = 3$, we may have a matrix like this:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

A n-by-n matrix represents a function from $\mathbb{R}^n \to \mathbb{R}^n$, given by *matrix multiplication*. For the matrix above, it is a function $\mathbb{R}^3 \to \mathbb{R}^3$, where

$$f_M(x, y, z) = M(x, y, z) = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} (x, y, z) = \begin{pmatrix} 1x + 2y + 3z \\ 4x + 5y + 6z \\ 7x + 8y + 9z \end{pmatrix}$$

Problem 1. Let M be a n-by-n matrix. Prove that the function associated to M (the function $f_M(x) = Mx$ on $\mathbb{R}^n$) is *linear*: that is, $f_M(x+y) = f_M(x) + f_M(y)$ and $f_M(rx) = rf_M(x)$, for $x, y \in \mathbb{R}^n$ and $r \in \mathbb{R}$.

Problem 2. Prove that if f, g are linear, then $f + g$, $f \circ g$ are linear.

It is a theorem (that we will not prove) that every linear function $f : \mathbb{R}^n \to \mathbb{R}^n$ is actually f_M for some n-by-n matrix M. This allows us to figure out how to multiply matrices. Let M, N be n-by-n matrices and f_M, f_N be the corresponding functions. As we just showed, $f_M \circ f_N$ is linear, so it should be represented by some matrix, which we will *define* to be MN. For any $x \in \mathbb{R}^n$, $f_M(f_N(x)) = f_M(Nx) = M(Nx)$, and we define $MN(x)$ to be $M(Nx)$.

Example 1. Let's calculate $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$ using this approach. Let $x = (x_1, x_2) \in \mathbb{R}^2$. By definition,

$$\begin{aligned} \left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} \right) (x_1, x_2) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \left(\begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} (x_1, x_2) \right) \\ &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1x_1 \\ 3x_2 \end{pmatrix} \\ &= \begin{pmatrix} -1x_1 + 3x_2 \\ 3x_2 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 3 \\ 0 & 3 \end{pmatrix} (x_1, x_2) \end{aligned}$$

Thus, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 3 \\ 0 & 3 \end{pmatrix}$.

Problem 3. Calculate $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$ using this approach.

Let's prove that this definition of matrix multiplication is exactly the same as the one you may know if you are familiar with linear algebra.

Problem 4. Show that if M, N are n-by-n matrices, then the entry in the i-th row and j-th column of MN, denoted $(MN)_{ij}$, can be found by the following formula:

$$(MN)_{ij} = \sum_{k=1}^{n} M_{ik}N_{kj}, 1 \leq i, j \leq n$$

One of the most common uses of matrices is solving systems of linear equations. Let's say we want to solve

$$\begin{cases} 2x + 3y + 4z & = 4 \\ x - y + 2z & = 0 \\ -x + 4y + z & = -7 \end{cases}$$

We can write this as $\begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 2 \\ -1 & 4 & 1 \end{pmatrix}(x, y, z) = (4, 0, -7)$. If we write $M = \begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 2 \\ -1 & 4 & 1 \end{pmatrix}$, then we want to find $(x, y, z) = M^{-1}(4, 0, -7)$, if such an inverse function exists. Let's figure out how we can figure out what M^{-1} is. We will use the process of *row reduction.* Think about how you would normally solve an equation like this. You would probably start by isolating one variable, substituting into the other equations, and eventually finding numbers for all the variables. Let's figure out how to do that with matrices.

$$\begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 2 \\ -1 & 4 & 1 \end{pmatrix}$$

First, we want the first number in the first row to be 1, so we multiply the first row through by $\frac{1}{2}$.

$$\begin{pmatrix} 1 & \frac{3}{2} & 2 \\ 1 & -1 & 2 \\ -1 & 4 & 1 \end{pmatrix}$$

Then, we subtract the first row the second row, and add it to the third row, so that all rows other than the first one have a 0 in the first column.

$$\begin{pmatrix} 1 & \frac{3}{2} & 2 \\ 0 & -\frac{5}{2} & 0 \\ -1 & 4 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & \frac{3}{2} & 2 \\ 0 & -\frac{5}{2} & 0 \\ 0 & \frac{11}{2} & 3 \end{pmatrix}$$

Now that the first row starts with a 1 every other row starts with a 0, we move to the second row. We want the second number in the second row to be 1, so we multiply through by $-\frac{2}{5}$.

$$\begin{pmatrix} 1 & \frac{3}{2} & 2 \\ 0 & 1 & 0 \\ 0 & \frac{11}{2} & 3 \end{pmatrix}$$

Then, we subtract $\frac{3}{2}$ times the second row from the first, and $\frac{11}{2}$ times the second row from the third.

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Now, finally, we move to the third row. We divide through by 3.

