A QUICK INTRODUCTION TO MATRICES AND DETERMINANTS

ADDITIONAL MULTIVARIABLE CALCULUS MATERIAL: HANDOUT 1

Please review material about vectors, dot and cross products and local optimization of functions of several variables from the textbook in its entirety before attempting to read this material. We have already encountered matrices and determinants when computing *cross products* in $\mathbb{R}^3$ and more recently as a part of the *discriminant* formula when we locally optimized functions of several variables. In this handout we will cover matrices in a somewhat more general sense. Please note that matrices have a *huge* and *complicated* theory which is covered in a branch of mathematics called Linear Algebra. In this handout we will only cover a very small portion of this wonderful theory.

◇ ——— ◇

1. The Basics

We have already seen that a matrix is a *rectangular grid* of objects. Typically, these objects are numbers, but (as we have already seen), there is no restriction on using functions instead of numbers, or even other matrices. Here is an example of a matrix of numbers:

$$M = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

We say that M has **two rows** and **three columns**, or that A *is a* 2×3 *matrix*. The *rows* of M can be thought of as vectors: there are only two of them, namely $R_1 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ and $R_2 = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$; these are vectors in $\mathbb{R}^3$. On the other hand there are three *column* vectors, namely $C_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, $C_2 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ and $C_3 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$. Each of these is a vector in $\mathbb{R}^2$.

In order to refer to the actual components of M, we index them by row and column number. For instance, M_{12} is the number in the first row and second column of M, namely 2. Similarly, $M_{21} = 4$ and so on.

So we can either write M as a vertical vector of its two row vectors, like so: $M = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$ or as a horizontal vector of its three column vectors: $M = \begin{bmatrix} C_1 & C_2 & C_3 \end{bmatrix}$. It is important to know that even aside from our experiences with cross products, *matrices are not new*. We have seen matrices before and manipulated them regularly. For example, every number is a 1×1 matrix. A less silly example: each vector in $\vec{v}$ in $\mathbb{R}^k$ can be thought of as a $1 \times k$ matrix $\begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix}$ or as a $k \times 1$ matrix $\begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}$. In the past, we have not emphasized the difference between "row vectors" and "column vectors", but we will be careful here.

Fact 1.1. Any matrix M is a rectangular grid of size $m \times n$. Moreover,

(1) m is the *number of rows* of M.
(2) n is the *number of columns* of M.
(3) M has m row vectors. Each row vector of M is n dimensional.
(4) M has n column vectors. Each column vector of M is m dimensional.
(5) M may be seen as a vertical vector of its m row vectors or a horizontal vector of its n column vectors.
(6) Every number is a 1×1 matrix.
(7) Each vector of dimension k may be seen as a $1 \times k$ (row) or a $k \times 1$ (column) matrix.

And here are some *extremely simple* exercises to test these basic concepts.

1.1. **Exercises.** Throughout these exercises, set $M = \begin{bmatrix} 3 & 5 & -9 & 1 \\ 1 & 7 & 9 & 0 \\ 2 & 3 & 5 & 6 \end{bmatrix}$.

1. How many rows and columns does M have?

2. What are the row vectors of M?

3. What are the column vectors of M?

4. What are M_{32}, M_{11} and M_{23}?

5. Give an example of a 4×2 matrix N. Do problems 2 and 3 for N instead of M.

6. Rewrite $\begin{bmatrix} 1 & 2 & 4 & 3 \end{bmatrix}$ as a column vector.

◇ ◇

2. Matrix Algebra

There are three important matrix operations that we will consider here. In increasing order of difficulty, they are: *scaling*, *addition* and *multiplication*.

Scaling of matrices is very similar to the scaling of vectors. Given a *number* r and a matrix M, the matrix rM has the same size as M but each individual component in the matrix is multiplied by r. For example, if $M = \left[\begin{smallmatrix}1&2\\3&4\end{smallmatrix}\right]$ then $2M = \left[\begin{smallmatrix}2&4\\6&8\end{smallmatrix}\right]$. Thus the operation of scaling requires as *input* a number and a matrix, and its *output* is a matrix of the same size as the input matrix.

