Differential Equations with Linear Algebra

Ben Woodruff[1]

Typeset on January 21, 2013

[1]Mathematics Faculty at Brigham Young University–Idaho, woodruffb@byui.edu

Contents

Chapter 1

Review

This chapter covers the following ideas.

1. Graph basic functions by hand. Compute derivatives and integrals, in particular using the product rule, quotient rule, chain rule, integration by u-substitution, and integration by parts (the tabular method is useful for simplifying notation). Explain how to find a Laplace transform.
2. Explain how to verify a function is a solution to an ODE, and illustrate how to solve separable ODEs.
3. Explain how to use the language of functions in high dimensions and how to compute derivatives using a matrix. Illustrate the chain rule in high dimensions with matrix multiplication.
4. Graph the gradient of a function together with several level curves to illustrate that the gradient is normal to level curves.
5. Explain how to test if a differential form is exact (a vector field is conservative) and how to find a potential.

1.1 Basics

We need to review our ability to graph functions with multiple inputs and/or outputs. The next few problems ask you to practice some skills that will be crucial as the course progresses.

Problem 1.1 Construct graphs of the following functions. Explain how to obtain each graph by transforming and rescaling the first. Then state the amplitude and period of the function.

1. $y = \sin(x)$
2. $y = 5\sin(x) + 1$
3. $y = 4\sin(3(x - \pi)) + 2$
4. $y = 4\sin(3x - \pi) + 2$

Problem 1.2 Consider the function $f(x) = e^{-x}$.

1. Construct graphs of $y = f(x)$ and $y = 2f(-(x + 3)) - 1$.

2. State $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$ from your graph.

3. Compute $\lim_{x\to\infty} xf(x)$ and $\lim_{x\to\infty} x^2 f(x)$. [Hint: L'Hopital's rule will help.]

As the semester progresses, we'll need the functions

$$\cosh x = \frac{e^x + e^{-x}}{2} \qquad \text{and} \sinh x = \frac{e^x - e^{-x}}{2}.$$

These functions are the hyperbolic trig functions, and we say the hyperbolic sine of x when we write $\sinh x$. These functions are very similar to sine and cosine functions, and have very similarly properties.

$\boxed{\textbf{Problem 1.3}}$ Three useful facts about the trig functions are (1) $\frac{d}{dx}\sin x = \cos x$, (2) $\frac{d}{dx}\cos x = -\sin x$, and (3) $\cos^2 x + \sin^2 x = 1$. Use the definitions above to show the following:

1. $\frac{d}{dx}\sinh x = \cosh x$,

2. $\frac{d}{dx}\cosh x = \sinh x$, and

3. $\cosh^2 x - \sinh^2 x = 1$.

[Hint: Start by replacing the hyperbolic function with its definition in terms of exponentials. Then perform the computations.]

$\boxed{\textbf{Problem 1.4}}$ The three facts from the previous problem are crucial tools need to prove that $\frac{d}{dx}\tan x = \sec^2 x$.

1. Use the quotient rule to give a formula for $\frac{d}{dx}\tanh x$ in terms of hyperbolic trig functions.

2. Similarly obtain a formula for the derivative of $\operatorname{sech} x = \frac{1}{\cosh x}$.

3. What is $\frac{d}{dx}\operatorname{csch} x$?

You might ask why these function are called the hyperbolic trig functions. What does a hyperbola have to do with anything?

$\boxed{\textbf{Problem 1.5}}$ Each pair of parametric equations traces out a curve in the xy plane. Given a Cartesian equation of the curve by eliminating the parameter t, and then graph the curve.

1. $x = \cos t$, $y = \sin t$, $-2\pi < t < 2\pi$.

2. $x = \cosh t$, $y = \sinh t$, $-\infty < t < \infty$.

Give a reason as to why do we call cosh the hyperbolic cosine.

$\boxed{\textbf{Problem 1.6}}$ Use implicit differentiation to find the derivative of $y = sinh^{-1}x$. Your answer should not involve any hyperbolic trig functions, and should be in terms of x. [Hint: First write $x = sinh(y)$, and then implicitly differentiate both sides. You'll need the key identity from a few problems above to help you finish.]

The problems above asked you to review your differentiation skills. You'll want to make sure you can use the basic rules of differentiation (such as the power, product, quotient, and chain rules). The next few problems will help you review your integration techniques, and you will apply them to two new ideas.

Problem 1.7 Compute the three integrals

$$\int xe^{-x^2}dx \qquad \text{and} \qquad \int_0^1 xe^{-x^2}dx \qquad \text{and} \qquad \int_0^\infty xe^{-x^2}dx.$$

If you have never used the tabular method to perform integration-by-parts, I strongly suggest that you open the online text and read a few examples (see the bottom of page 2).

Problem 1.8 Compute $\int x\sin(5x)dx$ and $\int x^2\sin(5x)dx$.

Problem 1.9 Compute $\int \tanh^{-1} x dx$. The derivative of $\tanh^{-1} x$ is $\dfrac{1}{1-x^2}$.

1.2 Laplace Transforms

Definition 1.1: The Laplace Transform. Let $f(t)$ be a function that is defined for all $t \geq 0$. Using the function $f(t)$, we define the Laplace transform of f to be a few function F where for each s we obtain the value by computing the integral

$$F(s) = \mathscr{L}\{f(t)\} = \int_0^\infty e^{-st}f(t)dt.$$

The domain of F is the set of all s such that the improper integral above converges. The function $f(t)$ is called the inverse Laplace transform of $F(s)$, and we write $f(t) = \mathscr{L}^{-1}(F(s))$.

Note that the Laplace transform of a function with independent variable t is another function with a different independent variable s. After integration, all t's will be removed from $F(s)$. You can of course use any other letters besides t and s.

We will use the Laplace transform throughout the semester to help us solve many problems related to mechanical systems, electrical networks, and more. The mechanical and electrical engineers in this course will use Laplace transforms in many future courses. Our goal in the problems that follow is to practice integration-by-parts. As an extra bonus, we'll learn the Laplace transforms of some basic functions.

Problem 1.10 Compute the integral $\int_0^\infty e^{-st}dt$, and state for which s the integral converges. What is the Laplace transform of $f(t) = 1$? (If the last question seems redundant, then horray.)

Problem 1.11 Compute the Laplace transform of $f(t) = e^{2t}$, and state the domain. Then compute the Laplace transform of $f(t) = e^{3t}$ and state the domain. Finally, compute the Laplace transform of $f(t) = e^{at}$ for any a, and state the domain.

Problem 1.12 Suppose $s > 0$ and n is a positive integer. Explain why

$$\lim_{t\to\infty} \frac{t^n}{e^{st}} = 0.$$

Use this fact to prove that the Laplace transform of t^2 is

$$\mathscr{L}\{t^2\}\frac{2}{s^3}.$$

[You'll need to do integration-by-parts twice, try the tabular method.]

Problem 1.13 In the previous problems, you showed that

$$\mathscr{L}\{t^0\} = \frac{1}{s^1} \qquad \text{and} \qquad \mathscr{L}\{t^2\} = \frac{2}{s^3}.$$

Show that the Laplace transform of t is $\mathscr{L}\{t^1\} = \dfrac{1}{s^2}$. Then compute the Laplace transforms of t^3, t^4, and so on until you see a pattern. Use this pattern to state the Laplace transform of t^10 and t^n, provided n is a positive integer. [Hint: Try the tabular method of integration-by-parts. After evaluating at 0 and ∞, all terms but 1 will be zero.]

Theorem 1.2. *Since integration can be done term-by-term, and constants can be pulled out of the integral, we have the crucial fact that*

$$\mathscr{L}\{af(t) + bg(t)\} = a\mathscr{L}\{f(t)\} + b\mathscr{L}\{g(t)\}$$

for functions f, g and constants a, b.

Problem 1.14 Without integrating, rather using the results above, compute the Laplace transform $L(3 + 5t^2 - 6e^{8t})$, and state the domain.

Problem 1.15 Recall that $\cosh t = \dfrac{e^t + e^{-t}}{2}$ and $\sinh t = \dfrac{e^t - e^{-t}}{2}$. Use this to prove that

$$\mathscr{L}\{\cosh at\} = \frac{s}{s^2 - a^2} \qquad \text{and} \qquad \mathscr{L}\{\sinh at\} = \frac{a}{s^2 - a^2}.$$

1.3 Ordinary Differential Equations

A differential equation is an equation which involves derivatives (of any order) of some function. For example, the equation $y'' + xy' + \sin(xy) = xy^2$ is a differential equation. An **ordinary differential equation (ODE)** is a differential equation involving an unknown function y which depends on only one independent variable (often x or t). A partial differential equation involves an unknown function y that depends on more than one variable (such as $y(x, t)$). The order of an ODE is the order of the highest derivative in the ODE. A solution to an ODE on an interval (a, b) is a function $y(x)$ which satisfies the ODE on (a, b).

Example 1.3. The first order ODE $y'(x) = 2x$, or just $y' = 2x$, has unknown function y with independent variable x. A solution on $(-\infty, \infty)$ is the function $y = x^2 + C$ for any constant C. We obtain this solution by simply integrating both sides. Notice that there are infinitely many solutions to this ODE.

Typically a solution to an ODE involves an arbitrary constant C. There is often an entire family of curves which satisfy a differential equation, and the constant C just tells us which curve to pick. A **general solution** of an ODE is an infinite class of solutions of the ODE. A **particular solution** is one of the infinitely many solutions of an ODE.

