- M9227 - Week 1 Homework -

1 Problems

We don't expect you to do all of these problems since they vary widely in difficulty. We've left some keywords to look things up online, but believe me, you'll get a lot less of the thrill if you look everything up. Enjoy!

Problem 1. (Linear Maps) Which of the following operations are linear?

(a) $x \mapsto 2x$

(b) $x \mapsto x^2$

(c) $(x, y) \mapsto (y, x)$

(d) $(x, y) \mapsto xy$

(e) $(x, y) \mapsto x + y$

(f) $(x, y) \mapsto (x + y, y)$

Problem 2. (Multiplication.) Compute the following:

(a)

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

(b)

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^n$$

(c)

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

This is a nice example of the non-commutativity of matrices.

Problem 3. (Bases I) If I have a vector (a, b), I can express this as $a \cdot (1, 0) + b \cdot (0, 1)$. Generalizing, any vector can be expressed as the sum $\sum_i a_i e_i$ where a_i are numbers and e_i are the unit vectors with 1 in the ith spot and 0 elsewhere.

(a) Show that any 2-vector (a, b) can be expressed as a combination of $(1, -1)$ and $(1, 1)$.

(b) Show that any 3-vector can be expressed as a combination of $(1, -1, 1)$, $(0, 1, 1)$ and $(0, 0, 2)$. Show that this isn't true for the vectors $(1, -1, 1)$, $(0, 1, 1)$ and $(1, 0, 2)$. What's the difference between the first set of vectors and the second set of vectors? (Hint: try to express the third vector in terms of the first two).

(c) Now that you've done that stuff, it's easy to check that every 2-vector (a, b) can be written as a combination of $(1, 1), (1, -1)$ and $(1, 2)$. Do I need all of these vectors to make (a, b)? Can I make any (a, b) with just one vector? Can I make it with just two?

Problem 4. (Bases II) A **basis** is a set of vectors that can be used to express any other vector and which is *minimal*, in the sense that if you take away any of the vectors in the set then there are some vectors that you cannot make anymore. You've played with them a bit in the question above.

(a) How many vectors need to be in a basis for n-dimensional space, $\mathbb{R}^n$? That is, what is the fewest vectors needed to make any vector $(v_1, \ldots, v_n)$?

(b) If I have a basis $b_1, \ldots, b_n$ in n-space and I express a vector $v = (v_1, \ldots, v_n)$ in terms of that basis, I see that $v = \sum_i v_i' b_i$ where v_i' and v_i are numbers that are *n*ot necessarily the same. So I can write any vector v as a new list, i.e the list $(v_1', \ldots, v_n')$. This is called a **change of basis**.

(c) Given a vector (a, b), compute its new vector (a', b') in terms of the basis $(1, 1), (1, -1)$.

(d) Notice that changing basis is linear, so we can express it as a matrix, $(a', b') = M(a, b)$. This M is called a **change of basis matrix**. Compute the change of basis matrix for the new basis $(1, 1), (1, -1)$.

(e) We can also reverse the change of basis by going back to our original basis. This takes the vector $(a', b') = M(a, b)$ and makes it into (a, b) again. This can also be expressed as a matrix, namely as the inverse of M, M^{-1}. Find M^{-1} for the case given above.

Problem 5. (Bases III)

(a) But what happens to a matrix under a basis change? Well, suppose we have a vector v and a matrix A. If we change basis with a change of basis matrix M, then v goes to $v' = Mv$ and Av goes to $(Av)' = MAv$, right? So when we change basis, A should go to a new matrix A' so that $A'v' = (Av)'$. This way, the change of basis preserves the way that A and v interact after the change of basis.

(b) Show that $A' = MAM^{-1}$ makes the equation $A'v' = (Av)'$ work.

(c) So a linear map/matrix A changes to MAM^{-1} under a change of basis.

(d) Let v, A and M be defined as so:

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

Compute M^{-1}, then compute $A' = MAM^{-1}$ and $v' = Mv$. Finally show that $A'v' = MAv$ by explicit computation. Remember that the inverse is the matrix satisfying $MM^{-1} = Id$, where:

$$Id = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

. Id is called the identity matrix, and it sends any vector to itself.

Problem 6. (Abstract algebra.) Let X, Y be a 2×2 matrices. Prove that if $XY = 0$ then there is some power of YX such that $YX = 0$. (We say that a matrix is equal to zero when all of its entries are zero.) Is our assumption on the size of the matrices necessary?

Suppose $XY = YX = I = XZ = ZX$. Prove that $Y = Z$.

Problem 7. (Lie groups preserving a bilinear form and the orthogonal group.) Let I be the matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The transpose of a matrix

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

is

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

In other words, we flip the matrix across the diagonal.