Addition of matrices is also very similar to adding vectors; everything happens componentwise. Obviously this requires the sizes of the two matrices being added to be the same. So for example, if we have $M = \left[\begin{smallmatrix}1&2&3\\4&5&6\\7&8&9\end{smallmatrix}\right]$, $N = \left[\begin{smallmatrix}5&1&-3\\-7&2&0\\3&3&1\end{smallmatrix}\right]$ and $P = \left[\begin{smallmatrix}-4&5&1\\3&6&-9\end{smallmatrix}\right]$ then we can add M and N to each other, but neither M nor N can be added to P because the number of columns don't match. In any case, adding M and N gives a new matrix $M + N$ which has the same size as M and N. Here is that matrix:

$$M + N = \begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix} + \begin{bmatrix}5&1&-3\\-7&2&0\\3&3&1\end{bmatrix} = \begin{bmatrix}6&3&0\\-3&7&6\\10&11&10\end{bmatrix}$$

As you can see, each component of the matrix $M + N$ is the sum of the corresponding components from M and N. Thus, the operation of adding matrices takes as *input* two matrices of the same size $m \times n$, and the *output* is another matrix of size $m \times n$.

Matrix multiplication is much trickier than scaling or adding, and we will deal with it later. First, here are some more exercises on adding and scaling.

2.1. **Exercises.** Let $M = \left[\begin{smallmatrix}1&-2&1&6\\3&-4&5&2\\-5&1&2&5\end{smallmatrix}\right]$ for these exercises.

1. Compute $2M$ and $0M$.

2. Which of the following matrices are we allowed to add to M?

(1) $N = \left[\begin{smallmatrix}5&1&0&-3\\-7&4&2&0\end{smallmatrix}\right]$

(2) $P = \left[\begin{smallmatrix}-4&5\\3&6\\3&8\end{smallmatrix}\right]$

(3) $Q = \left[\begin{smallmatrix}2&1&8&-4\\10&-1&4&6\\-3&-1&0&9\end{smallmatrix}\right]$

3. Compute the sum of M with each matrix from Exercise 2 that can be added to M.

4. Compute $M - 2Q$ and $2M + Q$. Remember that you have already computed $2M$ in Exercise 1 and that Q is defined in Problem 2.

◇ ◇

3. Matrix Multiplication

Computationally, multiplication is the hardest of the matrix operations that we will consider. Unfortunately, it is also the most useful.

We represent the product of two matrices A and B as AB. There are a few rules first: *the number of columns of A must be equal to the number of rows of B*. We will define the matrix product AB using *dot products*. So if A has size $m \times n$ then we require B to have size $n \times p$. Note that the n must be common to the sizes of both A and B. Write $A = \begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix}$ as a collection of row vectors. Also write $B = \begin{bmatrix} C_1 & \dots & C_p \end{bmatrix}$ as a collection of column vectors. Note that all vectors here, $R_1, \dots, R_m, C_1, \dots C_n$ are n-dimensional. Now, AB is the $m \times p$ matrix defined by

$$AB = \begin{bmatrix} R_1 \cdot C_1 & \dots & R_1 \cdot C_p \\ \vdots & \ddots & \vdots \\ R_m \cdot C_1 & \dots & R_m \cdot C_p \end{bmatrix}$$

Here each component is a dot product. For instance, AB_{11} is the dot product of R_1 with C_1. A quick reminder about the size of the product matrix:

$$\boxed{(m \times n)\text{-matrix } \textbf{times } (n \times p)\text{-matrix} = (m \times p)\text{-matrix}}$$

3.1. **A Real Example.**

We will multiply the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ with the matrix $B = \begin{bmatrix} 2 & 1 \\ 3 & 4 \\ -1 & 3 \end{bmatrix}$. First check that this is possible: the size of A is 2×3 and the size of B is 3×2. As required by the *rule*, the number of columns of A equals the number of rows of B since both quantities equal 3. We also know that the size of AB will be 2×2. Let's try to compute this matrix: we will find values of a through d in the following equation.

