Mathematics for Social Scientists

Answers to the end-of-chapter exercises

Jonathan Kropko

February 17, 2016

Contents

A Note

To Students: you will find here the answers to every exercise that appears at the end of one of the chapters of *Mathematics for Social Scientists*. Remember, the social sciences have become very quantitative. If you are able to perform the underlying math, you will feel confident using quantitative methods later. And that feeling will put you far ahead of many of your peers in the social sciences. Don't cheat yourself out of that advantage by using this guide to cut corners.

If you have worked on a exercise, thought about the problems it presents, and given it real effort, then this guide can help you to evaluate your work. The answers here do not simply list the answer. They work through the entire problem from beginning to end. It is my hope that this guide can be just as effective at teaching as the main text. Be careful reading through problems in the later chapters, however. For the sake of space, I sometimes skip describing the steps for techniques that are covered in an earlier chapter. I might not take the time to explain a square root in an exercise from the chapter on integrals. If you find that I've skipped over describing an important step, check to see whether it's covered in an earlier chapter.

There are often many different ways to solve the same problem, so copying the particular steps outlined here again and again will be transparent to your instructor, who also has a copy of this answer guide. If you try the problem and find yourself stuck, then give yourself some time to think about a solution for yourself before referring to this guide. You will find the answer more satisfying if you do.

To Instructors: please be aware that your students may be able to acquire a copy of this guide and take steps to make it less easy for a student to copy the steps exactly from this guide. You might perhaps slightly alter the problems. You might want to discuss this guide with the students at the beginning of the course to let them know you have this guide and that – since there are often many ways to get to the same answer – you will be able to identify work that copies from this guide.

It is important that, before assigning exercises from this book, you read through the answer guide to get a sense of how much work and what level of work each exercise entails. Some of the exercises are simply stated but deceptively complex.

To Everyone: it is very likely that I've made errors in this guide that I've been unable to catch before it became available to you. If you identify an error, I would be grateful if you contacted the publisher, or contacted me through the information provided by the publisher online, to let me know.

I hope that, no matter what your role in a math class, you find this guide instructive and useful, and that it helps you to become a more confident user of applied mathematics in the social sciences. Thank you for reading *Mathematics for Social Scientists* and this guide.

–JMK

1 Algebra Review

1. First I consider the parentheses, calculating division first,

$$5 \times (3 - 2) + 6 \div 2 - 7 \times 0,$$

then subtraction,

$$5 \times 1 + 6 \div 2 - 7 \times 0.$$

Now that the parentheses are solved, I perform multiplication and division next,

$$5 + 6 \div 2 - 7 \times 0,$$

$$5 + 3 - 7 \times 0,$$

$$5 + 3 - 0,$$

which evaluates to 8.

2. (a) Here is one way to proceed:

$$\begin{aligned} &19800 \\ &= 100 \times 198 \\ &= (10 \times 10) \times 198 \\ &= (2 \times 5) \times (2 \times 5) \times 198 \\ &= 2 \times 2 \times 5 \times 5 \times (2 \times 99) \\ &= 2 \times 2 \times 2 \times 5 \times 5 \times (9 \times 11) \\ &= 2 \times 2 \times 2 \times 5 \times 5 \times (3 \times 3) \times 11 \\ &= 2 \times 2 \times 2 \times 3 \times 3 \times 5 \times 5 \times 11. \end{aligned}$$

(b) From the previous problem, we know that

$$\sqrt{19800} = \sqrt{2 \times 2 \times 2 \times 3 \times 3 \times 5 \times 5 \times 11}.$$

We use the "jailbreak method" to remove pairs of prime factors from the square root. There is a pair of 2s, of 3s, and of 5s to remove, leaving a 2 and the 11 inside:

$$= (2 \times 3 \times 5)\sqrt{2 \times 11}.$$

$$= 30\sqrt{22}.$$

3. (a) The lowest common denominator for 8 and 12 is 24. So we rewrite both fractions to have a denominator of 24. For the first fraction, we multiply the top and bottom by 3, and for the second fraction we multiply the top and bottom by 2:

$$\frac{5}{8} + \frac{1}{12} = \frac{15}{24} + \frac{2}{24} = \frac{17}{24}.$$

(b) We can cross-cancel a factor of 3 from both 3 and 21, and we can cross-cancel a factor of 5 from both 5 and 10. These cancelations leave us with

$$\frac{1}{1} \times \frac{2}{7} = \frac{2}{7}.$$

(c) We take the reciprocal of the denominator and write the problem as multiplication:

$$\frac{2}{7} \times \frac{8}{3}.$$

Neither 2 and 3 nor 7 and 8 share common factors. So we just multiply the numerators and denominators together:

$$\frac{2 \times 8}{7 \times 3} = \frac{16}{21}.$$

4. (a) First we can distribute the exponent 3 to each variable inside the parentheses,

$$x^3y^3x^2,$$

and we can rewrite the expression as

$$x^3x^2y^3.$$

Finally, we add the exponents for the factors that have a base of x:

$$x^5y^3.$$

(b) First we take the reciprocal of the bottom fraction and write the expression as a multiplication problem:

$$\frac{w^3z^4}{(w+1)(z-3)} \times \frac{(w-2)(z-3)}{(wz)^3} = \frac{w^3z^4(w-2)(z-3)}{(w+1)(z-3)(wz)^3}.$$

The factor $(z-3)$ is on the top and bottom and cancels out:

$$\frac{w^3z^4(w-2)}{(w+1)(wz)^3}.$$

We distribute the exponent of 3 to each variable inside (wz):

$$\frac{w^3z^4(w-2)}{(w+1)w^3z^3}.$$

The factor w^3 is on the top and bottom and cancels out:

$$\frac{z^4(w-2)}{(w+1)z^3},$$

and we subtract the exponents on z, leaving

$$\frac{z(w-2)}{(w+1)}.$$

(c) We can take a number of steps in any order. Here's one solution. First, distribute the exponent 3 to the x^2 and to y:

$$\frac{\frac{(x^2)^3}{y^3}x^{-2}}{xy^2}$$

For the $(x^2)^3$ term, we multiply the exponents together to get x^6:

$$\frac{\frac{x^6}{y^3}x^{-2}}{xy^2}.$$

Next we can remove all the fractions by multiplying the exponents of terms in denominators by -1. So the y^3 becomes y^{-3}, x becomes x^{-1}, and y^2 becomes y^{-2}:

$$x^6x^{-2}x^{-1}y^{-3}y^{-2}.$$

We can add the exponents of factors that share the same base, leaving

$$x^3y^{-5}.$$

Finally we can move y^{-5} back to the denominator in order to write it with a positive exponent:

$$\frac{x^3}{y^5}.$$

(d) The "jailbreak" method works for cube-roots too, except three repetitions of the same prime factor are needed in order to bring one factor outside the root. The prime factorization of 189 is

$$3 \times 63$$
$$3 \times (3 \times 21)$$
$$3 \times 3 \times 3 \times 7.$$

So the problem becomes

$$_3\sqrt{3 \times 3 \times 3 \times 7}.$$

The 3 repetitions of the prime factor 3 allow one 3 to leave the root, leaving only 7 inside the root:

$$3_3\sqrt{7}.$$

(e) We can rewrite the problem as

$$x^{\frac{1}{3}}x^{\frac{1}{5}},$$

and then we can add the exponents:

$$x^{\frac{1}{3}+\frac{1}{5}}.$$

In order to add two fractions with different denominators, we find the lowest common denominator for 3 and 5, which is 15. Rewriting each fraction the expression becomes

$$x^{\frac{5}{15}+\frac{3}{15}} = x^{\frac{8}{15}}.$$

We're not quite done yet. Fractional exponents are somewhat messy, so we can rewrite the exponent as

$$(x^8)^{\frac{1}{15}},$$

which means that the expression becomes

$$_{15}\sqrt{x^8}.$$

5. (a)

$$2 \times 2 \times 2 \times \ldots \times 2 \text{ (128 times)}$$
$$= 2^{128}$$

Stata tells me this number is approximately 3.403e+38, which is

$$340,300,000,000,000,000,000,000,000,000,000,000,000$$

(b)

$$\frac{340,300,000,000,000,000,000,000,000,000,000,000,000 \text{ keys}}{1,000,000,000,000,000 \text{ keys per second}}$$
$$= 340,300,000,000,000,000,000,000 \text{ seconds}$$
$$\text{Divide by 60: } = 5,672,000,000,000,000,000,000 \text{ minutes}$$
$$\text{Divide by 60: } = 94,530,000,000,000,000,000 \text{ hours}$$
$$\text{Divide by 24: } = 3,939,000,000,000,000,000 \text{ days}$$
$$\text{Divide by 365: } = 10,790,000,000,000,000 \text{ years,}$$

which is approximately 1000 times longer than the age of the universe.

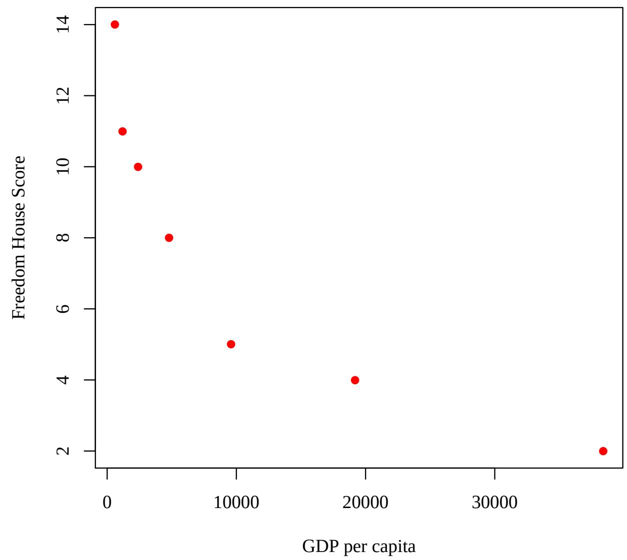
(c)

$$\sqrt{2^{128}} = (2^{128})^{.5} = 2^{128 \times .5} = 2^{64}$$
$$\text{Approximately } 18,450,000,000,000,000,000$$

(d)

$$\frac{2^{64} \text{ keys}}{1,000,000,000,000,000 \text{ keys per second}} \approx 18447 \text{ seconds}$$
$$\text{Divide by 60: } \approx 307 \text{ minutes}$$
$$\text{Divide by 60: } \approx 5.1 \text{ hours}$$

6. (a) The function is an exponential function, and no straight line fits the data very well.



(b) These values can be represented in terms of g as follows:

Guinea-Bissau: $\log_2(1200) = \log_2(600 \times 2) = \log_2(600) + \log_2(2) = g + 1.$

Kyrgyzstan: $\log_2(2400) = \log_2(600 \times 4) = \log_2(600) + \log_2(2^2) = \log_2(600) + 2\log_2(2) = g + 2.$

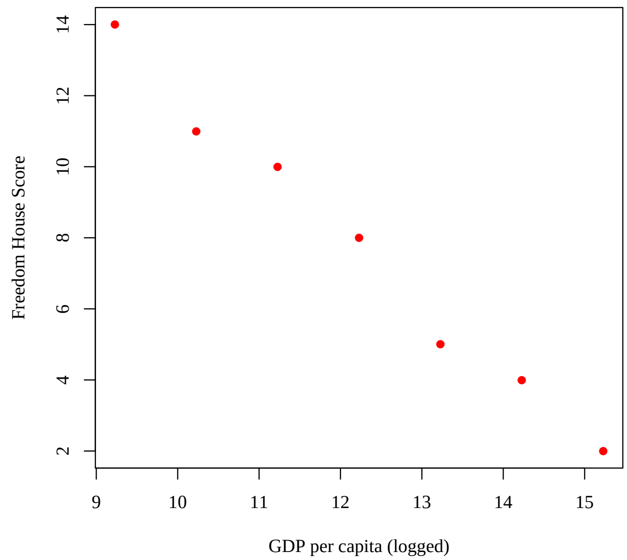
Honduras: $\log_2(4800) = \log_2(600 \times 8) = \log_2(600) + \log_2(2^3) = \log_2(600) + 3\log_2(2) = g + 3.$

Jamaica: $\log_2(9600) = \log_2(600 \times 16) = \log_2(600) + \log_2(2^4) = \log_2(600) + 4\log_2(2) = g + 4.$

Latvia: $\log_2(19200) = \log_2(600 \times 32) = \log_2(600) + \log_2(2^5) = \log_2(600) + 5\log_2(2) = g + 5.$

Denmark: $\log_2(38400) = \log_2(600 \times 64) = \log_2(600) + \log_2(2^6) = \log_2(600) + 6\log_2(2) = g + 6.$

(c) The new graph is



(d) The data are now much more linear. In general, you may want to take the logarithm of a variable when it's distribution is exponential – that is, when values tend to double or triple instead of increasing by additive amounts. Microeconomists, for example, almost always take the logarithm of income when they consider the income of randomly drawn people from the US population since the top income earners have many, many times the more income than people closer to the median.

