

Hidden Structure and Computation

2021 pre-REU packet

prepared by Sean Howe

June 1-25, 2021

10:00am-5:00pm mountain time on weekdays

Meetings on gather.town

Program staff:

Sean Howe (sean.howe@utah.edu) - Director

Peter Wear (wear@math.utah.edu) - Codirector

Amanda Alexander (alexander@math.utah.edu) - Assistant

Matthew Bertucci (bertucci@math.utah.edu) - Assistant

Carlos Ospina (ospina@math.utah.edu) - Assistant

Contents

What is this, anyway?	ii
Acknowledgements	ii
Week 1. Clubs and Vectors	1
1. Definitions	1
1.1. Counting	1
1.2. Vectors	1
1.3. Linear independence	3
2. Problems	4
2.1. Clubtown, eventown, and oddtown	4
2.2. Vectors and the dot product	6
2.3. Linear independence	8
3. Challenge problems	10
Week 2. Matrices and Transformations	11
1. Definitions	11
1.1. Matrices	11
1.2. Linear transformations	12
1.3. The field with two elements	12
1.4. Reduced row echelon form and rank	13
2. Problems	14
2.1. Matrices	14
2.2. Matrices as geometric transformations	15
2.3. Matrices as data transformations	16
2.4. Systems of equations, RREF, and rank	18
2.5. The field with two elements and the Oddtown theorem	19
2.6. More fields and isotropic vectors	20
3. Challenge problems	21
Week 3. Polynomials, permutations, determinants, and finding eigenvalues	22
1. Definitions	22
1.1. Polynomials	22
1.2. Permutations	22
1.3. Determinants	24
1.4. Bases and the matrix of a linear transformation	25
2. Problems	27
2.1. Polynomials	27
2.2. Permutations	29
2.3. Determinants and characteristic polynomials	31
2.4. Bases and the matrix of a linear transformation	33
3. Challenge problems	34

What is this, anyway?

This is a packet for the *whole program*, split up by weeks. Each weekend we'll update it with the material for the following week (and fix any typos we find!). Each week contains three sections:

- (1) “Definitions and Facts” contain just the barebones – it’s meant as a reference, so that you can refer back to it after we cover some ideas while you’re working on problems. Facts are statements that you can take to be true and use to solve problems; if you’re curious, justifications for most of these facts are worked out eventually in the problems, though we might not get to all of them as a group!
- (2) “Problems” is the meat. They’re broken up by topic, roughly in the order we plan to get to them, but we may skip around. A problem with a \star might be a bit harder, depending on your background. A problem with a \square is meant to be explored using a computer. There are lots of problems – we definitely won’t go over all of them or even most of them, and different people will work on different problems!

Some problems ask you to “show”, “explain”, or “justify” something. This means that you should give as precise and logically rigorous of an explanation as you can, starting from facts we already know. This might start with working out a few simple examples to observe a pattern, but you should try to push beyond that afterwards to try and put into words the underlying phenomenon! If you’ve taken a proofs-based course, try to write a proof; if not, just try to be as careful and compelling as you can with your reasoning!

- (3) “Challenge problems” – we won’t discuss these at all as a group, though you might end up working on them if you get bored, and you are welcome to discuss them with the other participants in the program in the afternoons. Sometimes they’re closely related to things we are doing, sometimes they aren’t. Some of them are pretty hard, but all of them are really interesting and have beautiful solutions. If you do solve one, write up your solution as carefully as you can and then send it to Sean in an email so that we can discuss!

Acknowledgements. A lot of the material in this program was originally modeled off of Laci Babai’s Apprentice Program at the University of Chicago REU. Thanks Laci!

Clubs and Vectors

1. Definitions

1.1. Counting.

DEFINITION 1.1.1 (Factorials). For n a positive integer, $n! := n(n-1)(n-2)\dots(3)(2)(1)$. We say out loud “ n factorial.” It is convenient to also declare that $0! = 1$.

EXAMPLE 1.1.2. The number $n!$ also counts the number of different ways we can arrange n distinct things in a row (why?) – for example, $3! = 3 \cdot 2 \cdot 1 = 6$, and we can arrange an apple, a banana, and a chicken in any of the six following orders:

$$(a, b, c), (a, c, b), (b, a, c), (b, c, a), (c, a, b), \text{ or } (c, b, a).$$

As another example, $52!$ is the number of different ways a deck of cards can be ordered after shuffling.

DEFINITION 1.1.3 (Binomial coefficients). For n a positive integer and $0 \leq k \leq n$,

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}.$$

We say out loud n choose k (why?), and also call this a *binomial coefficient* (why?).

1.2. Vectors.

DEFINITION 1.2.1 (Row and column vectors). A *row vector of length n* is a list of n real numbers next to each other

$$\vec{v} = [a_1 \ a_2 \ \dots \ a_n], \ a_i \in \mathbb{R}$$

A *column vector of length n* is a stack of n real numbers on top of each other

$$\vec{v} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

DEFINITION 1.2.2 (Transpose of vectors). We write

$$[a_1 \ a_2 \ \dots \ a_n]^t = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \text{ and } \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}^t = [a_1 \ a_2 \ \dots \ a_n]$$

Out loud we say \vec{v}^t as “ v transpose.”

DEFINITION 1.2.3. \mathbb{R}^n is the set of column vectors with n real entries,

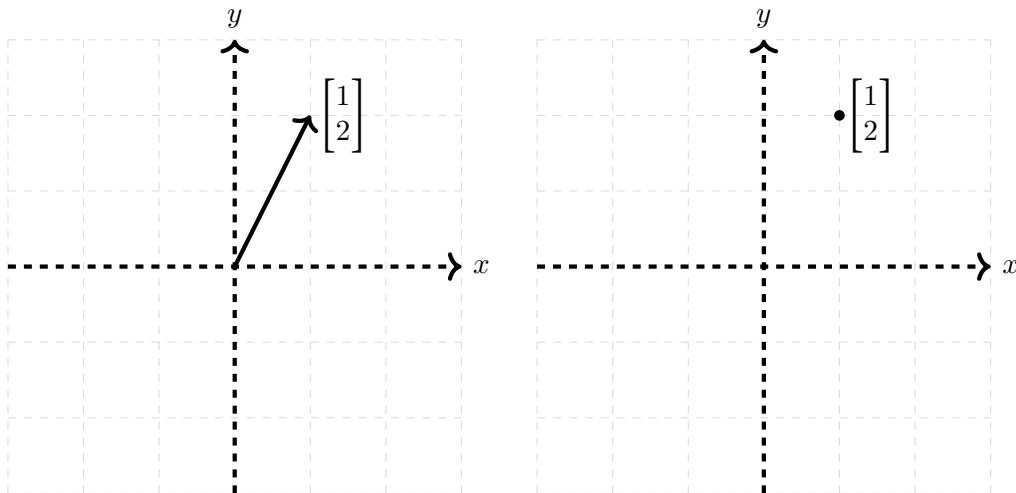
$$\mathbb{R}^n = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix}, a_i \in \mathbb{R} \right\}.$$

DEFINITION 1.2.4. The *zero vector* $\vec{0} \in \mathbb{R}^n$ is the vector

$$\vec{0} := [0 \ 0 \ \dots \ 0]^t.$$

EXAMPLE 1.2.5 (Geometric interpretation of vectors). We can think of a vector

$$\begin{bmatrix} x \\ y \end{bmatrix} = [x \ y]^t \in \mathbb{R}^2 \text{ as an arrow or point in the plane,}$$



Similarly, we can think of a vector

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = [x \ y \ z]^t \in \mathbb{R}^3 \text{ as an arrow or point in three dimensional space.}$$

And so on! (Of course, it's a lot harder to draw pictures in $n \geq 4$ dimensions!).

1.2.6. Operations with vectors.

DEFINITION 1.2.7. We can add two column vectors of the same height, or multiply a column vector by a scalar λ (scalar = a real number), and similarly for row vectors.

$$\begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \dots \\ a_n + b_n \end{bmatrix}, \quad \lambda \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} = \begin{bmatrix} \lambda a_1 \\ \lambda a_2 \\ \dots \\ \lambda a_n \end{bmatrix}.$$

DEFINITION 1.2.8. If

$$\vec{v} = [a_1 \ a_2 \ \dots \ a_n]^t \in \mathbb{R}^n$$

then the *length* of \vec{v} , $\|\vec{v}\|$, is defined by

$$\|\vec{v}\| = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}.$$

DEFINITION 1.2.9. The *dot product* of two vectors in \mathbb{R}^n is the scalar (real number) given by

$$[a_1 \ a_2 \ \dots \ a_n]^t \cdot [b_1 \ b_2 \ \dots \ b_n]^t = a_1 b_1 + a_2 b_2 + \dots + a_n b_n.$$

FACT 1.2.10. If $\vec{v}, \vec{w} \in \mathbb{R}^n$ are two non-zero vectors, and θ is the angle between the corresponding arrows out of the origin

$$\vec{v} \cdot \vec{w} = \cos(\theta) \|\vec{v}\| \|\vec{w}\|.$$

DEFINITION 1.2.11. Two vectors \vec{v} and \vec{w} in \mathbb{R}^n are *orthogonal* if $\vec{v} \cdot \vec{w} = 0$. We say a collection of vectors $\vec{v}_1, \dots, \vec{v}_k$ is *pairwise orthogonal* if for each $i \neq j$, \vec{v}_i and \vec{v}_j are orthogonal (i.e. $\vec{v}_i \cdot \vec{v}_j = 0$).

1.3. Linear independence.

DEFINITION 1.3.1.