Often an ODE comes with an **initial condition** $y(x_0) = y_0$ for some values x_0 and y_0. We can use these initial conditions to find a particular solution of the ODE. An ODE, together with an initial condition, is called an **initial value problem (IVP)**.

Example 1.4. The IVP $y' = 2x$, $y(2) = 1$, has the general solution $y = x^2 + C$ from the previous problem. Since $y = 1$ when $x = 2$, we have $1 = 2^2 + C$ which means $C = -3$. Hence the solution to our IVP is $y = x^2 - 3$.

Problem 1.16 Consider the ordinary differential equation $y'' + 9y = 0$. By computing derivatives, show that $y(t) = A\cos(3t) + B\sin(3t)$ is a general solution to the ODE, where A and B are arbitrary constants. If we know that $y(0) = 1$ and $y'(0) = 2$, determine the values of A and B.

Problem 1.17 Consider the ordinary differential equation $y\dfrac{dy}{dx} = x^2$. Find a general solution to this ODE by integrating both sides with respect to x. State an interval on which your solution is valid.

They could introduce the entire method of separation by parts without me telling them what to do. I just need to ask them to do an integral. Afterward, I could ask them to solve an ODE. Put it in the same problem.

Problem 1.18 Consider the ODE given by $y' = 4ty$. Find a general solution to this ODE. [Hint: Rewrite y' as $\dfrac{dy}{dt}$. Then put all the terms that involve y on one side of the equation, and the terms that involve t on the other. Then it should be similar to the previous problem.]

Problem 1.19 Solve the IVP given by $y' = \dfrac{x^2 - 1}{y^4 + 1}$, where $y(0) = 1$.

1.4 General Functions and Derivatives

Recall that to compute partial derivatives, we hold all but one variable constant and then differentiate with respect to that variable. Partial derivatives can be organized into a matrix Df where columns represents the partial derivative of f with respect to each variable. This matrix, called the derivative or total derivative, takes us into our study of linear algebra. Some examples of functions and their derivatives appear in Table 1.1. When the output dimension is one, the matrix has only one row and the derivative is often called the gradient of f, written ∇f.

In multivariate calculus, we focused our time on learning to graph, differentiate, and analyze each of the types of functions in the table above. The next few problems ask you to review this.

Function	Derivative
$f(x)=x^2$	$Df(x)=[2x]$
$\vec{r}(t)=(3\cos(t),2\sin(t))$	$D\vec{r}(t)=\begin{bmatrix}-3\sin t\\ 2\cos t\end{bmatrix}$
$\vec{r}(t)=(\cos(t),\sin(t),t)$	$D\vec{r}(t)=\begin{bmatrix}-\sin t\\ \cos t\\ 1\end{bmatrix}$
$f(x,y)=9-x^2-y^2$	$Df(x,y)=\nabla f(x,y)=[-2x \quad -2y]$
$f(x,y,z)=x^2+y+xz^2$	$Df(x,y,z)=\nabla f(x,y,z)=\begin{bmatrix}2x+z^2 & 1 & 2xz\end{bmatrix}$
$\vec{F}(x,y)=(-y,x)$	$D\vec{F}(x,y)=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$
$\vec{F}(r,\theta,z)=(r\cos\theta,r\sin\theta,z)$	$D\vec{F}(r,\theta,z)=\begin{bmatrix}\cos\theta & -r\sin\theta & 0\\ \sin\theta & r\cos\theta & 0\\ 0 & 0 & 1\end{bmatrix}$
$\vec{r}(u,v)=(u,v,9-u^2-v^2)$	$D\vec{r}(u,v)=\begin{bmatrix}1 & 0\\ 0 & 1\\ -2u & -2v\end{bmatrix}$

Table 1.1: The table above shows the (matrix) derivative of various functions. Each column of the matrix corresponds a partial derivative of the function. When the output of a function is a vector, partial derivatives are vectors which are placed in columns of the matrix. The order of the columns matches the order in which you list the variables.

Problem 1.20 Let $\vec{r}(t)=\left\langle t^2-1,2t+3\right\rangle$. Construct a graph of $\vec{r}(t)$, and compute the derivative $D\vec{r}(t)$.

Problem 1.21 Let $f(x,y)=4-x^-y^2$. Construct a 3D graph of $z=f(x,y)$. Also construct a graph of several level curves. Then compute the derivative $Df(x,y)$.

Recall that a level curve of $z=f(x,y)$ is curve in the xy plane where the output z is constant.

Problem 1.22 Let $\vec{r}(t)=\langle 3\cos t,2\sin t,t\rangle$. Construct a 3D graph of $\vec{r}(t)$, and compute the derivative $D\vec{r}(t)$.

Problem 1.23 Let $\vec{F}(x,y)=(y,-2x)$. Construct a 2D graph of this vector field, and compute the derivative $D\vec{F}(x,y)$.

1.4.1 The General Chain Rule

The chain rule in first semester calculus states that

$$(f\circ g)'(x)=f'(g(x))g'(x).$$

You may remember this as "the derivative of the outside function times the derivative of the inside function." In multivariable calculus, most textbooks use a tree rule to develop the formula

$$\frac{df}{dt} = f_x x_t + f_y y_t$$

for a function $f(x, y)$, where x and y depend on t (so that $\vec{r}(t) = (x(t), y(t))$ is a curve in the xy plane). Written in matrix form, the chain rule is simply

$$\frac{df}{dt} = \begin{bmatrix} f_x & f_y \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} = Df \cdot Dr,$$

which is the (matrix) product of the derivatives, just as it was in first semester calculus. You are welcome to tackle the following problems by using the tree rule or matrix product.

Problem 1.24 Suppose that $f(x, y) = x^2 + 3xy$, where $x = t^2 + 1$ and $y = \sin t$, so we could write $\vec{r}(t) = (t^2 + 1, \sin t)$.

1. Compute $Df(x, y)$, $\dfrac{dx}{dt}$, and $D\vec{r}(t)$. (You should have two matrices.)
2. Compute $\frac{df}{dt}$.

Problem 1.25 Suppose that $f(x, y) = x + 3y$ and that $\dfrac{dx}{dt} = \cos t$ and $\dfrac{dy}{dt} = e^t$. Compute $\frac{df}{dt}$.

Problem 1.26 Suppose that $z = f(x, y)$ and that $\dfrac{\partial f}{\partial x} = 3x^2 y$ and $\dfrac{\partial f}{\partial y} = x^3 y - e^y$. Also suppose that $x = \sqrt{t}$ and $y = \ln t$. Compute $\frac{df}{dt}$.

Problem 1.27 Suppose that $z = f(x, y)$ is a differential function of two variables. Suppose that $\vec{r}(t)$ is a parametrization of a level curve of f. We can write the level curve in vector form as $\vec{r}(t) = (x(t), y(t))$, or in parametric form $x = x(t)$ and $y = y(t)$.

1. If $f(\vec{r}(0)) = 7$, then what is $f(\vec{r}(2))$?
2. Why does $\frac{df}{dt} = \nabla f(x, y) \cdot \dfrac{d\vec{r}}{dt}$?
3. Why is the gradient of f normal to level curves?

Recall that the word normal means there is a 90 degree angle between the gradient and the level curve.

Before proceeding, let's practice with an examples to visually remind us that the gradient is normal to level curves. This key fact will help us solve most of the differential equations we encounter in the course.

Problem 1.28 Consider the function $f(x, y) = x^2 - y$. Start by computing the gradient. Then construct a graph which contains several level curves of f, as well as the gradient at several points on each level curve.

1.5 Potentials of Vector Fields and Differential Forms

When the output dimension of a function is one, so we would write $f : \mathbb{R}^n \to \mathbb{R}^1$, then we call the derivative the gradient and write $\vec{\nabla} f = (f_x, f_y, f_z)$. Notice that this is a vector field. Taking a derivative gives us a vector field. Is every vector field the derivative of some function? Hopefully you remember that the answer to this question is "No."

If a vector field $\vec{F} = (M, N)$ (or in 3D $\vec{F} = (M, N, P)$) is the gradient of some some function f (so that $\vec{\nabla} f = \vec{F}$), then we say that the vector field $\vec{F}$ is a gradient field (or conservative vector field). We say that f is a potential for the vector field $\vec{F}$ when $\nabla f = \vec{F}$. In this section, we'll review how to determine if a vector field has a potential, as well as how to find a potential.

Problem 1.29 Let $\vec{F} = (M, N) = (2x + y, x + 4y)$. Find a potential for $\vec{F}$ by doing the following.

1. If we suppose $M = 2x + y$ is the partial of f with respect to x, then $f_x = 2x + y$. Find a function f whose partial with respect to x is M.
2. If we suppose $N = x + 4y$ is the partial of f with respect to y, then $f_y = x + 4y$. Find a function f whose partial with respect to y is N.
3. What is a potential for $\vec{F}$? Prove your answer is correct by computing the gradient of your answer.

By taking derivatives, there is a test that tells you if a function will have a potential. Some textbooks call it the test for a conservative field.

Problem 1.30: Test for a conservative vector field. Let's prove the test for a conservative vector field in both 2 and 3 dimensions.