7. (a)

$$
\begin{aligned}
&= \ln(x^2 z) - \ln\left(\sqrt{e^y}\right) \\
&= \ln(x^2) + \ln(z) - \ln\left((e^y)^{1/2}\right) \\
&= 2\ln(x) + \ln(z) - \ln\left(e^{y/2}\right) \\
&= 2\ln(x) + \ln(z) - \frac{y}{2}\ln(e) \\
&= 2\ln(x) + \ln(z) - \frac{y}{2}.
\end{aligned}
$$

(b) Writing out the terms of the long-products in the numerator and denominator:

$$\frac{2^1 \times 2^2 \times 2^3 \times \ldots \times 2^{99} \times 2^{100}}{2^2 \times 2^3 \times \ldots \times 2^{99} \times 2^{100}}.$$

Every factor is shared by the numerator and the denominator, except for $2^1 = 2$. Therefore every factor except 2 cancels, and the expression reduces to 2.

(c) Writing out the individual terms in the summation:

$$(5^1 - 5^0) + (5^2 - 5^1) + (5^3 - 5^2) + (5^4 - 5^3) + \ldots + (5^{N-1} - 5^{N-2}) + (5^N - 5^{N-1}).$$

If we drop the parentheses and change the order in which the terms are added and subtracted, we can rewrite the sum as

$$= -5^0 + 5^1 - 5^1 + 5^2 - 5^2 + 5^3 - 5^3 + \ldots + 5^{N-2} - 5^{N-2} + 5^{N-1} - 5^{N-1} + 5^N.$$

Every power of 5 is both added and subtracted, except for $-5^0 = -1$ and 5^N, which as the first and last terms to appear in the summation only appear once. Since every other term cancels out, we are left with

$$5^N - 1.$$

(d) First, observe that multiplication inside a logarithm can be rewritten as addition outside a logarithm. The same logic can be extended – long-products inside a logarithm can be rewritten as summations outside a logarithm:

$$\sum_{i=1}^{N} \ln(2e^{a_i}).$$

The log can be broken up into the sum of the log of each factor,

$$\sum_{i=1}^{N} \Big(\ln(2) + \ln(e^{a_i}) \Big),$$

the natural log cancels the exponential base of the second logarithm,

$$\sum_{i=1}^{N} \Big(\ln(2) + a_i \Big),$$

the summation can be distributed to each term,

$$\sum_{i=1}^{N} \ln(2) + \sum_{i=1}^{N} a_i,$$

and since the first logarithmic term does not contain the summation index i, this term is the same thing added to itself N times, so we can write it as

$$N \ln(2) + \sum_{i=1}^{N} a_i.$$

8. (a)

$$\begin{aligned} f(y) &= \log\Big(p^y\Big) + \log\Big((1-p)^{1-y}\Big) \\ &= y\log(p) + (1-y)\log(1-p). \end{aligned}$$

(b) A logarithm is only defined when the expression inside the parentheses is greater than zero. The expression

$$y \log(p) + (1 - y) \log(1 - p)$$

contains two logarithms, both of which contain p. The first logarithm is defined only when $p > 0$, and the second logarithm is defined only when $(1 - p) > 0$, which implies that $p < 1$. These two conditions are both true only when $0 < p < 1$.

9. (a) Multiply both sides by $x + 2$:

$$(3x - 4)(x + 2) = 2x^2 - 4x - 13.$$

Using FOIL on the left-hand side:

$$3x^2 + 2x - 8 = 2x^2 - 4x - 13.$$

Bringing all the terms to one side of the equation and combining like terms:

$$x^2 + 6x + 5 = 0.$$

Two numbers that add to 6 and multiply to 5 are 5 and 1, so the equation factors to

$$(x + 5)(x + 1) = 0.$$

The solutions are therefore

$$x = -5 \text{ and } x = -1.$$

Note that had $x = -2$ been a solution we would have had to disregard it as it would have placed a 0 in the denominator of the right-hand side of the original equation.

(b) To begin with, notice that we will have to disregard any solution that is less than $x = -13/4$, as this would place a negative value underneath the square root. First, square both sides of the equation:

$$4x + 13 = (x + 2)^2,$$

$$4x + 13 = x^2 + 4x + 4.$$

Combining like terms:

$$0 = x^2 - 9.$$

Note that $x^2 - 9$ is a difference of squares, which factors to

$$0 = (x + 3)(x - 3),$$

so the solutions are

$$x = -3 \text{ and } x = 3.$$

$x = -3$ is a valid solution since it is not less than -13/4 = -3.25.

(c) First, use the rule of exponents that says to add the exponents of two factors that share the same base:

$$10^{3x^2+x} = 100.$$

Next, cancel the base of 10 by taking the common-log (the log with a base of 10) of both sides.

$$\log\left(10^{3x^2+x}\right) = \log(100),$$

$$3x^2 + x = \log(100).$$

Note that $100 = 10^2$, so the left-hand side reduces to

$$3x^2 + x = \log\left(10^2\right),$$

$$3x^2 + x = 2\log(10),$$
$$3x^2 + x = 2.$$

Bring all the terms to one side:

$$3x^2 + x - 2 = 0.$$

Check to see if the right-hand side factors neatly. In this quadratic expression, $A = 3$, $B = 1$, and $C = -2$. First calculate that $A \times C = -6$. The factor pairs of -6 are (1 and -6), (2 and -3), (3 and -2), and (6 and -1). The pair (3 and -2) adds to $B = 1$. So we break the middle term into

$$3x^2 + 3x - 2x - 2 = 0,$$

and group the expression into

$$(3x^2 + 3x) + (-2x - 2) = 0$$

$3x$ factors out of the first set of parentheses and -2 factors out of the second:

$$3x(x + 1) - 2(x + 1) = 0.$$

We can now factor $(x + 1)$ out of both terms:

$$(3x - 2)(x + 1) = 0.$$

So the solutions are $x = 2/3$ and $x = -1$.

(d) Multiply out the left-hand side:

$$x - x^2 < -6$$

In order to write the x^2 with a positive coefficient, multiply both sides by -1. Note that this action requires that we flip the inequality sign:

$$x^2 - x > 6$$

Subtract 6 from both sides:

$$x^2 - x - 6 > 0.$$

Two numbers that add to -1 and multiply to -6 are 2 and -3, so the left-hand side of the inequality factors to:

$$(x + 2)(x - 3) > 0.$$

The left-hand side is 0 when $x = -2$ or $x = 3$, but the question requires us to find the region in which the $(x + 2)(x - 3)$ is negative. We need to test values of x that are (1) less than -2, (2) between -2 and 3, and (3) greater than 3. First we consider $x = -3$:

$$(-3 + 2)(-3 - 3) = -1 \times -6 = 6,$$

so numbers in this region satisfy the inequality. Next we try $x = 0$:

$$(0 + 2)(0 - 3) = 2 \times -3 = -6,$$

so x values between -2 and 3 do not satisfy the inequality. Finally, we try $x = 4$:

$$(4+2)(4-3) = 6 \times 1 = 6,$$

so numbers in this region satisfy the inequality. Therefore the solution set consists of any x which is less than -2 or greater than 3.

10. (a) Here is one way to proceed. We start with

$$\ell(\mu) = \ln\left(\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-.5(x_i-\mu)^2}\right).$$

Remember that $\log(ab) = \log(a) + \log(b)$, which means that multiplication inside a logarithm becomes addition outside a logarithm. That remains true even if we are multiplying a lot of terms together, and it still remains true if we represent this multiplication with a long product. So a long-product inside a logarithm becomes a summation outside the logarithm:

$$\ell(\mu) = \sum_{i=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi}} e^{-.5(x_i-\mu)^2}\right).$$

Now we can deal with the terms inside the logarithm. First we break the logarithm up into

$$\ell(\mu) = \sum_{i=1}^{N} \left[\ln\left(\frac{1}{\sqrt{2\pi}}\right) + \ln\left(e^{-.5(x_i-\mu)^2}\right)\right].$$

One property of summations is that $\sum_i (a_i + b_i) = \sum_i a_i + \sum_i b_i$. That's just another way of saying that we can rearrange terms that are being added together in any order we like. So we can break up the summation as follows:

$$\ell(\mu) = \sum_{i=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi}}\right) + \sum_{i=1}^{N} \ln\left(e^{-.5(x_i-\mu)^2}\right).$$

Let's work with the second term first. The natural-log cancels an exponential base of e:

$$\ell(\mu) = \sum_{i=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi}}\right) + \sum_{i=1}^{N} \left(-.5(x_i-\mu)^2\right),$$

and by the same distribution property that says $(ab + ac) = a(b + c)$, the -.5 can be brought outside the sum

$$\ell(\mu) = \sum_{i=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi}}\right) - .5 \sum_{i=1}^{N} (x_i-\mu)^2.$$

Now let's work with the first term. We can rewrite the square root and the reciprocal as a power of $-1/2$:

$$\ell(\mu) = \sum_{i=1}^{N} \ln\left((2\pi)^{-1/2}\right) - .5 \sum_{i=1}^{N} (x_i-\mu)^2.$$

The exponent can be brought outside the logarithm, and also outside the summation:

$$\ell(\mu) = -\frac{1}{2} \sum_{i=1}^{N} \ln(2\pi) - .5 \sum_{i=1}^{N} (x_i-\mu)^2.$$

Notice that $\ln(2\pi)$ is a constant, and a constant that is added to itself N times is equal to N times the constant. So we can write the function as

$$\ell(\mu) = -\frac{1}{2}N\ln(2\pi) - .5\sum_{i=1}^{N}(x_i - \mu)^2,$$

which can be rewritten as

$$\ell(\mu) = \frac{-\ln(2\pi)}{2}N - .5\sum_{i=1}^{N}(x_i - \mu)^2.$$

(b) The logarithm turned the long-product into an expression that only contains a summation. The logarithm also canceled out the exponential base. In general, it is much easier to deal with addition than multiplication, and it is easier to deal with factors than exponents. Researchers prefer to take the log of their likelihood functions precisely because the logarithm turns multiplication into addition, and turns exponents into factors.

2 Sets and Functions

1. (a) A function maps each input to one and only one output. The domain is the set of possible inputs, and the range is the set of possible outputs. Remember that a set is a collection of objects that are often, *but not necessarily*, numbers. In this case, the domain is the set of all humans, and the range is the set of all zombies. In the social sciences (in subfields other than the geopolitics of zombie apocalypses) we can use functions to map non-numbers to other non-numbers. For example, in game theoretical economics we map an actor's preferences to the actor's actions, in political science we can map a country's governmental regime type to its level of journalistic freedom, and in psychology we map certain compositions of neurochemicals to particular psychological disorders. Perhaps, for precision, we might measure these concepts with numbers, but there is nothing inherently numeric about these functions.