- (1) A *linear combination* of vectors $\vec{v}_1, \dots, \vec{v}_k$ in \mathbb{R}^n is a sum

$$a_1\vec{v}_1 + \dots + a_k\vec{v}_k \text{ for } a_i \in \mathbb{R}.$$

It is called *trivial* if $a_i = 0$ for all $1 \leq i \leq k$, otherwise it is non-trivial.

- (2) A collection of vectors $\vec{v}_1, \dots, \vec{v}_k$ in \mathbb{R}^n is *linearly dependent* if $\vec{0}$ can be written as a non-trivial linear combination of the vectors, i.e. if there are $a_i \in \mathbb{R}$ **not all zero** such that

$$a_1\vec{v}_1 + \dots + a_k\vec{v}_k = \vec{0}.$$

Such a linear combination is called a *linear dependence*.

If the vectors are not linearly dependent, then they are *linearly independent*. Equivalently, vectors $\vec{v}_1, \dots, \vec{v}_k$ are linearly independent if any nontrivial linear combination is nonzero (i.e. not equal to $\vec{0}$).

FACT 1.3.2 (The first miracle of linear algebra). We state this fact in two equivalent ways:

- (1) If $\vec{v}_1, \dots, \vec{v}_k$ are linearly independent vectors in \mathbb{R}^n , then $k \leq n$.
- (2) If $\vec{w}_1, \dots, \vec{w}_m$ are vectors in \mathbb{R}^n with $m > n$, then they are linearly dependent.

2. Problems

2.1. Clubtown, eventown, and oddtown. The residents of Clubtown, population N , decide to form clubs, subject to the following rules:

- (1) Two clubs are considered to be the same if they have the exact same members¹.
- (2) A club may have any number of members, including zero.

EXERCISE 2.1.1. How many distinct clubs can be formed?

EXERCISE 2.1.2. How many clubs are there with

- A. An even number of members?
- B. An odd number of members?

EXERCISE 2.1.3. For $0 \leq k \leq N$, how many clubs can be formed with k members?

EXERCISE 2.1.4. * Use Exercise 2.1.3 to explain the result of Exercise 2.1.1 by evaluating $(1+t)^N$ at $t=1$.

EXERCISE 2.1.5 (Clubtown). The mayor of Clubtown decides to cut down on the proliferation of clubs. They change the second rule and add a third:

- (1) Two clubs are considered to be the same if they have the exact same members.
- (2) ~~A club may have any number of members, including zero.~~
Any club must have at least two members.
- (3) Any two clubs can share at most one member.

A. Find some *maximal* systems of clubs for the new rules². What's the largest maximal system you can find? The smallest? Experiment – think about examples with some small values of N (the population), then see if you can generalize.

B. “*Still too many clubs!*” rages the Mayor, who updates the third rule:

- (3) ~~Any two clubs can share at most one member.~~
Any two clubs must share **exactly** one member.

What's the largest maximal system you can find now?

EXERCISE 2.1.6 (Eventown). Some residents of Clubtown, upset with the new rules, move away to found Eventown. They keep the first rule, but change the second and third rules:

- (1) Two clubs are considered to be the same if they have the exact same members.
- (2) Each club has to have an **even** number of members.
- (3) Each pair of clubs has to have an even number of members in common.

A. Find a simple way to make $2^{\lfloor N/2 \rfloor}$ clubs³.

B. * When $N=8$, find a fundamentally different way to make $16=2^4$ clubs.

EXERCISE 2.1.7 (Oddtown). Eventown is no paradise either, so others move away to found Oddtown. They keep the same first rule as Clubtown and Eventown and the same third rule as Eventown, but the second rule is changed again:


- (1) Two clubs are considered to be the same if they have the exact same members.
- (2) Each club has to have an **odd** number of members.
- (3) Each pair of clubs has to have an even number of members in common.

¹For example, if the Dog Lovers Club and the Cat Lovers Club have the same members, then they merge to become the Dog and Cat Lovers Club. In other words, a club is determined uniquely by its members.

²A system of clubs that satisfies these rules is called maximal if it is not possible to add in another club without breaking the rules.

³ $\lfloor N/2 \rfloor$ means the largest positive integer that is $\leq N/2$; e.g. $\lfloor 5 \rfloor = 5$, $\lfloor 4.999 \rfloor = 4$, $\lfloor 3/2 \rfloor = 1$.

Experiment with finding some valid systems of clubs. What's the largest you can find?

EXERCISE 2.1.8 (). Write a program to find maximal valid systems of clubs in Clubtown, Eventown, and Oddtown, then use it to look for patterns related to the above questions.

2.2. Vectors and the dot product.

EXERCISE 2.2.1. The i th standard basis vector in \mathbb{R}^n is the vector

$$\vec{e}_i = \underbrace{\begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots \end{bmatrix}^t}_{\text{1 in } i\text{th spot, 0 everywhere else}}$$

Show that if $\vec{v} = [a_1 \ a_2 \ \dots \ a_n]^t$ then

$$\vec{v} = a_1\vec{e}_1 + a_2\vec{e}_2 + \dots + a_n\vec{e}_n.$$

EXERCISE 2.2.2. For $\lambda \in \mathbb{R}$, explain why $\|\lambda\vec{v}\| = |\lambda|\|\vec{v}\|$ in two ways: geometrically and algebraically.

EXERCISE 2.2.3. For $\vec{v}, \vec{w}, \vec{u} \in \mathbb{R}^n$, and $\lambda \in \mathbb{R}$, show

- A. $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$.
- B. $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$.
- C. $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$
- D. $\vec{u} \cdot (\lambda\vec{v}) = (\lambda\vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v})$.

EXERCISE 2.2.4. Assuming Fact 1.2.10, explain (briefly!) why two nonzero vectors are orthogonal (in the sense of Definition 1.2.11) if and only if they make a right angle.

EXERCISE 2.2.5. Suppose $\vec{v}_1, \dots, \vec{v}_k$ is a set of non-zero pairwise orthogonal vectors in \mathbb{R}^n and each \vec{v}_i has length 1 (i.e. $\|\vec{v}_i\| = 1$). If $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ are scalars and

$$\vec{w} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \dots + \lambda_k\vec{v}_k$$


then show $\vec{w} \cdot \vec{v}_i = \lambda_i$. Afterwards, compare with Exercise 2.2.1.

EXERCISE 2.2.6. Suppose $\|\vec{w}\| = 1$.

- A. Show that $\vec{v}^\perp = \vec{v} - (\vec{v} \cdot \vec{w})\vec{w}$ is orthogonal to \vec{w} , then draw a picture illustrating the identity

$$\vec{v} = (\vec{v} \cdot \vec{w})\vec{w} + \vec{v}^\perp.$$

Hint: For the picture, start by thinking about what happens in \mathbb{R}^2 if $\vec{w} = e_1$.

- B.  Write a program that generates random vectors \vec{v} and \vec{w} in \mathbb{R}^3 with $\|\vec{w}\| = 1$ then plots the picture showing $\vec{v} = (\vec{v} \cdot \vec{w})\vec{w} + \vec{v}^\perp$.

EXERCISE 2.2.7. Draw a detailed diagram to explain Fact 1.2.10 via the definition of cosine in terms of right triangles.


Hint: Rescale so that $\|w\| = 1$, then consider the picture in the previous exercise.

EXERCISE 2.2.8.

- A. Find a set of n nonzero pairwise orthogonal vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ in \mathbb{R}^n .
- B. For each n , explain a way to find infinitely many different sets of vectors as in (1).
- C. For $n \geq 2$, explain a way to find infinitely many different sets of vectors as in (1) with the added condition that $\|\vec{v}_i\| = 1$ for all $1 \leq i \leq n$.

EXERCISE 2.2.9.

- A. Suppose a streetlight shines from two meters above the center of a parking lot. A mosquito is at coordinates (x, y, z) , which means x meters east of the center, y meters north of the center, and z meters above the center. Assume the mosquito is flying above the ground ($z > 0$) and below the top of the light ($z < 2$). Where will the mosquito's shadow fall on the parking lot? Express your answer as the multiplication of a matrix and a vector.

- B.** Suppose there's a wall going northwest through a point 1 meter east of the center of the parking lot. Where will the mosquito's shadow fall on the wall? (Assuming it's in a spot where it will cast a shadow on the wall – what are those conditions on x, y, z ?). Express your answer as the multiplication of a matrix and a vector.
- C.**  Write a program that finds and plots the shadow of a mosquito at a random spot from a light at another random spot onto a random wall.

2.3. Linear independence.

EXERCISE 2.3.1. Find a non-trivial linear dependence between the vectors

$$\vec{v}_1 = [1 \ 2]^t, \vec{v}_2 = [1 \ 4]^t, \text{ and } \vec{v}_3 = [1 \ 8]^t,$$

i.e. an equation $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = 0$ with $c_1, c_2,$ and c_3 real numbers not all equal to zero.

EXERCISE 2.3.2.

A. Without doing any computations, show that

$$\vec{v}_1 = \begin{bmatrix} 1 \\ \pi \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} \sqrt{2} \\ 7 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0.2 \\ 300 \end{bmatrix}.$$

are linearly dependent vectors in \mathbb{R}^2 . *Hint: Fact 1.3.2.*

B. Find a linear dependence between them. *Hint: Set up a system of equations and then solve it! This will be a little messy.*

EXERCISE 2.3.3. Are the vectors

$$\vec{v}_1 = [1 \ 2 \ 3]^t, \vec{v}_2 = [4 \ 5 \ 6]^t, \text{ and } \vec{v}_3 = [7 \ 8 \ 9]^t$$

linearly independent?

