Linear algebra notes

C. Eric Overton-Walker

These notes are written for a 24 day summer course (1.5 hour lectures). I wrote the syllabus to include two exams and numerous in-class quizzes, so that leaves 22 days of roughly 1 to 1.5 hour lectures. I aim to complete about 1 section of the textbook [LDP06] each lecture, though if I get ahead I'll bank that time. I'm to cover chapters 1 through 6 and these notes do so (skipping a few sections), but if at the end of the semester I'm not done then that's okay; we'll just test over what we completed.

We freely use [LDP06] for course structure and most examples. Any errors are my own.

Contents

How to think about this class . . . 1
Tips for success . . . 2
Day 01 of 24 – §1.1 Systems of linear equations . . . 4
Day 02 of 24 – §1.2 Row echelon form . . . 9
Day 03 of 24 – §1.3 Matrix arithmetic . . . 12
Day 04 of 24 – §1.4 Matrix algebra . . . 18
Day 05 of 24 – §1.5 Elementary matrices . . . 23
Day 06 of 24 – §2.1 The determinant of a matrix . . . 26
Day 07 of 24 – §2.2 Properties of determinants . . . 29
Day 08 of 24 – §3.1 Vector spaces . . . 33
Day 09 of 24 – §3.2 Subspaces . . . 36
Day 10 of 24 – §3.3 Linear indepedence . . . 42
Day 11 of 24 – §3.4 Basis and dimension . . . 47
Day 12 of 24 – §3.5 Change of basis . . . 49
Day 13 of 24 – §3.6 Row space and column space . . . 53
Day 14 of 24 – §4.1 Linear transformations . . . 58
Day 15 of 24 – §4.2 Matrix representations of transformations . . . 62
Day 16 of 24 – §4.3 Similarity . . . 65
Day 17 of 24 – §5.1 The scalar product . . . 69
Day 18 of 24 – §5.2 Orthogonal subspaces . . . 74
Day 19 of 24 – §5.4 Inner product spaces . . . 79
Day 20 of 24 – §5.5 Orthonormal sets . . . 83
Day 21 of 24 – §6.1 Eigenvalues and eigenvectors . . . 85
Day 22 of 24 – §6.3 Diagonalization . . . 89
Appendix: Characterizations of nonsingular matrices . . . 93
Appendix: Definitions . . . 94
Appendix: Homeworks . . . 97
Appendix: Notation . . . 97
Appendix: Reference . . . 98

How to think about this class

Welcome to the most important math class you'll ever take.

I believe I can say that with confidence, no matter your background and no matter your intentions for the future. From the purest mathematician to the most applied scientist, linear algebra is a critical pillar of your education and an absolutely necessary tool to have in your toolbox.

As I currently write this (April 2023), it's been close to a decade since I've really thought about any sort of applied mathematics. Yet I use linear algebra basically every day. The same is true for the colleague I share an office with, who is at the other extreme: working in a very applied domain, in tandem with the data science department. Understanding linear algebra is necessary to both of our jobs.

What's more, it requires a deep understanding; it's not enough to solely be able to compute examples. Everyone needs to know this tool like the back of their hand, and that means not only a fluency in working with examples but an understanding of why the tools work. We'll see theory in this class, do some proofs together, and I'll expect you to do the same in submitted work (albeit in a smaller, controlled environment – so don't worry if you're still new to proofs).

Why is linear algebra so important to us, and why specifically the theory? It's because this will be essentially the only math class you ever take which is anywhere close to being a "solved subject." People can and do work in active research in every math class you've taken and every math class you will take. Often their work looks dramatically different to what you cover during class, but they're still pushing the needle. Moreover, it's often very easy, even as a student in those classes, to ask a question, innocent-seeming though it might be, and fully in the language of the class itself – no esoteric terms – for which the answer requires weeks to unpack, or a graduate-level education, or is in fact still an open problem. For instance:

- Are there infinitely many primes separated by 2?
- Given an elementary function, calculate its antiderivative in terms of elementary functions.
- What are all the roots of the function $\zeta(s) = 1 + 1/2^s + 1/3^s + 1/4^s + 1/5^s + \cdots$?
- Are there any sets which are bigger than $\mathbf{N}$ but smaller than $\mathbf{R}$?
- Are various natural complexity classes distinct?

But linear algebra, or at least the linear algebra of finite-dimensional vector spaces and their linear transformations (we will learn these words!), is much more thoroughly understood. By the end of the semester, we'll have close to a complete understanding of these topics. And yes, you can still ask questions for which answers aren't known, but I wager it's a lot harder to do – try me, and see if you can!

It's this reason that makes linear algebra so important. It's so well-understood! That means it's an incredibly useful tool. Moreover, if you are handed a really hard problem *outside* of this class, but you can reduce it to a linear algebra problem, then you've made the problem that much more tractable. Whole domains in mathematics are progressed in this way. And of course scientists gain the same benefits: any question which can be reduced to a linear system, any function which can be linearly approximated, any concept which yields to linear algebra, is a concept that can be understood. Computers basically *only* do linear algebra, and they do it very, very, very well. A CPU (or GPU) is just a machine that can calculate a bunch of linear algebra problems billions of times a second.

Tips for success □

Because linear algebra is such an important class, it is worth your effort to master it. On top of that, this class is often a student's first exposure to upper level math – the only prerequisite for this course at our university is calculus 1 (not proof-based). So this course often requires an adjustment from the "plug-and-chug" experiences you've had in all your past math classes. Here are some general words of wisdom I wish I had heard (or, that I *did* hear, but wish I had taken to heart).

You'll have to invest a lot of work outside of lectures. That's especially true given the fast paced nature of a summer course, but it's true even for the typically paced 16 week series. Students often feel like they follow along in lecture, which is great! But please don't stop there complacently. The homework is incredibly important in this regard; math is not a spectator sport and only by *doing* do you figure out what you have and have not yet mastered.

The rule of thumb is that you should expect to spend about double to triple the time on material outside of class that you do inside it. We only meet for 36 hours all semester, so expect to spend over 100 hours throughout the summer on linear algebra. This need not all be spent on homework, and please note that that time should be productively spent. It's normal to struggle a bit at first (everyone does – no one comes out of the womb knowing this material!) but if you are staring at a question for 30 minutes without knowing where to start, you should readjust. Ask yourself:

- What information has this problem given me? Write it down.
- What am I being asked to do? What would a correct final answer look like? Write an example. When you are finished, compare your answer to what you wrote here.

- Do I know all the definitions of the words involved? Write them down.
- Are there any theorems which I can use that relate the inputs I am given to the outputs I am seeking? Write them down.
- Have I seen similar problems in the lecture notes or textbook? Pull them up and compare those problems to your current problem; depending on how similar they are, would the solutions also be similar? Write a solution to your problem which mirrors those solutions. Does it work? If not, where must they differ?

This is likely your first exposure to proofs, which can be intimidating (but very rewarding too! They inspire my love of math). In general, following a logical argument can be complicated, but some advice below may make it easier. When you read a theorem, ask yourself:

- What are the hypotheses?
- What is the conclusion?
- What are all the definitions?
- What is an example of an object that meets the hypotheses? What does the conclusion say about the object? Can you work that out by hand to confirm? An example does not a proof make, but are there aspects about your example that feel universal – like they would work for any example? If you're unsure, repeat this bullet point with an example that's as different as you can manage while still meeting the hypotheses.
- Before reading the proof, attempt a proof yourself. After your attempt, compare to the notes or textbook. There are often many ways to prove something, but were your approaches similar? Does the proof do something or address a feature that your attempt did not anticipate?
- When you read the proof, can you follow the order of steps? Ask yourself: if I currently know "X," then we should do "Y," and therefore "Z," ...
- After reading the proof, take an example of an object that meets the hypotheses, and apply every step of the proof to that example. Confirm that after you're done, your example satisfies the conclusion.
- There will never be an extraneous hypothesis in our class (to the best of my knowledge). Where is each hypothesis used in the proof?
- What happens if you remove a hypothesis? What part of the argument fails? Can you construct a specific counterexample without that hypothesis where the conclusion fails?

In addition to the homework, it's very productive to read the textbook, both in advance and after the fact. Before a lecture, read through the corresponding section. At this stage, pay particular attention to definitions and examples. Namely:

- For each definition, can you write down an example without looking at the book?
- Double check the textbook's examples. Confirm that everything they say is an "X" actually is and that all the arithmetic they do is correct. (Often, the textbook chooses to skip some detail checking for expedience or to defer to the exercises – *do this detail checking!*)
- When you get to theorems, take the advice above, but be aware that it's okay if something doesn't click the first time. Focus mostly on the theorem statement and examples. You should attempt a proof and read the textbook's, but it may take time to fully internalize.
- Flag any concepts you're uncomfortable with, and pay particular attention to those during lecture.

After class, read the textbook again with an eye for self-reflection and detail-scrubbing. In particular:

- Did you make any mistakes when you first read?
- Were there nuances pointed out in lecture that didn't come across when you first read?
- Returning to the theorems, how does the lecture's proof compare to the book's, and how do both compare to your first attempt?
- Do you now confidently understand the concepts you flagged? If not, follow up.
- Can you explain the section? Could you produce a lecture to your peers? How about to people unfamiliar with the material? A "yes" here is strong evidence of your mastery (but be honest with yourself!). Put it to the test and lecture to your roommate, pet, rubber duck, ...

Day 01 of 24 – §1.1 Systems of linear equations □

Let's start with a high school level problem.

Example 1. Solve the system of equations

$$\begin{cases} y = \phantom{-\tfrac{1}{2}}2x - 1 \\ y = -\tfrac{1}{2}x + 4 \end{cases}$$

For some notational consistency with future stuff, we'll rewrite:

$$\begin{cases} -2x + y = -1 \\ x + 2y = 8 \end{cases}$$

We know how to do this. Solving the system means finding an ordered pair (x, y) such that when you plug in that point, both equations spit out true statements. It's easy to solve, for instance, by multiplying the bottom equation by 2:

$$\begin{cases} -2x + y = -1 \\ 2x + 4y = 16 \end{cases}$$

combining like terms

$$\begin{aligned} 0x + 5y &= 15 \\ y &= 3, \end{aligned}$$

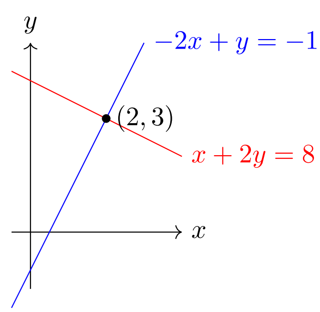
and backsubstituting to solve for x

$$\begin{aligned} x + 2y &= 8 \\ x + 2(3) &= 8 \\ x &= 2. \end{aligned}$$

Thus the solution is $(2, 3)$. You can easily check:

$$\begin{aligned} -2(2) + 3 &= -1 \checkmark \\ 2 + 2(3) &= 8 \checkmark \end{aligned}$$

Remark 2. Recall that, graphically, plotting an equation in x and y is shading all the ordered pairs (x, y) in the plane that make the equation true. Thus if a solution is a point that makes all equations true, pictorially it is a point that lies on both graphs, hence an intersection point. Thus the graphical solution to our example is



What we've just seen is the main content of our course! Of course, we need not restrict to a system of *two* linear equations, nor for that matter systems in the xy-plane.

Definition 3. A **linear equation in n variables** is

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $a_1, \ldots, a_n, b \in \mathbf{R}$ and $x_1, \ldots, x_n$ are indeterminates. A **linear system of m equations** is a collection of m linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad \vdots \qquad\qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

We'll call this an **$m \times n$ system** for short. A **solution to an $m \times n$ system** is a ordered n-tuple $(x_1, x_2, \ldots, x_n)$ that satisfies all the equations (i.e., when you plug in you get true statements).

Example 4. Let's consider two more systems.

$$\begin{cases} 2x_1 + 3x_2 = 1 \\ 4x_1 + 6x_2 = 3 \end{cases} \qquad\qquad \begin{cases} 3x_1 - 2x_2 = 1 \\ 6x_1 - 4x_2 = 2 \end{cases}$$

Solve them:

$$\begin{cases} 4x_1 + 6x_2 = 2 \\ 4x_1 + 6x_2 = 3 \end{cases} \qquad\qquad \begin{cases} 6x_1 - 4x_2 = 2 \\ 6x_1 - 4x_2 = 2 \end{cases}$$

$$0x_1 + 0x_2 = 1 \qquad\qquad 0x_1 + 0x_2 = 0$$

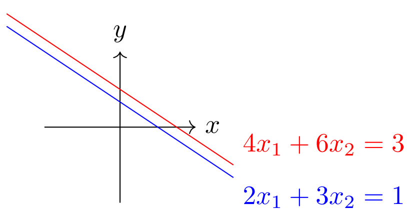
$$0 = 1 \qquad\qquad 0 = 0$$

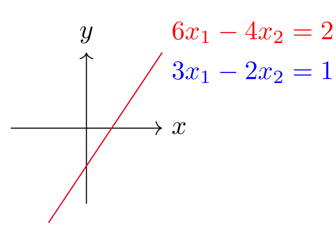
That's weird. On the first, we got a false statement, and on the second, a true statement. The first must have no solutions, since there's no ordered pair (x_1, x_2) which will ever make $0 = 1$. And the second will have infinitely many solutions. As soon as you pick any x_1, you just choose x_2 to be

$$\begin{aligned} 3x_1 - 2x_2 &= 1 \\ -2x_2 &= 1 - 3x_1 \\ x_2 &= \frac{1 - 3x_1}{2}. \end{aligned}$$

So any $(x_1, (1 - 3x_1)/2)$ works. Like $(0, 1/2)$, $(1, -1)$, $(2, -5/2)$, $(3, -4)$, etc.

Graphically, we just solved these two systems:





If solutions are intersections, we can see the first must have 0 solutions, and the second has infinitely many.

Definition 5. An $m \times n$ linear system with no solutions $(x_1, x_2, \ldots, x_n)$ is called **inconsistent**. It's **consistent** otherwise – one or more solutions.

Remark 6. A natural question (maybe *the* natural question): how do we know if a system is inconsistent? And if it does have solutions, how many will there be? We've already seen 0 solutions (inconsistent), 1 solution, and infinitely many.

To answer this question, we need to systematize how to solve $m \times n$ systems, not just 2×2s. So what are the algebra steps we use? What's allowed?

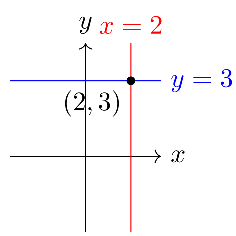
Notice that we get our final answer of (n_1, n_2) when our arithmetic leads us to

$$\begin{cases} x_1 = n_1 \\ x_2 = n_2. \end{cases}$$

At its core, that's exactly what we were doing in **Example 1**. The algebra gave us

$$\begin{cases} y = 3 \\ x = 2, \end{cases}$$

and of course those lines intersect at $(2, 3)$.



Definition 7. Given an $m \times n$ system, any other system which also has n indeterminates and has the exact same set of solutions as the first is called **equivalent** to the first. Note: this is *not equal*! Equal matrices have identical entries, but equivalent just needs the same solutions.

Example 8. From **Example 1**,

$$\begin{cases} -2x + y = -1 \\ x + 2y = 8 \end{cases}$$

and the system

$$\begin{cases} y = 3 \\ x = 2 \end{cases}$$

are equivalent. Again, this is not equal, so we don't use "=." We write

$$\begin{cases} -2x + y = -1 \\ x + 2y = 8 \end{cases} \sim \begin{cases} y = 3 \\ x = 2 \end{cases}$$

So, asking how to solve systems is the same as asking what arithmetic we can do that produces equivalent systems. Here are some obvious things:

I. Swap the order of equations.

Example 9.

$$\begin{cases} x_1 + x_2 = 2 \\ 2x_1 - x_2 = 3 \end{cases} \sim \begin{cases} 2x_1 - x_2 = 3 \\ x_1 + x_2 = 2 \end{cases}$$

II. Multiply (or divide) both sides of an equation by any nonzero number.

Example 10.

$$\begin{cases} 2x_1 - x_2 = 3 \\ x_1 + x_2 = 2 \end{cases} \sim \begin{cases} 2x_1 - x_2 = 3 \\ -2x_1 - 2x_2 = -4 \end{cases}$$

III. Add (or subtract) two equations.

Example 11.

$$\begin{cases} 2x_1 - x_2 = 3 \\ -2x_1 - 2x_2 = -4 \end{cases} \sim \begin{cases} 2x_1 - x_2 = 3 \\ 0x_1 - 3x_2 = -1 \end{cases} \qquad (*)$$

Let's go ahead and finish solving this system. Use **II.** on the second equation by dividing by -3:

$$\sim \begin{cases} 2x_1 - x_2 = 3 \\ 0x_1 + x_2 = 1/3 \end{cases}$$

Use **III.**:

$$\sim \begin{cases} 2x_1 + 0x_2 = 10/3 \\ 0x_1 + x_2 = 1/3 \end{cases}$$

Finally use **II.** by dividing by 2:

$$\sim \begin{cases} x_1 + 0x_2 = 10/6 \\ 0x_1 + x_2 = 1/3 \end{cases}$$

Thus the solution is $(10/6, 1/3)$.

Remark 12. In fact, **I.**, **II.**, and **III.** are *all* you need. In fact, we were basically home free when we got to $(*)$. In fact, any time we're at $(*)$ and the number of equations m and the number of indeterminates n are the same, we have the following.

Definition 13. An $m \times n$ system is said to be **in strict triangular form / upper triangular form** if the ith equation has 0s for all the coefficients up to x_i, but x_i doesn't have a 0 coefficient.

Example 14. Here's some systems in upper triangular form:

$$\begin{cases} 2x_1 - x_2 = 3 \\ 0x_1 - 3x_2 = -1 \end{cases} \qquad (*)$$

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 1 \\ \qquad\quad x_2 - x_3 = 2 \\ \qquad\qquad\quad 2x_3 = 4 \end{cases}$$

$$\begin{cases} 2x_1 - x_2 + 3x_3 - 2x_4 = 1 \\ \qquad\quad x_2 - 2x_3 + 3x_4 = 2 \\ \qquad\qquad\quad 4x_3 + 3x_4 = 3 \\ \qquad\qquad\qquad\quad 4x_4 = 4 \end{cases}$$

The reason we're home free is because you can solve this via *backsubstitution*. Let's demonstrate with the second example:

$$\begin{aligned} 2x_3 &= 4 \\ x_3 &= 2, \\ x_2 - 2 &= 2 \\ x_2 &= 4, \\ 3x_1 + 2(4) + 2 &= 1 \\ 3x_1 &= -9 \\ x_1 &= -3. \end{aligned}$$

The solution is $(-3, 4, 2)$.

There's a more efficient way to convey the information of a linear system:

Definition 15. An $m \times n$ **matrix** (pl. matrices) is an array of numbers

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \left[a_{ij}\right]_{1 \le i \le m, 1 \le j \le n}.$$

When the a_{ij}s are coefficients of a linear system, we call A a **coefficient matrix**. For a given $m \times n$ system

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{cases},$$

we **augment** the coefficient matrix and produce

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

We could also write $[A \mid B]$, if we let B be the $m \times 1$ matrix

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Remark 16. Notice that this conveys the exact same information as the linear system. Therefore, all the same things we can do to linear systems, we can do to matrices. In particular, solving a system given by an augmented matrix $[A \mid B]$ means

I. Swapping two rows.
II. Multiplying (or dividing) rows by a nonzero number.
III. Add (or subtract) two rows.

Example 17. Convert the system

$$\begin{cases} x_1 + 2x_2 + x_3 = 3 \\ 3x_1 - x_2 - 3x_3 = -1 \\ 2x_1 + 3x_2 + x_3 = 4 \end{cases}$$

to a matrix and solve.

$$\left[\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ 3 & -1 & -3 & -1 \\ 2 & 3 & 1 & 4 \end{array}\right]$$

The goal is to get to upper triangular. We call $a_{11} = 1$ the pivot and eliminate a_{21} and a_{31}, then repeat.

$$\left[\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ 3 & -1 & -3 & -1 \\ 2 & 3 & 1 & 4 \end{array}\right] \overset{R_2-3R_1}{\sim} \left[\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 2 & 3 & 1 & 4 \end{array}\right] \overset{R_3-2R_1}{\sim} \left[\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & -1 & -1 & -2 \end{array}\right]$$

Our new pivot is $a_{22} = -7$. Repeat:

$$\overset{R_2-7R_3}{\sim} \left[\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

Now we can solve via backsubstitution:

$$\begin{aligned} x_3 &= 4 \\ -7x_2 - 6(4) &= -10 \\ -7x_2 &= 14 \\ x_2 &= -2 \\ x_1 + 2(-2) + 4 &= 3 \\ x_1 &= 3 \end{aligned}$$

The solution is $(3, -2, 4)$.

Homework 1. §1.1: 1b, 2b, 3d, 4d, 6a, 10

Day 02 of 24 – §1.2 Row echelon form

Example 18. Solve the system

$$\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 \\ -2 & -2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 1 & 3 \end{array}\right].$$

$$\overset{R_1+R_2}{\sim} \left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ -2 & -2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 1 & 3 \end{array}\right] \overset{R_3+2R_1}{\sim} \left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 2 & 5 \\ 0 & 0 & 1 & 1 & 3 \end{array}\right]$$

Notice that we can't get strict triangular form. But we keep going, with the pivot now $a_{23} = 1$:

$$\overset{R_3-2R_2}{\sim} \left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 3 \end{array}\right] \overset{R_4-R_2}{\sim} \left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

The last two rows are equations of the form $0x_1 + 0x_2 + 0x_3 + 0x_4 = 1$ and thus the system is inconsistent.

Example 19. Solve the system

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 \\ -2 & -2 & 0 & 0 & 2 \\ 0 & 0 & 1 & 1 & 2 \end{array}\right].$$

We see

$$\begin{aligned}
&\overset{R_1+R_2}{\sim} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ -2 & -2 & 0 & 0 & 2 \\ 0 & 0 & 1 & 1 & 2 \end{array}\right] \\
&\overset{R_3+2R_1}{\sim} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 2 & 4 \\ 0 & 0 & 1 & 1 & 2 \end{array}\right] \\
&\overset{R_3-2R_2}{\sim} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 \end{array}\right] \\
&\overset{R_4-R_2}{\sim} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]
\end{aligned}$$

In contrast, now we have equations of the form $0x_1 + 0x_2 + 0x_3 + 0x_4 = 0$, which implies infinitely many solutions. Solving this system means finding a 4-tuple with

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_3 + x_4 &= 2 \end{aligned}$$

Definition 20. The variables x_1 and x_3 in **Example 19** are called **lead variables**, and the remaining are called **free variables**.

Finishing **Example 19**, solving this system means

$$\begin{aligned} x_3 &= 2 - x_4 \\ x_1 &= 1 - x_2 - x_3 - x_4 \\ x_1 &= 1 - x_2 - 2. \end{aligned}$$

You get infinitely many solutions. You're allowed to choose whatever you like for the free variables, and then the lead variables are forced upon you. The general solution is

$$(1 - x_2 - 2, x_2, 2 - x_4, x_4),$$

and some specific examples are:

$$\begin{array}{ll} x_2 = 1,\ x_4 = 0: & (-2, 1, 2, 0), \\ x_2 = 3,\ x_4 = -1: & (-4, 3, 3, -1), \end{array}$$

etc.

Definition 21. In both **Example 18** and **Example 19**, the matrices weren't in strict triangular form but instead **row echelon form**. Using **I.**, **II.**, and **III.** to reach row echelon form is called **Gaussian elimination**.

Example 22. Here's some examples of row echelon form:

$$\begin{bmatrix} 1 & 4 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example 23. Here's some nonexamples:

$$\begin{bmatrix} 2 & 4 & 6 \\ 0 & 3 & 5 \\ 0 & 0 & 4 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Definition 24. An $m \times n$ system is **overdetermined** if $m > n$; i.e., if there are more equations than variables. It is **underdetermined** if $m < n$; i.e., if there are more variables than equations.

Remark 25. An underdetermined system can never have a unique solution. Consider a sample row echelon form:

$$m\left\{\left[\begin{array}{ccccc|c} \multicolumn{5}{c}{\overbrace{\hphantom{1 \quad * \quad * \quad * \quad *}}^{n}} & \\ 1 & * & * & * & * & * \\ 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & ! \end{array}\right]\right.\left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\}r$$

You'll have $r \leq m$ nonzero rows. So there will be r lead variables, and $n - r$ free variables, but since $r \leq m < n$,

$$n - r > 0.$$

So either the system is inconsistent and there are no solutions (e.g., in the case that the ! is nonzero), or it's consistent, we have free variables to play with, and there are infinitely many solutions. But never a unique solution.

Example 26. The underdetermined system

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 4 & 2 & 3 \end{array}\right] \overset{R_2 - 2R_1}{\sim} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

is inconsistent. (Visualize parallel planes in $\mathbf{R}^3$.)

Example 27. The underdetermined system

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 2 & 2 & 3 \\ 1 & 1 & 1 & 2 & 3 & 2 \end{array}\right] \sim \cdots \sim \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

has two free variables x_2 and x_3, and thus infinitely many solutions. We can find them explicitly by continuing:

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right] \overset{R_2-R_3}{\sim} \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right]$$
$$\overset{R_1-R_3}{\sim} \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right]$$
$$\overset{R_1-R_2}{\sim} \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

Thus

$$\begin{aligned} x_1 &= 1 - x_2 - x_3 \\ x_4 &= 2 \\ x_5 &= -1 \end{aligned}$$

so solutions are the points $(1 - x_2 - x_3, x_2, x_3, 2, -1)$ for any x_2 and x_3.

Definition 28. We say a matrix is in **reduced row echelon form** when it's in row echelon form and also each leading 1 is the only entry in its column. Using **I.**, **II.**, and **III.** to reach reduced row echelon form is called **Gauss-Jordan reduction**.

Example 29. Here's some more matrices in reduced row echelon form:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Definition 30. An $m \times n$ system $[A \mid B]$ is **homogeneous** if $B = 0$; e.g.,

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & 0 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & 0 \end{array}\right].$$

Remark 31. A homogenous system can never be inconsistent. Why? Because $(x_1, x_2, \ldots, x_n) = (0, 0, \ldots, 0)$ is always a solution. We call it the trivial solution.

If a homogeneous system has a unique solution, then it has to be $(0, \ldots, 0)$. But it doesn't have to have a unique solution. For instance:

Theorem 32. *An underdetermined homogeneous system has nontrivial solutions.*

Proof. Underdetermined systems must have either no solutions or infinitely many, by **Remark 25**. But since there is at least one solution, namely $(0, \ldots, 0)$, there must be infinitely many. □

Homework 2. §1.2: 2, 3, 4, 5e, 6c, 7

Day 03 of 24 – §1.3 Matrix arithmetic □

We want to understand matrices more deeply than just for solving linear systems. First, some language:

Definition 33. A matrix with one row or one column is called a **vector**, and is in bijection with a tuple. That is, the **row vector**

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

corresponds to the point $(1, 2, 3)$, as does the **column vector**

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

We care about vectors because they can express points, and so in particular express solutions to $m \times n$ systems. We can also express matrices in terms of its row or column vectors.

Example 34. The matrix

$$A = \begin{bmatrix} 3 & 2 & 5 \\ -1 & 8 & 4 \end{bmatrix}$$

can be written

$$A = \begin{bmatrix} \overline{a_1} & \overline{a_2} & \overline{a_3} \end{bmatrix}, \qquad \overline{a_1} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \qquad \overline{a_2} = \begin{bmatrix} 2 \\ 8 \end{bmatrix}, \qquad \overline{a_3} = \begin{bmatrix} 5 \\ 4 \end{bmatrix},$$

or as

$$A = \begin{bmatrix} \overline{a_1} \\ \overline{a_2} \end{bmatrix}, \qquad \overline{a_1} = \begin{bmatrix} 3 & 2 & 5 \end{bmatrix}, \qquad \overline{a_2} = \begin{bmatrix} -1 & 8 & 4 \end{bmatrix}.$$

There's no standard notation; your book uses bold for column vectors and arrows for row vectors. At first, I'll use bars for all vectors, rows or columns, but as the semester progresses we might drop the bar when it's clear from context.

We can do arithmetic with matrices:

Definition 35. If A is a matrix and α is a scalar, the **scalar multiplication** αA is the matrix whose entries are all multiplied by α.

Definition 36. If A and B are two matrices of the same size, then the **matrix addition** $A + B$ is the matrix of added termwise entries.