1. Suppose that $\vec{F}(x, y) = (M, N)$ is a continuously differentiable vector field on the entire plane. Suppose further that $\vec{F}$ has a potential f. The derivative of $\vec{F}$ is
$$D\vec{F}(x, y) = \begin{pmatrix} M_x & M_y \\ N_x & N_y \end{pmatrix}.$$
Some of the entries in this matrix must be equal? Which ones? Explain. [If you're not sure, try taking the derivative of the problem above.]
2. Suppose that $\vec{F}(x, y, z) = (M, N, P)$ is a continuously differentiable vector field on all of space. Suppose further that $\vec{F}$ has a potential f. State the derivative of $\vec{F}$, and then state which pairs of entries must be equal.

The test for a conservative vector field states states more than what you showed in this problem. It states that if $\vec{F}$ is a continuously differentiable vector field on a simply connected domain, then (1) if $\vec{F}$ has potential, then certain pairs of partials must be equal, and (2) if those pairs of partial derivatives are equal, then the $\vec{F}$ has a potential. We will not prove part (2).

Problem 1.31 For each vector field below, either give a potential, or explain why no potential exists.

1. $\vec{F} = (4x + 5y, 5x + 6y)$
2. $\vec{F} = (2x - y, x + 3y)$
3. $\vec{F} = \left(4x + \dfrac{2y}{1 + 4x^2}, \arctan(2x)\right)$
4. $\vec{F} = (3y + 2yz, 3x + 2xz + 6z, 2xy + 6y)$

We'll finish by introducing the vocabulary of differential forms. We'll use this vocabulary throughout the semester as we study differential equations. The vocabulary of vector fields parallels the vocabulary of differential forms.

Definition 1.5: Differential Forms. Assume that f, M, N, P are all functions of three variables x, y, z. Similar definitions hold in all dimensions.

- A differential form is an expression of the form $Mdx + Ndy + Pdz$ (just as a vector field is a function $\vec{F} = (M, N, P)$).
- The differential of a function f is the expression $df = f_x dx + f_y dy + f_z dz$ (just as the gradient is $\vec{\nabla} F = (f_x, f_y, f_z)$).
- If a differential form is the differential of a function f, then the differential form is said to be exact (just as we say a vector field is a gradient field). Again, the function f is called a potential for the differential form.

A differential form is exact precisely when the corresponding vector field is a gradient field.

Notice that $Mdx + Ndy + Pdz$ is exact if and only if $\vec{F} = (M, N, P)$ is a gradient field. The language of differential forms is practically the same as the language of conservative vector fields. Why do we have different sets of words for the same idea? That happens all the time when different groups of people work on seeming different problems, only to discover years later that they have been working on the same problem. If both sets of vocabulary stick, it's often because both have advantages. We have many different notations for the derivative (such as y', $\frac{dy}{dx}$, and Df), and each notation has advantages. The language of differential forms is best suited when studying differential equations.

Problem 1.32 For each differential form below, state if the differential form is exact. If it is exact, give a potential.

1. $(2x + 3y)dx + (4x + 5y)dy$
2. $(2x - y)dx + (3y - x)dy$
3. $\left(4x + \dfrac{3y}{1 + 9x^2}\right) dx + \arctan(3x)dy$
4. $(3y + 2yz)dx + (3x + 2xz + 6z)dy + (2xy + 5y)dz$

1.6 Wrap Up

Once you have finished the problems in the section and feel comfortable with the ideas, create a short one page lesson plan that contains examples of the key ideas. You will get a chance to teach from this lesson plan prior to taking the exam. Then log on to I-Learn and download the quiz. Once you have taken the quiz, upload your work to I-Learn and then download the key to see how you did. If you still need to work on mastering some of the ideas, please do so and then demonstrate your mastery though the quiz corrections.

Chapter 2

Linear Algebra Arithmetic

This chapter covers the following ideas.

1. Be able to use and understand matrix and vector notation, addition, scalar multiplication, the dot product, matrix multiplication, and matrix transposing.
2. Use Gaussian elimination to solve systems of linear equations. Define and use the words homogeneous, nonhomogeneous, row echelon form, and reduced row echelon form.
3. Find the rank of a matrix. Determine if a collection of vectors is linearly independent. If linearly dependent, be able to write vectors as linear combinations of the preceding vectors.
4. For square matrices, compute determinants, inverses, eigenvalues, and eigenvectors.
5. Illustrate with examples how a nonzero determinant is equivalent to having independent columns, an inverse, and nonzero eigenvalues. Similarly a zero determinant is equivalent to having dependent columns, no inverse, and a zero eigenvalue.

The next unit will focus on applications of these ideas. The main goal of this unit is to familiarize yourself with the arithmetic involved in linear algebra.

2.1 Basic Notation

Most of linear algebra centers around understanding vectors, with matrices being functions which transform vectors from one vector space into vectors in another vector space. This chapter contains a brief introduction to the arithmetic involved with matrices and vectors. The next chapter will show you many of the uses of the ideas we are learning. You will be given motivation for all of the ideas learned here, as well as real world applications of these ideas, before the end of the next chapter. For now, I want you become familiar with the arithmetic of linear algebra so that we can discuss how all of the ideas in this chapter show up throughout the course.

Definition 2.1. A matrix of size m by n has m rows and n columns. We

Matrix size is row by column.

normally write matrices using capital letters, and use the notation

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = [a_{jk}],$$

where a_{jk} is the entry in the jth row, kth column.

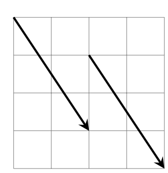
- We say two matrices A and B are equal if $a_{jk} = b_{jk}$ for all j and k.
- We add and subtract matrices of the same size entry wise. So we write $A + B = C$ where $c_{jk} = a_{jk} + b_{jk}$. If matrices do not have the same size, then we cannot add them.
- We can multiply a matrix A by a scalar C to obtain a new matrix cA. We do this multiplying every entry in the matrix A by the scalar c.
- If the number of rows and columns are equal, then we say the matrix is square.
- The main diagonal of a square $(n \times n)$ matrix consists of the entries $a_{11}, a_{22}, \ldots, a_{nn}$.
- The trace of a square matrix is the sum of the entries on the main diagonal $(\sum a_{jj})$.
- The transpose of a matrix $A = [a_{jk}]$ is a new matrix $B = A^T$ formed by interchanging the rows and columns of A, so that $b_{jk} = a_{kj}$. If $A^T = A$, then we say that A is symmetric.

Problem 2.1 Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -1 \\ 0 & 4 \end{bmatrix}$. Compute $2A - 3B$, and find the trace of both A and B.

Problem 2.2 Write down a 3 by 2 matrix, and compute the transpose of that matrix. Then give an example of a 3 by 2 symmetric matrix, or explain why it is not possible.

Vectors represent a magnitude in a given direction. We can use vectors to model forces, acceleration, velocity, probabilities, electronic data, and more. We can use matrices to represent vectors. A row vector is a $1 \times n$ matrix. A column vector is an $m \times 1$ matrix. Textbooks often write vectors using bold face font. By hand (and in this book) we add an arrow above them. The notation $\mathbf{v} = \vec{v} = \langle v_1, v_2, v_3 \rangle$ can represent either row or column vectors. Many different ways to represent vectors are used throughout different books. In particular, we can represent the vector $\langle 2, 3 \rangle$ in any of the following forms

$$\langle 2, 3 \rangle = 2\mathbf{i} + 3\mathbf{j} = (2, 3) = \begin{bmatrix} 2 & 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{pmatrix} 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$



Both vectors represent $\langle 2, -3 \rangle$, regardless of where we start.

The notation $(2, 3)$ has other meanings as well (like a point in the plane, or an open interval), and so when you use the notation $(2, 3)$, it should be clear from the context that you are working with a vector. To draw a vector $\langle v_1, v_2 \rangle$, one option is to draw an arrow from the origin (the tail) to the point (v_1, v_2) (the head). However, the tail does not have to be placed at the origin.

The principles of addition and subtraction of matrices apply to vectors (which can be though of as row or column matrices). We will most often think of vectors as column vectors.

Definition 2.2. The magnitude (or length) of the vector $\vec{u} = (u_1, u_2)$ is $|\vec{u}| = \sqrt{u_1^2 + u_2^2}$. In higher dimensions we extend this as

$$|\vec{u}| = \sqrt{u_1^2 + u_2^2 + u_3^2 + \cdots u_n^2} = \sqrt{\sum_{i=1}^{n} u_i^2}.$$

A unit vector is a vector with length 1. In many books unit vectors are written with a hat above them, as $\hat{\mathbf{u}}$.

A unit vector $\hat{\mathbf{u}}$ has length $|\vec{u}| = 1$

We will need to be able to find vectors of any length that point in a given direction.

Problem 2.3 Find a vector of length 12 that points in the same direction as the vector $\vec{v} = (1, 2, 3, 4)$. Then give a general formula for finding a vector of length c that points in the direction of $\vec{u}$.

The simplest vectors in 2D are a one unit increment in either the x or y direction, and we write these vectors in any of the equivalent forms

$$\mathbf{i} = \vec{i} = \langle 1, 0 \rangle = (1, 0) \quad \text{and} \quad \mathbf{j} = \vec{j} = \langle 0, 1 \rangle = (0, 1).$$

We call these the standard basis vectors in 2D. In 3D we include the vector $\mathbf{k} = \vec{j} = \langle 0, 0, 1 \rangle$ as well as add a zero to both $\vec{i}$ and $\vec{j}$ to obtain the standard basis vectors. The word basis suggests that we can base other vectors on these basis vectors, and we typically write other vectors in terms of these standard basis vectors. Using only scalar multiplication and vector addition, we can obtain the other vectors in 2D from the standard basis vectors.

