# Linear Algebra—Math 325.01

David Meredith
Department of Mathematics
San Francisco State University
San Francisco, CA 94132

Fall 2001—MWF 10-11—HUM 408

# Contents

# Chapter 1 Introduction

When it comes to making practical, difficult calculations, linear algebra is the most important branch of applied mathematics. Calculus gives us the theory; linear algebra provides the answers. Whether you are computing the cross-section of a proton or a black hole; a dna molecule or a bridge cable, you will use linear algebra. Linear algebra is the mathematics of computer graphics. This course will introduce you to the basic ideas of the subject: vectors, matrices and linear transformations. You will learn to perform basic calculations by hand and with software, and you will learn the geometric meaning of these operations.

## 1.1 Instructor

David Meredith
TH 933
Office Hours: MWF 11-12
E-mail: meredith@sfsu.edu
URL: http://online.sfsu.edu/˜meredith

## 1.2 Web Site and E-mail

Assignments, class notes, handouts, syllabus changes, grade sheet, etc. can be found at the class web site, which you can access by going to my web site and clicking on **Linear Algebra.** I will not distribute any more paper handouts to you. All information will be distributed via the class web site.

When you cannot get to my office hour or wait for the next class, you may submit questions to me by e-mail. If you cannot make it to class, you can submit projects by e-mail.

## 1.3 Text

Liptchutz, *Linear Algebra (2nd Edition),* Schaum's Outline Series, McGraw Hill, New York, 1991.

## 1.4 Software

### 1.4.1 Mathematical Software

In this course you will learn to perform matrix calculations with software. There are several programs you can use. I will instruct the class to use *Mathematica,* but I will try to help you individually if you choose to use another program.

#### 1.4.1.1 *Mathematica*

*Mathematica* computes numerically and symbolically. It will invert a $50 \times 50$ matrix, and it will evaluate $\int e^{4x} \sin 5x\, dx$. With *Mathematica,* you can work interactively or you can write subroutines. The Math Dept. at SFSU has decided to use this program in many of its courses. *Mathematica* is freely available M-F, 10-3, in the Math/Stat Computing Lab, TH404. Buying your own copy is not necessary, but it might be a good investment. If you learn to use *Mathematica* effectively, you will find it very valuable in all your science, engineering and math courses. The bookstore will sell you a copy for $140.

Unfortunately, the user-interface for matrix calculations in *Mathematica* is one of the program's worst features. It is usable but weird.

#### 1.4.1.2 *Maple*

*Maple* is a symbolic and numerical powerhouse similar to *Mathematica* but not as widely used outside of academia.

#### 1.4.1.3 *Matlab*

*Matlab* is the classic linear algebra program with the best numerical algorithms. A student version is available at the bookstore. If you are going into engineering, it might pay to learn this program. *Matlab* now includes *Maple,* so some symbolic calculations can be performed with *Matlab.* If your focus is symbolic work, *Mathematica* and *Maple* are better choices.

#### 1.4.1.4 *X(PLORE)*

Truth in advertising: I wrote *X(PLORE).* It is freely available from my web site and in the Math/Stat Computing Lab. You can do all the linear algebra calculations for this course in *X(PLORE).* The advantage is the cost; the disadvantage is the lack of symbolic computation.

#### 1.4.1.5 *TestGiver*

*TestGiver* is the program that you will use for your weekly homework. It incorporates an adequate matrix calculator that you can use for class computations.

#### 1.4.1.6 Graphing Calculators

These toys can be used for the simple calculations and graphs. The most recent models will perform symbolic integration and matrix calculations. On the other hand, calculators are much harder to use than computer software. The keystroke combinations are harder to remember, the screen is much harder to read, and the capabilities are much less. I have never seen a professional scientist or engineer use a graphing calculator. Professionals use computers; you should start learning to use a computer too. It is possible to do all the computer homework in this course on the most powerful graphing calculators (TI 89, HP 49), but it won't be fun.

### 1.4.2 Word Processing Software

You can write mathematical reports with *Mathematica,* but not easily. Moreover, the results are ugly. Much better tools are any standard word processor with equation-writing capability (*Word* or *WordPerfect* will do nicely, but you may have to install the equation module for *Word*) or a special scientific work processor like *Scientific WorkPlace* ($600) or *Scientific Notebook*($75). I highly recommend *Scientific Workplace* and *Scientific Notebook.* They include Maple, so you can do symbolic and matrix computations inside the word processor. The output looks good, much better than the output from either *Mathematica* or *Word.* Sometimes the bookstore stocks *Scientific Notebook.* More information about these products can be found at. http://www.mackichan.com.

This document was produced with *Scientific Workplace.*

## 1.5 Class procedures

### 1.5.1 Schedule

| | **Topics** | **Chapter** | **Due Date** |
|---|---|---|---|
| Homework | Vector Geometry in $\mathbb{R}^2$ | 1 | 9/10 |
| Homework | Vector Geometry in $\mathbb{R}^3$ | 1 | 9/17 |
| Homework | Vector Geometry in $\mathbb{R}^n$ | 1 | 9/24 |
| **Project 1** | **Centers of Tetrahedra** | | **9/26** |
| Homework | Rectangular Matrices | 2 | 10/1 |
| Homework | Solving Linear Equations | 3.1-3.6 | 10/8 |
| **Midterm 1** | | | **10/15** |
| Homework | Linear Equations and Matrices | 3.7-3.11 | 10/19 |
| Homework | Square Matrices and Inverses | 3.12 | 10/24 |
| Homework | Determinants | 8 (parts) | 10/29 |
| **Project 2** | **Projections, Reflections and Rotations in $\mathbb{R}^3$** | | **10/31** |
| Homework | Eigenvalues and Eigenvectors | 9 | 11/5 |
| Homework | Diagonalization and Symmetric Matrices | 10 | 11/12 |
| **Midterm 2** | | | **11/14** |
| **Project 3** | | | **11/21** |
| Homework | Subspaces | 4 | 11/26 |
| Homework | Subspaces | 4 | 12/3 |
| Homework | Subspaces | 4 | 12/10 |
| **Project 4** | | | **12/12** |
| **Final** | **Friday,12/21, 8:00-10:30** | | |

### 1.5.2 Homework

You will have two kinds of homework in this course. There will be **problems** and **projects**. Homework is an important part of this course. Fifty percent of your grade comes from the homework.

#### 1.5.2.1 Problems

Homework problems will be assigned weekly. They are similar to the exercises in the text, but the format is new. I've developed an electronic homework system called *TestGiver* that you will use.

1. You begin by downloading the program *TestGiver* from the class web site and installing it on your computer at home. Or you can use *TestGiver* in the Math/Stat Lab (TH 404).

   (a) *TestGiver* includes a massive help file, and the folder that contains your installation also contains a pdf file with a manual for the program. You should read the section titled "Overview" in the help file or the manual.

2. Then every week, **at least seven days before the assignment is due,** you download the weekly assignment file. This will be a file with the extension TGV.

   (a) Load your test into *TestGiver* using the menu item `File/Open New Test`. Random items will be given values, and the questions will be displayed on the `Test` page.

      i. If the test is not displayed, click the tab marked "Test" at the top of the window.

      ii. You should see the questions and spaces for your answers.

   iii. You may want to print your test and work away from the computer. Wise students will keep printed copies of their completed tests. At any time during the test-taking, a printed copy of the test with all answers entered so far can be created by selecting the menu item `File/Print Test`.

(b) When you enter an answer it will be checked instantly, and you will be told if it is right or wrong.

   i. Most questions allow you to keep trying answers until you get the right answer. But questions with a yellow background give you only one try. Think carefully before answering a one-try question.
   Usually one-try questions are like true/false questions. They have a limited number of possible answers, and the one-try rule is there to keep test takers from just trying all possible answers without actually doing the problem.
   If you enter the wrong answer to a one-try question, you can submit a written correction that explains the problem thoroughly in complete sentences. Begin with a statement of the problem.

(c) At any time you can save a partially completed test and reload it later into *TestGiver*. All randomly produced parameters and all answers, right and wrong, are saved. Nothing is changed when the test is reloaded for further work. Saving is done with the menu item `File/Save Test` and reloading is done with the menu item `File/Open Old Test`. The recommended file extension for saved tests is TST.

   i. WARNING: always reload a TST file that you have saved. If you reload a TGV file, all answers will be lost and all random parameters will be recalculated. Loading a TGV file with `File/Open New Test` always restarts a test. Loading a TST file with `File/Open Old Test` allows you to continue a test you have already started.

(d) When you finish the assignment, send a report of your work to me over the internet.

   i. Get connected to the internet. If you are in the lab, you are already connected.
   ii. Send your report by using the menu item `File/Send Report.`
   iii. Reports must be send on or before the due date. Late papers will be rejected by the recording software.
      A. Falling behind is one of the main reasons students fail calculus. This rule is designed to help you keep up. It is better to submit a partially completed assignment and go on to the next one than to fall behind while trying to complete an old assignment.
      B. If you have a good reason for turning in an assignment late, please see me.

(e) If you have questions about your homework, you can send me an email by clicking `Help/Send e-mail to Dr. Meredith.`

#### 1.5.2.2 Projects

1. Projects are formal reports. I will give you a problem, usually with a lot of parts, and you will write a report based on the problem.

   (a) Reports must be written in correct mathematical English. Your textbook is a model of correct mathematical writing.

      i. The test of good writing is this: a student in another linear algebra class at another college should be able to read your paper and understand what problem you were doing and how you answered it.

(b) The most important criteria for a good report is that it is clear and without errors. It is better to write a grammatical, well-organized report that correctly answers part of the question than to write a report that purports to answer the entire question but is either fragmentary or contains mistakes.

   i. Think of getting one point for each correct, well-written part and -1 point for each part that is grammatically or mathematically wrong. You can get a lot of points just by not turning in wrong answers.
   ii. There is no partial credit for mistakes.
   iii. There is lots of partial credit for correctly stating and answering part of a problem. If you cannot do a problem as stated, then:
      A. do a simpler version. If the problem asks about $n \times n$ matrices, you might solve it for $3 \times 3$ matrices.
      B. do an example. If the problem asks you to prove something about symmetric matrices, you can show that the property holds for one symmetric matrix.
      C. discuss the problem. Prove what you can and explain why you are stuck.

      The important thing is to write a well-organized report that is internally correct, even if it does not completely answer the problem posed.

(c) You do not have to take up the problem parts in the order I present them. You can organize your report any way you like, so long as it is clear.

   i. You can use a part that you cannot prove to prove another part, so long as you are clear about what you are doing.

2. Rules for Writing Projects

   (a) Things to do

      i. Your paper should read from left to right and top to bottom. If your readers have to follow a wandering path down the page, they will be confused.
         A. Do not put little clusters of words on the page.
         B. Start each paragraph at the left margin, but do not start each sentence at the left margin. Write paragraphs, not lists of sentences.
      ii. Every word on your page must be part of a sentence.
         A. Pick a word at random on your page. Can you find the capital letter that starts the sentence containing it? Can you find the period that ends the sentence? Can you find the subject and verb?
      iii. Each part of your project should begin with a statement of what you will prove.
         A. Do not assume that your reader has a copy of the problem statement. Although you can refer to theorems in the text by number (say: "by Lipschutz, 5.3.2"). you have to state completely what each part of your report proves.
         B. Do not simply copy the problem statement. The problem statement asks the reader to do something, which is not what you want. You want to say what you will calculate or prove, so rewrite the problem statement as a positive statement of what you will do.
         C. If you are only doing part of a problem or a special case, then your introductory statement should say exactly what you will do (not what you won't do). If the problem asks you to prove something for $n \times n$ matrices and you do the problem for $3 \times 3$ matrices, then your statement should start: "Let $M$ be a $3 \times 3$ matrix. Then ...".
      iv. Finish each part with a conclusion.
      v. If you type, you must still use standard mathematical notation. If you cannot get your word processor to print fractions, exponents, etc., then write the formulas in by hand.

vi. Good mathematical writing is clear but succinct. Don't overdo long calculations. Put in enough steps so another student could repeat your calculation, but no more.

(b) Things to Avoid

i. NEVER use arrows. They are not well-defined mathematical symbols. They have no meaning. Use words and mathematical formulas to convey your meaning.

A. If one system of equations is derived from another, don't just put an arrow to one from the other. Put in a sentence that says how the systems are related, for example: "Eliminating $x$ from all but the first equation, we obtain the system:.."

ii. Never use a triangle of three dots to replace the word "therefore".

A. Never use a triangle of three dots for anything else either.

iii. Do not use abbreviations or other informal grammatical elements. Prepare a manuscript that you would be proud to see published.

iv. Never turn in a paper that includes unfinished problems, fragments, scratchwork, or other clutter. The reader will try to read everything on your page, so make sure that everything on your page is clear and complete.

A. Do not turn in first drafts. Solve problems first on scratch paper, then write your final answer on the paper you turn in.

3. Projects may be turned in by groups of up to three students.

   (a) If two projects contain significant copied parts, both will receive zeros. I do not care who copied from whom.

4. Projects will be graded on a scale of 0-4

   (a) 0: not submitted, no attempt to write complete sentences, or no correct parts.
   (b) 1: many mathematical or grammatical errors
   (c) 2: some parts done correctly, not too many mathematical or grammatical errors
   (d) 3: most parts done correctly and few mistakes, writing is grammatically correct
   (e) 4: almost all parts done correctly and no mistakes, writing is clear and well-organized

5. Late papers will not be accepted except by prior arrangement.

   (a) If you cannot come to class, have a friend bring your work or send it by email.
   (b) Write your work up as you go along. If something happens and you cannot finish, you will have a partially completed assignment to turn in.

#### 1.5.3 Exams

Exam problems will be similar to the easier homework problems. Exams must be take individually. Bring plenty of blank $8\frac{1}{2}$" $\times$ 11" paper.

##### 1.5.3.1 Midterms

There will be two midterms: October 10 and November 21. You may use one $8\frac{1}{2}$" $\times$ 11" page of notes (both sides) for each test.

#### 1.5.3.2 Final

The final will be in the regular lecture room. You may use two $8\frac{1}{2}$" $\times$ 11" page of notes (both sides) for the final. The final exam will cover the entire course.

### 1.5.4 Grading

| **Grading System** | |
|---|---|
| Problems | 25% |
| Projects | 25% |
| Midterms | 25% |
| Final Exam | 25% |

Final grades will be assigned according to a scale no harsher than: $A \geq 85\%$, $B \geq 70\%$, $C \geq 60\%$.

# *Chapter 2 Projects*

## 2.1 Project 1: Due September 26

Construct four points in $\mathbb{R}^3$ by taking your Social Security number (scramble the digits if you wish), add three digits so you have 12 digits, add a minus sign to every other digit, and divide the 12 digits into four groups, with three digits in each group. Thus each group is a point. Call the four points $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}$. These points are the vertices of a tetrahedron.

1. Find the volume of your tetrahedron. If the volume is 0, change some of the digits so you get a non-zero volume.
2. Find the center and radius of the sphere passing through all four vertices. The center will be the point equidistant from the vertices; the radius will be the distance from the center to any of the vertices. Check your answer by checking that all four distances are equal.
3. Find the center and radius of the sphere tangent to the four faces (sides) of the tetrahedron. The projections from the center to the four faces will be the same length, and the radius will be the length of any of these projections. Check that all four projections are the same length.

**Extra Credit:** use computer graphics to sketch your answer.
I know this is a hard problem. We can talk about it in class if anyone asks.

## 2.2 Project 2: Due October 31

This project is entirely in $\mathbb{R}^3$. Let $\Pi_1$ and $\Pi_2$ be distinct planes through the origin. Let $L$ be the line of intersection of the planes. Let $Q_1$ be the reflection matrix for $\Pi_1$, and let $Q_2$ be the reflection matrix for $\Pi_2$.

1. Explain why $Q_2Q_1$ is a rotation about $L$ by an angle equal to twice the shortest angle from $\Pi_1$ to $\Pi_2$. This is hard. Even if you cannot do this part, you can use it to do the next part.
2. Let $L$ be a line parallel to a vector $(a, b, c)$, where $a, b, c$ are the first three digits from your Social Security number (scrambled if you wish). Let $d$ be the two-digit number formed from the fourth and fifth digits of your SSN. Find the matrix $Q$ which rotates space around $L$ by $d$ degrees.
   (a) Check that $Q$ is correct by choosing a point $\mathbf{p}$ more or less at random. Make sure that $\mathbf{p}$ is not on $L$. Let $\mathbf{q}$ be the projection of $\mathbf{p}$ onto $L$. Show
      i. $Q\mathbf{q} = \mathbf{q}$
      ii. The projection of $Q\mathbf{p}$ onto $L$ is equal to $\mathbf{q}$.
      iii. The angle between $\mathbf{p} - \mathbf{q}$ and $Q\mathbf{p} - \mathbf{q}$ is $d$ degrees.

## 2.3 Project 3: Due November 21

This project involves a lot of computer calculating, which you can do with *C++* or *Mathematica* or *TestGiver.* Either way you must write a report separate from your computer code that explains the questions you were studying and the results you found. Computer code should be relegated to an appendix.

Here is the theory you will use. Let $A$ be an $n \times n$ ;symmetric matrix, so all the eigenvalues of $A$ are real. Assume that the largest eigenvalue of $A$ (largest in absolute value) has multiplicity one. Then the following scheme is a practical way to find the largest eigenvalue.

1. Pick a vector $\mathbf{x}_0 \in \mathbb{R}^n$ at random.
2. Define $\mathbf{x}_{i+1} = \dfrac{A^2\mathbf{x}_i}{||A^2\mathbf{x}_i||}$.
3. Keep generating vectors $\mathbf{x}_i$ until $\mathbf{x}_i \approx \mathbf{x}_{i+1}$. Then $\mathbf{x}_{i+1}$ is an eigenvector with largest eigenvalue.

You are not expected to prove that this procedure works. Here is what you have to do.

1. Pick a random $5 \times 5$ matrix $B$ and let $A = B + B^T$, so $A$ is symmetric..
2. Pick a random starting vector $\mathbf{x}_0 \in \mathbb{R}^5$
3. Compute vectors $\mathbf{x}_{i+1} = \dfrac{A^2\mathbf{x}_i}{||A^2\mathbf{x}_i||}$ until $\mathbf{x}_i \approx \mathbf{x}_{i+1}$. Now you have your eigenvector $\mathbf{x}$. (Computer science students—here's a chance to show off your programming skills. Write a program in *Mathematica* or *C++* that does the job.)
4. Once you have an eigenvector $\mathbf{x}$, compare $\mathbf{x}$ and $A\mathbf{x}$ to find the eigenvalue. Check that the eigenvalue you get is the same as the largest eigenvalue of $A$ as found by *Mathematica* or *TestGiver*.
5. **Optional part:** if you have succeeded in finding the largest eigenvalue, you can use the method to find the other eigenvalues.
   (a) Graph the characteristic polynomial of $A$ and estimate the roots, which are the eigenvalues of $A$.
   (b) Prove: $\mathbf{x}$ is an eigenvector of $A$ with eigenvalue $\lambda$ if and only if $\mathbf{x}$ is an eigenvector of $A - \alpha I$ with eigenvalue $\lambda - \alpha$.
   (c) If $A$ is invertible, prove: $\mathbf{x}$ is an eigenvector of $A$ with eigenvalue $\lambda$ if and only if $\mathbf{x}$ is an eigenvector of $A^{-1}$ with eigenvalue $\lambda^{-1}$.
   (d) Prove: if $\alpha$ is close to an eigenvalue $\lambda$, then $\lambda - \alpha$ will be the smallest (in absolute value) eigenvalue of $A - \alpha I$, so $\dfrac{1}{\lambda - \alpha}$ will be the largest eigenvalue of $(A - \alpha I)^{-1}$.
   (e) To estimate an eigenvalue $\lambda$ of $A$, pick $a$ close to $\lambda$ (estimate from the graph of the characteristic polynomial) and then find the eigenvalue $\lambda$ by using the power method to find the largest eigenvalue $\dfrac{1}{\lambda - a}$ of $(A - aI)^{-1}$ . Do this for all the remaining eigenvalues of $A$.

This method is called the **power method** for finding eigenvalues. It is widely used.

## 2.4 Skew-symmetric Matrices—due December 12

Every matrix in this problem is real unless stated otherwise.

A matrix $A$ is skew-symmetric if $A^T = -A$. Write an essay about skew-symmetric matrices, including the following points. The points you must cover are listing in a jumbled order. You should reorganize the points to make a coherent essay. It might be a good idea to look up skew-symmetric matrices in the library to get some more theorems to include. Don't copy the exact words you find in books. Just copy the ideas, include the books in a bibliography, and put in a footnote connecting the idea to the reference. You are not expected to discover original mathematics on your own, but you are expected to write proofs in your own words.

1. If possible, say something about the determinant of a skew-symmetric $3 \times 3$ matrix.
2. The eigenvalues of a skew-symmetric matrix are purely imaginary (multiples of $i$).
3. The determinant of a $2 \times 2$ skew-symmetric matrix is non-negative. ($\det A \geq 0$).
4. If $A$ and $B$ are skew-symmetric, then so is $A + B$ and $cA$
5. Every square matrix is the sum of a symmetric and a skew-symmetric matrix. (Hint: $A = \dfrac{A + A^T}{2} + \dfrac{A - A^T}{2}$.)
6. If $A$ and $B$ are skew-symmetric, $AB$ need not be skew-symmetric.
7. The diagonal elements of a skew-symmetric matrix are all 0.
8. If $A$ is $2 \times 2$ and skew-symmetric and $\det A = 0$ then $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

# Chapter 3 Lectures

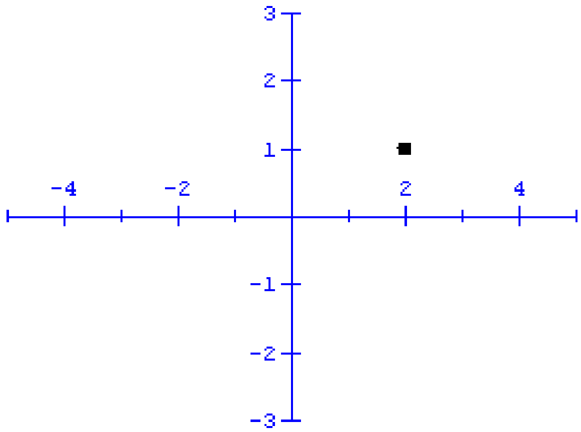
## 3.1 Introduction to TestGiver

1. I've written a computer program **TestGiver** to "help" you do your homework.
   (a) Truth in advertising: **TestGiver** helps me by automatically grading and recording your homework
   (b) **TestGiver** helps you in two ways:
      i. **TestGiver** tells you if your answer is right or wrong.
         A. Usually you can keep trying for a right answer
      ii. **TestGiver** includes computational and graphical tools
2. You will download **TestGiver** from the class website and install it on your PC, or you can use it in the Math/Stat lab in TH 409.
3. You will download weekly homework assignments from the class website.
   (a) You can work on an assignment, save your work, and continue at a later time.
   (b) You can print your homework, work out the answers, then go a computer to enter them. You do not have to do all your work on the computer.
   (c) When you have finished your homework, you send it to me over the internet with one click of the mouse.
4. If using this system causes a hardship for you, please see me. We'll work out a different homework system for you.

## 3.2 Two-Vectors

### 3.2.1 What is a two-vector?

1. A 2-vector is an array of two numbers presented vertically: $\begin{bmatrix} 3 \\ 7 \end{bmatrix}$
   (a) But sometimes, for reasons of typography, the vector will be presented as a row $[3, 7]$ or $[3\ 7]$ or $(3, 7)$.
   (b) In *Mathematica,* a vector is an array within braces `{3,7}`
      i. *Mathematica* has a very sophisticated way of managing vectors and matrices, which I will talk more about later.
   (c) In *TestGiver,* you should always write a vector as matrix with one column `[3;7]` using semicolons to separate the elements.
2. The set of 2-vectors is denoted $\mathbb{R}^2$
   (a) So $\mathbb{R}^2$ is the set of pairs, which can be thought of as representing the plane
3. You can think of vectors two ways:
   (a) as points in space

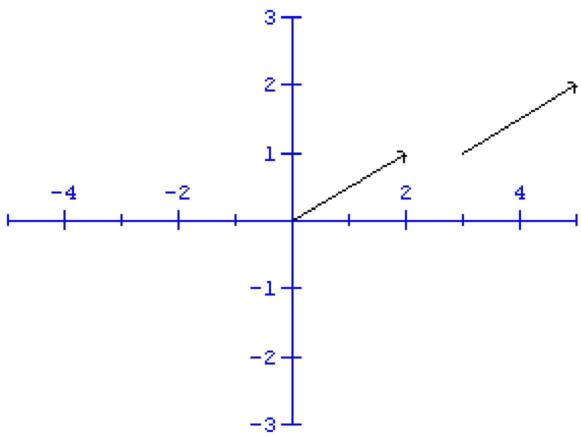



The vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

(b) as arrows between points

i. The same vector can start and end at different places

ii. a point is the same as an arrow from the origin to the point




The vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

(c) however you illustrate a vector, it **is** a column of numbers.

4. Special vector, the **zero vector**

$$\mathbf{0} = \mathbf{0}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

### 3.2.2 Vector Arithmetic

1. Vectors can be added, subtracted and multiplied by scalars (numbers):

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

$$\begin{bmatrix} 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$$

$$4\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$$

2. A **linear combination** of vectors is a sum of scalars times vectors:

$$3\begin{bmatrix} 1 \\ 3 \end{bmatrix} - 2\begin{bmatrix} 0 \\ -1 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 13 \end{bmatrix}$$

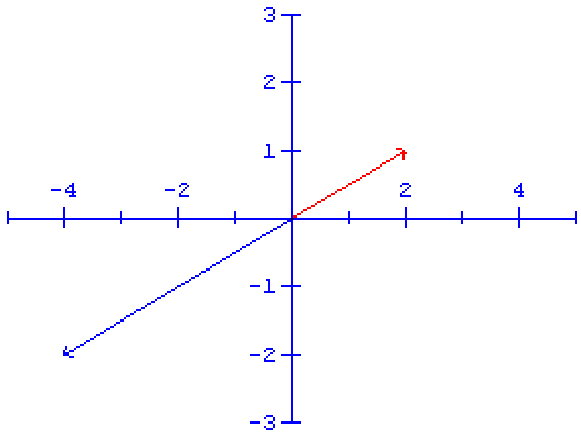
3. $\mathbf{0}+\mathbf{a}=\mathbf{a}+\mathbf{0}=\mathbf{a}$ for any vector $\mathbf{a}$.

4. **Class**: $\mathbf{a}=(3,5)$, $\mathbf{b}=(2,-4)$. Find $2\mathbf{a}-4\mathbf{b}$.

5. These operations have a geometric interpretation

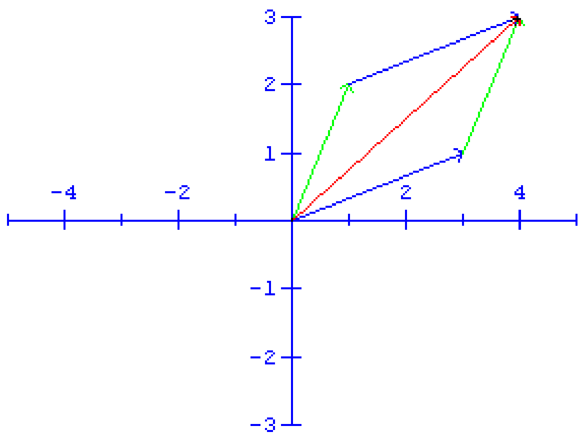
   (a) Multiplication is stretching

      i. multiplication by a negative reversed direction too



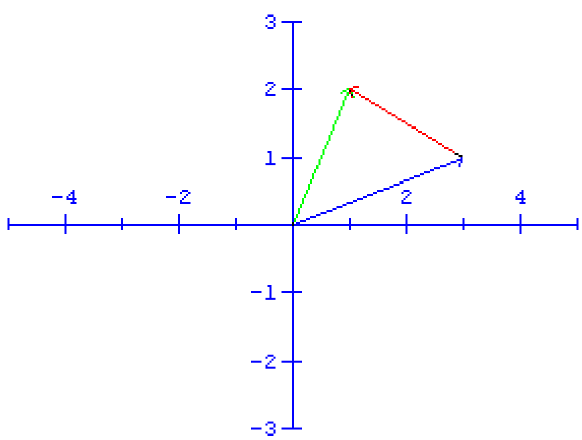

blue = $-2\times$ red

   (b) Addition adds distances from the origin




blue + green = red

   (c) Subtraction can be represented geometrically too.




red = green-blue

6. **Class**: continue example above. Sketch $\mathbf{a}=(3,5)$, $\mathbf{b}=(2,-4)$ and $\mathbf{a}+\mathbf{b}$ and $\mathbf{a}-\mathbf{b}$.

7. You can carry out these operations in software:

   (a) *TestGiver*

```
a = [2;3];
b = [4;-5];
c = 3*a - 2*b
```

Input

```
a : MATRIX(INTEGER)
a =
[2;
   3]
b : MATRIX(INTEGER)
b =
[4;
   -5]
c : MATRIX(INTEGER)
c =
[-2;
   19]
```

Output

   (b) *Mathematica*

```
In[4]:= a = {2, 3}
        b = {4, -5}
        c = 3 a - 2 b

Out[4]= {2, 3}

Out[5]= {4, -5}

Out[6]= {-2, 19}
```

### 3.2.3 Anatomy of a Vector

1. A 2-vector contains two numbers.

   (a) It is a nuisance to always give names to these numbers like this

$$\mathbf{a} = \begin{bmatrix} x \\ y \end{bmatrix}$$

   (b) Better to say that the top number is $\mathbf{a}[1]$ and the bottom number $\mathbf{a}[2]$.

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}[1] \\ \mathbf{a}[2] \end{bmatrix}$$

   (c) The formulas for arithmetic become

      i. addition: $(\mathbf{a} + \mathbf{b})[i] = \mathbf{a}[i] + \mathbf{b}[i]$

      ii. subtraction: $(\mathbf{a} - \mathbf{b})[i] = \mathbf{a}[i] - \mathbf{b}[i]$

      iii. scalar product: $(a\mathbf{b})[i] = a(\mathbf{b}[i])$.

2. **Class**: if $\mathbf{a} = (3, 5)$, $\mathbf{b} = (2, -4)$

   (a) what is $\mathbf{a}[1]\mathbf{b}[2] - \mathbf{a}[2]\mathbf{b}[1]$ ?

   (b) if $\mathbf{c} = 3\mathbf{a} - 4\mathbf{b}$, find $\mathbf{c}[2]$ quickly.

### 3.2.4 Inner or Dot Product

If $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ then

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}[1]\mathbf{b}[1] + \mathbf{a}[2]\mathbf{b}[2]$$

1. Sometimes the dot product is denoted $\mathbf{a} \cdot \mathbf{b} = \langle \mathbf{a}, \mathbf{b} \rangle$
2. If $\mathbf{a} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, then

$$\mathbf{a} \cdot \mathbf{b} = -1 \times 2 + 2 \times 3 = 4$$

3. **Class** find $\begin{bmatrix} 4 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -3 \end{bmatrix}$
4. *TestGiver*

```
a = [4;1];
b = [3;-3];
x = a.b
```

Input

```
a : MATRIX(INTEGER)
a =
[4;
   1]
b : MATRIX(INTEGER)
b =
[3;
   -3]
x : INTEGER
x = 9
```

Output

5. *Mathematica*

```
In[9]:= a = {4, 1}
        b = {3, -3}
        x = a.b

Out[9]= {4, 1}

Out[10]= {3, -3}

Out[11]= 9
```

#### 3.2.4.1 Perpendicular or Orthogonal Vectors:

1. $\mathbf{a} \perp \mathbf{b}$ if and only if $\mathbf{a} \cdot \mathbf{b} = 0$
   (a) **Class**: find a non-zero vector perpendicular to $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$.
2. In general, a vector orthogonal to $\begin{bmatrix} a \\ b \end{bmatrix}$ is $\begin{bmatrix} b \\ -a \end{bmatrix}$.

#### 3.2.4.2 Norm or Length of a Vector

1. If $\mathbf{a} \in \mathbb{R}^2$ then $||\mathbf{a}|| = \sqrt{\mathbf{a}[1]^2 + \mathbf{a}[2]^2}$ . This is the Pythagorean theorem.

2. $\mathbf{a} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $||\mathbf{a}|| = \sqrt{1+9} = \sqrt{10}$

3. The norm can be expressed with the dot product.

$$||\mathbf{a}|| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$$

   (a) Example:

$$\begin{aligned} ||\begin{bmatrix} 1 & 3 \end{bmatrix}|| &= \sqrt{\begin{bmatrix} 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 3 \end{bmatrix}} \\ &= \sqrt{10} \end{aligned}$$

4. **Class:** find $\left|\left|\begin{bmatrix} 2 \\ 5 \end{bmatrix}\right|\right|$

5. *TestGiver* calculates norms two ways.

```
a = [1;3];
x = sqrt(a.a);
x = Norm(a,2)
```

Input

```
a : MATRIX(INTEGER)
a =
[1;
   3]
x : REAL
x = 3.16228
x = 3.16228
```

Output

6. *Mathematica* calculates norms with square roots.

```
In[7]:= a = {1, 3}
        x = Sqrt[a.a]

Out[7]= {1, 3}

Out[8]= √10
```

#### 3.2.4.3 Unit Vectors

1. A **unit vector** is a vector of length 1.

   (a) So $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is not a unit vector, but $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ is a unit vector.