EXERCISE 2.3.4. Show that if $\vec{v}_i, i = 1, \dots, k$ is a set of vectors in \mathbb{R}^n such that $\vec{v}_i = \vec{0}$ for some i then the vectors are linearly dependent.

EXERCISE 2.3.5. Show that $\vec{v}_1, \dots, \vec{v}_k$ is linearly dependent if and only if there is some $1 \leq i \leq k$ such that \vec{v}_i can be written as a linear combination of the other vectors, i.e.

$$\vec{v}_i = a_1\vec{v}_1 + \dots + a_{i-1}\vec{v}_{i-1} + a_{i+1}\vec{v}_{i+1} + \dots + a_k\vec{v}_k.$$

EXERCISE 2.3.6. Find a set of n vectors in \mathbb{R}^n which are linearly independent.

EXERCISE 2.3.7. Suppose $\vec{v}_1, \dots, \vec{v}_k$ are a set of pairwise orthogonal non-zero vectors in \mathbb{R}^n . Show that they are linearly independent.

EXERCISE 2.3.8. Find two vectors in \mathbb{R}^2 which are linearly independent but not orthogonal. Justify as carefully as you can why the two vectors you find really are linearly independent.

EXERCISE 2.3.9. Suppose $\vec{v}_1, \dots, \vec{v}_k$ are vectors in \mathbb{R}^n such that

- (1) $\vec{v}_i \cdot \vec{v}_j = 1$ for $i \neq j$
- (2) $\|\vec{v}_i\| \geq \sqrt{2}$.

A. For $a_1, \dots, a_k \in \mathbb{R}$ not all zero, show that $\|\sum_{i=1}^k a_i \vec{v}_i\| > 0$.

Hint: Expand the dot product of $\sum_{i=1}^k a_i \vec{v}_i$ with itself.

B. Deduce that $\vec{v}_1, \dots, \vec{v}_k$ are linearly independent.

C. Using Fact 1.3.2, deduce that $k \leq n$.

D. Recall that in Clubtown any club has at least two members, and any two clubs must share exactly one member. Show that the membership vectors⁴ of any valid system of clubs satisfy the above hypotheses, then use C. to justify

THEOREM (Clubtown Theorem). *If Clubtown, with rules as in Exercise 2.1.5-B., has n residents, then there can be at most n clubs.*

⁴The membership vector of a club C is obtained by numbering the residents 1 through n , then putting a 1 in the i th entry if resident i is a member of club C and a 0 in the i th entry otherwise.

- E.** * Can you give some geometric intuition for why the vectors $\vec{v}_1, \dots, \vec{v}_k$ are linearly independent? (Hint: what if I instead require $\|\vec{v}_i\| \geq c$ for some positive real number c ? How small can c be for the argument in A. to still work? What happens to the picture if I let c go to infinity?).

3. Challenge problems

Remember some of these might be hard! In particular, don't feel bad if you spend hours thinking about one and don't solve it, but also don't let this warning stop you from trying!

EXERCISE 3.0.1 (The envelope problem⁵). I have two pieces of paper, each with a positive whole number written on it. The numbers are not the same. I let you pick one piece of paper, and you look at the number on it. You then must guess whether you have seen the larger or smaller number. Find a way to guess that gives you better than a 50 percent chance of success, no matter what the two numbers are.

EXERCISE 3.0.2. Consider an 8 by 8 chessboard, with two opposite corners removed. Is it possible to cover this with non-overlapping dominoes? (A single domino covers two adjacent squares).

EXERCISE 3.0.3. What if we take an n by n chessboard, with n not divisible by 3, and remove one square. Is it possible to cover this with non-overlapping triominoes (covers three in a row)?

EXERCISE 3.0.4. Consider the three pairwise adjacent faces of an $n \times n \times n$ cube. For what values of n is it possible to tile the cube with 3×1 bandaids? A bandaid can wrap around an edge.

⁵This problem is **not** a paradox. It has a perfectly reasonable mathematical solution, even though you might think it looks like a paradox at first!

Matrices and Transformations

1. Definitions

1.1. Matrices.

DEFINITION 1.1.1. A $m \times n$ (real) matrix A is a table with m rows, n columns, and entries in \mathbb{R}

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & \dots & \\ A_{31} & \dots & & \\ \dots & & & \end{bmatrix}.$$

Here A_{ij} denotes the entry in the i th row and j th column.

We can add together to $m \times n$ matrices A and B by adding their corresponding entries, or multiply a matrix by a scalar by multiplying all of the entries:

$$(A + B)_{ij} = A_{ij} + B_{ij}, \quad (\lambda A)_{ij} = \lambda A_{ij}.$$

DEFINITION 1.1.2. If A is an $m \times n$ matrix, the transpose matrix A^t is the $n \times m$ matrix obtained by swapping the rows and columns of A . In terms of entries,

$$(A^t)_{ij} = A_{ji}.$$

EXAMPLE 1.1.3. A $m \times 1$ matrix is a length m column vector. A $1 \times n$ matrix is a length n row vector, with addition and scalar multiplication as defined before.

DEFINITION 1.1.4. The $n \times n$ matrix with ones along the diagonal and zeroes everywhere else is called the $n \times n$ identity matrix

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \dots & \ddots & \dots \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \text{i.e. } (I_n)_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

DEFINITION 1.1.5. We can multiply an $r \times s$ matrix A and an $s \times t$ matrix B to get an $r \times t$ matrix AB with entries

$$(AB)_{ij} = \sum_{k=1}^s a_{ik}b_{kj}.$$

Matrix multiplication is associative: $A(BC) = (AB)C$ when all of the multiplications make sense (i.e. A , B , and C are the right sizes to multiply!).

Warning: Matrix multiplication is not usually commutative – in general $AB \neq BA$ even when both make sense (i.e. both A and B are $n \times n$ matrices for the same n).

EXAMPLE 1.1.6. For any $m \times n$ matrix A ,

$$AI_n = A \text{ and } I_m A = A.$$

DEFINITION 1.1.7. An $n \times n$ matrix A is *invertible* if there is another $n \times n$ matrix A^{-1} such that $A^{-1}A = I_n$, where I_n is the $n \times n$ identity matrix.

FACT 1.1.8 (Matrix inverses are two-sided). $A^{-1}A = I_n$ if and only if $AA^{-1} = I_n$.

DEFINITION 1.1.9. For A an $n \times n$ matrix, a **nonzero** vector $\vec{v} \in \mathbb{R}^n$ is an *eigenvector of A with eigenvalue $\lambda \in \mathbb{R}$* if

$$A\vec{v} = \lambda\vec{v}.$$

A scalar $\lambda \in \mathbb{R}$ is an eigenvalue of A if there is an eigenvector of A with eigenvalue λ .

1.2. Linear transformations.

DEFINITION 1.2.1. A linear transformation T is a function from \mathbb{R}^n to \mathbb{R}^m such that

$$T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$$

for any $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$, and

$$T(\lambda\vec{v}) = \lambda T(\vec{v})$$

for any $\vec{v} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$.

EXAMPLE 1.2.2. If A is an $m \times n$ matrix, then

$$T(\vec{v}) = A\vec{v}$$

is a linear transformation from \mathbb{R}^n to \mathbb{R}^m .

FACT 1.2.3. Any linear transformation is of this form for some matrix A .

1.3. The field with two elements.

DEFINITION 1.3.1. A *field* is a mathematical setting where you can multiply, add, subtract, and divide satisfying all of the expected rules. In particular, there are distinguished elements 0 and 1 that behave as you expect, and if \mathbb{K} is a field and k is a number in \mathbb{K} that is not equal to zero then it admits a *multiplicative inverse*, i.e. an element $1/k$ such that $k(1/k) = 1$.

EXAMPLE 1.3.2. The rational numbers \mathbb{Q} are a field; so are the real numbers \mathbb{R} and the complex numbers \mathbb{C} . The integers \mathbb{Z} are not a field (why?).

DEFINITION 1.3.3. The field with two elements \mathbb{F}_2 has just two numbers, called 0 and 1. Addition is described by

$$0 + 0 = 0, 0 + 1 = 1 + 0 = 1, 1 + 1 = 0$$

and multiplication is described by

$$(0)(0) = 0, (1)(0) = (0)(1) = 0, (1)(1) = 1$$

DEFINITION 1.3.4. It makes sense to talk about vectors and matrices with entries in any field, and linear independence, the dot product and matrix multiplication still make sense. In particular, \mathbb{F}_2^n is the set of column vectors of length n with entries in \mathbb{F}_2 .

DEFINITION 1.3.5. For p a prime number, the field with p elements, \mathbb{F}_p (sometimes also written $\mathbb{Z}/p\mathbb{Z}$), has p numbers, called

$$0, 1, 2, \dots, p-1$$

with addition and multiplication defined by performing addition/multiplication of integers then taking the remainder when dividing by p . This turns out to be a field!

1.4. Reduced row echelon form and rank.

DEFINITION 1.4.1. A $m \times n$ matrix A is in *reduced row echelon form* (RREF for short) if:

- (1) The first non-zero entry in each row is a 1. We call these *leading ones*. (Note: the whole row can be zero, so each row either has a leading one or is zero.)
- (2) The leading ones proceed from left to right as we go down the rows.
- (3) Any column with a leading one is zero in all other entries.

EXAMPLE 1.4.2. The matrix

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

is in RREF, while

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

are not (which conditions do they fail?).

DEFINITION 1.4.3. Given a $m \times n$ matrix A consider the following *elementary row operations*:

- (1) Swap two rows.
- (2) Replace a row of A with a non-zero multiple of itself.
- (3) Add a multiple of one row to a different row.