The standard basis vectors in 3D
$\mathbf{i} = \vec{i} = \langle 1, 0, 0 \rangle = (1, 0, 0)$
$\mathbf{j} = \vec{j} = \langle 0, 1, 0 \rangle = (0, 1, 0)$
$\mathbf{k} = \vec{k} = \langle 0, 0, 1 \rangle = (0, 0, 1)$

Problem 2.4 Write the vector $(2, 3)$ in the form $(2, 3) = c_1\vec{i} + c_2\vec{j}$.

If instead we use the non-standard basis vectors $\vec{u}_1 = (1, 2)$ and $\vec{u}_2 = (-1, 4)$, then write the vector $(2, 3)$ in the form $(2, 3) = c_1\vec{u}_1 + c_2\vec{u}_2$.

Definition 2.3. A linear combination of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is an expression of the form $c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_n\vec{v}_n$, where c_i is a constant for each i.

A linear combination of vectors is simply a sum of scalar multiples of the vectors. We start with some vectors, stretch each one by some scalar, and then sum the result. Much of what we will do this semester (and in many courses to come) relates directly to understanding linear combinations.

Problem 2.5 The force acting on an object is $\vec{F} = (-3, 2)$ N. The object is in motion and has velocity vector $\vec{v} = (1, 1)$ and acceleration vector $\vec{a} = (-1, 2)$. Write the force as a linear combination of the velocity and acceleration vectors.

Problem 2.6 Write the vector $(2, 3, 1)$ as a linear combination of the standard basis vectors in $\mathbb{R}^3$. Then write $(2, 3, 1)$ as a linear combination of the vectors $(1, 0, 0)$, $(1, 1, 0)$, and $(1, 1, 1)$.

One of the key applications of linear combinations we will make throughout the semester is matrix multiplication. Let's introduce the idea with an example.

Example 2.4. Consider the three vectors $\begin{bmatrix}1\\0\\1\end{bmatrix}$, $\begin{bmatrix}0\\2\\-3\end{bmatrix}$, and $\begin{bmatrix}2\\1\\1\end{bmatrix}$. Let's multiply the first vector by 2, the second by -1, and the third by 4, and then sum the result. This gives us the linear combination

$$2\begin{bmatrix}1\\0\\1\end{bmatrix} - 1\begin{bmatrix}0\\2\\-3\end{bmatrix} + 4\begin{bmatrix}2\\1\\1\end{bmatrix} = \begin{bmatrix}10\\2\\9\end{bmatrix}$$

We will define matrix multiplication so that multiplying a matrix on the right by a vector corresponds precisely to creating a linear combination of the columns of A. We now write the linear combination above in matrix form

$$\begin{bmatrix}1 & 0 & 2\\0 & 2 & 1\\1 & -3 & 1\end{bmatrix}\begin{bmatrix}2\\-1\\4\end{bmatrix} = \begin{bmatrix}10\\2\\9\end{bmatrix}.$$

Definition 2.5: A matrix times a vector. We define the matrix product $A\vec{x}$ (a matrix times a vector) to be the linear combination of columns of A where the components of $\vec{x}$ are the scalars in the linear combination. For this to make sense, notice that the vector $\vec{x}$ must have the same number of entries as there are columns in A. We can make this definition more precise as follows. Let $\vec{v}_i$ be the ith column of A so that $A = \begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n\end{bmatrix}$, and let $\vec{x} = \begin{bmatrix}x_1\\x_2\\\cdots\\x_n\end{bmatrix}$.
Then the matrix product is the linear combination

The product $A\vec{x}$ gives us linear combinations of the columns of A.

$$A\vec{x} = \begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\cdots\\x_n\end{bmatrix} = \vec{a}_1x_1 + \vec{a}_2x_2 + \cdots + \vec{a}_nx_n.$$

The definition above should look like the dot product. If you think of A as a vector of vectors, then $A\vec{x}$ is just the dot product of A and $\vec{x}$.

Problem 2.7 Write down a 2 by 4 nonzero matrix, and call it A (fill the matrix with some integers of your choice). Then write down a vector $\vec{x}$ such that the matrix product $A\vec{x}$ makes sense (again, fill the vector with integers of your choice). Then use the definition above to obtain the product $A\vec{x}$.

Definition 2.6: A matrix times a matrix. Let $\vec{b}_j$ represent the jth column of B (so $B = \begin{bmatrix}\vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_n\end{bmatrix}$). The product AB of two matrices $A_{m\times n}$ and $B_{n\times p}$ is a new matrix $C_{m\times p} = [c_{ij}]$ where the jth column of C is the product $A\vec{b}_j$. To summarize, the matrix product AB is a new matrix whose jth column is a linear combinations of the columns of A using the entries of the jth column of B to perform the linear combinations.

Problem 2.8 Let $A = \begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix}$ and $B = \begin{bmatrix}0 & 1 & 4\\-1 & 2 & -3\end{bmatrix}$. Use the definition given above to compute both AB and BA. Be prepared to show the class how you used linear combinations to get the matrix product. (If you are used to using the row dotted by column approach, then this problem asks you to do the matrix product differently.)

We introduced matrix multiplication in terms of linear combinations of column vectors. My hope is that by doing so you immediately start thinking of linear combinations whenever you encounter matrix multiplication (as this is what it was invented to do). There are many alternate ways to think of matrix multiplication. Here are two additional methods.

1. "Row times column approach." The product AB of two matrices $A_{m\times n}$ and $B_{n\times p}$ is a new matrix $C_{m\times p} = [c_{ij}]$ where $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$ is the dot product of the of the ith row of A and the jth column of B. Wikipedia has an excellent visual illustration of this approach.

2. Rephrase everything in terms of rows (instead of columns). We form linear combinations of rows using rows. The matrix product $\vec{x}B$ (notice the order is flopped) is a linear combination of the rows of B using the components of x as the scalars. For the product AB, let $\vec{a}_i$ represent the ith row of A. Then the ith row of AB is the product $\vec{a}_iB$. We'll most often use the column definition instead of this, because we use the function notation $f(x)$ from calculus, and later we will use the notation $A(\vec{x})$ instead of $(\vec{x})A$ to describe how matrices act as functions.

Problem 2.9 Let $A = \begin{bmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{bmatrix}$ and $B = \begin{bmatrix}0 & 1 & 4\\ -1 & 2 & -3\end{bmatrix}$. Use the two alternate definitions above to compute AB. Be prepared to show the class how you used both alternate definitions (You'll need to show your intermediate steps).

Problem 2.10 Do each of the following:

1. Solve the system of equations $x + 2y = 3$, $4x + 5y = 6$.
2. Write the vector $\begin{pmatrix}3\\6\end{pmatrix}$ as a linear combination of $\begin{pmatrix}1\\4\end{pmatrix}$ and $\begin{pmatrix}2\\5\end{pmatrix}$.
3. Let $A = \begin{pmatrix}1 & 2\\ 4 & 5\end{pmatrix}$ and $\vec{b} = \begin{pmatrix}3\\6\end{pmatrix}$. Find a vector $\vec{x}$ so that $A\vec{x} = \vec{b}$. This matrix A is called the coefficient matrix of the system in the first part.

How are these three questions related?

Prior to introducing Gaussian elimination, let's solve a system of equations using an elimination method. If $2x + 3y = 4$ and $5x + 7y = 0$, then we can eliminate x from the second equation by multiplying both sides of the first equation by 5, and both sides of the second equation by 2, and then subtracting. This would give us the equations $10x + 15y = 20$ and $10x + 14y = 1$. The first equation minus the second then gives $(10 - 10)x + (15 - 14)y = (20 - 0)$, or more simply $y = 20$. Similarly, you could multiply the first equation by 7, and the second by 3, to eliminate y.

Problem 2.11 Solve the system of equations

$$\begin{aligned}2x + 3y - 4z &= 4\\ 3x + 4y - 3z &= 8\\ 7x + 12y - 12z &= 19.\end{aligned}$$

Use elimination to find your solution. Eliminate x from the 2nd and 3rd equations (which will give you two equations that do not involve x). Then use one of these simplified equations to eliminate y from the other simplified equation. At this point you should have an equation that only involves z. Then use back substitution to give y and x.

Problem 2.12 Answer the following.

1. Suppose that $ax + by = c$ and $dx + ey = f$, where a, b, c, d, e, f are all constants. This is a system of equations with 2 equations and 2 unknowns. Each equation represents a line in the plane. How many solutions are there to this system? (You should have a few different cases.)

2. Suppose that $a_{11}x + a_{12}y + a_{13}z = b_1$, $a_{21}x + a_{22}y + a_{23}z = b_2$ and $a_{31}x + a_{32}y + a_{33}z = b_1$, where each a_{ij} is a constant. This is a system of equations with 3 equations and 3 unknowns. Each equation represents a plane in space. How many solutions are there to this system? (You should have a few different cases.)

3. Suppose that $a_{11}x + a_{12}y + a_{13}z = b_1$ and $a_{21}x + a_{22}y + a_{23}z = b_2$, where each a_{ij} is a constant. This is a system of equations with 2 equations and 3 unknowns. Each equation represents a plane in space. How many solutions are there to this system? (You should have a few different cases.)