(b) Originally, there is 1 zombie. After 1 day, there are 3. On day 2 there are $9 = 3 \times 3 = 3^2$. On day 3 there are $27 = 3 \times 3 \times 3 = 3^3$. In general, after x days, there are

$$f(x) = 3^x$$

zombies.

(c) We simply plug 28 into the function from part (b):

$$f(28) = 3^{28} = 22,876,790,000,000 \text{ [22.8 \textit{trillion}] zombies.}$$

(d) We set $f(x) = 318,851,733$, and we solve for x:

$$3^x = 318,851,733.$$

We take the logarithm, base-3, of both sides in order to cancel out the exponential base of 3:

$$\log_3(3^x) = \log_3(318,851,733),$$

$$x = \log_3(318,851,733).$$

In order to get a computer or calculator to evaluate the value of this logarithm, we can convert it to a natural logarithm as follows:

$$x = \frac{\ln(318,851,733)}{\ln(3)} \approx 17.8 \text{ days}.$$

2. (a)

$$f(x) = \sqrt{\frac{x}{2}},$$

we plug in x^2 for x:

$$f(x^2) = \sqrt{\frac{(x^2)}{2}},$$

$$f(x^2) = \frac{\sqrt{x^2}}{\sqrt{2}},$$

$$f(x^2) = \frac{x}{\sqrt{2}}.$$

This answer is fine, but if you want to remove the square root term from the denominator, you can multiply the top and bottom of this fraction by $\sqrt{2}$:

$$f(x^2) = \frac{\sqrt{2}x}{2}.$$

(b)

$$f(x) = \sqrt{\frac{x}{2}}.$$

First we replace $f(x)$ with y:

$$y = \sqrt{\frac{x}{2}},$$

then we interchange x and y:

$$x = \sqrt{\frac{y}{2}},$$

and we solve for y:

$$x^2 = \frac{y}{2},$$

$$y = 2x^2.$$

Finally we substitute $f^{-1}(x)$ in for y:

$$f^{-1}(x) = 2x^2.$$

(c)

$$\begin{aligned}(f \circ g)(y) &= f(g(y)) = f(y^2 - 2y + 4) \\ &= \sqrt{\frac{y^2 - 2y + 4}{2}}.\end{aligned}$$

(d)

$$\begin{aligned}(g \circ f)(x) &= g(f(x)) = g\left(\sqrt{\frac{x}{2}}\right) \\ &= \left(\sqrt{\frac{x}{2}}\right)^2 - 2\left(\sqrt{\frac{x}{2}}\right) + 4 \\ &= \frac{x}{2} - 2\sqrt{\frac{x}{2}} + 4.\end{aligned}$$

(e)

$$\begin{aligned}(f \circ f)(x) &= f(f(x)) = f\left(\sqrt{\frac{x}{2}}\right) \\ &= \sqrt{\frac{\sqrt{\frac{x}{2}}}{2}} \\ &= \sqrt{\frac{\sqrt{\frac{x}{2}}}{2}}.\end{aligned}$$

But square roots within square roots are pretty messy. So let's convert everything to exponents, remembering that exponents in the denominator can be expressed as negative exponents:

$$\begin{aligned}&\left(\frac{\sqrt{\frac{x}{2}}}{2}\right)^{1/2} \\ &= \left(\frac{(\frac{x}{2})^{1/2}}{2}\right)^{1/2} \\ &= \left(\frac{x^{1/2}2^{-1/2}}{2}\right)^{1/2} \\ &= \left(x^{1/2}2^{-1/2}\right)^{1/2} 2^{-1/2} \\ &= x^{1/4}2^{-1/4}2^{-1/2} \\ &= 2^{-3/4}x^{1/4} \\ &= \sqrt[4]{\frac{x}{2^3}} \\ &= \sqrt[4]{\frac{x}{8}}.\end{aligned}$$

You didn't necessarily have to simplify this much, but I wanted to show how much the question could possibly be simplified.

(f)

$$\begin{aligned}(h \circ k)(w) &= h(k(w)) = h(2\ln(w)) \\ &= e^{-2\ln(w)/2} \\ &= e^{-\ln(w)}.\end{aligned}$$

Recall that for any logarithm, we have the property that $-\ln(w) = \ln(1/w)$. So the equation becomes

$$e^{\ln(1/w)},$$

and now the exponential base and the natural logarithm cancel, leaving us with

$$\frac{1}{w}.$$

(g)

$$\begin{aligned}(k \circ h)(z) &= k(h(z)) = k(e^{-z/2}) \\ &= 2\ln(e^{-z/2}).\end{aligned}$$

The natural logarithm and the exponential base cancel each other out and we are left with

$$= 2(-z/2)$$

$$= -z.$$

3. (a) The function contains a logarithm, a square root, and a fraction. Only positive numbers can be placed inside a logarithm (you can't take the log of 0), and only non-negative numbers can be placed inside a square root (unlike a log, you can take the square-root of 0). Values of x that are strictly greater than 3 are allowed inside this particular logarithm, and values of x which are 5 or less are allowed inside this square root. Finally, we cannot have a fraction with 0 in the denominator, which occurs when $x = 5$, so 5 cannot be included in the domain. So all together, the function accepts real number values of x that are greater than 3 and less than 5. In set-builder notation, this statement is

$$\{x \in \mathbf{R} \mid 3 < x < 5\}.$$

(b) A fraction is 0 when the numerator is equal to 0. So we set

$$\ln(x-3) = 0.$$

We cancel out the natural logarithm by setting both sides as exponents with a base of e:

$$e^{\ln(x-3)} = e^0,$$

$$x - 3 = 1,$$
$$x = 4.$$

We have to check whether this value of x exists in the domain, and by our answer in part (a), it does. An alternative way to solve the problem would be to observe that $\ln(1) = 0$ and then to set $x - 3 = 1$.

(c) TRUE. I plug values of x which are closer and closer to 5 into the function:

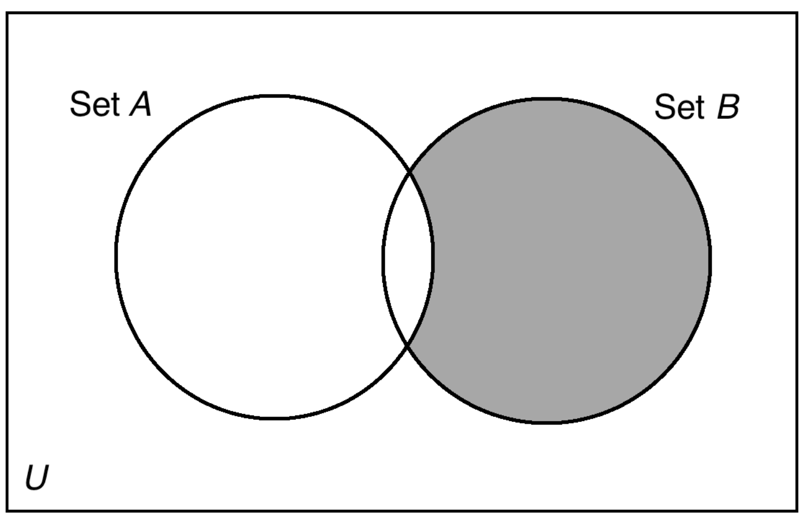
x	$f(x)$
4.9	2.03
4.99	6.88
4.999	21.90
4.9999	69.31
4.99999	219.19
4.999999	693.15
4.99999999999	219192.38

So it appears that the values of the function are definitely approaching ∞ as x gets closer to 5. On the other side, I hit the limit on the number of digits the spreadsheet program I'm using can handle before rounding, but the closest I can get to 3 is:

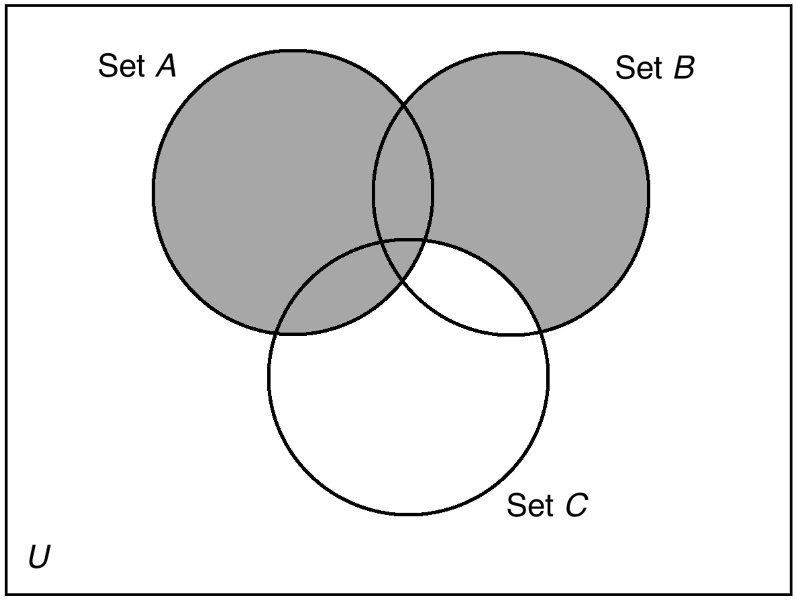
x	$f(x)$
3.1	-1.67
3.01	-3.26
3.001	-4.89
3.0001	-6.51
3.00001	-8.14
3.00000000000001	-22.78

While -22.78 is not indicative of a function whose values are exploding towards $-\infty$, there is also nothing in the behavior of the function to indicate that the decrease is leveling off as I approach 3. If I didn't hit the point where the spreadsheet program rounded by input to 3, I could have made this number as low as I liked.

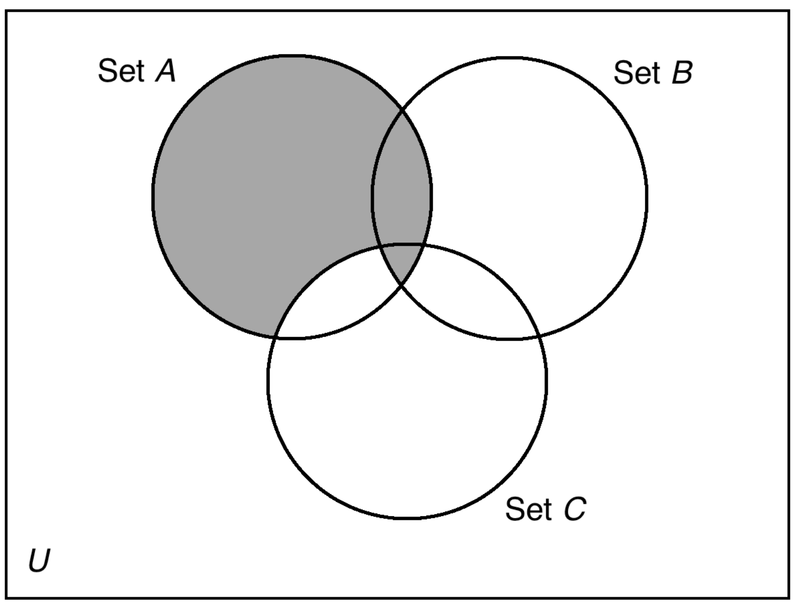
4. (a) As described in the text, $\tilde{A} \cap B$ is everything in set B that's also outside of set A:



(b) $A \cup (B \cap \tilde{C})$ is everything in set A, as well as everything that is both in B and outside of C:

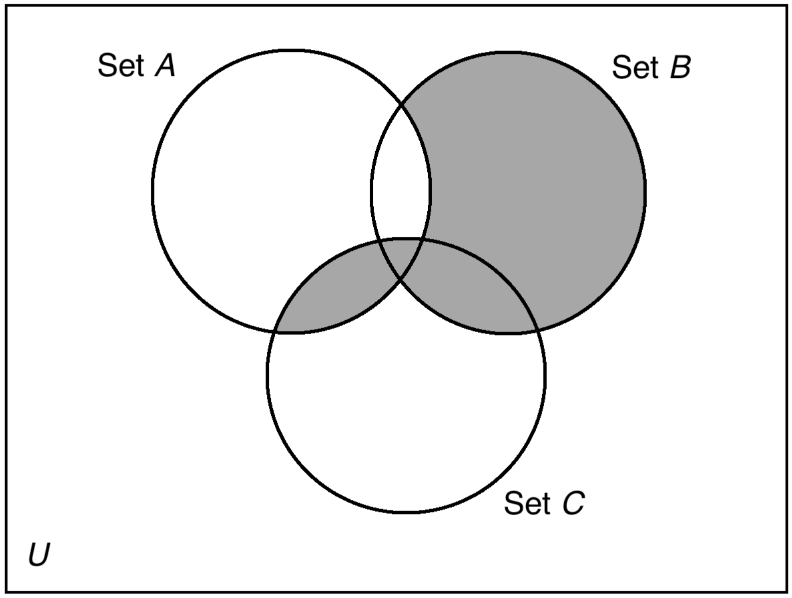


(c) $A \cap (B \cup \tilde{C})$ is everything that is both in set A and either in set B or outside C:



Here, the one part of set A that is not shaded is neither in set B, nor is outside C, so while it is inside A it does not satisfy the second requirement.