FACT 1.4.4. If A is a $m \times n$ matrix, then it can be put into reduced row echelon form using elementary row operations, and the resulting matrix in RREF is unique.

DEFINITION 1.4.5. The *rank* of a matrix A is the number of leading ones when it is put in reduced row echelon form.

DEFINITION 1.4.6. For A an $m \times n$ matrix, we define the *kernel* (or *nullspace*) of A to be

$$\ker A = \{\vec{v} \in \mathbb{R}^n \mid A\vec{v} = \mathbf{0}\}$$

2. Problems

2.1. Matrices.

EXERCISE 2.1.1. Give an example of 2×2 matrices A and B such that $AB = 0$ and $BA \neq 0$ (here 0 means the matrix with all entries 0).

EXERCISE 2.1.2. If $\vec{v} \in \mathbb{R}^n$, and A is an $m \times n$ matrix, what type of object is $A\vec{v}$?

EXERCISE 2.1.3. Find all of the eigenvectors and eigenvalues for the matrix $\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$.

EXERCISE 2.1.4. A permutation on n elements is a one-to-one map $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$.

- A. Given a permutation σ , describe a matrix A_σ such that $A_\sigma \vec{e}_i = \vec{e}_{\sigma(i)}$ where \vec{e}_i denotes the i th standard basis vector (see Exercise 2.2.1).
- B. If σ and τ are two permutations, describe the matrix $A_\sigma A_\tau$.
- C. * For σ a permutation, describe all of the eigenvectors for A_σ .

EXERCISE 2.1.5.

- A. Express the rows of AB as linear combinations of the rows of B .
- B. Express the columns of AB as linear combinations of the columns of A .

Note: It can be really handy sometimes to think about matrix multiplication like this: each row of A is a set of instructions on how to combine the rows of B , or each column of B is a set of instructions on how to combine the columns of A .

EXERCISE 2.1.6. Show that $(AB)^t = B^t A^t$.

EXERCISE 2.1.7. If $\vec{v}, \vec{w} \in \mathbb{R}^n$, show

$$\vec{v}^t \vec{w} = \vec{v} \cdot \vec{w},$$

where the lefthand side is matrix multiplication of a $1 \times n$ matrix and a $n \times 1$ matrix and the righthand side is the dot product of two column vectors.

2.2. Matrices as geometric transformations.

EXERCISE 2.2.1.

A. Show that if T is a linear transformation from \mathbb{R}^m to \mathbb{R}^n , then

$$T(\vec{0}) = \vec{0}.$$

B. Show that $T(\begin{bmatrix} x & y \end{bmatrix}^t) = \begin{bmatrix} x+1 & y \end{bmatrix}^t$ is not a linear transformation from \mathbb{R}^2 to \mathbb{R}^2 .

EXERCISE 2.2.2. For each of the following matrices A , give a geometric description of the linear transformation $T(\vec{v}) = A\vec{v}$ and use this to describe the eigenvalues and eigenvectors.

A. $\begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$ B. $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ C. $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ D. $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

EXERCISE 2.2.3. Given a linear transformation $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$, how can you compute the i th column in the matrix A giving rise to T as in Fact 1.2.3?

EXERCISE 2.2.4. For $\theta \in [0, 2\pi)$, let $R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the transformation of counter-clockwise rotation about the origin by angle θ (in radians). This is a linear transformation (you may assume this without justification).

A. Find the matrix for R_θ .

B. What are the eigenvalues and eigenvectors of this matrix, if any?

Hint: think geometrically! Your answer will depend on θ .

EXERCISE 2.2.5. \star For $\theta \in [0, 2\pi)$, consider the linear transformation from \mathbb{R}^3 to \mathbb{R}^3 that rotates a vector \vec{v} counter-clockwise by angle θ around the ray from the origin in the direction of $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^t$.

A. Sketch a picture illustrating this, then find the matrix A for this linear transformation.

B. \square Write a computer program that finds the matrix for rotation by angle θ around a ray in any direction in \mathbb{R}^3 .

EXERCISE 2.2.6. Consider a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

A. Compute $A\vec{e}_1$ and $A\vec{e}_2$.

B. Sketch a parallelogram with these two vectors as sides.

C. Explain why the area of this parallelogram is $|ad - bc|$

Hint: Exercise 3.2.6-A. from week one might be helpful here!

D. Give a geometric explanation of the sign (positive or negative) of $ad - bc$.

2.3. Matrices as data transformations.

EXERCISE 2.3.1. The Fibonacci sequence is the sequence

$$0, 1, 1, 2, 3, 5, 8, 13, \dots$$

defined by the *recurrence relation*

$$f_0 = 0, f_1 = 1, \text{ and } f_n = f_{n-1} + f_{n-2} \text{ for } n \geq 2.$$

A. Find a 2×2 matrix A such that for $n \geq 2$,

$$\begin{bmatrix} f_n \\ f_{n-1} \end{bmatrix} = A \begin{bmatrix} f_{n-1} \\ f_{n-2} \end{bmatrix}$$

B. Explain why f_n is the first entry of $A^{n-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

C. Compute the eigenvalues of this matrix.

D. Express $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as a linear combination of eigenvectors.

E. Use this to give a closed formula for f_n .

F. \square Experiment with sequences defined by other recurrence relations (maybe involving more than just the previous two terms, and starting at different values!). Write computer programs to solve them explicitly and to find closed formulas for their terms and compare the results.

EXERCISE 2.3.2. We have a pond filled with lily pads, numbered 1 through n . A frog is jumping between the pads, once every minute. If a frog is on the j th lily pad, then the probability it will jump to the i th lily pad is a real number p_{ij} . We also write $p_{ij}^{(s)}$ for the probability that a frog who starts on lily pad j will end up on lily pad i after s jumps. We write T for the $n \times n$ matrix with entries

$$T_{ij} = p_{ij}.$$

A. Explain why

$$(T^s)_{ij} = p_{ij}^{(s)}.$$

Here the lefthand side is the ij th entry of the matrix

$$T^s = \underbrace{(T)(T)(T) \dots (T)}_{\text{multiplying together } s \text{ copies of } T}.$$

B. \square Write a computer program which generates some transition probabilities p_{ij} (make sure $\sum_{i=1}^n p_{ij} = 1$ for each j (why?)), and computes T^s for some very large s . What do you notice?

EXERCISE 2.3.3 (A Random Hop). In the setting of Exercise 2.3.2, suppose we have n lily pads in our pond and that a frog is equally likely to hop from any one to any other (but will not stay on the lily pad it is on).

(1) Write down the transition matrix T .

(2) Show that

$$\vec{v}_{\text{stab}} = \begin{bmatrix} \frac{1}{n} & \dots & \frac{1}{n} \end{bmatrix}$$

is an eigenvector for T . What is its eigenvalue?

(3) Show that if

$$\vec{w} = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix}^t$$

is such that $\sum_{i=1}^n a_i = 0$, then \vec{w} is an eigenvector for T . What is its eigenvalue?

- (4) For each $1 \leq i \leq n$, write \vec{e}_i as a linear combination

$$\vec{e}_i = \vec{v}_{\text{stab}} + \vec{w}$$

where \vec{w} is as in the previous part.

- (5) Deduce that $\lim_{N \rightarrow \infty} T^N \vec{e}_i = \vec{v}_{\text{stab}}$ (here taking a limit of vectors just means taking a limit in each coordinate).
- (6) What does this mean in terms of the frog hopping around?
- (7) Show the refined estimate

$$|(T^N)_{ij} - 1/n| \leq \left(\frac{1}{n-1}\right)^{N-1} \left(\frac{1}{n}\right),$$

which gives a precise estimate of how quickly $T^N \vec{e}_i$ converges to \vec{v}_{stab} .

2.4. Systems of equations, RREF, and rank.

EXERCISE 2.4.1. Consider a system of m linear equations in n unknowns x_1, \dots, x_n

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= c_1 \\a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= c_2 \\&\dots = \dots \\a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= c_n\end{aligned}$$

Rewrite this system of equations as a single matrix equation $A\vec{x} = \vec{c}$. In particular, note that $\ker A$ (see Definition 1.4.6) is identified with the set of solutions of the *homogeneous* system of equations $A\vec{v} = \vec{0}$.

EXERCISE 2.4.2. For A an $m \times n$ matrix, show that

- A. $\vec{0} \in \ker A$
- B. If $\vec{v}_1, \vec{v}_2 \in \ker A$, then $\vec{v}_1 + \vec{v}_2 \in \ker A$.
- C. If $\vec{v} \in \ker A$ and $\lambda \in \mathbb{R}$ then $\lambda\vec{v} \in \ker A$.

EXERCISE 2.4.3.

- A. If A is a $m \times n$ matrix, and R is the reduced row echelon form of A , explain why

$$\ker A = \ker R.$$

That is, explain why the solutions to $A\vec{x} = 0$ are the same as the solutions to $R\vec{x} = 0$.

- B. Suppose R is a matrix in RREF. Explain how to find the solutions to $R\vec{x} = \vec{0}$.
- C. When, in terms of the rank of A , is there a non-zero solution to $A\vec{x} = \vec{0}$?

EXERCISE 2.4.4. Compute the rank of the following $n \times n$ matrices $A = (a_{ij})$.

- A. $a_{ij} = i + j$
- B. $a_{ij} = ij$
- C. $a_{ij} = i^2 + j^2$
- D. $a_{ij} = (i + j)^2$
- E. $a_{ij} = 2^{i+j}$
- F. $a_{ij} = 2^{ij}$.