Definition 2.7. We say that a system of linear equation is consistent, if it has at least one solution. We say it is inconsistent if there is no solution.

2.2 Gaussian Elimination

Gaussian elimination is an efficient algorithm we will use to solve systems of equations. This is the same algorithm implemented on most computers systems. The main idea is to eliminate each variable from all but one equation/row (if possible), using the following three operations (called elementary row operations):

1. Multiply an equation (or row of a matrix) by a nonzero constant,

2. Add a nonzero multiple of any equation (or row) to another equation,

3. Interchange two equations (or rows).

These three operations are the operations learned in college algebra when solving a system using a method of elimination. Gaussian elimination streamlines elimination methods to solve generic systems of equations of any size. The process involves a forward reduction and (optionally) a backward reduction. The forward reduction creates zeros in the lower left corner of the matrix. The backward reduction puts zeros in the upper right corner of the matrix. We eliminate the variables in the lower left corner of the matrix, starting with column 1, then column 2, and proceed column by column until all variables which can be eliminated (made zero) have been eliminated. Before formally stating the algorithm, let's look at a few examples.

Example 2.8. Let's start with a system of 2 equations and 2 unknowns. I will write the augmented matrix representing the system as we proceed. To solve

$$\begin{aligned} x_1 - 3x_2 &= 4 \\ 2x_1 - 5x_2 &= 1 \end{aligned} \qquad \left[\begin{array}{cc|c} 1 & -3 & 4 \\ 2 & -5 & 1 \end{array}\right]$$

we eliminate the $2x_1$ in the 2nd row by adding -2 times the first row to the second row.

$$\begin{aligned} x_1 - 3x_2 &= 4 \\ x_2 &= -7 \end{aligned} \qquad \left[\begin{array}{cc|c} 1 & -3 & 4 \\ 0 & 1 & -7 \end{array}\right]$$

The matrix at the right is said to be in **row echelon form**. row echelon form

Definition 2.9: Row Echelon Form. We say a matrix is in row echelon form (ref) if

- each nonzero row begins with a 1 (called a leading 1),
- the leading 1 in a row occurs further right than a leading 1 in the row above, and
- any rows of all zeros appear at the bottom.

The position in the matrix where the leading 1 occurs is called a pivot. The column containing a pivot is called a pivot column. pivot column

At this point in our example, we can use "back-substitution" to get $x_2 = -7$ and $x_1 = 4+3x_2 = 4-21 = -17$. Alternatively, we can continue the elimination process by eliminating the terms above each pivot, starting on the right and working backwards. This will result in a matrix where all the pivot columns contain all zeros except for the pivot. If we add 3 times the second row to the first row, we obtain.

$$\begin{aligned} x_1 &= -17 \\ x_2 &= -7 \end{aligned} \qquad \left[\begin{array}{cc|c} 1 & 0 & -17 \\ 0 & 1 & -7 \end{array}\right]$$

The matrix on the right is said to be in **reduced row echelon form** (or just rref). We can easily read solutions to systems of equations directly from a matrix which is in reduced row echelon form.

Definition 2.10: Reduced Row Echelon Form. We say that a matrix is in reduced row echelon form (rref) if reduced row echelon form - rref

- the matrix is in row echelon form, and
- each pivot column contains all zeros except for the pivot (leading one).

Example 2.11. Let's now solve a nonhomogeneous (meaning the right side is not zero) system with 3 equations and 3 unknowns:

$$\begin{aligned} 2x_1 + x_2 - x_3 &= 2 \\ x_1 - 2x_2 &= 3 \\ 4x_2 + 2x_3 &= 1 \end{aligned} \qquad \left[\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 1 & -2 & 0 & 3 \\ 0 & 4 & 2 & 1 \end{array}\right].$$

We'll encounter some homogeneous systems later on. To simplify the writing, we'll just use matrices this time. To keep track of each step, I will write the row operation next to the row I will replace. Remember that the 3 operations are (1)multiply a row by a nonzero constant, (2)add a multiple of one row to another, (3) interchange any two rows. If I write $R_2 + 3R1$ next to R_2, then this means I will add 3 times row 1 to row 2. If I write $2R_2 - R1$ next to R_2, then I have done two row operations, namely I multiplied R_2 by 2, and then added (-1) times $R1$ to the result (replacing $R2$ with the sum). The steps below read left to right, top to bottom. In order to avoid fractions, I wait to divide until the last step, only putting a 1 in each pivot at the very end.

$$\Rightarrow^{(1)} \left[\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 1 & -2 & 0 & 3 \\ 0 & 4 & 2 & 1 \end{array}\right] \begin{array}{c} \\ 2R_2 - R_1 \\ \\ \end{array} \quad \Rightarrow^{(2)} \left[\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & -5 & 1 & 4 \\ 0 & 4 & 2 & 1 \end{array}\right] \begin{array}{c} \\ \\ 5R_3 + 4R_2 \end{array}$$

$$\Rightarrow^{(3)} \left[\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & -5 & 1 & 4 \\ 0 & 0 & 14 & 21 \end{array}\right] \begin{array}{c} \\ \\ R_3/7 \end{array} \quad \Rightarrow^{(4)} \left[\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & -10 & 2 & 8 \\ 0 & 0 & 2 & 3 \end{array}\right] \begin{array}{c} 2R_1 + R_3 \\ R_2 - R_3 \\ \\ \end{array}$$

$$\Rightarrow^{(5)} \left[\begin{array}{ccc|c} 4 & 2 & 0 & 7 \\ 0 & -10 & 0 & 5 \\ 0 & 0 & 2 & 3 \end{array}\right] \begin{array}{c} \\ R_2/5 \\ \\ \end{array} \quad \Rightarrow^{(6)} \left[\begin{array}{ccc|c} 4 & 2 & 0 & 7 \\ 0 & -2 & 0 & 1 \\ 0 & 0 & 2 & 3 \end{array}\right] \begin{array}{c} R_1 + R_2 \\ \\ \\ \end{array}$$

$$\Rightarrow^{(7)} \left[\begin{array}{ccc|c} 4 & 0 & 0 & 8 \\ 0 & -2 & 0 & 1 \\ 0 & 0 & 2 & 3 \end{array}\right] \begin{array}{c} R_1/4 \\ R_2/-2 \\ R_3/2 \end{array} \quad \Rightarrow^{(8)} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & 1 & 3/2 \end{array}\right]$$

Writing the final matrix in terms of a system, we have the solution $x_1 = 2, x_2 = -1/2, x_3 = 3/2$. Remember that this tells us (1) where three planes intersect, (2) how to write the 4th column $\vec{b}$ in our original augmented matrix as a linear combination of the columns of the coefficient matrix A, and (3) how to solve the matrix equation $A\vec{x} = \vec{b}$ for $\vec{x}$.

The following steps describe the Gaussian elimination algorithm that we used above. Please take a moment to compare what is written below with the example above. Most of the problems in this unit can be solved using Gaussian elimination, so we will practice it as we learn a few new ideas.

1. Forward Phase (row echelon form) - The following 4 steps should be repeated until you have mentally erased all the rows or all the columns. In step 1 or 4 you will erase a column and/or row from the matrix.

   (a) Consider the first column of your matrix. Start by interchanging rows (if needed) to place a nonzero entry in the first row. If all the elements in the first column are zero, then ignore that column in future computations (mentally erase the column) and begin again with the smaller matrix which is missing this column. If you erase the last column, then stop.

   Computer algorithms place the largest (in absolute value) nonzero entry in the first row. This reduces potential errors due to rounding that can occur in later steps.

   (b) Divide the first row (of your possibly smaller matrix) row by its leading entry so that you have a leading 1. This entry is a pivot, and the column is a pivot column. [When doing this by hand, it is often convenient to skip this step and do it at the very end so that you avoid fractional arithmetic. If you can find a common multiple of all the terms in this row, then divide by it to reduce the size of your computations.]

   (c) Use the pivot to eliminate each nonzero entry below the pivot, by adding a multiple of the top row (of your smaller matrix) to the nonzero lower row.

   (d) Ignore the row and column containing your new pivot and return to the first step (mentally cover up or erase the row and column containing your pivot). If you erase the last row, then stop.

   Ignoring rows and columns is equivalent to incrementing row and column counters in a computer program.

2. Backward Phase (reduced row echelon form - often called Gauss-Jordan elimination) - At this point each row should have a leading 1, and you should have all zeros to the left and below each leading 1. If you skipped step 2 above, then at the end of this phase you should divide each row by its leading coefficient to make each row have a leading 1.

   (a) Starting with the last pivot column. Use the pivot in that column to eliminate all the nonzero entries above it, by adding multiples of the row containing the pivot to the nonzero rows above.

   (b) Work from right to left, using each pivot to eliminate the nonzero entries above it. Nothing to the left of the current pivot column changes. By working right to left, you greatly reduce the number of computations needed to fully reduce the matrix.

Example 2.12. As a final example, let's reduce $\left[\begin{array}{cccc|c} 0 & 1 & 1 & -2 & 7 \\ 1 & 3 & 5 & 1 & 6 \\ 2 & 0 & 4 & 3 & -8 \\ -2 & 1 & -3 & 0 & 5 \end{array}\right]$ to reduced row echelon form (rref). The first step involves swapping 2 rows. We swap row 1 and row 2 because this places a 1 as the leading entry in row 1.