   (b) If $\mathbf{u}$ is any non-zero vector, $\frac{1}{||\mathbf{u}||}\mathbf{u}$ is a unit vector pointing the same direction as $\mathbf{u}$.

   (c) **Class:** find a unit vector parallel to $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

#### 3.2.4.4 Angle Between Vectors

1. If $\theta$ is the angle between vectors $\mathbf{a}$ and $\mathbf{b}$ then $\mathbf{a} \cdot \mathbf{b} = ||\mathbf{a}||\,||\mathbf{b}||\cos\theta$

   (a) The angle between $\mathbf{a} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ is found by:

$$\begin{aligned} \cos\theta &= \frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{a}||\,|\mathbf{b}|} \\ &= \frac{4}{\sqrt{5}\sqrt{13}} \\ \theta &= \cos^{-1}\left(\frac{4}{\sqrt{5}\sqrt{13}}\right) \\ & \quad 1.0517 \text{ (radians)} \end{aligned}$$

   (b) **Class**: find the angle between $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ -3 \end{bmatrix}$.

   (c) Notice that the angle is acute iff the dot product is positive.

   (d) *TestGiver*

```
a = [4;1];
b = [3;-3];
theta = acos(a.b/(sqrt(a.a)*sqrt(b.b)))
```

Input

```
a : MATRIX(INTEGER)
a =
[4;
   1]
b : MATRIX(INTEGER)
b =
[3;
   -3]
theta : REAL
theta = 1.03038
```

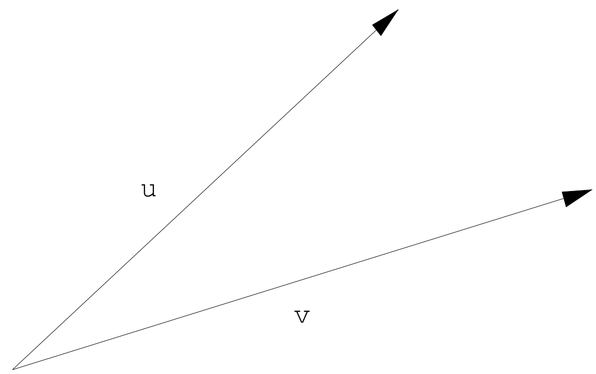
Output

   (e) *Mathematica*

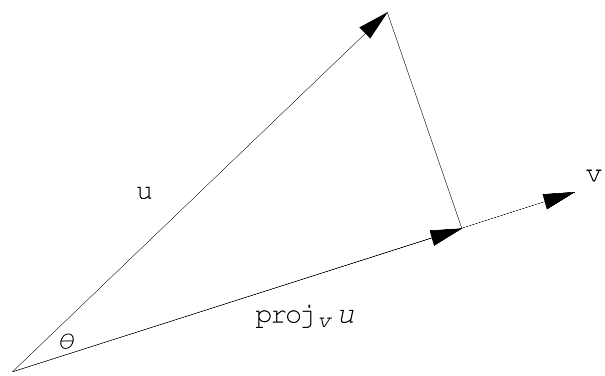
```
In[15]:= a = {4, 1}
         b = {3, -3}
         θ = ArcCos[a.b / (Sqrt[a.a] * Sqrt[b.b])]
         N[θ]

Out[15]= {4, 1}

Out[16]= {3, -3}

Out[17]= ArcCos[3/√34]

Out[18]= 1.03038
```

#### 3.2.4.5 Projections

1. Suppose we have two vectors $\mathbf{u}$ and $\mathbf{v}$:




The projection of **u** onto **v** is a vector parallel to **v** with length $||\mathbf{u}|| \cos\theta$, where $\theta$ is the angle between the vectors.




2. The formula for the projection vector is

$$\text{proj}_{\mathbf{v}}\mathbf{u} = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$$

   (a) Proof: since $\text{proj}_{\mathbf{v}}\mathbf{u}$ is a vector of length $||\mathbf{u}|| \cos\theta$ parallel to **v**, and since $\frac{1}{||\mathbf{v}||}\mathbf{v}$ is a unit vector parallel to **v**.

$$\text{proj}_{\mathbf{v}}\mathbf{u} = (||\mathbf{u}|| \cos\theta)\,\mathbf{v}$$

   But this expression can be simplified:

$$\begin{aligned}
\text{proj}_{\mathbf{v}}\mathbf{u} &= (||\mathbf{u}|| \cos\theta)\left(\frac{1}{||\mathbf{v}||}\mathbf{v}\right) \\
&= \left(||\mathbf{u}|| \frac{\mathbf{u}.\mathbf{v}}{||\mathbf{u}||\,||\mathbf{v}||}\right)\left(\frac{1}{||\mathbf{v}||}\mathbf{v}\right) \\
&= \left(\frac{\mathbf{u}\cdot\mathbf{v}}{||\mathbf{v}||^2}\right)\mathbf{v} \\
&= \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}
\end{aligned}$$

   (b) **Class:** find the projection of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ onto $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Draw a picture.

#### 3.2.4.6 Various inequalities with norms

1. Schwarz or Cauchy-Schwarz: $|\mathbf{u}\cdot\mathbf{v}| \leq ||\mathbf{u}||\,||\mathbf{v}||$

   (a) Of course! $||\mathbf{u}\cdot\mathbf{v}|| = ||\mathbf{u}||\,||\mathbf{v}|| \cos\theta$ and $\cos\theta \leq 1$

   (b) Proof by *Mathematica.* We must show

$$\begin{aligned}
||\mathbf{u}||^2\,||\mathbf{v}||^2 - (\mathbf{u}\cdot\mathbf{v})^2 &\geq 0 \\
(\mathbf{u}.\mathbf{u})^2\,(\mathbf{v}.\mathbf{v})^2 - (\mathbf{u}\cdot\mathbf{v})^2 &\geq 0
\end{aligned}$$

```
In[1]:= u = {u1, u2}; v = {v1, v2};
        FullSimplify[
          (u.u) (v.v) - (u.v)^2]

Out[2]= (u2 v1 - u1 v2)^2
```
i.

2. Minkowski: $||\mathbf{u}+\mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$.

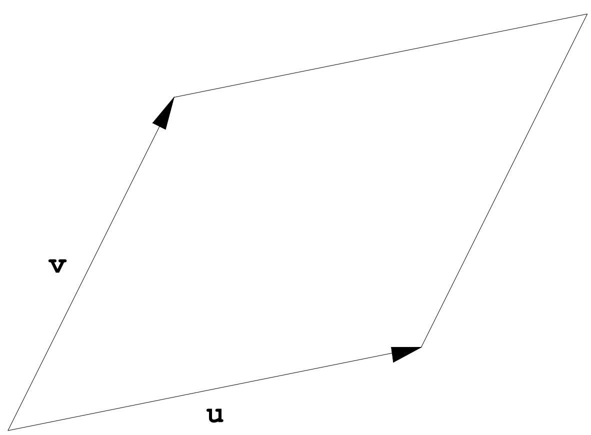
   (a) This is called the triangle inequality in elementary geometry

   (b) Direct proof:

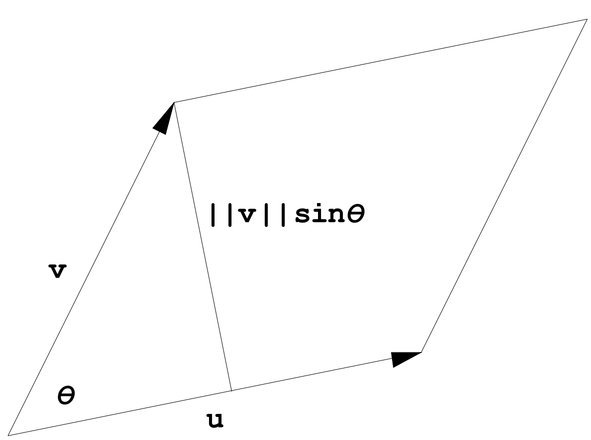
$$\begin{aligned}
(||\mathbf{u}|| + ||\mathbf{v}||)^2 &= ||\mathbf{u}||^2 + 2\,||\mathbf{u}||\,||\mathbf{v}|| + ||\mathbf{v}||^2 \\
&= \mathbf{u}.\mathbf{u} + 2\,||\mathbf{u}||\,||\mathbf{v}|| + \mathbf{v}\cdot\mathbf{v} \\
&\geq \mathbf{u}.\mathbf{u} + 2\mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{v} \\
&= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) \\
&= ||\mathbf{u}+\mathbf{v}||^2
\end{aligned}$$

### 3.2.5 Cross Product and Areas

1. Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$ determine a parallelogram.




   (a) The area is:




$$\text{area} = ||\mathbf{u}||\,||\mathbf{v}||\sin\theta$$

   (a) To find the area of the triangle formed by $\mathbf{u}$ and $\mathbf{v}$, take half the area of the parallelogram.

2. It turns out that there is an easy way to calculate this area.

   (a) In $\mathbb{R}^2$, define the **cross product** of **u** and **v**:

   $$\mathbf{u} \times \mathbf{v} = \mathbf{u}[1]\mathbf{v}[2] - \mathbf{u}[2]\mathbf{v}[1]$$

      i. Example:

   $$\begin{bmatrix} 3 \\ 2 \end{bmatrix} \times \begin{bmatrix} 5 \\ -1 \end{bmatrix} = 3 \times (-1) - 2 \times 5$$
   $$= -13$$

   (b) The area of the parallelogram is $|\mathbf{u} \times \mathbf{v}|$.

      i. If **u** and **v** have integer coordinates, then the parallelogram they span has integer area.

   (c) The sign of $\mathbf{u} \times \mathbf{v}$ is just as important than the numerical value.

      i. If $\mathbf{u} \times \mathbf{v} > 0$, then the angle from **u** to **v** is counterclockwise or positive

         A. The pair $\mathbf{u}, \mathbf{v}$ is **positively oriented**

      ii. If $\mathbf{u} \times \mathbf{v} < 0$, then the angle from **u** to **v** is clockwise or negative

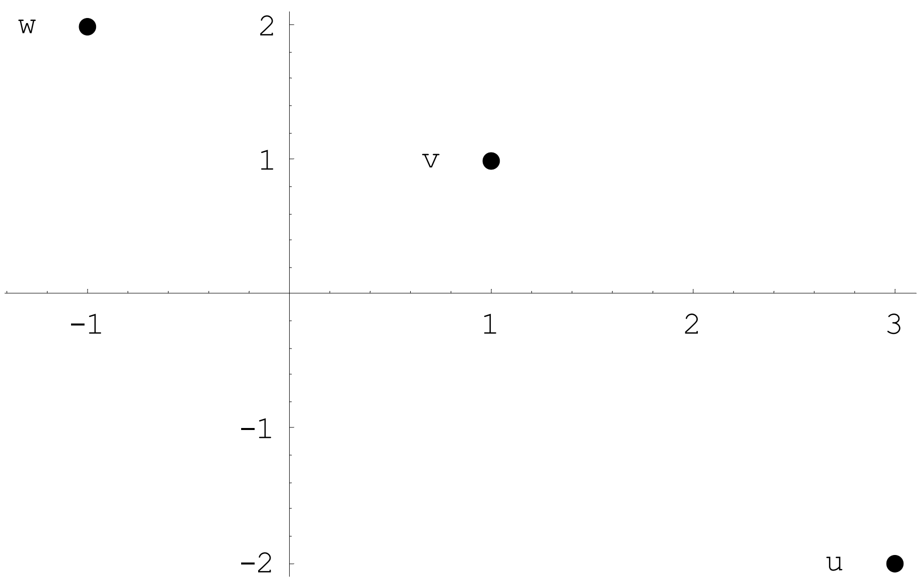
         A. The pair $\mathbf{u}, \mathbf{v}$ is **negatively oriented**

      iii. In either case the pair $\mathbf{u}, \mathbf{v}$ is **linearly independent**

      iv. If $\mathbf{u} \times \mathbf{v} = 0$, then **u** and **v** are parallel, pointing either in the same or opposite directions

         A. The pair $\mathbf{u}, \mathbf{v}$ is **linearly dependent**

3. **Class:** let $\mathbf{u} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.




   (a) Find the area of the triangle **uvw**.

      i. Find two sides of the triangle

      ii. Take their cross product

   (b) Show that if you go from **u** to **v** to **w** , you take a left turn at **v**

      i. Show that the pair $\mathbf{v} - \mathbf{u}, \mathbf{w} - \mathbf{v}$ is positively oriented

4. Algebra of cross products

   (a) $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$

   (b) $(a\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (a\mathbf{v}) = a(\mathbf{u} \times \mathbf{v})$

   (c) $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$

#### 3.2.6 Lines

1. There are three ways to represent a line in $\mathbb{R}^2$.
   (a) by two distinct points
   (b) by one point and a direction vector
   (c) by an linear equation like $3x - 2y = 1$
2. The real problem is to go from one representation to another.
3. **Class:** how many different changes of representation are possible? How many do you have to be able to do?
4. Changing representations
   (a) $(c) \longrightarrow (a)$: If the equation is $3x - 2y = 1$ how can you find two points on the line?
      i. Pick two $x$-values and solve for corresponding $y$-values. You can try $x = 0$ and $x = 1$.
      ii. For example $\begin{bmatrix} 0 \\ -1/2 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$
      iii. **Class:** find two points on the line with equation $2x + y = 3$
   (b) $(a) \longrightarrow (b)$: If the points are $\begin{bmatrix} 0 \\ -1/2 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, how do you find one point and a direction vector?
      i. To find one point on the line, pick one of the two given points. To find the direction vector, take the difference of the two given points.
      ii. In our example, take the base point to be $\begin{bmatrix} 0 \\ -1/2 \end{bmatrix}$ and the direction vector to be $\begin{bmatrix} 3 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ -1/2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3/2 \end{bmatrix}$.
      iii. This can be confusing, because in this method vectors are both points and directions. Actually, vectors are columns of numbers that can represent both points and directions.
      iv. **Class:** find a base point and direction vector for the line through the points $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.
   (c) $(b) -> (c)$: If the point is $\begin{bmatrix} 0 \\ -1/2 \end{bmatrix}$ and the direction vector is $\begin{bmatrix} 3 \\ 3/2 \end{bmatrix}$, how do we find the equation?
      i. A vector perpendicular to the direction vector gives the left-hand side of the equation. If the direction vector is $\begin{bmatrix} 3 \\ 3/2 \end{bmatrix}$, the perpendicular vector could be $\begin{bmatrix} 3/2 \\ -3 \end{bmatrix}$ and the left side of the equation would be:
      $$\frac{3}{2}x - 3y =?.$$
      ii. Evaluating the left-hand side at the given point gives the constant for the right-hand side:
      $$\frac{3}{2}(0) - 3\left(\frac{-1}{2}\right) = \frac{3}{2}$$
      iii. So the equation is:
      $$\frac{3}{2}x - 3y = \frac{3}{2}$$
      or equivalently
      $$3x = 6y = 1$$

iv. **Class:** if a line has base point $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and direction vector $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$, find the equation of the line.

5. The method for going $(b) \longrightarrow (c)$ is not obvious. Let's see why it works.

   (a) Using vector language: suppose we have a line given by a base point $\mathbf{p}$ and a direction vector $\mathbf{d}$

   (b) $\mathbf{x}$ is a point on the line if and only if $\mathbf{x} - \mathbf{p}$ is parallel to the direction vector $\mathbf{d}$

   (c) That is, $\mathbf{x}$ is a point on the line if and only if

   $$(\mathbf{x} - \mathbf{p}) \times \mathbf{d} = \mathbf{0}$$

   (d) We can rewrite this equation as:

   $$\mathbf{x} \times \mathbf{d} = \mathbf{p} \times \mathbf{d}$$

   (e) If $\mathbf{p} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\mathbf{d} = \begin{bmatrix} c \\ d \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, the equation becomes:

   $$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix} \times \begin{bmatrix} c \\ d \end{bmatrix} &= \begin{bmatrix} a \\ b \end{bmatrix} \times \begin{bmatrix} c \\ d \end{bmatrix} \\ dx - cy &= da - cb \end{aligned}$$

   This is just the formula given above.

6. **Class:** given a triangle with vertices $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

   (a) sketch the triangle

   (b) find the equation of the line from $\mathbf{u}$ to the midpoint of $\mathbf{vw}$

      i. before finding the equation, find another representation of the line, eg find two points or a point and a direction.

   (c) find the equation of the line from $\mathbf{u}$ perpendicular to $\mathbf{vw}$.

      i. before finding the equation, find another representation of the line, eg find two points or a point and a direction.

   (d) (Hard) find the equation of the line through $\mathbf{u}$ bisecting the angle at $\mathbf{u}$.

      i. before finding the equation, find another representation of the line, eg find two points or a point and a direction.

#### 3.2.6.1 Parametric Representation of lines

1. If a line has base point $\mathbf{b}$ and direction vector $\mathbf{d}$, then every point $\mathbf{x}$ on the line can be expressed as $\mathbf{x} = \mathbf{b} + t\mathbf{d}$ for some $t \in \mathbb{R}$.

   (a) This is called the parametric equation for the line. The parameter is $t$.

2. **Class:** is $\begin{bmatrix} 5 \\ 4 \end{bmatrix}$ on the line with base point $\mathbf{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and direction vector $\mathbf{d} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$?

   (a) Can you solve $\begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 2 \\ 1 \end{bmatrix}$?

3. Given two points $\mathbf{u}$ and $\mathbf{v}$, let $\mathbf{d} = \mathbf{v} - \mathbf{u}$. Then the interval from $\mathbf{u}$ to $\mathbf{v}$ is $\{\mathbf{u} + t\mathbf{d} : 0 \le t \le 1\}$.

#### 3.2.6.2 Intersection of Lines

1. Given two lines, you can find their intersection (if they intersect) by simultaneously solving their equations.

2. Example: given lines $3x - 2y = 7$ and $2x + y = 3$, solve the equations simultaneously:

$$\begin{aligned} 3x - 2y &= 7 \\ 2x + y &= 3 \end{aligned}$$

$$\begin{aligned} x - \frac{2}{3}y &= \frac{7}{3} \\ x + \frac{1}{2}y &= \frac{3}{2} \end{aligned}$$

$$\begin{aligned} x - \frac{2}{3}y &= \frac{7}{3} \\ \left(\frac{1}{2} + \frac{2}{3}\right) y &= \frac{3}{2} - \frac{7}{3} \end{aligned}$$

$$\begin{aligned} x - \frac{2}{3}y &= \frac{7}{3} \\ \frac{7}{6}y &= \frac{-5}{6} \end{aligned}$$

$$\begin{aligned} x - \frac{2}{3}y &= \frac{7}{3} \\ y &= \frac{-5}{7} \end{aligned}$$

$$\begin{aligned} x &= \frac{7}{3} + \frac{2}{3} \times \frac{-5}{7} \\ y &= \frac{-5}{7} \end{aligned}$$

$$\begin{aligned} x &= \frac{13}{7} \\ y &= \frac{-5}{7} \end{aligned}$$

   Don't forget to check by putting the answer back into the equations.

3. **Class:** given a triangle with vertices $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, find the intersection of the perpendicular bisectors of the sides.

### 3.2.7 Projection from point to line

## 3.3 Three-Vectors

Almost everything we said about two-vectors carries over to three-vectors. I will discuss mostly what is different.

### 3.3.1 What is a three-vector?

1. A 3-vector is an array of three numbers presented vertically: $\begin{bmatrix} 3 \\ 7 \\ 5 \end{bmatrix}$
   (a) But sometimes, for reasons of typography, the vector will be presented as a row $[3, 7, 5]$ or $[3\ 7\ 5]$ or $(3, 7, 5)$.
   (b) In *Mathematica,* a vector is an array within braces {`3,7,5`}
      i. *Mathematica* has a very sophisticated way of managing vectors and matrices, which I will talk more about later.
   (c) In *TestGiver,* you should always write a vector as matrix with one column `[3;7;5]` using semicolons to separate the elements.
2. The set of 3-vectors is denoted $\mathbb{R}^3$
   (a) So $\mathbb{R}^3$ is the set of triples, which can be thought of as representing the space
3. You can think of vectors two ways: as points in space or as arrows between points.
   (a) However you picture a vector, it **is** a column of numbers.
4. Special vector, the **zero vector**

$$\mathbf{0} = \mathbf{0}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

### 3.3.2 Vector Arithmetic

1. Vectors can be added, subtracted and multiplied by scalars (numbers):

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

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

$$4 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 4 \end{bmatrix}$$

2. A **linear combination** of vectors is a sum of scalars times vectors:

$$3 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 13 \\ 2 \end{bmatrix}$$

3. $\mathbf{0} + \mathbf{a} = \mathbf{a} + \mathbf{0} = \mathbf{a}$ for any vector $\mathbf{a}$.
4. **Class**: $\mathbf{a} = (3, 5, 7)$, $\mathbf{b} = (2, -4, 6)$. Find $2\mathbf{a} - 4\mathbf{b}$.
5. These operations have a geometric interpretation
   (a) Multiplication is stretching
   (b) Addition adds distances from the origin
   (c) Subtraction can be represented geometrically too.
6. You can carry out these operations in software with *TestGiver* and *Mathematica*

### 3.3.3 Anatomy of a Vector

1. A 3-vector contains three numbers.
   (a) It is a nuisance to always give names to these numbers like this

$$\mathbf{a} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

   (b) Better to say that the top number is $\mathbf{a}[1]$, the next number $\mathbf{a}[2]$ and the bottom number $\mathbf{a}[3]$.
   (c) The formulas for arithmetic become just like before
      i. addition: $(\mathbf{a}+\mathbf{b})[i] = \mathbf{a}[i] + \mathbf{b}[i]$
      ii. subtraction: $(\mathbf{a}-\mathbf{b})[i] = \mathbf{a}[i] - \mathbf{b}[i]$
      iii. scalar product: $(a\mathbf{b})[i] = a(\mathbf{b}[i])$.
2. **Class**: if $\mathbf{a} = (3,5,7)$, $\mathbf{b} = (2,-4,6)$
   (a) what is $\mathbf{a}[1]\mathbf{b}[2] - \mathbf{a}[2]\mathbf{b}[3]$ ?
   (b) if $\mathbf{c} = 3\mathbf{a} - 4\mathbf{b}$, find $\mathbf{c}[2]$ quickly.

### 3.3.4 Inner or Dot Product

If $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ then

$$\mathbf{a}\cdot\mathbf{b} = \mathbf{a}[1]\mathbf{b}[1] + \mathbf{a}[2]\mathbf{b}[2] + \mathbf{a}[3]\mathbf{b}[3]$$

1. Sometimes the dot product is denoted $\mathbf{a}\cdot\mathbf{b} = \langle \mathbf{a}, \mathbf{b} \rangle$
2. If $\mathbf{a} = \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$, then

$$\mathbf{a}\cdot\mathbf{b} = -1\times 2 + 2\times 3 - 3\times 4 = -8$$

3. **Class** find $\begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -3 \\ 2 \end{bmatrix}$
4. *TestGiver* and *Mathematica* will computed dot products.

#### 3.3.4.1 Perpendicular or Orthogonal Vectors:

1. $\mathbf{a} \perp \mathbf{b}$ if and only if $\mathbf{a}\cdot\mathbf{b} = 0$
   (a) **Class**: find a non-zero vector perpendicular to $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$.
   (b) **Class:** how many answers are possible

#### 3.3.4.2 Norm or Length of a Vector

1. If $\mathbf{a} \in \mathbb{R}^3$ then $||\mathbf{a}|| = \sqrt{\mathbf{a}[1]^2 + \mathbf{a}[2]^2 + \mathbf{a}[3]^2}$ . This is the Pythagorean theorem.
2. $\mathbf{a} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$, $||\mathbf{a}|| = \sqrt{1+9+4} = \sqrt{14}$
3. The norm can be expressed with the dot product.

$$||\mathbf{a}|| = \sqrt{\mathbf{a}\cdot\mathbf{a}}$$

(a) Example:

$$\begin{aligned} \left|\left|\begin{bmatrix} 1 & 3 & 2 \end{bmatrix}\right|\right| &= \sqrt{\begin{bmatrix} 1 & 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 3 & 2 \end{bmatrix}} \\ &= \sqrt{14} \end{aligned}$$

4. **Class:** find $\left|\left|\begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix}\right|\right|$

5. *TestGiver* and *Mathematica* calculate norms

#### 3.3.4.3 Unit Vectors

1. A **unit vector** is a vector of length 1.

   (a) So $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ is not a unit vector, but $\begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}$ is a unit vector.

   (b) If $\mathbf{u}$ is any non-zero vector, $\frac{1}{||\mathbf{u}||}\mathbf{u}$ is a unit vector pointing the same direction as $\mathbf{u}$.

   (c) **Class:** find a unit vector parallel to $\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$.

#### 3.3.4.4 Angle Between Vectors

1. If $\theta$ is the angle between vectors $\mathbf{a}$ and $\mathbf{b}$ then $\mathbf{a} \cdot \mathbf{b} = ||\mathbf{a}|| \, ||\mathbf{b}|| \cos\theta$

   (a) The angle between $\mathbf{a} = \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ is found by:

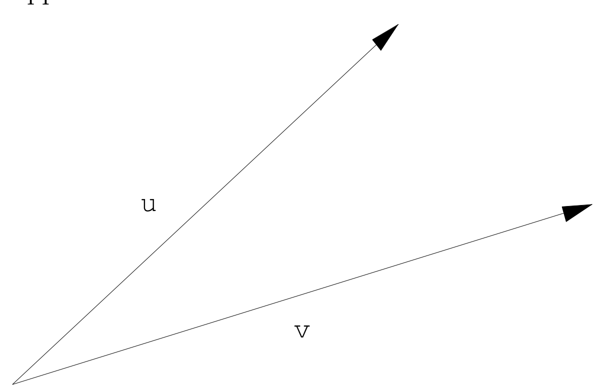
$$\begin{aligned} \cos\theta &= \frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{a}|| \, |\mathbf{b}|} \\ &= \frac{-8}{\sqrt{14}\sqrt{29}} \\ \theta &= \cos^{-1}\left(\frac{-8}{\sqrt{14}\sqrt{29}}\right) \\ & \quad 1.9791 \text{ (radians)} \end{aligned}$$

   (b) **Class**: find the angle between $\begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ -3 \\ 1 \end{bmatrix}$.

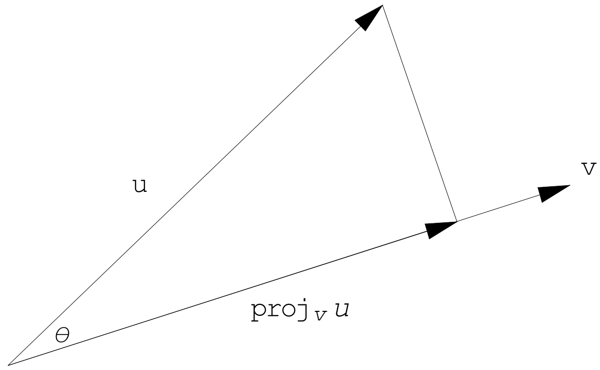
   (c) The angle is acute iff the dot product is positive.

#### 3.3.4.5 Projections

1. Suppose we have two vectors $\mathbf{u}$ and $\mathbf{v}$:



The projection of $\mathbf{u}$ onto $\mathbf{v}$ is a vector parallel to $\mathbf{v}$ with length $||\mathbf{u}|| \cos\theta$, where $\theta$ is the angle between the vectors.




2. The formula for the projection vector is

$$\mathrm{proj}_{\mathbf{v}}\mathbf{u} = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$$

   (a) Proof: exactly the same as before. Nothing in that proof assumed that the vectors were two-vectors.

   (b) **Class:** find the projection of $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ onto $\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

#### 3.3.4.6 Various inequalities with norms

Exactly the same as before.

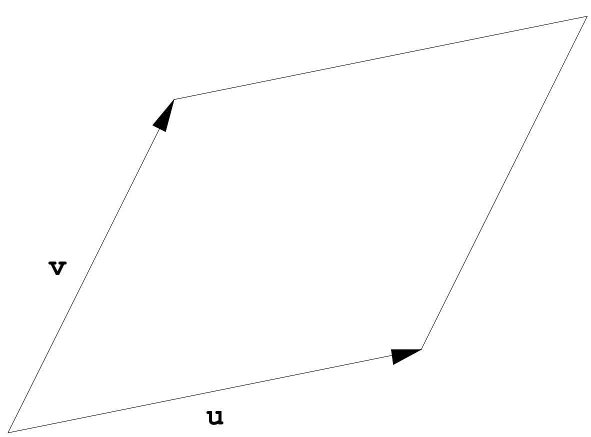
1. Schwarz or Cauchy-Schwarz: $|\mathbf{u}\cdot\mathbf{v}| \le ||\mathbf{u}||\,||\mathbf{v}||$
2. Minkowski: $||\mathbf{u}+\mathbf{v}|| \le ||\mathbf{u}|| + ||\mathbf{v}||$.

   (a) This is called the triangle inequality in elementary geometry

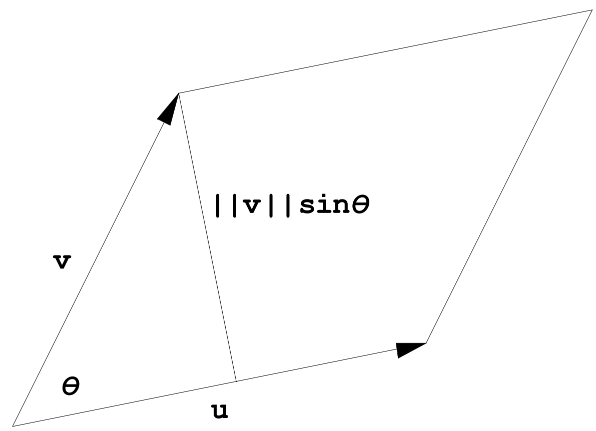
### 3.3.5 Cross Product and Areas and Volumes

**Watch out. Here comes some new stuff that separates space from the plane.**

1. Two vectors in $\mathbb{R}^3$ determine a parallelogram.




   (a) If the vectors are $\mathbf{u}$ and $\mathbf{v}$, then the area is:




$$\text{area} = ||\mathbf{u}||\,||\mathbf{v}||\sin\theta$$

(a) To find the area of the triangle formed by **u** and **v**, take half the area of the parallelogram.

2. It turns out that there is an easy way to calculate this area.

   (a) In $\mathbb{R}^3$, define the **cross product** of **u** and **v**:

$$\mathbf{u}\times\mathbf{v} = \begin{bmatrix} \mathbf{u}[2]\mathbf{v}[3]-\mathbf{u}[3]\mathbf{v}[2] \\ \mathbf{u}[3]\mathbf{v}[1]-\mathbf{u}[1]\mathbf{v}[3] \\ \mathbf{u}[1]\mathbf{v}[2]-\mathbf{u}[2]\mathbf{v}[1] \end{bmatrix}$$

   i. Note that in $\mathbb{R}^3$ the cross product is a vector, while in $\mathbb{R}^2$ the cross product is a number

   ii. Example:

$$\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} \times \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2\times 2-(-1)\times(-1) \\ (-1)\times 5-3\times 2 \\ 3\times(-1)-2\times 5 \end{bmatrix}$$
$$= \begin{bmatrix} 3 \\ -11 \\ -13 \end{bmatrix}$$

   (b) The area of the parallelogram is $||\mathbf{u}\times\mathbf{v}||$.

   i. In the example, the area is $\sqrt{3^2+(-17)^2+(-13)^2} = 21.61$

   ii. It is **not** true that if **u** and **v** have integer coordinates, then the parallelogram they span has integer area.

   (c) We cannot talk about the sign of $\mathbf{u}\times\mathbf{v}$.

   i. Instead we talk about the direction of $\mathbf{u}\times\mathbf{v}$.

   ii. $\mathbf{u}\times\mathbf{v}$ is perpendicular to both **u** and **v**. If you point the fingers of your right hand in the direction **u** so that your palm faces **v**, then your thumb points toward $\mathbf{u}\times\mathbf{v}$.

   iii. If $\mathbf{u}\times\mathbf{v}$ =0, then **u** and **v** are parallel, pointing either in the same or opposite directions

3. If you have three vectors, $\mathbf{u},\mathbf{v},\mathbf{w}$, they span a parallelopiped whose volume is $|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|$.

   (a) This is called the **triple product** of $\mathbf{u},\mathbf{v},\mathbf{w}$.

   (b) Up to sign, it doesn't matter in what order you take the vectors

4. **Class:** let $\mathbf{u} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$

(a) Find the volume of the tetrahedron with these vertices.

  i. Find three edges of the tetrahedron emanating from a single vertex

  ii. Take their triple product. The volume of the tetrahedron is $1/6$ of this value.

5. Algebra of cross products

   (a) $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$

   (b) $(a\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (a\mathbf{v}) = a(\mathbf{u} \times \mathbf{v})$

   (c) $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$

### 3.3.6 Planes

1. There are three principal ways to represent a plane in $\mathbb{R}^3$

   (a) three non-colinear points (not on one line)

   (b) a point and a vector normal to the plane

   (c) a linear equation like $2x - 3y + z = 6$

2. The real problem is to go from one representation to another.

3. **Class:** how many different changes of representation are possible? How many do you have to be able to do?