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and then subtract 2 times the third row from the first.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Now, this matrix is called the *identity matrix* because, if we denote it I, then $I(x, y, z) = (x, y, z)$. If M had an inverse, we would want $M^{-1} \cdot M = I$, because we want M^{-1} to "undo" M. We've shown that we can manipulate the rows of M to get I, so M must have an inverse! For each step we did, we can associate a *elementary matrix*: for example, if we add row 1 to row 2, our elementary matrix is

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

because the product EM is the matrix M except with the first row added to the second.

Problem 5. Calculate EM in order to verify that E does what we claim it does.

Problem 6. Find the elementary matrices for each operation we did so far.

Now that we have our elementary matrices, we can denote them $E_1, E_2, \ldots, E_i$, where E_1 is the first one we applied (in this case multiplying row 1 by $\frac{1}{2}$). Then, we see that

$$E_i E_{i-1} \ldots E_2 E_1 M = I$$

as we just showed, so surely $M^{-1} = E_i E_{i-1} \ldots E_2 E_1$!

Problem 7. Explicitly write out M^{-1} is and verify that it is an inverse to M (ie. multiply $M^{-1} \cdot M$ and see that it is in fact I).

This process is known as *row reduction* and lets us figure out how to build up matrices in terms of simple building blocks, our elementary matrices.

2. Deriving the Determinant

For a permutation $\sigma \in S_n$ (the symmetric group on n objects), recall that we can decompose σ as a product of transpositions. Then, we can define the *sign* of σ, denoted $\operatorname{sgn}(\sigma)$, as $(-1)^m$, where m is the number of transpositions in the decomposition of σ.

Problem 8. Calculate the following expressions:
1. $\operatorname{sgn}((123))$
2. $\operatorname{sgn}((12)(34)(56))$
3. $\operatorname{sgn}((12)(12))$

Recall that a *cyclic* permutation (also known as a *cycle*), written $(a_1 a_2 \dots a_i)$, is the permutation given by $\sigma(a_1) = a_2, \sigma(a_2) = a_3, \dots, \sigma(a_i) = a_1$. For any permutation $\pi \in S_n$, we can decompose it as a product of *disjoint cycles.* Two cycles $(a_1 a_2 \dots a_i)$ and $(b_1 b_2 \dots b_j)$ are *disjoint* if they share no elements. Formally, these cycles are disjoint if $a_k \neq b_\ell$ for every $1 \leq k \leq i, 1 \leq \ell \leq j$. When we learned about the symmetric group and permutations, we learned that every permutation can be decomposed into a product of disjoint cycles.

Problem 9. Show that the sign of a cycle $(a_1 a_2 \dots a_i)$ is equal to $(-1)^{i-1}$.

Problem 10. Show that $\operatorname{sgn}(\sigma_1 \sigma_2) = \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2)$.

Now, using our knowledge of disjoint cycle decomposition, we can now calculate the sign of any permutation.

Example 2. Let $\sigma = [456123] \in S_6$. Remember that when we use square brackets, we mean that $\sigma(1) = 4, \sigma(2) = 5, \dots$. We can rewrite σ in cycle notation: $\sigma = (14)(25)(36)$. Then,

$$\begin{aligned}
\operatorname{sgn}(\sigma) &= \operatorname{sgn}((14)(25)(36)) \\
&= \operatorname{sgn}((14))\operatorname{sgn}((25))\operatorname{sgn}((36)) \\
&= (-1)^1 \cdot (-1)^1 \cdot (-1)^1 \\
&= -1
\end{aligned}$$

Now, we can finally define the determinant.

Definition 2. Let M be a square matrix of size n-by-n. We define the *determinant* of M to be

$$\det(M) := \sum_{\sigma \in S_n} (\operatorname{sgn}(\sigma)) M_{1,\sigma(1)} M_{2,\sigma(2)} \dots M_{n,\sigma(n)}$$

Example 3. Let's look at the case when $n = 2$. We have $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and we want to find $\det(M)$. We know that $S_2 = \{e, (12)\}$, so we have

$$\begin{aligned}
\det(M) &= \sum_{\sigma \in S_2} (\operatorname{sgn}(\sigma)) M_{1,\sigma(1)} M_{2,\sigma(2)} \\
&= (\operatorname{sgn}(e)) M_{1,e(1)} M_{2,e(2)} + (\operatorname{sgn}((12))) M_{1,(12)(1)} M_{2,(12)(2)} \\
&= (1) M_{1,1} M_{2,2} + (-1) M_{1,2} M_{2,1} \\
&= ad - bc
\end{aligned}$$

You may recognise this formula as the determinant of a 2-by-2 matrix!

If you are more familiar with linear algebra, you may also know that

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

Problem 11. Prove this formula using our definition of the determinant.

3. Determinant Properties

For this entire section, remembering row reduction and elementary matrices will be very helpful.

Problem 12. Prove that for any two n-by-n matrices M and N, $\det(MN) = \det(M)\det(N)$.