(1) Get a nonzero entry in upper left

$$\Rightarrow \left[\begin{array}{cccc|c} 0 & 1 & 1 & -2 & 7 \\ 1 & 3 & 5 & 1 & 6 \\ 2 & 0 & 4 & 3 & -8 \\ -2 & 1 & -3 & 0 & 5 \end{array}\right] \begin{array}{l} R_1 \leftrightarrow R_2 \\ \\ \\ \\ \end{array}$$

(2) Eliminate entries in 1st column

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 3 & 5 & 1 & 6 \\ 0 & 1 & 1 & -2 & 7 \\ 2 & 0 & 4 & 3 & -8 \\ -2 & 1 & -3 & 0 & 5 \end{array}\right] \begin{array}{l} \\ \\ R_3 - 2R_1 \\ R_4 + 2R_1 \end{array}$$

(3) Eliminate entries in 2nd column

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 3 & 5 & 1 & 6 \\ 0 & 1 & 1 & -2 & 7 \\ 0 & -6 & -6 & 1 & -20 \\ 0 & 7 & 7 & 2 & 17 \end{array}\right] \begin{array}{l} \\ \\ R_3 + 6R_2 \\ R_4 - 7R_2 \end{array}$$

(4) Make a leading 1 in 4th column

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 3 & 5 & 1 & 6 \\ 0 & 1 & 1 & -2 & 7 \\ 0 & 0 & 0 & -11 & 22 \\ 0 & 0 & 0 & 16 & -32 \end{array}\right] \begin{array}{l} \\ \\ R_3/(-11) \\ R_4/16 \end{array}$$

(5) Eliminate entries in 4th column

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 3 & 5 & 1 & 6 \\ 0 & 1 & 1 & -2 & 7 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 & -2 \end{array}\right] \begin{array}{l} \\ \\ \\ R_4 - R_3 \end{array}$$

(6) Row Echelon Form

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 3 & 5 & 1 & 6 \\ 0 & 1 & 1 & -2 & 7 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

At this stage we have found a row echelon form of the matrix. Notice that we eliminated nonzero terms in the lower left of the matrix by starting with the first column and working our way over column by column. Columns 1, 2, and 4 are the pivot columns of this matrix. We now use the pivots to eliminate the other nonzero entries in each pivot column (working right to left).

Recall that a matrix is in reduced row echelon (rref) if:

1. Nonzero rows begin with a leading 1.
2. Leadings 1's on subsequent rows appear further right than previous rows.
3. Rows of zeros are at the bottom.
4. Zeros are above and below each pivot.

(7) Eliminate entries in 4th column

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 3 & 5 & 1 & 6 \\ 0 & 1 & 1 & -2 & 7 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} R_1 - R3 \\ R_2 + 2R_3 \\ \\ \\ \end{array}$$

(8) Eliminate entries in 2nd column

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 3 & 5 & 0 & 8 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} R_1 - 3R_2 \\ \\ \\ \\ \end{array}$$

(9) Reduced Row Echelon Form

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

(10) Switch to system form

$$\Rightarrow \begin{aligned} x_1 + 2x_3 &= -1 \\ x_2 + x_3 &= 3 \\ x_4 &= -2 \\ 0 &= 0 \end{aligned}$$

We have obtained the reduced row echelon form. When we write this matrix in the corresponding system form, notice that there is not a unique solution to the system. Because the third column did not contain a pivot column, we can write every variable in terms of x_3 (the redundant equation $x_3 = x_3$ allows us to write x_3 in terms of x_3). We are free to pick any value we want for x_3 and still obtain a solution. For this reason, we call x_3 a free variable, and write our infinitely many solutions in terms of x_3 as

Free variables correspond to non pivot columns. Solutions can be written in terms of free variables.

$$\begin{aligned} x_1 &= -1 - 2x_3 \\ x_2 &= 3 - x_3 \\ x_3 &= x_3 \\ x_4 &= -2 \end{aligned} \quad \text{or by letting } x_3 = t \quad \begin{aligned} x_1 &= -1 - 2t \\ x_2 &= 3 - t \\ x_3 &= t \\ x_4 &= -2 \end{aligned} .$$

By choosing a value (such as t) for x_3, we can write our solution in so called parametric form. We have now given a parametrization of the solution set, where t is an arbitrary real number.

parametric form

Problem 2.13 Each of the following augmented matrices requires one row operation to be in reduced row echelon form. Perform the required row operation, and then write the solution to the corresponding system of equations in terms of the free variables.

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

2. $\left[\begin{array}{ccc|c} 1 & 2 & 0 & -4 \\ 0 & 0 & 1 & 3 \\ -3 & -6 & 0 & 12 \end{array}\right]$

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

4. $\left[\begin{array}{ccccc|c} 0 & 1 & 0 & 7 & 0 & 3 \\ 0 & 0 & 1 & 5 & -3 & -10 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

Problem 2.14 Use Gaussian elimination to solve

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

by row reducing the matrix to reduced row echelon form. [Hint: Start by interchanging row 1 and row 2.]

Problem 2.15 Use Gaussian elimination to solve

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 4 \\ -x_1 + 2x_2 + 3x_3 &= 8 \\ 2x_1 - 4x_2 + x_3 &= 5 \end{aligned}$$

by row reducing the matrix to reduced row echelon form. [Hint: You should end up with infinitely many solutions. State your solution by writing each variable in terms of the free variable(s).]

Problem 2.16 Use Gaussian elimination to solve

$$\begin{aligned} x_1 + 2x_3 + 3x_4 &= -7 \\ 2x_1 + x_2 + 4x_4 &= -7 \\ -x_1 + 2x_2 + 3x_3 &= 0 \\ x_2 - 2x_3 - x_4 &= 4 \end{aligned}$$

by row reducing the matrix to reduced row echelon form.

2.3 Rank, Linear Independence, Inverses, and Determinants

Definition 2.13.
- The rank of a matrix is the number of pivot columns of the matrix. To find the rank of a matrix, you reduce the matrix using Gaussian elimination until you discover the pivot columns.
- The span of a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is all possible linear combinations of the vectors. In terms of matrices, the span of a set of vectors is all possible vectors $\vec{b}$ such that $A\vec{x} = \vec{b}$ for some vector $\vec{x}$, where the vectors $\vec{v}_i$ are placed in the columns of A.
- We say that a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is linearly independent if the only solution to the homogeneous system $c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_n\vec{v}_n = \vec{0}$ is the trivial solution $c_1 = c_2 = \cdots = c_n = 0$. Otherwise we say the vectors are linearly dependent, and it is possible to write one of the vectors as a linear combination of the others. We say the vectors are dependent because one of them depends on (can be obtained as a linear combination of) the others.
- In terms of spans, we say vectors are linearly dependent when one of them is in the span of the other vectors.

As we complete each of the following problems in class, we'll talk about the span of the vectors, and the rank of the corresponding matrix. The key thing we need to focus on is learning to use the words "linearly independent" and "linearly dependent."

Problem 2.17 Are the vectors $\vec{v}_1 = (1, 3, 5)$, $\vec{v}_2 = (-1, 0, 1)$, and $\vec{v}_3 = (0, 3, 1)$ linearly independent? Solve the system $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$ to answer this question. If they are dependent, then write one of the vectors as a linear combination of the others.

Problem 2.18 Are the vectors $\vec{v}_1 = (1, 2, 0)$, $\vec{v}_2 = (2, 0, 3)$, and $\vec{v}_3 = (3, -2, 6)$ linearly independent? Solve the system $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$ to answer this question. If they are dependent, then write one of the vectors as a linear combination of the others.

Problem 2.19 Answer each of the following:

1. Suppose you have row reduced a 3 by 3 matrix, and discovered that the rank of the matrix is 2. Are the columns of the matrix independent or dependent? What if the rank was 3?
2. Now suppose you have row reduced a 7 by 7 matrix. If the columns are independent, what possible options do you have for the rank.
3. Now suppose you have row reduced a 7 by 5 matrix. If the columns are independent, what must the rank be.
4. Now suppose you have row reduced a 5 by 7 matrix. Explain why the columns cannot be independent.
5. If you have n vectors placed in the columns of a matrix, what must the rank of the matrix be in order to guarantee that the vectors are independent?

Problem 2.20 Is the vector $[2,0,1,-5]$ in the span of

$$\{[1,0,-1,-2],[1,2,3,0],[0,1,-1,2]\}?$$

If it is, then write it as a linear combination of the others. If it is not, then explain why it is not.

Problem 2.21 Find the reduced row echelon form of the matrix

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

Use your result to answer the following questions.

1. Write both $(1,0)$ and $(0,1)$ as linear combinations of $(2,1)$ and $(-1,1)$.
2. Write $\begin{pmatrix}2\\0\end{pmatrix}$ as a linear combination of $\begin{pmatrix}2\\1\end{pmatrix}$ and $\begin{pmatrix}-1\\1\end{pmatrix}$. Then write $\begin{pmatrix}8\\0\end{pmatrix}$ as a linear combination of $\begin{pmatrix}2\\1\end{pmatrix}$ and $\begin{pmatrix}-1\\1\end{pmatrix}$.
3. Let $A=\begin{bmatrix}2 & -1\\ 1 & 1\end{bmatrix}$. Find vectors $\vec{x}$ and $\vec{y}$ so that $A\vec{x}=\begin{pmatrix}1\\0\end{pmatrix}$ and $A\vec{y}=\begin{pmatrix}0\\1\end{pmatrix}$.
4. Find a matrix B so that $AB=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$.