Example 37.

$$2 \begin{bmatrix} 3 & 4 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ -2 & 4 \end{bmatrix}$$

Example 38.

$$\begin{bmatrix} 3 & 5 & 1 \\ 2 & 0 & -3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & -4 \\ 10 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 5 & -3 \\ 12 & 2 & 0 \end{bmatrix}.$$

Definition 39. If A and B are the same size, define **matrix subtraction** $A - B$ to be $A + (-1)B$, which is indeed termwise subtraction.

We can also multiply matrices. The starting idea is that we want to compactly write an $m \times n$ system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

as a product of matrices

$$A\overline{x} = \overline{b},$$

just like we can write a single linear equation as $ax = b$.

If we write

$$\overline{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

then we have the right hand side. For the left, write

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad \overline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

and define the product $A\overline{x}$:

Definition 40. The **product of an $m \times n$ matrix A and an $n \times 1$ matrix $\overline{x}$** is an $m \times 1$ matrix computed by adding row×column:

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}.$$

Notice that when you set that product equal to

$$\overline{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

you get the $m \times n$ system we started with.

Example 41.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{bmatrix}$$

Example 42.

$$\begin{bmatrix} -3 & 1 \\ 2 & 5 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -3 \cdot 2 + 1 \cdot 4 \\ 2 \cdot 2 + 5 \cdot 4 \\ 4 \cdot 2 + 2 \cdot 4 \end{bmatrix} = \begin{bmatrix} -2 \\ 24 \\ 16 \end{bmatrix}$$

Remark 43. Notice that we can factor $A\overline{x}$:

$$\begin{aligned} A\overline{x} &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \\ &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}. \end{aligned}$$

So the equation $A\overline{x} = \overline{b}$ can be expressed in terms of the column vectors $\overline{a_1}, \overline{a_2}, \ldots, \overline{a_n}$:

$$A\overline{x} = x_1\overline{a_1} + x_2\overline{a_2} + \cdots + x_n\overline{a_n} = \overline{b}.$$

Example 44. We can write the system

$$\begin{cases} 2x_1 + 3x_2 - 2x_3 = 5 \\ 5x_1 - 4x_2 + 2x_3 = 6 \end{cases}$$

as

$$\begin{aligned} \begin{bmatrix} 2 & 3 & -2 \\ 5 & -4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} 5 \\ 6 \end{bmatrix} \\ \begin{bmatrix} 2x_1 + 3x_2 - 2x_3 \\ 5x_1 - 4x_2 + 2x_3 \end{bmatrix} &= \begin{bmatrix} 5 \\ 6 \end{bmatrix} \\ x_1 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ -4 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 2 \end{bmatrix} &= \begin{bmatrix} 5 \\ 6 \end{bmatrix}. \end{aligned}$$

Definition 45. If $\overline{v_1}, \overline{v_2}, \ldots, \overline{v_n}$ are vectors and $\alpha_1, \alpha_2, \ldots, \alpha_n$ are scalars, then a **linear combination of** $\overline{v_1}, \ldots, \overline{v_n}$ is a sum

$$\alpha_1\overline{v_1} + \alpha_2\overline{v_2} + \cdots + \alpha_n\overline{v_n}.$$

Example 46. This definition means that solving an $m \times n$ system $A\overline{x} = \overline{b}$ is finding n scalars $x_1, x_2, \ldots, x_n$ such that $\overline{b}$ is written as a linear combination of the columns of A and the scalars $x_1, \ldots, x_n$. For instance, recalling **Example 44**, see that

$$\begin{aligned} 2 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 3 \begin{bmatrix} 3 \\ -4 \end{bmatrix} + 4 \begin{bmatrix} -2 \\ 2 \end{bmatrix} &\overset{?}{=} \begin{bmatrix} 5 \\ 6 \end{bmatrix} \\ \begin{bmatrix} 4 + 9 - 8 \\ 10 - 12 + 8 \end{bmatrix} &\overset{?}{=} \begin{bmatrix} 5 \\ 6 \end{bmatrix} \\ \begin{bmatrix} 5 \\ 6 \end{bmatrix} &\overset{\checkmark}{=} \begin{bmatrix} 5 \\ 6 \end{bmatrix}. \end{aligned}$$

This means that the system in **Example 44** is consistent, and a solution is

$$\overline{x} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}.$$

In fact, this always works:

Theorem 47. *$A\overline{x} = \overline{b}$ is consistent, i.e., you can find a solution $\overline{x}$, if and only if you can write $\overline{b}$ as a linear combination of the column vectors of A.*

Proof. **Remark 43**. □

Now we want to upgrade matrix multiplication from $A\overline{x}$ to AB where B might not be a column vector anymore. The trick is to use our definition of $A\overline{x}$ from **Definition 40** on each column of B one at a time.

Definition 48. The **product of an $m \times n$ matrix A and an $n \times r$ matrix B** is an $m \times r$ matrix computed by working one column at a time on B:

$$AB = A\begin{bmatrix}\overline{b_1} & \overline{b_2} & \cdots & \overline{b_n}\end{bmatrix} = \begin{bmatrix}A\overline{b_1} & A\overline{b_2} & \cdots & A\overline{b_n}\end{bmatrix}.$$

Remark 49. Notice that the number of columns of A, n, has the match the number of rows of B, and that the output has the same number of rows as A, m, and columns as B, r.

$$\overset{m\times n}{A}\ \overset{n\times r}{B} = \overset{m\times r}{AB}.$$

Example 50. Let

$$A = \begin{bmatrix}3 & -2\\ 2 & 4\\ 1 & -3\end{bmatrix}, \qquad B = \begin{bmatrix}-2 & 1 & 3\\ 4 & 1 & 6\end{bmatrix}.$$

We get

$$\begin{aligned}AB &= \begin{bmatrix}A\overline{b_1} & A\overline{b_2} & A\overline{b_3}\end{bmatrix}\\ &= \begin{bmatrix}\begin{bmatrix}3 & -2\\ 2 & 4\\ 1 & -3\end{bmatrix}\begin{bmatrix}-2\\ 4\end{bmatrix} & \begin{bmatrix}3 & -2\\ 2 & 4\\ 1 & -3\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix} & \begin{bmatrix}3 & -2\\ 2 & 4\\ 1 & -3\end{bmatrix}\begin{bmatrix}3\\ 6\end{bmatrix}\end{bmatrix}\\ &= \begin{bmatrix}3(-2)-2(4) & 3(1)-2(1) & 3(3)-2(6)\\ 2(-2)+4(4) & 2(1)+4(1) & 2(3)+4(6)\\ 1(-2)-3(4) & 1(1)-3(1) & 1(3)-3(6)\end{bmatrix}\\ &= \begin{bmatrix}-14 & 1 & -3\\ 12 & 6 & 30\\ -14 & -2 & -15\end{bmatrix}.\end{aligned}$$

Example 51. Check this shit out:

$$\begin{aligned}BA &= \begin{bmatrix}\begin{bmatrix}-2 & 1 & 3\\ 4 & 1 & 6\end{bmatrix}\begin{bmatrix}3\\ 2\\ 1\end{bmatrix} & \begin{bmatrix}-2 & 1 & 3\\ 4 & 1 & 6\end{bmatrix}\begin{bmatrix}-2\\ 4\\ -3\end{bmatrix}\end{bmatrix}\\ &= \begin{bmatrix}-2(3)+1(2)+3(1) & -2(-2)+1(4)+3(-3)\\ 4(3)+1(2)+6(1) & 4(-2)+1(4)+6(-3)\end{bmatrix}\\ &= \begin{bmatrix}-1 & -1\\ 20 & -22\end{bmatrix}.\end{aligned}$$

So $AB \neq BA$ – they aren't even the same size! In fact, it gets worse:

Example 52. Let

$$A = \begin{bmatrix}3 & 4\\ 1 & 2\end{bmatrix}, \qquad B = \begin{bmatrix}1 & 2\\ 4 & 5\\ 3 & 6\end{bmatrix}.$$

Now, we can't even multiply A times B, because A is a 2×2 matrix and B is a 3×2, and the number of columns of A, 2, does not equal the number of rows of B, 3.

$$\overset{2\times 2}{A}\ \overset{3\times 2}{B}\ ☹$$

But, we can do BA.

$$\overset{3\times 2}{B}\ \overset{2\times 2}{A} = \overset{3\times 2}{BA}\ ☺$$

We get

$$\begin{aligned} BA &= \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1(3)+2(1) & 1(4)+2(2) \\ 4(3)+5(1) & 4(4)+5(2) \\ 3(3)+6(1) & 3(4)+6(2) \end{bmatrix} \\ &= \begin{bmatrix} 5 & 8 \\ 17 & 26 \\ 15 & 24 \end{bmatrix}. \end{aligned}$$

Even in the best case scenario where they're the same size, you can't guarantee AB is BA:

Example 53. Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}.$$

We compute:

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 1(1)+1(2) & 1(1)+1(2) \\ 0(1)+0(2) & 0(1)+0(2) \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & 0 \end{bmatrix} \\ BA &= \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1(1)+1(0) & 1(1)+1(0) \\ 2(1)+2(0) & 2(1)+2(0) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}. \end{aligned}$$

And clearly $AB \neq BA$.

Definition 54. Given an $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

define the **transpose of** A to be the $n \times m$ matrix A^T that swaps rows with columns:

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}.$$

Example 55.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \qquad A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

$$B = \begin{bmatrix} -3 & 2 & 1 \\ 4 & 3 & 2 \\ 1 & 2 & 5 \end{bmatrix}, \qquad B^T = \begin{bmatrix} -3 & 4 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 5 \end{bmatrix}.$$

$$C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \qquad C^T = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}.$$

Definition 56. When a $n \times n$ square matrix A is equal to its transpose ($A = A^T$) we say A is **symmetric**.

Example 57. Some symmetric matrices are C from **Example 55** and

$$\begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix}, \qquad \begin{bmatrix} 2 & 3 & 4 \\ 3 & 1 & 5 \\ 4 & 5 & 3 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & -2 \\ 2 & -2 & -3 \end{bmatrix}.$$

Homework 3. §1.3: 2ab, 4b, 9, 11, 12

Day 04 of 24 – §1.4 Matrix algebra □

So, we can add, subtract, multiply, and scale matrices, but multiplication is not commutative; $AB \neq BA$. What all *can* we do?

Proposition 58. *Let A, B, and C be matrices. Let α and β be scalars. The following are true, if they are defined:*

1. $A + B = B + A$.
2. $(A + B) + C = A + (B + C)$. *Write* $A + B + C$.
3. $(AB)C = A(BC)$. *Write* ABC.
4. $A(B + C) = AB + AC$.
5. $(A + B)C = AC + BC$.
6. $(\alpha\beta)A = \alpha(\beta A)$.
7. $\alpha(AB) = (\alpha A)B = A(\alpha B)$.
8. $(\alpha + \beta)A = \alpha A + \beta A$.
9. $\alpha(A + B) = \alpha A + \alpha B$.

Example 59. If

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 1 \\ -3 & 2 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix},$$

confirm that $A(B+C) = AB + AC$.

$$\begin{aligned}
A(B+C) &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \left(\begin{bmatrix} 2 & 1 \\ -3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \right) \\
&= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 1(3)+2(-1) & 1(1)+2(3) \\ 3(3)+4(-1) & 3(1)+4(3) \end{bmatrix} \\
&= \begin{bmatrix} 1 & 7 \\ 5 & 15 \end{bmatrix}. \\
AB + AC &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1(2)+2(-3) & 1(1)+2(2) \\ 3(2)+4(-3) & 3(1)+4(2) \end{bmatrix} + \begin{bmatrix} 1(1)+2(2) & 1(0)+2(1) \\ 3(1)+4(2) & 3(0)+4(1) \end{bmatrix} \\
&= \begin{bmatrix} -4 & 5 \\ -6 & 11 \end{bmatrix} + \begin{bmatrix} 5 & 2 \\ 11 & 4 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 7 \\ 5 & 15 \end{bmatrix}.
\end{aligned}$$

Definition 60. Given a matrix A and $k \in \mathbf{N}$, define $\underline{A^k}$ to be

$$A^k = \underbrace{AA \cdots A}_{k \text{ times}}$$

Example 61. Let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

We calculate

$$\begin{aligned}
A^2 = AA &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1(1)+1(1) & 1(1)+1(1) \\ 1(1)+1(1) & 1(1)+1(1) \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}. \\
A^3 = AAA = AA^2 &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 1(2)+1(2) & 1(2)+1(2) \\ 1(2)+1(2) & 1(2)+1(2) \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}. \\
A^4 = AAAA = AA^3 &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} = \begin{bmatrix} 1(4)+1(4) & 1(4)+1(4) \\ 1(4)+1(4) & 1(4)+1(4) \end{bmatrix} = \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix}.
\end{aligned}$$

In general, for any power $k \in \mathbf{N}$,

$$A^k = \begin{bmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{bmatrix}.$$

Remark 62. In the same way that $a + 0 = 0 + a = a$ for any a, it's easy to see that $A + 0 = 0 + A = A$ for any matrix A, where 0 is the matrix with all entries 0. Is there an analogous matrix that corresponds to the multiplicative identity 1? That is, since $1a = a1 = a$, is there a matrix where, when you multiply by it, the output doesn't change? Let's call that matrix I and we want

$$AI = IA = A.$$

Notice that I has to be square and the same size as A, since we're swapping order of matrix multiplication:

$$\overset{n\times n}{A}\ \overset{n\times n}{I} = \overset{n\times n}{I}\ \overset{n\times n}{A} = \overset{n\times n}{A}.$$

Definition 63. The $n \times n$ **identity matrix** I is

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

It has 1s on the diagonal and 0s everywhere else.

Example 64. We can check:

$$\begin{bmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3(1)+4(0)+1(0) & 3(0)+4(1)+1(0) & 3(0)+4(0)+1(1) \\ 2(1)+6(0)+3(0) & 2(0)+6(1)+3(0) & 2(0)+6(0)+3(1) \\ 0(1)+1(0)+8(0) & 0(0)+1(1)+8(0) & 0(0)+1(0)+8(1) \end{bmatrix} = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{bmatrix} = \begin{bmatrix} 1(3)+0(2)+0(0) & 1(4)+0(6)+0(1) & 1(1)+0(3)+0(8) \\ 0(3)+1(2)+0(0) & 0(4)+1(6)+0(1) & 0(1)+1(3)+0(8) \\ 0(3)+0(2)+1(0) & 0(4)+0(6)+1(1) & 0(1)+0(3)+1(8) \end{bmatrix} = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{bmatrix}.$$

Example 65. In complete generality for 2×2s:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a(1)+b(0) & a(0)+b(1) \\ c(1)+d(0) & c(0)+d(1) \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1(a)+0(c) & 1(b)+0(d) \\ 0(a)+1(c) & 0(b)+1(d) \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Remark 66. Now, every $a \neq 0$ has a multiplicative inverse. It's a number b where $ab = 1$. It's easy to find:

$$\begin{aligned} ab &= 1 \\ b &= \frac{1}{a}. \end{aligned}$$

So the multiplicative inverse of a is $1/a$. Notice that $a \neq 0$, or else you're dividing by 0!

Is there a multiplicative inverse for matrices? That is, if you start with A, is there a matrix B where $AB = I$?

Definition 67. An $n \times n$ matrix A is **nonsingular / invertible** if it has an **inverse** A^{-1}; i.e., there is a matrix A^{-1} such that

$$AA^{-1} = A^{-1}A = I.$$

We say A is **singular / not invertible** if it doesn't have an inverse.

Remark 68. This doesn't yet answer our question; it's just new language. Which matrices are nonsingular?

Example 69. We can confirm that if

$$A = \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix},$$

then A is invertible and

$$A^{-1} = \begin{bmatrix} \frac{-1}{10} & \frac{2}{5} \\ \frac{3}{10} & \frac{-1}{5} \end{bmatrix},$$

because

$$AA^{-1} = \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} \frac{-1}{10} & \frac{2}{5} \\ \frac{3}{10} & \frac{-1}{5} \end{bmatrix} = \begin{bmatrix} 2\left(\frac{-1}{10}\right) + 4\left(\frac{3}{10}\right) & 2\left(\frac{2}{5}\right) + 4\left(\frac{-1}{5}\right) \\ 3\left(\frac{-1}{10}\right) + 1\left(\frac{3}{10}\right) & 3\left(\frac{2}{5}\right) + 1\left(\frac{-1}{5}\right) \end{bmatrix} = \begin{bmatrix} \frac{-1}{5} + \frac{6}{5} & \frac{4}{5} - \frac{4}{5} \\ \frac{-3}{10} + \frac{3}{10} & \frac{6}{5} - \frac{1}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$A^{-1}A = \begin{bmatrix} \frac{-1}{10} & \frac{2}{5} \\ \frac{3}{10} & \frac{-1}{5} \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} \frac{-1}{10}(2) + \frac{2}{5}(3) & \frac{-1}{10}(4) + \frac{2}{5}(1) \\ \frac{3}{10}(2) - \frac{1}{5}(3) & \frac{3}{10}(4) - \frac{1}{5}(1) \end{bmatrix} = \begin{bmatrix} \frac{-1}{5} + \frac{6}{5} & \frac{-2}{5} + \frac{2}{5} \\ \frac{3}{5} - \frac{3}{5} & \frac{6}{5} - \frac{1}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Example 70. The matrix

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

is not invertible. To see this, suppose we multiply A by *anything*:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1(a) + 0(c) & 1(b) + 0(d) \\ 0(a) + 0(c) & 0(b) + 0(d) \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}.$$

No matter what, that can't be

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

because of the 0 in row 2, column 2.

Remark 71. How do we know if a matrix is nonsingular, besides just being handed the inverse? That's an important question. We'll answer it later, but for now, let's get comfortable with inverses:

Theorem 72. *If A and B are nonsingular $n \times n$ matrices, then so is AB, and in fact, we know the inverse explicitly:*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

In general, if we have a whole bunch of nonsingular $n \times n$ matrices $A_1, A_2, \ldots, A_k$, then so is $A_1A_2\cdots A_k$ and

$$(A_1A_2\cdots A_k)^{-1} = A_k^{-1}\cdots A_2^{-1}A_1^{-1}.$$

Proof. Let's prove the claim for just two matrices A and B; the second claim is an "induction argument." We can just check if the thing we claim is $(AB)^{-1}$ actually gets us I when we multiply:

$$AB(AB)^{-1} = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I. \checkmark$$
$$(AB)^{-1}AB = B^{-1}A^{-1}AB = B^{-1}IB = B^{-1}B = I. \checkmark$$

□

Remark 73. Now, here's an answer for when 2×2s are nonsingular. If

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

then A is nonsingular if $ad - bc \neq 0$ and the inverse is

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

To check, see that

$$
\begin{aligned}
AA^{-1} &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \left(\frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \right) \\
&= \frac{1}{ad-bc} \begin{bmatrix} a(d)+b(-c) & a(-b)+b(a) \\ c(d)+d(-c) & c(-b)+d(a) \end{bmatrix} \\
&= \frac{1}{ad-bc} \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \\
A^{-1}A &= \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \\
&= \frac{1}{ad-bc} \begin{bmatrix} d(a)-b(c) & d(b)-b(d) \\ -c(a)+a(c) & -c(b)+a(d) \end{bmatrix} \\
&= \frac{1}{ad-bc} \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\end{aligned}
$$

Note that this is only a partial answer – there are matrices other than 2×2s! We will see a bigger picture later.

Proposition 74. *Recall the transpose matrix A^T. The following are true:*

1. $(A^T)^T = A$.
2. $(\alpha A)^T = \alpha A^T$.
3. $(A+B)^T = A^T + B^T$.
4. $(AB)^T = B^T A^T$.

Remark 75. Notice that **Proposition 74** #4 looks similar to the rule for inverses in **Theorem 72**. Careful – these are separate concepts. Of course, they have to be, because every matrix has a transpose, but only nonsingular (and thus square!) matrices even have an inverse.

Example 76. Let

$$
A = \begin{bmatrix} 1 & 2 \\ 3 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}.
$$

Checking **Proposition 74** #4:

$$
\begin{aligned}
(AB)^T &= \left(\begin{bmatrix} 1 & 2 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \right)^T = \begin{bmatrix} 1(1)+2(2) & 1(0)+2(1) \\ 3(1)+3(2) & 3(0)+3(1) \end{bmatrix}^T = \begin{bmatrix} 5 & 2 \\ 9 & 3 \end{bmatrix}^T = \begin{bmatrix} 5 & 9 \\ 2 & 3 \end{bmatrix}. \\
B^T A^T &= \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 2 \\ 3 & 3 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1(1)+2(2) & 1(3)+2(3) \\ 0(1)+1(2) & 0(3)+1(3) \end{bmatrix} = \begin{bmatrix} 5 & 9 \\ 2 & 3 \end{bmatrix}.
\end{aligned}
$$

Homework 4. §1.4: 3, 4, 10, 16

Day 05 of 24 – §1.5 Elementary matrices ☐

Remark 77. We know how to solve an $m \times n$ system $A\overline{x} = \overline{b}$ via row operations, but we'd like to solve it via matrix multiplication, because that's analogous to how you solve $ax = b$:

$$\begin{aligned} ax &= b \\ a^{-1}ax &= a^{-1}b \\ x &= a^{-1}b. \end{aligned}$$

This is naïvely optimistic for matrices, because we know that for $A\overline{x} = \overline{b}$, there need not exist A^{-1}. So what we'll do is multiply by nonsingular $m \times m$ matrices, even if A is singular. This gives equivalent systems, because if $\overline{x}$ solves

$$A\overline{x} = \overline{b},$$

then the same $\overline{x}$ clearly solves

$$MA\overline{x} = M\overline{b},$$

but the reverse is also true because we can multiply by M^{-1}:

$$\begin{aligned} M^{-1}MA\overline{x} &= M^{-1}M\overline{b} \\ A\overline{x} &= \overline{b}. \end{aligned}$$

So what nonsingular $m \times m$ Ms should we use? They will be matrices that make the system progressively simpler:

> **Definition 78.** An <u>**elementary matrix**</u> is an $m \times m$ matrix E which is exactly one row operation away from I. There are three kinds:
> **I.** Swap two rows of I.
> **II.** Multiply (or divide) a row of I by a scalar.
> **III.** Add (or subtract) one row of I with another.
> Also note that elementary matrices are nonsingular, and the inverse is of the same kind (check!).

Example 79. For example,

$$E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad E' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \qquad E'' = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

are type **I.**, **II.**, and **III.**, respectively.

Remark 80. The notation is not a coincidence. Look at what multiplying by elementary matrices does:

$$\begin{aligned} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} &= \begin{bmatrix} 0(a)+1(d)+0(g) & 0(b)+1(e)+0(h) & 0(c)+1(f)+0(i) \\ 1(a)+0(d)+0(g) & 1(b)+0(e)+0(h) & 1(c)+0(f)+0(i) \\ 0(a)+0(d)+1(g) & 0(b)+0(e)+1(h) & 0(c)+0(f)+1(i) \end{bmatrix} \\ &= \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix}. \end{aligned}$$

Multiplying by E swapped R_1 and R_2. That's a type **I.** row operation.

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} &= \begin{bmatrix} 1(a)+0(d)+0(g) & 1(b)+0(e)+0(h) & 1(c)+0(f)+0(i) \\ 0(a)+1(d)+0(g) & 0(b)+1(e)+0(h) & 0(c)+1(f)+0(i) \\ 0(a)+0(d)+3(g) & 0(b)+0(e)+3(h) & 0(c)+0(f)+3(i) \end{bmatrix} \\ &= \begin{bmatrix} a & b & c \\ d & e & f \\ 3g & 3h & 3i \end{bmatrix}. \end{aligned}$$

Multiplying by E' scaled $3R_3$. That's a type **II.** row operation.

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} 1(a)+0(d)+1(g) & 1(b)+0(e)+1(h) & 1(c)+0(f)+1(i) \\ 0(a)+1(d)+0(g) & 0(b)+1(e)+0(h) & 0(c)+1(f)+0(i) \\ 0(a)+0(d)+1(g) & 0(b)+0(e)+1(h) & 0(c)+0(f)+1(i) \end{bmatrix}$$
$$= \begin{bmatrix} a+g & b+h & c+i \\ d & e & f \\ g & h & i \end{bmatrix}.$$

Multiplying by E'' added $R_1 + R_3$. That's a type **III.** row operation.

Definition 81. A matrix A is **row equivalent** to B if

$$B = E_k \cdots E_2 E_1 A$$

where $E_1, E_2, \ldots, E_k$ are elementary matrices.

Since elementary matrices correspond to elementary row operations, we have:

Theorem 82. *Let A be an $n \times n$ matrix. The following are equivalent.*
1. *A is row equivalent to I.*
2. *A is nonsingular.*
3. *The homogeneous system $A\overline{x} = \overline{0}$ has only the trivial solution $\overline{x} = \overline{0}$.*

Proof. We prove $1 \Rightarrow 2$, $2 \Rightarrow 3$, and $3 \Rightarrow 1$.

$1 \Rightarrow 2$: If A is row equivalent to I, then

$$I = E_k \cdots E_2 E_1 A,$$

or equivalently

$$E_1^{-1} E_2^{-1} \cdots E_k^{-1} = A.$$

So we claim $E_k \cdots E_2 E_1 = A^{-1}$. To check:

$$A^{-1}A = E_k \cdots E_2 E_1 E_1^{-1} E_2^{-1} \cdots E_k^{-1} = I. \checkmark$$
$$AA^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1} E_k \cdots E_2 E_1 = I. \checkmark$$

So A is nonsingular, and we know its inverse.

$2 \Rightarrow 3$: If A is nonsingular, then there is an A^{-1}. Suppose we have a solution $\overline{x}$ of $A\overline{x} = \overline{0}$. Then:

$$\overline{x} = A^{-1}A\overline{x} = A^{-1}\overline{0} = \overline{0},$$

so $\overline{x}$ had to be $\overline{0}$ – that's the only solution.

$3 \Rightarrow 1$: We need to show A is row equivalent to I. A is certainly row equivalent to its reduced row echelon form. But if that's not I, then A must have a free variable, which is impossible since $A\overline{x} = \overline{0}$ has one solution, not infinitely many. Thus A is row equivalent to I.

□

We can use this to solve nonhomogeneous systems too:

Corollary 83. *The $n \times n$ system $A\overline{x} = \overline{b}$ has a unique solution if and only if A is nonsingular.*

Proof. One direction is easy: if A is nonsingular, then

$$\begin{aligned} A\overline{x} &= \overline{b} \\ A^{-1}A\overline{x} &= A^{-1}\overline{b} \\ \overline{x} &= A^{-1}\overline{b}. \end{aligned}$$

So $\overline{x}$ can only be one thing, $A^{-1}\overline{b}$.

For the other direction, if $A\overline{x} = \overline{b}$ has a unique solution $\overline{x}$, look what happens if A were singular: By **Theorem 82**, $A\overline{x} = \overline{0}$ would have a solution that isn't $\overline{0}$. Call it $\overline{y}$. But then

$$A(\overline{x} + \overline{y}) = A\overline{x} + A\overline{y} = \overline{b} + \overline{0} = \overline{b},$$

so $\overline{x} + \overline{y}$ is another solution of the system, but we already said that we have only one solution $\overline{x}$. This contradiction means that A couldn't be singular. $\square$

Remark 84. Elementary matrices give us a way to calculate the inverse of a nonsingular matrix. By **Theorem 82**, we can calcuate A^{-1} by starting at I and using the same elementary matrices that take us from A to I. The idea is to write an augmented matrix $[A \mid I]$, then do the row operations that take the left A to I. This then forces the right I to go to A^{-1}.

Example 85. Find A^{-1} if

$$A = \begin{bmatrix} 1 & 4 \\ -1 & -2 \end{bmatrix}.$$

We calculate $[A \mid I] \sim [I \mid A^{-1}]$, giving us A^{-1}.

$$\left[\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ -1 & -2 & 0 & 1 \end{array}\right] \overset{R_1+R_2}{\sim} \left[\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{array}\right] \overset{R_1-2R_2}{\sim} \left[\begin{array}{cc|cc} 1 & 0 & -1 & -2 \\ 0 & 2 & 1 & 1 \end{array}\right] \overset{\frac{1}{2}R_2}{\sim} \left[\begin{array}{cc|cc} 1 & 0 & -1 & -2 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right].$$

Therefore,

$$A^{-1} = \begin{bmatrix} -1 & -2 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

To check this, you can do two things: just multiply AA^{-1} and $A^{-1}A$ and see that you get I, or use **Remark 73**, which says what the inverse of a 2×2 is:

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{1(-2) - 4(-1)}\begin{bmatrix} -2 & -4 \\ 1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -2 & -4 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \checkmark$$

Example 86. We can use **Example 85** to factor A and A^{-1} in terms of elementary matrices:

$$I = \overset{\frac{1}{2}R_2}{\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}} \overset{R_1-2R_2}{\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}} \overset{R_1+R_2}{\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}} A,$$

so

$$\begin{aligned} A^{-1} &= \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \\ A &= \left(\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\right)^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \end{aligned}$$

Example 87. If we want to find the solution to the system

$$\begin{cases} x_1 + 4x_2 = 3 \\ -x_1 - 2x_2 = 1 \end{cases}$$

$$\begin{bmatrix} 1 & 4 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix},$$

we can, because we know the inverse from **Example 85**:

$$\begin{bmatrix} -1 & -2 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 4 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1(3) - 2(1) \\ \frac{1}{2}(3) + \frac{1}{2}(1) \end{bmatrix} = \begin{bmatrix} -5 \\ 2 \end{bmatrix}.$$

Homework 5. §1.5: 10ef, 11, 12ab

Day 06 of 24 – §2.1 The determinant of a matrix ☐

Since nonsingular matrices are so important, we want to develop more tests to check if a given matrix is nonsingular. Let's give a few examples:

Example 88. A 1×1 matrix is just $[a]$. This is the same as just the number a, and so $[a]$ is invertible if and only if a is. That means $a \neq 0$.