4. Changing representations

   (a) $(c) \longrightarrow (a)$: If the equation is $3x - 2y + z = 1$ how can you find three points on the plane?

      i. Pick to $x$-values and $y$-values and solve for corresponding $z$-values. Usually you can use $x = 1, y = 0$ and $x = 0, y = 1$, $x = y = 1$

      ii. For example $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

      iii. If this doesn't work, try other values for $x$ and $y$. Just be sure that your points are not colinear by checking that $\mathbf{v} - \mathbf{u}$ is not parallel to $\mathbf{w} - \mathbf{u}$. Check that $(\mathbf{v} - \mathbf{u}) \times (\mathbf{w} - \mathbf{u}) \neq \mathbf{0}$

$$\left(\begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}\right) \times \left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}\right)$$
$$= \begin{bmatrix} -3 \\ 2 \\ -1 \end{bmatrix}$$

      iv. **Class:** find three points on the plane with equation $2x + y - z = 3$. Check that your points are not colinear.

   (b) $(a) \longrightarrow (b)$: If the points are $\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, how do you find one point and a normal vector?

      i. To find one point on the plane, pick one of the three given points. To find the normal vector, take the cross product of two differences.

ii. In our example, take the base point to be $\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$ and the direction vector to be

$$\begin{aligned} & \left( \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} \right) \times \left( \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} \right) \\ = & \begin{bmatrix} -3 \\ 2 \\ -1 \end{bmatrix} \end{aligned}$$

.

iii. This can be confusing, because in this method vectors are both points and directions. Actually, vectors are columns of numbers that can represent both points and directions.

iv. **Class:** find a base point and direction vector for the plane through the three points you found on the plane $2x + y - z = 3$.

(c) $(b) -> (c)$: If the point is $\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$ and the normal vector is $\begin{bmatrix} -3 \\ 2 \\ -1 \end{bmatrix}$, how do we find the equation?

i. The normal vector gives the left-hand side of the equation:

$$-3x + 2y - z = ?.$$

ii. Evaluating the left-hand side at the given point gives the constant for the right-hand side:

$$-3\,(1) + 2\,(0) - (-2) = -1$$

iii. So the equation is:

$$-3x + 2y - z = -1$$

or equivalently

$$3x - 2y + z = 1$$

iv. **Class:** if a line has plane has base point and direction vector as found in your previous calculation, find the equation of the plane.

5. **Important:** For a plane, the coefficients of the equation are a normal vector.

6. The method for going $(b) \longrightarrow (c)$ is not obvious, but the proof that it is correct is almost same as the proof in the two-dimensional case.

(a) Using vector language: suppose we have a line given by a base point $\mathbf{p}$ and a normal vector $\mathbf{n}$

(b) $\mathbf{x}$ is a point on the plane if and only if $\mathbf{x} - \mathbf{p}$ is perpendicular to the normal vector $\mathbf{n}$

(c) That is, $\mathbf{x}$ is a point on the line if and only if

$$(\mathbf{x} - \mathbf{p}) \cdot \mathbf{n} = \mathbf{0}$$

(d) We can rewrite this equation as:

$$\mathbf{x} \cdot \mathbf{n} = \mathbf{p} \cdot \mathbf{n}$$

(e) If $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ and $\mathbf{p} = \begin{bmatrix} d \\ e \\ f \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, the equation becomes:

$$\begin{aligned} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} &= \begin{bmatrix} d \\ e \\ f \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ ax + by + cz &= ad + be + cf \end{aligned}$$

This is just the formula given above.

#### 3.3.6.1 Parametric Representation of Planes

1. If a plane is given by three non-colinear points $\mathbf{u}, \mathbf{v}, \mathbf{w}$, define $\mathbf{d} = \mathbf{v} - \mathbf{u}$ and $\mathbf{e} = \mathbf{w} - \mathbf{u}$.
   (a) Every point on the plane has the form $\mathbf{u}+s\mathbf{d}+t\mathbf{e}$.for some scalars $s, t \in \mathbb{R}$.
   (b) This is called the parametric equation for the plane. The parameters are $s$ and $t$.
2. **Class:** if a plane contains three points $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ find a fourth point on the plane.
3. The triangle with three given vertices $\mathbf{u}, \mathbf{v}, \mathbf{w}$ is the set of points $\{\mathbf{u}+s\mathbf{d}+t\mathbf{v} : 0 \le s, t; s+t \le 1\}$

### 3.3.7 Lines

1. Just like in $\mathbb{R}^2$, there are three ways to represent a line in $\mathbb{R}^3$.
   (a) by two distinct points
   (b) by one point and a direction vector
   (c) by two linear equations like $3x - y + z = 0$ and $x + y + z = 1$.
2. The real problem is to go from one representation to another.
3. **Class:** how many different changes of representation are possible? How many do you have to be able to do?
4. Changing representations
   (a) $(c) \longrightarrow (a)$: If the equations are $3x - y + z = 0$ and $x + y + z = 1$, how can you find two points on the line?
      i. Pick two $x$-values and solve for corresponding $y$- and $z$-values
      ii. For example $x = 0$ and $x = 1$.
         A. If $x = 0$ we have to find $y$ and $z$ such that

$$\begin{aligned} 3(0) - y + z &= 0 \\ 0 + y + z &= 1 \\ y &= z = 1/2 \end{aligned}$$

$$\begin{aligned} 3(1) - y + z &= 0 \\ 1 + y + z &= 1 \\ y &= 3/2 \\ z &= -3/2 \end{aligned}$$

         B. The two points are $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 3/2 \\ -3/2 \end{bmatrix}$.
      iii. **Class:** find two points on the line with equations $2x + y + z = 3$, $x - y + z = 0$. (Just set up the equations.)

$$\begin{aligned} y + z &= 3 \\ -y + z &= 0 \end{aligned}$$

, Solution is: $\left\{y = \frac{3}{2}, z = \frac{3}{2}\right\}$

$$\begin{aligned} 2 + y + z &= 3 \\ 1 - y + z &= 0 \end{aligned}$$

, Solution is: $\{y = 1, z = 0\}$

(b) $(a) \longrightarrow (b)$: If the points are $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 3/2 \\ -3/2 \end{bmatrix}$, how do you find one point and a direction vector?

i. To find one point on the line, pick one of the two given points. To find the direction vector, take the difference of the two given points.

ii. In our example, take the base point to be $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$ and the direction vector to be $\begin{bmatrix} 1 \\ 3/2 \\ -3/2 \end{bmatrix} - \begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix} =. \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$

iii. This can be confusing, because in this method vectors are both points and directions. Actually, vectors are columns of numbers that can represent both points and directions.

iv. **Class:** find a base point and direction vector for the line through the points $\begin{bmatrix} 0 \\ 3/2 \\ 3/2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

(c) $(b) -> (a)$: If the point is $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$ and the direction vector is $\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, how do we find two points.

i. We already have one point. That point plus the direction vector is a second point.

ii. In the example, the two points are $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{3}{2} \\ -\frac{3}{2} \end{bmatrix}$

(d) $(a) \longrightarrow (c)$ Given two points, how do we find two equations? (The equations are not unique—many choices are possible because the line can be the intersection of many different pairs of planes.)

i. Pick a point not on the line.Now you have three non-colinear points. They define a plane that has an equation.

ii. Pick a point not on the plane you just defined. This point plus the two points on the line are non-colinear and define a second plane that gives you your second equation.

iii. Example: suppose the two given points are $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ \frac{3}{2} \\ -\frac{3}{2} \end{bmatrix}$.

A. If you let $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ be the third point, the equation is of the plane through $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$, $\begin{bmatrix} 1 \\ \frac{3}{2} \\ -\frac{3}{2} \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ is $3x - y + z = 0$.

B. If you let $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ be the third point, the equation is of the plane through $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$, $\begin{bmatrix} 1 \\ \frac{3}{2} \\ -\frac{3}{2} \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is $x + y + z = 1$.

C. Two equations defining the line through $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ \frac{3}{2} \\ -\frac{3}{2} \end{bmatrix}$ are $3x - y + z = 0$ and $x + y + z = 1$

5. **Class:** find the line through the point $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ orthogonal to the plane $3x - 2y + z = 3$. What is the easiest way to represent the line?

#### 3.3.7.1 Parametric Representation of lines

1. If a line has base point $\mathbf{b}$ and direction vector $\mathbf{d}$, then every point $\mathbf{x}$ on the line can be expressed as $\mathbf{x} = \mathbf{b} + t\mathbf{d}$ for some $t \in \mathbb{R}$.

   (a) This is called the parametric equation for the line. The parameter is $t$.

   (b) If the base point is $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$ and the direction vector is $\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, then every point on the line can be expressed as

$$\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} t \\ 1/2 + t \\ 1/2 - 2t \end{bmatrix}$$

2. If a line is given by two points $\mathbf{u}$ and $\mathbf{v}$, and if $\mathbf{d} = \mathbf{v} - \mathbf{u}$, then the interval from $\mathbf{u}$ to $\mathbf{v}$ is $\{\mathbf{u} + t\mathbf{d} : 0 \le t \le 1\}$.

### 3.3.8 Intersections of Lines and Planes

1. The intersection of three planes usually defines a point. You can find the point by solving the three equations simultaneously.

   (a) example: suppose three planes have the equations:

$$\begin{aligned} 3x - 2y + z &= -2 \\ x + y - 2z &= 5 \\ 2x - y - z &= 1 \end{aligned}$$

   To find the common point, solve the equations:

$$\begin{aligned} x + y - 2z &= 5 \\ 5y + 7z &= -17 \\ -3y + 3z &= -9 \end{aligned}$$

$$\begin{aligned} x + y - 2z &= 5 \\ y - z &= 3 \\ 5y + 7z &= -17 \end{aligned}$$

$$\begin{aligned} x + y - 2z &= 5 \\ y - z &= 3 \\ 12z &= -32 \end{aligned}$$

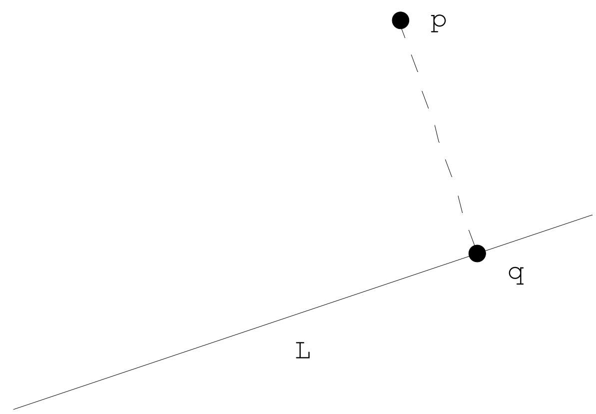
$$\begin{aligned} z &= -8/3 \\ y &= 1/3 \\ x &= -2/3 \end{aligned}$$

   The point is $\begin{bmatrix} -2/3 \\ 1/3 \\ -8/3 \end{bmatrix}$.

2. The intersection of a line and a plane is usually a point. Again, finding the point can be reduced to solving a system of three equations in three unknowns. The line gives two equations and the plane gives one equation.

### 3.3.9 Projection from a point to a Line or Plane

1. Given a point $\mathbf{p}$ and a line $L$, if you drop a perpendicular from $\mathbf{p}$ to $L$, how can you compute the point $\mathbf{q}$ where the perpendicular meets $L$?

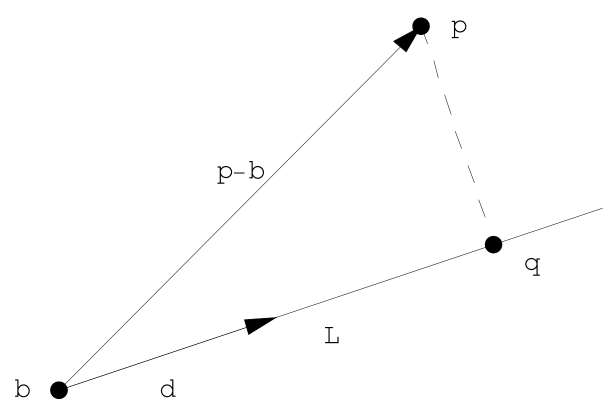
   (a) 
   

   (b) Before you can compute an answer, the question has to be more specific. The line $L$ and point $\mathbf{p}$ must be specified.

   (c) Suppose $L$ is given by a base point $\mathbf{b}$ and a direction vector $\mathbf{d}$.

   (d) Then

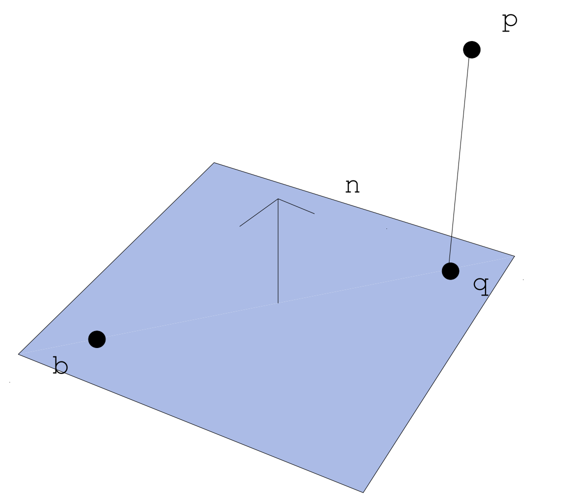
$$\mathbf{q}-\mathbf{b} = \frac{(\mathbf{p}-\mathbf{b})\cdot\mathbf{d}}{\mathbf{d}\cdot\mathbf{d}}\mathbf{d}$$
$$\mathbf{q} = \mathbf{b}+\frac{(\mathbf{p}-\mathbf{b})\cdot\mathbf{d}}{\mathbf{d}\cdot\mathbf{d}}\mathbf{d}$$




2. Given a point $\mathbf{p}$ and a plane $P$, find the point $\mathbf{q}$ where a perpendicular from the $\mathbf{p}$ meets $P$.

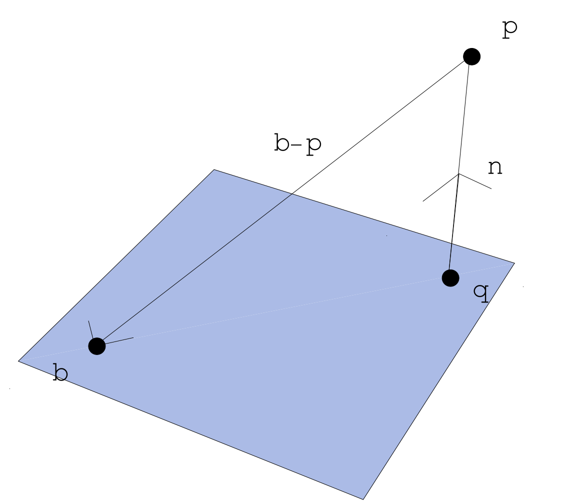
   (a) Before you can compute an answer, the question has to be more specific. The plane $P$ and point $\mathbf{p}$ must be specified.

   (b) Suppose the plane is given by a base point $\mathbf{b}$ and a normal vector $\mathbf{n}$.




(c) $\mathbf{q}-\mathbf{p}$ is the projection of $\mathbf{b}-\mathbf{p}$ onto $\mathbf{n}$.

$$\begin{aligned} \mathbf{q}-\mathbf{p} &= \frac{(\mathbf{b}-\mathbf{p})\cdot\mathbf{n}}{\mathbf{n}\cdot\mathbf{n}}\mathbf{n} \\ \mathbf{q} &= \mathbf{p}+\frac{(\mathbf{b}-\mathbf{p})\cdot\mathbf{n}}{\mathbf{n}\cdot\mathbf{n}}\mathbf{n} \end{aligned}$$




3. **Class:** find the projection of $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ onto the plane $x-y+z=0$ .

   (a) What would be a good base point for this plane? Can you find a normal vector?

## 3.4 Vectors in General, $n$-vectors

Almost everything we said about two-vectors and three-vectors carries over to $n$-vectors. I will discuss mostly what is different.

### 3.4.1 What is an $n$-vector?

1. An **vector,** or $n$-vector is an array of $n$ numbers presented vertically: $\begin{bmatrix} 3 \\ 7 \\ \vdots \\ 5 \end{bmatrix}$

   (a) But sometimes, for reasons of typography, the vector will be presented as a row $[3, 7, \ldots, 5]$ or $[3\ 7\ \cdots\ 5]$ or $(3, 7, \ldots, 5)$.

   (b) In *Mathematica,* a vector is an array within braces `{3,7,4,5}`

      i. *Mathematica* has a very sophisticated way of managing vectors and matrices, which I will talk more about later.

   (c) In *TestGiver,* you should always write a vector as matrix with one column `[3;7;4;5]` using semicolons to separate the elements.

2. The set of $n$-vectors is denoted $\mathbb{R}^n$

3. Special vector, the **zero vector**

$$\mathbf{0} = \mathbf{0}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

### 3.4.2 Vector Arithmetic

1. Vectors can be added, subtracted and multiplied by scalars (numbers)

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

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

$$4 \begin{bmatrix} 1 \\ -1 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 4 \\ 12 \end{bmatrix}$$

2. A **linear combination** of vectors is a sum of scalars times vectors:

$$3 \begin{bmatrix} 1 \\ 3 \\ 1 \\ 2 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \\ 13 \\ 2 \\ 7 \end{bmatrix}$$

3. $\mathbf{0} + \mathbf{a} = \mathbf{a} + \mathbf{0} = \mathbf{a}$ for any vector $\mathbf{a}$.

4. **Class**: $\mathbf{a} = (3, 5, 7, 6)$, $\mathbf{b} = (2, -4, 6, -1)$. Find $2\mathbf{a} - 4\mathbf{b}$.

5. These operations have a geometric interpretation

   (a) Multiplication is stretching

   (b) Addition adds distances from the origin

   (c) Subtraction can be represented geometrically too.

6. You can carry out these operations in software with *TestGiver* and *Mathematica*

### 3.4.3 Anatomy of a Vector

1. A vector contains $n$ numbers.

   (a) It is a nuisance to always give names to these numbers like this

$$\mathbf{a} = \begin{bmatrix} x \\ y \\ \vdots \\ z \end{bmatrix}$$

   (b) Better to say that the top number is $\mathbf{a}[1]$, the next number $\mathbf{a}[2]$ and the bottom number $\mathbf{a}[n]$.

   (c) The formulas for arithmetic become just like before

      i. addition: $(\mathbf{a} + \mathbf{b})[i] = \mathbf{a}[i] + \mathbf{b}[i]$

      ii. subtraction: $(\mathbf{a} - \mathbf{b})[i] = \mathbf{a}[i] - \mathbf{b}[i]$

      iii. scalar product: $(a\mathbf{b})[i] = a(\mathbf{b}[i])$.

2. **Class**: if $\mathbf{a} = (3, 5, 7, 4)$, $\mathbf{b} = (2, -4, 6, 2)$

   (a) what is $\mathbf{a}[1]\mathbf{b}[2] - \mathbf{a}[2]\mathbf{b}[4]$ ?

   (b) if $\mathbf{c} = 3\mathbf{a} - 4\mathbf{b}$, find $\mathbf{c}[3]$ quickly.

### 3.4.4 Inner or Dot Product

If $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ then

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= \mathbf{a}[1]\mathbf{b}[1] + \mathbf{a}[2]\mathbf{b}[2] + \cdots + \mathbf{a}[n]\mathbf{b}[n] \\ &= \sum_{i=1}^{n} \mathbf{a}[i]\mathbf{b}[i] \end{aligned}$$

1. Sometimes the dot product is denoted $\mathbf{a} \cdot \mathbf{b} = \langle \mathbf{a}, \mathbf{b} \rangle$

2. **Class** find $\begin{bmatrix} 4 \\ 1 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -3 \\ 2 \\ -1 \end{bmatrix}$

3. *TestGiver* and *Mathematica* will computed dot products.

#### 3.4.4.1 Perpendicular or Orthogonal Vectors:

1. $\mathbf{a} \perp \mathbf{b}$ if and only if $\mathbf{a} \cdot \mathbf{b} = 0$

   (a) **Class**: find a non-zero vector perpendicular to $\begin{bmatrix} 1 \\ -2 \\ 1 \\ -1 \end{bmatrix}$.

   (b) **Class:** how many answers are possible

#### 3.4.4.2 Norm or Length of a Vector

1. If $\mathbf{a} \in \mathbb{R}^n$ then $||\mathbf{a}|| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$ . This is the Pythagorean theorem.

2. **Class:** find $\left\| \begin{bmatrix} 2 \\ 5 \\ 3 \\ -1 \end{bmatrix} \right\|$

3. *TestGiver* and *Mathematica* calculate norms

#### 3.4.4.3 Unit Vectors

1. A **unit vector** is a vector of length 1.

   (a) So $\begin{bmatrix} 1 \\ 2 \\ -1 \\ 2 \end{bmatrix}$ is not a unit vector, but $\begin{bmatrix} 1/\sqrt{10} \\ 2/\sqrt{10} \\ -1/\sqrt{10} \\ 2/\sqrt{10} \end{bmatrix}$ is a unit vector.

   (b) If $\mathbf{u}$ is any non-zero vector, $\frac{1}{||\mathbf{u}||}\mathbf{u}$ is a unit vector pointing the same direction as $\mathbf{u}$.

   (c) **Class:** find a unit vector parallel to $\begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}$.

#### 3.4.4.4 Elementary Vectors

1. An elementary vector is a vector with all zeros except for a single 1

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

2. An elementary vector of length $n$ with the 1 in the $i^{th}$ place is denoted

$$\mathbf{e}_{n,i} = \begin{bmatrix} 0_1 \\ \vdots \\ 1_i \\ \vdots \\ 0_n \end{bmatrix}$$

   (a) If the length $n$ is clear from context, then we use the abreviated notation $\mathbf{e}_i$

3. Elementary vectors are unit vectors

4. Elementary vectors of the same length are either the same or orthogonal

#### 3.4.4.5 Orthonormal Sets of Vectors

1. A set of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_t$ is **orthonormal** if and only if

   (a) $||\mathbf{x}_i|| = 1$ all $i$

   (b) $\mathbf{x}_i \cdot \mathbf{x}_j = 0$ if $i \neq j$

   (c) Equivalently we can say $\mathbf{x}_i \cdot \mathbf{x}_j = \delta_{ij}$

      i. The Kronecker delta $\delta_{ij} = 0$ if $i \neq j$ and $= 1$ if $i = j$.

2. Example:

$$\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/2 \\ 1/2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

3. A set of distinct elementary vectors of the same length is orthonormal

4. Orthonormal sets of vectors are sometimes useful because they are like elementary (coordinate) vectors

#### 3.4.4.6 Angle Between Vectors

1. If $\theta$ is the angle between vectors $\mathbf{a}$ and $\mathbf{b}$ then $\mathbf{a} \cdot \mathbf{b} = ||\mathbf{a}||\,||\mathbf{b}|| \cos\theta$

   (a) The angle between $\mathbf{a} = \begin{bmatrix} -1 \\ 2 \\ -3 \\ 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix}$ is found by:

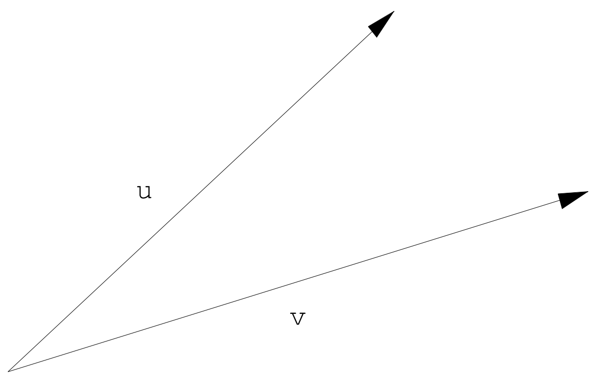
$$\begin{aligned} \cos\theta &= \frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{a}||\,|\mathbf{b}|} \\ &= \frac{-6}{\sqrt{18}\sqrt{30}} \\ \theta &= \cos^{-1}\left(\frac{-6}{\sqrt{18}\sqrt{30}}\right) \\ & \quad 1.8320 \text{ (radians)} \end{aligned}$$

   (b) **Class**: find the angle between $\begin{bmatrix} 4 \\ 1 \\ 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ -3 \\ 1 \\ 2 \end{bmatrix}$.

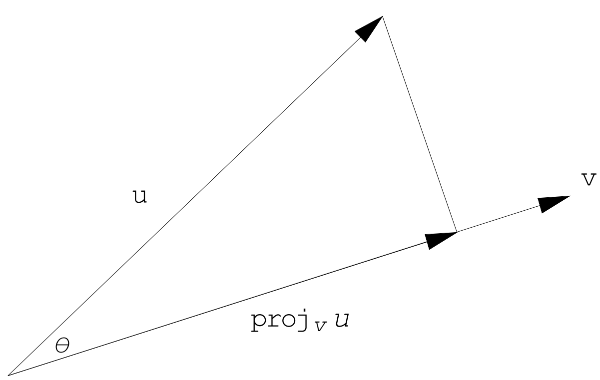
   (c) The angle is acute iff the dot product is positive.

#### 3.4.4.7 Projections

1. Suppose we have two vectors $\mathbf{u}$ and $\mathbf{v}$:




The projection of $\mathbf{u}$ onto $\mathbf{v}$ is a vector parallel to $\mathbf{v}$ with length $||\mathbf{u}|| \cos\theta$, where $\theta$ is the angle between the vectors.




2. The formula for the projection vector is

$$\text{proj}_{\mathbf{v}}\mathbf{u} = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\right) \mathbf{v}$$

   (a) Proof: exactly the same as before. Nothing in that proof assumed that the vectors were two-vectors.

(b) **Class:** find the projection of $\begin{bmatrix} 1 \\ 2 \\ -1 \\ 2 \end{bmatrix}$ onto $\begin{bmatrix} 3 \\ 1 \\ 2 \\ -1 \end{bmatrix}$.

#### 3.4.4.8 Various inequalities with norms

Exactly the same as before.

1. Schwarz or Cauchy-Schwarz: $|\mathbf{u} \cdot \mathbf{v}| \leq ||\mathbf{u}||\, ||\mathbf{v}||$
2. Minkowski: $||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$.
   (a) This is called the triangle inequality in elementary geometry

### 3.4.5 Areas and Volumes and Linear Independence

1. In $\mathbb{R}^n$
   (a) a single vector has length
   (b) two vectors determine a parallelogram which has area
   (c) three, four,$\ldots, n$ vectors span a parallelopiped with volume appropriate for its internal dimension. The volume of the intermediate cases is difficult to compute, but if there are $n$ vectors then we can find the volumne of the parallelopiped by using the determinant. See later.
2. The important qualification here is that, for $i$ vectors to span a parallelopiped of dimension $i$, the vectors must be **linearly independent.**
   (a) For two vectors to span a parallelogram, then cannot be colinear
   (b) for three vectors to span a parallelopiped, they cannot be coplanar.
   (c) We will give the general definition later.

## 3.5 Rectangular Matrices

### 3.5.1 Definition

1. A matrix is a rectangle of numbers, usually enclosed in brackets.

$$\begin{bmatrix} -85 & -55 & -37 & -35 \\ 97 & 50 & 79 & 56 \\ 49 & 63 & 57 & -59 \end{bmatrix}$$

   (a) Matrices come in different sizes
   (b) An $r \times c$ matrix is a matrix with $r$ rows and $c$ columns.
   (c) The example above is $3 \times 4$.
   (d) A matrix is called **square** if it has the same number of rows as columns, for example

$$\begin{bmatrix} -85 & -55 & -37 \\ -35 & 97 & 50 \\ 79 & 56 & 49 \end{bmatrix}$$

2. The element in row $i$, column $j$ of a matrix $A$ is denoted $A[i, j]$ in this class.
   (a) Elsewhere you may see $A_{ij}$ or $a_{ij}$
   (b) The submatrix consisting of rows $i_1 \cdots i_2$ and columns $j_1 \cdots j_2$ is denoted $A[i_1..i_2, j_1..j_2]$.

(c) Row $i$ of a $r \times c$ matrix $A$ can be denoted $row_i A$ or $A[i, 1..c]$

i. This is a $1 \times c$ matrix

(d) Column $j$ of a $r \times c$ matrix $A$ can be denoted $col_j A$ or $A[1..r, j]$

i. This is a $r \times 1$ matrix

3. Class: if
$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$
Find $A[2,1]$, $row_1 A$, $col_3 A$

4. Any matrix can be multiplied by a scalar. Matrices *of the same size* can be added, subtracted and put into linear combinations, just like vectors.

(a) Just work position by position
$$2\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix} - 3\begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} -4 & 1 \\ -11 & 3 \end{bmatrix}$$

(b) Class do
$$3\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

5. Something important to observe.

(a) Consider any of the objects $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{R}^{17}$, $2 \times 2$ matrices, $3 \times 4$ matrices, any other Euclidean space $\mathbb{R}^n$ or any other set of $r \times c$ matrices

(b) Each of these sets has addition, a zero, negation and scalar multiplication.

(c) Addition is commutative: $a + b = b + a$

(d) Scalar multiplication is associative: $\alpha(\beta a) = (\alpha\beta) a$

(e) scalar multiplication is distributive: $\alpha(a + b) = \alpha a + \alpha b$ and $(\alpha + \beta) a = \alpha a + \beta a$

(f) Here's the imprecise idea: these objects, and any others with the "same properties," are all called **vector spaces**.

6. What distinguishes matrices from vectors is matrix multiplication, a (vast) generalization of the dot product

(a) If $A$ is a row and $B$ is a column with the same number of elements, then $A \cdot B$ is their dot product, a scalar
$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 6 \\ 5 \\ 4 \end{bmatrix} = 28$$

(b) Class:
$$\begin{bmatrix} 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

(c) If $A$ is $r \times 1$ and $B$ is $1 \times c$, then $AB$ is an $r \times c$ matrix constructed as follows:
$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_c \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_c \\ a_2b_1 & a_2b_2 & \cdots & a_2b_c \\ \vdots & \vdots & \ddots & \vdots \\ a_rb_1 & a_rb_2 & \cdots & a_rb_c \end{bmatrix}$$

i. This is a matrix product, not a dot product

ii. That is, $(AB)[i,j] = A[i]B[j] = (row_i A)(col_j B)$

iii. Example:

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

iv. Class do

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

(d) If $C$ is a column with $r$ entries, and $R$ is a row with $c$ entries,

i. $CR$ is always defined, and is a $r \times c$ matrix

ii. $RC$ is defined if and only if $r = c$, in which case $RC$ is $1 \times 1$ or a scalar

(e) If $A$ is $r \times n$ and $B$ is $n \times c$ then we can define the product $AB$, an $r \times c$ matrix.

i. Write $A = \begin{bmatrix} A_1 \\ \vdots \\ A_r \end{bmatrix}$ and $B = \begin{bmatrix} B_1 & \cdots & B_c \end{bmatrix}$, where each $A_i$ is a row vector with $n$ entries, and each $B_i$ is a column vector with $n$ entries.

A. Therefore $A_i \cdot B_j$ is a scalar

ii. The product is

$$\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_r \end{bmatrix} \begin{bmatrix} B_1 & B_2 & \cdots & B_c \end{bmatrix} = \begin{bmatrix} A_1 \cdot B_1 & A_1 \cdot B_2 & \cdots & A_1 \cdot B_c \\ A_2 \cdot B_1 & A_2 \cdot B_2 & \cdots & A_2 \cdot B_c \\ \vdots & \vdots & \ddots & \vdots \\ A_r \cdot B_1 & A_r \cdot B_2 & \cdots & A_r \cdot B_c \end{bmatrix}$$

iii. That is, $(AB)[i,j] = (row_i A)(col_j B)$

A. $(AB)[i,j]$ is defined for $1 \le i \le r$ and $1 \le j \le c$ so $AB$ is $r \times c$.

B. An equivalent definition is $(AB)[i,j] = \sum_{k=1}^{n} A[i,k]\,B[k,j]$

iv. When $AB$ is defined, when $cols(A) = rows(B)$, we say that $A$ and $B$ are *conformable* for multiplication.

v. Example:

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

$$= \begin{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} & \begin{bmatrix} 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ \begin{bmatrix} 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} & \begin{bmatrix} 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ \begin{bmatrix} 1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} & \begin{bmatrix} 1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{bmatrix}$$

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

vi. Another point of view:

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

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

vii. Class: calculate

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

viii. Class: if $A$ is $3 \times 4$ and $B$ is $4 \times 2$, what is the size of $AB$? What is the size of $BA$?

(f) Doing it on the computer:

i. To enter a matrix in *TestGiver*, surround it with brackets `[..]` and separate the elements with commas, the rows with semicolons.

ii. *TestGiver*

```
A = [1,-1;2,0;1,-2];
B = [2,1;3,-1];
C = A*B
```

```
A : MATRIX(INTEGER)
A =
[1, -1;
  2, 0;
  1, -2]
B : MATRIX(INTEGER)
B =
[2, 1;
  3, -1]
C : MATRIX(REAL)
C =
[-1, 2;
  4, 2;
  -4, 3]
```

iii. *Mathematica*

```
a = {{1, -1}, {2, 0}, {1, -2}};
b = {{2, 1}, {3, -1}};
(c = a.b) // MatrixForm
```

```
//MatrixForm=
( -1  2 )
(  4  2 )
( -4  3 )
```

A. Semicolons at ends of lines suppress output

B. Use dot for matrix product, not space or '*'.

C. The optional command `MatrixForm` makes *Mathematica* display output nicely, but be careful.
If you say `c = a.b // MatrixForm` , the variable `c` will not be correctly defined
You must surround the assignment with parentheses for `c` to be defined correctly as the result matrix.