$$AB = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & 4 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Let's use the definition above. First, the decomposition of A into row vectors gives us $R_1 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ and of course $R_2 = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$. Similarly, the decomposition of B into column vectors yields $C_1 = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$ and

$C_2 = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$. Now $a = AB_{11} = R_1 \cdot C_1$. This dot product equals $(1 \times 2 + 2 \times 3 + 3 \times -1) = 5$. Similarly, $b = AB_{12}$ is the dot product of $R_1 \cdot C_2 = 18$. Similarly, $c = AB_{21} = R_2 \cdot C_1 = 17$ and $d = R_2 \cdot C_2 = 6$. So,

$$AB = \begin{bmatrix} 5 & 18 \\ 17 & 42 \end{bmatrix}$$

Note that computing BA is a *very different* proposition. Again we have to check that the product is even possible. Since B has size 3×2 and A has size 2×3, we know that BA will have size 3×3. BA does not even have the same *size* as AB, so clearly there is no hope of getting AB to be the same matrix as BA. Applying the same principles as above, we see that

$$BA = \begin{bmatrix} 6 & 9 & 12 \\ 11 & 26 & 31 \\ 11 & 13 & 15 \end{bmatrix}$$

As a sample computation, BA_{11} is the dot product of $\begin{bmatrix} 2 & 1 \end{bmatrix}$ with $\begin{bmatrix} 1 & 4 \end{bmatrix}$ which equals 6.

3.2. **Matrices as Linear Functions from $\mathbb{R}^m$ to $\mathbb{R}^n$.** Let's pick a matrix like our old friend $M = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$. The size is of course 2×3, so if I wanted to multiply this matrix on the left with a vector $\vec{v}$ on the right, the vector would have to have size 3×1. The 3 comes from wanting to be able to multiply by M, so the number of rows of our vector must equal the number of columns of our matrix M, i.e., 3. The 1 comes from the fact that, well, $\vec{v}$ is a vector after all, so it can't have more than one row *and* more than one column! So pick some vector $\vec{v} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ for example. And now let's multiply; the answer should be a 2×1 matrix, like so:

$$M\vec{v} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 - 2 + 0 \\ 4 - 5 + 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

But why be so specific? Let's multiply M by a completely arbitrary vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in $\mathbb{R}^3$. We get $M \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+3z \\ 4x+5y+6z \end{bmatrix}$. So we can define a function f from $\mathbb{R}^3$ to $\mathbb{R}^2$ which describes "multiplication by M":

$$f(x, y) = (x + 2y + 3z, 4x + 5y + 6z)$$

This is a *complete description* of the matrix M as a function from $\mathbb{R}^3$ to $\mathbb{R}^2$! This also shows that any such matrix M is a *linear function* because no powers higher than 1 will ever show up in any component of the output; there are no x^2 terms, for instance. So a matrix of size

2×3 takes as input vectors in $\mathbb{R}^3$ and the output is a vector in $\mathbb{R}^2$. This rule generalizes to any matrix of any size!

An $m \times n$ matrix takes as input an n-vector and outputs an m-vector

Here are some exercises involving matrix multiplication. I strongly urge you to do these exercises, because in addition to testing the concepts defined above they will provide intuition about distributive laws of matrix multiplication, identity matrices and so on.

3.3. Exercises.

1. For each pair of matrices A and B below, find both AB and BA whenever possible. If multiplication is not possible, explain why not.

(1) $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.
(2) $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 1 \end{bmatrix}$.
(3) $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$.
(4) $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$.

2. Compute AB and BA when $A = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 1 & 4 \\ -2 & 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & 1 & 0 \end{bmatrix}$. Confirm that even when both AB and BA are defined, they need not be equal to each other.