(d) $(\tilde{A} \cap B) \cup (A \cap C)$ consists of everything in B that's outside of A and everything in the intersection of A and C:



5. (a) All vegetables that aren't red (sorry tomatoes!)

   (b) All countries that are either non-democratic or economically underdeveloped, or both

   (c) All real numbers from 5 to 7, including 5, but not including 7

   (d) All real numbers strictly greater than 3

   (e) Integers that are at least as big as 2

   (f) A technical and literal interpretation is: all real numbers such that the number divided by 4 is an integer. A better, plain language is interpretation is: all real numbers that are divisible by 4.

6. (a) $\{x \in \mathbf{R} \mid -5 \leq x < 4\}$

   (b) $\{x \in \mathbf{R} \mid x > 12\}$

   (c) An integer is divisible by 3 if the fraction of that integer over 3 is also an integer (so no decimals): $\{x \in \mathbf{Z} \mid \frac{x}{3} \in \mathbf{Z}\}$

(d) Note, it's not actually necessary to solve this equation in order to express the set of solutions in set-builder notation. Simply put, this set is the set of real numbers that solve the equation $x^3 - 7x + 6 = 0$: $\{x \in \mathbf{R} \mid x^3 - 7x + 6 = 0\}$

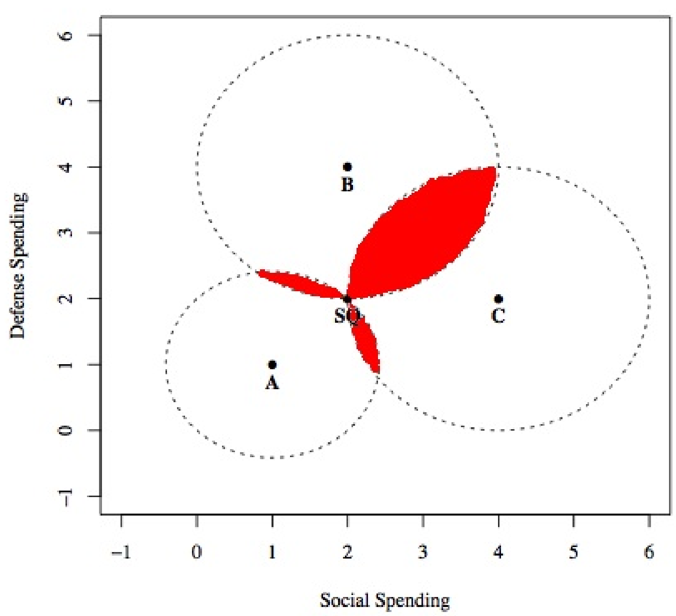
7. (a) A proposal will pass if it is voted for by:
A and B, or A and C, or B and C, or A and B and C.

We translate the above statement into set notation by using parentheses to separately denote the cases in which a bill will pass,
(A and B), or (A and C), or (B and C), or (A and B and C),

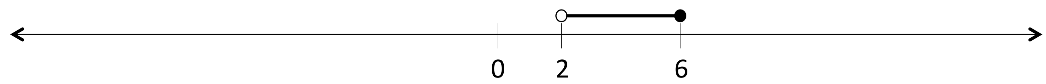
and then by replacing the "and" statements with intersections and the "or" statements with unions:
$(A \cap B) \cup (A \cap C) \cup (B \cap C) \cup (A \cap B \cap C)$.

Since the only bill that all three legislators can support $(A \cap B \cap C)$ is the same as the status quo, another acceptable answer is just
$(A \cap B) \cup (A \cap C) \cup (B \cap C)$.

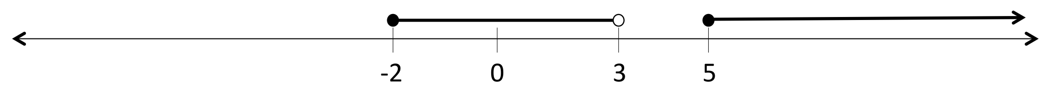
(b) The win-set is shaded in the graph below:



8. (a) This interval contains all real numbers between 2 and 6, not including 2 but including 6:



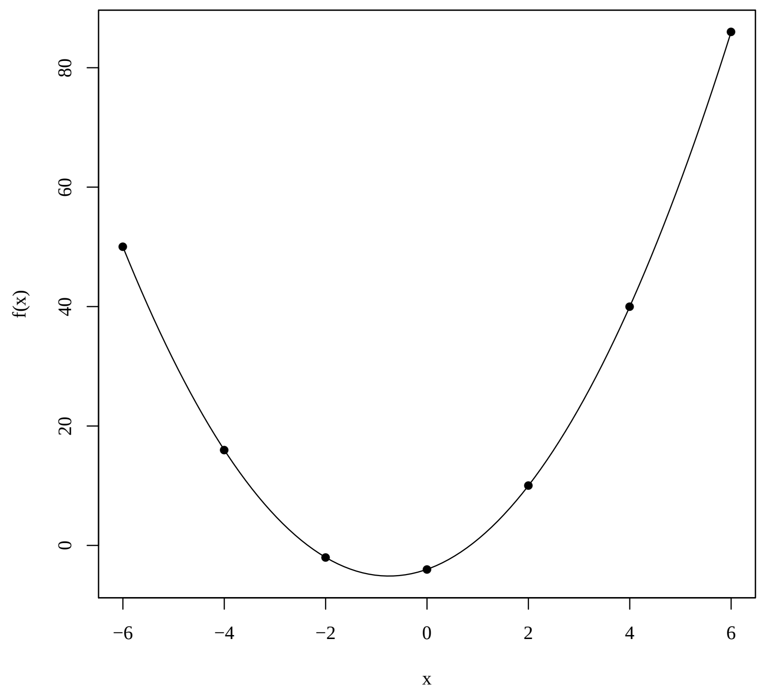
(b) This interval contains all real numbers between -2 and 3, including -2 but not including 3, as well as all real numbers greater than or equal to 5:



(c) For the following 4 graphs, it helps to list some points in a table by choosing a set of candidate x values and plugging them into the function. These points can be plotted, and these points suggest the overall shape of the graph. Some points that are on the graph of $f(x) = 2x^2 + 3x - 4$ are