7. The elementary vectors are $\mathbf{e}_{i,n} = \begin{bmatrix} 0_1 \\ \vdots \\ 1_i \\ \vdots \\ 0_n \end{bmatrix}$. If $A$ is $r \times c$ then $A\mathbf{e}_{c,i} = col_i(A)$.

(a) Example: $r = 3$, $c = 4$, $i = 2$

$$\begin{aligned} A\mathbf{e}_{4,2} &= col_2(A) \\ \begin{bmatrix} 8 & 4 & -5 & -5 \\ 3 & -5 & 8 & 5 \\ -1 & 0 & 3 & -4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} &= \begin{bmatrix} 4 \\ -5 \\ 0 \end{bmatrix} \end{aligned}$$

8. Matrix multiplication obeys reasonable rules. Assuming all the products are defined:

   (a) $(AB)C = A(BC)$
   (b) $A(B+C) = AB + AC$
   (c) $(A+B)C = AC + BC$
   (d) But $AB$ is not always the same as $BA$.
      i. One product can even be defined when the other is not, for example if $A$ is $2 \times 2$ and $B$ is $2 \times 3$.
      ii. Even if both products are defined, they need not be equal. **Class:** calculate

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

9. Matrix multiplication is most often used to define **linear functions (linear transformations)** from one vector space to another.

   (a) If $A$ is $r \times c$ and $\mathbf{v} \in \mathbb{R}^c$ then $A\mathbf{v} \in \mathbb{R}^r$
   (b) That is, if $A$ is $r \times c$ and $\mathbf{v}$ is $c \times 1$ then $A\mathbf{v}$ is $r \times 1$.
   (c) So an $r \times c$ matrix $A$ defines a function

$$\mathbb{R}^c \xrightarrow{A} \mathbb{R}^r$$

   (d) Suppose we have two matrices, $A$ is $r \times n$ and $B$ is $n \times c$
      i. Then $AB$ is defined and is $r \times c$ and:

$$\begin{array}{ccccc} & & \mathbb{R}^n & & \\ & B \nearrow & & \searrow A & \\ \mathbb{R}^c & & \xrightarrow{AB} & & \mathbb{R}^r \end{array}$$

      This diagram "commutes" because $A(B\mathbf{v}) = (AB)\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^c$.
      ii. The product matrix corresponds to the composition of functions

10. Special matrices and matrix multiplication

   (a) zero matrix—any matrix with all zeros is denoted 0 or $0_{r,c}$ if we have to designate the size
      i. For any $r \times c$ matrix $A$ we have $0_{n,r}A = 0_{n,c}$ and $A0_{c,n} = 0_{r,n}$
      ii. Briefly, $0A = 0$ and $A0 = 0$
      iii. But I don't want to say $A0 = 0A$ because these might be different sizes!
      iv. If $A$ is $r \times c$ and $A\mathbf{x} = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^c$, then $A = 0$
         A. Proof: $col_i(A) = A\mathbf{e}_{c,i} = \mathbf{0}$. Since every column of $A$ is all zeros, $A$ is all zeros.
   (b) identity matrix

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

      i. denoted $I_n$ or $I$ if the size is obvious.
      ii. For any $r \times c$ matrix $A$ we have $I_rA = AI_c = A$
      iii. Briefly, $IA = AI = A$
      iv. **class:** check $I_2$ times some $2 \times 2$
      v. $col_i(I_n) = \mathbf{e}_{n,i}$
      vi. If $A$ is $n \times n$ and $A\mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$ then $A = I_n$.
         A. Proof: $col_i(A) = A\mathbf{e}_{n,i} = \mathbf{e}_{n,i}$ so $A = I_n$.

### 3.5.2 Transpose matrix

1. For any $r \times c$ matrix $A$ we have $A^T$ is a $c \times r$ matrix with the rows and columns of $A$ exchanged.

   (a) The rows (columns) of $A$ are the columns (rows) of $A^T$

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

   (b) Class find

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ a & b & c \end{bmatrix}^T$$

   (c) Formally, $A^T[i,j] = A[j,i]$

2. Note if $A$ and $B$ are matrices of the same size and $a$ is a scalar:

   (a) $A^{TT} = A$

      i. Proof:

$$\begin{aligned} A^{TT}[i,j] &= A^T[j,i] \\ &= A[i,j] \end{aligned}$$

      so $A^{TT}$ and $A$ agree in every position.

   (b) $(A+B)^T = A^T + B^T$

      i. Proof:

$$\begin{aligned} (A+B)^T[i,j] &= (A+B)[j,i] \\ &= A[j,i] + B[j,i] \\ &= A^T[i,j] + B^T[i,j] \\ &= \left(A^T + B^T\right)[i,j] \end{aligned}$$

      so $(A+B)^T$ and $\left(A^T + B^T\right)$ agree in every position.

   (c) $(aA)^T = a\left(A^T\right)$

      i. Class prove

3. If $A$ and $B$ are conformable for multiplication ($A$ is $r \times n$ and $B$ is $n \times c$) then $B^T$ and $A^T$can be multiplied, and

$$(AB)^T = B^T A^T$$

   (a) Proof: The sizes are right. $AB$ is $r \times c$ so $(AB)^T$ is $c \times r$. $B^T$ is $c \times n$ and $A^T$ is $n \times r$ so $B^T C^T$ is $c \times r$, the same size as $(AB)^T$. To show $(AB)^T = B^T A^T$, we need only show that the two matrices are the same in every position, since we know they are the same size:

$$\begin{aligned} (AB)^T[i,j] &= (AB)[j,i] \\ &= (row_j A) \cdot (col_i B) \\ &= \left(row_i B^T\right) \cdot \left(col_j A^T\right) \\ &= \left(B^T A^T\right)[i,j] \end{aligned}$$

   (b) **Class:** check

$$\left(\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\right)^T = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^T$$

### 3.5.3 Triangular Matrices

1. A matrix $A$ is **upper triangular** if $i > j \Longrightarrow A[i,j] = 0$.

   (a) Examples:
   $$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

2. A matrix $A$ is **lower triangular** if $i < j \Longrightarrow A[i,j] = 0$

   (a) Examples:
   $$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

3. Observe that the lower triangular examples are just the transposes of the upper triangular examples.

   (a) **Theorem**: $A$ is upper triangular if and only if $A^T$ is lower triangular

   (b) **Proof**: suppose $i > j$. Then $A[i,j] = 0$ if and only if $A^T[j,i] = 0$. Thus all the elements below the diagonal of $A$ are zero if and only if all the elements of $A^T$ above the diagonal are zero.

4. **Theorem:** If $A$ and $B$ are upper triangular and conformable, then $AB$ is upper triangular.

   (a) Same for lower triangular.

   (b) **Proof**: Suppose $A$ is $r \times n$ and $B$ is $n \times c$. Suppose $i > j$. We will show $AB[i,j] = 0$.

$$\begin{aligned} AB[i,j] &= row_i(A) \cdot col_j(B) \\ &= \begin{bmatrix} A[i,1] & \cdots & A[i,i-1] & A[i,i] & \cdots & A[i,n] \end{bmatrix} \begin{bmatrix} B[1,j] \\ \vdots \\ B[j,j] \\ B[j+1,j] \\ \vdots \\ B[n,j] \end{bmatrix} \\ &= \begin{bmatrix} 0 & \cdots & 0 & A[i,i] & \cdots & A[i,n] \end{bmatrix} \begin{bmatrix} B[1,j] \\ \vdots \\ B[j,j] \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ &= 0 \text{ because } i > j \end{aligned}$$

### 3.5.4 Orthogonal Matrices

1. A matrix $Q$ is **orthogonal** if and only if $Q^T Q = I$.

2. **Class:** check

   (a) The identity matrix is orthogonal

(b) The matrix

$$\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \end{bmatrix}$$

is orthogonal

3. **Class:** prove that if $Q_1$and $Q_2$ are orthogonal and conformable for multiplication, then $Q_1Q_2$ is orthogonal.

4. A matrix is *orthogononal* if and only if the *columns* are *orthonormal.*

### 3.5.5 Symmetric matrices

1. A *square* matrix $A$ is symmetric if $A = A^T$

   (a) Example:

$$\begin{bmatrix} 63 & 57 & 45 & 92 \\ 57 & -59 & -8 & 43 \\ 45 & -8 & -93 & -62 \\ 92 & 43 & -62 & 77 \end{bmatrix}$$

   The elements across the diagonal are equal.

2. **Theorem:** if $A$ is a square matrix then $A + A^T$ is symmetric.

3. **Class:** if $A$ is a square matrix $A$ then $AA^T$ is symmetric.

### 3.5.6 Projections and Reflections and Rotations

1. Let $\mathbf{u} \in \mathbb{R}^3$ be a unit vector. That means $\mathbf{u}^T\mathbf{u}$ =1. Let $S = \mathbf{u}\mathbf{u}^T$, a $3 \times 3$ matrix. Let $L$ be the line through the origin parallel to $\mathbf{u}$.

   (a) Earlier we showed that the orthogonal projection $\mathbf{q}$ from any point $\mathbf{p}$ to $L$ is

$$\mathbf{q} = (\mathbf{u} \cdot \mathbf{p})\,\mathbf{u}$$

   using $\mathbf{b} = \mathbf{0}$, $\mathbf{d} = \mathbf{u}$ and $\mathbf{u} \cdot \mathbf{u}$ =1.

   (b) The projection formula can be expressed in more matrix oriented language

$$\begin{aligned} \mathbf{q} &= (\mathbf{u} \cdot \mathbf{p})\,\mathbf{u} \\ &= \left(\mathbf{u}^T\mathbf{p}\right)\mathbf{u} \\ &= \mathbf{u}\left(\mathbf{u}^T\mathbf{p}\right) \\ &= \left(\mathbf{u}\mathbf{u}^T\right)\mathbf{p} \\ &= S\mathbf{p} \end{aligned}$$

   (c) We say that $S$ is the orthogonal projection matrix onto $L$.

   (d) If

$$\mathbf{u} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

   Then

$$\begin{aligned} S &= \mathbf{u}\mathbf{u}^T \\ &= \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \end{aligned}$$

(e) If $\mathbf{p} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ then

$$\begin{aligned} \mathbf{q} &= S\mathbf{p} \\ &= \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} \end{aligned}$$

and $\mathbf{q}$ is a vector on $L$.

2. Let $P = I - \mathbf{u}\mathbf{u}^T$, a $3 \times 3$ matrix. Let $\Pi$ be the plane through the origin orthgonal to $\mathbf{u}$.

   (a) Earlier we showed that the orthogonal projection $\mathbf{q}$ from any point $\mathbf{p}$ to $\Pi$ is

$$\mathbf{q} = \mathbf{p} - (\mathbf{p} \cdot \mathbf{u})\,\mathbf{u}$$

   using $\mathbf{b} = \mathbf{0}$ and $\mathbf{n} = \mathbf{u}$ and $\mathbf{u} \cdot \mathbf{u} = 1$

   (a) The projection formula can be expressed in more matrix oriented langauge

$$\begin{aligned} \mathbf{q} = \mathbf{p} - & \quad (\mathbf{p} \cdot \mathbf{u})\,\mathbf{u} \\ &= I\mathbf{p} - \left(\mathbf{u}^T\mathbf{p}\right)\mathbf{u} \\ &= I\mathbf{p} - \mathbf{u}\left(\mathbf{u}^T\mathbf{p}\right) \\ &= I\mathbf{p} - \left(\mathbf{u}\mathbf{u}^T\right)\mathbf{p} \\ &= \left(I - \mathbf{u}\mathbf{u}^T\right)\mathbf{p} \\ &= P\mathbf{p} \end{aligned}$$

   (b) We say that $P$ is the orthogonal projection matrix onto $\Pi$.

   (c) If

$$\mathbf{u} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

   then the equation for $\Pi$ is $x + y + z = 0$ and

$$\begin{aligned} P &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \\ &= \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \end{aligned}$$

   (d) Check: $P \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ is on $\Pi$.

   (e) Show for any unit vector $\mathbf{u}$ that $P$ is symmetric and $P^2 = P$. Remember, $\mathbf{u}^T\mathbf{u} = 1$.

$$\begin{aligned} P^T &= \left(I - \mathbf{u}\mathbf{u}^T\right)^T \\ &= I^T - \mathbf{u}^{TT}\mathbf{u}^T \\ &= I - \mathbf{u}\mathbf{u}^T \\ &= P \end{aligned}$$

$$
\begin{aligned}
P^2 &= \left(I - \mathbf{uu}^T\right)\left(I - \mathbf{uu}^T\right) \\
&= I - 2\mathbf{uu}^T + \mathbf{uu}^T\mathbf{uu}^T \\
&= I - 2\mathbf{uu}^T + \mathbf{uu}^T \\
&= I - \mathbf{uu}^T \\
&= P
\end{aligned}
$$

3. Let $Q = I - 2\mathbf{uu}^T$.We will show that $Q$ is a reflection across the plane $\Pi$ through the origin orthogonal to $\mathbf{u}$.

   (a) Let $\mathbf{p} \in \mathbb{R}^3$ and let $\mathbf{q} = P\mathbf{p}$ be the projection of $\mathbf{p}$ onto $\Pi$. Let $\mathbf{r}$ be the reflection of $\mathbf{p}$ across $\Pi$. Then

$$
\begin{aligned}
\mathbf{r} &= \mathbf{p} + 2\left(\mathbf{q} - \mathbf{p}\right) \\
&= \mathbf{p} + 2\left(P\mathbf{p} - \mathbf{p}\right) \\
&= (I + 2P - 2I)\mathbf{p} \\
&= (2P - I)\,\mathbf{p} \\
&= \left(2\left(I - \mathbf{uu}^T\right) - I\right)\mathbf{p} \\
&= \left(I - 2\mathbf{uu}^T\right)\mathbf{p} \\
&= Q\mathbf{p}
\end{aligned}
$$

   (b) Show: $Q$ is symmetric

$$
\begin{aligned}
Q^T &= I^T - \left(2\mathbf{uu}^T\right)^T \\
&= I - 2\mathbf{u}^{TT}\mathbf{u}^T \\
&= I - 2\mathbf{uu}^T \\
&= Q
\end{aligned}
$$

   (c) Show: $Q$ is orthogonal because $Q^TQ = Q^2 = I$

      i. Since $Q^T = Q$, $Q^TQ = Q^2$. Since $\mathbf{u}^T\mathbf{u} = 1$,

$$
\begin{aligned}
Q^2 &= \left(I - 2\mathbf{uu}^T\right)\left(I - 2\mathbf{uu}^T\right) \\
&= I - 4\mathbf{uu}^T + 4\mathbf{uu}^T\mathbf{uu}^T \\
&= I - 4\mathbf{uu}^T + 4\mathbf{uu}^T \\
&= I
\end{aligned}
$$

## 3.6 Solving Systems of Linear Equations

1. A *linear equation* is an equation where the variables are multiplied and constants, and the sum of the resulting products is constant.

$$3x - 2y + z = 5$$

2. A *system of linear equations* is a finite set of linear equations.

$$
\begin{aligned}
-2y + z &= 5 \\
2x + y - z &= 7 \\
-x - y + z &= 3
\end{aligned}
$$

3. You have solved such systems in high school; here we want to develop a systematic method of solving equations that works for all possible systems.

4. The easiest systems are those in *row-echelon* form:

(a) If the variables are ordered $x_1, x_2, \cdots, x_n$ then each equation starts with a later variable than the previous equation and the leading coefficient (coefficient of the first variable) is 1. We allow one line of the form $0 = 1$ and possibly some lines at the bottom of the form $0 = 0$.

(b) Example:

$$\begin{aligned} x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 &= \frac{7}{2} \\ x_2 - \frac{1}{2}x_3 &= \frac{-5}{2} \\ x_3 &= 21 \end{aligned}$$

$$\begin{aligned} x_1 - \frac{1}{2}x_2 + \frac{1}{2}x_3 &= \frac{3}{2} \\ x_2 - x_3 &= \frac{3}{5} \\ 0 &= 0 \end{aligned}$$

$$\begin{aligned} x_1 - \frac{1}{2}x_2 + \frac{1}{2}x_3 &= \frac{3}{2} \\ x_2 - x_3 &= \frac{3}{5} \\ 0 &= 1 \end{aligned}$$

5. The leading variables are called *pivot variables.* The others are *free variables.*

(a) **Class** identify in examples above.

6. So long as a row-echelon system does not include $0 = 1$ it can be solved by *back substitution*, making the pivot variables functions of the free variables

(a)

$$\begin{aligned} x_3 &= 21 \\ x_2 &= \frac{-5}{2} + \frac{1}{2}21 = 8 \\ x_1 &= \frac{7}{2} + \frac{1}{2}21 - \frac{1}{2}8 = 10 \end{aligned}$$

(b)

$$\begin{aligned} x_2 &= \frac{3}{5} + x_3 \\ x_1 &= \frac{3}{2} - \frac{1}{2}x_3 + \frac{1}{2}\left(\frac{3}{5} + x_3\right) = \frac{9}{5} \end{aligned}$$

7. **Class** try:

$$\begin{aligned} x + 3y - z &= 5 \\ z &= 7 \end{aligned}$$

$$\begin{aligned} x - 3y + 2z &= 5 \\ y - z &= 3 \end{aligned}$$

$$\begin{aligned} x - y + z &= 2 \\ y - 3z &= 5 \\ z &= 7 \end{aligned}$$

8. What transformations can you apply to a system of linear equations without changing the solutions?

   (a) exchange the position of two equations

   (b) mulitply or divide one equation by a non-zero constant

   (c) change an equation by adding or subtracting a multiple of a *different* equation from it.

9. These operations are called **row operations.**

10. By using row operations you can change any system of linear equations into a row-echelon system with the same solutions. Here is the method.

    (a) Find an equation that uses the first variable and put it on top

    (b) Divide the first equation by the leading coefficient

    (c) Subtract multiples of the first equation from the remaining equations so that the first variable is eliminated from all but the first equation:

    (d) Repeat with all but the top equation and fewer variables

    The result will be a system of equations in *row-echelon* form that can be solved by *back substitution.*

    (a) This is a recursive algorithm, but it stops because at each stage there are fewer equations and fewer variables.

11. Here is the method illustrated by an example

$$\begin{aligned} -2y+z &= 5 \\ 2x+y-z &= 7 \\ -x-y+z &= 3 \end{aligned}$$

    (a) Find an equation that uses the first variable and put it on top

$$\begin{aligned} 2x+y-z &= 7 \\ -2y+z &= 5 \\ -x-y+z &= 3 \end{aligned}$$

    (b) Divide the first equation by the leading coefficient

$$\begin{aligned} x+\frac{1}{2}y-\frac{1}{2}z &= \frac{7}{2} \\ -2y+z &= 5 \\ -x-y+z &= 3 \end{aligned}$$

    (c) Subtract multiples of the first equation from the remaining equations so that the first variable is eliminated from all but the first equation:

$$\begin{aligned} x+\frac{1}{2}y-\frac{1}{2}z &= \frac{7}{2} \\ -2y+z &= 5 \\ \frac{-1}{2}y+\frac{1}{2}z &= \frac{13}{2} \end{aligned}$$

    (d) Repeat with all but the top equation and fewer variables

$$\begin{aligned} -2y+z &= 5 \\ \frac{-1}{2}y+\frac{1}{2}z &= \frac{13}{2} \end{aligned}$$

(e) Find an equation that uses the first variable and put it on top

$$\begin{aligned} -2y + z &= 5 \\ \frac{-1}{2}y + \frac{1}{2}z &= \frac{13}{2} \end{aligned}$$

(f) Divide the first equation by the leading coefficient

$$\begin{aligned} y - \frac{1}{2}z &= \frac{-5}{2} \\ \frac{-1}{2}y + \frac{1}{2}z &= \frac{13}{2} \end{aligned}$$

(g) Subtract multiples of the first equation from the remaining equations so that the first variable is eliminated from all but the first equation:

$$\begin{aligned} y - \frac{1}{2}z &= \frac{-5}{2} \\ \frac{1}{4}z &= \frac{21}{4} \end{aligned}$$

(h) Repeat with all but the top equation and fewer variables.

$$\frac{1}{4}z = \frac{21}{4}$$

(i) Find an equation that uses the first variable and put it on top

$$\frac{1}{4}z = \frac{21}{4}$$

(j) Divide the first equation by the leading coefficient

$$z = 21$$

(k) Subtract multiples of the first equation from the remaining equations so that the first variable is eliminated from all but the first equation:

$$z = 21$$

(l) The system is now in row-echelon form:

$$\begin{aligned} x + \frac{1}{2}y - \frac{1}{2}z &= \frac{7}{2} \\ y - \frac{1}{2}z &= \frac{-5}{2} \\ z &= 21 \end{aligned}$$

You can use *back substitution* to find the solutions

$$\begin{aligned} z &= 21 \\ y &= 8 \\ x &= 10 \end{aligned}$$

12. **Class** try

$$\begin{aligned} 3x - y + z &= 2 \\ 2x + 2y - z &= 3 \\ -x + y + 2z &= 3 \end{aligned}$$

13. Does every system of equation have a solution?

(a) Of course not:

$$\begin{aligned} x+y &= 3 \\ x+y &= 4 \end{aligned}$$

14. Can a system of equations have more than one solution

    (a) of course

    $$x+y=3$$

    (b) This system has infinite solutions.

15. To review: a system of equations can have no solutions, one solution, or infinite solutions. Can a system have a more than one solution but only a finite number?

    (a) No, but we'll have to see why.

16. The method described above can always reduce a system to *row echelon form.*

    (a) This is the one we did above

    $$\begin{aligned} 2x_1+x_2-x_3 &= 7 \\ -2x_2+x_3 &= 5 \\ -x_1-x_2+x_3 &= 3 \end{aligned}$$

    reduces to

    $$\begin{aligned} x_1+\frac{1}{2}x_2-\frac{1}{2}x_3 &= \frac{7}{2} \\ x_2-\frac{1}{2}x_3 &= \frac{-5}{2} \\ x_3 &= 21 \end{aligned}$$

    This system has a unique solution

    (b) Class check this one:

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

    reduces to

    $$\begin{aligned} x_1-\frac{1}{2}x_2+\frac{1}{2}x_3 &= \frac{3}{2} \\ x_2-x_3 &= \frac{3}{5} \\ 0 &= 0 \end{aligned}$$

    This system has an infinite number of solutions, one for each value of $x_3$

    (c) **Class** check this one—just change one number in previous calculation:

    $$\begin{aligned} 2x_1-x_2+x_3 &= 3 \\ x_1+2x_2-2x_3 &= 3 \\ 3x_1+x_2-x_3 &= 7 \end{aligned}$$

    reduces to

    $$\begin{aligned} x_1-\frac{1}{2}x_2+\frac{1}{2}x_3 &= \frac{3}{2} \\ x_2-x_3 &= \frac{3}{5} \\ 0 &= 1 \end{aligned}$$

    This system cannot be solved.

(d) A system that cannot be solved is *inconsistent*. A system that has solutions is *consistent*.

   i. The only way a system can be inconsistent is to reduce to one with a line $0 = 1$.

(e) If a system is consistent then:

   i. either there are as many non-trivial equations as unknowns, every variable is a pivot variable, and the solution is unique

   ii. or there are fewer non-trivial equations than unknowns, and there are free variables

      A. You can assign any value you want to the free variables and the remaining pivot variables are determined.

      B. Since the free variables can be assigned values an infinite number of ways, there are infinite solutions when there are free variables.

(f) Thus every system of equations has no solutions, one solution, or infinite solutions

17. A system is in *row canonical* or *row reduced echelon* form if it is in row echelon form and each pivot variable appears in only one equation.

   (a) Once a system is reduced to echelon form, it is easy to reduce it to *row canonical* or *row reduced echelon* form by eliminating the coefficients above the pivot variables.

$$\begin{aligned} x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 &= \frac{7}{2} \\ x_2 - \frac{1}{2}x_3 &= \frac{-5}{2} \\ x_3 &= 21 \end{aligned}$$

   reduces to

$$\begin{aligned} x_1 \qquad &= 10 \\ x_2 \quad &= 8 \\ x_3 &= 21 \end{aligned}$$

18.

$$\begin{aligned} x_1 - \frac{1}{2}x_2 + \frac{1}{2}x_3 &= \frac{3}{2} \\ x_2 - x_3 &= \frac{3}{5} \\ 0 &= 0 \end{aligned}$$

reduces to

$$\begin{aligned} x_1 \qquad &= \frac{9}{5} \\ x_2 - x_3 &= \frac{3}{5} \\ 0 &= 0 \end{aligned}$$

   (a) Solution is:

$$\begin{aligned} x_1 &= \frac{9}{5} \\ x_2 &= x_3 + \frac{3}{5} \end{aligned}$$

19. Another example of system in row canonical form:

$$\begin{aligned} x_1 \quad + 2x_3 \quad &= 0 \\ x_2 + 3x_3 \quad &= 5 \\ x_4 &= 7 \end{aligned}$$

   (a) **Class** find pivot variables, free variables, general solution

20. **Class** try

$$\begin{aligned} x_1 + 2x_2 \quad - x_4 &= 3 \\ x_3 + 2x_4 &= 4 \end{aligned}$$

21. What can you say if you just know how many equations and how many variables there are?

   (a) Not much.
   (b) If there are more variables than non-trivial equations then either there are no solutions or infinite solutions.
   (c) Random coefficients—the usual case
      i. If there are fewer equations than variables *and if the coefficients are "random"* then there are infinite solutions
      ii. If there are as many equations as variables *and if the coefficients are "random"* then there is a unique solution
      iii. If there are more equations than variables *and if the coefficients are "random"* then there are no solutions

### 3.6.1 Homogeneous Systems of Linear Equations

1. Review: given a system of linear equations:

   (a) you can use row operations to reduce it to echelon form
   (b) then you can tell if there are any solutions
      i. you can also tell if there is a unique solution
      ii. always there are no solutions, a unique solution, or infinite solutions
   (c) if there are solutions, you can find them all by back substitution

2. A **homogeneous** system of linear equations is a system where all the constants are 0.

   (a) For example

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

   (b) A homogeneous system always has at least one solution, namely $(0, 0, \ldots, 0)$
   (c) The only interesting question about a homogeneous system is whether or not it has a non-zero solution
      i. which is the same as having infinite solutions
      ii. because if there is a non-zero solution then there are at least two solutions so there are infinite solutions

(d) The example reduces to:

$$\begin{aligned} x - \frac{2}{3}y + \frac{1}{3}z &= 0 \\ y - \frac{11}{7}z &= 0 \end{aligned}$$

The solutions are:

$$\begin{aligned} y &= \frac{11}{7}z \\ x &= \frac{2}{3}y - \frac{1}{3}z \\ &= \frac{5}{7}z \end{aligned}$$

3. A homogeneous system with fewer equations than unknowns **always** has non-zero solutions.

4. **Class**: is there a non-zero solution for:

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

## 3.7 Matrices and Linear Equations

1. Given a system of linear equations

$$\begin{aligned} 3x - 2y + z &= 5 \\ x \quad - 4z &= 7 \end{aligned}$$

we can derive two matrices from the system

(a) The **coefficient matrix of the system:**

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

(b) the **augmented matrix of the system:**

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

2. **Class**: find coefficient matrix and augmented matrix for

$$\begin{aligned} 2x + y - 2z &= 7 \\ 2y + z &= 5 \end{aligned}$$

3. Using matrices we can write a system of equation in two more equivalent forms

(a) system form:

$$\begin{aligned} 3x - 2y + z &= 5 \\ x \quad - 4z &= 7 \end{aligned}$$

(b) vector form:

$$x \begin{bmatrix} 3 \\ 1 \end{bmatrix} + y \begin{bmatrix} -2 \\ 0 \end{bmatrix} + z \begin{bmatrix} 1 \\ -4 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

i. This form is important when we study linear independence.

(c) matrix form:

$$\begin{bmatrix} 3 & -2 & 1 \\ 1 & 0 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

i. Notice that the coefficient matrix shows up in the matrix form.

ii. If $A$ is an $r \times c$ matrix and $\mathbf{b}$ is a $r \times 1$ matrix, then the equation

$$A\mathbf{x} = \mathbf{b}$$

has a $c \times 1$ matrix $\mathbf{x}$ as its solution. If we think of the entries of $\mathbf{x}$ as $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_c \end{bmatrix}$, then the equation $A\mathbf{x} = \mathbf{b}$ is the matrix form of a system of $r$ equations in $c$ variables.

iii. Very often, as the course progresses, I will say: "consider the system of linear equations $A\mathbf{x} = \mathbf{b}$". The matrix form will become the most usual form for a system of linear equations in this class.

(d) **class:** write the system

$$\begin{aligned} 3x - 2y &= 4 \\ 2x + y &= 6 \end{aligned}$$

in vector and matrix form.

(e) **class:** If $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, write the system $A\mathbf{x} = \mathbf{b}$ in system form.

4. Consider the parallels between a system and its augmented matrix.

(a) If you knew one, you would know the other (almost: the matrix doesn't tell you the names of the variables, but it tells you everything else)

(b) A matrix is in **row echelon form** if the first non-zero element in each row is 1, and the leading 1 in one row comes before the leading 1 in the next row.

i. leading ones are called **pivots**. Their locations are the **pivot positions**. The columns containing pivots are **pivot columns.**

(c) a matrix is in **row canonical** or **row reduced echelon** form if it is in echelon form and everything in a pivot column except the pivot is 0.

(d) Row operations can be done to the matrix:

| | |
|---|---|
| exchange equations | exchange rows |
| multiply equation by non-zero constant | multiply row by non-zero constant |
| add multiple of one equation to another | add multiple of one row to another |

5. Any matrix can be row reduced to echelon form by the following process:

(a) Find the leftmost non-zero column and swap rows so it has a non-zero element on top

(b) divide the first row by the first non-zero element so the first non-zero element is 1

(c) subtract multiples of the first row from lower rows, so the first non-zero column looks like $\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

(d) Repeat with the rows below the first row

6. Example:

$$\begin{bmatrix} 0 & -2 & 1 & 5 \\ 2 & 1 & -1 & 7 \\ -1 & -1 & 1 & 3 \end{bmatrix}$$

(a) The first non-zero column is the first column. Exchange first two rows.

$$\begin{bmatrix} 2 & 1 & -1 & 7 \\ 0 & -2 & 1 & 5 \\ -1 & -1 & 1 & 3 \end{bmatrix}$$

(b) Divide first row by 2.

$$\begin{bmatrix} 1 & 1/2 & -1/2 & 7/2 \\ 0 & -2 & 1 & 5 \\ -1 & -1 & 1 & 3 \end{bmatrix}$$

(c) Add 0 times first row to second row and 1 times first row to third row.

$$\begin{bmatrix} 1 & 1/2 & -1/2 & 7/2 \\ 0 & -2 & 1 & 5 \\ 0 & -1/2 & 1/2 & 13/2 \end{bmatrix}$$

(d) Repeat with bottom two rows. Second column is first non-zero column. No need to exchange

(e) Divide row 2 by $-2$:

$$\begin{bmatrix} 1 & 1/2 & -1/2 & 7/2 \\ 0 & 1 & -1/2 & -5/2 \\ 0 & -1/2 & 1/2 & 13/2 \end{bmatrix}$$

(f) Add 1/2 times row 2 to row 3 :

$$\begin{bmatrix} 1 & 1/2 & -1/2 & 7/2 \\ 0 & 1 & -1/2 & -5/2 \\ 0 & 0 & 1/4 & 21/4 \end{bmatrix}$$

(g) Repeat with the bottom row. Divide the bottom row by 1/4

$$\begin{bmatrix} 1 & 1/2 & -1/2 & 7/2 \\ 0 & 1 & -1/2 & -5/2 \\ 0 & 0 & 1 & 21 \end{bmatrix}$$

7. We could continue, starting with the bottom row and working up to put the matrix into row canonical form:

(a) Add 1/2 times row 3 to row 2 and row 1

$$\begin{bmatrix} 1 & 1/2 & 0 & 14 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 21 \end{bmatrix}$$

(b) Add $-1/2$ times row 2 to row 1

$$\begin{bmatrix} 1 & 0 & 0 & 10 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 21 \end{bmatrix}$$

8. the row echelon form is not unique

$$\begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$$
$$\rightarrow \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 \\ 0 & -1 \end{bmatrix}$$
$$\rightarrow \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

   (a) Hard theorem: the row canonical form is unique. In the example above, both results reduce further to $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

9. If you row reduce a random square matrix to row canonical form, you get $\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$, the **identity matrix**

   (a) If you row reduce a random matrix with more rows than columns, you get an identity matrix followed by rows of zeros

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

   (b) If you row reduce a random matrix with more columns than rows, you get an identity matrix followed by random non-zero values

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

### 3.7.1 Row Operations and Equation Solving

1. Way back when, we started with the system of equations

$$\begin{aligned} 2x + y - z &= 7 \\ -2y + z &= 5 \\ -x - y + z &= 3 \end{aligned}$$

   and we saw that it could be reduced to

$$\begin{aligned} x + \frac{1}{2}y - \frac{1}{2}z &= \frac{7}{2} \\ y - \frac{1}{2}z &= \frac{-5}{2} \\ z &= 21 \end{aligned}$$

   and further reduced to:

$$\begin{aligned} x &= 10 \\ y &= 8 \\ z &= 21 \end{aligned}$$

2. The augmented matrix of this system is $\begin{bmatrix} 0 & -2 & 1 & 5 \\ 2 & 1 & -1 & 7 \\ -1 & -1 & 1 & 3 \end{bmatrix}$

and above we just reduced this to

$$\begin{bmatrix} 1 & 1/2 & -1/2 & 7/2 \\ 0 & 1 & -1/2 & -5/2 \\ 0 & 0 & 1 & 21 \end{bmatrix}$$

and further to

$$\begin{bmatrix} 1 & 0 & 0 & 10 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 21 \end{bmatrix}$$

3. So we see that operating with equations or matrices yields the same result.
   (a) Matrices allow us to eliminate the pesky variables.
   (b) from now on, when you have a system of equations, you should replace it by its augmented matrix and reduce the matrix.
   (c) This is especially easy in software.
      i. *Testgiver*: the matrix above is

      $$[0, -2, 1, 5; 2, 1, -1, 7; -1, -1, 1, 3]$$

      ii. To reduce this matrix to row canonical form, use the command

```
A = [0,-2,1,5;2,1,-1,7;-1,-1,1,3])
rref(A)
```

      Demonstrate

4. Most important special case: suppose a system of equations has the same number of equations as unknowns. The coefficient matrix $C$ is square, and augmented matrix $[C|\,\mathbf{b}]$. If $C \longrightarrow I$ then $[C|\,\mathbf{b}] \longrightarrow [I|\,\mathbf{c}]$ for some column $\mathbf{c}$, and the solution to $C\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = \mathbf{c}$.