Example 89. What about 2×2 matrices? We saw in **Remark 73** that if

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

then

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

That's only defined if $ad - bc \neq 0$.

Remark 90. So there appears to be a scalar, a in the case of 1×1, $ad - bc$ in the case of 2×2, which tells you your matrix is nonsingular when it is not 0. What about 3×3 and up?

Example 91. Let's try to work out the scalar for 3×3 explicitly. A matrix is nonsingular if we can do row operations to get to I (**Theorem 82**), so what does that look like?

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \overset{R_2 - \frac{d}{a}R_1}{\sim} \begin{bmatrix} a & b & c \\ 0 & e - \frac{bd}{a} & f - \frac{cd}{a} \\ g & h & i \end{bmatrix} \overset{R_3 - \frac{g}{a}R_1}{\sim} \begin{bmatrix} a & b & c \\ 0 & e - \frac{bd}{a} & f - \frac{cd}{a} \\ 0 & h - \frac{bg}{a} & i - \frac{cg}{a} \end{bmatrix} = \left[\begin{array}{c|cc} a & b & c \\ \hline 0 & \frac{ae-bd}{a} & \frac{af-cd}{a} \\ 0 & \frac{ah-bg}{a} & \frac{ai-cg}{a} \end{array}\right].$$

To do those steps, we must assume $a \neq 0$.

Now we've got a smaller block:

$$\begin{bmatrix} \frac{ae-bd}{a} & \frac{af-cd}{a} \\ \frac{ah-bg}{a} & \frac{ai-cg}{a} \end{bmatrix}$$

and that block needs to get to I. But it's a 2×2, and we know what's required for a 2×2:

$$\begin{aligned}
0 &\neq \left(\frac{ae-bd}{a}\right)\left(\frac{ai-cg}{a}\right) - \left(\frac{af-cd}{a}\right)\left(\frac{ah-bg}{a}\right) \\
&= \frac{a^2ei - aceg - abdi + bcdg - (a^2fh - abfg - acdh + bcdg)}{a^2} \\
&= \frac{a^2ei - aceg - abdi - a^2fh + abfg + acdh}{a^2} \\
&= \frac{aei - ceg - bdi - afh + bfg + cdh}{a}.
\end{aligned}$$

Dividing by a won't make something 0 unless it already was, so really we just need the numerator:

$$aei - ceg - bdi - afh + bfg + cdh \neq 0. \tag{$*$}$$

Now, bear in mind we had to assume $a \neq 0$. But that's not always the case. If $a = 0$, then we have three possibilities to consider:

1. d and g are also 0.
2. $d \neq 0$.
3. $g \neq 0$.

#1 means we have

$$\begin{bmatrix} 0 & b & c \\ 0 & e & f \\ 0 & h & i \end{bmatrix}$$

which has to be singular. Notice that plugging in $a = d = g = 0$ to $(*)$ gives you 0, which is promising because we want a number which is 0 when you're singular.

For #2, by swapping row 1 and 2, you get

$$\begin{bmatrix} d & e & f \\ 0 & b & c \\ g & h & i \end{bmatrix},$$

but now you know the top left is nonzero, so we have, by the same arithmetic as before:

$$\sim \left[\begin{array}{c|cc} d & e & f \\ \hline 0 & b & c \\ 0 & \frac{dh-eg}{d} & \frac{di-fg}{d} \end{array}\right]$$

$$b\left(\frac{di-fg}{d}\right) - c\left(\frac{dh-eg}{d}\right) = \frac{bdi - bfg - cdh + ceg}{d},$$

so

$$bdi - bfg - cdh + ceg \neq 0. \tag{$**$}$$

Notice $(**)$ is the same thing you'd get if you plugged $a = 0$ into $(*)$, but off by -1. Still, multiplying by -1 doesn't change if something is 0 or not.

The exact same argument as #2 works for #3, but we'll skip for brevity. If you check this, note that the row swap again introduces a -1 compared to $(*)$. Regardless, in all three cases, $(*)$ is the formula. It's 0 if and only if your 3×3 is singular.

Remark 92. How might you do this for even bigger matrices? Notice that we want from 2×2 to 3×3 by doing some row reductions and getting to a smaller block matrix. That's the idea. We get to larger $n \times n$ matrices by reducing the calculation to smaller $(n-1) \times (n-1)$ matrices.

Definition 93. The **determinant,** $\det A$, of an $n \times n$ matrix A is

$$\begin{aligned}
\det[a] &= a \\
\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} &= ad - bc \\
\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} &= a \det \begin{bmatrix} \times & \times & \times \\ \times & e & f \\ \times & h & i \end{bmatrix} - b \det \begin{bmatrix} \times & \times & \times \\ d & \times & f \\ g & \times & i \end{bmatrix} + c \det \begin{bmatrix} \times & \times & \times \\ d & e & \times \\ g & h & \times \end{bmatrix} \\
&= a(ei - fh) - b(di - fg) + c(dh - eg) \\
&= aei - afh - bdi + bfg + cdh - ceg. \qquad (*)
\end{aligned}$$

In general, for an $n \times n$:

$$\det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = a_{11} \det \begin{bmatrix} \times & \times & \cdots & \times \\ \times & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ \times & a_{n2} & \cdots & a_{nn} \end{bmatrix} - a_{12} \det \begin{bmatrix} \times & \times & \cdots & \times \\ a_{21} & \times & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & \times & \cdots & a_{nn} \end{bmatrix} + \cdots + (-1)^{n+1} a_{nn} \det \begin{bmatrix} \times & \times & \cdots & \times \\ a_{21} & a_{22} & \cdots & \times \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & \times \end{bmatrix}.$$

In fact, this worked for 2×2s too:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = a \det \begin{bmatrix} \times & \times \\ \times & d \end{bmatrix} - b \begin{bmatrix} \times & \times \\ c & \times \end{bmatrix} = ad - bc.$$

We call deleting the ith row and jth column and calculating the determinant of that smaller block the **minor,** $\det M_{ij}$. We say we have a **cofactor** A_{ij} when you keep track of the ± 1; i.e., write $A_{ij} = (-1)^{i+j} \det M_{ij}$. Thus the determinant is the **cofactor expansion**

$$\begin{aligned}
\det A &= a_{11} A_{11} + a_{12} A_{12} + \cdots + a_{1n} A_{1n} \\
&= a_{11} \det M_{11} - a_{12} \det M_{12} + \cdots + (-1)^{1+n} \det M_{1n}.
\end{aligned}$$

Proposition 94. *You don't have to do a cofactor expansion along the first row. You can do a cofactor expansion along any row or any column.*

Example 95. Calculate

$$\det \begin{bmatrix} 0 & 2 & 3 & 0 \\ 0 & 4 & 5 & 0 \\ 0 & 1 & 0 & 3 \\ 2 & 0 & 1 & 3 \end{bmatrix}.$$

Since **Proposition 94** says we can use any row/column, let's use the one with the most 0s:

$$\begin{aligned}
\det \begin{bmatrix} 0 & 2 & 3 & 0 \\ 0 & 4 & 5 & 0 \\ 0 & 1 & 0 & 3 \\ 2 & 0 & 1 & 3 \end{bmatrix} &= 0 \det \begin{bmatrix} 4 & 5 & 0 \\ 1 & 0 & 3 \\ 0 & 1 & 3 \end{bmatrix} - 0 \det \begin{bmatrix} 2 & 3 & 0 \\ 1 & 0 & 3 \\ 0 & 1 & 3 \end{bmatrix} + 0 \det \begin{bmatrix} 2 & 3 & 0 \\ 4 & 5 & 0 \\ 0 & 1 & 3 \end{bmatrix} - 2 \det \begin{bmatrix} 2 & 3 & 0 \\ 4 & 5 & 0 \\ 1 & 0 & 3 \end{bmatrix} \\
&= -2 \left(0 \det \begin{bmatrix} 4 & 5 \\ 1 & 0 \end{bmatrix} - 0 \det \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix} + 3 \det \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \right) \\
&= -2 \Big(3 \big(2(5) - 3(4) \big) \Big) \\
&= -2(3)(-2) \\
&= 12.
\end{aligned}$$

Remark 96. We know that $\det A \neq 0$ is equivalent to A is nonsingular when A is 1×1, 2×2, and 3×3. Is it always equivalent? Yes, but TBD (**Theorem 106**).

Proposition 97. $\det(A) = \det(A^T)$.

Example 98.

$$\det \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = 1(4) - 2(3) = -2.$$

$$\det \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^T = \det \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = 1(4) - 3(2) = -2. \checkmark$$

Proposition 99. *If A is triangular (upper or lower), then $\det(A)$ is the product of the diagonal.*

Example 100.

$$\begin{aligned} \det \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix} &= 1 \det \begin{bmatrix} 3 & 0 \\ 5 & 6 \end{bmatrix} - 0 \det \begin{bmatrix} 2 & 0 \\ 4 & 6 \end{bmatrix} + 0 \det \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \\ &= 3(6) - 0(5) \\ &= 18. \end{aligned}$$

Product of the diagonal: $1 \cdot 3 \cdot 6 = 18$. ✓

Proposition 101.
1. *If A has a row/column of 0s, then* $\det A = 0$.
2. *If A has two identical rows/columns, then* $\det A = 0$.

Homework 6. §2.1: 2, 3, 5, 6

Day 07 of 24 – §2.2 Properties of determinants ☐

To understand determinants better, let's see what row operations do to the determinant of a matrix.

Example 102. Type **I.**: swapping two rows. Let's see what happens with 2×2 and 3×3:

$$\begin{aligned} \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} &= ad - bc. \\ \det \begin{bmatrix} c & d \\ a & b \end{bmatrix} &= cb - da \\ &= -(ad - bc) \\ &= -\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \end{aligned}$$

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

You can use this idea to work your way up to any $n \times n$ matrix. Swapping two rows always just introduces a minus sign. (We saw this already in **Example 91**!)

Let E be the elementary matrix that swaps rows i and j. Swapping two rows of A is the same as multiplying EA. Notice that

$$\det E = -\det I = -\underbrace{(1 \cdot 1 \cdots 1)}_{\textbf{Proposition 99}} = -1.$$

So in this case

$$\det(EA) = -\det A = \det E \det A.$$

But be careful! It is not yet a given that $\det AB = \det A \det B$. So we had to cook it up ad hoc for this specific E.

Example 103. Type **II.**: scaling a row by α. If we scale up the ith row, and then do a cofactor expansion on that same row:

$$\det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ \alpha a_{i1} & \alpha a_{i2} & \cdots & \alpha a_{in} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \pm \alpha a_{i1} \det \begin{bmatrix} \times & a_{12} & \cdots & a_{1n} \\ \times & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ \times & \times & \cdots & \times \\ \vdots & & & \vdots \\ \times & a_{n2} & \cdots & a_{nn} \end{bmatrix} \mp \alpha a_{i2} \det \begin{bmatrix} a_{11} & \times & \cdots & a_{1n} \\ a_{21} & \times & \cdots & a_{2n} \\ \vdots & & & \vdots \\ \times & \times & \cdots & \times \\ \vdots & & & \vdots \\ a_{n1} & \times & \cdots & a_{nn} \end{bmatrix} \pm \cdots \pm \alpha a_{in} \det \begin{bmatrix} a_{11} & a_{12} & \cdots & \times \\ a_{21} & a_{22} & \cdots & \times \\ \vdots & & & \vdots \\ \times & \times & \cdots & \times \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & \times \end{bmatrix}$$

We can write that as

$$\begin{aligned} &= \alpha a_{i1} A_{i1} + \alpha a_{i2} A_{i2} + \cdots + \alpha a_{in} A_{in} \\ &= \alpha \left(a_{i1} A_{i1} + a_{i2} A_{i2} + \cdots + a_{in} A_{in} \right) \\ &= \alpha \det A. \end{aligned}$$

So scaling a row by α scales the determinant by α.

Let E be the elementary matrix that scales row i by α. Notice

$$\det E = \alpha \det I = \alpha.$$

So once again

$$\det(EA) = \alpha \det A = \det E \det A.$$

Curious. . .

Example 104. Type **III.**: add row i to row j. Do a cofactor expansion on row j:

$$\begin{aligned} \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \vdots \\ a_{i1} + a_{j1} & a_{i2} + a_{j1} & \cdots & a_{in} + a_{j1} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} &= (a_{i1} + a_{j1}) A_{j1} + (a_{i2} + a_{j2}) A_{j2} + \cdots + (a_{in} + a_{jn}) A_{jn} \\ &= a_{i1} A_{j1} + a_{j1} A_{j1} + a_{i2} A_{j2} + a_{j2} A_{j2} + \cdots + a_{in} A_{jn} + a_{jn} A_{jn} \\ &= \det A + a_{i1} A_{j1} + a_{i2} A_{j2} + \cdots + a_{in} A_{jn}. \end{aligned}$$

For that mixed cofactor expansion, notice that

$$
\begin{aligned}
a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} &= \det \begin{array}{cl} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} & \begin{array}{l} \\ \\ \\ \\ j\text{th row} \\ \\ \\ \end{array} \end{array} \\
&= 0
\end{aligned}
$$

since the determinant is 0 when you have identical rows (**Proposition 101**). So ultimately adding row i to row j did not change the determinant at all.

If E is the elementary matrix adding row i to row j, we have

$$\det E = \det I = 1$$

and

$$\det(EA) = \det A = \det E \det A.$$

Hmm...

Remark 105. For all types of elementary matrices, we saw that

$$\det(EA) = \det E \det A,$$

and to summarize:

$$\det E = \begin{cases} -1 & \textbf{I.} \\ \alpha & \textbf{II.} \\ 1 & \textbf{III.} \end{cases}$$

Theorem 106. *For any size $n \times n$, A is singular if and only if* $\det A = 0$.

Proof. Using Gaussian elimination, turn A into row echelon form U:

$$U = E_k \cdots E_2 E_1 A.$$

By **Remark 105**,

$$\det U = \det(E_k \cdots E_2 E_1 A) = \det E_k \cdots \det E_2 \det E_1 \det A.$$

All the $\det E_i$ are nonzero. So $\det A = 0$ if and only if $\det U = 0$. But since $\det U$ is in row echelon form:

- A being singular means U has a row of 0s, hence by **Proposition 101**, $\det U = 0$.
- A being nonsingular means U is upper triangular with 1s on the diagonal and so by **Proposition 99**, $\det U = 1 \cdot 1 \cdots 1 = 1 \neq 0$.

□

Remark 107. Something to point out is that calculating $\det A$ is therefore easy if you can reduce to a triangular matrix U, because then

$$\det A = \frac{\det U}{\det E_k \cdots \det E_2 \det E_1},$$

and **Remark 105** tells us how to calculate all the $\det E_i$ and **Proposition 99** tells us how to calculate U (multiply the diagonal).

Example 108. What is

$$\det \begin{bmatrix} 2 & 1 & 3 \\ 4 & 2 & 1 \\ 6 & -3 & 4 \end{bmatrix}?$$

We calculate

$$\begin{aligned}
\begin{bmatrix} 2 & 1 & 3 \\ 4 & 2 & 1 \\ 6 & -3 & 4 \end{bmatrix} &\overset{R_2-2R_1}{\sim} \begin{bmatrix} 2 & 1 & 3 \\ 0 & 0 & -5 \\ 6 & -3 & 4 \end{bmatrix} \quad \textbf{III.} \\
&\overset{R_3-3R_1}{\sim} \begin{bmatrix} 2 & 1 & 3 \\ 0 & 0 & -5 \\ 0 & -6 & -5 \end{bmatrix} \quad \textbf{III.} \\
&\sim \begin{bmatrix} 2 & 1 & 3 \\ 0 & -6 & -5 \\ 0 & 0 & -5 \end{bmatrix} \quad \textbf{I.}
\end{aligned}$$

Thus

$$\begin{aligned}
\det \begin{bmatrix} 2 & 1 & 3 \\ 4 & 2 & 1 \\ 6 & -3 & 4 \end{bmatrix} &= \frac{\det \begin{bmatrix} 2 & 1 & 3 \\ 0 & -6 & -5 \\ 0 & 0 & -5 \end{bmatrix}}{1 \cdot 1 \cdot (-1)} \\
&= \frac{2(-6)(-5)}{-1} \\
&= -60.
\end{aligned}$$

Theorem 109. *For any $n \times n$s, not just elementaries,*

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

Proof. Either A is singular or it is not.

1. If A is singular, then by **Theorem 106**, $\det A = 0$. Exercise: A singular $\Rightarrow$ AB singular. But then we have

$$\begin{aligned}
\det(AB) &\overset{?}{=} \det A \det B \\
0 &\overset{\checkmark}{=} 0 \cdot \det B.
\end{aligned}$$

2. If A is nonsingular, then $A = E_k \cdots E_2 E_1$. Thus using **Remark 105** a bunch,

$$\begin{aligned}
\det(AB) &= \det(E_k \cdots E_2 E_1 B) \\
&= \det E_k \cdots \det E_2 \det E_1 \det B \\
&= \det(E_k \cdots E_2 E_1) \det B \\
&= \det A \det B. \checkmark
\end{aligned}$$

$\square$

Homework 7. §2.2: 1, 4, 7, 13

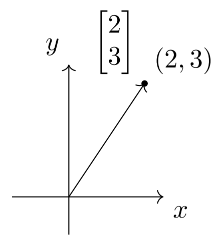
Day 08 of 24 – §3.1 Vector spaces

The linear algebra we've done so far has just involved real numbers and tuples of real numbers, but this is not mandatory. We're going to develop the theory in more generality. At its core, the arithmetic we've been doing has just required a few basic ideas.

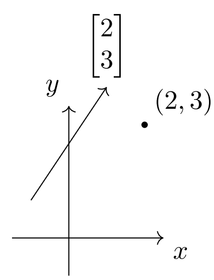
Example 110. We've been dealing so far with $\mathbf{R}^m$. Recall a vector

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

is in correspondence with a point $(x_1, x_2, \ldots, x_m)$. What is this correspondence? For instance, $(2, 3)$ corresponds to



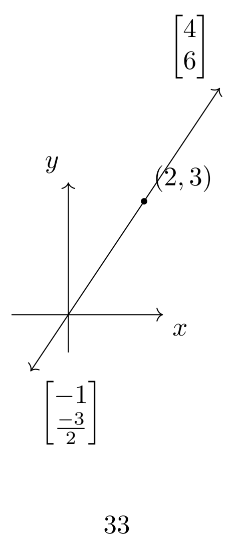
But the vector is the arrow, and is still the same vector even if it doesn't begin at the origin.



We know we can scale vectors, and graphically we can see:

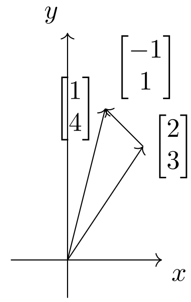
$$2\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$

$$-\frac{1}{2}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{-3}{2} \end{bmatrix}.$$

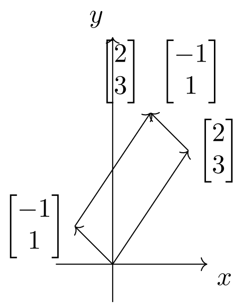


We also know we can add vectors; graphically this is:

$$\begin{bmatrix}2\\3\end{bmatrix}+\begin{bmatrix}-1\\1\end{bmatrix}=\begin{bmatrix}1\\4\end{bmatrix}.$$



And here's a visual proof that addition is commutative:



Example 111. Another example is matrices. In fact, **Example 110** is also matrices, since a (column) vector is $m \times 1$. But even a general $m \times n$ matrix works. We write $\mathbf{R}^m$ for the $m \times 1$ vectors, so we write $\mathbf{R}^{m\times n}$ for the $m \times n$ vectors. Here they aren't necessarily arrows anymore, just matrices, but notice we can still scale a matrix αA and add two matrices $A + B = B + A$.

Example 112. Here's an out-there example: quadratic and smaller polynomials. Recall

$$f = ax^2 + bx + c.$$

If $a = 0$, we have a linear polynomial, $a = b = 0$ is a constant polynomial, and $a = b = c = 0$ is the zero polynomial.

We can still scale:

$$\begin{aligned}\alpha f &= \alpha(ax^2 + bx + c)\\ &= \alpha ax^2 + \alpha bx + \alpha c,\end{aligned}$$

and we can add:

$$\begin{aligned}f + g &= (ax^2 + bx + c) + (dx^2 + ex + f)\\ &= (a + d)x^2 + (b + e)x + (c + d).\end{aligned}$$

(And order doesn't matter.)

Definition 113. Let V be a set. Suppose for every scalar α and for all $v, w \in V$, you can scale by α or you can add $v + w$, and you'll still get an element of V. That means
C1. $\alpha v \in V$.
C2. $v + w \in V$.

We use "C" because these conditions say that V is **closed** under addition and scalar multiplication; you will not leave V if you do them.

We then call $(V, \cdot, +)$ a **vector space** if the following additional axioms are satisfied:

A1. $v + w = w + v$.
A2. $(v + w) + x = v + (w + x)$.
A3. There exists $0 \in V$ such that $v + 0 = 0 + v = v$.
A4. For every v, there is a $-v$ such that $v + (-v) = -v + v = 0$.
A5. $\alpha(v + w) = \alpha v + \alpha w$.
A6. $(\alpha + \beta)v = \alpha v + \beta v$.
A7. $(\alpha\beta)v = \alpha(\beta v)$.
A8. $1v = v$.

Example 114. We've seen in **Example 110**, **Example 111**, and **Example 112** three examples of vector spaces (although you should go back and check *all* the axioms are satisfied; we did *not* do this!). Let's now see what happens when axioms fail.

Let $W = \{(a, 1) \in \mathbf{R}^2\}$ with usual addition and scalar multiplication. This is a set with elements like $(2, 1)$, $(5, 1)$, etc. See that **C2.** fails:

$$(2, 1) + (5, 1) = (7, 2) \notin W.$$

Also **C1.** fails, and a few others. So W is not a vector space.

Example 115. Let's consider $(\mathbf{R}, \cdot, \max)$. Here, scalar multiplication is normal, but instead of $+$ we take the max. A lot of things do work:

C1. $\alpha a \in \mathbf{R}$. ✓
C2. $\max\{a, b\} \in \mathbf{R}$. ✓
A1. $\max\{a, b\} = \max\{b, a\}$. ✓
A2. $\max\{\max\{a, b\}, c\} = \max\{a, \max\{b, c\}\} = \max\{a, b, c\}$. ✓
A3. $\max\{a, z\} = \max\{z, a\} \stackrel{?}{=} a$? We need a number z that's smaller than every $a \in \mathbf{R}$. Well that'd be $-\infty \notin \mathbf{R}$. So no **A3.** $\times$

Example 116. What about $C[a, b] = \{f : [a, b] \to \mathbf{R} \mid f \text{ is continuous}\}$? Define addition by

$$(f + g)(x) = f(x) + g(x)$$

and scalar multiplication by

$$(\alpha f)(x) = \alpha f(x).$$

C1. $\alpha f \in C[a, b]$. ✓
C2. $f + g \in C[a, b]$. ✓
A1. $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$. ✓
A2. ✓
A3. 0 is the zero function. $(f + 0)(x) = f(x) + 0 = f(x) = 0 + f(x) = (0 + f)(x)$. ✓
A4. If f is continuous, so is $-f$. ✓
$\vdots$

Yes, $C[a, b]$ is a vector space.

Example 117. Let P_n be the set of polynomials of degree less than n. For example, P_3 is **Example 112**. Just like P_3, P_n for any n is a vector space.

Theorem 118. *If V is a vector space, then*

1. $0v = 0$.
2. $v + w = 0 \Rightarrow w = -v$.

3. $(-1)v = -v$.

Proof. Be careful!!! These are obvious for $(\mathbf{R}, \cdot, +)$, but in a random vector space, the only things we know are **C1.**, **C2.**, and **A1.–A8.**. So while it might "feel obvious," we have to prove 1, 2, and 3 using only those axioms.

1. First, since $0 = 0 + 0$:

$$0v = (0+0)v = 0v + 0v. \qquad (\mathbf{A6.})$$

Now, add an inverse to $0v$ to both sides, which must exist by **A4.**:

$$\begin{aligned} -0v + 0v &= -0v + 0v + 0v \\ 0 &= 0 + 0v && (\mathbf{A4.}) \\ 0 &= 0v. && (\mathbf{A3.}) \end{aligned}$$

2. If $v + w = 0$, add an inverse $-v$ to both sides:

$$\begin{aligned} -v + v + w &= -v + 0 \\ w &= -v + 0 && (\mathbf{A4.}) \\ w &= -v. && (\mathbf{A3.}) \end{aligned}$$

3. By 2, we know that for any v, the additive inverse $-v$ is unique. So we've done it if we can show that $(-1)v$ serves the same function as $-v$, which means adding v and $(-1)v$ needs to give us 0. Let's compute:

$$v + (-1)v \overset{\mathbf{A8.}}{=} 1v + (-1)v \overset{\mathbf{A6.}}{=} (1-1)v = 0v \overset{\#1}{=} 0.$$

$\square$

Homework 8. §3.1: 8, 9, 10, 11

Day 09 of 24 – §3.2 Subspaces ☐

Example 119. The set

$$S = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid y = 2x \right\}$$

is clearly a subset of $\mathbf{R}^2$. But is it more than just a set? Is it a vector space too? More precisely, is it a vector space with the same $\cdot$ and $+$ as $(\mathbf{R}^2, \cdot, +)$?

Notice that every element of S looks like

$$\begin{bmatrix} x \\ 2x \end{bmatrix}.$$

If we scale:

$$\alpha \begin{bmatrix} x \\ 2x \end{bmatrix} = \begin{bmatrix} \alpha x \\ \alpha 2x \end{bmatrix} = \begin{bmatrix} \alpha x \\ 2(\alpha x) \end{bmatrix}$$

which is still in S. Also

$$\begin{bmatrix} x \\ 2x \end{bmatrix} + \begin{bmatrix} x' \\ 2x' \end{bmatrix} = \begin{bmatrix} x + x' \\ 2x + 2x' \end{bmatrix} = \begin{bmatrix} x + x' \\ 2(x + x') \end{bmatrix}$$

which is also still in S.

Example 120. This doesn't work for *every* subset though. Consider

$$T = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid y = x^2 \right\} \subseteq \mathbf{R}^2.$$

Scaling:

$$\alpha \begin{bmatrix} x \\ x^2 \end{bmatrix} = \begin{bmatrix} \alpha x \\ \alpha x^2 \end{bmatrix} \neq \begin{bmatrix} \alpha x \\ (\alpha x)^2 \end{bmatrix}. \times$$

Adding:

$$\begin{bmatrix} x \\ x^2 \end{bmatrix} + \begin{bmatrix} x' \\ x'^2 \end{bmatrix} = \begin{bmatrix} x + x' \\ x^2 + x'^2 \end{bmatrix} \neq \begin{bmatrix} x + x' \\ (x + x')^2 \end{bmatrix}. \times$$

Definition 121. Let V be a vector space. We say that $S \subseteq V$ is a **subspace** if:

0. $S \neq \emptyset$.
1. $\alpha S \in S$ for all scalars α.
2. $s + t \in S$ for all $s, t \in S$.

#1 and #2 are just saying S is **closed** under the same $\cdot$ and $+$ as V.

Remark 122.

1. Every subspace is in fact a vector space – you can check **A1.–A8.**.
2. Every vector space V must have 0 by **A3.**, and the set $\{0\}$ is a subspace of V. Also every V is a subspace of itself.
3. In **Example 119**, S is a subspace of $\mathbf{R}^2$, but in **Example 120**, T is not.