What are all the 2×2 matrices A with the property that $A^t I A = I$? (Hint: write out this out in terms of matrices, multiply the three matrices on the left; this should give four equations – simplify them as much as you can!) Can you describe what these matrices *do*? Think about the examples of matrices that we gave you. What if impose the condition that $\det A = 1$ as well, i.e. $ad - bc = 1$?

Problem 8. (Holomorphic and anti-holomorphic splitting.) Let J be a 2×2 matrix such that $J^2 = -1$. (In particular, we can pick

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

Check this property for J. Prove that any 2×2. matrix A, can be written as a sum of matrices $A = B + C$ such that $BJ = JB$ and $CJ = -JC$. (We say that A *commutes* with J and B *anti-commutes* with J. Can you prove anything about C^2?

Problem 9. (Complex numbers form a field.) We said that for any 2-vector (x, y), we can associate the matrix $xI + yJ$ and use this correspondence to multiply 2-vectors. Namely, we say that

$$(x, y) \cdot (z, w) = \left(x \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + y \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right) \begin{bmatrix} z \\ w \end{bmatrix}.$$

Check that this "multiplication" distributes over addition of vectors, and that every element has a multiplicative inverse. (How do you compute it? Think about the problem geometrically, and remember that you can think of $xI + yJ$ as a rotation.)

Problem 10. (Changing coordinates in phase space.) Can you characterize all 2×2 matrices A such that $A^t JA = J$ with a nice equation?

Problem 11. (So you think you know linear algebra?) Let A be a matrix. Suppose there is a matrix B such that $BA = I$. Then $AB = I$. (If you look at the internet, you lose!)

Problem 12. (Algebraic groups and Iwasawa decomposition.) Let $SL_2(\mathbb{R})$ be the collection of 2×2 matrices A with $\det A = 1$. We say that A is equivalent to B when there is a 2×2 matrix C such that $B = CAC^{-1}$, where C^{-1} is some matrix such that $CC^{-1} = C^{-1}C = I$. Let K be the rotations

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

A be the scalings

$$\begin{bmatrix} r & 0 \\ 0 & 1/r \end{bmatrix},$$

and N be the shears

$$\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}.$$

Prove that any matrix in $SL_2(\mathbb{R})$ is equivalent to a matrix in K, A, or N. Prove that every element of $SL_2(\mathbb{R})$ is the product of a rotation, a scaling, and a shear – in otherwirds, that $SL_2(\mathbb{R}) = KAN$.

We call K the elliptic elements, A the hyperbolic elements, and N the parabolic elements. This is worth thinking about in as many different ways as you can possibly imagine.

2 Nonsense Formalism

First, some formal definitions. We avoided them in class, but it is nice to say them at some point. Most beginning classes in pure math spend a lot of time emphasizing the formalism; we will introduce it as we go, because that's more fun! Skip over this stuff if it seems too boring or confusing. The actual exercises are farther below.

Definition 1. A *set* is just a collection of objects. We often denote sets by letters like X or S. If x is an element inside the set X, i.e. one of the elements of the collection, we will will say that x is in S, and we may abbreviate this symbolically as $x \in S$. (Aside: $\in$ comes from ϵ, which comes from abbreviating *est* when written in Greek, which is Latin for "to be".) If we want to write down the elements of a set, we may do this as following: the natural numbers 1 through 6 could be denoted by $\{1, 2, 3, 4, 5, 6\}$. Likewise, the set of all natural numbers might be indicated by $\{1, 2, 3, 4, 5, \ldots\}$, where the dots are supposed to imply that the reader is smart enough to guess what we mean. The natural numbers are sometimes denoted by $\mathbb{N}$.

Definition 2. A *function* is something that, given an element of a set, assigns a well-defined element of another set. For instance, the operation $x \mapsto 2x$, i.e. multiplication by 2, defines a function from $\mathbb{N}$ to $\mathbb{N}$: to $1 \in \mathbb{N}$, it assigns $2 \in \mathbb{N}$, to $2 \in \mathbb{N}$, it assigns $4 \in \mathbb{N}$, etc. To every moment in time, I can assign the number of milliseconds that have passed since the beginning of 1970; this is a perfectly good function on the set of moments in time. (This is how computers measure time: google "timestamp".) We can define functions in ways that make them somewhat difficult to compute: for instance, to any natural number n, we can assign the n-th prime number. We since this function takes a natural number and gives a natural number, we might want to denote this by writing $f : \mathbb{N} \to \mathbb{N}$. We denote the value that f assigns to $n \in \mathbb{N}$ by $f(n)$. For instance, if f is the "multiply by 2" function, then $f(2) = 4$. Sometimes we call functions *maps*; we might say that "f maps 2 to 4".