### 3.7.2 Matrices and Homogeneous Systems of Linear Equations

1. The only important matrix for a homogeneous system is the coefficient matrix..
2. If the coefficient matrix reduces to the identity matrix (possibly followed by 0's), then the system has only the zero solution
   (a) The system

   $$\begin{aligned} 3x + 2y - 5z &= 0 \\ x + y + z &= 0 \\ x - y - 7z &= 0 \end{aligned}$$

   has a coefficient matrix:

   $$\begin{bmatrix} 3 & 2 & -5 \\ 1 & 1 & 1 \\ 1 & -1 & -7 \end{bmatrix}$$

   which reduces to

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

   so the system has only the zero solution.

(b) The system

$$\begin{aligned} 3x + 2y - 5z &= 0 \\ x + y + z &= 0 \\ x - 7z &= 0 \end{aligned}$$

has a coefficient matrix:

$$\begin{bmatrix} 3 & 2 & -5 \\ 1 & 1 & 1 \\ 1 & 0 & -7 \end{bmatrix}$$

which reduces to

$$\begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 8 \\ 0 & 0 & 0 \end{bmatrix}$$

so the system has solutions:

$$\begin{aligned} x &= 7t \\ y &= -8t \\ z &= t \end{aligned}$$

for any value of $t$.

3. We want to develop a more regular way of writing the solutions to a homogeneous system when there are non-zero solutions.

   (a) One approach is to give a *set* of **basic solutions**.

   (b) Assign the value 0 to all the free variables but one, and assign the value 1 to that free variable.

   (c) The resulting solution is a basic solution.

   (d) There is one basic solution for each free variable.

4. In the example above, there is only one free variable, so there is one basic solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7 \\ -8 \\ 1 \end{bmatrix}$$

5. Consider the system with coefficient matrix:

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

   (a) There are three equations and five variables, so we expect three pivot variables and two free variables

   (b) The matrix reduces to:

$$\begin{bmatrix} 1 & 0 & 0 & -\frac{74}{27} & -\frac{7}{3} \\ 0 & 1 & 0 & -\frac{5}{27} & -\frac{1}{3} \\ 0 & 0 & 1 & -\frac{22}{9} & -2 \end{bmatrix}$$

   (c) Therefore the equations would have reduced to

$$\begin{aligned} x_1 \qquad\qquad - \frac{74}{27}x_4 - \frac{7}{3}x_5 &= 0 \\ x_2 \quad - \frac{5}{27}x_4 - \frac{1}{3}x_5 &= 0 \\ x_3 - \frac{22}{9}x_4 - 2x_5 &= 0 \end{aligned}$$

(d) Suppose the variables are $x_1, x_2, x_3, x_4, x_5$. The free variables are $x_4$ and $x_5$. There are two basic solutions.

i. If $x_4 = 1$ and $x_5 = 0$ the basic solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 74/27 \\ 5/27 \\ 22/9 \\ 1 \\ 0 \end{bmatrix}$$

ii. if $x_4 = 0$ and $x_5 = 1$ the basic solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 7/3 \\ 1/3 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

6. Let $\mathbf{x}_1, \ldots, \mathbf{x}_t$ be solutions to a homogeneous system of equations $A\mathbf{x} = \mathbf{0}$. Then any linear combination $\mathbf{y} = a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t$ is also a solution.

(a) Proof:

$$\begin{aligned} A\mathbf{y} &= A\left(a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t\right) \\ &= a_1 A\mathbf{x}_1 + \cdots + a_t A\mathbf{x}_t \\ &= a_1\mathbf{0} + \cdots + a_t\mathbf{0} \\ &= \mathbf{0} \end{aligned}$$

7. Let $\mathbf{x}_1, \ldots, \mathbf{x}_t$ be the basic solutions to $A\mathbf{x} = \mathbf{0}$. Then any solution $\mathbf{y}$ is a unique linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_t$

(a) Here's what we have to prove.

i. If $A\mathbf{y} = \mathbf{0}$ then there exists scalars $a_1, \ldots, a_t$ such that $\mathbf{y} = a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t$

ii. If $a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t = b_1\mathbf{x}_1 + \cdots + b_t\mathbf{x}_t$ then $a_i = b_i$ all $i$.

(b) Here is the key fact. Since there are $t$ basic solutions, there are $t$ free variables. Call the positions $c_1, \ldots, c_t$. Then $\mathbf{x}_j\left[c_i\right] = \delta_{ij}$.

(c) Here's how we prove it.

i. Now for the hard idea. Let $a_i = \mathbf{y}[c_i]$Consider the vector

$$\begin{aligned} \mathbf{z} &= \mathbf{y} - \left(a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t\right) \\ \mathbf{z}[c_i] &= \mathbf{y}[c_i] - \left(a_1\mathbf{x}_1[c_i] + \cdots + a_t\mathbf{x}_t[c_i]\right) \\ &= a_i - \left(a_1\delta_{i1} + \cdots + a_t\delta_{it}\right) \\ &= a_i - a_i \\ &= 0 \end{aligned}$$

That is, $\mathbf{z}$ is a vector that has a 0 in every free variable position. Since $\mathbf{z}$ is a linear combination of solutions, $\mathbf{z}$ is a solution, or $A\mathbf{z} = \mathbf{0}$. But the only solution for which every free variable is 0 is the solution $\mathbf{z} = \mathbf{0}$. Therefore $\mathbf{y} = a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t$

ii. Let $\mathbf{y} = a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t = b_1\mathbf{x}_1 + \cdots + b_t\mathbf{x}_t$. For $1 \leq i \leq t$,

$$\begin{aligned} \mathbf{y}[c_i] &= \sum_{j=1}^{t} a_j\mathbf{x}_j[c_i] \\ &= \sum_{j=1}^{t} a_j\delta_{ij} \\ . &= a_i \end{aligned}$$

Similarly $\mathbf{y}\left[c_i\right] = b_i$, so $a_i = b_i$.

### 3.7.3 The Fundamental Theorem of Linear Algebra

1. **Equation Form:** Suppose you have a homogeneous system of linear equations $A\mathbf{x} = \mathbf{0}$. No matter how you reduce the system, the pivot variables are always the same.

2. **Matrix Form:** Suppose you have a matrix $A$. No matter how you reduce $A$, the pivot variables are always the same.

   (a) This follows immediately from the equation form of the theorem.

3. **Corollary (the biggee):** no matter how you reduce a matrix, you always get the same number of pivot variables (which is the number of non-zero rows in the reduced matrix). This number is called the **rank** of the matrix.

   (a) Example:

$$\begin{aligned} A &= \begin{bmatrix} -3725 & -240 & -9905 & 2650 & -415 \\ -1533 & 924 & -5329 & 678 & 505 \\ 4353 & 1461 & 10\,129 & -3573 & 1265 \\ 3423 & -231 & 9655 & -2253 & 83 \end{bmatrix} \\ RREF\,(A) &= \begin{bmatrix} 1 & 0 & \frac{3699}{1351} & -\frac{926}{1351} & \frac{93}{1351} \\ 0 & 1 & -\frac{4964}{4053} & -\frac{545}{1351} & \frac{2678}{4053} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ rank\,(A) &= 2 \end{aligned}$$

4. **Proof** of Fundamental Theorem (Equation version) Let the variables in the system be $x_1, \ldots, x_n$. A variable $x_i$ is *free* if and only if there are two solutions

$$\begin{gathered} \ldots, a_i, a_{i+1}, \ldots, a_n \\ \ldots, a_i^*, a_{i+1}, \ldots, a_n \end{gathered}$$

   with $a_i \neq a_i^*$.This is just a property of the solution set independent of the way that the system is reduced.

### 3.7.4 Rank and Systems of Equations

1. If $rank\,(A) = rows\,(A)$ then every system of equations $A\mathbf{x} = \mathbf{b}$ can be solved.

   (a) Proof: If $A$ row reduces to a reduced matrix $B$ then $B$ has a pivot in every row and the augmented matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ row reduces to $\begin{bmatrix} B & \mathbf{c} \end{bmatrix}$ for some final column $\mathbf{c}$. This reduced form has a solution because $B$ has no zero rows, so the original equation has a solution.

2. If $rank\,(A) = cols\,(A)$ then every system of equations $A\mathbf{x} = \mathbf{b}$ has no more than one solution.

   (a) If $A$ row reduces to a reduced matrix $B$ then $B$ has a pivot in every column and the augmented matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ row reduces to $\begin{bmatrix} B & \mathbf{c} \end{bmatrix}$ for some final column $\mathbf{c}$. This reduced form has no more than one solution because $B$ has no free variables, so the original equation has no more than one solution

3. If $A$ is square and $rank\,(A) = rows\,(A) = cols\,(A)$ then every system $A\mathbf{x} = \mathbf{b}$ has a unique solution.

   (a) This is the case where $A \longrightarrow I$.

   (b) Actually a stronger result holds. Let $A$ be a $n \times n$ matrix. The following four statements are equivalent. If one is true then all the others are true as well.

i. $A \longrightarrow I$
ii. $rank\,(A) = n$
iii. Every system $A\mathbf{x} = \mathbf{b}$ has a solution for any right-hand-side $\mathbf{b}$.
iv. The homogeneous system $A\mathbf{x} = \mathbf{0}$ has only the solution $\mathbf{x} = \mathbf{0}$.

### 3.7.5 Matrix Equations

1. We've seen how to solve $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ is a column matrix or vector. If $A$ is $n \times m$ then $\mathbf{b}$ has to be $n \times 1$ and $\mathbf{x}$ has to be $m \times 1$. How about $AX = B$ where $A$ is $n \times m$ and $B$ is $n \times t$. The solution $X$ must be $m \times t$, but how to find it.

    (a) Easy: solve $A\,(col_i X) = col_i B$ for $1 \leq i \leq t$.

    (b) If $A$ is square of maximal rank, so $A \longrightarrow I$, then you can solve $AX = B$ by row reducing the augmented matrix $\begin{bmatrix} A & B \end{bmatrix} \longrightarrow \begin{bmatrix} I & C \end{bmatrix}$ and the solution is $X = C$.

## 3.8 Square Matrices

### 3.8.1 Invertible Matrices

1. If $A$ and $B$ are square matrices, and if $AB = I$ then $B$ is called the **inverse** of $A$.

    (a) We say $B = A^{-1}$

    (b) $B$ is also the inverse of $A$ if $BA = I$

    (c) In fact, if $A$ and $B$ are square matrices of the same size, then $AB = I$ if and only if $BA = I$.

    i. This is surprising, since in general $AB \neq BA$ for matrix multiplication
    ii. **Proof:** Suppose $A$ and $B$ are square of the same size, and $AB = I$. We will show that $BA\mathbf{x} = \mathbf{x}$ for all vectors $\mathbf{x}$, which implies that $BA = I$.

    A. The equation $A\mathbf{x} = \mathbf{b}$ can be solved for any right-hand-sides $\mathbf{b}$, by taking $\mathbf{x} = B\mathbf{b}$. Then

$$\begin{aligned} A\mathbf{x} &= A\,(B\mathbf{b}) \\ &= (AB)\,\mathbf{b} \\ &= I\mathbf{b} \\ &= \mathbf{b}. \end{aligned}$$

    B. Therefore if $A\mathbf{x} = \mathbf{0}$ we must have $\mathbf{x} = \mathbf{0}$ by the theorem above.

    C. Moreover, for any vector $\mathbf{x}$

$$\begin{aligned} A\,(BA\mathbf{x} - \mathbf{x}) &= ABA\mathbf{x} - \mathbf{A}\mathbf{x} \\ &= A\mathbf{x} - A\mathbf{x} \\ &= \mathbf{0} \end{aligned}$$

    Therefore, for any vector $\mathbf{x}$, $BA\mathbf{x} - \mathbf{x} = \mathbf{0}$ or $BA\mathbf{x} = \mathbf{x}$.

    (d) If you had a course in logic, you might think I'm only half-way through the proof. I showed that $AB = I$ implies $BA = I$. Don't I have to prove $BA = I$ implies $AB = I$, since I'm trying to prove that $AB = I$ if and only if $BA = I$? **NO.** I proved that if the product of two square matrices is $I$ then the reversed product is also $I$. I don't have to prove that twice.

2. Some square matrices have inverses and some do not

(a) $I \times I = I$ so $I^{-1} = I$

(b) $0 \times B = 0 \neq I$, so 0 cannot have an inverse. 0 is not invertible.

(c) Big questions: how can you tell if a matrix has an inverse or not, and if it does how do you calculate it?

3. **Theorem:** Let $A$ be an $n \times n$ matrix. Then $A$ is invertible if and only if $rank\,(A) = n$.

   (a) **Proof:** For this theorem we do have to go both ways.

   i. Suppose $A$ is invertible. Then there exists a matrix $B$ such that $BA = I$. Suppose $A\mathbf{x} = \mathbf{0}$. We will show that $\mathbf{x} = \mathbf{0}$, which proves that $rank\,(A) = n$.

$$\begin{aligned} A\mathbf{x} &= \mathbf{0} \\ BA\mathbf{x} &= B\mathbf{0} \\ I\mathbf{x} &= \mathbf{0} \\ \mathbf{x} &= \mathbf{0} \end{aligned}$$

   ii. Conversely. if $rank\,(A) = n$ then every equation $A\mathbf{x} = \mathbf{b}$ can be solved. Therefore every matrix equation $AX = B$ can be solved. In particular the equation $AX = I$ can be solved, so $A$ has an inverse.

4. The last proof tells us how to find the inverse of a square matrix $A$. Just solve the equation $AX = I$. Of course $A$ has to satisfy $rank\,(A) = n$, or $A$ has to row reduce to $I$.

   (a) Construct the augmented matrix $\begin{bmatrix} A & I \end{bmatrix}$.

   (b) Row reduce to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$

   (c) Example: $A = \begin{bmatrix} -5 & -2 & -1 \\ -9 & 5 & 8 \\ 4 & -3 & 8 \end{bmatrix}$

$$\left[\begin{array}{ccc|ccc} -5 & -2 & -1 & 1 & 0 & 0 \\ -9 & 5 & 8 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{64}{535} & -\frac{19}{535} & \frac{11}{535} \\ 0 & 1 & 0 & -\frac{104}{535} & \frac{36}{535} & -\frac{49}{535} \\ 0 & 0 & 1 & -\frac{7}{535} & \frac{23}{535} & \frac{43}{535} \end{array}\right]$$

   i. Check:

$$\begin{bmatrix} -5 & -2 & -1 \\ -9 & 5 & 8 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} -\frac{64}{535} & -\frac{19}{535} & \frac{11}{535} \\ -\frac{104}{535} & \frac{36}{535} & -\frac{49}{535} \\ -\frac{7}{535} & \frac{23}{535} & \frac{43}{535} \end{bmatrix}$$

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

5. Class: find inverse of $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ and check.

6. Important special cases:

   (a) **Diagonal matrix:** a square matrix is diagonal if all the off-diagonal elements are 0. Some of the on-diagonal elements can be 0 too.

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

i. An diagonal matrix is invertible if and only if all the diagonal elements are non-zero. In that case

$$D^{-1} = \begin{bmatrix} d_1^{-1} & 0 & \cdots & 0 \\ 0 & d_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^{-1} \end{bmatrix}$$

(b) **Triangular matrix:** an upper (or lower) triangular square matrix is invertible if and only if all the diagonal elements are non-zero. In that case the diagonals of the inverse are the inverses of the diagonals. Here's an upper-triangular example:

$$T = \begin{bmatrix} d_1 & * & \cdots & * \\ 0 & d_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

$$T^{-1} = \begin{bmatrix} d_1^{-1} & * & \cdots & * \\ 0 & d_2^{-1} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^{-1} \end{bmatrix}$$

(c) **Orthogonal matrix:** if $Q$ is orthogonal and square then $Q^TQ = I$ so $Q^{-1} = Q^T$.

i. Square orthogonal matrices are the easiest of all to invert.

ii. Remarkable fact: if $Q$ is square and orthogonal, then the rows as well as the columns are orthonormal

A. Because $Q^T$ is also orthogonal since $\left(Q^T\right)^T Q^T = QQ^T = I$.

7. New inverses from old

(a) If $A$ is invertible, then $\left(A^{-1}\right)^{-1} = A$

i. Proof: If $A$ and $B$ are square matrices of the same size, to show that $B$ is the inverse of a $A$ you must show either $AB = I$ or $BA = I$.

ii. To show that $\left(A^{-1}\right)^{-1} = A$ you must show that $A$ is the inverse of $A^{-1}$, or that $AA^{-1} = I$, (which is pretty obvious).

(b) If $A$ and $B$ are invertible of the same size, then $(AB)^{-1} = B^{-1}A^{-1}$

i. We must show that $B^{-1}A^{-1}$ is the inverse of $AB$, or that $\left(B^{-1}A^{-1}\right)(AB) = I$

$$\begin{aligned} \left(B^{-1}A^{-1}\right)(AB) &= B^{-1}\left(A^{-1}A\right)B \\ &= B^{-1}IB \\ &= B^{-1}B \\ &= I \end{aligned}$$

(c) If $A$ is invertible then $A^T$ is invertible and $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

i. Proof: We must show that $\left(A^{-1}\right)^T$ is the inverse of $A^T$, or that $\left(A^{-1}\right)^T A^T = I$.

$$\begin{aligned} \left(A^{-1}\right)^T A^T &= \left(AA^{-1}\right)^T \\ &= I^T \\ &= I \end{aligned}$$

## 3.9 Everything You Really Need to Know About Determinants

1. Suppose you have four points $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ in space at the corners of a parallelogram.

   (a) Let $\mathbf{v} = \mathbf{b} - \mathbf{a}$ and $\mathbf{w} = \mathbf{d} - \mathbf{a}$, non-parallel sides of the parallelogram

   (b) If $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} c \\ d \end{bmatrix}$ then the area of the parallelogram is $|\mathbf{v} \times \mathbf{w}| = |ad - bc|$.

   (c) We say $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$, the **determinant** of $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$.

   (d) Class find $\begin{vmatrix} 3 & 4 \\ 6 & 5 \end{vmatrix}$

   (e) Class find $\begin{vmatrix} x-3 & 2 \\ -1 & x+1 \end{vmatrix}$

2. In general, any square matrix has a determinant, which the computer knows how to find.

   (a) Using TestGiver

   **input**

```
a = MakeMatrix(RandomInteger(20)-10;j,1,5;k,1,5);
det(a)
```

   **output**

```
a :  MATRIX(INTEGER)
a =
[5, 2, 4, 9, -8;
 -6, 3, -4, -5, -4;
 5, -6, -9, -2, -3;
 -3, -4, -6, -1, -8;
 9, 2, -3, -1, 7]
-9810
```

   (b) *Mathematica*

3. In $\mathbb{R}^n$, $n$ vectors form the corner of a parallelopiped. These vectors also form a square matrix. The absolute value of the determinant of the matrix is the volume of the parallelopiped. The sign of the determinant is determined by the order of the vectors.

(a) In three space, if the vectors are listed in order $\mathbf{u}, \mathbf{v}, \mathbf{w}$, the determinant is positive if the vectors form a right-handed set.

(b) In fact, $\det\begin{bmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{bmatrix} = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$

4. To calculate a $3 \times 3$ determinant, there are three alternatives. To calculate $\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

(a)

$$\begin{aligned} & \left(\begin{bmatrix} a \\ d \\ g \end{bmatrix} \times \begin{bmatrix} b \\ e \\ h \end{bmatrix}\right) \cdot \begin{bmatrix} c \\ f \\ i \end{bmatrix} \\ = & \begin{bmatrix} dh - eg \\ gb - ah \\ ae - db \end{bmatrix} \cdot \begin{bmatrix} c \\ f \\ i \end{bmatrix} \\ = & \; cdh - ceg + fgb - fah + iae - idb \end{aligned}$$

(b)

$$\begin{aligned} & \begin{matrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{matrix} \\ = & \; aei + bfg + cdh - ceg - afh - bdi \end{aligned}$$

(c)

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

(d) ONLY THE THIRD METHOD GENERALIZES TO LARGER MATRICES!!!!

(e) Class calculate

$$\det\begin{bmatrix} 4 & -1 & -4 \\ -1 & -1 & -4 \\ -2 & 3 & 4 \end{bmatrix} = 40$$

$$\det\begin{bmatrix} x+4 & -1 & -4 \\ -1 & x-1 & -4 \\ -2 & 3 & x+4 \end{bmatrix} = x^3 + 7x^2 + 11x + 40$$

5. For triangular matrices, the determinant is easy to calculate:

$$\det\begin{bmatrix} a_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_n \end{bmatrix} = a_1 \cdots a_n$$

6. There are two ways to calculate a general determinant:

(a) The first and best is to use row reduction.

i. This is what a computer does

ii. swapping two rows, the determinant changes by $-1$

iii. mulitplying a row by $c$ multiplies the determinant by $c$

iv. adding a multiple of a row to another does not change the determinant.

v. Example:

$$\begin{aligned}
\det\begin{bmatrix}5&1&4\\3&1&5\\1&2&2\end{bmatrix} &\to (-1)\det\begin{bmatrix}1&2&2\\3&1&5\\5&1&4\end{bmatrix}\\
&\to (-1)\det\begin{bmatrix}1&2&2\\0&-5&-1\\0&-9&-6\end{bmatrix}\\
&\to (-1)\det\begin{bmatrix}1&2&2\\0&-5&-1\\0&0&-\frac{21}{5}\end{bmatrix}\\
\det\begin{bmatrix}5&1&4\\3&1&5\\1&2&2\end{bmatrix} &= (-1)(1)(-5)\left(-\frac{21}{5}\right)\\
&= -21
\end{aligned}$$

vi. Class try with $\det\begin{bmatrix}4&-1&-4\\-1&-1&-4\\-2&3&4\end{bmatrix}$

(b) A method that is sometimes theoretically useful is row and column expansion.

i. Let $A$ be a matrix. The $i,j$ cofactor, $A^{i,j}$, is what you get if you delete row $i$, column $j$, from $A$ and take the determinant of what is left.

ii.

$$\begin{aligned}
\det A &= a_{11}A^{11} - a_{12}A^{12} + \cdots + (-1)^n a_{1n}A^{1n}\\
&= a_{11}A^{11} - a_{21}A^{21} + \cdots + (-1)^n a_{n1}A^{n1}\\
&= \sum_{j=1}^{n} (-1)^{i+j} a_{ij}A^{ij} \text{ for any row } i\\
&= \sum_{i=1}^{n} (-1)^{i+j} a_{ij}A^{ij} \text{ for any column } j
\end{aligned}$$

iii.

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

iv. As the matrices get bigger, row and column expansion becomes very expensive. Row reduction is much more efficient.

v. Class try $\det\begin{bmatrix}4&-1&-4\\-1&-1&-4\\-2&3&4\end{bmatrix} = 40$ going down the third column.

7. Volume properties of determinant

(a) If $A$ is a $n\times n$ matrix, we can think of $A:\mathbb{R}^n \longrightarrow \mathbb{R}^n$. That is, $A$ is a function.

(b) If $\mathcal{R}$ is a region in $\mathbb{R}^n$ with volumn $v$, then $A(\mathcal{R})$ is also a region in $\mathbb{R}^n$, and its volume is $|\det A|\, v$

(c) Example:

i. $n = 2$ and $A = \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}$. $\det(A) = 2$ .

ii. Let $\mathcal{R}$ be the triangle with vertices $(1,2)$, $(3,1)$ and $(2,4)$.

iii. The volume of $\mathcal{R}$ is $v = \frac{1}{2}(2,-1) \times (1,2) = \frac{5}{2}$

iv. $A(\mathcal{R})$ is the triangle with vertices $(5,14)$, $(5,12)$ and $(10,28)$. Its volume is $\frac{1}{2}(0,-2) \times (5,14) = 5 = |\det A|\, v$.

8. Multiplicative properties of determinants of an $n \times n$ matrix $A$.

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

i. This is hard to prove completely, but it is suggested by the volume property of the determinant.

ii. If $A$ and $B$ are $n \times n$ matrices then we have the diagram

$$\begin{array}{ccccc} & & \mathbb{R}^n & & \\ & B \nearrow & & \searrow A & \\ \mathbb{R}^n & & \xrightarrow{AB} & & \mathbb{R}^n \end{array}$$

iii. If $\mathcal{R}$ is a region in $\mathbb{R}^n$ with volume $v(\mathcal{R})$, then $(AB)(\mathcal{R}) = A(B(\mathcal{R}))$.

iv. Thus

$$\begin{aligned} |\det(AB)|\, v(\mathcal{R}) &= v(AB(\mathcal{R})) \\ &= v(A(B(\mathcal{R}))) \\ &= |\det A|\, v(B(\mathcal{R})) \\ &= |\det A|\, |\det B|\, v(\mathcal{R}) \\ |\det(AB)| &= |\det A|\, |\det B| \end{aligned}$$

(b) $\det\left(A^{-1}\right) = (\det A)^{-1}$

i. $\det\left(AA^{-1}\right) = (\det A)\left(\det\left(A^{-1}\right)\right) = \det I = 1$

(c) $\det(cA) = c^n \det A$

i. $\det(cA) = \det((cI)A) = \det(cI)\det A = c^n \det A$

(d) $\det(-A) = (-1)^n \det A$

i. Take $c = -1$

(e) $\det(A + B) \neq \det A + \det B$

(f) Class try: $\det \begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix} \det \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \det\left(\begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}\right)$

9. If $A$ is an invertible matrix then we have seen that $\det A \det\left(A^{-1}\right) = 1$.

(a) Therefore $\det A \neq 0$

(b) The converse is true. If $A$ is a square matrix and $\det A \neq 0$, then $A$ is invertible.

(c) There is even a horrible formula for the inverse: $A^{-1}[i,j] = \dfrac{(-1)^{i+j} A^{j,i}}{\det A}$

i. This is useful for $2 \times 2$ but no larger.

ii. For larger matrices, use row reduction: $\begin{bmatrix} A & I \end{bmatrix} \longrightarrow \begin{bmatrix} I & A^{-1} \end{bmatrix}$

iii. Nevertheless the formula shows that $A$ has an inverse whenever $\det A \neq 0$.

10. The **trace** of a square matrix $A$ is the sum of the diagonal elements of $A$:

$$tr(A) = \sum_{i=1}^{n} A[i,i]$$

(a) $tr\left(\begin{bmatrix} -5 & 8 & 5 \\ -1 & 0 & 3 \\ -4 & -5 & -2 \end{bmatrix}\right) = -7$

(b) $tr\left(A^T A\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} A\left[i,j\right]^2$ so $tr\left(A^T A\right) = 0$ if and only if $A = 0$

i. **Proof:**

$$\begin{aligned} tr\left(A^T A\right) &= \sum_{j=1}^{n} \left(A^T A\right)[j,j] \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} A^T[j,i] A[i,j] \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} A[i,j] A[i,j] \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} A[i,j]^2 \end{aligned}$$

ii. This is an easy computer test to see if a matrix is all zeros

(c) If $A$ and $B$ are $n \times n$ matrices, then $tr(AB) = tr(BA)$.

$$\begin{aligned} tr(AB) &= \sum_{i=1}^{n} (AB)[i,i] \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} A[i,j]\, B[j,i] \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} B[j,i]\, A[i,j] \\ &= \sum_{j=1}^{n} (BA)[j,j] \\ &= tr(BA) \end{aligned}$$

11. If $A$ is a matrix, you can form a new matrix $A - XI$

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

The **determinant** of this matrix is called the **characteristic polynomial** $ch_A(X)$ of the matrix $A$.

(a) Class: evaluate this characteristic polynomial.

12. The characteristic polynomial of an $n \times n$ matrix is always a polynomial $ch_A(X)$ of degree $n$ with leading coefficient $(-1)^n$

$$ch_A(X) = (-1)^n X^n + a_{n-1} X^{n-1} + \cdots + a_0$$

13. $ch_A(0) = a_0 = \det A$ and $a_{n-1} = (-1)^{n-1}\, tr(A)$

(a) Class check these formulas for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

i. find characteristic polynomial of $A$

ii. find determinant of $A$
iii. find trace of $A$
iv. check formulas

14. There are many applications where it is useful to consider the family of matrices $A - XI$ for different values of the variable $X$
    (a) We want to consider $A$, $A - I$, $A + I$, $A - 0.5I$, etc.

15. Which of these matrices are not invertible? For which values of $X$ is $A - XI$ not invertible.
    (a) when $\det(A - XI) = 0$
    (b) when $X$ is a root of the characteristic polynomial of $A$.
    (c) We call these numbers the **eigenvalues** or **characteristic values** of $A$.

16. **Fundamental Theorem of Algebra**: every polynomial of degree $n$ can be factored uniquely (up to order) into $n$ linear factors over the complex numbers.
    (a) Gauss, 1799 (his Ph.D thesis).
    (b) so for an $n \times n$ matrix $A$ we have

$$ch_A(X) = (-1)^n (X - e_1) \cdots (X - e_n)$$

    where $e_1, \ldots, e_n$ are complex numbers that are the roots of $ch_A(X)$.
        i. If $A$ is real, then the coefficients of $ch_A(X)$ are real and the complex roots appear in conjugate pairs
        ii. The roots may be repeated.
            A. The **algebraic multiplicity** of an eigenvalue is the number of times it is repeated in the factorization of the characteristic polynomial.
            B. the sum of the algebraic multiplicities of all the different eigenvalues is $n$, the size of the matrix.
            C. So the largest number of different eigenvalues possible is $n$
    (c) The product of the eigenvalues is $\det A$ and the of the eigenvalues is $tr(A)$.
        i. Because the constant term of the factored polynomial is

$$\begin{aligned} \det A &= (-1)^n (-e_1) \cdots (-e_n) \\ &= e_1 \cdots e_n \end{aligned}$$

        and the coefficient of $X^{n-1}$ is

$$\begin{aligned} (-1)^{n-1} tr(A) &= (-1)^n (-e_1 - \cdots - e_n) \\ &= (-1)^{n-1} (e_1 + \cdots + e_n) \end{aligned}$$

17. Example: $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$
    (a) Characteristic polynomial: $ch_A(X) = (1 - X)(3 - X) - 4 = X^2 - 4X - 1$
    (b) By the quadratic formula, the roots are $X = \frac{4 \pm \sqrt{16+4}}{2} = 2 \pm \sqrt{5}$
        i. These are the eigenvalues of $A$
    (c) The factorization is $ch_A(X) = \left(X - \left(2 + \sqrt{5}\right)\right)\left(X - \left(2 - \sqrt{5}\right)\right)$ so each eigenvalue has algebraic multiplicity 1.
        i. The sum of the algebraic multiplicities is 2, the size of the matrix.
    (d) The product of the eigenvalues is $-1 = \det A$ and the sum is $4 = tr(A)$

18. Example: $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$

   (a) Characteristic polynomial: $ch_A(X) = (3 - X)(3 - X) - 0 = X^2 - 6X + 9$

   (b) By the quadratic formula, the roots are $X = \frac{6 \pm \sqrt{36-36}}{2} = 3 \pm 0$

      i. These are the eigenvalues of $A$

   (c) The factorization is $ch_A(X) = (X - 3)(X - 3)$ so the unique eigenvalue 3 has multiplicity 2

      i. The sum of the algebraic multiplicities is 2, the size of the matrix.

   (d) The product of the eigenvalues is $9 = \det A$ and the sum is $6 = tr(A)$

19. Example: $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

   (a) Characteristic polynomial: $ch_A(X) = (1 - X)(1 - X) + 1 = X^2 - 2X + 2$

   (b) By the quadratic formula, the roots are $X = \frac{2 \pm \sqrt{4-8}}{2} = 1 \pm i$

      i. These are the eigenvalues of $A$

      ii. they are a conjugate pair

   (c) The factorization is $ch_A(X) = (X - (1 + i))(X - (1 - i))$ so each eigenvalue has algebraic multiplicity 1.

      i. The sum of the algebraic multiplicities is 2, the size of the matrix.

   (d) The product of the eigenvalues is $(1 - i)(1 + i) = 2 = \det A$ and the sum is $(1 - i) + (1 + i) = 2 = tr(A)$

20. Class find eigenvalues of $\begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$ and multiplicities. Check the product and sum of eigenvalues.