Example 123. Let

$$S = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid y = x \right\}.$$

0. $S \neq \emptyset$ because $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T \in S$. ✓
1. $\alpha \begin{bmatrix} x & x & z \end{bmatrix}^T = \begin{bmatrix} \alpha x & \alpha x & \alpha z \end{bmatrix}^T \in S$. ✓
2. $\begin{bmatrix} x & x & z \end{bmatrix}^T + \begin{bmatrix} x' & x' & z' \end{bmatrix}^T = \begin{bmatrix} x + x' & x + x' & z + z' \end{bmatrix}^T \in S$. ✓

Example 124. Let

$$S = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid y = 1 \right\}.$$

0. $S \neq \emptyset$ because $\begin{bmatrix} 1 & 1 \end{bmatrix}^T \in S$. ✓
1. $\alpha \begin{bmatrix} x & 1 \end{bmatrix}^T = \begin{bmatrix} \alpha x & \alpha \end{bmatrix}^T \notin S$. $\times$

Example 125. Here are a bunch of examples of subspaces (you should check):

1. $$\left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid c = -b \right\} \subseteq \mathbf{R}^{2 \times 2}.$$
2. $$\{ f \in P_n \mid f(0) = 0 \} \subseteq P_n.$$
3. $$C^n[a, b] = \left\{ f \in C[a, b] \mid f^{(n)} \text{ exists and is continuous on } [a, b] \right\} \subseteq C[a, b].$$
4. $$\left\{ f \in C^2[a, b] \mid f''(x) + f(x) = 0 \right\} \subseteq C^2[a, b].$$

Definition 126. Here's an important one, called the **kernel / null space**. Let A be an $m \times n$ matrix. Let

$$N(A) = \operatorname{nul}(A) = \ker(A) = \left\{\overline{x} \in \mathbf{R}^n \mid A\overline{x} = \overline{0}\right\} \subseteq \mathbf{R}^n.$$

0. $\ker A \neq \emptyset$ because $\overline{0} \in \ker A$: $A\overline{0} = \overline{0}$. ✓
1. If $A\overline{x} = \overline{0}$, then $A(\alpha\overline{x}) = \alpha A\overline{x} = \alpha\overline{0} = \overline{0}$. ✓
2. If $A\overline{x} = \overline{0}$ and $A\overline{y} = \overline{0}$, then $A(\overline{x} + \overline{y}) = A\overline{x} + A\overline{y} = \overline{0} + \overline{0} = \overline{0}$. ✓

Example 127. Calculate the null space of

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{bmatrix}.$$

Let's solve $A\overline{x} = \overline{0}$:

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \end{array}\right] \overset{2R_1 - R_2}{\sim} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 0 \end{array}\right] \overset{R_1 - R_2}{\sim} \left[\begin{array}{cccc|c} 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 \end{array}\right]$$

So we have free variables x_3 and x_4 and lead variables

$$\begin{aligned} x_1 &= x_3 - x_4 \\ x_2 &= -2x_3 + x_4. \end{aligned}$$

A general solution looks like

$$\begin{bmatrix} \alpha - \beta \\ -2\alpha + \beta \\ \alpha \\ \beta \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Thus

$$\operatorname{nul} A = \left\{ \alpha \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \mid \alpha, \beta \in \mathbf{R} \right\} \subseteq \mathbf{R}^4.$$

Here is one theorem which tells us something cool about the kernel.

Theorem 128. *Let $A\overline{x} = \overline{b}$ be consistent with solution $\overline{x}$. $\overline{y}$ is also a solution if and only if $\overline{y} = \overline{x} + \overline{z}$ where $\overline{z} \in \ker A$.*

Proof. There are two things to show:

1. If $\overline{x}$ is a solution and $\overline{z} \in \ker A$, then $\overline{x} + \overline{z}$ is a solution:

$$A(\overline{x} + \overline{z}) = A\overline{x} + A\overline{z} = \overline{b} + \overline{0} = \overline{b}. ✓$$

2. Every solution $\overline{y}$ ends up being $\overline{x}$ plus something in $\ker A$. If we want to show $\overline{y} = \overline{x} + \overline{z}$, then we can show $\overline{y} - \overline{x} = \overline{z}$ is in $\ker A$:

$$A(\overline{y} - \overline{x}) = A\overline{y} - A\overline{x} = \overline{b} - \overline{b} = \overline{0}. ✓$$

□

Definition 129. Let $v_1, v_2, \ldots, v_n \in V$ and let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be scalars. A **linear combination of** $v_1, \ldots, v_n$ is

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.$$

The set of all linear combinations of $v_1, \ldots, v_n$ is the **span**

$$\operatorname{Span}(v_1, v_2, \ldots, v_n).$$

Example 130. In **Example 127**,

$$\operatorname{nul} A = \operatorname{Span}\left(\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}\right).$$

Example 131. We write $\overline{e_i}$ for the vector

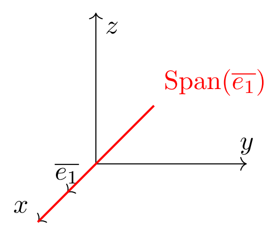
$$\begin{array}{cc} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} & \begin{array}{c} \\ \\ \\ i\text{th row} \\ \\ \\ \\ \end{array} \end{array}.$$

In $\mathbf{R}^3$, we have

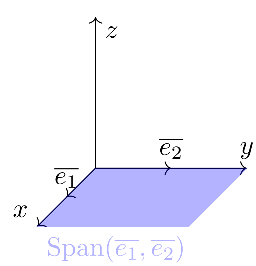
$$\overline{e_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \overline{e_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \overline{e_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

We calculate and draw:

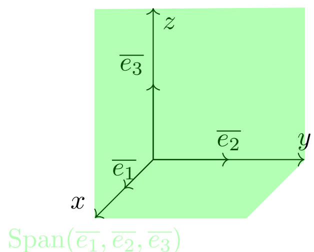
$$\operatorname{Span}(\overline{e_1}) = \{\alpha \overline{e_1}\} = \left\{\begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix}\right\}.$$



$$\operatorname{Span}(\overline{e_1}, \overline{e_2}) = \{\alpha \overline{e_1} + \beta \overline{e_2}\} = \left\{\begin{bmatrix} \alpha \\ \beta \\ 0 \end{bmatrix}\right\}.$$



$$\operatorname{Span}(\overline{e_1}, \overline{e_2}, \overline{e_3}) = \{\alpha\overline{e_1} + \beta\overline{e_2} + \gamma\overline{e_3}\} = \left\{ \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \right\} = \mathbf{R}^3.$$



Theorem 132. *If $v_1, \ldots, v_n \in V$, then $\operatorname{Span}(v_1, \ldots, v_n)$ is a subspace of V.*

Proof. An element of $\operatorname{Span}(v_1, \ldots, v_n)$ looks like $\alpha_1 v_1 + \cdots + \alpha_n v_n$.

0. $0 = 0v_1 + \cdots + 0v_n \in \operatorname{Span}(v_1, \ldots, v_n)$, so $\operatorname{Span}(v_1, \ldots, v_n) \neq \emptyset$. ✓
1. $\beta(\alpha_1 v_1 + \cdots + \alpha_n v_n) = (\beta\alpha_1)v_1 + \cdots + (\beta\alpha_n)v_n \in \operatorname{Span}(v_1, \ldots, v_n)$. ✓
2. $\alpha_1 v_1 + \cdots + \alpha_n v_n + \widetilde{\alpha_1} v_1 + \cdots + \widetilde{\alpha_n} v_n = (\alpha_1 + \widetilde{\alpha_1})v_1 + \cdots + (\alpha_n + \widetilde{\alpha_n})v_n \in \operatorname{Span}(v_1, \ldots, v_n)$. ✓

□

Definition 133. Let $v_1, \ldots, v_n \in V$. We say that the set $\{v_1, \ldots, v_n\}$ is a **spanning set for V** if $\operatorname{Span}(v_1, \ldots, v_n) = V$.

Example 134. In **Example 131**, $\operatorname{Span}(\overline{e_1}, \overline{e_2}, \overline{e_3}) = \mathbf{R}^3$, so $\{\overline{e_1}, \overline{e_2}, \overline{e_3}\}$ is a spanning set for $\mathbf{R}^3$. For that matter, so is $\{\overline{e_1}, \overline{e_2}, \overline{e_3}, v, w, \ldots\}$ for any extra vectors $v, w, \ldots$, because you can write any vector

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \in \mathbf{R}^3$$

as

$$\alpha\overline{e_1} + \beta\overline{e_2} + \gamma\overline{e_3} + 0v + 0w + \cdots.$$

Example 135. We can check that

$$S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

is a spanning set for $\mathbf{R}^3$. To write any vector

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

as a linear combination

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

by **Theorem 47** we must solve the system

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}. \tag{$*$}$$

We know from **Corollary 83** that there is a unique solution if and only if

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

is nonsingular. In fact it is; by permuting rows, we can write it lower triangular:

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

and thus $\det A = -(1 \cdot 1 \cdot 1) \neq 0$, so A is nonsingular. The solution to $(*)$ is found via backsubstitution:

$$\begin{aligned} \alpha &= c \\ \beta &= b - \alpha = b - c \\ \gamma &= a - \alpha - \beta = a - c - (b - c) = a - b. \end{aligned}$$

Thus

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = c \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + (b - c) \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + (a - b) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

so S is a spanning set for $\mathbf{R}^3$.

Example 136. But

$$S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} \right\}$$

is not a spanning set of $\mathbf{R}^3$. The coefficient matrix

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

is singular;

$$\begin{aligned} \det A &= 1\Big((1)(1) - (-1)(3)\Big) - 2\Big((2)(1) - (-1)(4)\Big) + 4\Big((2)(3) - (1)(4)\Big) \\ &= 4 - 2(6) + 4(2) \\ &= 0. \end{aligned}$$

We can check explicitly:

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & a \\ 2 & 1 & -1 & b \\ 4 & 3 & 1 & c \end{array}\right] \sim \cdots \sim \left[\begin{array}{ccc|c} 1 & 2 & 4 & a \\ 0 & 3 & 9 & 2a - b \\ 0 & 0 & 0 & 2a - 3c + 5b \end{array}\right]$$

So if $2a - 3c + 5b \neq 0$, the system is inconsistent, meaning no solutions. Thus, many vectors

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbf{R}^3$$

are not in $\operatorname{Span}(S)$; for example,

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \notin \operatorname{Span} S,$$

because $2(1) - 3(1) + 5(1) = 2 - 3 + 5 = 4 \neq 0$.

Homework 9. §3.2: 1, 4ab, 5, 11

Day 10 of 24 – §3.3 Linear indepedence

Remark 137. Recall from **Example 134** that the set

$$\{\overline{e_1}, \overline{e_2}, \overline{e_3}, v, w, \dots, \}$$

spans $\mathbf{R}^3$, but any extra vectors $v, w, \dots$ were redundant. In fact, you could do something dumb and span $\mathbf{R}^3$ with $\mathrm{Span}(\mathbf{R}^3)$ – the linear combination is a coefficient of 1 in front of the vector you want, and 0s everywhere else. But we'd like to find *minimal* spanning sets, because that turns dealing with *infinitely* many vectors in $\mathbf{R}^3$ into dealing with linear combinations of *finitely* many vectors.

Example 138. Consider

$$u = \begin{bmatrix} -1 \\ 3 \\ 8 \end{bmatrix}, \qquad v = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \qquad w = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}.$$

You can check that

$$u = 3v + 2w.$$

That means that if we want to write a linear combination of u, v, and w, we can rewrite it just in terms of v and w:

$$\alpha u + \beta v + \gamma w = \alpha(3v + 2w) + \beta v + \gamma w = (\beta + 3\alpha)v + (\gamma + 2\alpha)w.$$

Thus $\mathrm{Span}(u, v, w) = \mathrm{Span}(v, w)$. But we can't go any further; there's no way to express v in terms of w or w in terms of v:

$$\alpha v = \begin{bmatrix} \alpha \\ -\alpha \\ 2\alpha \end{bmatrix} \stackrel{?}{=} \beta w = \begin{bmatrix} -2\beta \\ 3\beta \\ \beta \end{bmatrix}.$$

This can only happen if $\alpha = \beta = 0$. Another way to say this is, if you set it equal to 0:

$$\alpha v + \beta w = 0$$

can only happen if α and β are both 0.

So that means $\mathrm{Span}(v) \subsetneq \mathrm{Span}(v, w)$ (and similarly $\mathrm{Span}(w) \subsetneq \mathrm{Span}(v, w)$).

Remark 139. This works more generally. If we have $v_1, v_2, \dots, v_n$ and you can write v_1 as a linear combination of the others, then $\mathrm{Span}(v_1, v_2, \dots, v_n) = \mathrm{Span}(v_2, \dots, v_n)$. And writing v_1 as a linear combination of the others is the same as being able to write

$$c_1v_1 + c_2v_2 + \dots + c_nv_n = 0$$

where not every c_i is 0 (because you can solve for v_1).

Definition 140. We say that $v_1, v_2, \dots, v_n$ are **linearly dependent** if you can write

$$c_1v_1 + c_2v_2 + \dots + c_nv_n = 0$$

and not every c_i was 0. On the other hand, we say that $v_1, \dots, v_n$ are **linearly independent** if the only way to write

$$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = 0$$

is when every c_i is 0.

Example 141. In **Example 138**, u, v, and w are linearly dependent:

$$-u + 3v + 2w = 0,$$

but v and w are linearly independent.

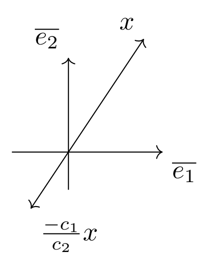
Example 142. We can understand what linear (in)dependence looks like graphically. First, an example in $\mathbf{R}^2$. If x and y are linearly dependent, then

$$c_1 x + c_2 y = 0$$

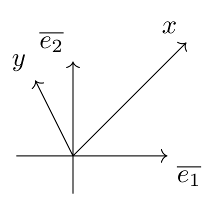
and c_1, c_2 aren't both 0. If we suppose $c_2 \neq 0$, then

$$\begin{aligned} c_2 y &= -c_1 x \\ y &= \frac{-c_1}{c_2} x. \end{aligned}$$

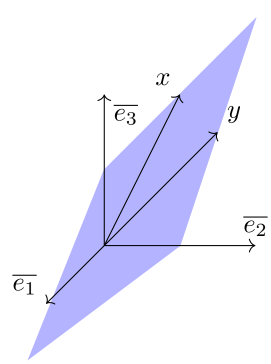
So in $\mathbf{R}^2$, linear dependence means y is a scalar multiple of x – on the same line through 0.



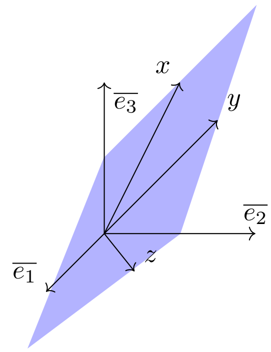
And linear independence means they aren't on the same line.



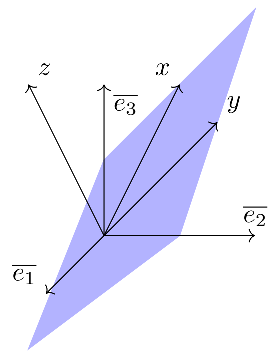
Example 143. In $\mathbf{R}^3$, two linearly independent vectors again don't lie on the same line through 0. They determine a plane through 0 in $\mathbf{R}^3$.



For a third vector z, if the set $\{x, y, z\}$ is linearly dependent, then z must lie in this plane.



And if $\{x, y, z\}$ is linearly independent, z does not lie on this plane.



Example 144. Is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

linearly independent?

If it is, we have to show that

$$c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

forces $c_1 = c_2 = c_3 = 0$. Using **Theorem 47** we must solve the system

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We know from **Theorem 82** that there is a unique solution if and only if the coefficient matrix is nonsingular, and in fact its determinant is -1. So there is a unique solution, which for a homogeneous system must be the trivial solution $c_1 = c_2 = c_3 = 0$.

Note: compare to **Example 135**!

Example 145. Is

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} \right\}$$

linearly independent?

Again, we need

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

to be nonzero, but recall from **Example 136** that it is 0, so by **Theorem 82** there are non-unique solutions, which means there's c_1, c_2, c_3, not all 0, that solve

$$c_1 \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + c_3 \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

For instance,

$$\begin{aligned} 2 \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} - 3 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} &= \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix} + \begin{bmatrix} -6 \\ -3 \\ -9 \end{bmatrix} + \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \end{aligned}$$

Theorem 146. *Let $v_1, \ldots, v_n \in \mathbf{R}^n$. Let $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$. The set $\{v_1, \ldots, v_n\}$ is linearly independent if and only if A is nonsingular.*

Proof. Use **Theorem 47** to rewrite

$$c_1 v_1 + \cdots + c_n v_n = 0$$

as

$$A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = 0,$$

and then use **Theorem 82** to get a unique solution, which must be $c_1 = \cdots = c_n = 0$, if and only if A is nonsingular. □

Remark 147. Theorem 146 only works when the number of vectors, n, equals the dimension. (This ensures A is square.) What do you do if you're trying to find out if a smaller set is linearly independent?

You can still produce a matrix A out of the column vectors $v_1, \ldots, v_k$, but it won't be square, so singular/nonsingular doesn't make sense. It will still represent a homogeneous system $[A \mid 0]$ though, and solutions will give you values of $c_1, \ldots, c_k$ that write

$$c_1 v_1 + \cdots + c_k v_k = 0$$

By **Theorem 32**, there will be a nontrivial solution $c_1, \ldots, c_k$ not all 0 if and only if you have free variables.

Example 148. Is

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 7 \\ 7 \end{bmatrix} \right\}$$

linearly independent?

Form a matrix of column vectors, and row reduce:

$$\left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ -1 & 3 & 0 & 0 \\ 2 & 1 & 7 & 0 \\ 3 & -2 & 7 & 0 \end{array}\right] \overset{R_1+R_2}{\sim} \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 2 & 1 & 7 & 0 \\ 3 & -2 & 7 & 0 \end{array}\right]$$
$$\overset{R_3-2R_1}{\sim} \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 5 & 5 & 0 \\ 3 & -2 & 7 & 0 \end{array}\right]$$
$$\overset{R_4-3R_1}{\sim} \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 5 & 5 & 0 \\ 0 & 4 & 4 & 0 \end{array}\right]$$
$$\sim \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Here c_3 is a free variable, so the vectors must be linearly dependent.

Theorem 149. *Any vector $w \in \mathrm{Span}(v_1, \ldots, v_n)$ can be written* uniquely *as a linear combination*

$$w = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

if and only if $\{v_1, \ldots, v_n\}$ is linearly independent.

Proof. We'll prove one half; see the book for the other half.
If $\{v_1, \ldots, v_n\}$ is linearly independent, then

$$c_1 v_1 + \cdots + c_n v_n = 0$$

forces $c_1 = \cdots = c_n = 0$. Now suppose you write

$$w = \alpha_1 v_1 + \cdots + \alpha_n v_n$$
$$w = \beta_1 v_1 + \cdots + \beta_n v_n.$$

Subtract and you get

$$0 = (\alpha_1 - \beta_1) v_1 + \cdots + (\alpha_n - \beta_n) v_n$$

but by linear independence this forces

$$\begin{array}{ccc} \alpha_1 - \beta_1 = 0 & \Rightarrow & \alpha_1 = \beta_1 \\ \vdots & & \vdots \\ \alpha_n - \beta_n = 0 & \Rightarrow & \alpha_n = \beta_n, \end{array}$$

so w's linear combination was unique. □

Homework 10. §3.3: 1, 5, 6

Day 11 of 24 – §3.4 Basis and dimension ☐

> **Definition 150.** Let V be a vector space. Let $v_1, \dots, v_n \in V$. If
> 1. $v_1, \dots, v_n$ are linearly independent, and
> 2. $\mathrm{Span}(v_1, \dots, v_n) = V$,
>
> then we say $\{v_1, \dots, v_n\}$ is a **basis** (pl. bases) for V.

Example 151. The "standard basis" for $\mathbf{R}^3$ (resp. $\mathbf{R}^n$) is $\{\overline{e_1}, \overline{e_2}, \overline{e_3}\}$ (resp. $\{\overline{e_1}, \dots, \overline{e_n}\}$). But a basis is not unique. You can check that

$$\left\{ \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}2\\0\\1\end{bmatrix} \right\}$$

or

$$\left\{ \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix} \right\}$$

are also bases, because they are linearly independent and span $\mathbf{R}^3$. But notice that in all cases, a basis of $\mathbf{R}^3$ has had 3 elements...

Example 152. The set

$$\left\{ \begin{bmatrix}1&0\\0&0\end{bmatrix}, \begin{bmatrix}0&1\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\1&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix} \right\}$$

is a basis for $\mathbf{R}^{2\times 2}$.

1. $$c_1\begin{bmatrix}1&0\\0&0\end{bmatrix} + c_2\begin{bmatrix}0&1\\0&0\end{bmatrix} + c_3\begin{bmatrix}0&0\\1&0\end{bmatrix} + c_4\begin{bmatrix}0&0\\0&1\end{bmatrix} = \begin{bmatrix}c_1&c_2\\c_3&c_4\end{bmatrix} = \begin{bmatrix}0&0\\0&0\end{bmatrix}$$

   if and only if $c_1 = c_2 = c_3 = c_4 = 0$, so linear independence ✓

2. $$\mathrm{Span}\left(\begin{bmatrix}1&0\\0&0\end{bmatrix}, \begin{bmatrix}0&1\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\1&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix} \right) = \left\{ a\begin{bmatrix}1&0\\0&0\end{bmatrix} + b\begin{bmatrix}0&1\\0&0\end{bmatrix} + c\begin{bmatrix}0&0\\1&0\end{bmatrix} + d\begin{bmatrix}0&0\\0&1\end{bmatrix} = \begin{bmatrix}a&b\\c&d\end{bmatrix} \right\} = \mathbf{R}^{2\times 2} \checkmark$$

We wonder: would any basis of $\mathbf{R}^{2\times 2}$ have 4 elements?

Theorem 153. *If* $\mathrm{Span}(v_1, \dots, v_n) = V$*, then any collection* $w_1, \dots, w_m \in V$ *with* $m > n$ *is linearly dependent.*

Proof. Since $\mathrm{Span}(v_1, \dots, v_n) = V$, every w_i can be written

$$w_i = a_{i1}v_1 + \cdots + a_{in}v_n.$$

To check if $\{w_1, \dots, w_m\}$ is linearly dependent, we need to write a sum

$$c_1w_1 + \cdots + c_mw_m = 0$$

without all of the c_i being 0. But we can do this:

$$\begin{aligned} c_1 w_1 + \cdots + c_m w_m &= 0 \\ c_1 \left(a_{11} v_1 + \cdots + a_{1n} v_n\right) + \cdots + c_m \left(a_{m1} v_1 + \cdots + a_{mn} v_n\right) &= 0 \\ \left(c_1 a_{11} + \cdots + c_m a_{m1}\right) v_1 + \cdots + \left(c_1 a_{1n} + \cdots + c_m a_{mn}\right) v_n &= 0 \end{aligned}$$

The coefficients form a linear system; we need to find a nontrivial solution to show that $\{w_1, \ldots, w_m\}$ is linearly dependent.

$$\begin{array}{ccccccc} c_1 a_{11} & + & \cdots & + & c_m a_{m1} & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ c_1 a_{1n} & + & \cdots & + & c_m a_{mn} & = & 0 \end{array}$$

$$\begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Since $m > n$, there are more columns than rows, hence more variables than equations, so by **Theorem 32**, there is a nontrivial solution $c_1, \ldots, c_m$ not all 0. Thus

$$c_1 w_1 + \cdots + c_m w_m = 0$$

and not all c_i are 0. □

Corollary 154. *If $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_m\}$ are bases for V, then $m = n$.*

Proof. By definition, $\operatorname{Span}(v_1, \ldots, v_n) = V$ and $\{w_1, \ldots, w_m\}$ is linearly independent. So by **Theorem 153**, $m \leq n$.

But now just swap: $\operatorname{Span}(w_1, \ldots, w_m) = V$ and $\{v_1, \ldots, v_n\}$ is linearly independent. By **Theorem 153**, $n \leq m$.

Thus $m = n$. □

So all bases of a vector space *are* the same size!

Definition 155. If V has a basis of size n, we say the **dimension of V is n,** $\dim V = n$. By fiat, $\dim\{0\} = 0$.

Theorem 156. *Let $\dim V = n > 0$. The set $\{v_1, \ldots, v_n\}$ is linearly independent if and only if it spans V.*

Proof. If $\{v_1, \ldots, v_n\}$ is linearly independent, then by **Theorem 153**, $\{v_1, \ldots, v_n, w\}$ for any $w \in V$ is not. So

$$c_1 v_1 + \cdots + c_n v_n + cw = 0$$

and not all coefficients are 0. In fact, $c \neq 0$, or else $\{v_1, \ldots, v_n\}$ would be linearly dependent. So

$$\begin{aligned} c_1 v_1 + \cdots + c_n v_n &= -cw \\ \frac{-c_1}{c} v_1 + \cdots + \frac{-c_n}{c} v_n &= w \end{aligned}$$

so w is a linear combination of $\{v_1, \ldots, v_n\}$, hence $w \in \operatorname{Span}(v_1, \ldots, v_n)$. But w was *any* vector in V, so $\operatorname{Span}(v_1, \ldots, v_n) = V$.

On the other hand, if $\operatorname{Span}(v_1, \ldots, v_n) = V$, suppose $\{v_1, \ldots, v_n\}$ *wasn't* linearly independent. One of the v_is can be written as a linear combination of the others. Eliminate it and the remaining vectors still span V. You may need to remove more vectors, but eventually you'll get a linearly independent spanning set, i.e., a basis, of size $k < n$. But, $\dim V = n$ and **Corollary 154** says that your basis *must* be size n. So it was wrong to suppose $\{v_1, \ldots, v_n\}$ was linearly dependent. It must be linearly independent. □

Proposition 157. *Let* $\dim V = n > 0$.
1. $\{v_1, \ldots, v_k\}$ *with* $k < n$ *cannot span* V.
2. *If* $\{v_1, \ldots, v_k\}$ *is linearly independent, you can add more vectors until you get a basis of* V.
3. *If* $\operatorname{Span}(v_1, \ldots, v_m) = V$ *and* $m > n$, *you can remove vectors until you get a basis of* V.

Homework 11. §3.4: 4, 7, 17

Day 12 of 24 – §3.5 Change of basis □

Remark 158. We care about bases because they convey the (often infinite) data of a vector space into a finite combination of basis vectors. For example, any $v \in \mathbf{R}^2$ can be expressed (uniquely, by **Theorem 149**) as

$$v = a\overline{e_1} + b\overline{e_2}.$$

But remember, while bases have the same size (**Corollary 154**), bases are not unique, so given a different basis $\{\overline{b_1}, \overline{b_2}\}$, you can still express v uniquely as

$$v = c\overline{b_1} + d\overline{b_2}.$$

The coefficients are likely different now.

Definition 159. Given a basis $\{\overline{b_1}, \ldots, \overline{b_n}\}$ and a vector v, when you write v uniquely as

$$v = c_1\overline{b_1} + \cdots + c_n\overline{b_n},$$

then the vector of the coefficients

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

is called the **<u>coordinate vector of v</u>**.