EXERCISE 2.4.5. If A is an $m \times n$ matrix, explain why

- A. $\text{rank}(A) \leq m$, and
- B. $\text{rank}(A) \leq n$.

EXERCISE 2.4.6.

- A. For each elementary row operation on a $m \times n$ matrix, find a matrix E such that performing the elementary row operation on a $m \times n$ matrix A can be expressed as the matrix multiplication EA . *Hint: recall Exercise 2.1.5.*
- B. Does this give you another way to understand 2.4.3-A.?

EXERCISE 2.4.7. Explain why $\ker A$ can be interpreted as the set of all linear dependences between the columns of A (*Hint: recall Exercise 2.1.5*). Deduce that the columns of A are linearly independent if and only if $\ker A = \{\vec{0}\}$.

EXERCISE 2.4.8 (The First Miracle/Fact 1.3.2). Suppose given k linearly independent vectors $\vec{v}_1, \dots, \vec{v}_k$ in \mathbb{R}^m . Combine Exercises 2.4.3-C, 2.4.5, and 2.4.7 to show that $k \leq m$. *Hint: form the $m \times k$ matrix A whose i th column is \vec{v}_i .*

2.5. The field with two elements and the Oddtown theorem.

EXERCISE 2.5.1. How many vectors are there in \mathbb{F}_2^n ?

EXERCISE 2.5.2.

A. For $\vec{v} \in \mathbb{F}_2^n$, show that

$$\vec{v} \cdot \vec{v} = \begin{cases} 1 & \text{if there are an odd number of ones in } \vec{v} \\ 0 & \text{if there are an even number of ones in } \vec{v} \end{cases}$$

B. For \vec{v} and \vec{w} in \mathbb{F}_2^n , when is $\vec{v} \cdot \vec{w} = 0$?

EXERCISE 2.5.3. Show that if $\vec{v}_1, \dots, \vec{v}_k$ are pairwise orthogonal vectors in \mathbb{F}_2^n (that is, $\vec{v}_i \cdot \vec{v}_j = 0$ for $i \neq j$) and $\vec{v}_i \cdot \vec{v}_i = 1$ for each $1 \leq i \leq k$, then $\vec{v}_1, \dots, \vec{v}_k$ are linearly independent.

EXERCISE 2.5.4. Recall that a club in Oddtown has an odd number of members, and any two clubs in Oddtown overlap in an even number of members. Use the previous two exercises, applied to the membership vectors, plus the first miracle of linear algebra¹, to deduce:

THEOREM (Oddtown Theorem). *Oddtown of population N can have at most N clubs.*

¹The justification for the first miracle given in the exercises in the previous section works just as well over \mathbb{F}_2 or any other field as it does over \mathbb{R} !

2.6. More fields and isotropic vectors.

EXERCISE 2.6.1. How many vectors are there in \mathbb{F}_p^n ?

EXERCISE 2.6.2.

- A. Find the multiplicative inverse of 2 in $\mathbb{F}_3, \mathbb{F}_5, \mathbb{F}_7, \mathbb{F}_{11}, \dots$
- B. Find the multiplicative inverse of 3 in $\mathbb{F}_2, \mathbb{F}_5, \mathbb{F}_7, \mathbb{F}_{11}, \dots$
- C. And so on...

EXERCISE 2.6.3. *Long division* says that if a and b be two integers then there are unique integers q and r such that

- (1) $a = qb + r$, and
- (2) $0 \leq r < |b|$.

- A. Compute q and r for some examples to see that this is really just long division (r stands for remainder!).

The Euclidean Algorithm says that if a and b are integers then there are integers s and t such that

$$sa + tb = \gcd(a, b)$$

where \gcd denotes the greatest common divisor. To compute them, we use long division as above to write

$$\begin{aligned} a &= q_0b + r_0 \\ b &= q_1r_0 + r_1 \\ r_0 &= q_2r_1 + r_2 \\ r_1 &= q_3r_2 + r_3 \\ &\dots = \dots \end{aligned}$$

the last non-zero remainder is the \gcd , and working backwards lets you find s and t .

- B. Carry out the above algorithm to compute the \gcd of 782 and 255, and express it as $s * 782 + t * 255$.
- C. Explain how the Euclidean algorithm can be used to compute multiplicative inverses in \mathbb{F}_p .
- D. \square Implement the Euclidean algorithm on a computer and use it to find multiplicative inverses in \mathbb{F}_p .

EXERCISE 2.6.4. Describe all isotropic vectors in \mathbb{C}^2 .

EXERCISE 2.6.5.

- A. For \mathbb{K} a field, show that \mathbb{K}^2 has a non-zero isotropic vector if and only if there is an element $k \in \mathbb{K}$ such that $k^2 = -1$. (Here -1 means the unique element in \mathbb{K} such that $1 + (-1) = 0$).
- B. (\square may help!) Experiment until you come up with a simple condition that determines for which prime numbers p the field \mathbb{F}_p satisfies the condition in **A**.

3. Challenge problems

Remember these are supposed to be hard! In particular, don't feel bad if you spend hours thinking about one and don't solve it, but also don't let this warning stop you from trying!

EXERCISE (Prime-exponent polynomials). Given any non-zero real polynomial $f(x)$, prove that there exists a non-zero real polynomial $g(x)$ such that $f(x)g(x)$ is a prime-exponent polynomial (that is, every exponent is prime number, like $x^{23} + 10x^7 - 0.2x^3 + \pi x^2$.)

EXERCISE (Thirteen weights). I have thirteen dumbbells (with integer weights) such that if I take any one of them away, I can split the remaining 12 up into two groups of six such that each group has the same total weight. Show that all thirteen weights are the same.

EXERCISE (Integer-length sides). A large rectangle is divided up into smaller rectangles (with sides parallel to the original rectangle). Suppose each of the smaller rectangles has at least one side of integer length. Show that the same is true of the larger rectangle.

Polynomials, permutations, determinants, and finding eigenvalues

1. Definitions

1.1. Polynomials.

DEFINITION 1.1.1. A *polynomial in the variable x with coefficients in a field \mathbb{K}* is a sum of the form

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$$

where the a_i are elements of \mathbb{K} . For example $\pi x^3 + \sqrt{2}x^2 + 5$ is a polynomial in the variable x with coefficients in \mathbb{R} . We can multiply and add polynomials in the “obvious” way.

- (1) The *degree* of a polynomial is the highest power of the variable appearing with a non-zero coefficient. The *leading coefficient* is the coefficient of this highest power. For example, $\pi x^3 + \sqrt{2}x^2 + 5$ has degree 3 and leading coefficient π . A polynomial is *monic* if its leading coefficient is 1
- (2) We say $f(x)|g(x)$ (the symbol $|$ means divides) if there is another polynomial $q(x)$ such that $g(x) = q(x)f(x)$.
- (3) A polynomial is *prime* if it is
 - (a) monic, and
 - (b) irreducible: if $f(x) = a(x)b(x)$ then one of $a(x)$ or $b(x)$ is constant (degree 0) and the other has the same degree as $f(x)$.

EXAMPLE 1.1.2. If $f(x)$ is a polynomial with coefficients in \mathbb{K} , we can plug in an element $k \in \mathbb{K}$ to get a new element $f(k)$ in \mathbb{K} . For example, if $f(x) = \pi x^3 + \sqrt{2}x^2 + 5$ then $f(2) = 8\pi + 4\sqrt{2} + 5$.

FACT 1.1.3 (Prime factorization of polynomials). Every polynomial $f(x)$ with coefficients in \mathbb{K} has a *factorization*

$$f(x) = cp_1(x)^{a_1} p_2(x)^{a_2} \dots p_m(x)^{a_m}$$

where the p_i are prime polynomials, the a_i are positive integers, and c is a constant. The factorization is uniquely determined up to reordering the prime polynomials p_i .

FACT 1.1.4 (The Fundamental Theorem of Algebra). A polynomial with coefficients in \mathbb{C} is prime if and only if it is monic of degree one, i.e. of the form $x - a$ for a a complex number. In particular, the prime factorization in this case is equivalent to the statement that any polynomial with coefficients in \mathbb{C} factors as a product of linear (degree one) terms.

1.2. Permutations.

DEFINITION 1.2.1. A *permutation* is a way to rearrange a set of things. In math, we usually number the things $1, 2, \dots, n$, then think of a permutation as a way to rearrange these numbers. In math-speak, a permutation is a one-to-one function

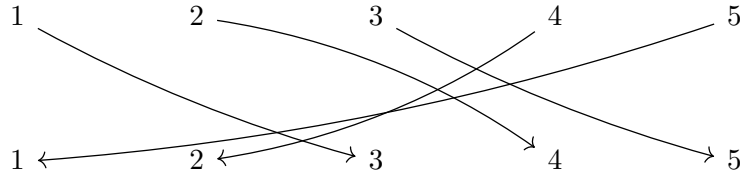
$$\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}.$$

We write S_n for the collection of all permutations of $\{1, \dots, n\}$ (also known as the symmetric group on n elements).

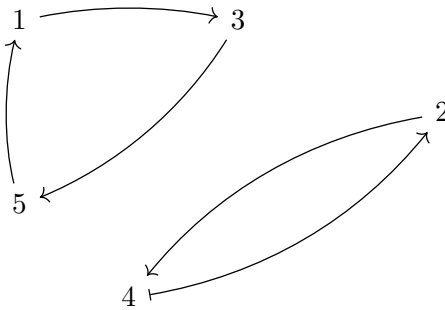
EXAMPLE 1.2.2. Here are some representations of the permutation defined by

$$\sigma(1) = 3, \sigma(2) = 4, \sigma(3) = 5, \sigma(4) = 2, \sigma(5) = 1.$$

- Braid representation



- Graph representation



- Cycle notation

$$\sigma = (1\ 3\ 5)(2\ 4)$$

This is read as “1 goes to 3 goes to 5 goes back to 1” and “2 goes to 4 goes back to 2” – i.e. you loop back at the end! This is the most compact notation so we usually use it.