21. Finding eigenvalues of large matrices is a major industry within numerical analysis. It is of great scientific, commercial and even military importance in solving problems about transportation networks and fluid and gas flows. This was one of the problems that gave rise to the early deveopment and funding of computers

22. There are some easy special cases.

   (a) If a matrix is triangular, then the eigenvalues are the diagonal elements.

      i. If the diagonal elements of a triangular matrix $T$ are $t_1, \ldots, t_n$ then $ch_T(X) = (-1)^n (X - t_1) \cdots (X - t_n)$.

   (b) This applies to diagonal matrices especially

   (c) Example: $A = \begin{bmatrix} -5 & 8 & 4 \\ 0 & 3 & -5 \\ 0 & 0 & -5 \end{bmatrix}$.

      i. $ch_A(X) = (-5 - X)(3 - X)(-5 - X)$.

      ii. The eigenvalues are $-5$ with algebraic multiplicity 2 and 3 with algebraic multiplicity 1.

   (d) In this case it is obvious that the sum of the eigenvalues is the trace and the product of the eigenvalues is the determinant.

## 3.10 Eigenvectors and Eigenvalues

### 3.10.1 Introduction

1. One of the questions we have pursued is finding solutions to equations $A\mathbf{x} = \mathbf{b}$.
2. A superficially similar questions might be this: given a square matrix $A$, find a non-zero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar $\lambda$.
    (a) Another way to ask the problem is to ask for a non-zero vector $\mathbf{x}$ such that $A\mathbf{x}$ is parallel to $\mathbf{x}$.
3. As we will see, this problem is much harder than solving linear equations.
    (a) It arises frequently in applications of linear algebra to statistics, engineering, differential equations, and other fields.
4. Fortunately, we have done all the hard work in the previous chapter. Let's see why.
5. $A\mathbf{x} = \lambda\mathbf{x}$ if and only if $(A - \lambda I)\,\mathbf{x} = \mathbf{0}$.
    (a) Non-zero solutions $\mathbf{x}$ exist if and only if $A - \lambda I$ is is not invertible, or singular.
    (b) So we will find non-zero solutions $\mathbf{x}$ if and only if $\lambda$ is an eigenvalue of $A$.
    (c) In that case, putting in the eigenvalue, it is easy to solve for non-zero vectors $\mathbf{x}$.
6. The vectors $\mathbf{x}$ are called *eigenvectors* associated with the eigenvalues $\lambda$.
    (a) Eigenvectors and eigenvalues come in pairs.
    (b) If somehow you know an eigenvector, it is easy to find the eigenvalue.
7. Class check:: $A = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$,
    (a) $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\lambda = 1$
    (b) $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\lambda = 0$
8. If you know the eigenvalues of a matrix, for example by finding the roots of the characteristic polynomial, you can find the eigenvectors by solving $(A - \lambda I)\,\mathbf{x} = \mathbf{0}$ , a problem we know how to solve.
    (a) You have to do this for each different eigenvalue $\lambda$. The solutions $\mathbf{x}$ are the eigenvectors associated with $\lambda$.
    (b) Generally this equation will not have a unique solution, but it will have a line of solutions, and you will take any representative one.
    (c) The number of "really different" eigenvectors associated with an eigenvalue is the number of free variables in the equation $(A - \lambda I)\,\mathbf{x} = \mathbf{0}$.
    (d) You can take the basic solutions of this system as the different eigenvectors associated with eigenvalue $\lambda$.
9. There are **two** kinds of problems you will be asked about eigenvectors
    (a) Prove a theorem about eigenvectors.
    (b) Given a matrix, find all the eigenvalues and eigenvectors.

The methods used in these two cases may be very different.

10. Example of proving a theorem: if $\mathbf{x}$ is an eigenvector of a matrix $A$ with eigenvalue $\lambda$, then $\mathbf{x}$ is an eigenvector of $A^2$ with eigenvalue $\lambda^2$.

    (a) We know that $A\mathbf{x} = \lambda\mathbf{x}$. Then

    $$\begin{aligned} A^2\mathbf{x} &= A(A\mathbf{x}) \\ &= A(\lambda\mathbf{x}) \\ &= \lambda(A\mathbf{x}) \\ &= \lambda(\lambda\mathbf{x}) \\ &= \lambda^2\mathbf{x} \end{aligned}$$

    so $\mathbf{x}$ is an eigenvector of $A^2$ with eigenvalue $\lambda^2$.

    (b) **Class**: prove that if $\mathbf{x}$ is an eigenvector of an invertible matrix $A$ with eigenvalue $\lambda$, then $\mathbf{x}$ is an eigenvector of $A^{-1}$ with eigenvalue $\lambda^{-1}$.

    (c) **Class**: prove that $A$ is not invertible if and only if 0 is an eigenvalue of $A$.

### 3.10.2 Calculating Eigenvectors and Eigenvalues

1. Using *TestGiver* to find the eigenvectors and eigenvalues for the matrix $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$.

   **Input**

   A=[1,3;2,2];

   v,e1 = Eig(A)

   A*v-v*e1

   **Output**

   A : MATRIX(INTEGER)

   A =

   [1, 3;

   2, 2]

   v : MATRIX(COMPLEX)

   e1 : MATRIX(COMPLEX)

   v,e1 =

   [0.83205, 0.707107;

   -0.5547, 0.707107],

   [-1, 0;

   0, 4]

   [-5.55112E-16, 4.44089E-16;

   -1.11022E-16, -4.44089E-16]

   (a) Eigenvectors are columns of `v` , eigenvalues are diagonals of `e1`.

   (b) Eigenvectors are normalized to length 1

   (c) Matrix*eigenvectors = eigenvectors*eigenvalues—see section on Diagonalization

2. Using *Mathematica* to find the eigenvectors and eigenvalues for the matrix $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$.

```
In[15]:= a = {{1., 3.}, {2., 2.}}; (* construct matrix *)
         {v, e} = Eigensystem[a];
         v
         e

Out[17]= {4., -1.}

Out[18]= {{-0.707107, -0.707107}, {-0.83205, 0.5547}}
```

Check

```
In[13]:= a.e[[1]] - v[[1]] e[[1]]
         a.e[[2]] - v[[2]] e[[2]]

Out[13]= {0, 0}

Out[14]= {0, 0}
```

(a) Eigenvalues returned as a list or vector of values

(b) Eigenvectors returned as rows of a matrix

i. Later we will see why *Matlab* and other serious matrix programs (including *TestGiver*) return the eigenvectors as columns.

(c) *Mathematica* tries to hide user from round-off "errors", which can be a problem if the small number that *Mathematica*. set to zero was really supposed to be non-zero.

3. Doing it by hand

(a) Note that if $A\mathbf{x} = \lambda\mathbf{x}$ then $(A - \lambda I)\,\mathbf{x} = \mathbf{0}$. Since $\mathbf{x} \neq \mathbf{0}$, $A - \lambda I$ is a matrix with a non-trivial kernel, so $\det(A - \lambda I) = 0$.

i. Remember, $\det(A - XI)$ is the characteristic polynomial of $A$. Eigenvalues $\lambda$ are roots of the characteristic polynomial.

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

So the eigenvalues of $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ are $\lambda = 4$ and $\lambda = -1$. The corresponding eigenvectors can be easily found:

i. If the eigenvalue is $\lambda = 4$ then $A - \lambda I = \begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix}$ and we can take the eigenvector $\mathbf{x} \in \ker(A - \lambda I)$ to be $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Other choices would be $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$ or $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. The last one may be best because it is a unit vector.

ii. If the eigenvalue is $\lambda = -1$ then $A - \lambda I = \begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix}$ and we can take the eigenvector $\mathbf{x} \in \ker(A - \lambda I)$ to be $\mathbf{x} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. Other choices would be $\begin{bmatrix} -3 \\ 2 \end{bmatrix}$ or $\begin{bmatrix} 3/\sqrt{13} \\ -2/\sqrt{13} \end{bmatrix}$, a unit eigenvector.

(b) Class find eigenvalues and eigenvectors for $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.

i. Find the characteristic polynomial
ii. Find the roots, which are the eigenvalues
iii. Find an eigenvector for each eigenvalue

4. Upper triangular matrices
   (a) The characteristic polynomial of $\begin{bmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & c \end{bmatrix}$ is $(a - X)(b - X)(c - X)$ so the eigenvalues are $a, b, c$
   (b) The eigenvectors must still be solved for, but the eigenvector associated with eigenvalue $a$ is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.
   (c) Class find eigenvalues and eigenvectors for $\begin{bmatrix} 2 & 0 & 1 \\ 0 & -1 & 3 \\ 0 & 0 & 4 \end{bmatrix}$

5. Let $A$ be an $n \times n$ matrix. The polynomial $\det(A - XI)$ with variable $X$ looks like $(-1) X^n +$ lower degree terms.
   (a) This polynomial is called the *characteristic polynomial* of $A$.
   (b) It has $n$ roots if you count repetitions. They are the eigenvalues of $A$.
      i. **Therefore an $n \times n$ has at most $n$ different eigenvalues.**
   (c) The product of the eigenvalues will always be the determinant of the matrix
   (d) The sum of the eigenvalues is the *trace* of the matrix, the sum of the diagonal elements.
   (e) In the characteristic polynomial

$$\det(A - XI) = (-1)^n X^n + a_{n-1}X^{n-1} + \cdots + a_0$$

$$\begin{aligned} a_{n-1} &= (-1)^{n-1} \mathit{trace}(A) \\ a_0 &= \det A \end{aligned}$$

   (f) **The computer does not use the characteristic polynomial to find the eigenvalues and eigenvectors. It uses an entirely different method based on upper-triangular matrices that is much faster and more accurate for large matrices (but not convenient for small matrices or paper-and-pencil calculations).**

6. Let $A$ be an $n \times n$ matrix. Then
   (a) A non-zero vector $\mathbf{x} \in \mathbb{R}^n$ and a scalar $\lambda$ form an eigenvector-eigenvalue pair if and only if $A\mathbf{x} = \lambda\mathbf{x}$.
   (b) $A$ has a characteristic polynomial $(-1)^n X^n + a_{n-1}X^{n-1} + \cdots + a_0$ whose roots are the eigenvalues of $A$, so $A$ has at most $n$ distinct eigenvalues
   (c) In your next liinear algebra course you will spend some time the following. Let $A$ be an $n \times n$ matrix.
      i. the *geometric multiplicity* of an eigenvalue $\lambda$ is the number of basic solutions of $(A - \lambda I)\mathbf{x} = \mathbf{0}$ associated with $\lambda$.
         A. most of the time there is basic solution associated with an eigenvalue, so the geometric multiplicity of most eigenvalues is 1
      ii. the *algebraic multiplicity* of an eigenvalue $\lambda$ is the number of times the eigenvalue is repeated as a root of the characteristic polynomial of $A$
         A. most of the time the roots of the characteristic polynomial are not repeated, so the algebraic multiplicity of most eigenvalues is 1.

iii. The geometric multiplicity of an eigenvalue is less than or equal to its algebraic multiplicity.

A. which is certainly the case most of the time, when both values are 1.

7. Two examples:

(a) $A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$

i. Characteristic polynomial det $\begin{bmatrix} 3-\lambda & 0 \\ 0 & 3-\lambda \end{bmatrix} = (3-\lambda)^2$

ii. Eigenvalue $\lambda = 3$ algebraic multiplicity 2

iii. Eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ geometric multiplicity of eigenvalue $\lambda$ =3 is 2

(b) $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$

i. Characteristic polynomial det $\begin{bmatrix} 3-\lambda & 1 \\ 0 & 3-\lambda \end{bmatrix} = (3-\lambda)^2$

ii. Eigenvalue $\lambda = 3$ algebraic multiplicity 2

iii. Eigenvector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ geometric multiplicity of eigenvalue $\lambda$ =3 is 1

8. Class: find characteristic polynomial, eigenvalues and their algebraic and geometric multiplicity for the matrix::

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

### 3.10.3 Diagonalization

1. Let $A$ be $2 \times 2$ with eigenvectors $\mathbf{u}, \mathbf{v}$ and eigenvalue $\lambda, \mu$. Then

$$\begin{aligned} A\mathbf{u} &= \lambda\mathbf{u} \\ A\mathbf{v} &= \mu\mathbf{v} \\ A\begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix} &= \begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix}\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix} \end{aligned}$$

2. Example: $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ has eigenvalues $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\lambda_1 = 4$ and $\mathbf{x}_2 = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$, $\lambda = -1$. Therefore we have:

$$\begin{aligned} U &= \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \\ D &= \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix} \end{aligned}$$

(a) Check:

$$\begin{aligned} U^{-1} &= \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ \frac{1}{5} & -\frac{1}{5} \end{bmatrix} \\ UDU^{-1} &= \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ \frac{1}{5} & -\frac{1}{5} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \\ &= A \end{aligned}$$

3. Class: find $U$ and $D$ for $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$

4. More generally, if an $n \times n$ matrix $A$ has eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_t$ and eigenvalues $\lambda_1, \ldots, \lambda_t$ then

$$A \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_t \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_t \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_t \end{bmatrix}$$

This is true even if $t$ is not the size of the matrix. For example, if $t = 1$ it says that $A\mathbf{x}_1 = \mathbf{x}_1 [\lambda_1]$.

(a) However, if $A$ has $n$ distiinct eigenvalues $\lambda_1, \ldots, \lambda_n$, then (unproven result) the corresponding eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent, and if we define:

$$\begin{aligned} U &= \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \\ D &= \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \end{aligned}$$

then $U$ is an invertible matrix and we have:

$$\begin{aligned} AU &= UD \\ U^{-1}AU &= D \\ A &= UDU^{-1} \end{aligned}$$

the *diagonalization of* $A$

(b) This can be done for almost all matrices but not all, in the same way that almost all square matrices are invertible but not all.

i. The reason is that the eigenvalues are the roots of the characteristic polynomial, and most polynomials do not have repeated roots.

5. This factorization is useful for a number of operations. For example it allows the rapid computation of powers of a matrix:

$$\begin{aligned} D^n &= \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_t^n \end{bmatrix} \\ A^n &= \left(UDU^{-1}\right)^n \\ &= UDU^{-1}UDU^{-1} \cdots UDU^{-1} \\ &= UD^nU^{-1} \end{aligned}$$

(a) Example using $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ from above.

$$\begin{aligned} A^{100} &= UD^{100}U^{-1} \\ &= \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & (-1) \end{bmatrix}^{100} \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ \frac{1}{5} & -\frac{1}{5} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 4^{100} & 0 \\ 0 & (-1)^{100} \end{bmatrix} \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ \frac{1}{5} & -\frac{1}{5} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1.6069 \times 10^{60} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ \frac{1}{5} & -\frac{1}{5} \end{bmatrix} \\ &= \begin{bmatrix} 6.4276 \times 10^{59} & 9.6414 \times 10^{59} \\ 6.4276 \times 10^{59} & 9.6414 \times 10^{59} \end{bmatrix} \end{aligned}$$

i. This looks like a matrix with two equal rows, so it cannot be invertible. But $A$ is invertible so $A^{100}$ is invertible.

A. That just means that $A^{100}$ is **nearly singular**—it rounds off to a singular matrix unless you keep an awful lot of decimal places.

B. Most invertible matrices have high powers that are nearly singular.

6. Class: find $\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}^{20}$

### 3.10.4 Symmetric Matrices

1. Let $A$ be an $n \times n$ symmetric matrix. Then $A$ has **real** eigenvalues and an **orthonormal** set of $n$ eigenvectors.

(a) Very important theorem. Simplifies eigenvalue analysis in many cases.

(b) called **spectral theorem** or **principal axis theorem**

(c) Typical matrix is the Hessian. If $f(x_1, \ldots, x_n)$ is a function we have matrix $\left[\dfrac{\partial^2 f}{\partial x_i x_j}\right]$ that is symmetric.

(d) $A = UDU^{-1} = UDU^T$

(e) Try it: $A = \begin{bmatrix} -3. & -3 & -1 & 1 \\ -3 & 4 & 0 & 5 \\ -1 & 0 & 5 & -2 \\ 1 & 5 & -2 & 3 \end{bmatrix}$, eigenvectors: $\left\{\begin{bmatrix} .248\,79 \\ -.384\,33 \\ -.887\,98 \\ 4.336\,6 \times 10^{-2} \end{bmatrix}\right\} \leftrightarrow$ $5.377\,8$, $\left\{\begin{bmatrix} -.538\,36 \\ .392\,88 \\ -.352\,95 \\ -.656\,69 \end{bmatrix}\right\} \leftrightarrow -.246\,49$, $\left\{\begin{bmatrix} -9.391\,4 \times 10^{-2} \\ .693\,69 \\ -.294\,78 \\ .650\,44 \end{bmatrix}\right\} \leftrightarrow 9.$ $094\,4$, $\left\{\begin{bmatrix} .799\,66 \\ .465\,54 \\ 4.032\,4 \times 10^{-3} \\ -.379\,21 \end{bmatrix}\right\} \leftrightarrow -5.225\,8$,

$$\begin{bmatrix} .248\,79 & -.384\,33 & -.887\,98 & 4.336\,6 \times 10^{-2} \\ -.538\,36 & .392\,88 & -.352\,95 & -.656\,69 \\ -9.391\,4 \times 10^{-2} & .693\,69 & -.294\,78 & .650\,44 \\ .799\,66 & .465\,54 & 4.032\,4 \times 10^{-3} & -.379\,21 \end{bmatrix} \times$$

$$\begin{bmatrix} .248\,79 & -.538\,36 & -9.391\,4 \times 10^{-2} & .799\,66 \\ -.384\,33 & .392\,88 & .693\,69 & .465\,54 \\ -.887\,98 & -.352\,95 & -.294\,78 & 4.032\,4 \times 10^{-3} \\ 4.336\,6 \times 10^{-2} & -.656\,69 & .650\,44 & -.379\,21 \end{bmatrix}$$

$$= \begin{bmatrix} 1.0 & 3.676\,6 \times 10^{-7} & -5.016\,4 \times 10^{-6} & 9.117\,9 \times 10^{-7} \\ 3.676\,6 \times 10^{-7} & 1.0 & 1.625\,6 \times 10^{-6} & -3.423\,1 \times 10^{-6} \\ -5.016\,4 \times 10^{-6} & 1.625\,6 \times 10^{-6} & .999\,99 & -8.499 \times 10^{-7} \\ 9.117\,9 \times 10^{-7} & -3.423\,1 \times 10^{-6} & -8.499 \times 10^{-7} & 1.0 \end{bmatrix}$$

(f) Class try: pick a $2 \times 2$ symmetric matrix and find eigenvalues and eigenvectors. Show that eigenvectors are orthogonal, and make them orthonormal. Show $A = UDU^T$

(g) Class try $\begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

i. This example shows that you can get a full set of eigenvectors even without a distinct eigenvalues.

2. Proof assuming $n$ distinct eigenvalues

(a) First we show that the eigenvalues of a symmetric matrix are real. If $A\mathbf{x} = \lambda\mathbf{x}$ then $\bar{A}\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}} = A\bar{\mathbf{x}}$. Thus $\bar{\mathbf{x}}^T A^T = \mathbf{x}^T\bar{\lambda} = \bar{\mathbf{x}}^T A$. Therefore $\bar{\mathbf{x}}^T A\mathbf{x} = \bar{\mathbf{x}}^T\lambda\mathbf{x} = \bar{\mathbf{x}}^T\bar{\lambda}\mathbf{x}$. or $\lambda\left|\mathbf{x}\right|^2 = \bar{\lambda}\left|\mathbf{x}\right|^2$. Since $\mathbf{x} \neq \mathbf{0}$, $\lambda = \bar{\lambda}$.

(b) Next we show that the eigenvectors with distinct eigenvalues are orthogonal. (You can always make them length 1). If $A\mathbf{x} = \lambda\mathbf{x}$ and $A\mathbf{y} = \mu\mathbf{y}$ and $\lambda \neq \mu$ then

$$\begin{aligned}\lambda\mathbf{x}^T\mathbf{y} &= (\lambda\mathbf{x})^T\mathbf{y} \\ &= (A\mathbf{x})^T\mathbf{y} \\ &= \mathbf{x}^T A^T\mathbf{y} \\ &= \mathbf{x}^T A\mathbf{y} \\ &= \mathbf{x}^T\mu\mathbf{y} \\ &= \mu\left(\mathbf{x}^T\mathbf{y}\right)\end{aligned}$$

Since $\mu \neq \lambda$, $\mathbf{x}^T\mathbf{y} = 0$.

(c) Finally, if $A$ has $n$ distinct eigenvalues (which we have shown must be real), then it has $n$ eigenvectors (which we have shown must be orthonormal).

### 3.10.5 Positive Definite Matrices

1. **Definition:** a matrix is called *positive definite* if it is symmetric and all its eigenvalues are positive.

   (a) A diagonal matrix is positive definite iff the diagonal elements are all positive

   (b) A diagonal matrix represents an expansion of space along the eigenvectors, by differing multiplication factors.(which can be $< 1$, so we might get a contraction, but no reversal of axes)

2. An $n \times n$ symmetric matrix $A$ is positive definite iff $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.

   (a) **Proof**: Suppose $A$ is positive definite. Let $A = UDU^T$ with $D = \begin{bmatrix}\lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n\end{bmatrix}$ and $\lambda_i > 0$ all $i$. Then $\mathbf{x}^T UDU^T\mathbf{x} = \left(U^T\mathbf{x}\right)^T D\left(U^T\mathbf{x}\right)$. Suppose $\mathbf{x} \neq \mathbf{0}$ and $U^T\mathbf{x} = \begin{bmatrix}a_1 \\ \vdots \\ a_n\end{bmatrix}$. At least one $a_i \neq 0$ since $U$ and therefore $U^T$ is invertible, and $\left(U^T\mathbf{x}\right)^T D\left(U^T\mathbf{x}\right) = \lambda_1 a_1^2 + \cdots + \lambda_n a_n^2 > 0$. since all $\lambda_i > 0$ and all $a_i^2 \geq 0$ and at least one $a_i > 0$.
   Conversely, if $A$ is symmetric and $\mathbf{x}^T A\mathbf{x} > \mathbf{0}$ for all $\mathbf{x} \neq \mathbf{0}$, suppose $\lambda$ is an eigenvalue of $A$ with eigenvector $\mathbf{x}$. We must show $\lambda > 0$. But $\mathbf{x} \neq \mathbf{0}$ so $\mathbf{x}^T A\mathbf{x} > \mathbf{0}$, and thus

$$\begin{aligned}0 &< \mathbf{x}^T A\mathbf{x} \\ &= \mathbf{x}^T\lambda\mathbf{x} \\ &= \lambda\left|\mathbf{x}\right|^2\end{aligned}$$

   so $\lambda > 0$.

   (b) In some sense a positive definite matrix $A$ is a new ruler on the space $\mathbb{R}^n$. Instead of definiting the length of $\mathbf{x}$ as $\sqrt{\mathbf{x}^T\mathbf{x}}$, we can define the length of $\mathbf{x}$ as $\sqrt{\mathbf{x}^T A\mathbf{x}}$.

3. A matrix is positive definite if and only if the upper determinants are all postive.

4. Class test $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$

   (a) The eigenvalues are : $-.515\,73, .170\,92, 11.345$.

5. Focus on $2x2$ matrices.

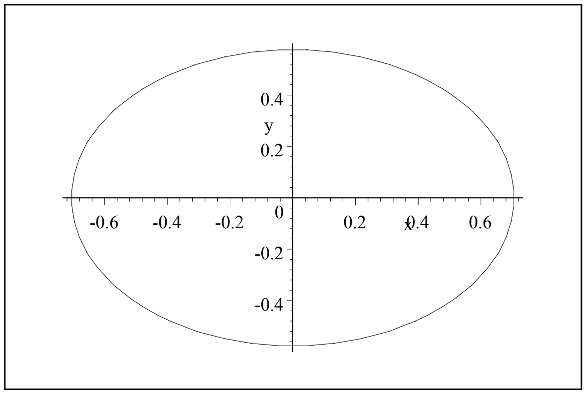
   (a) $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is positive definite iff $a > 0$ and $ac - b^2 > 0$.

### 3.10.6 Quadratic Forms

1. Problem: graph $2x^2 + 3y^2 = 1$

   (a) This is an ellipse

$$\begin{aligned}\left(\frac{x}{1/\sqrt{2}}\right)^2 + \left(\frac{y}{1/\sqrt{3}}\right)^2 &= 1\\ \left(\frac{x}{0.707}\right)^2 + \left(\frac{y}{0.577}\right)^2 &= 1\end{aligned}$$




$2x^2 + 3y^2 = 1$

2. How about

$$2\left(0.447x - 0.894y\right)^2 + 3\left(0.894x + 0.447y\right)^2 = 1$$

   (a) If we define

$$\begin{aligned}u &= 0.447x - 0.894y\\ v &= 0.894x + 0.447y\end{aligned}$$

   The equation becomes:

$$2u^2 + 3v^2 = 1$$

   (b) The expressions $0.447x - 0.894y$ and $0.894x + 0.447y$ are carefully chosen to be "orthogonal"

      i. That is if we restate the definition of $u$ and $v$:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0.447 & -0.894 \\ 0.894 & 0.447 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

      then $\begin{bmatrix} 0.447 & -0.894 \\ 0.894 & 0.447 \end{bmatrix}$ is an *orthogonal* matrix. Check:

$$\begin{aligned}&\begin{bmatrix} 0.447 & -0.894 \\ 0.894 & 0.447 \end{bmatrix}^T \begin{bmatrix} 0.447 & -0.894 \\ 0.894 & 0.447 \end{bmatrix}\\ =&\begin{bmatrix} 0.999\,05 & 0 \\ 0 & 0.999\,05 \end{bmatrix}\end{aligned}$$

      ii. Define

$$\begin{aligned}U &= \begin{bmatrix} 0.447 & -0.894 \\ 0.894 & 0.447 \end{bmatrix}^{-1}\\ &= \begin{bmatrix} 0.447 & -0.894 \\ 0.894 & 0.447 \end{bmatrix}^T\\ &= \begin{bmatrix} 0.447 & 0.894 \\ -0.894 & 0.447 \end{bmatrix}\end{aligned}$$

Then

$$\begin{aligned} \begin{bmatrix} u \\ v \end{bmatrix} &= U^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \\ \begin{bmatrix} x \\ y \end{bmatrix} &= U \begin{bmatrix} u \\ v \end{bmatrix} \end{aligned}$$

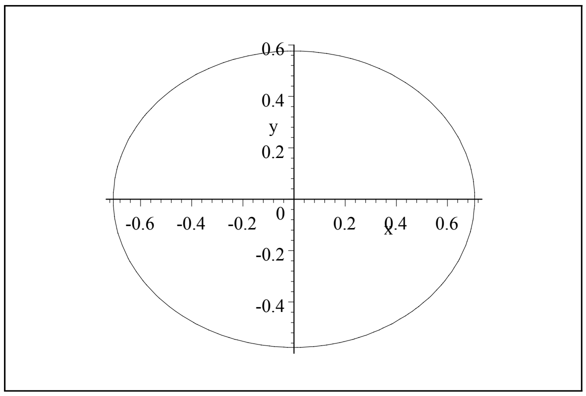
(c) Now is when it pays to think of the matrix $U$ as a map $U : \mathbb{R}^2_{(u,v)} \longrightarrow \mathbb{R}^2_{(x,y)}$

i. The big idea is this

A. The map $U$ pushes points in the $(u, v)$ plane back to points in the $(x, y)$ plane.

B. The map $U$ pushes the ellipse

$$\begin{aligned} 2u^2 + 3v^2 &= 1 \\ \left(\frac{u}{1/\sqrt{2}}\right)^2 + \left(\frac{v}{1/\sqrt{3}}\right)^2 &= 1 \end{aligned}$$




$2u^2 + 3v^2 = 1$

to the solution of

$$2\,(0.447x - 0.894y)^2 + 3\,(0.894x + 0.447y)^2 = 1$$

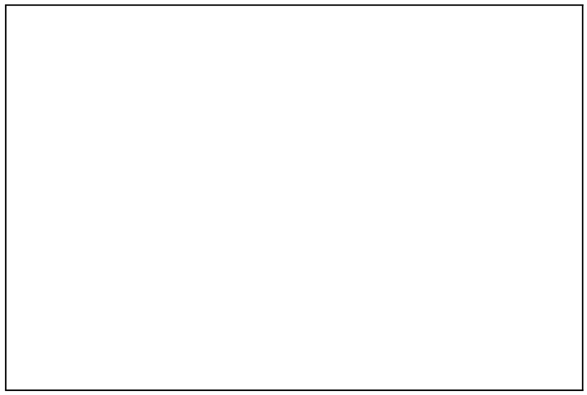
C. The solution of $2\,(0.447x - 0.894y)^2 + 3\,(0.894x + 0.447y)^2 = 1$ will be a rotated ellipse with vertices at:

$$\begin{aligned} U \begin{bmatrix} \pm 1/\sqrt{2} \\ 0 \end{bmatrix} &= \pm \begin{bmatrix} 0.316 \\ -0.632 \end{bmatrix} \\ U \begin{bmatrix} 0 \\ \pm 1/\sqrt{3} \end{bmatrix} &= \pm \begin{bmatrix} 0.516 \\ 0.258 \end{bmatrix} \end{aligned}$$

Notice that the vertices are orthogonal but not orthonormal. In fact the lengths of the semiaxes are:

$$\begin{aligned} \sqrt{0.316^2 + 0.632^2} &= 0.707 \\ &= 1/\sqrt{2} \\ \sqrt{0.516^2 + 0.258^2} &= 0.577 \\ &= 1/\sqrt{3} \end{aligned}$$

so the $xy$-ellipse is the $uv$-ellipse rotated.

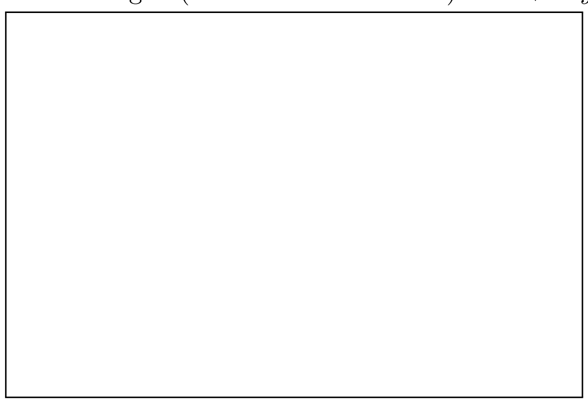


$2\left(0.447x-0.894y\right)^2+3\left(0.894x+0.447y\right)^2=1$

(d) If we expand algebraically, we get:

$$\begin{aligned} & 2\left(0.447x-0.894y\right)^2+3\left(0.894x+0.447y\right)^2 \\ = \; & 2.797\,3x^2+0.799\,24xy+2.197\,9y^2 \end{aligned}$$

so we have graphed $2.797\,3x^2+0.799\,24xy+2.197\,9y^2=1$ or, rounded off to three-digits (which we started with) $28x^2+8xy+22y^2=10$



$28x^2+8xy+22y^2=10$

3. Suppose we had started with $28x^2+8xy+22y^2=10$.

   (a) How could we graph this equation and find the vertices?

4. First some general ideas.

   (a) The equation $28x^2+8xy+22y^2=10$ can be recast as

$$\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 28 & 4 \\ 4 & 22 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=10$$

   (b) General theory: any **quadratic form** $ax^2+2bxy+cy^2$ can be written

$$\begin{aligned} & ax^2+2bxy+cy^2 \\ = \; & \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \\ = \; & \mathbf{x}^T A\mathbf{x} \end{aligned}$$

   for some symmetric matrix $A$.

   (c) If $A$ is positive definite then $f(x,y)>0$ for all $(x,y)\neq(0,0)$

      i. so $f(x,y)$ has a (unique) minimum at $(x,y)=(0,0)$ and the graphs of $f(x,y)=c$ are ellipses

   (d) Class: write $3x^2+6xy+2y^2$ in matrix form and determine if the coefficient matrix is positive definite.

(e) Class: is our example $A = \begin{bmatrix} 28 & 4 \\ 4 & 22 \end{bmatrix}$ positive definite?

5. If you have a positive definite $A$, how can you find the graph of $\mathbf{x}^T A\mathbf{x} = c$?

   (a) Write $A = UDU^T$. where $U$ contains orthonormal eigenvectors and $D$ contains the eigenvectors.

   (b) In our example $A = \begin{bmatrix} 28 & 4 \\ 4 & 22 \end{bmatrix}$ and $c = 10$.

$$\begin{aligned} U &= \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 0.894 & 0.447 \\ 0.447 & -0.894 \end{bmatrix} \\ D &= \begin{bmatrix} 30 & 0 \\ 0 & 20 \end{bmatrix} \end{aligned}$$

   (c) If we define

$$\begin{bmatrix} u \\ v \end{bmatrix} = U^T \begin{bmatrix} x \\ y \end{bmatrix}$$

   then we get two consequences:

$$\begin{aligned} U \begin{bmatrix} u \\ v \end{bmatrix} &= \begin{bmatrix} x \\ y \end{bmatrix} \\ \begin{bmatrix} u & v \end{bmatrix} &= \begin{bmatrix} x & y \end{bmatrix} U \end{aligned}$$

   and we have

$$\begin{aligned} 28x^2 + 8xy + 22y^2 &= \mathbf{x}^T A\mathbf{x} \\ &= \begin{bmatrix} x & y \end{bmatrix} UDU^T \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} u & v \end{bmatrix} D \begin{bmatrix} u \\ v \end{bmatrix} \\ &= 30u^2 + 20v^2 \end{aligned}$$

   (d) Therefore $U$ maps the ellipse $30u^2 + 20v^2 = 10$ into the ellipse $28x^2 + 8xy + 22y^2 = 10$, which is the problem we started with.