x	$f(x)$
-6	50
-4	16
-2	-2
0	-4
2	10
4	40
6	86

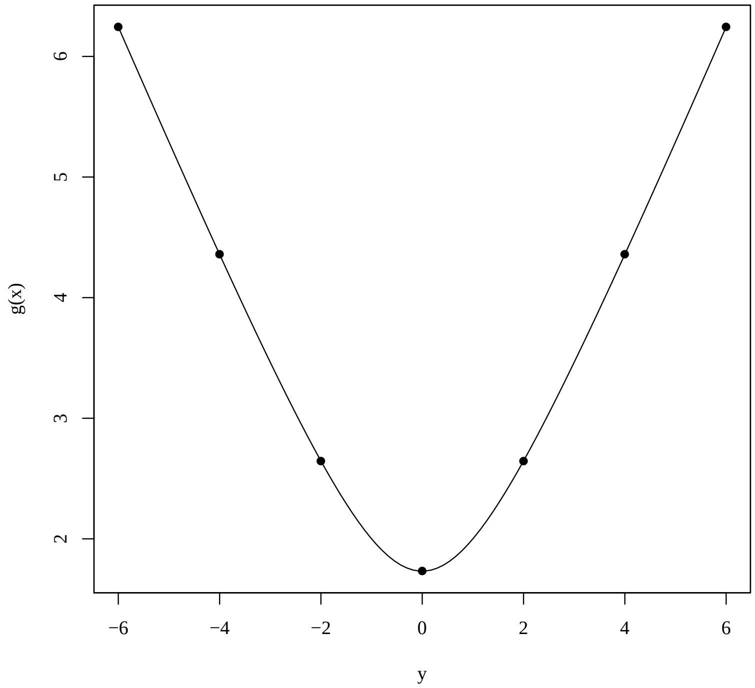
The graph of $f(x)$ is



(d) Some points that are on the graph of $g(y) = \sqrt{y^2 + 3}$ are

y	$g(y)$
-6	6.24
-4	4.36
-2	2.65
0	1.73
2	2.65
4	4.36
6	6.24

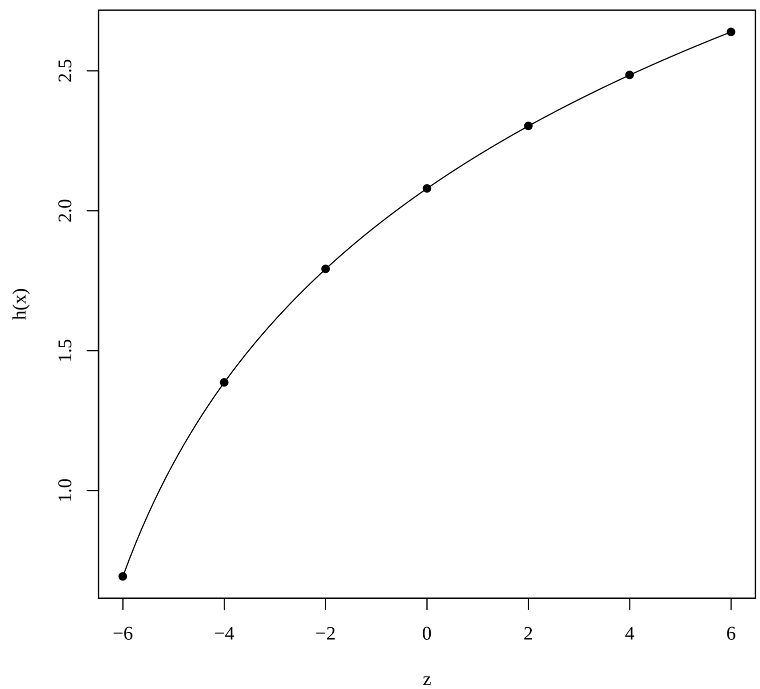
The graph of $g(y)$ is



(e) Some points that are on the graph of $h(z) = \ln(z + 8)$ are

z	$h(z)$
-6	0.69
-4	1.39
-2	1.79
0	2.08
2	2.30
4	2.48
6	2.64

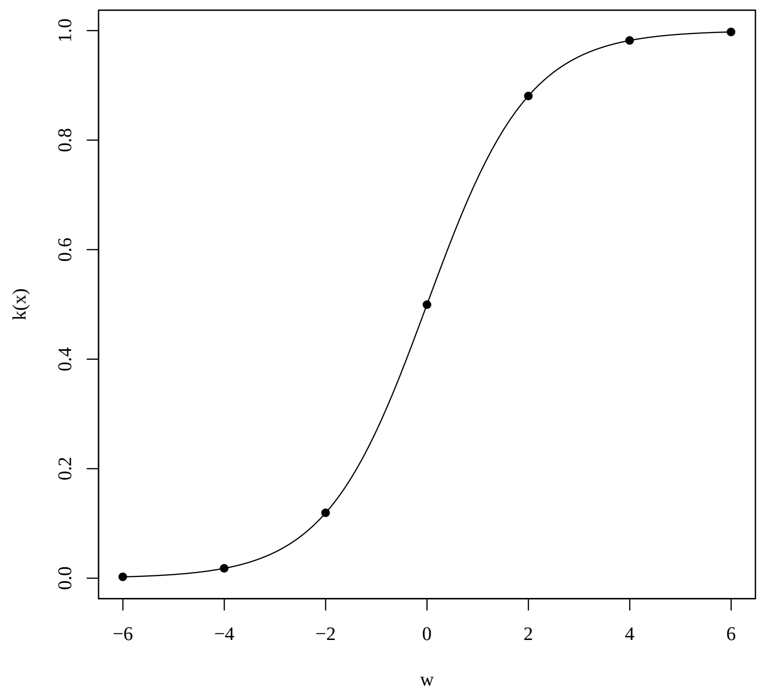
The graph of $h(z)$ is



(f) Some points that are on the graph of $k(w) = \frac{e^w}{1+e^w}$ are

w	$k(w)$
-6	0.00
-4	0.02
-2	0.12
0	0.50
2	0.88
4	0.98
6	1.00

The graph of $k(w)$ is



9. (a) The slope of a line that passes through the points (1,2) and (3,8) is

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{8-2}{3-1} = \frac{6}{2} = 3.$$

To find the y-intercept, we plug the slope and one of the points into

$$y = mx + b,$$

$$2 = (3)1 + b,$$

$$b = 2 - 3 = -1.$$

So this line is

$$y = 2x - 3.$$

(b) In this case, given a point and the y-intercept, we have enough information to plug into

$$y = mx + b$$

to solve for the slope. Plugging this information in gives us

$$1 = m(3) - 2,$$

$$3 = 3m,$$

$$m = 1.$$

So the line is

$$y = x - 2.$$

(c) In this case, given a point and the slope, we have enough information to plug into

$$y = mx + b$$

to solve for the y-intercept. Plugging this information in gives us

$$6 = -2(-3) + b,$$

$$6 = 6 + b,$$

$$b = 0.$$

So the line is

$$y = -2x.$$

(d) There are many examples of third degree polynomials with roots at -2, 1, and, 3. But the simplest example is

$$f(x) = (x + 2)(x - 1)(x - 3).$$

It's not necessary to multiply these terms out. Because three x terms are multiplied together, the largest power will be x^3. But to see this explicitly, begin by applying FOIL to the first two parenthetical terms,

$$f(x) = (x^2 + x - 2)(x - 3),$$

then distribute x and -3 to each of the quadratic terms,

$$f(x) = x^3 + x^2 - 2x - 3x^2 - 3x + 6,$$

and combine the like terms:

$$f(x) = x^3 - 2x^2 - 5x + 6.$$

(e) There are many examples of fifth degree polynomials with roots only at -4, 2, and, 5. But a simple example is

$$f(x) = (x + 4)^2(x - 2)^2(x - 5).$$

It's not necessary to multiply these terms out. Because two x^2 terms and one x term are multiplied together, the largest power will be x^5.

10. The symbols are individually translated as in the following table:

Symbol	Translation
U	The set of proposals that legislator 1 can make that will pass unanimously
$\equiv$	is equivalent to
$\{\ldots\}$	the set of
(x_2^1, x_3^1)	allocations to legislators 2 and 3 in the first proposal
$\mid$	such that
$x_j^1 \geq \delta v_j,\ j = 2, 3$	both legislator 2 and 3 receive allocations that are greater than or equal to the amount they expect from the next proposal, multiplied by the discount factor
$x_2^1 + x_3^1 \leq 1$	where the sum of legislator 2 and 3's allocations is less than or equal to 1
$\cap$	and
A	the allocations for legislators 1, 2, and 3 add up to exactly 1.

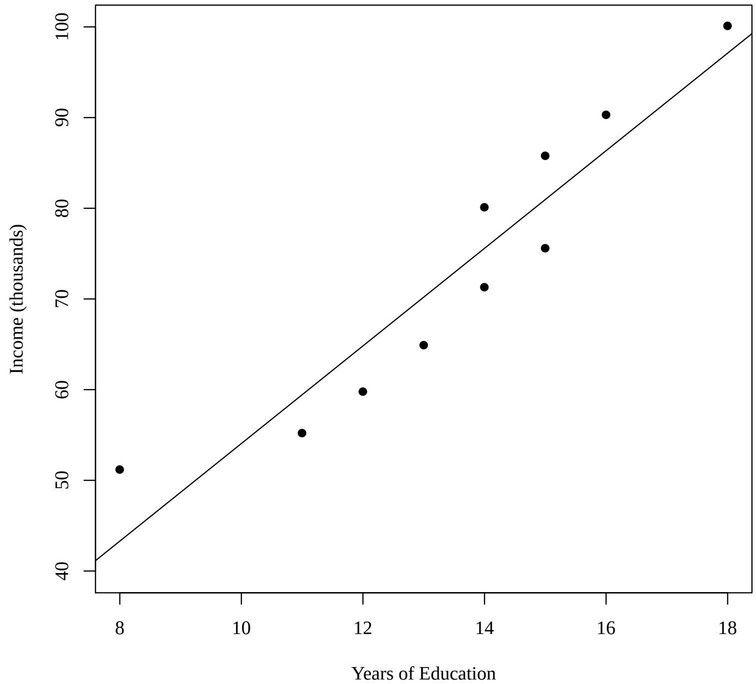
All together, the translation is

> "The set of proposals that legislator 1 can make that will pass unanimously is equivalent to the set of allocations to legislators 2 and 3 in the first proposal such that both legislator 2 and 3 receive allocations that are greater than or equal to the amount they expect from the next proposal multiplied by the discount factor, where the sum of legislator 2 and 3's allocations is less than or equal to 1, and the allocations for legislators 1, 2, and 3 add up to exactly 1."

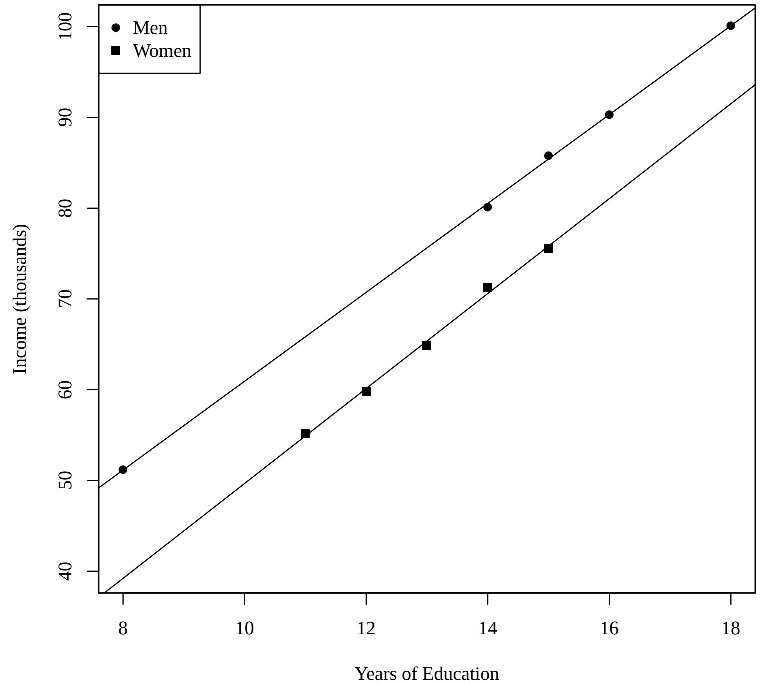
That's an awkward and wordy sentence. Like any English translation of a foreign language, direct word-for-word translations tend to be awkward and wordy. We can better express the spirit of the statement as something like

> "Any proposal by legislator 1 to allocate to legislators 2 and 3 at least as much as they would expect in later rounds, discounted for the penalty they incur by failing to pass a proposal right away, will pass the legislature unanimously provided that the allocations to legislators 2 and 3 do not exceed the amount of available resources and all available resources are allocated."

11. (a) The scatterplot of education and income, with the best fit line, is:



(b) Plotting the men with dots and the women with squares, and including separate best fit lines for men and for women:



The two lines are nearly parallel. After accounting for education, there appears to be a persistent wage gap between men and women that does not diminish with higher levels of education, although both men and women gain more income with more education.

(c) The constant is interpreted as follows: on average, a man (**Sex**=0) with no formal education (**Education**=0) makes $11,356.

The coefficient on **Education** is interpreted as follows: an increase of one year in the number of years of formal education is associated with a $4,939 increase in income, after accounting for the effect of sex on income.

The coefficient on **sex** is interpreted as follows: women, as compared to men, have incomes that are $10,174 lower, on average, after accounting for the number of years of formal education.

(d) A women with 17 years of formal education has the following data: **Sex** = 1 and **Education**=17. Plugging these values into the regression model gives us:

$$\textbf{Income}_i = 11{,}356 + 4{,}939(17) - 10{,}174(1) + \varepsilon_i,$$

$$\textbf{Income}_i = 85{,}145 + \varepsilon_i.$$

The error ε_i represents the variation that we cannot explain – all of the ways we can be wrong about this prediction. But these errors are just as likely to be too high as too low, so on average, we expect a woman with 17 years of formal education to make $85,145.

3 Probability

1. (a) This problem asks you to derive the probability of the union of two events, as implied by the fact that the question uses the word "or." The first event A is "the randomly drawn member of Congress is a Republican" and the second event B is "the randomly drawn member of Congress is a Senator." The probability of the union of two events is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

There are 533 total members of Congress, and 233+46=279 Republicans in Congress, so the probability of A is

$$P(A) = \frac{279}{533} = .523.$$

There are 100 Senators, so the probability of B is

$$P(B) = \frac{100}{533} = .188.$$

The intersection of A and B refers to the number of Republican Senators, of which there are 46. The probability of the intersection is

$$P(A \cap B) = \frac{46}{533} = .086.$$

So the probability that a randomly drawn member of Congress is a Republican or a Senator is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = .523 + .188 - .086 = .625.$$

(b) The question asks how many ways are there to choose a group of 13 out of a larger group of 433. Order doesn't matter because if we're not considering rank and chairmanship all committee members have equal standing. That means we are to find the following combination:

$$\binom{433}{13} = \frac{433!}{13!(433-13)!} = \frac{433!}{13!420!}$$

$$= \frac{433 \times 432 \times 431 \times 430 \times 429 \times 428 \times 427 \times 426 \times 425 \times 424 \times 423 \times 422 \times 421}{13 \times 12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}$$

$$\approx 2,518,055,000,000,000,000,000,000 \text{ (2.5 } \textit{Septillion}!)$$

2. (a) A student whose parents belong to a country club probably comes from a family with some wealth. Wealthier families can afford private SAT lessons for their children. Therefore these events are conditionally independent, and are related only insofar as they both depend on the wealth of the student's family.

(b) These events are probably dependent since better schools and libraries would directly cause a lower rate of illiteracy.

(c) Two mutually exclusive events share no outcomes. That means that if one event occurs, then the other cannot occur. Therefore the two events are dependent.