Problem: 21, revised Answer each of the following questions.

1. Find the reduced row echelon form of the matrix
$$B=\begin{bmatrix}2 & -1 & 1 & 0 & 2 & 0 & 2\\ 1 & 1 & 0 & 1 & 0 & 3 & 3\end{bmatrix}.$$
2. Write $(1,0)$ as a linear combination of $(2,1)$ and $(-1,1)$. Remember, that when writing $c_1(2,1)+c_2(-1,1)=(1,0)$, you must solve for the unknown constants. Feel free to row reduce the augmented matrix $\begin{bmatrix}2 & -1 & 1\\ 1 & 1 & 0\end{bmatrix}$.
3. Write $(0,1)$ as a linear combination of $(2,1)$ and $(-1,1)$. Remember, that when writing $c_1(2,1)+c_2(-1,1)=(0,1)$, you must solve for the unknown constants. Feel free to row reduce the augmented matrix $\begin{bmatrix}2 & -1 & 0\\ 1 & 1 & 1\end{bmatrix}$.
4. Continue to write each of $\begin{pmatrix}2\\0\end{pmatrix}$, $\begin{pmatrix}0\\3\end{pmatrix}$, and $\begin{pmatrix}2\\3\end{pmatrix}$ as a linear combination of $\begin{pmatrix}2\\1\end{pmatrix}$ and $\begin{pmatrix}-1\\1\end{pmatrix}$. [Hint:At some point, rather than row reducing $\begin{bmatrix}2 & -1 & \vec{v}\\ 1 & 1 & \end{bmatrix}$, ask yourself how you could use part 1 to answer this.]

5. The following matrix row reduces to give

$$\begin{bmatrix}1 & 0 & 2 & 4 & 5 & 8\\ 0 & 2 & 5 & 2 & -1 & 3\\ 0 & -2 & -1 & 0 & 2 & 1\end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix}1 & 0 & 0 & 3 & \frac{9}{2} & 6\\ 0 & 1 & 0 & -\frac{1}{4} & -\frac{9}{8} & -1\\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{4} & 1\end{bmatrix}.$$

   Use this to write both $(4,2,0)$ and $(5,-1,2)$ as a linear combination of the first three columns.

Definition 2.14. The identity matrix I is a square matrix so that if A is a square matrix, then $IA = AI = A$. The identity matrix acts like the number 1 when performing matrix multiplication.

If A is a square matrix, then the inverse of A is a matrix A^{-1} where we have $AA^{-1} = A^{-1}A = I$, provided such a matrix exists.

Problem Let $A = \begin{bmatrix}1 & 3\\ 3 & 4\end{bmatrix}$. We now develop an algorithm for computing the inverse A^{-1}. If an inverse matrix exists, then we know it's the same size as A, so we could let $A^{-1} = \begin{bmatrix}\vec{v}_1 & \vec{v}_2\end{bmatrix}$ be the inverse matrix, where $\vec{v}_1$ and $\vec{v}_2$ are the columns of A^{-1}.

1. We know that $AA^{-1} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$. Explain why $A\vec{v}_1 = \begin{pmatrix}1\\0\end{pmatrix}$ and $A\vec{v}_2 = \begin{pmatrix}0\\1\end{pmatrix}$.
2. Solve the matrix equations $A\vec{v}_1 = \begin{pmatrix}1\\0\end{pmatrix}$ and $A\vec{v}_2 = \begin{pmatrix}0\\1\end{pmatrix}$. (This involves row reducing $\begin{bmatrix}1 & 3 & 1\\ 3 & 4 & 0\end{bmatrix}$ and $\begin{bmatrix}1 & 3 & 0\\ 3 & 4 & 1\end{bmatrix}$).
3. What is the reduced row echelon form of $\begin{bmatrix}1 & 3 & 1 & 0\\ 3 & 4 & 0 & 1\end{bmatrix}$. How is this related to your previous work.
4. State the inverse of A.

The previous problem showed you how to obtain a matrix B so that $AB = I$. You just had to row reduce that matrix $\begin{bmatrix}A & I\end{bmatrix}$ to the matrix $\begin{bmatrix}I & A^{-1}\end{bmatrix}$. The inverse shows up instantly after row reduction.

Problem 2.22 Use the algorithm describe immediately before this problem to compute the inverse of

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

Then use your work to write each of the standard basis vectors $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ as a linear combination of the columns of A.

Problem 2.23 Let $A = \begin{bmatrix}a & b\\ c & d\end{bmatrix}$. Use Gaussian elimination to show that the inverse of A is

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

In computing the inverse of a 2 by 2 matrix, the number $ad - bc$ appears in the denominator. We call this number the determinant. If I asked you to compute the inverse of a 3 by 3 matrix, you would again see a number appear in the denominator. We call that number the determinant. This holds true in all dimensions.

Problem: Optional Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. Use Gaussian elimination to find the inverse of A, and show that the common denominator is $a(ei - hf) - b(di - gf) + c(dh - ge)$.

Definition 2.15: Determinants of 2 by 2 and 3 by 3 matrices. The determinant of a 2×2 and 3×3 matrix are the numbers

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

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= a \det \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \det \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \det \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ &= a(ei - hf) - b(di - gf) + c(dh - ge) \end{aligned}$$

We use vertical bars next to a matrix to state we want the determinant. Notice the negative sign on the middle term of the 3×3 determinant. Also, notice that we had to compute three determinants of 2 by 2 matrices in order to find the determinant of a 3 by 3.

In the examples above, we obtained the determinant of a 3 by 3 matrix by computing the determinant of several 2 by 2 matrices. We obtained each 2 by 2 matrix by removing a row and column from the original 3 by 3 matrix. We now add some language to extend the definition above to all dimensions.

Definition 2.16: Minors, Cofactors, and General determinants. Let A be an n by n matrix.

- The minor M_{ij} of a matrix A is the determinant of the the matrix formed by removing row i and column j from A.
- The cofactor C_{ij} is the product of the minor M_{ij} and $(-1)^{i+j}$, so we have $C_{ij} = (-1)^{i+j} M_{ij}$. So it's either the minor, or the opposite of the minor.
- To compute the determinant, first pick a row or column. We define the determinant to be $\sum_{k=1}^{n} a_{ik} C_{ik}$ (if we chose row i) or alternatively $\sum_{k=1}^{n} a_{kj} C_{kj}$ (if we chose column j).
- You can pick ANY row or ANY column you want, and then compute the determinant by multiplying each entry of that row or column by its cofactor, and then summing the results. (The fact that this works would require proof. That proof will be left to a course in linear algebra.)
- A sign matrix keeps track of the $(-1)^{j+k}$ term in the cofactor. All you have to do is determine if the first entry of your expansion has a plus or minus, and then alternate the sign as you expand.

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

sign matrix

Problem 2.24 Compute the determinant of the matrix $\begin{bmatrix}2 & 3 & -1\\ 1 & 0 & 0\\ 4 & 2 & 5\end{bmatrix}$ in 3 different ways. First, use a cofactor expansion using the first row (Definition 2.15). Then use a cofactor expansion using the 2nd row. Then finally use a cofactor expansion using column 3. Which of the was the quickest, and why?

Problem 2.25 Compute the determinants of the matrices

$$A = \begin{bmatrix}2 & 1 & -6 & 8\\ 0 & 3 & 5 & 4\\ 0 & 0 & 1 & 5\\ 0 & 0 & 0 & -4\end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix}3 & 2 & 5 & -1\\ 0 & 8 & 4 & 2\\ 0 & -1 & 0 & 0\\ 0 & -5 & 3 & -1\end{bmatrix}.$$

You can make these problems really fast if you use a cofactor expansion along a row or column that contains a lot of zeros.

Problem 2.26 Compute the determinant of

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

Then find the inverse of A (or explain why it does not exist). Are the columns of A linearly independent or linearly dependent?

Problem 2.27 Compute the determinant of $A = \begin{bmatrix}2 & 1 & 0\\ 0 & 2 & 1\\ 1 & 0 & 2\end{bmatrix}$. Does A have an inverse? Are the columns of A linearly independent or linearly dependent? Answer both of the previous questions without doing any row reduction. Then row reduce $\begin{bmatrix}A & I\end{bmatrix} = \begin{bmatrix}2 & 1 & 0 & 1 & 0 & 0\\ 0 & 2 & 1 & 0 & 1 & 0\\ 1 & 0 & 2 & 0 & 0 & 1\end{bmatrix}$ to confirm your answer.

After completing the previous two problems, you should see that there is a connection between the determinant, inverse, and linear independence. Make a conjecture about what this connection is. We'll learn a little more about determinants and inverses, and then you'll have a chance to state your conjecture, as well as prove it.

Problem 2.28 Start by writing the system of equations

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

as a matrix product $A\vec{x} = \vec{b}$. (What are A, $\vec{x}$ and $\vec{b}$?) Then find the inverse of A, and use this inverse to find $\vec{x}$. [Hint: If we just have numbesr, then to solve $ax = b$, we multiply both sides by $\frac{1}{a}$ to obtain $\frac{1}{a}ax = \frac{1}{a}b$ or just $x = \frac{1}{a}b$.]