Problem 13. Find an example of two matrices A and B where $\det(A+B) \neq \det(A) + \det(B)$.

Problem 14. Let M be an n-by-n matrix. Let $r \in \mathbb{R}$ be any real and let N be an n-by-n matrix where we multiply the any row of M by r. For example, if $n = 2$ and $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then N could be $\begin{pmatrix} a & b \\ rc & rd \end{pmatrix}$ or $\begin{pmatrix} ra & rb \\ c & d \end{pmatrix}$.

(1) Prove that $\det(N) = c\det(M)$.
(2) Prove that $\det(cM) = c^n \det(M)$.

Problem 15. Let M be an n-by-n matrix. Let $r \in \mathbb{R}$ be any real and let N be an n-by-n matrix where we add r times any row to any other row. For example, if $n = 2$ and $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then N could be $\begin{pmatrix} a+rc & b+rd \\ c & d \end{pmatrix}$ or $\begin{pmatrix} a & b \\ c+ra & d+rb \end{pmatrix}$. Prove that $\det(N) = \det(M)$.

Problem 16. Let M be an n-by-n matrix and N be an n-by-n matrix where we swap any two rows of M. Prove that $\det(N) = -\det(M)$.

Problem 17. Let M be an n-by-n matrix and M^t the transpose of M, defined by $(M^t)_{ij} = M_{ji}$. For example, if $n = 2$ and $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then M^t is $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Prove that $\det(M^t) = \det(M)$.

Note that after proving this statement, we will also see that, for example, we can swap columns or rows without changing the determinant, not just rows.

Problem 18. Let $\sigma \in S_n$ be a permutation and let M_σ be the matrix obtained by taking the columns of the identity matrix and permuting the columns according to σ. Prove $\operatorname{sgn}(\sigma) = \det(M_\sigma)$.
Hint: what is $(M_\sigma)_{ij}$*?*

4. Ultra-Hard Mega Difficult Very Challenging Challenge Problems

Problem 19. Prove that $\det(M \otimes N) = \det(M)\det(N) = \det(N \otimes M)$, where $\otimes$ represents the tensor product of linear transformations (which on the level of matrices is the Kronecker product).

Problem 20. Let F be a field of characteristic 0 and V a vector space over F. Let $T(V)$ be the *tensor algebra* of V. To define it, we start by defining $T^k(V) := V^{\otimes k} = \underbrace{V \otimes \cdots \otimes V}_{k \text{ times}}$. There is a canonical isomorphism $T^k(V) \otimes T^\ell(V) \to T^{k+\ell}(V)$ so we can define $T(V) := \bigoplus_{k=0}^{\infty} T^k(V)$ with the multiplication extended linearly to $T(V) \otimes T(V) \to T(V)$ by the above isomorphism. This makes $T(V)$ is a graded algebra over F. Define $\bigwedge(V) := T(V)/\langle v \otimes v | v \in V\rangle$. This is the *exterior algebra* over V, and it inherits a graded algebra structure from $T(V)$.

(1) Prove that if V is finite dimensional of dimension $n = \dim(V)$, then $\bigwedge^k(V)$ (defined as the k-th graded component of $\bigwedge(V)$) is trivial when $k > n$ or $k < 0$.
(2) Prove that $\dim\left(\bigwedge^k(V)\right) = \binom{n}{k}$ when $0 \le k \le n$.
(3) Recall that for any finite-dimensional vector space V over a field, $\dim(V) = \dim(V^*)$. Prove that $\det \in (\bigwedge^n V)^*$, and as an immediate corollary, that the unique basis of $(\bigwedge^n V)^*$ that takes value 1 on the identity is $\{\det\}$.

This problem lets us understand the determinant purely algebraically. We proved that the determinant is the unique up to scalar "alternating n-linear form" on V. We can also understand it analytically:

Problem 21. Let $F = \mathbb{R}$ and consider V a finite-dimensional real vector space with dimension $n = \dim(V)$. We can define a topology on V by pulling back along any isomorphism $V \cong \mathbb{R}^n$. (Exercise: prove that these topologies coincide.) We can do something similar for the real vector space $M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$, the vector space of n-by-n real matrices. Let $\det : M_n(\mathbb{R}) \to \mathbb{R}$ be the determinant and $\mathrm{tr} : M_n(\mathbb{R}) \to \mathbb{R}$ be the trace (sum of the diagonal entries).

(1) Prove that $\det, \mathrm{tr}$ are continuous.
(2) Prove that for any $M \in M_n(\mathbb{R})$,

$$\lim_{t \to 0} \det(I + tM) = \mathrm{tr}(M).$$

This second part is related to the fact that $\mathfrak{sl}_n(\mathbb{R}) = \{M \in M_n(\mathbb{R}) | \,\mathrm{tr}(M) = 0\}$.