This is a tricky question. Be sure not to confuse mutually exclusive with independent. An example of two mutually exclusive events might be $A =$"a person votes for the Democrat" and $B =$"a person votes for the Republican." These events are not independent because if a person votes for the Democrat he or she by definition does not vote for the Republican. In contrast, if two events are independent, then whether one occurs or not has no bearing on whether the other will occur or not.

3. There are many, many examples. One example is
$A =$"there are more than 200 visitors to Monticello (Thomas Jefferson's home, now a historical museum near Charlottesville, VA) today", and
$B =$"the Japanese stock market rises 100 points."
These events are almost certainly independent. Unless, of course, one of the visitors to Monticello is a Japanese stock broker on vacation who realizes that the elderberry wine produced in central Virginia is delicious and places an unprecedented order for elderberries on the Japanese commodities exchange and sets off a chain of events that results in the recovery of the Japanese and Virginian economies and the invention of Bluegrass Kabuki theater.

4. By the definition of a combination we know that

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

If we plug $n-k$ in for k in the above equation, we get

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

which is the same thing as before. Therefore

$$\binom{n}{k} = \binom{n}{n-k}.$$

5. (a) The event A is "at least 2 people in a class of 30 share a birthday," so the complement $\tilde{A}$ is "no one in a class of 30 shares a birthday with anyone else in the class." The sample space S consists of "every way the 30 people in the class can have a birthday."

(b) There are 365 days in a year. Consider the 30 students as if they are lined up. The first student can have any of 365 birthdays. The second student cannot share a birthday with the first student, but can have any of the 364 remaining possibilities. Likewise the 3rd student has 363 possibilities, and so on, until the 30th student has 336 possibilities. All together, multiplying these stages together, there are

$$365 \times 364 \times 363 \times \ldots \times 336 \approx 2.17 \times 10^{76} \text{ possibilities.}$$

An alternative way to think about this problem is that we are trying to choose 30 birthdays out of the set of 365 possible birthdays. Order matters here since if two students switch their birthdays, that's a different outcome. We use a permutation:

$$_{365}P_{30} = \frac{365!}{(365-30)!} = 365 \times 364 \times 363 \times \ldots \times 336 \approx 2.17 \times 10^{76} \text{ possibilities.}$$

(c) For the sample space, the event with all possible outcomes, we don't care whether two students share a birthday or not. Each student has 365 possible birthdays. So the 30 students have

$$365^{30} \approx 7.4 \times 10^{76}.$$

possible combinations of birthdays.

(d) The probability of event A, "at least 2 people in a class of 30 share a birthday," is

$$P(A) = 1 - P(\tilde{A}) = 1 - \frac{|\tilde{A}|}{|S|}.$$

In part (b) we derived

$$|\tilde{A}| = 365 \times 364 \times 363 \times \ldots \times 336 \approx 2.17 \times 10^{76},$$

and in part (c) we derived

$$|S| = 365^{30} \approx 7.4 \times 10^{76}.$$

So the probability of A is

$$1 - \frac{365 \times 364 \times 363 \times \ldots \times 336}{365^{30}} = 0.706,$$

which is a surprisingly high result.

6. (a) A TRUE response (X) happens when the respondent has taken a bribe (W) **AND** has spun A (Y), **OR** has never taken a bribe ($\tilde{W}$) **AND** has spun a B ($\tilde{Y}$).

If we replace the events with their symbols, the word AND with intersections, and the word OR with a union, the statement becomes

$$X = (W \cap Y) \cup (\tilde{W} \cap \tilde{Y}).$$

(b) The events W and Y are independent because the probability of taking a bribe does not affect the probability of spinning an A or B, and the probability of spinning A or B does not affect the probability of taking a bribe.

The events X and Y are NOT independent because the probability of stating TRUE depends on the probability of spinning A or B.

(c) We start with our answer in part (a):

$$X = (W \cap Y) \cup (\tilde{W} \cap \tilde{Y}).$$

Then the probability of X is

$$P(X) = P\Big((W \cap Y) \cup (\tilde{W} \cap \tilde{Y})\Big).$$

If we apply the formula for the probability of a union, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, we get

$$P(X) = P(W \cap Y) + P(\tilde{W} \cap \tilde{Y}) - P\Big((W \cap Y) \cap (\tilde{W} \cap \tilde{Y})\Big).$$

First consider the last term. The event $W \cap Y$ represents officials who've taken a bribe and spun A. The event $\tilde{W} \cap \tilde{Y}$ represents officials who've never taken a bribe and who've spun B. These two events are mutually exclusive because they share no elements. That means that their intersection is the empty set, and the probability of the empty set is 0:

$$P\Big((W \cap Y) \cap (\tilde{W} \cap \tilde{Y})\Big) = 0.$$

That leaves us with

$$P(X) = P(W \cap Y) + P(\tilde{W} \cap \tilde{Y}).$$

Finally, recall that in part (b) we found that events W and Y are independent. The same logic must apply to $\tilde{W}$ and $\tilde{Y}$. We apply the formula for the probability of the intersection of two independent events, $P(A \cap B) = P(A)P(B)$ and find

$$P(X) = P(W)P(Y) + P(\tilde{W})P(\tilde{Y}).$$

(d) We start with

$$P(X) = P(W)P(Y) + P(\tilde{W})P(\tilde{Y}).$$

We apply the rule for the probability of a complement, $P(\tilde{A}) = 1 - P(A)$, and find

$$P(X) = P(W)P(Y) + \Big(1 - P(W)\Big)\Big(1 - P(Y)\Big).$$

Multiplying this equation out, it becomes

$$P(X) = P(W)P(Y) + \Big(1 - P(W) - P(Y) + P(W)P(Y)\Big),$$

$$P(X) = 2P(W)P(Y) + 1 - P(W) - P(Y).$$

In order to solve for $P(W)$, we bring every term with $P(W)$ over to one side, and bring every term with $P(W)$ to the other side,

$$P(X) + P(Y) - 1 = 2P(W)P(Y) - P(W),$$

we factor out $P(W)$,

$$P(X) + P(Y) - 1 = P(W)\Big(2P(Y) - 1\Big),$$

and divide to isolate $P(W)$:

$$P(W) = \frac{P(X) + P(Y) - 1}{2P(Y) - 1}.$$

(e) If the spinner had equal sized areas for A and B, then the probability of spinning A would be $P(Y) = 0.5$. But plugging this value into

$$P(W) = \frac{P(X) + P(Y) - 1}{2P(Y) - 1}$$

places 0 in the denominator. Therefore it is only possible to solve for $P(W)$ in this way when the spinner has unequal areas for A and B.

(f) If 35% of respondents select TRUE, then $P(X) = 0.35$, and we know that $P(Y) = 0.75$. We simply plug these probabilities into the equation for $P(W)$:

$$P(W) = \frac{P(X) + P(Y) - 1}{2P(Y) - 1},$$

$$P(W) = \frac{0.35 + 0.75 - 1}{2(0.75) - 1} = \frac{0.1}{0.5} = 0.2.$$

So given these results, we conclude that 20% of government officials have taken bribes.

7. Let H be the event in which I have a headache and let F be the event in which I have the flu. The problem tells us that

- $P(H) = .1$, since I have a headache 1 out of 10 days in any event,
- $P(H|F) = .5$ since half of all flu sufferers have headaches,
- and $P(F) = 0.02$ since 2% of the population will come down with the flu.

What we are trying to find is $P(F|H)$, the probability of coming down with the flu given a headache. Bayes' rule tells us that

$$P(F|H) = \frac{P(H|F)P(F)}{P(H)}.$$

So we simply plug in the corresponding values:

$$P(F|H) = \frac{0.5 \times 0.02}{0.1} = 0.1.$$

8. (a) Let N be the event that an email contains the word "Nigeria" and let S be the event that an email is spam. The question tells us that

   - $P(N|S) = 0.05$ since 5% of all spam messages contain the word "Nigeria,"
   - $P(S) = 0.35$ since 35% of emails are spam,
   - $P(\tilde{S}) = 0.65$ since this is the complement event to S,
   - and $P(N|\tilde{S}) = 0.001$ since only 0.1% of legitimate emails contain the word "Nigeria."

   In order to find $P(S|N)$, the probability that an email that contains the word "Nigeria" is spam, we apply version 2 of Bayes' rule:

$$P(S|N) = \frac{P(N|S)P(S)}{P(N|S)P(S) + P(N|\tilde{S})P(\tilde{S})} = \frac{0.05 \times 0.35}{0.05 \times 0.35 + 0.001 \times 0.65} = 0.96.$$

(b) This question requires us to find the proportions of non-spam with the word "Nigeria," $P(N|\tilde{S})$, such that the probability that an email with "Nigeria" is spam is less than 0.95:

$$P(S|N) < 0.95.$$

First, substitute the Bayes' rule expression we found in part (a) in for $P(S|N)$:

$$\frac{P(N|S)P(S)}{P(N|S)P(S) + P(N|\tilde{S})P(\tilde{S})} < 0.95.$$

Next, plug in what we still know to be true – $P(S) = 0.35$, $P(\tilde{S}) = 0.65$, and $P(N|S) = 0.05$:

$$\frac{0.05 \times 0.35}{(0.05 \times 0.35) + 0.65P(N|\tilde{S})} < 0.95,$$

$$\frac{0.0175}{0.0175 + 0.65P(N|\tilde{S})} < 0.95.$$

Finally, we solve the inequality for $P(N|\tilde{S})$:

$$0.0175 < 0.95\Big(0.0175 + 0.65P(N|\tilde{S})\Big),$$

$$0.0175 < 0.0166 + 0.62P(N|\tilde{S}),$$

$$P(N|\tilde{S}) > \frac{0.0175 - 0.0166}{0.62} = .0015.$$

So if at least 0.15% of non-spam emails contain the word "Nigeria," then emails with this word will no longer be filtered.

(c) We know that, regardless of which version of the email is used, the probability that someone is gullible is $P(G) = 0.05$ and the probability that someone is not is $P(\tilde{G}) = 0.95$. Our goal is to calculate $P(G|R)$, the probability that someone is gullible given that they respond to the email. Let's first consider version 1 of the email. In this version, we know that

   - $P(R|G) = 0.4$ since 40% of gullible people will respond,

- and $P(R|\tilde{G}) = 0.2$ since 20% of non-gullible people respond.

We plug these values into version 2 of Bayes' rule for $P(G|R)$:

$$P(G|R) = \frac{P(R|G)P(G)}{P(R|G)P(G) + P(R|\tilde{G})P(\tilde{G})} = \frac{0.4 \times 0.05}{0.4 \times 0.05 + 0.2 \times 0.95} = 0.095.$$

For the second version of the email,

- $P(R|G) = 0.1$ since 10% of gullible people will respond,
- and $P(R|\tilde{G}) = 0.001$ since only 1 in 1000 non-gullible people respond.

Again, we plug these values into version 2 of Bayes' rule for $P(G|R)$:

$$P(G|R) = \frac{P(R|G)P(G)}{P(R|G)P(G) + P(R|\tilde{G})P(\tilde{G})} = \frac{0.1 \times 0.05}{0.1 \times 0.05 + 0.001 \times 0.95} = 0.84.$$

Therefore version 2 yields a much, much higher probability that the respondents will be gullible.