EXAMPLE 1.2.3. Sometimes we omit cycles of length one when n is understood from context; for example, if $\sigma \in S_5$ is the permutation $\sigma = (1\ 2)(3)(4)(5)$, we might shorten this by writing instead just $\sigma = (1\ 2)$.

DEFINITION 1.2.4 (Composition and inverses).

- (1) If σ and τ are both permutations of $\{1, 2, \dots, n\}$, then $\sigma\tau$ is the permutation $\sigma \circ \tau$, i.e. $\sigma\tau(j) = \sigma(\tau(j))$. In terms of braid diagrams, we stack the diagram for σ below the diagram for τ then straighten out the arrows.
- (2) We write Id for the identity permutation $\text{Id}(j) = j$.
- (3) If σ is a permutation of $\{1, 2, \dots, n\}$, then σ^{-1} is the inverse function. In the braid diagram, it is obtained by reversing the arrows; we have $\sigma\sigma^{-1} = \sigma^{-1}\sigma = \text{Id}$.

DEFINITION 1.2.5 (Inversions and sign). For σ a permutation of $\{1, \dots, n\}$:

- An *inversion* of σ is a pair of numbers number $i, j \in \{1, 2, \dots, n\}$ such that $i < j$ and $\sigma(i) > \sigma(j)$. Visually, σ has an inversion whenever two arrows cross in the braid diagram.
- The *sign* of σ is

$$\text{sgn}(\sigma) = (-1)^{\text{Number of inversions in } \sigma}$$

EXAMPLE 1.2.6. $(135)(24)$ inverts the following pairs: 1 and 4, 1 and 5, 2 and 4, 2 and 5, 3 and 4, 3 and 5, and 4 and 5. Thus, it has seven inversions – this can be seen quickly by counting the intersections in the braid diagram above – and $\text{sgn}(\sigma) = (-1)^7 = -1$.

FACT 1.2.7 (Signs multiply). $\text{sgn}(\sigma\tau) = \text{sgn}(\sigma)\text{sgn}(\tau)$.

1.3. Determinants.

DEFINITION 1.3.1. The *determinant* of an $n \times n$ matrix $A = (a_{ij})$ with entries in a field \mathbb{K} is

$$\det A = \sum_{\sigma \in S_n} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i\sigma(i)} \right) \quad (\text{this is an element of } \mathbb{K}!).$$

FACT 1.3.2 (Determinants multiply). $\det(AB) = \det(A)\det(B)$.

FACT 1.3.3 (Determinants and volume). The absolute value of the determinant of an $n \times n$ matrix A with real entries is the n -dimensional volume¹ of the n -dimensional parallelogram framed by the columns of A (considered as vectors in \mathbb{R}^n).

DEFINITION 1.3.4. An $n \times n$ matrix $A = (a_{ij})$ is *upper-triangular* if $a_{ij} = 0$ when $i > j$. In other words, A is of the form

$$\begin{bmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & * \end{bmatrix}$$

where $*$ denotes an arbitrary entry.

FACT 1.3.5. $\det(A)$ can be computed by the following algorithm:

- (1) Use elementary row operations to transform A into an upper triangular matrix U .
- (2) Multiply together the diagonal entries of U to get a number α (in fact, this number is just the determinant of U).
- (3) We then have

$$\det A = (-1)^{\text{number of swaps}} \left(\prod_i \frac{1}{\lambda_i} \right) \alpha$$

Here the number of swaps means the number of times the elementary row operation of swapping two rows was applied to get from A to U , and the λ_i are the scaling factors that occurred in each of the second type of elementary row operation that was used (where a row is replaced with a non-zero multiple of itself). Note that the third type of row operation, adding a multiple of one row to another, doesn't show up at all in this formula!

In fact, in addition to using elementary row operations, we can also use elementary column operations, which are defined analogously but using columns instead of rows. They effect the determinant in the same way. This added flexibility can be helpful in simplifying computations!

EXAMPLE 1.3.6. We will use the algorithm to compute

$$\det \begin{bmatrix} 4 & 2 \\ 5 & 3 \end{bmatrix}.$$

We can use the row operations

$$\begin{bmatrix} 4 & 2 \\ 5 & 3 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \xrightarrow{R_2-3R_1} \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix} \xrightarrow{R_1+2R_2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus $\alpha = (-1)(1) = -1$, we swapped rows once, and we scaled one row by $1/2$, so

$$\det A = (-1)^1 \left(\frac{1}{1/2} \right) (-1) = 2.$$

¹1-dimensional volume = length, 2-dimensional volume = area, 3-dimensional volume = volume, ...

A quicker way to do it, using column operations too, would be:

$$\begin{bmatrix} 4 & 2 \\ 5 & 3 \end{bmatrix} \xrightarrow{C_1 - 2C_2} \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} -1 & 3 \\ 0 & 2 \end{bmatrix}$$

Thus $\alpha = (-1)(2) = -2$, and we did one swap and no scalings, so

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

FACT 1.3.7 (Determinants, rank, and invertibility). The following are equivalent:

- (1) $\det A \neq 0$
- (2) A is invertible
- (3) $\text{rank} A = n$
- (4) $\ker A = \{0\}$.

DEFINITION 1.3.8. If A is an $n \times n$ matrix, the *characteristic polynomial* of A is

$$f_A(x) := \det(xI_n - A)$$

FACT 1.3.9. The roots λ of $f_A(x)$ are exactly the eigenvalues of A .

1.4. Bases and the matrix of a linear transformation.

DEFINITION 1.4.1. A set of vectors $\vec{f}_1, \dots, \vec{f}_m$ in \mathbb{K}^n is a *basis* if every vector $\vec{v} \in \mathbb{K}^n$ can be written as a linear combination

$$\vec{v} = a_1\vec{f}_1 + a_2\vec{f}_2 + \dots + a_n\vec{f}_m$$

in exactly one way. For any \vec{v} we remember the coefficients a_i in the linear combination as a vector by writing

$$[\vec{v}]_{\vec{f}_\bullet} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}.$$

EXAMPLE 1.4.2. The *standard basis* for \mathbb{K}^n is

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Indeed, this is a basis because for

$$\vec{v} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \text{ we have } \vec{v} = a_1\vec{e}_1 + a_2\vec{e}_2 + \dots + a_n\vec{e}_n$$

and this is the only way to write \vec{v} as a linear combination of the \vec{e}_\bullet . We can rewrite this by

$$[\vec{v}]_{\vec{e}_\bullet} = \vec{v}.$$

FACT 1.4.3 (any n linearly independent vectors are a basis for \mathbb{R}^n). $\vec{f}_1, \dots, \vec{f}_m$ are a basis for \mathbb{K}^n if and only if $m = n$ and they are linearly independent.

DEFINITION 1.4.4. If $\vec{f}_1, \dots, \vec{f}_m$ is a basis for \mathbb{K}^n , and $T : \mathbb{K}^n \rightarrow \mathbb{K}^n$ is a linear transformation, we consider the $n \times n$ matrix whose j th column is $[T(\vec{f}_j)]_{\vec{f}_\bullet}$:

$$[T]_{\vec{f}_\bullet} = \left[[T(\vec{f}_1)]_{\vec{f}_\bullet} \quad [T(\vec{f}_2)]_{\vec{f}_\bullet} \quad \dots \quad [T(\vec{f}_n)]_{\vec{f}_\bullet} \right].$$

If A is an $n \times n$ matrix, we will sometimes write $[A]_{\vec{f}_\bullet}$; in this case A should be interpreted as the linear transformation $\mathbb{K}^n \rightarrow \mathbb{K}^n$ sending \vec{v} to $A\vec{v}$.

FACT 1.4.5. For any vector \vec{v} ,

$$[T(\vec{v})]_{\vec{f}_\bullet} = [T]_{\vec{f}_\bullet} [\vec{v}]_{\vec{f}_\bullet}.$$

EXAMPLE 1.4.6. The *identity linear transformation* Id of \mathbb{K}^n is defined by $\text{Id}(\vec{v}) = \vec{v}$ for every vector \vec{v} . Thus, for any basis \vec{f}_\bullet of \mathbb{K}^n ,

$$[\text{Id}]_{\vec{f}_\bullet} = I_n.$$

DEFINITION 1.4.7. For \vec{f}_\bullet and \vec{g}_\bullet two bases of \mathbb{K}^n , the *change of basis matrix* $C_{\vec{f}_\bullet, \vec{g}_\bullet}$ is the matrix whose j th column is $[\vec{f}_j]_{\vec{g}_\bullet}$:

$$C_{\vec{f}_\bullet, \vec{g}_\bullet} = \left[[\vec{f}_1]_{\vec{g}_\bullet} \quad [\vec{f}_2]_{\vec{g}_\bullet} \quad \dots \quad [\vec{f}_n]_{\vec{g}_\bullet} \right]$$

FACT 1.4.8. For \vec{f}_\bullet and \vec{g}_\bullet two bases of \mathbb{K}^n :

(1) $C_{\vec{g}_\bullet, \vec{f}_\bullet} = C_{\vec{f}_\bullet, \vec{g}_\bullet}^{-1}$

(2) If $T : \mathbb{K}^n \rightarrow \mathbb{K}^n$ is a linear transformation, then

$$[T]_{\vec{g}_\bullet} = C_{\vec{f}_\bullet, \vec{g}_\bullet} [T]_{\vec{f}_\bullet} C_{\vec{f}_\bullet, \vec{g}_\bullet}^{-1}.$$

FACT 1.4.9 (Characteristic polynomial of a linear transformation). If $T : \mathbb{K}^n \rightarrow \mathbb{K}^n$ is a linear transformation, then the characteristic polynomial of the matrix $[T]_{\vec{f}_\bullet}$ is the same no matter which basis \vec{f}_\bullet we choose. We call it the characteristic polynomial of T and write it as $f_T(x)$.