Example 160. When $\mathbf{R}^2$ has basis $\{\overline{e_1}, \overline{e_2}\}$, then the coordinate vector of

$$v = \begin{bmatrix} 7 \\ 7 \end{bmatrix} = 7\overline{e_1} + 7\overline{e_2}$$

is

$$\begin{bmatrix} 7 \\ 7 \end{bmatrix}.$$

But if $\mathbf{R}^2$ has basis

$$\left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \end{bmatrix} \right\} = \{\overline{b_1}, \overline{b_2}\}$$

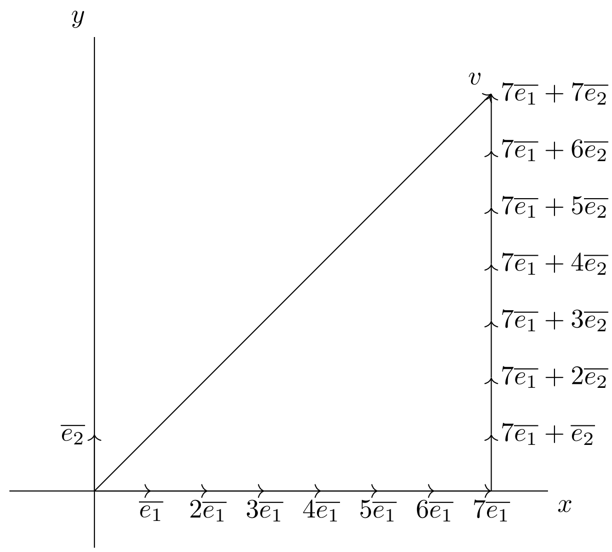
(which you should check forms a basis!), then the coordinate vector of

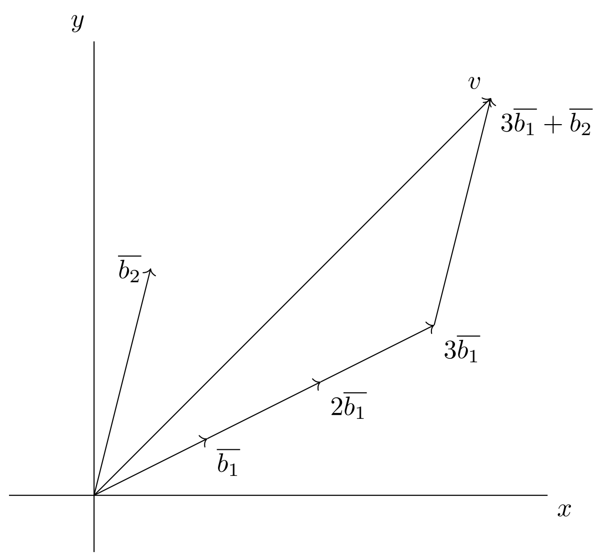
$$v = \begin{bmatrix} 7 \\ 7 \end{bmatrix} = 3\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

is

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

The geometric interpretation is that a basis determines your coordinate system:





Example 161. How does this work in general? If we want to express *any* $w \in \mathbf{R}^2$ in terms of the basis $\{\overline{b_1}, \overline{b_2}\}$ instead of $\{\overline{e_1}, \overline{e_2}\}$, or vice versa, what must we do?

1. Going from basis $\{\overline{b_1}, \overline{b_2}\}$ to $\{\overline{e_1}, \overline{e_2}\}$:
   Suppose you have a vector w given in coordinates $\{\overline{b_1}, \overline{b_2}\}$. So

   $$w = c_1\overline{b_1} + c_2\overline{b_2}.$$

   We can write the coordinate vectors for $\overline{b_1}$ and $\overline{b_2}$ in terms of $\overline{e_1}$ and $\overline{e_2}$. We have

   $$\overline{b_1} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2\overline{e_1} + \overline{e_2}$$
   $$\overline{b_2} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \overline{e_1} + 4\overline{e_2}.$$

Thus

$$\begin{aligned} w &= c_1\overline{b_1} + c_2\overline{b_2} \\ &= c_1\left(2\overline{e_1} + \overline{e_2}\right) + c_2\left(\overline{e_1} + 4\overline{e_2}\right) \\ &= \left(2c_1 + c_2\right)\overline{e_1} + \left(c_1 + 4c_2\right)\overline{e_2}. \end{aligned}$$

So the coordinate vector of w with respect to $\{\overline{e_1}, \overline{e_2}\}$ is

$$w = \begin{bmatrix} 2c_1 + c_2 \\ c_1 + 4c_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

In other words, we can write w in terms of the standard basis by multiplying its coordinate vector on the left by the matrix whose columns are the old basis. We say that this matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$$

is the *transition matrix* from basis $\{\overline{b_1}, \overline{b_2}\}$ to basis $\{\overline{e_1}, \overline{e_2}\}$.

2. Going from basis $\{\overline{e_1}, \overline{e_2}\}$ to $\{\overline{b_1}, \overline{b_2}\}$:
   Since the transition matrix from $\{\overline{b_1}, \overline{b_2}\}$ to $\{\overline{e_1}, \overline{e_2}\}$ has columns built from a basis, hence linearly independent columns, it is nonsingular, so it has an inverse. That inverse is

$$A^{-1} = \frac{1}{8-1}\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{4}{7} & \frac{-1}{7} \\ \frac{-1}{7} & \frac{2}{7} \end{bmatrix}.$$

   If $w = a_1\overline{e_1} + a_2\overline{e_2}$ and we want to solve for $w = c_1\overline{b_1} + c_2\overline{b_2}$, we know from part #1 that

$$\begin{aligned} \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \\ \begin{bmatrix} \frac{4}{7} & \frac{-1}{7} \\ \frac{-1}{7} & \frac{2}{7} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} \frac{4}{7} & \frac{-1}{7} \\ \frac{-1}{7} & \frac{2}{7} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \\ \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} \frac{4}{7} & \frac{-1}{7} \\ \frac{-1}{7} & \frac{2}{7} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}. \end{aligned}$$

   So the matrix A^{-1} is the transition matrix that turns a $w = a_1\overline{e_1} + a_2\overline{e_2}$ into $w = c_1\overline{b_1} + c_2\overline{b_2}$; it takes us from basis $\{\overline{e_1}, \overline{e_2}\}$ to basis $\{\overline{b_1}, \overline{b_2}\}$.

Example 162. We can double check **Example 160** now. If $v = 7\overline{e_1} + 7\overline{e_2}$, then we claimed the coordinates in terms of $\overline{b_1}$ and $\overline{b_2}$ were $v = 3\overline{b_1} + \overline{b_2}$. To see this:

$$\begin{aligned} \begin{bmatrix} \frac{4}{7} & \frac{-1}{7} \\ \frac{-1}{7} & \frac{2}{7} \end{bmatrix} \begin{bmatrix} 7 \\ 7 \end{bmatrix} &= \begin{bmatrix} \frac{4}{7}(7) - \frac{1}{7}(7) \\ \frac{-1}{7}(7) + \frac{2}{7}(7) \end{bmatrix} \\ &= \begin{bmatrix} 4-1 \\ -1+2 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 1 \end{bmatrix}. \checkmark \end{aligned}$$

Remark 163. What if we want to go from *any* basis $\{\overline{v_1}, \overline{v_2}\}$ to *any other* basis $\{\overline{w_1}, \overline{w_2}\}$? Easy!

$$\{\overline{v_1}, \overline{v_2}\} \longrightarrow \{\overline{e_1}, \overline{e_2}\} \longrightarrow \{\overline{w_1}, \overline{w_2}\}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} \longmapsto \begin{bmatrix} \overline{v_1} & \overline{v_2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \longmapsto \begin{bmatrix} \overline{w_1} & \overline{w_2} \end{bmatrix}^{-1} \begin{bmatrix} \overline{v_1} & \overline{v_2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Example 164. Let

$$\{\overline{v_1}, \overline{v_2}\} = \left\{ \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 7 \\ 3 \end{bmatrix} \right\}, \qquad \{\overline{w_1}, \overline{w_2}\} = \left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

Find the transition matrix from $\{\overline{v_1}, \overline{v_2}\}$ to $\{\overline{w_1}, \overline{w_2}\}$. Convert the coordinates $2\overline{v_1} - 3\overline{v_2}$ to $\{\overline{w_1}, \overline{w_2}\}$.

By **Remark 163**, the transition matrix is

$$\begin{aligned} \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} &= \frac{1}{3-2} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 4 \\ -4 & -5 \end{bmatrix}. \end{aligned}$$

To convert:

$$\begin{bmatrix} 3 & 4 \\ -4 & -5 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} -6 \\ 7 \end{bmatrix},$$

so $2\overline{v_1} - 3\overline{v_2} = -6\overline{w_1} + 7\overline{w_2}$.

Remark 165. All of this works in any finite dimensional vector space, not just $\mathbf{R}^2$.

Example 166. Let

$$\{\overline{v_1}, \overline{v_2}, \overline{v_3}\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ 4 \end{bmatrix} \right\}.$$

Convert $2\overline{v_1} + 3\overline{v_2} - \overline{v_3}$ into $\{\overline{e_1}, \overline{e_2}, \overline{e_3}\}$. Convert $\overline{e_1} - \overline{e_2} + \overline{e_3}$ into $\{\overline{v_1}, \overline{v_2}, \overline{v_3}\}$.

First,

$$\begin{aligned} \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 5 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} &= \begin{bmatrix} 2+6-1 \\ 2+9-5 \\ 2+6-4 \end{bmatrix} \\ &= \begin{bmatrix} 7 \\ 6 \\ 4 \end{bmatrix}, \end{aligned}$$

so $2\overline{v_1} + 3\overline{v_2} - \overline{v_3} = 7\overline{e_1} + 6\overline{e_2} + 4\overline{e_3}$.

Second,

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 3 & 5 & 0 & 1 & 0 \\ 1 & 2 & 4 & 0 & 0 & 1 \end{array}\right] \overset{R_2-R_1}{\sim} \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 4 & -1 & 1 & 0 \\ 1 & 2 & 4 & 0 & 0 & 1 \end{array}\right]$$

$$\overset{R_3-R_1}{\sim} \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 4 & -1 & 1 & 0 \\ 0 & 0 & 3 & -1 & 0 & 1 \end{array}\right]$$

$$\overset{R_1-2R_2}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & -7 & 3 & -2 & 0 \\ 0 & 1 & 4 & -1 & 1 & 0 \\ 0 & 0 & 3 & -1 & 0 & 1 \end{array}\right]$$

$$\overset{\frac{1}{3}R_3}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & -7 & 3 & -2 & 0 \\ 0 & 1 & 4 & -1 & 1 & 0 \\ 0 & 0 & 1 & \frac{-1}{3} & 0 & \frac{1}{3} \end{array}\right]$$

$$\overset{R_2-4R_3}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & -7 & 3 & -2 & 0 \\ 0 & 1 & 0 & \frac{1}{3} & 1 & \frac{-4}{3} \\ 0 & 0 & 1 & \frac{-1}{3} & 0 & \frac{1}{3} \end{array}\right]$$

$$\overset{R_1+7R_3}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{2}{3} & -2 & \frac{7}{3} \\ 0 & 1 & 0 & \frac{1}{3} & 1 & \frac{-4}{3} \\ 0 & 0 & 1 & \frac{-1}{3} & 0 & \frac{1}{3} \end{array}\right],$$

so

$$\begin{bmatrix} \frac{2}{3} & -2 & \frac{7}{3} \\ \frac{1}{3} & 1 & \frac{-4}{3} \\ \frac{-1}{3} & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} + 2 + \frac{7}{3} \\ \frac{1}{3} - 1 - \frac{4}{3} \\ \frac{-1}{3} - 0 + \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 0 \end{bmatrix},$$

and thus $\overline{e_1} - \overline{e_2} + \overline{e_3} = 5\overline{v_1} - 2\overline{v_2}$.

Homework 12. §3.5: 1, 2, 3, 7, 8

Day 13 of 24 – §3.6 Row space and column space ☐

Definition 167. Given an $m \times n$ matrix A, the subspace of $\mathbf{R}^{1\times n}$ spanned by the row vectors of A is the **row space of** A**,** row(A). The subspace of $\mathbf{R}^m$ spanned by the column vectors is the **column space of** A**,** col(A).

Example 168. Let

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

Here,

$$\begin{aligned} \operatorname{row} A = \operatorname{Span}\left(\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}\right) &= \left\{c_1 \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}\right\} \\ &= \left\{\begin{bmatrix} c_1 & c_2 & 0 \end{bmatrix}\right\}, \end{aligned}$$

a dimension 2 subspace of $\mathbf{R}^{1\times 3}$. Furthermore,

$$\begin{aligned} \operatorname{col} A = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) &= \operatorname{Span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\ &= \left\{c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\} \\ &= \left\{\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}\right\} \\ &= \mathbf{R}^2. \end{aligned}$$

Proposition 169. *If A is row equivalent to B, then* $\operatorname{row} A = \operatorname{row} B$.

Definition 170. $\dim(\operatorname{row}(A))$ is called the **rank of A,** $\operatorname{rank}(A)$.

Example 171. If

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

what is $\operatorname{rank} A$?

By **Proposition 169**, we can find $\operatorname{rank} A$ by reducing to row echelon form, and finding that row space.

$$\begin{bmatrix} 1 & -2 & 3 \\ 2 & -5 & 1 \\ 1 & -4 & -7 \end{bmatrix} \overset{2R_1 - R_2}{\sim} \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 1 & -4 & -7 \end{bmatrix} \overset{R_1 - R_2}{\sim} \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 2 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus

$$\operatorname{row} A = \operatorname{Span}\left(\begin{bmatrix} 1 & -2 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 5 \end{bmatrix}\right),$$

and hence $\operatorname{rank} A = 2$.

Theorem 172. *An $m \times n$ system $A\overline{x} = \overline{b}$ is consistent if and only if $\overline{b} \in \operatorname{col} A$.*

Proof. **Theorem 47** says $A\overline{x} = \overline{b}$ is consistent if and only if $\overline{b}$ is a linear combination of the columns $\overline{a_1}, \ldots, \overline{a_n}$, because you can rewrite

$$\begin{aligned} A\overline{x} &= \overline{b} \\ x_1\overline{a_1} + \cdots + x_n\overline{a_n} &= \overline{b}. \end{aligned}$$

But a linear combination of $\{\overline{a_1}, \ldots, \overline{a_n}\}$, by definition, lives in

$$\operatorname{Span}(\overline{a_1}, \ldots, \overline{a_n}) = \operatorname{col} A.$$

□

Corollary 173. *Let $A \in \mathbf{R}^{m\times n}$.*

1. *$A\overline{x} = \overline{b}$ is consistent for* every *$\overline{b} \in \mathbf{R}^m$ if and only if $\operatorname{col} A = \mathbf{R}^m$.*
2. *$A\overline{x} = \overline{b}$ has at most one solution for every $\overline{b} \in \mathbf{R}^m$ if and only if the column vectors of A are linearly independent.*

Proof.

1. By **Theorem 172**, $\overline{b} \in \operatorname{col} A$, but if that has to hold for *every* $\overline{b} \in \mathbf{R}^m$, then $\operatorname{col} A = \mathbf{R}^m$.
2. $A\overline{x} = \overline{0}$ has exactly one solution, the trivial solution $\overline{x} = \overline{0}$, if and only if the columns of A must be linearly independent. To show this holds for any $\overline{b}$, not just $\overline{0}$, see that if $\overline{y}$ and $\overline{z}$ are two solutions of $A\overline{x} = \overline{b}$, then

$$A(\overline{y} - \overline{z}) = A\overline{y} - A\overline{z} = \overline{b} - \overline{b} = \overline{0},$$

so:

$$\begin{aligned} \text{columns of } A \text{ are linearly independent} &\Leftrightarrow A\overline{x} = \overline{0} \text{ has a unique solution } \overline{x} = \overline{0} \\ &\Leftrightarrow \overline{y} - \overline{z} = 0 \\ &\Leftrightarrow \overline{y} = \overline{z} \\ &\Leftrightarrow A\overline{x} = \overline{b} \text{ has at most one solution.} \end{aligned}$$

□

Corollary 174. *A square matrix $A \in \mathbf{R}^{n\times n}$ is nonsingular if and only if the column vectors form a basis of $\mathbf{R}^n$.*

Now, we do THE MOST IMPORTANT THING OF YOUR LIFE:

Theorem 175 (Rank-Nullity)**.** *Let $A \in \mathbf{R}^{m\times n}$.*

$$\operatorname{rank} A + \dim \operatorname{nul} A = n.$$

It's called rank nullity because we say $\dim \operatorname{nul} A$ is the *nullity* of A. In words:

$$\underbrace{\text{The dimension of the row space}}_{\operatorname{rank} A} + \underbrace{\text{the dimension of the null space}}_{\dim \operatorname{nul} A} = \underbrace{\text{the number of columns}}_{n}.$$

Proof. Given any A, row reduce it to U. We know by **Proposition 169** that $\operatorname{row} A = \operatorname{row} U$. We also know that

$$[A \mid 0] \sim [U \mid 0].$$

Suppose $\operatorname{rank} A = r$. This means U will have r nonzero rows. Thus $[U \mid 0]$ will have r lead variables and $n - r$ free variables, and the number of free variables is $\dim \operatorname{nul} A$. Therefore

$$\operatorname{rank} A + \dim \operatorname{nul} A = r + n - r = n.$$

□

Example 176. Confirm rank-nullity for

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

by finding a basis for $\operatorname{row} A$ and a basis for $\operatorname{nul} A$.

We calculate the reduced row echelon form:

$$\begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 4 & -3 & 0 \\ 1 & 2 & 1 & 5 \end{bmatrix} \overset{2R_1 - R_2}{\sim} \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & 0 & 1 & 2 \\ 1 & 2 & 1 & 5 \end{bmatrix} \overset{R_3 - R_1}{\sim} \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \overset{R_1 + R_2}{\sim} \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Therefore

$$\operatorname{row} A = \operatorname{Span}\left(\begin{bmatrix}1 & 2 & 0 & 3\end{bmatrix}, \begin{bmatrix}0 & 0 & 1 & 2\end{bmatrix}\right),$$

so $\operatorname{rank} A = 2$. Furthermore, we learn that solving $[A \mid 0]$ is equivalent to

$$\begin{aligned} x_1 + 2x_2 + 3x_4 &= 0 \\ x_3 + 2x_4 &= 0, \end{aligned}$$

so

$$\begin{aligned} x_1 &= -2x_2 - 3x_4, \\ x_3 &= -2x_4. \end{aligned}$$

Therefore,

$$\operatorname{nul} A = \left\{ \begin{bmatrix} -2x_2 - 3x_4 \\ x_2 \\ -2x_4 \\ x_4 \end{bmatrix} \right\} = \left\{ x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ -2 \\ 1 \end{bmatrix} \right\} = \operatorname{Span}\left(\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ -2 \\ 1 \end{bmatrix} \right),$$

so $\dim \operatorname{nul} A = 2$. And yes, $\operatorname{rank} A + \dim \operatorname{nul} A = 2 + 2 = 4 = n$.

Remark 177. When A is row equivalent to B, the row spaces are the same, but the column spaces need not be. However, we do have the same dependency relationships. In **Example 176**, notice that for

$$\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \overline{u_1} & \overline{u_2} & \overline{u_3} & \overline{u_4} \end{bmatrix},$$

$\overline{u_1}$ and $\overline{u_3}$ are linearly independent, and $\overline{u_2} = 2\overline{u_1}$ and $\overline{u_4} = 3\overline{u_1} + 2\overline{u_3}$.

The same is true for the original

$$A = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 4 & -3 & 0 \\ 1 & 2 & 1 & 5 \end{bmatrix} = \begin{bmatrix} \overline{a_1} & \overline{a_2} & \overline{a_3} & \overline{a_4} \end{bmatrix};$$

$\overline{a_1}$ and $\overline{a_3}$ are linearly independent, and $\overline{a_2} = 2\overline{a_1}$ and $\overline{a_4} = 3\overline{a_1} + 2\overline{a_3}$.

This fact always holds.

Theorem 178. *If* $A \in \mathbf{R}^{m \times n}$*, then* $\dim \operatorname{row} A = \dim \operatorname{col} A$.

Note that by definition, $\operatorname{rank} A = \dim \operatorname{row} A$. This says you can calculate rank using columns, if you want.

Proof. Row reduce A to a matrix U.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{bmatrix} \sim U = \begin{bmatrix} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

U will have $\operatorname{rank} A = r$ leading 1s, and the columns that those 1s are in will be linearly independent. Let $\widehat{U}$ be the matrix where we delete all the columns with free variables from U. Write $\widehat{A}$ by deleting the same columns.

$$\widehat{A} = \begin{bmatrix} a_{11} & a_{12} & a_{14} \\ a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{bmatrix} \sim \widehat{U} = \begin{bmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Now the columns of $\widehat{U}$ are linearly independent, and since $\widehat{A} \sim \widehat{U}$, the columns of $\widehat{A}$ must also be linearly independent by **Remark 177**. But we've constructed $\widehat{A}$ to have r columns. Thus we've shown that $\dim \operatorname{col} A \geq \dim \operatorname{col} \widehat{A} = r$.

We're done if we can also show that $\dim \operatorname{col} A \leq r$, because that'll force $\dim \operatorname{col} A = r$, and remember: $r = \operatorname{rank} A = \dim \operatorname{row} A$.

To show $\dim \operatorname{col} A \leq r$, we just play the same game with A^T:

$$\dim \operatorname{col} A = \dim \operatorname{row} A^T = \dim \operatorname{col} \widehat{A^T} \leq \dim \operatorname{col} A^T = \dim \operatorname{row} A = r. \checkmark$$

$\square$

Finally, we round off with some examples.

Example 179. Find a basis for $\operatorname{col} A$ if

$$A = \begin{bmatrix} 1 & -2 & 1 & 2 \\ -1 & 3 & 2 & -2 \\ 0 & 1 & 3 & 4 \\ 1 & 2 & 13 & 5 \end{bmatrix}.$$

We row reduce:

$$\begin{bmatrix} 1 & -2 & 1 & 2 \\ -1 & 3 & 2 & -2 \\ 0 & 1 & 3 & 4 \\ 1 & 2 & 13 & 5 \end{bmatrix} \overset{R_1+R_2}{\sim} \begin{bmatrix} 1 & -2 & 1 & 2 \\ 0 & 1 & 3 & 0 \\ 0 & 1 & 3 & 4 \\ 1 & 2 & 13 & 5 \end{bmatrix}$$
$$\overset{R_4-R_1}{\sim} \begin{bmatrix} 1 & -2 & 1 & 2 \\ 0 & 1 & 3 & 0 \\ 0 & 1 & 3 & 4 \\ 0 & 4 & 12 & 3 \end{bmatrix}$$
$$\overset{R_3-R_2}{\sim} \begin{bmatrix} 1 & -2 & 1 & 2 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 4 & 12 & 3 \end{bmatrix}$$
$$\overset{R_4-4R_2}{\sim} \begin{bmatrix} 1 & -2 & 1 & 2 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & -2 & 1 & 2 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The linearly independent columns come from the leading 1s. Thus a basis for $\operatorname{col} A$ is

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 4 \\ 5 \end{bmatrix} \right\}.$$

Example 180. What is the dimension of

$$\operatorname{Span}\left(\begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ -3 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 8 \\ -5 \\ 4 \end{bmatrix} \right) \subseteq \mathbf{R}^4?$$

This span is the same as the column space of

$$A = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 2 & 5 & 4 & 8 \\ -1 & -3 & -2 & -5 \\ 0 & 2 & 0 & 4 \end{bmatrix},$$

which row reduces to:

$$\begin{bmatrix} 1 & 2 & 2 & 3 \\ 2 & 5 & 4 & 8 \\ -1 & -3 & -2 & -5 \\ 0 & 2 & 0 & 4 \end{bmatrix} \overset{R_2-2R_1}{\sim} \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 1 & 0 & 2 \\ -1 & -3 & -2 & -5 \\ 0 & 2 & 0 & 4 \end{bmatrix} \overset{R_1+R_3}{\sim} \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & -1 & 0 & -2 \\ 0 & 2 & 0 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

There are two leading 1s and thus $\dim \operatorname{col} A = 2$.

Homework 13. §3.6: 1, 2, 3, 4, 5

Day 14 of 24 – §4.1 Linear transformations ☐

We know vector spaces; we now want to know a special kind of function between vector spaces.

Definition 181. Let V and W be vector spaces. A function $L : V \to W$ is a **linear transformation** if

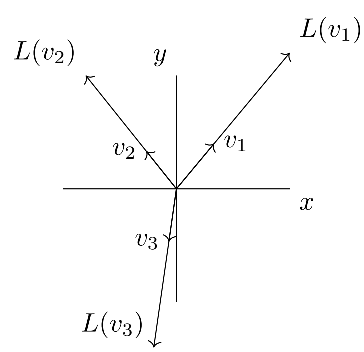
$$L(\alpha v_1 + \beta v_2) = \alpha L(v_1) + \beta L(v_2)$$

for all scalars α, β and vectors $v_1, v_2 \in V$. In words: L respects addition and scalar multiplication. If $L : V \to V$, we call L a **linear operator** on V.

Example 182. Let $L : \mathbf{R}^2 \to \mathbf{R}^2$ be $L(v) = 3v$. We check one at a time:

1. $L(\alpha v) = 3(\alpha v) = \alpha 3v = \alpha L(v)$. ✓
2. $L(v + w) = 3(v + w) = 3v + 3w = L(v) + L(w)$. ✓

Graphically, L stretches out $\mathbf{R}^2$ by a factor of 3:



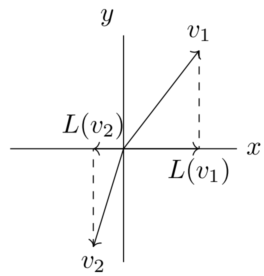
Example 183. Let $L : \mathbf{R}^2 \to \mathbf{R}^2$ be

$$L\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = a\overline{e_1} = \begin{bmatrix} a \\ 0 \end{bmatrix}.$$

We check all at once:

$$L\left(\alpha \begin{bmatrix} a \\ b \end{bmatrix} + \beta \begin{bmatrix} c \\ d \end{bmatrix}\right) = L\left(\begin{bmatrix} \alpha a + \beta c \\ \alpha b + \beta d \end{bmatrix}\right) = \begin{bmatrix} \alpha a + \beta c \\ 0 \end{bmatrix} = \alpha \begin{bmatrix} a \\ 0 \end{bmatrix} + \beta \begin{bmatrix} c \\ 0 \end{bmatrix} = \alpha L\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) + \beta L\left(\begin{bmatrix} c \\ d \end{bmatrix}\right).$$

Graphically, L takes any vector and projects it to the x-axis:



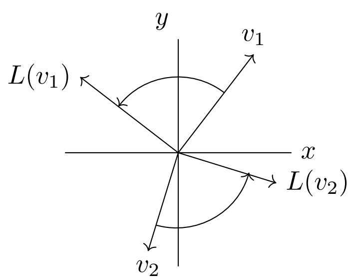
Example 184. Let $L : \mathbf{R}^2 \to \mathbf{R}^2$ be

$$L\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} -b \\ a \end{bmatrix}.$$

Again:

$$L\left(\alpha \begin{bmatrix} a \\ b \end{bmatrix} + \beta \begin{bmatrix} c \\ d \end{bmatrix}\right) = L\left(\begin{bmatrix} \alpha a + \beta c \\ \alpha b + \beta d \end{bmatrix}\right) = \begin{bmatrix} -(\alpha b + \beta d) \\ \alpha a + \beta c \end{bmatrix} = \alpha \begin{bmatrix} -b \\ a \end{bmatrix} + \beta \begin{bmatrix} -d \\ c \end{bmatrix} = \alpha L\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) + \beta L\left(\begin{bmatrix} c \\ d \end{bmatrix}\right).$$

Graphically:



Example 185. What about the distance formula $d : \mathbf{R}^2 \to \mathbf{R}$,

$$d\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \sqrt{x^2 + y^2}?$$

1.

$$\begin{aligned} d\left(\alpha \begin{bmatrix} x \\ y \end{bmatrix}\right) &= d\left(\begin{bmatrix} \alpha x \\ \alpha y \end{bmatrix}\right) \\ &= \sqrt{(\alpha x)^2 + (\alpha y)^2} \\ &= \sqrt{\alpha^2(x^2 + y^2)} \\ &= |\alpha|\sqrt{x^2 + y^2} \\ &\neq \alpha\sqrt{x^2 + y^2} \\ &= \alpha d\left(\begin{bmatrix} x \\ y \end{bmatrix}\right). \times \end{aligned}$$

Example 186. What about calculus? Let $D : P_{n+1} \to P_n$ be $D(f) = \frac{d}{dx}[f]$. Then the sum rule and constant multiple rules tell us

$$D(\alpha f + \beta g) = \frac{d}{dx}[\alpha f + \beta g] = \alpha \frac{d}{dx}[f] + \beta \frac{d}{dx}[g] = \alpha D(f) + \beta D(g). \checkmark$$

(Also, limits and integrals are linear transformations, and we need not restrict to P_{n+1}.)

Example 187. What about $L : \mathbf{R}^2 \to \mathbf{R}^3$ defined by

$$L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x \\ x+y \end{bmatrix}?$$

1.

$$L\left(\alpha \begin{bmatrix} x \\ y \end{bmatrix}\right) = L\left(\begin{bmatrix} \alpha x \\ \alpha y \end{bmatrix}\right) = \begin{bmatrix} \alpha y \\ \alpha x \\ \alpha x + \alpha y \end{bmatrix} = \alpha \begin{bmatrix} y \\ x \\ x+y \end{bmatrix} = \alpha L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right). \checkmark$$

2.

$$\begin{aligned} L\left(\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} z \\ w \end{bmatrix}\right) &= L\left(\begin{bmatrix} x+z \\ y+w \end{bmatrix}\right) \\ &= \begin{bmatrix} y+w \\ x+z \\ x+z+y+w \end{bmatrix} \\ &= \begin{bmatrix} y \\ x \\ x+y \end{bmatrix} + \begin{bmatrix} w \\ z \\ z+w \end{bmatrix} \\ &= L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) + L\left(\begin{bmatrix} z \\ w \end{bmatrix}\right). \checkmark \end{aligned}$$

Remark 188. Hey, check this out:
Let

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix},$$

and calculate

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0x+1y \\ 1x+0y \\ 1x+1y \end{bmatrix} = \begin{bmatrix} y \\ x \\ x+y \end{bmatrix}.$$

So the linear transformation L in **Example 187** and the A we just defined satisfy $L(\overline{v}) = A\overline{v}$. Cool.