9. (a) The expected utility of the juror for voting for acquittal is

$$\begin{aligned}\text{EU(acquit)} &= \text{U(acquit an innocent person)}P(\tilde{G}) + \text{U(acquit a guilty person)}P(G), \\ &= 0(1-\pi) - (1-z)\pi = -\pi(1-z).\end{aligned}$$

The expected utility of the juror for voting to convict is

$$\begin{aligned}\text{EU(convict)} &= \text{U(convict an innocent person)}P(\tilde{G}) + \text{U(convict a guilty person)}P(G). \\ &= -z(1-\pi) + 0\pi = -z(1-\pi).\end{aligned}$$

The juror will vote to convict when

$$\begin{aligned}\text{EU(convict)} &> \text{EU(acquit)}, \\ -z(1-\pi) &> -\pi(1-z), \\ -z + z\pi &> -\pi + z\pi, \\ -z &> -\pi, \\ z &< \pi.\end{aligned}$$

Substantively, this result means that the juror will only vote to convict when the juror's belief that the defendant is guilty outweighs her aversion to convicting an innocent person.

(b) We know from part (a) that $P(G) = \pi$ and $P(\tilde{G}) = 1 - \pi$. In addition, the problem tells us that $P(D|\tilde{G}) = p$ and $P(D|G) = q$. We simply plug these values into version 2 of Bayes' rule for $P(G|D)$:

$$P(G|D) = \frac{P(D|G)P(G)}{P(D|G)P(G) + P(D|\tilde{G})P(\tilde{G})} = \frac{q\pi}{q\pi + p(1-\pi)}.$$

(c) If $\pi = 1$ the juror is already convinced that the defendant is guilty before the trial has even taken place. Her posterior belief,

$$\frac{q(1)}{q(1) + p(1-1)} = \frac{q}{q} = 1,$$

is equal to the prior. That is, she remains convinced that the defendant is guilty. In this case, the signal didn't matter. Indeed, the fact that the defense presented a better case should have made the juror less certain about the guilt of the defendant, but because her prior biases were so strong she did not even consider the information revealed by the trial.

If $p = q$, her posterior belief becomes

$$\frac{p\pi}{p\pi + p(1-\pi)} = \frac{p\pi}{p\pi + p - p\pi} = \frac{p\pi}{p} = \pi.$$

In this case her posterior belief is equal to her prior belief in the defendant's guilt. But this result is true for any prior belief, not just the prior that $\pi = 1$. If $p = q$, then the defense would have presented the better case with the same probability regardless of whether the defendant is actually guilty or innocent. This situation can occur if the defense is far overmatched by the prosecution; regardless of the truth, the defense will lose because the lawyers are bad. Counter-intuitively, this situation can also occur if the defense is dominant; the defense will win regardless of the truth because the lawyers are excellent. In either case, the juror is receptive to the signal being sent by the trial. But because the performance of the defense appears to be independent of the truth of the case, the signal is too weak to alter the juror's belief.

Finally, if $p = 1$ and $q = 0$, the juror's posterior becomes

$$\frac{0\pi}{0\pi + 1(1-\pi)} = 0.$$

In this case, the defense is certain to present the better case if the truth is that the defendant is innocent, and is certain to present the weaker case if the defendant is guilty. Therefore the fact that the defense presented the stronger case sends the strongest possible signal to the juror about the defendant's innocence. As a result, the juror updates her belief to be certain of the defendant's innocence, regardless of her prior belief (as long as the prior is less than 1. If the prior where equal to 1 in this case, then the posterior would become 0/0, and the juror's head would simply explode).

10. (a) Each site will be cleared with probability p, where p is .95 if the Syrian government is compliant, and .6 if the government is non-compliant. Whether or not each site gets cleared can be thought of as a series of independent experiments with the same probability of success, so we can calculate the probability of clearing 7 sites using the binomial distribution. Denote the event that the government clears 7 sites as $X = 7$. If the government is compliant, then

$$P(X = 7|C) = \binom{10}{7} .95^7 .05^3 = 0.01,$$

and if the government is non-compliant, the probability is

$$P(X = 7|\tilde{C}) = \binom{10}{7} .6^7 .4^3 = 0.215.$$

(b) We want to find the probability that the government was compliant given that 7 sites were cleared. Since there are only two possible states of the world under consideration here – the one in which the Syrian government is compliant and the one in which the government is non-compliant – we can use the second version of Bayes' rule. The equation is

$$P(C|X=7) = \frac{P(X=7|C)P(C)}{P(X=7|C)P(C) + P(X=7|\tilde{C})P(\tilde{C})}.$$

Our prior belief that the government was compliant, as given in the problem, is $P(C) = .8$, implying that $P(\tilde{C}) = .2$. In part (a) we calculated that $P(X = 7|C) = .01$ and $P(X = 7|\tilde{C}) = .215$. We simply plug these numbers into the above equation:

$$P(C|X=7) = \frac{P(X=7|C)P(C)}{P(X=7|C)P(C) + P(X=7|\tilde{C})P(\tilde{C})} = \frac{.01 \times .8}{.01 \times .8 + .215 \times .2} = \frac{.008}{.051} = .157.$$

So our belief that the government was compliant is reduced from .8 to .157.

11. Let T be the event in which Obama uses the word "terrorism." We are trying to update our beliefs on Obama's ideology. The probability that Obama is far left given that he spoke about terrorism is

$$P(FL|T) = \frac{P(T|FL)P(FL)}{P(T)}.$$

We don't, however, know $P(T)$. What we do know are the probabilities of T given each of the six ideologies. We can calculate these from the table:

- $P(T|FL) = \frac{2}{10000} = 0.0002$
- $P(T|EL) = \frac{5}{10000} = 0.0005$
- $P(T|CL) = \frac{70}{10000} = 0.007$
- $P(T|CR) = \frac{50}{10000} = 0.005$
- $P(T|LR) = \frac{40}{10000} = 0.004$
- $P(T|FR) = \frac{120}{10000} = 0.0120$

Because these categories are a partition, version 3 of Bayes' rule tells us that we can replace $P(T)$ in the denominator with the sum of these conditional times the prior probabilities of each category:

$$P(T) = P(T|FL)P(FL) + P(T|EL)P(EL) + P(T|CL)P(CL)$$

$$+P(T|CR)P(CR) + P(T|LR)P(LR) + P(T|FR)P(FR)$$

Plugging in the probabilities, we get

$$P(T) = (0.0002 \times 0.2) + (0.0005 \times 0.25) + (0.007 \times 0.35)$$

$$+(0.005 \times 0.1) + (0.004 \times 0.05) + (0.0120 \times 0.05) = 0.003915$$

Now it is fairly straightforward to update our belief that Obama is far left:

$$P(FL|T) = \frac{P(T|FL)P(FL)}{P(T)} = \frac{0.0002 \times 0.2}{0.003915} = 0.01.$$

For economic left

$$P(EL|T) = \frac{P(T|EL)P(EL)}{P(T)} = \frac{0.0005 \times 0.25}{0.003915} = 0.03.$$

For center left,

$$P(CL|T) = \frac{P(T|CL)P(CL)}{P(T)} = \frac{0.007 \times 0.35}{0.003915} = 0.63.$$

For center right,

$$P(CR|T) = \frac{P(T|CR)P(CR)}{P(T)} = \frac{0.005 \times 0.12}{0.003915} = 0.13.$$

For libertarian right,

$$P(LR|T) = \frac{P(T|LR)P(LR)}{P(T)} = \frac{0.004 \times 0.05}{0.003915} = 0.05.$$

And for far right,

$$P(FR|T) = \frac{P(T|FR)P(FR)}{P(T)} = \frac{0.0120 \times 0.05}{0.003915} = 0.15.$$

4 Limits and Derivatives

1. If you draw out a few regular polygons with more sides than 6, you will see that these shapes more and more closely resemble a circle as the number of sides increases. Therefore the answer to this question is: a circle.

   A regular polygon with even a huge number of sides is technically *not* a circle, although it looks like one. A regular polygon only becomes a true circle with *infinitely many* sides, a quantity that can only be understood with a limit. For similar reasons, it's impossible to display a true circle on a television or computer screen, just a close approximation of one, because even on high-resolution screens a close-up examination of the circle reveals that the image is composed of tiny squares.

   For more discussion about the connection between circles and the concept of infinity, and how this connection was used by Archimedes to derive the first close approximation of π, see the article "Take It to the Limit" by Steven Strogatz which appeared in the *New York Times* on April 4, 2010.

2. (a) $\lim\limits_{x\to 5} 2x^2 - 5x + 7$

   For this limit, nothing prevents us from simply plugging 5 into the function. Therefore the limit is

   $$2(5)^2 - 5(5) + 7 = 50 - 25 + 7 = 32.$$

   (b) $\lim\limits_{y\to\infty} \frac{1}{y^6}$

   If we plug in larger and larger values of y, so that these values approach infinity, the value of the function gets closer and closer to 0 without ever exactly reaching it. Therefore the value of the limit is 0.

   In general,

   $$\lim_{x\to\infty} \frac{1}{x^p} = 0$$

   as long as $p > 0$. We will use this property to solve some of the limits below.

   (c) $\lim\limits_{z\to 0} \frac{1}{z^6}$

   We cannot simply plug 0 into the function because that would mean we divide by zero. If we plug in values that approach 0 from the right (such as z=.1, .01, .00001) we see that the values of the function get larger and larger without bound, so the limit from the right is ∞. If we plug in values that approach

0 from the left (such as z=-.1, -.01, -.00001) we also see that the function approaches ∞ (which happens because the even exponent on z causes negative values to become positive). Because the function approaches the same limit from the left and the right, this limit exists, and is equal to ∞.

(d) $\lim_{x\to\infty} \frac{2x+3}{5x^2}$

We could plug in larger and larger values of x and observe what happens to the function, but there's a more direct approach to solving this limit. Since there is addition in the numerator, we can break this function up into two fractions,

$$\lim_{x\to\infty} \frac{2x}{5x^2} + \frac{3}{5x^2}.$$

We can cancel a factor of x from the top and bottom of the first fraction,

$$\lim_{x\to\infty} \frac{2}{5x} + \frac{3}{5x^2},$$

break up the limit over addition,

$$\lim_{x\to\infty} \frac{2}{5x} + \lim_{x\to\infty} \frac{3}{5x^2},$$

and bring the constant factors outside each limit,

$$\frac{2}{5}\lim_{x\to\infty} \frac{1}{x} + \frac{3}{5}\lim_{x\to\infty} \frac{1}{x^2}.$$

Both limits are equal to 0 according to the property we derived in part (b) (we can also see that quickly by plugging in large values of x into each function). So the entire limit is

$$\frac{2}{5}(0) + \frac{3}{5}(0) = 0.$$

(e) $\lim_{y\to\infty} \frac{3y^7+4y^6-2y^5-8y^3-7y+1}{2y^7+y^3-8}$

Here too, we can plug in larger and larger values of y and observe what happens to the function, but there's a more direct approach to solving this limit. We can multiply the top and bottom of the fraction by the same thing, so let's multiply both sides by $1/y^7$:

$$\begin{aligned} &\lim_{y\to\infty} \frac{3y^7+4y^6-2y^5-8y^3-7y+1}{2y^7+y^3-8} \times \frac{\frac{1}{y^7}}{\frac{1}{y^7}} \\ &= \lim_{y\to\infty} \frac{3+\frac{4}{y}-\frac{2}{y^2}-\frac{8}{y^4}-\frac{7}{y^6}+\frac{1}{y^7}}{2+\frac{1}{y^4}-\frac{8}{y^7}}. \end{aligned}$$

According to the property we derived in part (b) the limit of each fraction approaches 0 as $y \to \infty$. Substituting 0 in for each of these fractions leaves us with a total limit of $\frac{3}{2}$.

(f) $\lim_{z\to 3} \frac{z^2-5z+6}{z-3}$

We cannot simply plug in 3 as that would make the denominator equal to 0. But observe that the numerator can be factored. Two numbers that add to -5 and multiply to 6 are -2 and -3, so the numerator factors to

$$\lim_{z\to 3}\frac{(z-2)(z-3)}{z-3}.$$

Now a factor of $z-3$ factors from the top and bottom, leaving us with a simpler limit in which we can plug in 3:

$$\lim_{z\to 3} z-2=1.$$

(g) $\displaystyle\lim_{x\to 5+}\frac{1}{x-5}$

We cannot simply plug in 5, but this problem asks us to only consider the limit from the right. We can do that by plugging in values of the function that are closer and closer to 5, all of which are *greater than* 5. I plug several values, such as $x = 5.1, 5.01, 5.00001$, into the function and find that the function at these values is respectively 10, 100, 100000. The limit from the right is equal to ∞.