DEFINITION 1.4.10. A linear transformation $T : \mathbb{K}^n \rightarrow \mathbb{K}^n$ is called *diagonalizable* if there is a basis of \mathbb{K}^n consisting of eigenvectors for T .

FACT 1.4.11 (First diagonalizability criterion). If $T : \mathbb{K}^n \rightarrow \mathbb{K}^n$ is a linear transformation and $f_T(x)$ has n distinct roots, then T is diagonalizable.

2. Problems

2.1. Polynomials.

EXERCISE 2.1.1. Let $f(x) = c_n x^n + c_{n-1} x^{n-1} + \dots + c_0$ be a polynomial with coefficients in \mathbb{K} .

- A. If $a \in \mathbb{K}$ is such that $(x - a) \mid f(x)$, show $f(a) = 0$.
- B. If $a \in \mathbb{K}$ is such that $f(a) = 0$, show $(x - a) \mid f(x)$ as follows:
 - i. Explain why

$$f(x) = f(x) - f(a) = c_n(x^n - a^n) + c_{n-1}(x^{n-1} - a^{n-1}) + \dots + c_1(x - a).$$

- ii. Compute $\frac{x^k - a^k}{x - a}$.
 - iii. Conclude.
- C. By A. and B., for $a \in \mathbb{K}$, $f(a) = 0$ is equivalent to $(x - a) \mid f(x)$. In either case, we say a is a *zero or root of $f(x)$* . Explain why the polynomial $f(x)$ of degree n has at most n distinct roots.

EXERCISE 2.1.2.

- A. Explain why $f(x) = x^2 + 1$, thought of as a polynomial with coefficients in \mathbb{R} , is prime. How does it factor as a polynomial with coefficients in \mathbb{C} ?
- B. Find the prime factorization of $x^4 - 2x^3 + 2x^2 - 2x + 1$, thought of as a polynomial with coefficients in \mathbb{R} .
Hint: to get started, guess and check with some simple numbers to find a root.

EXERCISE 2.1.3 (Lagrangian interpolation). Let t_0, \dots, t_n be distinct elements of a field \mathbb{K} .

- A. Show that if $f(x)$ and $g(x)$ are two polynomials of degree at most n with coefficients in \mathbb{K} such that $f(t_i) = g(t_i)$ for all $0 \leq i \leq n$ then $f(x) = g(x)$.
Hint: Use the result of Exercise 2.1.1-C.
- B. If a_0, \dots, a_n are arbitrary elements of \mathbb{K} , find a polynomial f of degree at most n such that $f(t_i) = a_i$ for all $0 \leq i \leq n$.
- C. \star One consequence of part A. is that, if \mathbb{K} is infinite, then we don't lose any information if we think of polynomials as being functions $\mathbb{K} \rightarrow \mathbb{K}$ instead of as formal sums of powers of x . Make this statement precise, then explain how it breaks down if \mathbb{K} is a finite field.

EXERCISE 2.1.4. *Polynomial long division* says that if $a(x)$ and $b(x)$ are two non-zero polynomials with coefficients in a field \mathbb{K} then there are unique polynomials $q(x)$ and $r(x)$ with coefficients in \mathbb{K} such that

- (1) $a(x) = q(x)b(x) + r(x)$, and
- (2) $\deg r(x) < \deg b(x)$.

- A. Compute $q(x)$ and $r(x)$ for some examples (this works just like long division for integers!)

The Polynomial Euclidean Algorithm Says that if $a(x)$ and $b(x)$ are polynomials with coefficients in \mathbb{K} then there are polynomials $s(x)$ and $t(x)$ with coefficients in \mathbb{K} such that

$$s(x)a(x) + t(x)b(x) = \gcd(a(x), b(x))$$

We can find $s(x)$ and $t(x)$ by using long division as above to write

$$\begin{aligned} a(x) &= q_0(x)b(x) + r_0(x) \\ b(x) &= q_1(x)r_0(x) + r_1(x) \\ r_0(x) &= q_2(x)r_1(x) + r_2(x) \\ r_1(x) &= q_3(x)r_2(x) + r_3(x) \\ &\dots\dots \end{aligned}$$

The last non-zero remainder is the gcd, and we get $s(x)$ and $t(x)$ by working backwards – just like the Euclidean algorithm for integers (see Exercise 2.6.3 in Week Two).

B. Carry out the polynomial Euclidean algorithm to compute the gcd of

$$a(x) = x^5 + x^4 + x^3 + 2x^2 + 1 \text{ and } b(x) = x^8 + x^7 + x^5 + x^3 + x^2 + 1.$$

and express it as $s(x)a(x) + t(x)b(x)$.

C. Implement the polynomial Euclidean algorithm in CoCalc.

EXERCISE 2.1.5.

A. Write a computer program to count all of the prime polynomials of degree d with coefficients in \mathbb{F}_p .

B. \star Can you predict a formula for this count? Can you prove it?

EXERCISE 2.1.6 (Growing fields). \star

A. If $f(x)$ is a degree $d \geq 1$ polynomial with coefficients in \mathbb{K} , explain how to do clock arithmetic on the set of polynomials of degree $< d$ with coefficients in \mathbb{K} by using polynomial long division (here the polynomial $f(x)$ is playing the role of a fixed integer n when we do clock arithmetic modulo n).

B. If f is a prime polynomial of degree d , explain why clock arithmetic modulo f makes the set of polynomials of degree $< d$ into a field.

C. Explain why the resulting field when $\mathbb{K} = \mathbb{R}$ and $f(x) = x^2 + 1$ is “the same” as \mathbb{C} .

D. Construct fields of order $2^2, 2^3, 2^4, \dots, 3^2, 3^3, \dots$ (go as far as you can!)

E. Is there a field of order p^n for every prime p and integer n ? (see Exercise 2.1.5).

2.2. Permutations.

EXERCISE 2.2.1. Write down some examples of permutations using their braid diagrams, graph representation, and cycle notation.

EXERCISE 2.2.2. Compute $\sigma\tau$ and $\tau\sigma$, σ^{-1} , and τ^{-1} for

$$\sigma = (1\ 3\ 5)(2\ 4) \text{ and } \tau = (1\ 2\ 3\ 4\ 5).$$

EXERCISE 2.2.3 (Cycle statistics). The number of k -cycles in a permutation σ is the number of cycles of length k when we write it in cycle notation. For example, the permutation $(1\ 3\ 5)(2\ 4)$ of $\{1, 2, \dots, 5\}$ contains one 3-cycle and one 2-cycle while the permutation $(1\ 4)(2\ 3)(5)$ of $\{1, 2, \dots, 5\}$ contains two 2-cycles and one 1-cycle.

Hint: \square If you have troubles getting started on any of the questions below, try using the Permutations package in CoCalc to experiment!

- A. How many total permutations are there of n things?
- B. How many permutations of n things consist of exactly one n -cycle?
- C. What is the probability that a randomly chosen permutation of n things consists of exactly one n -cycle?
- D. If I shuffle a deck of cards really well, how many cards would you expect, on average, to end up in the same spot in the deck (by which I mean same distance from the top) as they were before I shuffled?
- E. For $k \leq n$, what is the average number of k -cycles in a randomly chosen permutation of n things?
- F. What is the average number of cycles in a length n permutation? Approximate your answer as $n \rightarrow \infty$ using a simple function of n .
- G. (\star) Use your answer to compute the probability that a randomly chosen permutation of $\{1, 2, \dots, n\}$ contains a cycle of length greater than $n/2$. What is the limit of this probability as $n \rightarrow \infty$?

EXERCISE 2.2.4. A *transposition* is a permutation that consists of a single 2-cycle (e.g. $(2\ 5)$).

- A. Show that all transpositions are odd. If you can't come up with a mathematical argument, try writing a program in CoCalc that generates a random transposition and computes its sign to check this.
- B. Explain how to write a k -cycle as a product of $k - 1$ transpositions.
- C. If σ consists of a single k -cycle, what is its sign?
- D. Give a simple expression for the sign of any permutation σ in terms of the cycles it contains.

EXERCISE 2.2.5. \star Use the description of $\sigma\tau$ in terms of braid diagrams to explain Fact 1.2.7.

EXERCISE 2.2.6. Suppose σ and τ are permutations and in cycle notation

$$\tau = (a_1 \dots a_{k_1})(b_1 \dots b_{k_2}) \dots$$

- A. Show

$$\sigma\tau\sigma^{-1} = (\sigma(a_1) \dots \sigma(a_{k_1}))(\sigma(b_1) \dots \sigma(b_{k_2})) \dots$$

- B. Explain how to visualize this as a relabeling of the vertices in the graph representation.

EXERCISE 2.2.7. The *order* of a permutation σ is the smallest positive integer k such that $\sigma^k = \mathbf{1}$ – here $\mathbf{1}$ means the *identity permutation* $\mathbf{1}(i) = i$ for all $1 \leq i \leq n$, and σ^k means

$$\underbrace{\sigma\sigma\dots\sigma}_{k \text{ times}}$$

Describe the order of a permutation in terms of its cycles.