In fact, in general, all linear transformations can be written as matrices, and in fact all matrices are linear transformations. Wow! Let's prove the second claim, and the first later (**Theorem 196**).

Theorem 189. *If $A \in \mathbf{R}^{m\times n}$, then multiplication by A defines a linear transformation $L : \mathbf{R}^n \to \mathbf{R}^m$.*

Proof. Let $L(v) = Av$.
1. $L(\alpha v) = A(\alpha v) = \alpha Av = \alpha L(v)$. ✓
2. $L(v+w) = A(v+w) = Av + Aw = L(v) + L(w)$. ✓

□

Remark 190. If $L : V \to W$, then it's easy to see
1. $L(0_V) = 0_W$.
2. $L(c_1v_1 + \cdots + c_nv_n) = c_1L(v_1) + \cdots + c_nL(v_n)$.
3. $L(-v) = -L(v)$.

Definition 191. Let $L : V \to W$ be a linear transformation. We define the **kernel of L,** $\ker L$, to be

$$\ker L = \{v \in V \mid L(v) = 0_W\} \subseteq V.$$

It's everything in V that gets sent to 0.

Definition 192. Let $L : V \to W$. Let $S \subseteq V$. Define the **image of S, $L(S)$**, to be

$$L(S) = \{w \in W \mid w = L(s) \text{ for some } s \in S\} \subseteq W.$$

It's everything that S maps to. We call the image of V, $L(V)$, the **range of L**, or sometimes the **image of L, $\operatorname{im} L$**.

Example 193. Recall $L : \mathbf{R}^2 \to \mathbf{R}^2$ defined by

$$L\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} a \\ 0 \end{bmatrix}$$

from **Example 183**. First,

$$\begin{aligned}
\ker L &= \left\{\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbf{R}^2 \mid L\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} a \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\} \\
&= \left\{\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbf{R}^2 \mid a = 0\right\} \\
&= \left\{\begin{bmatrix} 0 \\ b \end{bmatrix} \in \mathbf{R}^2\right\} \\
&= \operatorname{Span}(\overline{e_2}) \subseteq \mathbf{R}^2.
\end{aligned}$$

Second,

$$\begin{aligned}
L\left(\mathbf{R}^2\right) &= \left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 \mid L\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} a \\ 0 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}\right\} \\
&= \left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 \mid a = x, y = 0\right\} \\
&= \left\{\begin{bmatrix} a \\ 0 \end{bmatrix} \in \mathbf{R}^2\right\} \\
&= \operatorname{Span}(\overline{e_1}) \subseteq \mathbf{R}^2.
\end{aligned}$$

Example 194. Let $L : \mathbf{R}^3 \to \mathbf{R}^2$ be

$$L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + y \\ y + z \end{bmatrix}.$$

The kernel is

$$\begin{aligned}
\ker L &= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbf{R}^3 \mid L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = \begin{bmatrix} x+y \\ y+z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\
&= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbf{R}^3 \mid x+y=0, y+z=0 \right\} \\
&= \left\{ \begin{bmatrix} z \\ -z \\ z \end{bmatrix} \in \mathbf{R}^3 \right\} \\
&= \operatorname{Span}\left(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right) \subseteq \mathbf{R}^3.
\end{aligned}$$

Let $S = \operatorname{Span}(\overline{e_1}, \overline{e_3}) \subseteq \mathbf{R}^3$. The image of S is

$$\begin{aligned}
L(S) &= \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbf{R}^2 \mid L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = \begin{bmatrix} x+y \\ y+z \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in S \right\} \qquad \text{(so } y=0\text{)} \\
&= \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbf{R}^2 \mid \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \right\} \\
&= \left\{ \begin{bmatrix} x \\ z \end{bmatrix} \in \mathbf{R}^2 \right\} \\
&= \mathbf{R}^2.
\end{aligned}$$

Notice that since $S \subseteq \mathbf{R}^3$ has image all of $\mathbf{R}^2$, this forces $L(\mathbf{R}^3) = \mathbf{R}^2$ as well.

Theorem 195. *If $L : V \to W$ and $S \subseteq V$ is a subspace, then*

1. $\ker L \subseteq V$ *is a subspace.*
2. $L(S) \subseteq W$ *is a subspace.*

In particular, $\operatorname{im} L = L(V)$ *is a subspace of* W.

Proof. We gotta check:

0. Is $\ker L \neq \emptyset$? Yes, by **Remark 190**, $L(0_V) = 0_W$, so $0_V \in \ker L$. ✓
1. If $v \in \ker L$, is $\alpha v \in \ker L$? Yes: $L(\alpha v) = \alpha L(v) = \alpha 0_W = 0_W$. ✓
2. If $v, v' \in \ker L$, is $v + v' \in \ker L$? Yes: $L(v + v') = L(v) + L(v') = 0_W + 0_W = 0_W$. ✓

So $\ker L$ is a subspace of V.

0. Is $L(S) \neq \emptyset$? Yes, since $0_V \in S$, $L(0_V) = 0_W \in L(S)$. ✓
1. If $w \in L(S)$, is $\alpha w \in L(S)$? Yes: since $w \in L(S)$, there is a $v \in S$ such that $L(v) = w$. Since S is a subspace, $\alpha v \in S$, and then $L(\alpha v) = \alpha L(v) = \alpha w$, so there is an $\alpha v \in S$ demonstrating $\alpha w \in L(S)$. ✓
2. If $w, w' \in L(S)$, is $w + w' \in L(S)$? Yes: since $w, w' \in L(S)$, there are $v, v' \in S$ such that $L(v) = w$ and $L(v') = w'$. Since S is a subspace, $v + v' \in S$, and then $L(v + v') = L(v) + L(v') = w + w'$, so there is a $v + v' \in S$ demonstrating $w + w' \in L(S)$. ✓

So $L(S)$ is a subspace of W. □

Homework 14. §4.1: 1, 5, 13, 17

Day 15 of 24 – §4.2 Matrix representations of transformations □

We saw in **Theorem 189** that every matrix determines a linear transformation, and we claimed the opposite is also true: you can write every linear transformation as multiplication by a matrix. Let's prove it!

Theorem 196. *If $L : \mathbf{R}^n \to \mathbf{R}^m$, then there exists $A \in \mathbf{R}^{m \times n}$ such that*

$$L(v) = Av$$

for all $v \in \mathbf{R}^n$.

Proof. We build A's columns:

$$A = \begin{bmatrix} \overline{a_1} & \overline{a_2} & \cdots & \overline{a_n} \end{bmatrix} = \begin{bmatrix} L(\overline{e_1}) & L(\overline{e_2}) & \cdots & L(\overline{e_n}) \end{bmatrix}.$$

And then we just check. Take any $v \in \mathbf{R}^n$ and write out its standard coordinate vector:

$$v = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = c_1\overline{e_1} + c_2\overline{e_2} + \cdots + c_n\overline{e_n}.$$

By **Remark 190**,

$$\begin{aligned} L(v) &= L(c_1\overline{e_1} + c_2\overline{e_2} + \cdots + c_n\overline{e_n}) \\ &= c_1L(\overline{e_1}) + c_2L(\overline{e_2}) + \cdots + c_nL(\overline{e_n}) \\ &= c_1\overline{a_1} + c_2\overline{a_2} + \cdots + c_n\overline{a_n} \\ &= A\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \\ &= Av. \checkmark \end{aligned}$$

$\square$

Example 197. Recall **Example 194** where $L : \mathbf{R}^3 \to \mathbf{R}^2$ was

$$L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ y+z \end{bmatrix}.$$

To find a corresponding matrix $A \in \mathbf{R}^{2\times 3}$, we have

$$L\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$L\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$L\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Therefore,

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

We can check our work:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1x+1y+0z \\ 0x+1y+1z \end{bmatrix} = \begin{bmatrix} x+y \\ y+z \end{bmatrix} = L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right). \checkmark$$

Remark 198. In fact, it's not just transformations $L : \mathbf{R}^n \to \mathbf{R}^m$ that can be turned into matrices. What about $L : V \to W$, for whatever V and W? It's the same idea. We just might not have the *standard* basis anymore.

Example 199. Let

$$L : \mathbf{R}^3 = \langle \overline{e_1}, \overline{e_2}, \overline{e_3} \rangle \to \mathbf{R}^2 = \left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\rangle$$

be defined by

$$L\left(a\overline{e_1} + b\overline{e_2} + c\overline{e_3}\right) = a \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (b+c) \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

To write the matrix $A \in \mathbf{R}^{2\times 3}$ that defines L, we plug in the basis of $\mathbf{R}^3$ and write the columns of A in terms of the basis of $\mathbf{R}^2$:

$$\begin{aligned} L(\overline{e_1}) &= 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ L(\overline{e_2}) &= 0 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ L(\overline{e_3}) &= 0 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \end{aligned}$$

So the matrix A is

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

Example 200. Let

$$L : \mathbf{R}^2 = \left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\rangle \to \mathbf{R}^2 = \left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\rangle$$

be defined by

$$L\left(\alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right) = (\alpha+\beta) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2\beta \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

We get:

$$\begin{aligned} L\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) &= 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ L\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right) &= 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \end{aligned}$$

so

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}.$$

Example 201. Recall (**Example 117** and **Example 186**) that P_{n+1}, polynomials of degree less than $n+1$, is a vector space, and differentiation $D : P_{n+1} \to P_n$ is a linear operator. Write

$$\begin{aligned} P_{n+1} &= \langle x^n, x^{n-1}, \ldots, x^2, x, 1 \rangle \\ P_n &= \langle x^{n-1}, x^{n-2}, \ldots, x^2, x, 1 \rangle. \end{aligned}$$

We calculate the value of the basis vectors:

$$\begin{aligned} D(x^n) &= nx^{n-1} \\ D(x^{n-1}) &= (n-1)x^{n-2} \\ &\vdots \\ D(x^2) &= 2x \\ D(x) &= 1 \\ D(1) &= 0 \end{aligned}$$

Therefore the matrix is

$$A = \begin{bmatrix} n & 0 & \cdots & 0 & 0 & 0 \\ 0 & n-1 & \cdots & 0 & 0 & 0 \\ & \vdots & & & \vdots & \\ 0 & 0 & \cdots & 2 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

So now check this shit out!!

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

and indeed,

$$\frac{d}{dx}\left[2x^4 + 1x^3 + 0x^2 - 1x + 3\right] = 8x^3 + 3x^2 + 0x - 1.$$

Wow! ☺

Homework 15. §4.2: 1, 2, 6, 13

Day 16 of 24 – §4.3 Similarity ☐

Remark 202. Recall that when we write the matrix $A \in \mathbf{R}^{m \times n}$ representing a linear transformation $L : V \to W$, A depends on the bases $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_m\}$. But of course V and W could have different bases. Changing bases doesn't change L, so there should be a relationship between the matrices.

Example 203. Let $L : \mathbf{R}^2 \to \mathbf{R}^2$,

$$L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x \\ x+y \end{bmatrix}.$$

The matrix with respect to $\{\overline{e_1}, \overline{e_2}\}$ is

$$\begin{aligned} L(\overline{e_1}) &= \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ L(\overline{e_2}) &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ A &= \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}. \end{aligned} \qquad (A \text{ represents } L : \langle \overline{e_1}, \overline{e_2} \rangle \to \langle \overline{e_1}, \overline{e_2} \rangle.)$$

If we have a different basis

$$\{\overline{b_1}, \overline{b_2}\} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\},$$

then we can write the two change of basis transition matrices from $\{\overline{e_1}, \overline{e_2}\}$ to $\{\overline{b_1}, \overline{b_2}\}$ and $\{\overline{b_1}, \overline{b_2}\}$ to $\{\overline{e_1}, \overline{e_2}\}$. The matrix from $\{\overline{b_1}, \overline{b_2}\}$ to $\{\overline{e_1}, \overline{e_2}\}$ is

$$C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

so from $\{\overline{e_1}, \overline{e_2}\}$ to $\{\overline{b_1}, \overline{b_2}\}$ is

$$C^{-1} = \frac{1}{1+1} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix}.$$

Now if we want to express $L : \langle \overline{b_1}, \overline{b_2} \rangle \to \langle \overline{b_1}, \overline{b_2} \rangle$, then we need to know what $L(\overline{b_1})$ and $L(\overline{b_2})$ are, in terms of $\overline{b_1}$ and $\overline{b_2}$. We can calculate:

$$\begin{aligned} L\left(\overline{b_1}\right) &= L\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \\ L\left(\overline{b_2}\right) &= L\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}. \end{aligned}$$

These outputs are in terms of $\{\overline{e_1}, \overline{e_2}\}$, so to finish:

$$\begin{aligned} C^{-1}L(\overline{b_1}) &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \\ C^{-1}L(\overline{b_2}) &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \end{aligned}$$

so thus

$$\begin{aligned} L(\overline{b_1}) &= 2\overline{b_1} + 0\overline{b_2} \\ L(\overline{b_2}) &= -1\overline{b_1} + 1\overline{b_2}, \\ B &= \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}. \qquad\qquad (B \text{ represents } L : \langle \overline{b_1}, \overline{b_2} \rangle \to \langle \overline{b_1}, \overline{b_2} \rangle.) \end{aligned}$$

How do we compare A and B? Notice:

$$\begin{tikzcd} \langle \overline{b_1}, \overline{b_2} \rangle \arrow[r, "B"] \arrow[d, "C"'] & \langle \overline{b_1}, \overline{b_2} \rangle \\ \langle \overline{e_1}, \overline{e_2} \rangle \arrow[r, "A"'] & \langle \overline{e_1}, \overline{e_2} \rangle \arrow[u, "C^{-1}"'] \end{tikzcd}$$

$$B = C^{-1}AC.$$

This always holds.

Proposition 204. *Let $L : V \to V$ and give V two bases $\{v_1, \ldots, v_n\}$ and $\{\widetilde{v_1}, \ldots, \widetilde{v_n}\}$. Let S be the transition matrix from $\{\widetilde{v_1}, \ldots, \widetilde{v_n}\}$ to $\{v_1, \ldots, v_n\}$. If A represents L on $\{v_1, \ldots, v_n\}$ and B represents L on $\{\widetilde{v_1}, \ldots, \widetilde{v_n}\}$, then*

$$B = S^{-1}AS.$$

$$\begin{array}{ccc}
\langle \widetilde{v_1}, \ldots, \widetilde{v_n} \rangle & \xrightarrow{B} & \langle \widetilde{v_1}, \ldots, \widetilde{v_n} \rangle \\
{\scriptstyle S}\big\downarrow & & \big\uparrow{\scriptstyle S^{-1}} \\
\langle v_1, \ldots, v_n \rangle & \xrightarrow[A]{} & \langle v_1, \ldots, v_n \rangle
\end{array}$$

Definition 205. Let $A, B \in \mathbf{R}^{n\times n}$. If there exists a nonsingular $S \in \mathbf{R}^{n\times n}$ such that

$$B = S^{-1}AS,$$

then A and B are **similar**.

Example 206. Let $L : \mathbf{R}^3 \to \mathbf{R}^3$ be defined by

$$L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix};$$

this matrix represents $L : \langle \overline{e_1}, \overline{e_2}, \overline{e_3} \rangle \to \langle \overline{e_1}, \overline{e_2}, \overline{e_3} \rangle$. Find a matrix representing

$$L : \left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\rangle \to \left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\rangle.$$

The transition matrix from the new basis to $\{\overline{e_1}, \overline{e_2}, \overline{e_3}\}$ is

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

and we can find the inverse:

$$
\left[\begin{array}{rrr|rrr} 1 & -2 & 1 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right]
\overset{R_1+R_2}{\sim}
\left[\begin{array}{rrr|rrr} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right]
$$

$$
\overset{R_2+R_3}{\sim}
\left[\begin{array}{rrr|rrr} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 3 & 1 & 1 & 1 \end{array}\right]
$$

$$
\overset{-R_2}{\sim}
\left[\begin{array}{rrr|rrr} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -1 & -1 & 0 \\ 0 & 0 & 3 & 1 & 1 & 1 \end{array}\right]
$$

$$
\overset{2R_2+R_1}{\sim}
\left[\begin{array}{rrr|rrr} 1 & 0 & -3 & -1 & -2 & 0 \\ 0 & 1 & -2 & -1 & -1 & 0 \\ 0 & 0 & 3 & 1 & 1 & 1 \end{array}\right]
$$

$$
\overset{R_1+R_3}{\sim}
\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & -2 & -1 & -1 & 0 \\ 0 & 0 & 3 & 1 & 1 & 1 \end{array}\right]
$$

$$
\overset{\frac{1}{3}R_3}{\sim}
\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & -2 & -1 & -1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}\right]
$$

$$
\overset{2R_3+R_2}{\sim}
\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & \frac{-1}{3} & \frac{-1}{3} & \frac{2}{3} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}\right],
$$

so

$$
S^{-1} = \begin{bmatrix} 0 & -1 & 1 \\ \frac{-1}{3} & \frac{-1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}.
$$

Thus

$$
\begin{aligned}
S^{-1}AS &= \begin{bmatrix} 0 & -1 & 1 \\ \frac{-1}{3} & \frac{-1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 2 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & -1 & 1 \\ \frac{-1}{3} & \frac{-1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 0 & -2 & 4 \\ 0 & 1 & 4 \\ 0 & 1 & 4 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}.
\end{aligned}
$$

Homework 16. §4.3: 1, 2, 4

Day 17 of 24 – §5.1 The scalar product

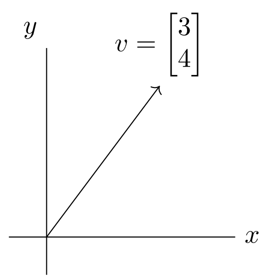
Definition 207. Let $v, w \in \mathbf{R}^{n\times 1}$. The **scalar product of v and w** is

$$v^T w = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n \in \mathbf{R}.$$

Example 208.

$$\begin{bmatrix} 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix} = 3(4) - 2(3) + 1(2) = 8.$$

Example 209. What is the length of



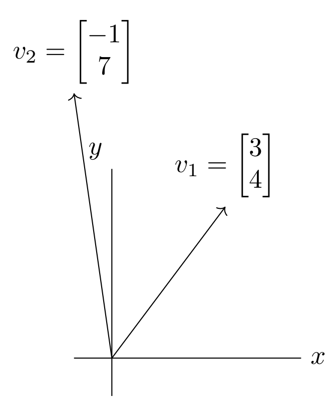
?

We know via the Pythagorean theorem that the length is

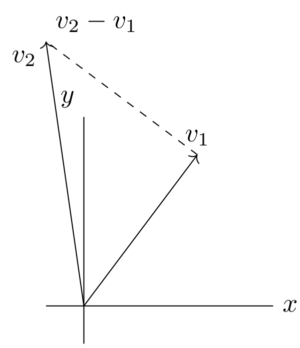
$$\sqrt{(3-0)^2 + (4-0)^2} = \sqrt{3^2 + 4^2} = \sqrt{3(3) + 4(4)} = \sqrt{\begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix}} = \sqrt{v^T v}.$$

Definition 210. We write $\|v\|$ for the **length** or **norm** of v. By above, $\|v\| = \sqrt{v^T v}$. (Consequently, $\|v\|^2 = v^T v$.)

Example 211. What about the distance between two vectors?



The distance between them is just the length of



So the distance from v_1 to v_2 is

$$\begin{aligned}
\|v_2 - v_1\| &= \sqrt{(v_2 - v_1)^T (v_2 - v_1)} \\
&= \sqrt{\begin{bmatrix} -1-3 & 7-4 \end{bmatrix} \begin{bmatrix} -1-3 \\ 7-4 \end{bmatrix}} \\
&= \sqrt{\begin{bmatrix} -4 & 3 \end{bmatrix} \begin{bmatrix} -4 \\ 3 \end{bmatrix}} \\
&= \sqrt{(-4)^2 + 3^2} \\
&= \sqrt{16 + 9} \\
&= \sqrt{25} \\
&= 5.
\end{aligned}$$

Definition 212. The distance between any v_1 and v_2 is $\underline{\|v_2 - v_1\|}$.

Proposition 213. *Let v_1, v_2 be vectors. If θ is the angle between v_1 and v_2, then*

$$v_1^T v_2 = \|v_1\| \|v_2\| \cos\theta.$$

Remark 214. You can use **Proposition 213** to find the angle between vectors:

$$\begin{aligned}
v_1^T v_2 &= \|v_1\| \|v_2\| \cos\theta \\
\frac{v_1^T v_2}{\|v_1\| \|v_2\|} &= \cos\theta.
\end{aligned}$$

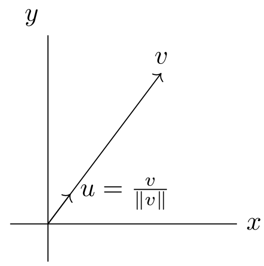
We say that

$$u_1 = \frac{v_1}{\|v_1\|}, \qquad u_2 = \frac{v_2}{\|v_2\|}$$

are *unit vectors* in the direction of v_1 and v_2, because their length is 1 and their direction is the same as v_1 and v_2. Indeed, for either u_1 or u_2,

$$\|u_i\| = \left\| \frac{v_i}{\|v_i\|} \right\| = \sqrt{\frac{v_i}{\|v_i\|}^T \frac{v_i}{\|v_i\|}} = \sqrt{\frac{1}{\|v_i\|^2} \cdot v_i^T v_i} = \frac{1}{\|v_i\|} \sqrt{v_i^T v_i} = \frac{1}{\|v_i\|} \|v_i\| = 1.$$

The direction is the same because all you're doing is scaling v_1 and v_2.



So we have

$$\cos\theta = \frac{v_1}{\|v_1\|}^T \frac{v_2}{\|v_2\|} = u_1^T u_2.$$

Example 215. Continuing with the same vectors as **Example 211**, we have

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{3^2+4^2}}\begin{bmatrix}3\\4\end{bmatrix} = \begin{bmatrix}\frac{3}{5}\\ \\ \frac{4}{5}\end{bmatrix}$$
$$u_2 = \frac{v_2}{\|v_2\|} = \frac{1}{\sqrt{(-1)^2+7^2}}\begin{bmatrix}-1\\7\end{bmatrix} = \begin{bmatrix}\frac{-1}{\sqrt{50}}\\ \\ \frac{7}{\sqrt{50}}\end{bmatrix},$$

so

$$\cos\theta = u_1^T u_2 = \begin{bmatrix}\frac{3}{5} & \frac{4}{5}\end{bmatrix}\begin{bmatrix}\frac{-1}{\sqrt{50}}\\ \\ \frac{7}{\sqrt{50}}\end{bmatrix} = \frac{3}{5}\left(\frac{-1}{\sqrt{50}}\right) + \frac{4}{5}\left(\frac{7}{\sqrt{50}}\right) = \frac{-3+28}{5\sqrt{50}} = \frac{25}{5\cdot\sqrt{25}\sqrt{2}} = \frac{1}{\sqrt{2}}.$$

Therefore, $\theta = \pi/4$.

Theorem 216 (Cauchy-Schwarz inequality)**.** *Let v, w be vectors. One has*

$$|v^T w| \leq \|v\|\|w\|,$$

and equality holds if and only if

1. *$v = 0$ or $w = 0$,*
2. *$v = \alpha w$ for some scalar α.*

Proof. By **Proposition 213**,

$$v^T w = \|v\|\|w\|\cos\theta$$
$$|v^T w| = \|v\|\|w\| \cdot |\cos\theta| \leq \|v\|\|w\| \cdot 1 = \|v\|\|w\|.$$

To see equality, it certainly holds if $v = 0$ or $w = 0$; we get $0 = 0$. For #2, suppose $v \neq 0$ and $w \neq 0$ but

$$|v^T w| = \|v\|\|w\|.$$

Then $|\cos\theta| = 1$, so $\cos\theta = \pm 1$, so $\theta = 0$ or π. That means either v and w point in the same direction or in opposite directions, but in both cases, one is a multiple of the other. ☐

Definition 217. If $v^T w = 0$, then by **Proposition 213**, either $v = 0$, $w = 0$, or $\cos\theta = 0$. If $\cos\theta = 0$, then $\theta = \pi/2$ or $3\pi/2$ (a right angle). We say that v and w are **orthogonal** if $v^T w = 0$.

Example 218. 0 is orthogonal to every v:

$$0^T v = 0 \cdot v_1 + \cdots + 0 \cdot v_n = 0.$$

Example 219.

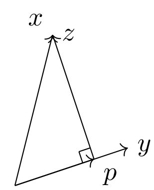
$$\begin{bmatrix} 2 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 2(1) - 3(1) + 1(1) = 0,$$

so

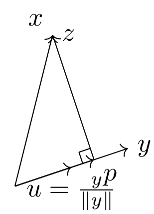
$$\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

are orthogonal.

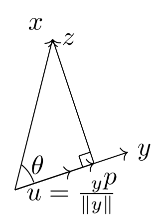
Remark 220. Now suppose you are handed vectors x, y, and you want to write x as $x = p + z$, where p is in the direction of y and z is orthogonal to y.



Let's first calculate the unit vector in the direction of y:



Now we want to find p and z such that $p = \alpha u$ and $z = x - \alpha u$, and $p^T z = 0$. That means we need to find α. Notice that if we let θ be the following angle:



then we must have

$$\begin{aligned} \cos\theta &= \frac{x^T p}{\|x\|\|p\|} \\ &= \frac{(p+z)^T p}{\|x\|\|p\|} \\ &= \frac{p^T p + z^T p}{\|x\|\|p\|} \\ &= \frac{\|p\|^2 + 0}{\|x\|\|p\|} \\ &= \frac{\|p\|}{\|x\|} \\ &= \frac{\alpha}{\|x\|}, \end{aligned}$$

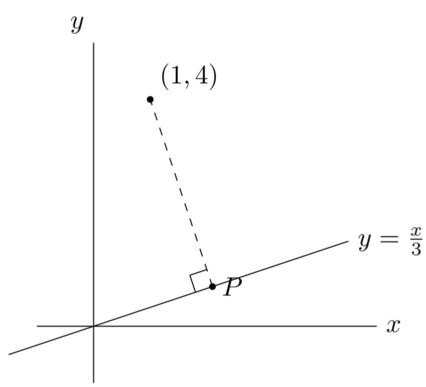
so $\alpha = \|x\| \cos\theta$. But we can simplify further:

$$\begin{aligned}
\alpha &= \|x\| \cos\theta \\
&= \|x\| \cos\theta \frac{\|y\|}{\|y\|} \\
&= \frac{\|x\|\|y\| \cos\theta}{\|y\|} \\
&= \frac{x^T y}{\|y\|}. && (\textbf{Proposition 213})
\end{aligned}$$

Definition 221. Let x and y be vectors.

1. The **scalar projection of x onto y** is
$$\alpha = \frac{x^T y}{\|y\|}.$$
2. The **vector projection of x onto y** is
$$p = \alpha u = \frac{x^T y}{\|y\|} \frac{y}{\|y\|} = \frac{x^T y}{y^T y} y.$$

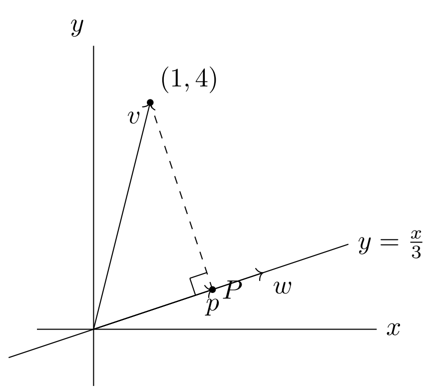
Example 222. Find the point P on the line $y = x/3$ closest to the point $(1, 4)$.



Let

$$v = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \qquad w = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \qquad u = \frac{w}{\|w\|},$$

and p the vector projection of v onto w. We are done when we find p, because the point P lies at the end of the vector P.



We calculate:

$$p = \frac{v^T w}{w^T w} w = \frac{\begin{bmatrix} 1 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix}}{\begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix}} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \frac{1 \cdot 3 + 4 \cdot 1}{3 \cdot 3 + 1 \cdot 1} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \frac{7}{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{21}{10} \\ \frac{7}{10} \end{bmatrix}.$$

So $P = (21/10, 7/10)$.