3. For the matrices A and B in Exercise 2 and $C = \begin{bmatrix} 0 & 3 & 0 \\ 1 & 0 & 2 \\ -1 & 1 & 0 \end{bmatrix}$, compute $(A + 2B)(C)$. Then compute $AC + 2(BC)$. Confirm that these two matrices are equal. *Multiplication distributes across addition.*

4. Given A and B from Exercise 2 and C from Exercise 3, confirm that $(AB)C = A(BC)$.

5. Let $M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and let $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. The matrix I_2 is called the *identity matrix* of size 2×2. I_2 contains 1's along the *diagonal* and 0's elsewhere. Check that $MI_2 = M$ and $I_2 M = M$.

6. What do you expect the identity matrix I_3 of size 3×3 to look like based on your experience in Exercise 5? Confirm that your guess is correct by checking that $NI_3 = N$ and $I_3 N = N$ for the matrix $N = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

7. Given $M = \begin{bmatrix} 3 & 2 \\ 1 & -1 \\ 2 & 5 \end{bmatrix}$ what is the input and output dimension for vectors that can be multiplied to this matrix on the right?

8. Write a function from $\mathbb{R}^2$ to $\mathbb{R}^3$ which describes the output of M from Exercise 7 when an input vector $\vec{v} = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ is multiplied to M on the right.

◇ ——— ◇

4. Determinants and Volumes

We have already encounted determinants when computing *cross products*. Determinants are defined only for *square* matrices, i.e., matrices for which the number of rows equals the number of columns. The determinant takes as *input* a square matrix and produces as *output* a single number. Some examples of such matrices are

(1) 1×1 matrices, which are just numbers. The determinant of a number k is just k itself.
(2) 2×2 matrices, which look like $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. The determinant $\det \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ equals $ad - bc$, as we have already seen when computing the discriminant, for instance.
(3) 3×3 matrices, which look like $\left[\begin{smallmatrix} a & b & c \\ d & e & f \\ g & h & i \end{smallmatrix}\right]$. We can reduce the computation of this determinant to computing *three* determinants of 2×2 matrices in the following way:

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

Beware the minus sign in the middle! The 2×2 matrix whose determinant we are multiplying by a in the first term is just the matrix that you would get by *blocking out* the row and column containing a. As you can see, there is a similar situation with the second term, etc.

Let's (gulp!) compute the determinant of a "random" 3×3 matrix. Here it is: $N = \left[\begin{smallmatrix} 1 & 2 & -1 \\ 0 & 2 & 1 \\ 4 & 2 & -3 \end{smallmatrix}\right]$. Applying our formula, we see that

$$\det N = 1 \cdot \det \left[\begin{smallmatrix} 2 & 1 \\ 2 & -3 \end{smallmatrix}\right] - 2 \cdot \det \left[\begin{smallmatrix} 0 & 1 \\ 4 & -3 \end{smallmatrix}\right] + (-1) \cdot \det \left[\begin{smallmatrix} 0 & 2 \\ 4 & 2 \end{smallmatrix}\right]$$

Now let's compute each of the three 2×2 determinants on the right side. This gives:

$$\det N = 1 \cdot (-6 - 2) - 2 \cdot (0 - 4) + (-1) \cdot (0 - 8)$$

We are very close to the answer now. Carry out the simple algebra, and get

$$\det N = -8 + 8 + 8 = 8$$

If we are being considerably more masochistic, we can *try* to compute the determinant of a simple 4×4 matrix. Here is an example

$$M = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

Here is the first step of the computation:

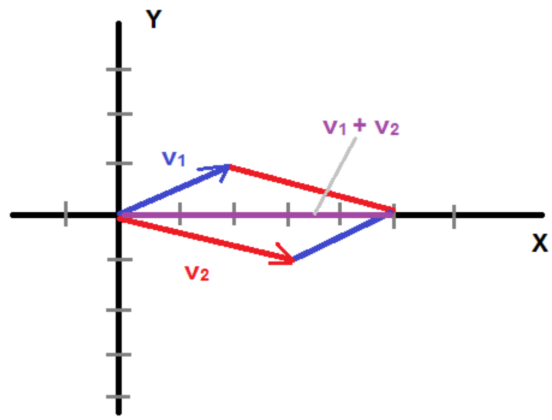
$$\det M = 1 \cdot \det \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix} - 1 \cdot \det \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} + 0 \cdot \det \begin{bmatrix} \text{who} \\ \text{cares,} \\ \text{really?} \end{bmatrix} - 1 \cdot \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Since we all *know* how to compute these 3×3 determinants, I will stop here, but you can see that this gets **extremely messy**. We will restrict our attention to 3×3 as a worst case. In linear algebra you will learn tricks to compute determinants of larger matrices much more efficiently via row reduction. We will not worry about that here.

4.1. **Relation to Areas and Volumes.** Now that we know how to compute determinants, we might ask: what is a determinant good for? The answer is straightforward to state, but rather intricate to prove. Let's say we have n vectors in $\mathbb{R}^n$. As an example, we can take two vectors in $\mathbb{R}^2$: let's pick $v_1 = \left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$ and $v_2 = \left[\begin{smallmatrix} 3 \\ -1 \end{smallmatrix}\right]$. We know that these two vectors represent a *parallelogram* with vertices $(0,0)$, $(2,1)$, $(3,-1)$ and $(5,0)^*$. We already know that the *area* of this parallelogram is the magnitude of the cross product between the vectors describing the edges: $(2,1,0)$ and $(3,-1,0)$. We have tacked on a z coordinate of 0 so that these vectors sit in $\mathbb{R}^3$ and can be crossed. Now the cross product equals $(2,1,0) \times (3,-1,0) = (-2-3)\vec{k} = -5\vec{k}$ which has a magnitude of 5.

---

*The last vertex is the sum of the two non-zero vertices by the *parallelogram law*. Remember?



Now let's compute the determinant of the vectors $\vec{v_1} = \left[\begin{smallmatrix}2\\1\end{smallmatrix}\right]$ and $\vec{v_2} = \left[\begin{smallmatrix}3,-1\end{smallmatrix}\right]$ shoved into a matrix in the obvious way. This gives us the matrix $M = \left[\begin{smallmatrix}2 & 3\\1 & -1\end{smallmatrix}\right]$. We can see immediately that the determinant of M is $\det M = -2-3 = -5$. Up to a minus sign, this is the area of the parallelogram that we have computed before!

If we have a box in $\mathbb{R}^3$ with parallel edges determined by three vectors $\vec{u_1}$, $\vec{u_2}$ and $\vec{u_3}$, then the determinant of the matrix $N = \left[\begin{smallmatrix}u_1 & u_2 & u_3\end{smallmatrix}\right]$ gives us the 3D volume of that box. We already knew this from learning about the *scalar triple product*. But if you take four vectors in $\mathbb{R}_4$, what is the 4D volume of the box bounded by these vectors? Cross products will not rescue us this time, but the determinant formula for volume will! Here is the take-away lesson about determinants and volumes: up to possibly a minus sign,

$$\boxed{\text{volume of box with edge vectors } \vec{v_1}, \cdots, \vec{v_n} = \det \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}}$$

And finally, here are some Exercises. As usual, some notions are introduced here so you should definitely try to solve them.

4.2. **Exercises.** Let $A = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 2 \\ 1 & 3 \end{bmatrix}$.

1. Compute $\det A$ and $\det B$.

2. Compute $\det AB$ and $\det BA$. It is a wonderful fact that $\det AB = \det A \cdot \det B = \det BA$, even though AB may not be equal to BA!

3. What is the 3D volume of the box in $\mathbb{R}^3$ with edge vectors $\vec{v_1} = \left[\begin{smallmatrix}1\\2\\1\end{smallmatrix}\right]$, $\vec{v_2} = \left[\begin{smallmatrix}3\\0\\3\end{smallmatrix}\right]$ and $\vec{v_3} = \left[\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right]$?