   (e) The vertices of $30u^2 + 20v^2 = 10$ are are $\pm \begin{bmatrix} 1/\sqrt{3} \\ 0 \end{bmatrix} = \pm \begin{bmatrix} 0.577 \\ 0 \end{bmatrix}$ and $\pm \begin{bmatrix} 0 \\ 1/\sqrt{2} \end{bmatrix} = \pm \begin{bmatrix} 0 \\ 0.707 \end{bmatrix}$

$30u^2 + 20v^2 = 10$

(f) The vertices of $28x^2+8xy+22y^2=10$ are

$$\begin{aligned} U\begin{bmatrix} \pm 1/\sqrt{3} \\ 0 \end{bmatrix} &= \pm\begin{bmatrix} 0.894 & 0.447 \\ 0.447 & -0.894 \end{bmatrix}\begin{bmatrix} 0.577 \\ 0 \end{bmatrix} \\ &= \pm\begin{bmatrix} 0.515 \\ 0.257 \end{bmatrix} \\ U\begin{bmatrix} 0 \\ \pm 1/\sqrt{2} \end{bmatrix} &= \pm\begin{bmatrix} 0.894 & 0.447 \\ 0.447 & -0.894 \end{bmatrix}\begin{bmatrix} 0 \\ 0.707 \end{bmatrix} \\ &= \pm\begin{bmatrix} 0.316 \\ -0.632 \end{bmatrix} \end{aligned}$$

These are the vertices for $28x^2+8xy+22y^2=10$ that we found earlier.

6. Class try $3x^2-2xy+3y^2=4$. Find the vertices.

---

## 3.11 Subspaces

### 3.11.1 What is a subspace?

1. **Class:** prove if $A$ is a matrix and $\mathbf{x},\mathbf{y}$ are vectors and $A\mathbf{x}=A\mathbf{y}=\mathbf{0}$ then $A(\mathbf{x}+\mathbf{y})=0$.
2. **Class:** prove if $A$ is a matrix and $\mathbf{x}$ is a vector and $c$ is a scalar and $A\mathbf{x}=\mathbf{0}$ then $A(c\mathbf{x})=0$.
3. **Definition:** if $A$ is a $r\times c$ matrix then the **kernal of** $A$ is the set of solutions to $A\mathbf{x}=\mathbf{0}$
$$\ker A=\{\mathbf{x}\in\mathbb{R}^c : A\mathbf{x}=\mathbf{0}_r\}$$
4. **Theorem:** let $A$ be a matrix and $\mathbf{x},\mathbf{y}$ be vectors. Then
   (a) if $\mathbf{x},\mathbf{y}\in\ker A$ then $\mathbf{x}+\mathbf{y}\in\ker A$
   (b) if $\mathbf{x}\in\ker A$ and $c\in\mathbb{R}$ then $c\mathbf{x}\in\ker A$
5. **Definition:** A **subspace** of $\mathbb{R}^n$ is a subset $V\subset\mathbb{R}^n$ such that:
   (a) $\mathbf{0}\in V$
   (b) if $\mathbf{u},\mathbf{v}\in V$ then $\mathbf{u}+\mathbf{v}\in V$
   (c) if $\mathbf{u}\in V$ and $a\in\mathbb{R}$ then $a\mathbf{v}\in V$.
6. subspaces are an important idea for understanding linear equations, orthogonality and eigenvalues.
7. Trivial (stupid) Examples
   (a) The set $\{\mathbf{0}\}$ called the zero-subspace
   (b) $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$
   (c) In fact, these two subspaces are called the **trivial subspaces** of $\mathbb{R}^n$. A **non-trivial subspace** is a subspace that is not one of these.
8. Important examples of non-trivial subspaces.
   (a) If $\mathbf{x}$ is a non-zero vector in $\mathbb{R}^2$, $V=\{r\mathbf{x}:r\in\mathbb{R}\}$ is a non-trivial subspace. This is a line.
   **Proof:**
      i. Let $r=0$, so $r\mathbf{x}=0\mathbf{x}=\mathbf{0}$, so $\mathbf{0}\in V$.

ii. If $\mathbf{u}, \mathbf{v} \in V$ then there exist scalars $r_1, r_2 \in \mathbb{R}$ such that $\mathbf{u} = r_1\mathbf{x}$ and $\mathbf{v} = r_2\mathbf{x}$. Then

$$\begin{aligned} \mathbf{u}+\mathbf{v} &= r_1\mathbf{x}+r_2\mathbf{y} \\ &= (r_1+r_2)\,\mathbf{x} \end{aligned}$$

which is the right form for an element of $V$ so $\mathbf{u}+\mathbf{v} \in V$..

iii. If $\mathbf{u} \in V$ and $a \in \mathbb{R}$ then there exists a scalar $r \in \mathbb{R}$ such that $\mathbf{u} = r\mathbf{x}$. Then

$$\begin{aligned} a\mathbf{u} &= a\,(r\mathbf{x}) \\ &= (ar)\,\mathbf{x} \end{aligned}$$

which is the right form for an element of $V$, so $a\mathbf{u} \in V$.

Therefore $V$ is a subspace of $\mathbb{R}^2$.

(b) If $\mathbf{x}$ is a non-zero vector in $\mathbb{R}^n$, $V = \{a\mathbf{x} : a \in \mathbb{R}\}$ is a non-trivial subspace. This is a line in $\mathbb{R}^n$.

i. Class: prove this is a subspace.

(c) If $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$, then $V = \{r_1\mathbf{x}_1+r_2\mathbf{x}_2 : r_1, r_2 \in \mathbb{R}\}$ is a subspace. If $\mathbf{x}$ and $\mathbf{y}$ are not parallel, this is a plane.

**Proof:**

i. Let $r_1 = r_2 = 0$, so $r_1\mathbf{x}_1 + r_2\mathbf{x}_2 = 0\mathbf{x}_1+0\mathbf{x}_2 = \mathbf{0}$ so $\mathbf{0} \in V$.

ii. If $\mathbf{u}, \mathbf{v} \in V$ then there exist scalars $r_1, r_2, s_1, s_2 \in \mathbb{R}$ such that $\mathbf{u} = r_1\mathbf{x}_1+r_2\mathbf{x}_2$ and $\mathbf{v} = s_1\mathbf{x}_1+s_2\mathbf{x}_2$. Then

$$\begin{aligned} \mathbf{u}+\mathbf{v} &= (r_1\mathbf{x}_1+r_2\mathbf{x}_2)+(s_1\mathbf{x}_1+s_2\mathbf{x}_2) \\ &= (r_1+s_1)\,\mathbf{x}_1+(r_2+s_2)\,\mathbf{x}_2 \end{aligned}$$

which is the right form for an element of $V$ so $\mathbf{u}+\mathbf{v} \in V$..

iii. If $\mathbf{u} \in V$ and $a \in \mathbb{R}$ then there exists a scalars $r_1, r_2 \in \mathbb{R}$ such that $\mathbf{u} = r_1\mathbf{x}_1+r_2\mathbf{x}_2$. Then

$$\begin{aligned} a\mathbf{u} &= a\,(r_1\mathbf{x}_1+r_2\mathbf{x}_2) \\ &= (ar_1)\,\mathbf{x}_1+(ar_2)\,\mathbf{x}_2 \end{aligned}$$

which is the right form for an element of $V$, so $a\mathbf{u} \in V$.

Therefore $V$ is a subspace of $\mathbb{R}^3$.

(d) If $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$, then $V = \{a\mathbf{x}_1+b\mathbf{x}_2 : a, b \in \mathbb{R}\}$ is a subspace.

i. Class: prove this is a subspace.

9. Class: explain why the following are not subspaces. Which of the three rules is broken. It is only necessary to find one.

(a) a single, non-zero vector

(b) a line or plane not through the origin.

(c) $\left\{\begin{bmatrix} x \\ y \end{bmatrix} : x > 0, y > 0\right\}$ first quadrant of $\mathbb{R}^2$

(d) $\left\{\begin{bmatrix} x \\ y \end{bmatrix} : xy \geq 0\right\}$ first and third quadrants of $\mathbb{R}^2$

10. **Theorem**: if $V \subset \mathbb{R}^n$ is a subspace then

(a) $\mathbf{v} \in V \Rightarrow -\mathbf{v} \in V$

(b) $\mathbf{u}, \mathbf{v} \in V \Rightarrow \mathbf{u}-\mathbf{v} \in V$

(c) $\mathbf{u}_1, \ldots, \mathbf{u}_t \in V$ and $r_1, \ldots, r_t \in \mathbb{R} \Rightarrow r_1\mathbf{u}_1 + \cdots + r_t\mathbf{u}_t \in V$

11. **Proof:**

   (a) Use property $(c)$ of subspaces and $r = -1$

   (b) By $(a)$ $-\mathbf{v} \in V$. Then by property $(b)$ of subspaces, $\mathbf{u} + (-\mathbf{v}) = \mathbf{u} - \mathbf{v} \in V$.

   (c) By property $(c)$ of subspaces, $r_1\mathbf{u}_1, \ldots, r_t\mathbf{u}_t \in V$. Thus by property $(b)$:

$$\begin{aligned} r_1\mathbf{u}_1 + r_2\mathbf{u}_2 &\in V \\ (r_1\mathbf{u}_1 + r_2\mathbf{u}_2) + r_3\mathbf{u}_3 &\in V \\ &\vdots \\ (r_1\mathbf{u}_1 + \cdots + r_{t-1}\mathbf{u}_{t-1}) + r_t\mathbf{u}_t &\in V \end{aligned}$$

12. How are subspaces created. One way is by taking the **span** of a set of vectors.

   (a) **Definition**: if $\mathbf{x}_1, \ldots, \mathbf{x}_t \in \mathbb{R}^n$ then

$$\begin{aligned} & \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t) \\ = & \{r_1\mathbf{x}_1 + \cdots + r_t\mathbf{x}_t : r_1, \ldots, r_t \in \mathbb{R}\} \end{aligned}$$

   (b) Example: If $\mathbf{x}_1, \mathbf{x}_2$ are non-parallel vectors in $\mathbb{R}^3$ then $\operatorname{span}(\mathbf{x}_1, \mathbf{x}_2)$ is a plane in $\mathbb{R}^3$.

   (c) **Theorem:** $\operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ is a subspace of $\mathbb{R}^n$. called the subspace spanned by $\mathbf{x}_1, \ldots, \mathbf{x}_t$.or the span of $\mathbf{x}_1, \ldots, \mathbf{x}_t$.

   (d) **Proof:**

      i. Let $r_1 = \cdots = r_t = 0$, so $r_1\mathbf{x}_1 + \cdots + r_t\mathbf{x}_t = 0\mathbf{x}_1 + \cdots + 0\mathbf{x}_t = \mathbf{0}$ so $\mathbf{0} \in \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$.

      ii. If $\mathbf{u}, \mathbf{v} \in \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ then there exist scalars $r_1, \ldots, r_t, s_1, \ldots, s_t \in \mathbb{R}$ such that $\mathbf{u} = r_1\mathbf{x}_1 + \cdots + r_t\mathbf{x}_t$ and $\mathbf{v} = s_1\mathbf{x}_1 + \cdots + s_t\mathbf{x}_t$. Then

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (r_1\mathbf{x}_1 + \cdots + r_t\mathbf{x}_t) + (s_1\mathbf{x}_1 + \cdots + s_t\mathbf{x}_t) \\ &= (r_1 + s_1)\mathbf{x}_1 + \cdots + (r_t + s_t)\mathbf{x}_t \end{aligned}$$

      which is the right form for an element of $\operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ so $\mathbf{u} + \mathbf{v} \in \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$..

      iii. If $\mathbf{u} \in V$ and $a \in \mathbb{R}$ then there exists a scalars $r_1, \ldots, r_t \in \mathbb{R}$ such that

$$\mathbf{u} = r_1\mathbf{x}_1 + \cdots + r_t\mathbf{x}_t$$

      Then

$$\begin{aligned} a\mathbf{u} &= a(r_1\mathbf{x}_1 + \cdots + r_t\mathbf{x}_t) \\ &= (ar_1)\mathbf{x}_1 + \cdots + (ar_t)\mathbf{x}_t \end{aligned}$$

      which is the right form for an element of $\operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$, so $a\mathbf{u} \in \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$.

   Therefore $\operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ is a subspace of $\mathbb{R}^n$.

   (e) Important observation:

$$\mathbf{x}_1, \ldots, \mathbf{x}_t \in \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$$

   **Class** prove

   (f) Example: in $\mathbb{R}^3$ what is the span of $\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$.

      i. The plane through $\mathbf{0}$ and these points. The equation is $x - y - 2z = 0$.

   (g) **Theorem:** If $V$ is a subspace and $\mathbf{x}_1, \cdots \mathbf{x}_t \in V$ then $\operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t) \subset V$.

i. **Class** prove.

(h) **Corollary:** if $\mathbf{x}_1, \cdots \mathbf{x}_s \in \operatorname{span}(\mathbf{y}_1, \ldots, \mathbf{y}_t)$ then $\operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_s) \subset \operatorname{span}(\mathbf{y}_1, \ldots, \mathbf{y}_t)$

13. What are the main questions about subspaces?

(a) Given a subspace $V \subset \mathbb{R}^n$ and a vector $\mathbf{x} \in \mathbb{R}^n$, is $\mathbf{x} \in V$?

i. The plane $3x - 2y + z = 0$ is a subspace of $\mathbb{R}^3$. Does it contains the vector $\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$?

(b) Given two subspaces $V, W \subset \mathbb{R}^n$, is $V \subset W$?

i. The plane $3x - 2y + z = 0$ and the line through the origin parallel to $\begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$ are both subspaces of $\mathbb{R}^3$. Is the line contained in the plane?

(c) Given a subspace $V \subset \mathbb{R}^n$ how can we measure the size of $V$?

i. For example, planes in $\mathbb{R}^3$ are "two dimensional" and lines are "one dimensional"

ii. Exact meaning of dimension to be explained later.

(d) Given a set of vectors, what is the smallest subspace containing the vectors?

i. What is the smallest subspace of $\mathbb{R}^3$ contaniing $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$?

### 3.11.2 The Four Fundamental Subspaces of a Matrix

1. Where to subspaces come from?

(a) When you get a question like "is the vector $\mathbf{x}$ in the subspace $V$?", you cannot answer it until you know which vector $\mathbf{x}$ and which subspace $V$.

(b) In the next section we develop methods for constructing subspaces.

(c) The concept of subspace is really just a useful tool for organizing our ideas about these concrete constructions.

#### 3.11.2.1 Null Spaces or Kernels

1. Let $A$ be an $n \times m$ matrix. Then $A : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ and we can try to solve the equation $A\mathbf{x} = \mathbf{0}_r$.

(a) The set of solutions is called the **null space** or **kernel** of $A$.

(b) There is always at least one solution, namely $\mathbf{x} = \mathbf{0}$

(c) The big question is to identify matrices that have non-zero solutions, and to somehow measure the size of the solution set.

(d) Class: $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$. Find a non-zero vector in $\ker A$.

2. **Theorem:** The kernel of $A$ is a subspace of $\mathbb{R}^m$

(a) **Proof:** We must show that $\ker A$ satisfies the three axioms for subspaces.

i. Since $A\mathbf{0} = \mathbf{0}$, $\mathbf{0} \in \ker A$

ii. If $\mathbf{u}, \mathbf{v} \in \ker A$ then $A\mathbf{u} = A\mathbf{v} = \mathbf{0}$ so $\mathbf{0} = A\mathbf{u} + A\mathbf{v} = A(\mathbf{u} + \mathbf{v})$ and thus $\mathbf{u} + \mathbf{v} \in \ker A$

iii. If $\mathbf{u} \in \ker A$ and $a \in \mathbb{R}$ then $A\mathbf{u} = \mathbf{0}$ so $\mathbf{0} = a(A\mathbf{u}) = A(a\mathbf{u})$ and thus $a\mathbf{u} \in \ker A$.

(b) Thus, if $\mathbf{u}_1, \cdots, \mathbf{u}_t \in \ker A$ and $a_1, \ldots, a_t \in \mathbb{R}$ then $a_1\mathbf{u}_1 + \cdots + a_t\mathbf{u}_t \in \ker A$

(c) Example: $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \in \ker A$ so $3\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix} \in \ker A$

(d) Class: check

3. **Old Theorem:** Let $A$ be a matrix. If you do some row operations on $A$ to get a new matrix $B$, then $\ker B = \ker A$.

   (a) **Class** prove

   (b) we've been using this fact for some time

4. **Old Theorems—class prove:**

   (a) $A$ has a pivot in every column if and only if $\ker A = \{\mathbf{0}\}$.

   (b) If $A$ has more columns than rows, then $\ker A \neq \{\mathbf{0}\}$.

   (c) $A$ is invertible if and only if $A$ is square and $\ker A = \{\mathbf{0}\}$.

#### 3.11.2.2 Column Spaces

1. The column space of a matrix $A$ is the subspace spanned by the columns of $A$

   (a) The column space is denoted $\operatorname{image}(A)$

   (b) the reason for the name "image" will be clear after the next theorem.

2. Example: $\operatorname{image}\begin{bmatrix} 1 & 3 \\ 2 & -4 \\ 5 & 7 \end{bmatrix}$ is a plane

   (a) **Class:** find the equation of the plane.

3. **Theorem:** Let $A$ be a $n \times m$ matrix. Then $\mathbf{u} \in \operatorname{image}(A)$ if and only if $\mathbf{u} = A\mathbf{x}$ for some vector $\mathbf{x} \in \mathbb{R}^m$.

   (a) **Remark:** If you think of $A : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ as a map, then the column space of $A$ consists of the vectors in the image of the map $A$.

   (b) **Proof:** Let $A = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_m \end{bmatrix}$. Then

$$\begin{aligned} \mathbf{y} \in \operatorname{image}(A) &\Longleftrightarrow \mathbf{y} = x_1\mathbf{c}_1 + \cdots + x_m\mathbf{c}_m \\ &\Longleftrightarrow \mathbf{y} = A\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \\ &\Longleftrightarrow \mathbf{y} = A\mathbf{x} \text{ for } \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \end{aligned}$$

4. Easy observation: if $A$ and $B$ are matrices with the same number of rows, then $\operatorname{image} A \subset \operatorname{image}\begin{bmatrix} A & B \end{bmatrix}$

5. Important questions to answer:

   (a) Given a vector $\mathbf{b}$ and a matrix $A$, is $\mathbf{b} \in \operatorname{image} A$

      i. is $\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \in \operatorname{image}\begin{bmatrix} 1 & 3 \\ 2 & -4 \\ 5 & 7 \end{bmatrix}$

(b) Given matrices $A$ and $B$, is image $A \subset$ image $B$

i. is image $\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \subset$ image $\begin{bmatrix} 1 & 3 \\ 2 & -4 \\ 5 & 7 \end{bmatrix}$

(c) Given matrices $A$ and $B$, does image $A =$ image $B$

i. does image $\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} =$ image $\begin{bmatrix} 1 & 3 \\ 2 & -4 \\ 5 & 7 \end{bmatrix}$

6. These questions are related:

(a) Question $(b)$ answers question $(c)$: image $A =$ image $B$ if and only if image $A \subset$ image $B$ and image $B \subset$ image $A$

(b) Question $(a)$ answers question $(b)$: image $A \subset$ image $B$ if and only if $col_i(A) \in$ image $B$ for all $i$.

(c) So if we can answer question $(a)$, we can answer the others.

7. **Theorem:** $\mathbf{b} \in$ image $A \iff$ rank $A =$ rank $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$

(a) **Proof:**

$$\begin{aligned} \mathbf{b} &\in \text{image } A \\ &\iff \text{there exists a vector } \mathbf{x} \text{ such that } A\mathbf{x} = \mathbf{b} \\ &\iff \text{rank } A = \text{rank} \begin{bmatrix} A & \mathbf{b} \end{bmatrix} \end{aligned}$$

(b) Another way to say this is $\mathbf{b} \in$ image $A \iff$ the last column of $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ is not a pivot column.

8. Example

$$\begin{aligned} A &= \begin{bmatrix} -85 & 97 & 49 \\ -55 & 50 & 63 \\ -37 & 79 & 57 \\ -35 & 56 & -59 \end{bmatrix} \\ \mathbf{b} &= \begin{bmatrix} 2503 \\ 2467 \\ 5443 \\ -2185 \end{bmatrix} \\ & RREF \begin{bmatrix} A & \mathbf{b} \end{bmatrix} \\ &= RREF \begin{bmatrix} -85 & 97 & 49 & 2503 \\ -55 & 50 & 63 & 2467 \\ -37 & 79 & 57 & 5443 \\ -35 & 56 & -59 & -2185 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 77 \\ 0 & 1 & 0 & 66 \\ 0 & 0 & 1 & 54 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

so $\mathbf{b} \in$ image $A$

9. **Corollary:** image $A \subset$ image $B \iff$ rank $B =$ rank $\begin{bmatrix} B & A \end{bmatrix} \iff$ there are no pivot columns in the $A$-part of $\begin{bmatrix} B & A \end{bmatrix}$

10. Example

$$B = \begin{bmatrix} -85 & 97 & 49 \\ -55 & 50 & 63 \\ -37 & 79 & 57 \\ -35 & 56 & -59 \end{bmatrix}$$

$$A = \begin{bmatrix} 2503 & 7039 \\ 2467 & 1382 \\ 5443 & 4529 \\ -2185 & 9318 \end{bmatrix}$$

$$RREF \begin{bmatrix} B & A \end{bmatrix} = RREF \begin{bmatrix} -85 & 97 & 49 & 2503 & 7039 \\ -55 & 50 & 63 & 2467 & 1382 \\ -37 & 79 & 57 & 5443 & 4529 \\ -35 & 56 & -59 & -2185 & 9318 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 77 & -5 \\ 0 & 1 & 0 & 66 & 99 \\ 0 & 0 & 1 & 54 & -61 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

so image $A \subset$ image $B$.

(a) If we change one number in $A$, we get a different result.

$$A = \begin{bmatrix} 2503 & 7039 \\ 2467 & 1382 \\ 5443 & 4529 \\ -2185 & 9319 \end{bmatrix}$$

$$RREF \begin{bmatrix} B & A \end{bmatrix} = RREF \begin{bmatrix} -85 & 97 & 49 & 2503 & 7039 \\ -55 & 50 & 63 & 2467 & 1382 \\ -37 & 79 & 57 & 5443 & 4529 \\ -35 & 56 & -59 & -2185 & 9319 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 77 & 0 \\ 0 & 1 & 0 & 66 & 0 \\ 0 & 0 & 1 & 54 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The first column of $A$ is in image $B$, but the second column is not.

11. **Corollary:** image $A =$ image $B \Longleftrightarrow \text{rank}\, A = \text{rank} \begin{bmatrix} A & B \end{bmatrix} = \text{rank}\, B$

(a) This requires using the fact that $\text{rank} \begin{bmatrix} A & B \end{bmatrix} = \text{rank} \begin{bmatrix} B & A \end{bmatrix}$.

$$\begin{aligned} \text{rank} \begin{bmatrix} A & B \end{bmatrix} &= \text{rank} \begin{bmatrix} A & B \end{bmatrix}^T \\ &= \text{rank} \begin{bmatrix} A^T \\ B^T \end{bmatrix} \\ &= \text{rank} \begin{bmatrix} B^T \\ A^T \end{bmatrix} \\ &= \text{rank} \begin{bmatrix} B^T \\ A^T \end{bmatrix}^T \\ &= \text{rank} \begin{bmatrix} B & A \end{bmatrix} \end{aligned}$$

(b) Just proving $\text{rank}\, A = \text{rank}\, B$ is not enough to show image $A =$ image $B$.

#### 3.11.2.3 Row Spaces

1. If $A$ is a matrix with $n$ columns, then the rows of $A$ are vectors in $\mathbb{R}^n$ and the span of all the columns is called the **row space** of $A$ and is denoted $\text{RowSpace}(A)$.

   (a) By the theorem and observation above, $\text{RowSpace}\, A$ is a subspace of $\mathbb{R}^n$ and the rows of $A$ are contained in $\text{RowSpace}\, A$.

   (b) Example: $\text{RowSpace}\begin{bmatrix} 1 & 2 & 5 \\ 3 & -4 & 7 \end{bmatrix}$ is a plane in $\mathbb{R}^3$, the plane spanned by $\begin{bmatrix} 1 & 2 & 5 \end{bmatrix}$ and $\begin{bmatrix} 3 & -4 & 7 \end{bmatrix}$.

   (c) **Class**: find equation of this plane.

2. **Class**: show that $\text{RowSpace}\, I_n = \mathbb{R}^n$.

3. **Class:** prove $\text{RowSpace}\, A = \text{image}\, A^T$

   (a) so any question about row spaces becomes a question about column spaces or images

   (b) $\mathbf{b} \in \text{RowSpace}\, A \iff \text{rank}\, A = \text{rank}\begin{bmatrix} A \\ \mathbf{b} \end{bmatrix}$

   (c) $\text{RowSpace}\, A \subset \text{RowSpace}\, B \iff \text{rank}\, B = \text{rank}\begin{bmatrix} B \\ A \end{bmatrix}$

   (d) $\text{RowSpace}\, A = \text{RowSpace}\, B \iff \text{rank}\, A = \text{rank}\begin{bmatrix} A \\ B \end{bmatrix} = \text{rank}\, B.$

4. **Class** prove these statements.

5. The four fundamental subspaces of a $r \times c$ matrix $A$ are

$$\begin{aligned} \ker A &\subset \mathbb{R}^c \\ \ker A^T &\subset \mathbb{R}^r \\ \text{RowSpace}\, A &= \text{image}\, A^T \subset \mathbb{R}^c \\ \text{image}\, A &= \text{RowSpace}\, A^T \subset \mathbb{R}^r \end{aligned}$$

   (a) **Class:** prove:

      i. If $\mathbf{x} \in \ker A$ and $\mathbf{y} \in \text{image}\, A^T$ then $\mathbf{x} \cdot \mathbf{y} = 0$

      ii. If $\mathbf{x} \in \ker A^T$ and $\mathbf{y} \in \text{image}\, A$ then $\mathbf{x} \cdot \mathbf{y} = 0$

### 3.11.3 Subspaces and Dimension

1. Our first goal is to show that every subspace of $\mathbb{R}^n$ is the span of a set of vectors.

   (a) This is the same as showing that every subspace of $\mathbb{R}^n$ is the column space of some matrix with $n$ rows

   (b) This will remove some of the mystery from the concept of subspace

2. We need some preliminary results. This is a bit long, since we are developing one of the key results of linear algebra, namely that all subspaces of $\mathbb{R}^n$ are finite dimensional.

   (a) Suppose we have a $r \times c$ matrix $A$ and a column vector $\mathbf{b} \in \mathbb{R}^r$, so we can form the augmented matrix $B = \begin{bmatrix} A & \mathbf{b} \end{bmatrix}$. Then $\text{rank}(A) = \text{rank}(B)$ if and only iff $\mathbf{b} \in \text{image}(A)$. Otherwise $\text{rank}(B) = \text{rank}(A) + 1$.

      i. Proof: We proved the first part already: $\text{rank}(A) = \text{rank}(B)$ if and only if $\mathbf{b} \in \text{image}(A)$. Now for the otherwise part. It is obvious that $B$ has no fewer pivots than $A$ and no more than one additional pivot, so clearly $\text{rank}(A) \leq \text{rank}(B) \leq \text{rank}(A+1)$. Thus, if $\text{rank}(A) \neq \text{rank}(B)$ then $\text{rank}(B) = \text{rank}(A) + 1$.

(b) Suppose we have column vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{t+1} \in \mathbb{R}^n$ such that $\mathbf{x}_{t+1} \notin \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$. Then $\operatorname{rank} \left[ \begin{array}{ccc} \mathbf{x}_1 & \cdots & \mathbf{x}_{t+1} \end{array} \right] = \operatorname{rank} \left[ \begin{array}{ccc} \mathbf{x}_1 & \cdots & \mathbf{x}_t \end{array} \right] + 1$.

i. **Proof:** let $A = \left[ \begin{array}{ccc} \mathbf{x}_1 & \cdots & \mathbf{x}_t \end{array} \right]$ and $\mathbf{b} = \mathbf{x}_{t+1}$. Recall $\operatorname{image} A = \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$. By hypothesis $\mathbf{b} \notin \operatorname{image}(A)$, so the previous result implies this one.

3. Now for the key **theorem:** suppose $V \subset \mathbb{R}^n$ is a subspace. We will we will prove that there exists a sequence of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_t$ spanning $V$.

(a) Choose $\mathbf{x}_1 = \mathbf{0} \in V$.

(b) Form a sequence of vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots \in V$ such that $\operatorname{span}(\mathbf{x}_1, \ldots \mathbf{x}_t) \subsetneqq \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_{t+1})$.

i. If you have constructed vectors $\mathbf{x}_1, \ldots, \mathbf{x}_t$ such that $\operatorname{span}(\mathbf{x}_1, \ldots \mathbf{x}_t) \subsetneqq V$, you can construct $\mathbf{x}_{t+1}$.

ii. I claim that the sequence $\mathbf{x}_1, \ldots, \mathbf{x}_t$ cannot have more than $n+1$ vectors.

iii. If so, then.for some $t \leq n+1$ it will be the case that $\operatorname{span}(\mathbf{x}_1, \ldots \mathbf{x}_t) = V$ and you have found the desired spanning set for $V$.

(c) Here is the reason you cannot have more than $n+1$ vectors in your sequence. Think of the $\mathbf{x}_t$ as column vectors, and let $A_t = \left[ \begin{array}{ccc} \mathbf{x}_1 & \cdots & \mathbf{x}_t \end{array} \right]$. Since $\mathbf{x}_1 = \mathbf{0}$, $\operatorname{rank} A_1 = 0$. By the preliminary results, $\operatorname{rank} A_{t+1} = \operatorname{rank} A_t + 1 = t$. In other words, $\operatorname{rank} A_t = t - 1$. Since $A_t$ has $n$ rows, $\operatorname{rank} A_t \leq n$.or $t - 1 \leq n$ or $t \leq n + 1$.

4. Class: find a matrix whose row space is the subspace $x - y + z = 0$ in $\mathbb{R}^3$.

5. We have already proven that $\operatorname{image} A = \operatorname{image} B$ implies $\operatorname{rank} A = \operatorname{rank} B$. Thus, if a subspace $V$ is the image of many different matrices (which will be the case), all the matices have the same rank.

6. **Very Important Definition:** the **dimension** of a subspace $V$ is the rank of any matrix $A$ such that $\operatorname{image}(A) = V$.

(a) That is the same as the rank of any matrix $A$ such that $\operatorname{image} A = V$.

7. **Class**: what is the dimension of $\mathbb{R}^n$.

(a) What is the dimension of the subspace $\{\mathbf{0}\}$?

8. **Class**: what is the dimension of the subspace $x - y + z = 0$ in $\mathbb{R}^3$.

9. **Theorem:** $\dim(\operatorname{RowSpace} A) = \operatorname{rank} A$

(a) **Proof:** since $\operatorname{rank} A = \operatorname{rank} A^T$ and $\operatorname{RowSpace} A = \operatorname{image} A^T$,

$$\begin{aligned} \dim(\operatorname{RowSpace} A) &= \dim\left(\operatorname{image} A^T\right) \\ &= \operatorname{rank} A^T \\ &= \operatorname{rank} A \end{aligned}$$

(b) Thus $\operatorname{image} A$ and $\operatorname{RowSpace} A$ have the same dimension, *even though they may be subspaces of different Euclidean spaces.*

i. If $A$ is $r \times c$ then $\operatorname{RowSpace} A \subset \mathbb{R}^c$ and $\operatorname{image} A \subset \mathbb{R}^r$.

10. **Theorem:** Suppose $V \subset W \subset \mathbb{R}^n$ are subspaces. Then

(a) $\dim V \leq \dim W$

(b) If $\dim V = \dim W$ then $V = W$.

(c) If $V \subset U \subset W$ is a subspace and $\dim W = \dim V + 1$ then $U = V$ or $U = W$.

(d) If $\dim V = r$ and $\dim W = s$ then there exists subspaces $V = U_r \subset U_{r+1} \subset \cdots \subset U_{s-1} \subset U_s = W$ such that $\dim U_i = i$.

11. **Proof:**

(a) let $V = \operatorname{image} A$ and $W = \operatorname{image} B$. If $C = \begin{bmatrix} A & B \end{bmatrix}$ then $W = \operatorname{image} C$. Thus $\dim V = \operatorname{rank} A \leq \operatorname{rank} \begin{bmatrix} A & B \end{bmatrix} = \dim W$.

(b) If $V \subset W$ and $\dim V = \dim W$ then $\operatorname{rank} A = \operatorname{rank} \begin{bmatrix} A & B \end{bmatrix} = \operatorname{rank} B$ so $V = W$.

(c) Since $\dim V \leq \dim U \leq \dim W = \dim V + 1$, either $\dim U = \dim V$ or $\dim U = \dim W$. Since $V \subset U \subset W$, either $U = V$ or $U = W$.