Homework 17. §5.1: 1, 2, 4, 5

Day 18 of 24 – §5.2 Orthogonal subspaces □

Example 223. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Let

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in N(A);$$

in other words, $Av = 0$. Since $Av = 0$,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The ith entry is the equation

$$0 = a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

So in other words, if we write r_i for the ith row of A, we have

$$r_i v = 0.$$

But notice

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \leftarrow i\text{th row of } A,\ r_i$$

$$\begin{array}{cc} & \begin{array}{cccccc} \phantom{a_{11}} & \phantom{a_{21}} & & \begin{array}{c} i\text{th col.}\\ \text{of } A^T,\ c_i\\ \downarrow \end{array} & & \phantom{a_{m1}} \end{array} \\ A^T = & \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{i1} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{i2} & \cdots & a_{m2} \\ \vdots & & & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{in} & \cdots & a_{mn} \end{bmatrix} \end{array}$$

So $r_i = c_i^T$, and therefore

$$\begin{aligned} r_i v &= 0 \\ c_i^T v &= 0; \end{aligned}$$

i.e., v is orthogonal to any column vector of A^T. Hence v is orthogonal to any linear combination of column vectors of A^T. But a linear combination of column vectors is just the column space $\operatorname{col}(A^T)$.

Thus, every vector in $N(A)$ is orthogonal to every vector in $\operatorname{col} A^T$.

Definition 224. Let $V, W \subseteq \mathbf{R}^n$ be subspaces. We say **V and W are orthogonal** and write $V \perp W$ if $v^T w = 0$ for all $v \in V$ and all $w \in W$.

Example 225. Let $V = \operatorname{Span}(\overline{e_1})$ and $W = \operatorname{Span}(\overline{e_3})$ in $\mathbf{R}^3$. Pick arbitrary $v \in V$ and $w \in W$:

$$v = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \qquad\qquad w = \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix}.$$

Now

$$v^T w = a \cdot 0 + 0 \cdot 0 + 0 \cdot c = 0,$$

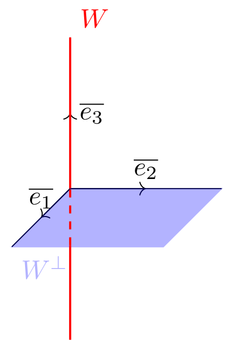
so $V \perp W$.

Definition 226. Let $V \subseteq \mathbf{R}^n$. The set of *all* vectors in $\mathbf{R}^n$ orthogonal to every $v \in V$ is **the orthogonal complement of** V,

$$V^\perp = \left\{ w \in \mathbf{R}^n \mid v^T w = 0 \text{ for all } v \in V \right\}.$$

Example 227. In **Example 225**, $V = \operatorname{Span}(\overline{e_1})$ and $W = \operatorname{Span}(\overline{e_3})$ are *not* orthogonal complements. The orthogonal complement of W in $\mathbf{R}^3$ is

$$\begin{aligned} W^\perp &= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbf{R}^3 \mid \begin{bmatrix} 0 & 0 & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \text{ for all } c \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbf{R}^3 \mid 0 \cdot x + 0 \cdot y + c \cdot z = 0 \text{ for all } c \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbf{R}^3 \mid cz = 0 \text{ for all } c \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \in \mathbf{R}^3 \right\} \\ &= \operatorname{Span}(\overline{e_1}, \overline{e_2}). \end{aligned}$$



Theorem 228.
1. *If $X \perp Y$, then $X \cap Y = \{0\}$.*
2. *If $X \subseteq \mathbf{R}^n$ is a subspace, then $X^\perp \subseteq \mathbf{R}^n$ is a subspace.*

Proof.
1. Let $v \in X \cap Y$. We calculate the length of v:
$$\|v\| = \sqrt{\underbrace{v^T}_{v \in X}\underbrace{v}_{v \in Y}} = \sqrt{\underbrace{0}_{\text{because } X \perp Y}} = 0.$$
So v's length is 0, and v must be 0.
2. We check:
   0. $X^\perp \neq \emptyset$? Yes: $0 \in X^\perp$, because for all $x \in X$, $x^T 0 = 0$. ✓
   1. If $v \in X^\perp$, is $\alpha v \in X^\perp$? Yes: $x^T(\alpha v) = \alpha(x^T v) = \alpha 0 = 0$. ✓
   2. If $v, w \in X^\perp$, is $v + w \in X^\perp$? Yes: $x^T(v+w) = x^T v + x^T w = 0 + 0 = 0$. ✓

□

Remark 229. Recall from **Theorem 172** that $A\overline{x} = \overline{b}$ if and only if $\overline{b} \in \operatorname{col} A$. By **Theorem 189**, we can think of $A\overline{x} = \overline{b}$ as $L(\overline{x}) = \overline{b}$ for $L : \mathbf{R}^n \to \mathbf{R}^m$ a linear transformation. That means that $\overline{b} \in \operatorname{col} A$ is the same as $\overline{b} \in L(\mathbf{R}^n)$, the range of L. We'll write $\operatorname{range}(A) = \operatorname{col}(A)$, or sometimes $R(A)$.

Theorem 230 (Fundamental subspace theorem)**.** *Let $A \in \mathbf{R}^{m \times n}$.*
1. $N(A) = R(A^T)^\perp$.
2. $N(A^T) = R(A)^\perp$.

Proof.
1. We saw in **Example 223** that $N(A) \perp \operatorname{col}(A^T) = N(A) \perp R(A^T)$. This means $N(A) \subseteq R(A^T)^\perp$, since $R(A^T)^\perp$ is the *biggest* subspace orthogonal to $R(A^T)$.
Now, if $v \in R(A^T)^\perp$, then v is orthogonal to every element of $R(A^T) = \operatorname{col}(A^T)$, so v is orthogonal to the column vectors of A^T, so $Av = 0$. Thus $v \in N(A)$, so $R(A^T)^\perp \subseteq N(A)$.
Thus $N(A) = R(A^T)^\perp$.
2. Play the same game, and replace A with A^T.

□

Example 231. Let
$$A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}.$$
We calculate
$$\operatorname{range}(A) = \operatorname{col}(A) = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$$
$$\operatorname{nul}(A^T) = \left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 \mid \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\} = \left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 \mid x + 2y = 0\right\} = \operatorname{Span}\left(\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right).$$

Observe that if $v \in \operatorname{range}(A)$ and $w \in \operatorname{nul}(A^T)$, then

$$v^T w = \begin{bmatrix} \alpha & 2\alpha \end{bmatrix} \begin{bmatrix} -2\beta \\ \beta \end{bmatrix} = \alpha(-2\beta) + 2\alpha\beta = 0. \checkmark$$

Furthermore,

$$\operatorname{range}(A^T) = \operatorname{col}(A^T) = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix} \right) = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \operatorname{Span}(\overline{e_1})$$

$$\operatorname{nul}(A) = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 \mid \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 \mid \begin{bmatrix} x \\ 2x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} 0 \\ y \end{bmatrix} \in \mathbf{R}^2 \right\} = \operatorname{Span}(\overline{e_2}).$$

If $v \in \operatorname{range}(A^T)$ and $w \in \operatorname{nul}(A)$, then

$$v^T w = \begin{bmatrix} \alpha & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \beta \end{bmatrix} = \alpha \cdot 0 + 0 \cdot \beta = 0. \checkmark$$

Theorem 232.

1. *If $S \subseteq \mathbf{R}^n$ is a subspace, then $\dim S + \dim S^\perp = n$.*
2. *If $\{\overline{v_1}, \dots, \overline{v_r}\}$ is a basis for S and $\{\overline{v_{r+1}}, \dots, \overline{v_n}\}$ is a basis for $S^\perp$, then $\{\overline{v_1}, \dots, \overline{v_r}, \overline{v_{r+1}}, \dots, \overline{v_n}\}$ is a basis for $\mathbf{R}^n$.*

Proof.

1. If $\dim S = r \leq n$, choose a basis $\{\overline{v_1}, \dots, \overline{v_r}\}$ of S. Let

$$A = \begin{bmatrix} \overline{v_1}^T \\ \vdots \\ \overline{v_r}^T \end{bmatrix} \in \mathbf{R}^{r \times n}.$$

We've built A so that $\operatorname{rank} A = r$ (the rows form a basis, hence are linearly independent, so there are r leading 1s in row echelon form), and $\operatorname{range}(A^T) = S$.
By **Theorem 230 (Fundamental subspace theorem)**,

$$\begin{aligned} S &= \operatorname{range}(A^T) \\ S^\perp &= \operatorname{range}(A^T)^\perp = N(A). \end{aligned}$$

By **Theorem 175 (Rank-Nullity)**,

$$\begin{array}{cccc} & \dim S & + & \dim S^\perp \\ = & r & + & \dim N(A) \\ = & n & & \end{array}$$

2. Thanks to **Theorem 156**, we only need to check that $\{\overline{v_1}, \dots, \overline{v_r}, \overline{v_{r+1}}, \dots, \overline{v_n}\}$ are linearly independent. Suppose

$$c_1\overline{v_1} + \cdots + c_r\overline{v_r} + c_{r+1}\overline{v_{r+1}} + \cdots + c_n\overline{v_n} = 0;$$

we want to show this forces all c_i to be 0. Label

$$\underbrace{c_1\overline{v_1} + \cdots + c_r\overline{v_r}}_{x} + \underbrace{c_{r+1}\overline{v_{r+1}} + \cdots + c_n\overline{v_n}}_{y} = 0$$

Thus $x + y = 0$, so $x = -y$. But since S and $S^\perp$ are subspaces,

$$\begin{array}{cc} x \in S & y \in S^\perp \\ -y \in S & -x \in S^\perp \\ y \in S & x \in S^\perp. \end{array}$$

So $x, y \in S \cap S^\perp$, and by **Theorem 228**, $x = y = 0$. This forces both $c_1 = \cdots = c_r = 0$ and $c_{r+1} = \cdots = c_n = 0$. Tada!

$\square$

Theorem 233. *If $S \subseteq \mathbf{R}^n$, then $(S^\perp)^\perp = S$.*

Proof. Recalling **Definition 226**,

$$\begin{aligned}
\left(S^\perp\right)^\perp &= \left\{w \in \mathbf{R}^n \mid v^T w = 0 \text{ for all } v \in S^\perp\right\} \\
&= \left\{w \in \mathbf{R}^n \mid v^T w = 0 \text{ for all } v \in \mathbf{R}^n \text{ such that } s^T v = 0 \text{ for all } s \in S\right\} \\
&= \left\{w \in \mathbf{R}^n \mid v^T w = 0 \text{ for all } v \in \mathbf{R}^n \text{ such that } v^T s = 0 \text{ for all } v \in \mathbf{R}^n\right\} \\
&= \{w \in \mathbf{R}^n \mid w = s,\ s \in S\} \\
&= \{s \in S\} \\
&= S.
\end{aligned}$$

□

Remark 234. Theorem 233 tells us that if T is the orthogonal complement to S, then S is the orthogonal complement to T:

$$\begin{aligned}
T &= S^\perp \\
T^\perp &= \left(S^\perp\right)^\perp \\
T^\perp &= S.
\end{aligned}$$

So we just say S and T are orthogonal complements, and which one gets the "$\perp$" doesn't matter.

Corollary 235. *Let $A \in \mathbf{R}^{m\times n}$ and let $\overline{b} \in \mathbf{R}^m$. Either:*
1. *there exists $\overline{x} \in \mathbf{R}^n$ such that $A\overline{x} = \overline{b}$, or*
2. *there exists $\overline{y} \in \mathbf{R}^m$ such that $A^T\overline{y} = \overline{0}$ and $\overline{y}^T\overline{b} \neq 0$.*

Proof. By **Theorem 230 (Fundamental subspace theorem)**, $\text{range}(A) = N(A^T)^\perp$. By **Theorem 228**, since $\text{range}(A) \perp N(A^T)$, we have $\text{range}(A) \cap N(A^T) = \{0\}$, so every nonzero $\overline{v} \in \mathbf{R}^n$ is either in $\text{range}(A)$ or $N(A^T)$ but not both. If $\overline{v} \in \text{range}(A)$, we're in case #1. If $\overline{v} \in N(A^T)$, we're in case #2. □

Example 236. Let

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

Find bases for $N(A)$, $\text{range}(A^T)$, $N(A^T)$, and $\text{range}(A)$.

We calculate:

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \overset{R_3 - R_1}{\sim} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \overset{R_1 - R_2}{\sim} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since

$$\begin{aligned}
\text{row}(A) &= \text{Span}\left(\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}\right), \\
\text{col}(A^T) = \text{range}(A^T) &= \text{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right).
\end{aligned}$$

Next, if

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in N(A),$$

then

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

so

$$\begin{aligned} x + z &= 0 \\ y + z &= 0, \end{aligned}$$

and hence an element of $N(A)$ looks like

$$\begin{bmatrix} -z \\ -z \\ z \end{bmatrix} \in \operatorname{Span}\left(\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right).$$

Now we calculate

$$A^T = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ 2 & 1 & 4 \end{bmatrix} \overset{R_2 - R_1}{\sim} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 2 & 1 & 4 \end{bmatrix} \overset{R_3 - 2R_1}{\sim} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since

$$\begin{aligned} \operatorname{row}(A^T) &= \operatorname{Span}\left(\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 \end{bmatrix} \right), \\ \operatorname{col}(A) = \operatorname{range}(A) &= \operatorname{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right). \end{aligned}$$

If

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in N(A^T),$$

then

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

so

$$\begin{aligned} x + z &= 0 \\ y + 2z &= 0, \end{aligned}$$

and thus

$$\begin{bmatrix} -z \\ -2z \\ z \end{bmatrix} \in \operatorname{Span}\left(\begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} \right) = N(A^T).$$

Homework 18. §5.2: 1, 2, 9, 14

Day 19 of 24 – §5.4 Inner product spaces ☐

Our goal is to generalize the scalar product $v^T w \in \mathbf{R}^n$ to any vector space.

Definition 237. Let V be a vector space. An **inner product** is a function $V \times V \to \mathbf{R}$, written $\langle v, w\rangle$ for $v, w \in V$, satisfying

1. $\langle v, v\rangle \geq 0$, and $\langle v, v\rangle = 0$ if and only if $v = 0$.
2. $\langle v, w\rangle = \langle w, v\rangle$.
3. $\langle \alpha v + \beta w, x\rangle = \alpha\langle v, x\rangle + \beta\langle w, x\rangle$.

Together we call $(V, \langle\ \cdot\ ,\ \cdot\ \rangle)$ an **inner product space**.

Example 238. The scalar product is an example. Let

$$\langle v, w\rangle = v^T w.$$

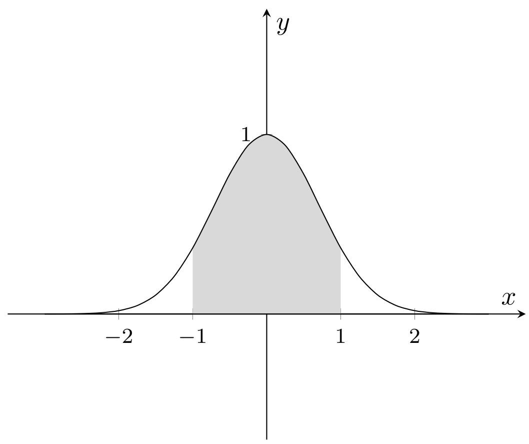
1. $\langle v, v\rangle = v^T v = \|v\|^2 \geq 0$ and it is 0 when $v = 0$. ✓
2. $\langle v, w\rangle = v^T w = w^T v = \langle w, v\rangle$. ✓
3. $\langle \alpha v + \beta w, x\rangle = (\alpha v + \beta w)^T x = (\alpha v)^T x + (\beta w)^T x = \alpha v^T x + \beta w^T x = \alpha\langle v, x\rangle + \beta\langle w, x\rangle$. ✓

Example 239. For $f, g \in C[a, b]$, define

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

1.

$$\langle f, f\rangle = \int_a^b f(x)^2 dx \geq 0.$$



Since an integral is area under the curve, the area is 0 when $f = 0$. ✓

2.

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx = \int_a^b g(x)f(x)dx = \langle g, f\rangle. ✓$$

3.

$$\begin{aligned}\langle \alpha f + \beta g, h\rangle &= \int_a^b (\alpha f(x) + \beta g(x))\, h(x)dx \\ &= \int_a^b (\alpha f(x)h(x) + \beta g(x)h(x))\, dx \\ &= \int_a^b \alpha f(x)h(x)dx + \int_a^b \beta g(x)h(x)dx \\ &= \alpha \int_a^b f(x)h(x)dx + \beta \int_a^b g(x)h(x)dx \\ &= \alpha\langle f, h\rangle + \beta\langle g, h\rangle. ✓\end{aligned}$$

Definition 240. Let $v \in (V, \langle \cdot\,,\,\cdot \rangle)$. The **length** of v is

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

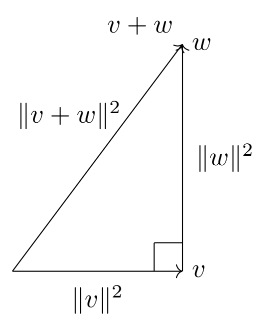
If $\langle v, w \rangle = 0$, then v and w are **orthogonal**.

Theorem 241 (Pythagorean Theorem)**.** *If $\langle v, w \rangle = 0$, then*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Proof.

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w \rangle \\
&= \langle v, v + w \rangle + \langle w, v + w \rangle && (\#3) \\
&= \langle v + w, v \rangle + \langle v + w, w \rangle && (\#2) \\
&= \langle v, v \rangle + \langle w, v \rangle + \langle v, w \rangle + \langle w, w \rangle && (\#3) \\
&= \langle v, v \rangle + \langle v, w \rangle + \langle v, w \rangle + \langle w, w \rangle && (\#2) \\
&= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \\
&= \langle v, v \rangle + 0 + \langle w, w \rangle \\
&= \|v\|^2 + \|w\|^2.
\end{aligned}$$



□

Proposition 242 (Cauchy-Schwarz inequality)**.**

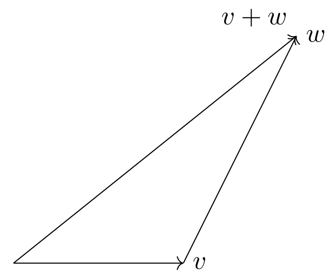
$$|\langle v, w \rangle| \leq \|v\| \|w\|,$$

and equality holds if and only if v and w are linearly dependent.

Definition 243. Let V be a vector space. A **norm** is a function $V \to \mathbf{R}$, written $\|v\|$ for $v \in V$, satisfying

1. $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$.
2. $\|\alpha v\| = |\alpha| \|v\|$.
3. $\|v + w\| \leq \|v\| + \|w\|$.

#3 is called the **triangle inequality**.



Example 244. We're using the same notation because the length $\|v\| = \sqrt{\langle v, v\rangle} = \sqrt{v^T v}$ is a norm.

1. $\|v\| = \sqrt{\langle v, v\rangle} \geq 0$ and is 0 if $v = 0$. ✓
2. $\|\alpha v\| = \sqrt{\langle \alpha v, \alpha v\rangle} = \sqrt{\alpha\langle v, \alpha v\rangle} = \sqrt{\alpha\langle \alpha v, v\rangle} = \sqrt{\alpha^2\langle v, v\rangle} = |\alpha|\sqrt{\langle v, v\rangle} = |\alpha|\|v\|$. ✓
3. See that

$$\begin{array}{lllllllll}\|v+w\|^2 & = & \langle v+w, v+w\rangle & = & \langle v, v\rangle + 2\langle v, w\rangle + \langle w, w\rangle & \leq & \|v\| + 2\|v\|\|w\| + \|w\| & = & (\|v\| + \|w\|)^2 \\ \|v+w\| & & & & & \leq & & & \|v\| + \|w\| \\ & & & & & \checkmark & & & \end{array}$$

Example 245. Define a norm on $\mathbf{R}^n$ by

$$\|v\|_p = \left(\sum_{i=1}^{n} |v_i|^p\right)^{1/p}.$$

This generalizes **Example 244** because when $p = 2$,

$$\|v\|_2 = \left(\sum_{i=1}^{n} |v_i|^2\right)^{1/2} = \sqrt{\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}}.$$

Example 246. Define

$$\|v\|_\infty = \max_i |v_i|.$$

Example 247. If

$$v = \begin{bmatrix} 4 \\ -5 \\ 3 \end{bmatrix},$$

then

$$\begin{aligned} \|v\|_1 &= |4| + |-5| + |3| = 12 \\ \|v\|_2 &= \sqrt{4^2 + (-5)^2 + 3^2} = \sqrt{50} \\ \|v\|_\infty &= \max\{|4|, |-5|, |3|\} = 5. \end{aligned}$$

Homework 19. §5.4: 1, 7a, 13, 14

Day 20 of 24 – §5.5 Orthonormal sets $\square$

Definition 248. Let $\{v_1, \dots, v_n\}$ be a set of nonzero vectors. If $\langle v_i, v_j \rangle = 0$ for all pairs $i \neq j$, then $\{v_1, \dots, v_n\}$ is an **orthogonal set**. When all the $v_1, \dots, v_n$ are length 1, we call $\{v_1, \dots, v_n\}$ an **orthonormal set**.

Example 249.

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\}$$

is an orothogonal set, because

$$\begin{aligned} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix} &= 1 \cdot 2 + 1 \cdot 1 + 1 \cdot (-3) = 0 \\ \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} &= 1 \cdot 4 + 1 \cdot (-5) + 1 \cdot 1 = 0 \\ \begin{bmatrix} 2 & 1 & -3 \end{bmatrix} \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} &= 2 \cdot 4 + 1 \cdot (-5) - 3 \cdot 1 = 0. \end{aligned}$$

But it's not orthonormal:

$$\begin{aligned} \left\| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\| &= \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3} \neq 1 \\ \left\| \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix} \right\| &= \sqrt{2^2 + 1^2 + (-3)^2} = \sqrt{14} \neq 1 \\ \left\| \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\| &= \sqrt{4^2 + (-5)^2 + 1^2} = \sqrt{42} \neq 1. \end{aligned}$$

We may make an orthonormal set by normalizing the vectors, since that only changes length and not direction:

$$\left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \frac{1}{\sqrt{42}} \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\}.$$

Theorem 250. *If $\{v_1, \dots, v_n\}$ is an orthogonal set, then $\{v_1, \dots, v_n\}$ are linearly independent.*

Proof. Let

$$c_1 v_1 + \cdots + c_n v_n = 0;$$

we show all c_i are 0. We take an inner product with one of the v_i:

$$0 = \langle 0, v_i \rangle = \langle c_1 v_1 + \cdots + c_n v_n, v_i \rangle = c_1 \langle v_1, v_i \rangle + \cdots + c_n \langle v_1, v_i \rangle \stackrel{*}{=} c_i \langle v_i, v_i \rangle = c_i \|v_i\|^2,$$

so since $\|v_i\|^2 \neq 0$, $c_i = 0$. But i was arbitrary, so all are 0. $\square$

Theorem 251. *Let $\{u_1, \dots, u_n\}$ be an orthonormal basis for V. If $v = c_1u_1 + \cdots + c_nu_n$, then $c_i = \langle v, u_i \rangle$.*

Proof.

$$\langle v, u_i \rangle = \langle c_1u_1 + \cdots + c_nu_n, u_i \rangle = c_1\langle u_1, u_i \rangle + \cdots + c_n\langle u_n, u_i \rangle = c_1 \cdot 0 + \cdots + c_i \cdot 1 + \cdots + c_n \cdot 1 = c_i.$$

□

Corollary 252. *Let $\{u_1, \dots, u_n\}$ be an orthonormal basis for V. If*

$$\begin{aligned} v &= \alpha_1u_1 + \cdots + \alpha_nu_n \\ w &= \beta_1u_1 + \cdots + \beta_nu_n, \end{aligned}$$

then $\langle v, w \rangle = \alpha_1\beta_1 + \cdots + \alpha_n\beta_n$.

Proof. By **Theorem 251**, $\langle u_i, w \rangle = \beta_i$, so

$$\begin{aligned} \langle v, w \rangle &= \langle \alpha_1u_1 + \cdots + \alpha_nu_n, w \rangle \\ &= \alpha_1\langle u_1, w \rangle + \cdots + \alpha_n\langle u_n, w \rangle \\ &= \alpha_1\beta_1 + \cdots + \alpha_n\beta_n. \end{aligned}$$

□

Corollary 253 (Parseval's formula)**.** *Let $\{u_1, \dots, u_n\}$ be an orthonormal basis for V. If $v = c_1u_1 + \cdots + c_nu_n$, then*

$$\|v\|^2 = c_1^2 + \cdots + c_n^2$$

Proof. By **Corollary 252**,

$$\|v\|^2 = \langle v, v \rangle = c_1^2 + \cdots + c_n^2.$$

□

Example 254.

$$\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} = \{u_1, u_2\} \qquad \text{and} \qquad \{\overline{e_1}, \overline{e_2}\}$$

are orthonormal bases of $\mathbf{R}^2$. Let

$$v = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

By **Theorem 251**, the coordinates of v with respect to $\{u_1, u_2\}$ are

$$\begin{aligned} c_1 &= \left\langle \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 4 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}} (4 \cdot 1 + 2 \cdot (-1)) = \frac{2}{\sqrt{2}} \\ c_2 &= \left\langle \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 4 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} (4 \cdot 1 + 2 \cdot 1) = \frac{6}{\sqrt{2}} \end{aligned}$$

By **Corollary 253 (Parseval's formula)**,

$$\begin{aligned} \|v\|^2 &= \left(\frac{2}{\sqrt{2}}\right)^2 + \left(\frac{6}{\sqrt{2}}\right)^2 = \frac{4}{2} + \frac{36}{2} = 2 + 18 = 20 && \text{(using } \{u_1, u_2\}\text{)} \\ \|v\|^2 &= 4^2 + 2^2 = 16 + 4 = 20 && \text{(using } \{\overline{e_1}, \overline{e_2}\}\text{).} \end{aligned}$$

Definition 255. Let $Q \in \mathbf{R}^{n\times n}$. Q is an **orthogonal matrix** if the columns of Q are an orthonormal set in $\mathbf{R}^n$.

Example 256.

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} \qquad \text{or} \qquad \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}.$$

Proposition 257. *$Q \in \mathbf{R}^{n\times n}$ is orthogonal if and only if $Q^TQ = I$. In other words, $Q^T = Q^{-1}$.*

Example 258.

$$\begin{aligned}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\cdot\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} &= \left(\frac{1}{\sqrt{2}}\right)^2\begin{bmatrix}1+1 & 1-1\\ 1-1 & 1+1\end{bmatrix}\\ &= \frac{1}{2}\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}\\ &= \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}.\end{aligned}$$

Theorem 259. *If Q is orthogonal, then $\langle v, w\rangle = \langle Qv, Qw\rangle$. In particular, $\|Qv\|^2 = \|v\|^2$, so multiplication by Q preserves lengths.*

Proof.

$$\langle Qv, Qw\rangle = (Qw)^TQv = w^TQ^TQv \overset{\textbf{Proposition 257}}{=} w^TIv = w^Tv = \langle v, w\rangle.$$

For the second claim,

$$\|Qv\|^2 = \langle Qv, Qv\rangle = \langle v, v\rangle = \|v\|^2,$$

so too $\|Qv\| = \|v\|$ and Q preserves lengths. $\square$

Homework 20. §5.5: 1, 2, 11, 14

Day 21 of 24 – §6.1 Eigenvalues and eigenvectors ☐

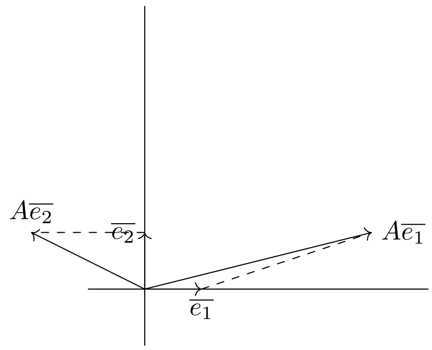
Example 260. We know a matrix $A \in \mathbf{R}^{m\times n}$ can be thought of as a linear transformation $\mathbf{R}^n \to \mathbf{R}^m$ which takes the standard basis $\{\overline{e_1}, \ldots, \overline{e_n}\}$ to the columns of A. For instance,

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

maps

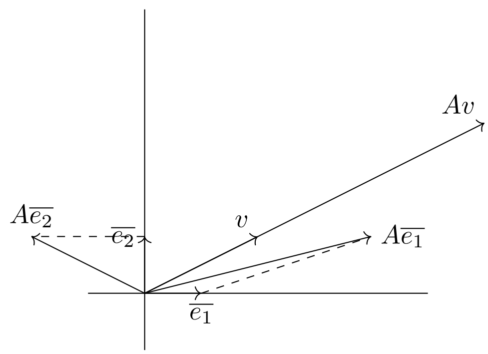
$$\begin{aligned}\begin{bmatrix}4 & -2\\ 1 & 1\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix} &= \begin{bmatrix}4\\ 1\end{bmatrix}\\ \begin{bmatrix}4 & -2\\ 1 & 1\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix} &= \begin{bmatrix}-2\\ 1\end{bmatrix},\end{aligned}$$

and the rest of $\mathbf{R}^2$ follows.