In the next problem, you'll prove that the determinant of a 2 by 2 matrix gives the area of a parallelogram whose edges are the columns of the matrix.

Problem 2.29 To find the area of the parallelogram with vertexes $O = (0, 0)$, $U = (a, c)$, $V = (b, d)$, and $P = (a + b, c + d)$, we would find the length of OU (the base b), and multiply it by the distance from V to OU. Complete the following:

1. Find the projection of $\vec{OV}$ onto $\vec{OU}$. (You may have to look up a formula from math 215.)
2. The vector $\vec{OV} - \text{proj}_{\vec{OV}}\vec{OU}$ is called the component of $\vec{OV}$ that is orthogonal to $\vec{OU}$. The length of this vector is precisely the distance from V to OU, which we'll call h. Find the length of this vector.
3. We now have the base $b = |OU|$ and height h of a parallelogram. Compute the product, and prove it equals $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = |ad - bc|$.

The result above extends to 3 dimensions. The determinant of a 3 by 3 matrix gives the volume of a parallelepiped whose edges are the columns of the matrix. We then use determinants to define nth dimensional volume.

Problem 2.30 Answer each of the following:

1. Let $\vec{u} = (2, 3)$. If you pick a vector $\vec{v}$ that is a linear combination of $\vec{u}$, what will the determinant of $\begin{bmatrix} \vec{u} & \vec{v} \end{bmatrix}$ equal? First explain how you know the answer (before you have even chosen a vector $\vec{v}$). Then give us an example by picking a vector that is a linear combination of $\vec{v}$.
2. Let $\vec{u} = (1, 0, 2)$ and $\vec{v} = (0, -1, 1)$. If $\vec{w}$ is a linear combination of $\vec{u}$ and $\vec{v}$, what will the determinant equal? Explain. Then show us an example to confirm your conjecture.
3. We already computed the determinant of $A = \begin{bmatrix} 2 & 1 & -6 & 8 \\ 0 & 3 & 5 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & -4 \end{bmatrix}$. Swap two columns of the matrix, and then compute the determinant. How does the determinant of your matrix with swapped columns relate to the determinant of the original matrix. If you swap two columns of a matrix, what happens to the determinant?

Problem 2.31 Construct a 2 by 2 matrix whose columns are linearly independent. What is the reduced row echelon form of your matrix? Compute the rank and the determinant, and finally find the inverse (if possible).

Now construct a 2 by 2 matrix whose columns are linear dependent. What is the reduced row echelon form of your matrix? Compute the rank and the determinant, and finally find the inverse (if possible).

Make a conjecture about the connection between (1) linear dependence, (2) rref, (3) rank, (4) determinant, and (5) inverses. Then use a computer to give two 3 by 3 examples similar to the examples above. You'll be asked to show us the computations on the computer in class.

Problem 2.32 Consider the matrix $A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 0 \\ 0 & 4 & 3 \end{bmatrix}$. Compute the determinant of A. Then create a matrix B so that the ijth entry of B is the cofactor C_{ij} (remove row i and column j, compute the determinant, and then times by an appropriate sign). This will require that you compute nine 2 by 2 determinants. Finally, compute the inverse of A (feel free to use a computer on this part). Make a conjecture about the connection between the determinant of A, this matrix B, and the inverse of A. We'll verify your conjecture is true on a 4 by 4 matrix in class.

2.4 Eigenvalues and Eigenvectors

The final computational skill we need to tackle is to compute eigenvalues and eigenvectors. Let's start by looking at an example to motivate the language we are about to introduce.

Example 2.17. Consider the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. When we multiply this matrix by the vector $\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we obtain $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\vec{x}$. Multiplication by the matrix A was miraculously the same as multiplying by the number 3. Symbolically we have $A\vec{x} = 3\vec{x}$. Not every vector $\vec{x}$ satisfies this property, as letting $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ gives the product $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, which is not a multiple of $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Our main goal in this section is to answer the following two questions:

1. For which nonzero vectors $\vec{x}$ (eigenvectors) is it possible to write $A\vec{x} = \lambda\vec{x}$?
2. Which scalars λ (eigenvalues) satisfy $A\vec{x} = \lambda\vec{x}$?

Now for some definitions.

Definition 2.18: Eigenvector and Eigenvalue. Let A be a square $n \times n$ matrix.

- An eigenvector of A is a nonzero vector $\vec{x}$ such that $A\vec{x} = \lambda\vec{x}$ for some scalar λ. (Matrix multiplication reduces to scalar multiplication.) We avoid letting $\vec{x}$ be the zero vector because $A\vec{0} = \lambda\vec{0}$ no matter what λ is.
- If $\vec{x}$ is an eigenvector satisfying $A\vec{x} = \lambda\vec{x}$, then we call λ and eigenvalue of A.

Problem 2.33 Use the definition above to determine with of the following are eigenvectors of $\begin{bmatrix} 3 & 1 \\ 4 & 6 \end{bmatrix}$:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, (1, 4), \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1, \\ -1 \end{pmatrix}, \begin{pmatrix} -2, \\ 2 \end{pmatrix}.$$

If the vector is an eigenvector, state the corresponding eigenvalue.

The next problem gives us an algorithm for computing eigenvalues and eigenvectors.

Problem 2.34: How to compute eigenvalues and eigenvectors Let A be a square matrix.

1. If λ is an eigenvalue, explain why we can find the eigenvectors by solving the equation $(A-\lambda I)\vec{x}=\vec{0}$. This means we can subtract λ from the diagonal entries of A, and then row reduce $\begin{bmatrix} A-\lambda I & \vec{0} \end{bmatrix}$ to obtain the eigenvectors. Note that you should always obtain infinitely many solutions.
2. Explain why we can obtain the eigenvalues of A by solving for when the determinant of $(A-\lambda I)$ is zero, i.e. solving the equation

$$\det(A-\lambda I)=0.$$

The algorithm above suggests the following definition.

Definition 2.19. If A is a square n by n matrix, then we call $\det(A-\lambda I)$ the characteristic polynomial of A. It is a polynomial in λ of degree n, and hence has n roots (counting multiplicity). These roots are the eigenvalues of A.

We now have an algorithm for finding the eigenvalues and eigenvectors of a matrix. We start by finding the characteristic polynomial of A. The zeros of this polynomial are the eigenvalues. To get the eigenvectors, we just have to row reduce the augmented matrix $\begin{bmatrix} A-\lambda I & \vec{0} \end{bmatrix}$. Finding eigenvalues and eigenvectors requires that we compute determinants, find zeros of polynomials, and then solve homogeneous systems of equations. You know you are doing the problem correctly if you get infinitely many solutions to the system $(A-\lambda I)\vec{x}=0$ for each lambda (i.e. there is at least one row of zeros along the bottom after row reduction). As another way to check your work, the following two facts can help.

- The sum of the eigenvalues equals the trace of the matrix (the sum of the diagonal elements).
- The product of the eigenvalues equals the determinant.

The trace and determinant are equal to the sum and product of the eigenvalues.

Problem 2.35 Consider the matrix $A=\begin{bmatrix} 3 & 1 \\ 4 & 6 \end{bmatrix}$ from problem 2.33.

1. Find the characteristic polynomial of A, and then find the zeros to determine the eigenvalues.
2. For each eigenvalue, find all corresponding eigenvectors.
3. Compute the trace and determinant of A.

Problem 2.36 Consider the matrix $A=\begin{bmatrix} 6 & 4 \\ 3 & 2 \end{bmatrix}$. Find the characteristic polynomial and eigenvalues of A. Then for each eigenvalue, find all corresponding eigenvectors. (Check your work by computing the trace and determinant of A.)

Problem 2.37 Consider the matrix $A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$. Find the characteristic polynomial and eigenvalues of A. Then for each eigenvalue, find all corresponding eigenvectors. (Check your work by computing the trace and determinant of A.)

Problem 2.38 Consider the matrix $A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}$. Find the characteristic polynomial and eigenvalues of A. Then for each eigenvalue, find all corresponding eigenvectors. (Check your work by computing the trace and determinant of A.)

2.5 Wrap Up

Once you have finished the problems in the section and feel comfortable with the ideas, create a short one page lesson plan that contains examples of the key ideas. You will get a chance to teach from this lesson plan prior to taking the exam. Then log on to I-Learn and download the quiz. Once you have taken the quiz, upload your work to I-Learn and then download the key to see how you did. If you still need to work on mastering some of the ideas, please do so and then demonstrate your mastery though the quiz corrections.

Chapter 3

Linear Algebra Applications

This chapter covers the following ideas. (These are subject to change, as I write the notes.)

1. Find the currents in electrical systems involving batteries and resistors, using both Gaussian elimination and Cramer's rule.
2. Find interpolating polynomials. Use the transpose and inverse of a matrix to solve the least squares regression problem of fitting a line to a set of data.
3. Find the partial fraction decomposition of a rational function. Utilize this decomposition to integrate rational functions.
4. Describe a Markov process. Explain how an eigenvector of the eigenvalue $\lambda = 1$ is related to the limit of powers of the transition matrix.
5. Explain how to generalize the derivative to a matrix. Use this generalization to locate optimal values of the function using the second derivative test. Explain the role of eigenvalues and eigenvectors in the second derivative test.