Hint: \square If you get stuck, experiment in CoCalc to come up with a simple answer and then see if you can explain it!

2.3. Determinants and characteristic polynomials.

EXERCISE 2.3.1. Use the definition of the determinant to compute

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

B. $\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

EXERCISE 2.3.2 (\star). Draw a picture that illustrates Fact 1.3.3 when $n = 2$: if $\vec{v}, \vec{w} \in \mathbb{R}^2$, then

$$|\det [\vec{v} \ \vec{w}]|$$

is the area of the parallelogram with vertices $\vec{0}$, \vec{v} , \vec{w} , and $\vec{v} + \vec{w}$.

EXERCISE 2.3.3. Show $\det A^t = \det A$.

EXERCISE 2.3.4.

- (1) Use the formula for the determinant of a 2×2 matrix to show that if A and B are 2×2 matrices then $\det(AB) = (\det A)(\det B)$. (This is Fact 1.3.2 in the case $n = 2$).
- (2) \star Using the definition of the determinant and multiplicativity of sign (Fact 1.2.7), show that Fact 1.3.2 holds in general.

EXERCISE 2.3.5. Use the algorithm in Fact 1.3.5 to compute

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

EXERCISE 2.3.6. Use the definition of the determinant to compute $\det A$ when $A = (a_{ij})$ is upper-triangular.

Hint: First explain why any non-identity permutation must have an inversion.

EXERCISE 2.3.7. \star Use the previous exercise, Week 2 - Exercise 2.4.6, and multiplicativity of the determinant to explain why the algorithm in Fact 1.3.5 works!

EXERCISE 2.3.8 (The Vandermonde determinant). For $x_1, x_2, \dots, x_n \in \mathbb{K}$, the *Vandermonde matrix* has i, j th entry x_i^{j-1} . In other words, it looks like:

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix}$$

where here we use the convention that any number to the zeroth power is 1.

- \star Compute the determinant of the Vandermonde matrix using row and column operations.
- \star Show that if $x_i = x_j$ for $i \neq j$ then the determinant is zero. Use this to compute the Vandermonde determinant in a different way.

EXERCISE 2.3.9. Compute $\det(M_n)$, where M_n is the $n \times n$ matrix

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

and the blank spots are all zero.

EXERCISE 2.3.10. * What does Fact 1.3.7 have to do with the interpretation of the determinant using volume in Fact 1.3.3?

EXERCISE 2.3.11. * Justify Fact 1.3.7.

Hint: use the algorithm for computing a determinant via elementary row operations and the fact that elementary row operations are given by matrix multiplication.

EXERCISE 2.3.12. Compute the characteristic polynomial of a 2×2 matrix.

EXERCISE 2.3.13. Show that if A is an $n \times n$ matrix then $f_A(x)$ is a degree n polynomial. Use Fact 1.3.9 to deduce that A has at most n eigenvalues.

EXERCISE 2.3.14. Justify Fact 1.3.9 carefully. *Hint:* Use Fact 1.3.7.

EXERCISE 2.3.15. Suppose $\vec{v}_1, \dots, \vec{v}_k$ are nonzero eigenvectors of A with distinct eigenvalues $\lambda_1, \dots, \lambda_k$. Show that they are linearly independent. Use this to give another justification of the fact that a $n \times n$ matrix A has at most n eigenvalues.

EXERCISE 2.3.16. Compute a simple formula for

$$\det \begin{bmatrix} a & b & b & \dots & b \\ b & a & b & \dots & b \\ b & b & a & \dots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \dots & a \end{bmatrix}$$

Hint: Think about eigenvalues and characteristic polynomials.

2.4. Bases and the matrix of a linear transformation.

EXERCISE 2.4.1. Consider the vectors in \mathbb{R}^2

$$\vec{f}_1 = \vec{e}_1 + \vec{e}_2, \vec{f}_2 = \vec{e}_1 - \vec{e}_2$$

- Show \vec{f}_1 and \vec{f}_2 form a basis for \mathbb{R}^2 .
- Compute the change of basis matrices $C_{\vec{e}_\bullet, \vec{f}_\bullet}$ and $C_{\vec{f}_\bullet, \vec{e}_\bullet}$.
- Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear transformation given by counter-clockwise rotation by 90° ($\pi/2$ radians) around the origin. Compute $[T]_{\vec{e}_\bullet}$ and $[T]_{\vec{f}_\bullet}$, then explain the result geometrically.
- If instead of \mathbb{R}^2 we work in \mathbb{F}_p^2 , are the vectors \vec{f}_1 and \vec{f}_2 still linearly independent? *Hint:* Your answer should depend on p .

EXERCISE 2.4.2. If σ is a permutation of $\{1, \dots, n\}$, consider the matrix A_σ with entries

$$(A_\sigma)_{ij} = \begin{cases} 1 & \text{if } i = \sigma(j) \\ 0 & \text{otherwise.} \end{cases}$$

- Show $A_\sigma \vec{e}_j = \vec{e}_{\sigma(j)}$.
- Show $A_\sigma A_\tau = A_{\sigma\tau}$.
- Show $A_{\text{Id}} = I_n$.
- Show $A_{\sigma^{-1}} = A_\sigma^{-1}$.
- Show $\det A_\sigma = \text{sgn}(\sigma)$.
- If $\vec{f}_i = \vec{e}_{\sigma(i)}$, show by direct computation that

$$[A_\tau]_{\vec{f}_\bullet} = A_{\sigma\tau\sigma^{-1}} = A_\sigma A_\tau A_{\sigma^{-1}} = A_\sigma A_\tau A_\sigma^{-1}.$$

Hint: See Exercise 2.2.6.

- Compare the previous part with Fact 1.4.8.

EXERCISE 2.4.3.

- Find a basis for \mathbb{R}^2 consisting of eigenvectors for the Fibonacci matrix

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

Hint: See Week 2 - Exercise 2.3.1.

- For the linear transformation $T(\vec{v}) = F\vec{v}$, compute the matrix of T in the basis from A.

EXERCISE 2.4.4. Suppose $T : \mathbb{K}^\times \rightarrow \mathbb{K}^\times$ is a linear transformation. If \vec{f}_\bullet is a basis consisting of eigenvectors for T , i.e. $T(\vec{f}_i) = \lambda_i \vec{f}_i$ with $\lambda_i \in \mathbb{K}$, write down the matrix $[T]_{\vec{f}_\bullet}$. This should explain why the term “diagonalizable” is used in Definition 1.4.10.

EXERCISE 2.4.5. Justify Fact 1.4.3

EXERCISE 2.4.6. Justify Fact 1.4.8

EXERCISE 2.4.7. Justify Fact 1.4.9

EXERCISE 2.4.8.

- Justify Fact 1.4.11.
- Give an example of a linear transformation T that is diagonalizable but the characteristic polynomial $f_T(x)$ has repeated roots.

3. Challenge problems

Remember these are supposed to be hard! In particular, don't feel bad if you spend hours thinking about one and don't solve it, but also don't let this warning stop you from trying!

EXERCISE (Polynomial hidden treasure). You land on a deserted island, looking for pirate treasure. Sure enough, there is an old, weathered treasure map with fiendishly complicated instructions written by a pirate who was a mathematician before following the lure of the sea. You realize that the treasure's location is encoded in the roots of some polynomial which starts

$$x^{100} + 5x^{99} + 13x^{98} + \dots$$

Unfortunately, the rest of the polynomial is lost to the sands of time. Show that not every root of this polynomial is real.

EXERCISE (Prison hats). I'm the warden of a prison and my budget just got slashed. I've got a *huge* overpopulation problem – it's time to take some drastic measures. But it's not like I can just execute 50 prisoners... well, not without giving them a "shot" first. Here's the game: I'll number the prisoners 1 to 50, and then put them in a line. In another room, I'll arrange 50 hats in a row on a table, and then randomly put slips of paper with the numbers 1 to 50 inside, one slip per hat. The prisoners will enter one at a time, and each prisoner can look at 25 hats to try to find the one with their number in it (but no moving the numbers or hats around!). If every single one of the prisoners finds their number, then they're all free! But if any one of them doesn't find their number, then it's the gallows for the lot of them. They can have an hour before the game starts to come up with a strategy, but once the game starts there's no communication between prisoners.

Wow, this warden is a jerk! Good thing some of the prisoners were part of a pre-REU program when they were younger... come up with a strategy that gives the prisoners a better than 30% chance of winning their freedom.

EXERCISE (Sam Lloyd's 15 puzzle). Arrange the numbers $1, \dots, 15$ on a 4 by 4 grid together with a blank square. You may alter the board by transposing the blank square with an adjacent square. To solve the puzzle, you must get the numbers into the order

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & \end{array}$$

Show that a random starting arrangement has a 50% chance of being feasible (having a solution).

EXERCISE (An independent curve). Find a curve in \mathbb{R}^n such that any n distinct points on the curve are linearly independent. Here by a curve we mean a parameterized curve – that is, a function from \mathbb{R} to \mathbb{R}^n

$$\gamma(t) = [x_1(t) \quad x_2(t) \quad \dots \quad x_n(t)]^t.$$

EXERCISE (gcd matrix). Let $D = (d_{ij})$ be the $n \times n$ matrix where $d_{ij} = \gcd(i, j)$. Then,

$$\det(D) = \phi(1)\phi(2)\dots\phi(n)$$

where ϕ is Euler's ϕ function

$$\phi(m) = \#\{k \mid 1 \leq k \leq m, \gcd(k, m) = 1.\}$$