Notice what happens to one vector in particular, $v = \begin{bmatrix} 2 & 1 \end{bmatrix}^T$:

$$\begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \cdot 2 - 2 \cdot 1 \\ 1 \cdot 2 + 1 \cdot 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}.$$



For this specific vector v, multiplying by A ended up being very simple: it just scaled v by 3. Convenient!

Definition 261. Let $A \in \mathbf{R}^{n \times n}$. A scalar λ is an **eigenvalue** of A if there exists a nonzero $v \in \mathbf{R}^n$ such that $Av = \lambda v$. When we have such a v, we call it an **eigenvector** associated to λ.

Remark 262. In **Example 260**, $\lambda = 3$ was an eigenvalue of A and $v = \begin{bmatrix} 2 & 1 \end{bmatrix}^T$ was a corresponding eigenvector.

Remark 263. We can rewrite $Av = \lambda v$ as the matrix equation

$$\begin{aligned} Av &= \lambda I v \\ Av - \lambda I v &= 0 \\ (A - \lambda I)v &= 0. \end{aligned}$$

We have an eigenvalue λ if and only if the homogeneous system $(A - \lambda I)v = 0$ has a nontrivial solution $v \neq 0$. By definition, the solutions to $(A - \lambda I)v = 0$ is $N(A - \lambda I)$, so there's a nontrivial solution if $N(A - \lambda I) \neq \{0\}$.

Definition 264. We call $N(A - \lambda I)$ the **eigenspace** associated to λ.

Remark 265. By **Theorem 82**, $(A - \lambda I)v = 0$ has nontrivial solutions if and only if $(A - \lambda I)$ is singular; i.e., $\det(A - \lambda I) = 0$.

Definition 266. If $A \in \mathbf{R}^{n\times n}$, then the expression $\det(A - \lambda I)$ is a polynomial in the variable λ of degree n, called the **characteristic polynomial**. The equation $\det(A - \lambda I) = 0$ is called the **characteristic equation**.

Example 267. Let

$$A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}.$$

What are the eigenvalues and eigenvectors?

We solve the characteristic equation:

$$\begin{aligned} 0 = \det(A - \lambda I) &= \det \begin{bmatrix} 3-\lambda & 2 \\ 3 & -2-\lambda \end{bmatrix} \\ &= (3-\lambda)(-2-\lambda) - 2\cdot 3 \\ &= -6 - 3\lambda + 2\lambda + \lambda^2 - 6 \\ &= \lambda^2 - \lambda - 12 \\ &= (\lambda - 4)(\lambda + 3), \end{aligned}$$

so the eigenvalues are $\lambda_1 = 4$ and $\lambda_2 = -3$.

When $\lambda_1 = 4$, we have

$$[A - 4I \mid 0] = \left[\begin{array}{rr|r} -1 & 2 & 0 \\ 3 & -6 & 0 \end{array}\right] \sim \left[\begin{array}{rr|r} 1 & -2 & 0 \\ 3 & -6 & 0 \end{array}\right] \sim \left[\begin{array}{rr|r} 1 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

so $x - 2y = 0$, i.e., $x = 2y$, and therefore any multiple of

$$v_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

is an eigenvector associated to $\lambda_1 = 4$.

When $\lambda_2 = -3$, we have

$$[A + 3I \mid 0] = \left[\begin{array}{rr|r} 6 & 2 & 0 \\ 3 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{rr|r} 3 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

so $3x + y = 0$, i.e, $y = -3x$, and therefore any multiple of

$$v_2 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

is an eigenvector associated to $\lambda_2 = -3$.

Remark 268. Notice that in **Example 267**,

$$\det A = \det \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix} = 3(-2) - 2(3) = -6 - 6 = -12,$$

which was the constant term of the characteristic polynomial $\det(A - \lambda I)$. This in fact always holds for any $A \in \mathbf{R}^{n\times n}$.

Moreover, notice that the sum of the eigenvalues was $4 + (-3) = 1$ which was the same as the sum of the diagonal $3 + (-2) = 1$. This too always holds.

Definition 269. Given a matrix $A \in \mathbf{R}^{n\times n}$, the **trace**, $\operatorname{tr}(A)$, is the sum of the diagonal.

Theorem 270. *If $A \in \mathbf{R}^{2\times 2}$, then*

$$\det(A - \lambda I) = \lambda^2 - \operatorname{tr} A\lambda + \det A.$$

Proof. Let

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

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} \\ &= (a-\lambda)(d-\lambda) - bc \\ &= ad - a\lambda - d\lambda + \lambda^2 - bc \\ &= \lambda^2 - (a+d)\lambda + ad - bc \\ &= \lambda^2 - \operatorname{tr} A\lambda + \det A. \end{aligned}$$

□

Example 271. Let

$$A = \begin{bmatrix} 5 & -18 \\ 1 & -1 \end{bmatrix}.$$

The characteristic equation is

$$\lambda^2 - 4\lambda + 13 = 0.$$

Using the quadratic formula,

$$\lambda_1, \lambda_2 = \frac{4 \pm \sqrt{(-4)^2 - 4(1)(13)}}{2(1)} = \frac{4 \pm \sqrt{16-52}}{2} = \frac{4 \pm \sqrt{-36}}{2} = 2 \pm 3i.$$

Theorem 272. *Let $A, B \in \mathbf{R}^{n\times n}$. If A is similar to B, then $\det(A - \lambda I) = \det(B - \lambda I)$, and thus A and B have the same eigenvalues.*

Proof. Recall A is similar to B if there exists a nonsingular S such that $B = S^{-1}AS$. Therefore,

$$\det(B - \lambda I) = \det(S^{-1}AS - \lambda I).$$

Notice that

$$S^{-1}(A - \lambda I)S = S^{-1}AS - S^{-1}\lambda IS = S^{-1}AS - \lambda S^{-1}IS = S^{-1}AS - \lambda I,$$

so

$$\begin{aligned} \det(S^{-1}AS - \lambda I) &= \det(S^{-1}(A - \lambda I)S) \\ &= \det(S^{-1}) \det(A - \lambda I) \det(S) \\ &= \frac{1}{\det(S)} \det(A - \lambda I) \det(S) \\ &= \det(A - \lambda I). \end{aligned}$$

□

Example 273. By construction,

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$$

are similar. B is

$$\begin{aligned} B &= \frac{1}{10-9} \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 13 & 8 \\ 9 & 6 \end{bmatrix} \\ &= \begin{bmatrix} -1 & -2 \\ 6 & 6 \end{bmatrix}. \end{aligned}$$

We see via **Theorem 270** that

$$\begin{aligned} \det(A - \lambda I) &= \lambda^2 - (2+3)\lambda + (2 \cdot 3 - 1 \cdot 0) \\ \det(B - \lambda I) &= \lambda^2 - (-1+6)\lambda + (-1 \cdot 6 + 2 \cdot 6) \end{aligned}$$

In both cases, it is $\lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3)$, with eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 3$.

Homework 21. §6.1: 1abcdefg, 2, 3, 4, 10, 17

Day 22 of 24 – §6.3 Diagonalization □

Remark 274. Exercise #2 in §6.1 is really cool; it tells us that if A is triangular, the eigenvalues are the elements of the diagonal. Today we'll study even nicer matrices: diagonal ones. We'd like to know: if you hand me a nondiagonal A, can I rewrite $A = SDS^{-1}$, where D is diagonal? That is, when is a matrix similar to a diagonal one?

Definition 275. $A \in \mathbf{R}^{n \times n}$ is <u>**diagonalizable**</u> if there exists a nonsingular S and a diagonal matrix D such that

$$S^{-1}AS = D.$$

So, when are matrices diagonalizable?

Theorem 276. *If $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of A with eigenvectors $v_1, \ldots, v_k$, then $\{v_1, \ldots, v_k\}$ are linearly independent.*

Proof. Let

$$\dim \operatorname{Span}(v_1, \ldots, v_k) = r.$$

We want to see that $r = k$ but suppose for the sake of contradiction that $r < k$, and that only the first r vectors $\{v_1, \ldots, v_r\}$ are linearly independent. That means $\{v_1, \ldots, v_r, v_{r+1}\}$ is linearly dependent and

$$c_1 v_1 + \cdots + c_r v_r + c_{r+1} v_{r+1} = 0 \tag{$*$}$$

where not all c_i are 0; in fact, $c_{r+1} \neq 0$. Thus $c_{r+1} v_{r+1} \neq 0$, so that means that there has to be some nonzero c_i among the others as well.

Next, multiply $(*)$ by A:

$$\begin{aligned} A\left(c_1v_1+\cdots+c_rv_r+c_{r+1}v_{r+1}\right) &= A0 \\ c_1Av_1+\cdots+c_rAv_r+c_{r+1}Av_{r+1} &= 0 \\ c_1\lambda_1v_1+\cdots+c_r\lambda_rv_r+c_{r+1}\lambda_{r+1}v_r &= 0. \end{aligned} \tag{$**$}$$

But now we calculate $(**)-\lambda_{r+1}(*)$:

$$\begin{array}{ccccccccc} (c_1\lambda_1-c_1\lambda_{r+1})v_1 & + & \cdots & + & (c_r\lambda_r-c_r\lambda_{r+1})v_r & + & (c_{r+1}\lambda_{r+1}-c_{r+1}\lambda_{r+1})v_{r+1} & = & 0 \\ (c_1\lambda_1-c_1\lambda_{r+1})v_1 & + & \cdots & + & (c_r\lambda_r-c_r\lambda_{r+1})v_r & & & = & 0 \\ c_1\underbrace{(\lambda_1-\lambda_{r+1})}_{\neq 0}v_1 & + & \cdots & + & c_r\underbrace{(\lambda_r-\lambda_{r+1})}_{\neq 0}v_r & & & = & 0 \end{array}$$

But this is a list of coefficients, not all 0, showing that $\{v_1,\ldots,v_r\}$ is linearly dependent. That's a contradiction, so it was wrong to suppose $r<k$. Thus $r=k$ as desired. □

Theorem 277. *$A\in\mathbf{R}^{n\times n}$ is diagonalizable if and only if A has n linearly independent eigenvectors $\{\overline{v_1},\ldots,\overline{v_n}\}$.*

Proof. Let

$$S=\begin{bmatrix}\overline{v_1} & \cdots & \overline{v_n}\end{bmatrix}.$$

We can calculate:

$$\begin{aligned} AS &= \begin{bmatrix}A\overline{v_1} & \cdots & A\overline{v_n}\end{bmatrix} \\ &= \begin{bmatrix}\lambda_1\overline{v_1} & \cdots & \lambda_n\overline{v_n}\end{bmatrix} \\ &= \begin{bmatrix}\overline{v_1} & \cdots & \overline{v_n}\end{bmatrix}\begin{bmatrix}\lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{bmatrix} \\ &= SD, \end{aligned}$$

where D is diagonal. Furthermore, S has n linearly independent columns, so S is nonsingular, and therefore

$$\begin{aligned} AS &= SD \\ S^{-1}AS &= D. \end{aligned}$$

Thus, if $\{\overline{v_1},\ldots,\overline{v_n}\}$ is linearly independent, then A is diagonalizable.

On the other hand, suppose A is diagonalizable. Thus

$$\begin{aligned} S^{-1}AS &= D \\ AS &= SD. \end{aligned}$$

Let $S=\begin{bmatrix}\overline{s_1} & \cdots & \overline{s_n}\end{bmatrix}$. We have

$$A\overline{s_i}=d_{ii}\overline{s_i}$$

where d_{ii} is the ith diagonal entry. But d_{ii} is just a scalar, so that's an eigenvector equation

$$A\overline{s_i}=\lambda_i\overline{s_i},$$

and so the columns of S are eigenvectors. But S was nonsingular, so those eigenvectors are linearly independent. □

Remark 278. Notice something else cool about diagonalizing: If $A = SDS^{-1}$, then

$$\begin{aligned}
A^2 &= \left(SDS^{-1}\right)^2 = SDS^{-1}SDS^{-1} = SD^2S^{-1} \\
A^3 &= A^2A = \left(SD^2S^{-1}\right)\left(SDS^{-1}\right) = SD^3S^{-1} \\
A^4 &= A^3A = \left(SD^3S^{-1}\right)\left(SDS^{-1}\right) = SD^4S^{-1} \\
&\vdots \\
A^k &= SD^kS^{-1}.
\end{aligned}$$

Calculating the powers of a diagonal matrix is really easy!

$$\begin{aligned}
D^2 &= \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_n^2 \end{bmatrix} \\
D^3 = D^2D &= \begin{bmatrix} \lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_n^2 \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1^3 & & & \\ & \lambda_2^3 & & \\ & & \ddots & \\ & & & \lambda_n^3 \end{bmatrix} \\
&\vdots \\
D^k &= \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix}.
\end{aligned}$$

Example 279. Let

$$A = \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}.$$

The characteristic equation is

$$\begin{aligned}
\lambda^2 - \operatorname{tr} A\lambda + \det A &= 0 \\
\lambda^2 - (2-5)\lambda + (2\cdot(-5) - (-3)\cdot 2) &= 0 \\
\lambda^2 + 3\lambda - 4 &= 0 \\
(\lambda+4)(\lambda-1) &= 0,
\end{aligned}$$

so $\lambda_1 = -4$, $\lambda_2 = 1$. The eigenvectors are

$$[A + 4I \mid 0] = \left[\begin{array}{cc|c} 6 & -3 & 0 \\ 2 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 2 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

so $2x - y = 0$, i.e., $y = 2x$, and $v_1 = \begin{bmatrix} 1 & 2 \end{bmatrix}^T$, and

$$[A - 1I \mid 0] = \left[\begin{array}{cc|c} 1 & -3 & 0 \\ 2 & -6 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -3 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

so $x - 3y = 0$, i.e., $x = 3y$, and $v_2 = \begin{bmatrix} 3 & 1 \end{bmatrix}^T$.

Let

$$S = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix}.$$

We calculate

$$S^{-1} = \frac{1}{1-6}\begin{bmatrix} 1 & -3 \\ -2 & 1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix},$$

so

$$\begin{aligned} S^{-1}AS &= \frac{1}{5}\begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \\ &= \frac{1}{5}\begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} -4 & 3 \\ -8 & 1 \end{bmatrix} \\ &= \frac{1}{5}\begin{bmatrix} -20 & 0 \\ 0 & 5 \end{bmatrix} \\ &= \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix} \\ &= D. \checkmark \end{aligned}$$

Now calculating something like

$$\begin{aligned} A^6 &= SD^6S^{-1} \\ &= \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} (-4)^6 & 0 \\ 0 & 1^6 \end{bmatrix}\frac{1}{5}\begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix} \\ &= \frac{1}{5}\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 4096 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix} \\ &= \frac{1}{5}\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} -4096 & 12288 \\ 2 & -1 \end{bmatrix} \\ &= \frac{1}{5}\begin{bmatrix} -4090 & 12285 \\ -8190 & 24575 \end{bmatrix} \\ &= \begin{bmatrix} -818 & 2457 \\ -1638 & 4915 \end{bmatrix} \end{aligned}$$

sure beats calculating $AAAAAA$. ☺

Definition 280. Recall the exponential function

$$e^a = 1 + a + \frac{1}{2!}a^2 + \frac{1}{3!}a^3 + \cdots .$$

Define the **matrix exponential**

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots .$$

Remark 281. If $A = SDS^{-1}$ is diagonalizable, then **Remark 278** tells us

$$\begin{aligned} e^A &= I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots \\ &= I + SDS^{-1} + \frac{1}{2!}\left(SDS^{-1}\right)^2 + \frac{1}{3!}\left(SDS^{-1}\right)^3 + \cdots \\ &= I + SDS^{-1} + \frac{1}{2!}SD^2S^{-1} + \frac{1}{3!}SD^3S^{-1} + \cdots \\ &= S\left(I + D + \frac{1}{2!}D^2 + \frac{1}{3!}D^3 + \cdots\right)S^{-1} \\ &= Se^DS^{-1}. \end{aligned}$$

We can calculate e^D:

$$\begin{aligned}
e^D &= I + D + \frac{1}{2!}D^2 + \frac{1}{3!}D^3 + \cdots \\
&= \begin{bmatrix} 1+\lambda_1+\frac{1}{2!}\lambda_1^2+\frac{1}{3!}\lambda_1^3+\cdots & & & \\ & 1+\lambda_2+\frac{1}{2!}\lambda_2^2+\frac{1}{3!}\lambda_2^3+\cdots & & \\ & & \ddots & \\ & & & 1+\lambda_n+\frac{1}{2!}\lambda_n^2+\frac{1}{3!}\lambda_n^3+\cdots \end{bmatrix} \\
&= \begin{bmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{bmatrix}.
\end{aligned}$$

Example 282. Working off of **Example 279**, we can calculate

$$\begin{aligned}
e^A &= Se^DS^{-1} \\
&= \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} e^{-4} & 0 \\ 0 & e \end{bmatrix} \frac{1}{5} \begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix} \\
&= \frac{1}{5} \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -e^{-4} & 3e^{-4} \\ 2e & -e \end{bmatrix} \\
&= \frac{1}{5} \begin{bmatrix} 6e - e^{-4} & 3e^{-4} - 3e \\ 2e - 2e^{-4} & 6e^{-4} - e \end{bmatrix}.
\end{aligned}$$

Homework 22. §6.3: 1ab, 2ab, 3ab, 18, 30ab, 31ab

Appendix: Characterizations of nonsingular matrices □

Theorem. *The following are equivalent:*

1. *$A \in \mathbf{R}^{n\times n}$ is nonsingular.* (**Definition 67**)
2. *A^{-1} is nonsingular.* (**Definition 67**)
3. *A is row equivalent to I.* (**Theorem 82**)
4. *A (and A^{-1}) is a finite product of elementary matrices.* (**Definition 81**)
5. *$Ax = 0$ if and only if $x = 0$.* (**Theorem 82**)
6. *$N(A) = \{0\}$.* (**Definition 126**)
7. *$Ax = b$ has a unique solution x.* (**Corollary 83**)
8. *The columns of A are linearly independent.* (**Theorem 146**)
9. *$Ax = b$ has at most one solution x.* (**Corollary 173**)
10. *The columns of A form a basis of $\mathbf{R}^n$.* (**Corollary 174**)
11. $\operatorname{rank} A = n$. (**Theorem 175 (Rank-Nullity)**)
12. *0 is not an eigenvalue of A.* (**Remark 265**)
13. $\det A \neq 0$. (**Theorem 106**)
14. $\det A^T \neq 0$. (**Proposition 97**)
15. *A^T is nonsingular.* (**Theorem 106**)
16. *$(A^T)^{-1}$ is nonsingular.* (**Definition 67**)
17. *A is column equivalent to I.* (**Theorem 82**)
18. *A^T (and $(A^T)^{-1}$) is a finite product of elementary matrices.* (**Definition 81**)
19. *$A^Tx = 0$ if and only if $x = 0$.* (**Theorem 82**)
20. *$N(A^T) = \{0\}$.* (**Definition 126**)
21. *$A^Tx = b$ has a unique solution x.* (**Corollary 83**)
22. *The rows of A are linearly independent.* (**Theorem 146**)
23. *$A^Tx = b$ has at most one solution x.* (**Corollary 173**)
24. *The rows of A form a basis of $\mathbf{R}^n$.* (**Corollary 174**)
25. $\dim \operatorname{col} A = n$. (**Theorem 178**)
26. *0 is not an eigenvalue of A^T.* (**Remark 265**)

Appendix: Definitions

A

Addition: .... **Definition 36**
Augment: .... **Definition 15**

B

Basis (pl. bases): .... **Definition 150**

C

Characteristic equation: .... **Definition 266**
Characteristic polynomial: .... **Definition 266**
Closed (subspace): .... **Definition 121**
Closed (vector space): .... **Definition 113**
Coefficient matrix: .... **Definition 15**
Cofactor: .... **Definition 93**
Cofactor expansion: .... **Definition 93**
Column space: .... **Definition 167**
Column vector: .... **Definition 33**
Coordinate vector: .... **Definition 159**
Consistent: .... **Definition 5**

D

Determinant: .... **Definition 93**
Diagonalizable: .... **Definition 275**
Dimension: .... **Definition 155**
Distance between two vectors: .... **Definition 212**

E

Eigenspace: .... **Definition 261**
Eigenvalue: .... **Definition 261**
Eigenvector: .... **Definition 261**
Elementary matrix: .... **Definition 78**
Equivalent: .... **Definition 7**
Exponential: .... **Definition 280**
Exponentiation: .... **Definition 60**

F

Free variable: .... **Definition 20**

G

Gaussian elimination: .... **Definition 21**
Gauss-Jordan reduction: .... **Definition 28**

H

Homogeneous: .... **Definition 30**

I

Identity matrix: .... **Definition 63**
Image (of a linear transformation): .... **Definition 192**
Image (of a subspace): .... **Definition 192**
Inconsistent: .... **Definition 5**
Inner product: .... **Definition 237**
Inner product space: .... **Definition 237**
Inverse: .... **Definition 67**
Invertible: .... **Definition 67**

K

Kernel (of a linear transformation): .... **Definition 191**
Kernel (of a matrix): .... **Definition 126**

L

Lead variable: .... **Definition 20**
Length (inner product): .... **Definition 240**
Length (scalar product): .... **Definition 210**
Linear combination (in an arbitrary vector space): .... **Definition 129**
Linear combination (in $\mathbf{R}^n$): .... **Definition 45**
Linear equation in n variables: .... **Definition 3**
Linear operator: .... **Definition 181**
Linear system of m equations: .... **Definition 3**
Linear transformation: .... **Definition 181**
Linearly dependent: .... **Definition 140**
Linearly independent: .... **Definition 140**

M

Matrix (pl. matrices): .... **Definition 15**
Matrix addition: .... **Definition 36**
Matrix exponential: .... **Definition 280**
Matrix exponentiation: .... **Definition 60**
Matrix multiplication: .... **Definition 48**
Matrix subtraction: .... **Definition 39**
Minor: .... **Definition 93**
Multiplication: .... **Definition 48**
$m \times n$ system: .... **Definition 3**
$m \times n$ matrix: .... **Definition 15**

N

Nonsingular: .... **Definition 67**
Norm (inner product): .... **Definition 243**
Norm (scalar product): .... **Definition 210**
Not invertible: .... **Definition 67**
Null space: .... **Definition 126**
$n \times n$ identity matrix: .... **Definition 63**

O

Orthogonal complement: .... **Definition 226**
Orthogonal set: .... **Definition 248**
Orthogonal (matrix): .... **Definition 255**
Orthogonal (subspaces): .... **Definition 224**
Orthogonal (vectors, inner product): .... **Definition 240**
Orthogonal (vectors, scalar product): .... **Definition 217**
Orthonormal set: .... **Definition 248**
Overdetermined: .... **Definition 24**

P

Product of an $m \times n$ matrix and an $n \times 1$ matrix: .... **Definition 40**
Product of an $m \times n$ matrix and an $n \times r$ matrix: .... **Definition 48**

R

Range: .... **Definition 192**
Rank: .... **Definition 170**
Reduced row echelon form: .... **Definition 28**
Row echelon form: .... **Definition 21**
Row equivalent: .... **Definition 81**
Row space: .... **Definition 167**
Row vector: .... **Definition 33**

S

Scalar multiplication: .... **Definition 35**
Scalar product: .... **Definition 207**
Scalar projection of x onto y: .... **Definition 221**
Similar: .... **Definition 205**
Singular: .... **Definition 67**
Solution to an $m \times n$ system: .... **Definition 3**
Span: .... **Definition 129**
Spanning set: .... **Definition 133**
Strict triangular form: .... **Definition 13**
Subspace: .... **Definition 121**
Subtraction: .... **Definition 39**
Symmetric: .... **Definition 56**

T

Trace: .... **Definition 269**
Transpose: .... **Definition 54**
Triangle inequality: .... **Definition 243**

U

Underdetermined: .... **Definition 24**
Upper triangular form: .... **Definition 13**

V

Vector: .... **Definition 33**
Vector projection of x onto y: .... **Definition 221**
Vector space: .... **Definition 113**

Appendix: Homeworks

§1.1	1b, 2b, 3d, 4d, 6a, 10
§1.2	2, 3, 4, 5e, 6c, 7
§1.3	2ab, 4b, 9, 11, 12
§1.4	3, 4, 10, 16
§1.5	10ef, 11, 12ab
§2.1	2, 3, 5, 6
§2.2	1, 4, 7, 13
§3.1	8, 9, 10, 11
§3.2	1, 4ab, 5, 11
§3.3	1, 5, 6
§3.4	4, 7, 17
§3.5	1, 2, 3, 7, 8
§3.6	1, 2, 3, 4, 5
§4.1	1, 5, 13, 17
§4.2	1, 2, 6, 13
§4.3	1, 2, 4
§5.1	1, 2, 4, 5
§5.2	1, 2, 9, 14
§5.4	1, 7a, 13, 14
§5.5	1, 2, 11, 14
§6.1	1abcdefg, 2, 3, 4, 10, 17
§6.3	1ab, 2ab, 3ab, 18, 30ab, 31ab

Appendix: Notation

A, B, etc	Matrices
$A \sim B$	"A is equivalent to B"
$[A \mid B]$	The block matrix of A augmented by B
a_{ij}	The entry in row i and column j of matrix A
A_{ij}	The row i, column j cofactor of a matrix A
A^T	The transpose of A
A^{-1}	The inverse of A
$\langle b_1, \ldots, b_n \rangle$	The vector space generated by the basis $\{b_1, \ldots, b_n\}$
$C[a,b]$	The vector space of continuous functions $f : [a,b] \to \mathbf{R}$
$C^n[a,b]$	The vector space of continuous, nth order differentiable functions $f : [a,b] \to \mathbf{R}$
$\operatorname{col} A$	The column space of A
D	A diagonal matrix
$\det A$	The determinant of A
$\det M_{ij}$	The row i, column j minor of a matrix
$\dim V$	The dimension of V
E	An elementary matrix
$\{\overline{e_1}, \ldots, \overline{e_n}\}$	The standard basis of $\mathbf{R}^n$
e^A	The matrix exponential
f, g, etc	A function or polynomial
I	The identity matrix
$\operatorname{im} L$	The image of L
$\ker L$	The kernel of L
L	A linear transformation
$L(S)$	The image of a subspace S
m, n, etc	Integers
$N(A)$, $\operatorname{nul} A$, $\ker A$	The kernel / null space of A
p	The vector projection of x onto y
P_n	The vector space of polynomials of degree less than n
Q	An orthogonal matrix
$\operatorname{range} A$, $R(A)$	The range of A
$\operatorname{rank} A$, $\operatorname{rnk} A$, $\operatorname{rk} A$	The rank of A
$\operatorname{row} A$	The row space of A
R_i	The ith row of a matrix
$\mathbf{R}^n$	The collection of n-vectors / n-tuples
$\mathbf{R}^{m \times n}$	The collection of $m \times n$ matrices
S, S^{-1}	A nonsingular matrix
S, T, etc	Subspaces

Span$(v_1, \ldots, v_n)$ .... The span of $\{v_1, \ldots, v_n\}$
tr A .... The trace of A
u .... A unit / normalized vector (length 1)
U .... A row reduced / triangular matrix
V, W, $(V, \cdot, +)$, etc .... Vector spaces
$(V, \langle \cdot, \cdot \rangle)$ .... An inner product space
$V^\perp$ .... The orthogonal complement of V
$V \perp W$ .... "The subspace V is orthogonal to the subspace W"
$\overline{v}$, $\overline{w}$, $\overline{x}$, $\overline{b}$, v, w, etc .... A (row / column) vector
$\|v\|$ .... The length / norm of v
$\langle v, w \rangle$ .... The inner product of v and w
x_i, x, y, z, etc .... An indeterminate / variable
α, β, γ, c_i, etc .... Coefficients / scalars
θ, ϑ .... An angle
λ .... An eigenvalue
$\emptyset$, $\varnothing$ .... The empty set

Appendix: Reference

[LDP06] S. J. Leon, L. G. De Pillis, *Linear algebra with applications, 9th edition*, Pearson Prentice Hall Upper Saddle River, NJ, (2006).