Remark 1. Let's say we have a computer, which we program by giving it a finite list of numbers between 1 and N. (Have you heard of "Turing machines"?) The computer executes the program until it stops. Some programs might cause the computer never to stop computing. If our list has m elements, we might reasonably ask, "Out of all the programs of length m, how long does the program that runs longest but does not run forever, run?". (This number can be much bigger than m, since the program might have loops.) The answer to this question is called the m-th Busy Beaver number, or $BB(m)$. This is a perfectly good function on the natural numbers! However, it is *impossible to write a computer program that computes* $BB(m)$. If we could, then, given a program of length m, we could compute $BB(m)$, run the program for $BB(m)+1$ steps, and if the computer has not yet finished executing, we know that the program must run forever. In other words, if we could compute $BB(m)$, then we could solve the "Halting Problem": we could tell whether a program will terminate without running it until we die. You can find a proof on the internet explaining why you can't solve this problem – or just ask one of us over email. Moreover, $BB(m)$ must grow faster than any function f that we could compute with a computer! (Otherwise, we could wait $f(m)$ steps rather than $BB(m)$ steps.)

If you ever want to win to the "who can name the biggest number" game, just say $BB(99)$. =)

Definition 3. If X, Y are two sets, we say that the *product* of X and Y, denoted by $X \times Y$, is the set of elements (x, y), with x ranging over X and y ranging over Y. For example, $\mathbb{N} \times \mathbb{N}$ is the set of pairs of real numbers. We may also use the notation X^n to denote $X \times X \times \ldots \times X$, where there are n Xs. For instance, $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$.

Remark 2. We like to use $\mathbb{R}$ for to denote the set of real numbers. For instance, $\mathbb{R}^2$ is the set of of pairs of real numbers. This set is useful: we can use it to identity points on a blackboard, as in the lecture.

Now comes a whole lot of junk.

Definition 4. A *ring* is a set that we can add, subtract, and multiply in, but not necessarily divide in. For instance, we can add and multiply any two integers (which we will denote by the letter $\mathbb{Z}$, for the German "$\mathbb{Z}$integer" – just kidding!), but we can't necessarily divide any two integers, because $2/3$ is not an integer. How do we write this down?

Well, we have a set A of elements. Addition takes two elements and gives another one, so we have a function $+ : A \times A \to A$. Multiplication also takes two elements and gives another one, so we have a function $\cdot : A \times A \to A$. We have a special element 0 such that for any $x \in A$, $x + 0 = 0$. We also have a special element 1 such that for any $x \in A$, $1 \cdot x = x$. For any $x \in A$, there's a special element $(-x)$ with the property that $x + (-x) = 0$. In other words, negative numbers should exist. If we want to subtract y from x, we just add $-y$ to x. Finally, we have a bunch of rules, corresponding to distributivity and commutativity: for any $x, y, z \in A$, we have that

$$\cdot x + y = y + x, \tag{1}$$

$$x \cdot (y + z) = x \cdot y + x \cdot z. \tag{2}$$

.

Problem 13. Prove that $1 \cdot 0 = 0$.

Problem 14. Suppose $1 = 0$. Prove that A only has one element.

Definition 5. In the integers, $x \cdot y = y \cdot x$; this is a useful property, and we call any ring with that property a *commutative* ring.

Example 1. All polynomials in one variable with real coefficients form a commutative ring. We can add and multiply them, after all!

Definition 6. The rational numbers also form a ring: but we can also divide in the rational numbers. In other words, for every rational x, we have an element x^{-1} with the property that $x \cdot x^{-1} = 1$. Then we divide by x by multiplying by x^{-1}. Any commutative ring with this property is called a "field". We also require that $1 \neq 0$ for a field.

Example 2. All rational functions in one variable form a field.

Remark 3. We should think of a field as a "measurement" system, like the real numbers. The real numbers are a little better: we can tell which of two numbers is bigger than another. (It is not so obvious how you could do this with the field of rational functions.) Nonetheless, if you can add, multiply, and divide, it's probably a pretty good system to measure things with.

Now, we need a formal notion to capture the idea of something "compatible with a measurement system". Remember the vectors on the blackboard? We could add them and we could scale them by real numbers. Real numbers are sort of the natural way to measure the vectors, so this makes sense. Here's a formal notion:

Definition 7. A *vector space* is a set V with an operation $+ : V \times V \to V$ such that $x + y = y + x$ for all $x, y \in V$. V should also have a special element called 0 such that $x + 0 = x$ for any $x \in V$. Finally, there should be a field F together with a function $\cdot : F \times V \to V$ with the property that $1 \cdot v = v$. This "multiplication" should distribute over addition, i.e. for $x \in F$, $y, z \in V$, $x \cdot (y + z) = x \cdot y + x \cdot z$.

This generalizes the properties of the arrows on our blackboard. Whew! That was exhausting! If this confused you, that's all right! It confused the mathematical world for many years as well. It's really not that important for our purposes.