(d) Start the sequence $U_i$ be defining $U_r = V$. If you have constructed $U_r \subset \cdots \subset U_i \subset W$ and $i < s$, you can construct $U_{i+1}$. Since $\dim U_i = i < s = \dim W$, $U_i \subsetneqq W$. Choose a vector $\mathbf{x} \in W \setminus U_i$. If $U_i = \operatorname{RowSpace} A$, let $U_{i+1} = \operatorname{RowSpace} \begin{bmatrix} A \\ \mathbf{x} \end{bmatrix}$. Then $U_i \subset U_{i+1} \subset W$ and

$$\begin{aligned} \dim U_{i+1} &= \operatorname{rank} \begin{bmatrix} A \\ \mathbf{x} \end{bmatrix} \\ &= \operatorname{rank} A + 1 \text{ because } \mathbf{x} \notin \operatorname{RowSpace} A \\ &= \dim U_i + 1 \\ &= i + 1 \end{aligned}$$

12. Class: if planes are two-dimensional subsets of $\mathbb{R}^3$, show that one plane cannot be properly contained in another.

### 3.11.4 Linear Independence and Bases of Subspaces

1. A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_t \in \mathbb{R}^n$ is **linearly dependent** if there exists scalars $a_1, \ldots, a_t$, not all zero, such that $a_1\mathbf{v}_1 + \ldots + a_t\mathbf{v}_t = \mathbf{0}$. A set of vectors that is not linearly dependent is **linearly independent.**

(a) a set of vectors is linearly independent if, when put into columns of a matrix, the rank of the matrix is equal to the number of columns.

i. **Proof**: We prove the contrapositive: let $\mathbf{v}_1, \ldots, \mathbf{v}_t$ be a set of (column) vectors. The vectors are linearly dependent if and only if $\operatorname{rank} \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_t \end{bmatrix} < t$. Let $A = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_t \end{bmatrix}$. The vectors are linearly dependent if and only if there exists scalars $a_1, \ldots, a_t$, not all zero, such that $a_1\mathbf{v}_1 + \ldots + a_t\mathbf{v}_t = \mathbf{0}$. But $a_1\mathbf{v}_1 + \ldots + a_t\mathbf{v}_t = A \begin{bmatrix} a_1 \\ \vdots \\ a_t \end{bmatrix}$. Let $\mathbf{x} = \begin{bmatrix} a_1 \\ \vdots \\ a_t \end{bmatrix}$. Then the columns of $A$ are linearly independent if and only if there exists a non-zero vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$. Such a non-zero vector exists if and only if $\operatorname{rank} A < t$.

(b) a set of vectors is linearly independent if, when put into rows of a matrix, the rank of the matrix is equal to the number of rows.

(c) **Class** prove: if a matrix has more columns than rows, then the columns are linearly dependent.

(d) **Class** prove: a set of more than $n$ vectors from $\mathbb{R}^n$ is linearly dependent.

i. Is a set of $n$ or fewer vectors from $\mathbb{R}^n$ necessarily linearly independent?

2. A **spanning set** for a subspace is a set of vectors that span the subspace.

   (a) If the subspace is $\text{RowSpace}(A)$, then the rows of $A$ are a spanning set

   (b) If the subspace is $\text{image}(A)$ then the columns of $A$ are a spanning set

   (c) If the subspace is $\ker A$, then it is harder to find a spanning set.

      i. We proved (7) that the basic solutions to the linear system $A\mathbf{x} = \mathbf{0}$ span $\ker A$.

3. A **basis** for a subspace is a spanning set that is linearly independent.

   (a) The rows of a matrix $A$ are a basis for $\text{RowSpace}(A)$ if and only if $\text{rank}\, A = \mathit{rows}(A)$

   (b) The columns of a matrix $A$ are a basis for $\text{image}(A)$ if and only if $\text{rank}\, A = \mathit{cols}(A)$

4. To find a basis for a subspace,

   (a) first find a spanning set. We've shown that every subspace has one.

   (b) make the spanning set the rows of a matrix $A$, so the given subspace is $\text{RowSpace}(A)$

   (c) Take the non-zero rows of $RREF(A)$. These non zero rows span the same space as the rows of $A$ and they form a reduced matrix, so they are linearly independent.

5. To find the basis for $\text{RowSpace}(A)$: take the non-zero rows of $RREF(A)$

6. To find a basis for $\text{image}\, A$ : take the non-zero rows of $RREF\left(A^T\right)$

7. Class: find a basis for the row space and image of $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 5 & 6 \end{bmatrix}$

8. To find a basis for $\ker A$, take the basic solutions to $A\mathbf{x} = \mathbf{0}$.

   (a) We showed in (7) that every vector in $\ker A$ is a unique linear combination. of the basic solutions. Suppose the basic solutions are $\mathbf{x}_1, \ldots, \mathbf{x}_t$. To show that the basic solutions are linearly independent, we must show that **if** $\mathbf{0} = a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t$ **then** $a_1 = \cdots = a_t = 0$. But

$$\begin{aligned} \mathbf{0} &= a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t \\ &= 0\mathbf{x}_1 + \cdots + 0\mathbf{x}_t \end{aligned}$$

   Since linear combinations of basic solutions are unique, $a_1 = \cdots = a_t = 0$.

9. **Important classical theorem:** If $V$ is a subspace, then all bases of $V$ have the same number of vectors, and that number is the dimension of $V$.

   (a) **Proof:** Let $V$ be a subspace, and suppose $V$ has a basis $\mathbf{x}_1, \ldots, \mathbf{x}_t$. If you put the vectors into the rows of a matrix $A$, then $V = \text{RowSpace}\, A$ (because the rows span $V$) and $\text{rank}\, A = \mathit{rows}(A) = t$ (since the $\mathbf{x}_i$ are linearly independent). Therefore $\dim V = \text{rank}\, A = t.\backslash$

10. **Corollary:** $\dim(\ker A) = \mathit{cols}(A) - \text{rank}(A)$.

   (a) Proof: $\dim(\ker A)$ is the number of free variables in the system $A\mathbf{x} = \mathbf{0}$, which is the number of variables less the number of pivots, or $\mathit{cols}(A) - \text{rank}(A)$.

11. Class: If $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 5 & 6 \end{bmatrix}$, what is the dimension of $\ker A$?

12. **Theorem:** if $A$ is a square matrix and $\lambda_1, \ldots, \lambda_t$ are **distinct** eigenvalues of $A$ with eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_t$ then $\{\mathbf{x}_1, \ldots, \mathbf{x}_t\}$ form a linearly independent set.

    (a) The proof is cute, and has two steps.

        i. Suppose $\mathbf{x}_1, \ldots, \mathbf{x}_t$ are eigenvalues of $A$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_t$. If there exist constants $a_1, \ldots, a_t$ such that $a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t = \mathbf{0}$, then there exists constants $b_1, \ldots, b_{t-1}$ such that $b_i \neq 0$ iff $a_i \neq 0$ and $b_1\mathbf{x}_1 + \cdots + b_{t-1}\mathbf{x}_{t-1} = \mathbf{0}$.

$$\begin{aligned} \mathbf{0} &= A(\mathbf{0}) \\ &= A(a_1\mathbf{x}_1 + \ldots + a_t\mathbf{x}_t) \\ &= a_1 A\mathbf{x}_1 + \cdots + a_t A\mathbf{x}_t \\ &= a_1\lambda_1\mathbf{x}_1 + \cdots + a_t\lambda_t\mathbf{x}_t \end{aligned}$$

$$\begin{aligned} \mathbf{0} &= \lambda_t\mathbf{0} \\ &= \lambda_t(a_1\mathbf{x}_1 + \ldots + a_t\mathbf{x}_t) \\ &= a_1\lambda_t\mathbf{x}_1 + \cdots + a_t\lambda_t\mathbf{x}_t \end{aligned}$$

        Subtracting, we get:

$$a_1(\lambda_1 - \lambda_t)\mathbf{x}_1 + \cdots + a_{t-1}(\lambda_{t-1} - \lambda_t)\mathbf{x}_{t-1} = \mathbf{0}$$

        Let $b_i = (\lambda_i - \lambda_t)a_i$. Since $\lambda_i - \lambda_t \neq 0$ for $i = 1, \ldots, t-1$, $b_i \neq 0$ iff $a_i \neq 0$.

        ii. The proof concludes by contradiction. If $\{\mathbf{x}_1, \ldots, \mathbf{x}_t\}$ are linearly dependent, then there exist constants $a_1, \ldots, a_t$ not all zero such that $a_1\mathbf{x}_1 + \cdots + a_t\mathbf{x}_t = \mathbf{0}$. We can renumber the terms so that $a_1 \neq 0$. But by applying the first part repeatedly, we can get $c \neq 0$ such that $c\mathbf{x}_1 = \mathbf{0}$. This is impossible since the eigenvector $\mathbf{x}_1 \neq \mathbf{0}$. Thus $\{\mathbf{x}_1, \ldots, \mathbf{x}_t\}$ is a linearly independent set.

### 3.11.5 The Main Problem About Representing Subspaces

1. Two big problems about matrices

    (a) Given a matrix $A$ find a matrix $B$ such that $\ker A = \text{image}\, B$

    (b) Given a matrix $A$ find a matrix $B$ such that $\text{image}\, A = \ker B$

2. Subspaces frequently always arise as either the image or the kernal of a matrix. And, given one represention, it is often necessary to find the other.

3. To solve the first problem: given $A$ find $B$ such that $\ker A = \text{image}\, B$.

    (a) This is the problem we began the semester with

    (b) Let $B$ be a matrix whose columns are the basic solutions of $A\mathbf{x} = \mathbf{0}$.

    (c) We have already shown that this solution is correct (8).

4. How about the second problem: given a matrix $A$ find a matrix $B$ such that $\text{image}\, A = \ker B$

    (a) Can you believe: $\text{image}\, A = \ker B$ if and only if $\ker A^T = \text{image}\, B^T$

    (b) If so, then given $A$ we want to find $B$ such that $\ker A^T = \text{image}\, B^T$

    (c) $B^T$ is the matrix whose columns are the basic solutions to $A^T\mathbf{x} = \mathbf{0}$.

5. If $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \\ 5 & 1 & 6 \end{bmatrix}$, find a matrix $B$ such that $\text{image}\, A = \ker B$.

(a) image $A = \ker B$.if and only if image $B^T = \ker A^T = \ker \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 5 & 6 \end{bmatrix}$.

(b) Row reducing $A^T$ we get $\begin{bmatrix} 1 & 0 & -1 & -2 & -3 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. The basic solutions to $A^T\mathbf{x} = \mathbf{0}$ are the columns of $\begin{bmatrix} 1 & 2 & 3 \\ -2 & -3 & -4 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. If we call this matrix $B^T$ then $\ker A^T = \text{image } B^T$. Thus image $A = \ker B = \ker \begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 2 & -3 & 0 & 1 & 0 \\ 3 & -4 & 0 & 0 & 1 \end{bmatrix}$

(c) Check: $BA = 0$ so every column of $B$ is in the kernal of $A$. Thus image $B \subset \ker A$. Also $\dim \ker A = 5 - \text{rank } A = 5 - 2 = 3$ and $\dim \text{image } B = \text{rank } B = \text{rank } B^T = 3$, so image $B = \ker A$.

6. There is an algorithm for doing all this. Given a matrix $A$ we want to find $B$ such that $\ker A = \text{image } B$

   (a) Form the matrix $\begin{bmatrix} A^T & I_n \end{bmatrix}$ and row reduce to $\begin{bmatrix} U & * \\ 0 & B^T \end{bmatrix}$, where $\begin{bmatrix} U \\ 0 \end{bmatrix}$ is the row-reduced form of $A^T$.

      i. If there is no 0 part, if the row-reduced form of $A^T$ has no zero rows, (if rank $A = cols\,(A)$) , then $\ker A = \{\mathbf{0}\}$ and you have to take $B = 0_{n,1}$. This is a special case since $\{\mathbf{0}\}$ is not the image of any non-zero matrix.

      ii. Otherwise you get a matrix $B$ which is the answer.

   (b) Example: $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Adjoin $I$ and row reduce.

$$\left[\begin{array}{ccc|ccc} 1 & 4 & 7 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 0 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{array}\right]$$
$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & -2 & \frac{5}{3} \\ 0 & 1 & 2 & 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right]$$

   so $\ker \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \text{image} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$.

   (c) Given $A$ find $B$ such that image $A = \ker B$.

      i. solve the equivalent problem: given $A$ find $B$ such that $\ker A^T = \text{image } B^T$.

      ii. Form the augmented matrix $\begin{bmatrix} A & I_n \end{bmatrix}$ and row reduce to $\begin{bmatrix} U & * \\ 0 & B \end{bmatrix}$

   (d) Example:

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

$$\begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{bmatrix}$$
$$\longrightarrow \begin{matrix} 1 & 0 & -1 & 0 & -\frac{8}{3} & \frac{5}{3} \\ 0 & 1 & 2 & 0 & \frac{7}{3} & -\frac{4}{3} \\ 0 & 0 & 0 & 1 & -2 & 1 \end{matrix}$$

$$\text{so image} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \ker \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$$

### 3.11.6 Orthogonal Subspaces

1. Recall: two vectors $\mathbf{x}$ and $\mathbf{y}$ are **orthogonal** if and only if $\mathbf{x} \cdot \mathbf{y} = 0$

   (a) $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is orthogonal to $\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$.

2. A set of vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_t\}$ is **orthonormal** if and only if $\mathbf{x}_i \cdot \mathbf{x}_j = \delta_{ij}$

   (a) example: $\left\{ \begin{bmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}, \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \\ 0 \end{bmatrix} \right\}$

3. Recall: a matrix $A$ is **orthogonal** if and only if $A^T A = I$. In that case the columns of $A$ form an orthonormal set

   (a) Example: $\begin{bmatrix} 1/\sqrt{14} & 2/\sqrt{5} \\ 2/\sqrt{14} & -1/\sqrt{5} \\ 3/\sqrt{14} & 0 \end{bmatrix}$

4. A square matrix $B$ is an **orthogonal projection** if and only if $B = B^T = B^2$

   (a) An orthogonal projection is not, in general, an orthogonal matrix.

   (b) But if $A$ is an orthogonal matrix then $B = AA^T$ is an orthogonal projection matrix.

      i. $B^T = \left(AA^T\right)^T = A^{TT}A^T = AA^T = B$

      ii. $B^2 = \left(AA^T\right)^2 = AA^TAA^T = AIA^T = AA^T = B$

   (c) If $A = \begin{bmatrix} 1/\sqrt{14} & 2/\sqrt{5} \\ 2/\sqrt{14} & -1/\sqrt{5} \\ 3/\sqrt{14} & 0 \end{bmatrix}$ then

$$\begin{aligned} B &= \begin{bmatrix} 1/\sqrt{14} & 2/\sqrt{5} \\ 2/\sqrt{14} & -1/\sqrt{5} \\ 3/\sqrt{14} & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{14} & 2/\sqrt{5} \\ 2/\sqrt{14} & -1/\sqrt{5} \\ 3/\sqrt{14} & 0 \end{bmatrix}^T \\ &= \begin{bmatrix} \frac{61}{70} & -\frac{9}{35} & \frac{3}{14} \\ -\frac{9}{35} & \frac{17}{35} & \frac{3}{7} \\ \frac{3}{14} & \frac{3}{7} & \frac{9}{14} \end{bmatrix} \end{aligned}$$

   (d) If $B$ is an orthogonal projection, so is $C = I - B$

      i. $C^T = (I-B)^T = I^T - B^T = I - B = C$

      ii. $C^2 = (I-B)^2 = I - 2B + B^2 = I - 2B + B = I - B = C$

5. Suppose $B$ is an orthogonal projection and $C = I - B$. Then

   (a) $BC = CB = 0$

      i. $BC = B(I-B) = B - B^2 = 0$

      ii. $CB = (I-B)B = B - B^2 = 0$

   (b) If $\mathbf{x} \in image(B)$ and $\mathbf{y} \in image(C)$ then $\mathbf{x} \perp \mathbf{y}$.

      i. $\mathbf{x} = B\mathbf{s}$ and $\mathbf{y} = C\mathbf{t}$ for some vectors $\mathbf{s}$ and $\mathbf{t}$. Then

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= (B\mathbf{s}) \cdot (C\mathbf{t}) = \\ & \quad (B\mathbf{s})^T (C\mathbf{t}) \\ &= \mathbf{s}^T B^T C\mathbf{t} \\ &= \mathbf{s}^T BC\mathbf{t} \\ &= \mathbf{s}^T 0\mathbf{t} \\ &= 0 \end{aligned}$$

(c) If $B$ and $C$ are $n \times n$ and $\mathbf{z} \in \mathbb{R}^n$ then there exists $\mathbf{x} \in image\,(B)$, $\mathbf{y} \in image\,(C)$ such that $\mathbf{x}+\mathbf{y}=\mathbf{z}$.

i. This is too easy. Since $B+C=I$, $\mathbf{z}=B\mathbf{z}+C\mathbf{z}$. Take $\mathbf{x}=B\mathbf{z} \in image\,(B)$ and $\mathbf{y}=C\mathbf{z} \in image\,(C)$.

(d) $image\,(B)=\ker(C)$ and $image\,(C)=\ker(B)$.

i. It is only necessary to prove the first statement. First we show $image\,(B) \subset \ker(C)$.

$$\begin{aligned} \mathbf{x} &\in image\,(B) \Longrightarrow \mathbf{x}=B\mathbf{y} \text{ for some vector } \mathbf{y} \\ &\Longrightarrow C\mathbf{x}=CB\mathbf{y}=0\mathbf{y}=\mathbf{0} \\ &\Longrightarrow \mathbf{x} \in \ker(C) \end{aligned}$$

Next we show $\ker(C) \subset image\,(B)$:

$$\begin{aligned} \mathbf{x} &\in \ker(C) \Longrightarrow C\mathbf{x}=\mathbf{0} \\ &\Longrightarrow (I-B)\,\mathbf{x}=\mathbf{0} \\ &\Longrightarrow \mathbf{x}=B\mathbf{x} \\ &\Longrightarrow \mathbf{x} \in image\,(B) \end{aligned}$$

Therefore $image\,(C)=\ker(B)$.

6. Suppose $A$ is an orthogonal matrix and $B=AA^T$. Then $image\,(B)=image\,(A)$.

(a) First we show $image\,(B) \subset image\,(A)$:

$$\begin{aligned} \mathbf{u} &\in image\,(B) \Longrightarrow \mathbf{u}=AA^T\mathbf{x} \text{ for some vector } \mathbf{x} \\ &\Longrightarrow \mathbf{u}=A\mathbf{y} \text{ for } \mathbf{y}=A^T\mathbf{x} \\ &\Longrightarrow \mathbf{u} \in image\,(A) \end{aligned}$$

Next we show $image\,(A) \subset image\,(B)$:

$$\begin{aligned} \mathbf{u} &\in image\,(A) \Longrightarrow \mathbf{u}=A\mathbf{x} \text{ for some vector } \mathbf{x} \\ &\Longrightarrow \mathbf{u}=A\left(A^TA\right)\mathbf{x}=BA\mathbf{x} \\ &\Longrightarrow \mathbf{u}=B\mathbf{y} \text{ for } \mathbf{y}=A\mathbf{x} \\ &\Longrightarrow \mathbf{u} \in image\,(B)\,. \end{aligned}$$

Therefore $image\,(A)=image\,(B)$.

7. Let $V \subset \mathbb{R}^n$ be a subspace. Then there exists an orthgonal matrix $A$ and an orthogonal projection matrix $B$ such that $V=image(A)=image\,(B)$.

(a) Can you believe that there exists an orthonormal basis $\{\mathbf{x}_1,\ldots,\mathbf{x}_t\}$ for $V$. Then $A=\begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_t \end{bmatrix}$ is orthogonal, and $V=image\,(A)$. By (4b) and (6), $B=AA^T$ is an orthogonal projection matrix and $V=image\,(B)$.

8. Two subspaces $V,W \subset \mathbb{R}^n$ are **orthogonal** if and only if $\mathbf{x} \in V$, $\mathbf{y} \in W$ implies $\mathbf{x}\cdot\mathbf{y}=0$.

(a) In words, every vector in $V$ is orthogonal to every vector in $W$.

(b) Example. two orthogonal lines in $\mathbb{R}^2$ or $\mathbb{R}^3$, *e.g.* the lines spanned by $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$.

(c) Example: the plane $3x-2y+z=0$ and the line spanned by $\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$.

i. Proof: if $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ is in the plane and $\mathbf{y} = \begin{bmatrix} 3t \\ -2t \\ t \end{bmatrix}$ is on the line then $\mathbf{x} \cdot \mathbf{y} = 3tx - 2ty + tz = t(3x - 2y + z) = 0$.

(d) We say $V \perp W$ when $V$ is orthogonal to $W$.

9. **Theorem**: if $V \perp W$ then $V \cap W = \{\mathbf{0}\}$.

   (a) **Proof**: if $\mathbf{x} \in V \cap W$ then $\mathbf{x} \in V$ and $\mathbf{x} \in W$ so $\mathbf{x} \cdot \mathbf{x} = 0$. Thus $|\mathbf{x}|^2 = 0$ so $\mathbf{x} = \mathbf{0}$.

10. Here are some useful facts about about any two subspaces whose intersection is 0.

    (a) **Theorem:** Suppose $V, W \subset \mathbb{R}^n$ are subspaces and $V \cap W = \{\mathbf{0}\}$.

    i. Suppose further that $\mathbf{x}_1, \mathbf{x}_2 \in V$ and $\mathbf{y}_1, \mathbf{y}_2 \in W$ and $\mathbf{x}_1 + \mathbf{y}_1 = \mathbf{x}_2 + \mathbf{y}_2$. Then $\mathbf{x}_1 = \mathbf{x}_2$ and $\mathbf{y}_1 = \mathbf{y}_2$.

    ii. If $\{\mathbf{x}_1, \ldots, \mathbf{x}_s\} \subset V$ is linearly independent, if $\{\mathbf{y}_1, \ldots, \mathbf{y}_t\} \subset W$ is linearly independent then $\{\mathbf{x}_1, \ldots, \mathbf{x}_s, \mathbf{y}_1, \ldots, \mathbf{y}_t\} \subset \mathbb{R}^n$ is linearly independent.

    iii. $\dim V + \dim W \leq n$.

    (b) Here are the **proofs:**

    i. $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{y}_2 - \mathbf{y}_1 \in V \cap W = \{\mathbf{0}\}$, so $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{y}_2 - \mathbf{y}_1 = \mathbf{0}$.

    ii. Suppose $a_1\mathbf{x}_1 + \cdots + a_s\mathbf{x}_s + b_1\mathbf{y}_1 + \cdots + b_t\mathbf{y}_t = \mathbf{0}$. We must show $a_1 = \cdots = a_s = b_1 = \cdots = b_t = 0$. But

$$\begin{aligned} a_1\mathbf{x}_1 + \cdots + a_s\mathbf{x}_s &\in V \\ b_1\mathbf{y}_1 + \cdots + b_t\mathbf{y}_t &\in W \\ & (a_1\mathbf{x}_1 + \cdots + a_s\mathbf{x}_s) \\ & + (b_1\mathbf{y}_1 + \cdots + b_t\mathbf{y}_t) \\ &= \mathbf{0} + \mathbf{0} \end{aligned}$$

    By the first part, $a_1\mathbf{x}_1 + \cdots + a_s\mathbf{x}_s = \mathbf{0}$ and $b_1\mathbf{y}_1 + \cdots + b_t\mathbf{y}_t = \mathbf{0}$. By linear independence, $a_1 = \cdots = a_s = 0$ and $b_1 = \cdots = b_t = 0$.

    iii. Let $s = \dim V$ and $t = \dim W$. Choose bases $\{\mathbf{x}_1, \ldots, \mathbf{x}_s\}$ for $V$ and $\{\mathbf{y}_1, \ldots, \mathbf{y}_t\}$ for $W$. By the second part, $\{\mathbf{x}_1, \ldots, \mathbf{x}_s, \mathbf{y}_1, \ldots, \mathbf{y}_t\}$ is linearly independent in $\mathbb{R}^n$, so altogether there cannot be more than $n$ vectors. That is $\dim V + \dim W = s + t \leq n$.

11. Let $A$ be a $r \times c$ matrix. Then we can think of $A$ and $A^T$ as maps

$$\mathbb{R}^c \underset{A^T}{\overset{A}{\rightleftarrows}} \mathbb{R}^r$$

    Note that both $\ker A$, $image\left(A^T\right) \subset \mathbb{R}^c$ and $\ker A^T, image\,(A) \subset \mathbb{R}^r$.

12. **Theorem:** $\ker A \perp image\left(A^T\right)$ and $\ker A^T \perp image\,(A)$.

    (a) **Proof:** it is only necessary to prove one of these statements. Let $\mathbf{x} \in \ker A$ and $\mathbf{y} \in image(A^T)$. We must show $\mathbf{y} \cdot \mathbf{x} = 0$. Since $\mathbf{y} \in image\,(A)$, $\mathbf{y} = A^T\mathbf{z}$ for some vector $\mathbf{z}$. Therefore

$$\begin{aligned} \mathbf{y} \cdot \mathbf{x} &= \mathbf{y}^T\mathbf{x} \\ &= \left(A^T\mathbf{z}\right)^T \mathbf{x} \\ &= \mathbf{z}^T A\mathbf{x} \\ &= \mathbf{z}^T\mathbf{0} \\ &= 0 \end{aligned}$$

13. If $B$ is an orthogonal projection matrix and $C = I - B$ then $image\,(B) \perp image\,(C)$ (5b).

14. **Definition:** if $V \subset \mathbb{R}^n$ is a subspace, then $V^\perp = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{y} = 0, \forall \mathbf{y} \in V\}$.

    (a) $V^\perp$ is called the **orthogonal complement** of $V$ in $\mathbb{R}^n$.

    (b) **Example**: the orthogonal complement of a line in $\mathbb{R}^3$ is a plane in $\mathbb{R}^3$.

    (c) **Proposition:** $V^\perp$ is a subspace of $\mathbb{R}^n$

    i. **Proof**: for all $\mathbf{y} \in V$, $\mathbf{0} \cdot \mathbf{y} = 0$. Thus $\mathbf{0} \in V^\perp$.

    ii. If $\mathbf{x}_1, \mathbf{x}_2 \in V^\perp$ then for all $\mathbf{y} \in V$ we have $\mathbf{y} \cdot (\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{y} \cdot \mathbf{x}_1 + \mathbf{y} \cdot \mathbf{x}_2 = 0 + 0 = 0$. Therefore $\mathbf{x}_1 + \mathbf{x}_2 \in V^\perp$.

    iii. If $\mathbf{x} \in V^\perp$ and $a \in \mathbb{R}$ then for all $\mathbf{y} \in V$ we have $\mathbf{y} \cdot (a\mathbf{x}) = a(\mathbf{y} \cdot \mathbf{x}) = a0 = 0$. Therefore $a\mathbf{x} \in V^\perp$.

    iv. Therefore $V^\perp$ is a subspace of $\mathbb{R}^n$.

    (d) $V \perp V^\perp$ so $V \cap V^\perp = \{\mathbf{0}\}$.

    (e) **Theorem**: If $V = image\,(B)$ for some orthogonal projection matrix $B$, and if $C = I - B$ then $V^\perp = image\,(C)$.

    i. This tells you how to construct a spanning set for $V^\perp$ from a spanning set for $V$.

    ii. **Proof**: We have already proven (5b) that $image\,(B) \perp image\,(C)$, so $image\,(C) \subset V^\perp$. It remains to show the opposite inclusion. Let $\mathbf{z} \in V^\perp$. I claim that $B\mathbf{z} = \mathbf{0}$. Since the columns of $B$ are vectors in $V$, $col_i\,(B) \cdot \mathbf{z} = 0$ for all $i$. Therefore $B^T\mathbf{z} = \mathbf{0}$. But $B = B^T$ so $B\mathbf{z} = \mathbf{0}$. By (5c) there exists $\mathbf{x} \in image\,(B) = \ker(C)$ and $\mathbf{y} \in image\,(C)$ such that $\mathbf{z} = \mathbf{x} + \mathbf{y}$. I claim $\mathbf{z} = \mathbf{y}$ so $\mathbf{z} \in image\,(C)$. Recall $C\mathbf{x} = \mathbf{0}$, and let $\mathbf{y} = C\mathbf{t}$ for some vector $\mathbf{t}$. Then

$$\begin{aligned}
\mathbf{z} &= (B + C)\,\mathbf{z} \\
&= C\mathbf{z} \text{ because } B\mathbf{z} = \mathbf{0} \\
&= C\,(\mathbf{x} + \mathbf{y}) \\
&= C\mathbf{x} + C\,(C\mathbf{t}) \\
&= C^2\mathbf{t} \\
&= C\mathbf{t} \\
&= \mathbf{y}
\end{aligned}$$

    (f) If $\mathbf{z} \in \mathbb{R}^n$ then there exists unique vectors $\mathbf{x} \in V$, $\mathbf{y} \in V^\perp$ such that $\mathbf{z} = \mathbf{x} + \mathbf{y}$.

    i. The uniqueness of the summands follows from (10a). To prove the existance of $\mathbf{x}$ and $\mathbf{y}$, choose an orthogonal projection matrix $B$ such that $V = image\,(B)$ (7). If $C = I - B$ then $V^\perp = image\,(C)$. By (5c), $\mathbf{z} = \mathbf{x} + \mathbf{y}$ with $\mathbf{x} \in image\,(B) = V$ and $\mathbf{y} \in image\,(C) = V^\perp$.

    (g) $\dim V^\perp = n - \dim V$.

    i. Let $s = \dim V$ and $t = \dim V^\perp$. We will show that $s + t = n$. Choose a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_s\}$ for $V$ and a basis $\{\mathbf{y}_1, \ldots, \mathbf{y}_t\}$ for $V^\perp$. Since $V \cap V^\perp = \{\mathbf{0}\}$, $\{\mathbf{x}_1, \ldots, \mathbf{x}_s, \mathbf{y}_1, \ldots, \mathbf{y}_t\}$ is a linearly independent set. But $\{\mathbf{x}_1, \ldots, \mathbf{x}_s, \mathbf{y}_1, \ldots, \mathbf{y}_t\}$ spans $\mathbb{R}^n$. (Proof: if $\mathbf{z} \in \mathbb{R}^n$ then $\mathbf{z} = \mathbf{x} + \mathbf{y}$ with $\mathbf{x} \in V$ and $\mathbf{y} \in W$. But $\mathbf{x} = a_1\mathbf{x}_1 + \cdots + a_s\mathbf{x}_s$ for some scalars $a_i$, and $\mathbf{y} = b_1\mathbf{y}_1 + \cdots + b_t\mathbf{y}_t$ for some scalars $b_j$. Thus $\mathbf{z} = a_1\mathbf{x}_1 + \cdots + a_s\mathbf{x}_s + b_1\mathbf{y}_1 + \cdots + b_t\mathbf{y}_t$, and $\{\mathbf{x}_1, \ldots, \mathbf{x}_s, \mathbf{y}_1, \ldots, \mathbf{y}_t\}$ spans $\mathbb{R}^n$.) Thus $\{\mathbf{x}_1, \ldots, \mathbf{x}_s, \mathbf{y}_1, \ldots, \mathbf{y}_t\}$ is a basis of $\mathbb{R}^n$ and $s + t = n$.

15. **Theorem:** if $V, W \subset \mathbb{R}^n$ are subspaces and $V \perp W$ then $\dim V + \dim W = n$ if and only if $W = V^\perp$.

(a) **Proof:** since $W \perp V$, $W \subset V^{\perp}$. If $\dim V + \dim W = n$ then $\dim W = n - \dim V = \dim V^{\perp}$. Therefore $W = V^{\perp}$. Conversely, if $W = V^{\perp}$ then $\dim V + \dim W = \dim V + \dim V^{\perp} = n$.

16. **Corollary:** let $V \subset \mathbb{R}^n$ be a subspace. Then $V = V^{\perp\perp}$.

    (a) **Proof:** $\dim V + \dim V^{\perp} = n = \dim V^{\perp} + \dim V^{\perp\perp}$ so $\dim V = \dim V^{\perp\perp}$. If we show $V \subset V^{\perp\perp}$, we are done. If $\mathbf{x} \in V$ then for all $\mathbf{y} \in V^{\perp}$ we have $\mathbf{x} \cdot \mathbf{y} = 0$. Thus $\mathbf{x} \in V^{\perp\perp}$.

    (b) Very important advanced note. This argument cannot be reversed. You cannot show directly that $V^{\perp\perp} \subset V$. In infinite dimensional spaces $V^{\perp\perp}$ can be larger than $V$. Only in finite dimensional spaces does $V = V^{\perp\perp}$, so finite dimensionality is essential for this proof.

    (c) **Alternate proof:** Let $V = image\,(B)$ for some orthogonal projection matrix $B$. Then $V^{\perp} = image\,(I - B)$ and so

$$\begin{aligned} V^{\perp\perp} &= image\,(I - (I - B)) \\ &= image\,(B) \\ &= V \end{aligned}$$

17. **Corollary:** the four fundamental subspaces form complementary pairs. For any matrix $A$, $\ker(A)^{\perp} = image\left(A^T\right)$ and $image\,(A)^{\perp} = \ker\left(A^T\right)$.

    (a) Let $A$ have $n$ columns, so $\ker(A)^{\perp}, image\left(A^T\right) \subset \mathbb{R}^n$. We have already seen that $\ker(A) \perp image\left(A^T\right)$. Moreover $\dim \ker A + \dim image\,(A) = n - rank\,(A) + rank\,(A) = n$, so by the theorem $\ker(A)^{\perp} = image\left(A^T\right)$.