(h) $\displaystyle\lim_{y\to 7}\frac{12}{y-7}$

We cannot plug in 7, so let's consider the limits from the right and left. From the right, we plug in values such as $y = 7.1, 7.01, and 7.00001$ and observe that the function gets larger and larger, approaching ∞. From the left, we plug in values such as $y = 6.9, 6.99, and 6.99999$ and observe that the function gets smaller and smaller, approaching $-\infty$. Since the function approaches different limits from the left and the right, this limit does not exist.

(i) $\displaystyle\lim_{z\to\infty}\left(1+\frac{1}{z}\right)^{2z}$

In order to solve this limit exactly, we employ a trick. Remember that the definition of the number e is

$$e=\lim_{z\to\infty}\left(1+\frac{1}{z}\right)^{z}\approx 2.718282\ldots$$

We can rewrite the function to reveal the definition of e inside the limit. Remember the following property of exponents:

$$(x^a)^b = x^{ab}.$$

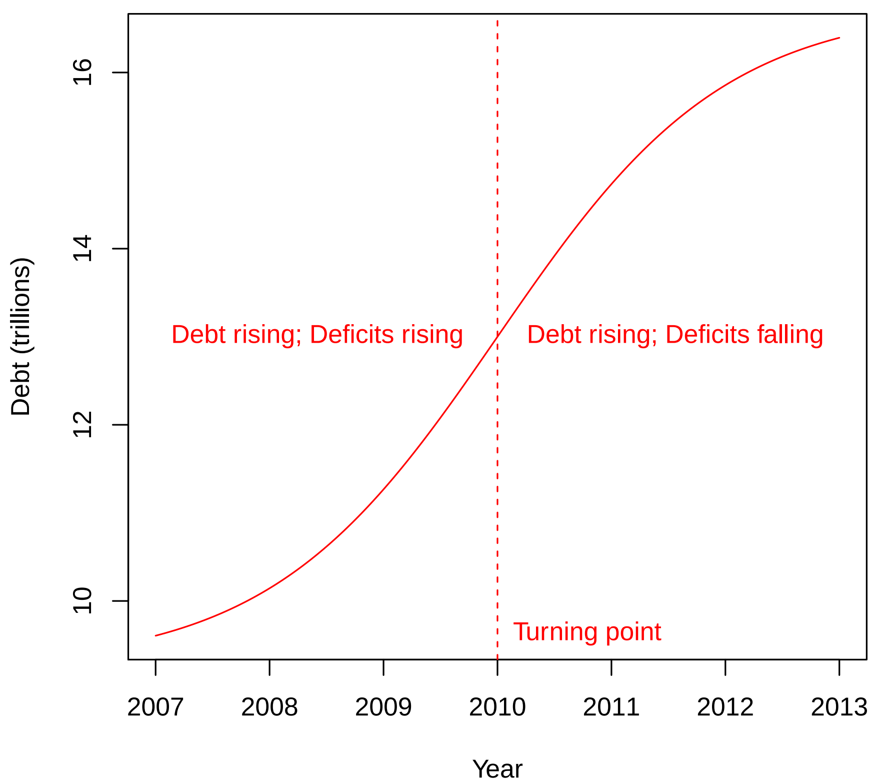
In other words, when an exponent contains two factors, we can rewrite that expression with two levels of exponents given by the two factors. In this case, we can rewrite

$$\begin{aligned}&\lim_{z\to\infty}\left(1+\frac{1}{z}\right)^{2z}\\ &=\lim_{z\to\infty}\left(\left(1+\frac{1}{z}\right)^{z}\right)^{2}\\ &=\lim_{z\to\infty}e^2\approx 7.389056\ldots\end{aligned}$$

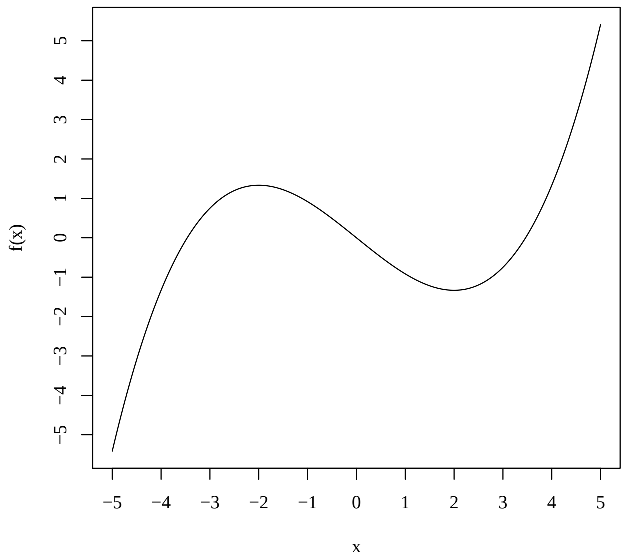
3. (a) If t is the year, and $f(t)$ is the size of the national debt at year t, then $f'(t)$ is the *change in the debt from one year to the next*, which is just another way to phrase the definition of a deficit. To say that the deficit is changing is to refer to the derivative of the function for the deficit, which is itself the derivative of $f(t)$. So, "our deficits are falling at the fastest rate in 60 years" means that

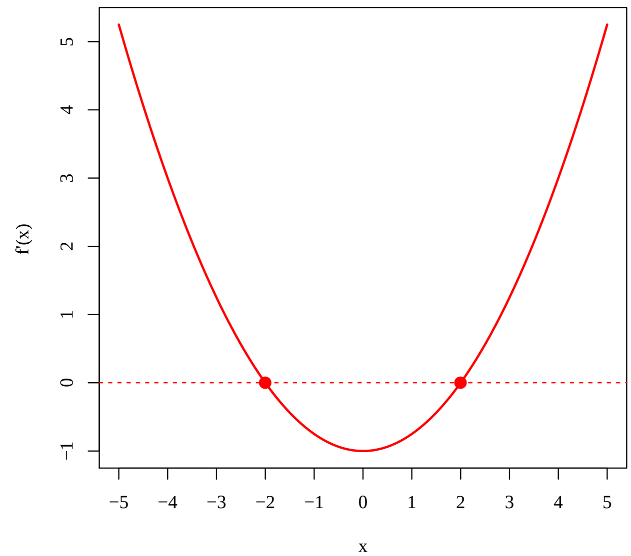
   - $f''(t)$ is currently equal to a negative number (indicating that the deficit is *falling*), and
   - that number is less than (more negative than) any number it has been for the last 60 years.

   (b) Here is one example. Note that the graph is positive everywhere (we always have debt), and it's slope is positive everywhere (we always have deficits instead of surpluses), but the deficits are growing before 2010, and shrinking after 2010.



4. First, notice that the slope of the graph is zero at -2 (the top of the hill), and at 2 (the bottom of the valley). So we can start by plotting the points (-2,0) and (2,0) on the graph of the derivative. Next, observe that the slope is negative between $x = -2$ and $x = 2$, and positive everywhere else. Finally, note that the farther away the graph gets from $x = -2$ and $x = 2$, the the further the slope gets from 0. We can use these observations to construct the graph on the right below:





5. (a) $f(x) = x^6 + 5x^5 - 2x^2 + 8$

We begin by breaking the derivative up over addition and subtraction and bringing constant factors outside each derivative,

$$\begin{aligned} f'(x) &= \frac{d}{dx}\left(x^6 + 5x^5 - 2x^2 + 8\right) \\ &= \frac{d}{dx}(x^6) + \frac{d}{dx}(5x^5) - \frac{d}{dx}(2x^2) + \frac{d}{dx}(8), \\ &= \frac{d}{dx}(x^6) + 5\frac{d}{dx}(x^5) - 2\frac{d}{dx}(x^2) + \frac{d}{dx}(8). \end{aligned}$$

We apply the power rule of differentiation to the first three differentiation terms, and we replace the derivative of the constant in the fourth term with 0. The total derivative is

$$f'(x) = 6x^5 + 25x^4 - 4x.$$

(b) $g(y) = 3e^y - \sqrt{y}$

We begin by breaking the derivative up over addition and subtraction and bringing constant factors outside each derivative,

$$\begin{aligned} g'(y) &= \frac{d}{dy}\left(3e^y - \sqrt{y}\right) \\ &= \frac{d}{dy}(3e^y) - \frac{d}{dy}(\sqrt{y}) \\ &= 3\frac{d}{dy}(e^y) - \frac{d}{dy}(\sqrt{y}). \end{aligned}$$

According to the rules of differentiation in section 4.8, the derivative of the exponential function e^y is simply e^y again, and the derivative of $\sqrt{y}$ is $\frac{1}{2\sqrt{y}}$. Substituting these derivatives into the overall function gives us

$$g'(y) = 3e^y - \frac{1}{2\sqrt{y}}.$$

(c) $h(z) = \ln(z) + \dfrac{1}{z} + 3^z$

We break the derivative up over addition and subtraction,

$$h'(z) = \frac{d}{dz}\left(\ln(z) + \frac{1}{z} + 3^z\right)$$

$$= \frac{d}{dz}(\ln(z)) + \frac{d}{dz}(\frac{1}{z}) + \frac{d}{dz}(3^z),$$

and apply the rules of differentiation in section 4.8. The derivative of a natural logarithm is $\frac{1}{z}$, the derivative of the inverse function $\frac{1}{z}$ is $\frac{-1}{z^2}$, and the derivative of the exponential function 3^z is $\ln(3)3^z$. Substituting these derivatives into the overall function gives us

$$h'(z) = \frac{1}{z} - \frac{1}{z^2} + \ln(3)3^z.$$

(d) $j(x) = (x+3)^7(3x^4 - 2x^2 - 8)$

We could multiply this polynomial out, but that would be a lengthy process since there is an exponent of 7, and we would have to FOIL 7 times. Instead, let's apply the product rule. First, let $g(x) = (x+3)^7$ and let $h(x) = 3x^4 - 2x^2 - 8$. Then $j(x) = g(x)h(x)$, and the product rule for derivatives tells us that

$$j'(x) = g'(x)h(x) + h'(x)g(x).$$

The derivative of the first function technically requires the chain rule:

$$g(x) = A^7, \qquad A = x + 3$$

$$\frac{dg}{dx} = \frac{dg}{dA}\frac{dA}{dx}$$

$$= 7A^6 \times 1 = 7(x+3)^6.$$

The derivative of the second function is

$$\frac{dh}{dx} = \frac{d}{dx}\left(3x^4 - 2x^2 - 8\right)$$

$$= 3\frac{d}{dx}(x^4) - 2\frac{d}{dx}(x^2) - \frac{d}{dx}(8),$$

since derivatives break up across addition and subtraction, and since constant factors can be brought outside a derivative. Then this derivative is

$$12x^3 - 4x.$$

Plugging both functions and both derivatives into the product rule, we get

$$j'(x) = \left(7(x+3)^6\right)\left(3x^4 - 2x^2 - 8\right) + \left((x+3)^7\right)\left(12x^3 - 4x\right).$$

We could simplify this derivative, but this is a correct answer, so let's stop here.

(e) $k(y) = e^{\sqrt{y}}$

This function has layers, so it requires the chain rule. Let's rewrite the function as

$$k(y) = e^A, \qquad A = \sqrt{y}.$$

The derivative of the outer layer is just

$$\frac{d}{dA}\left(e^A\right) = e^A = e^{\sqrt{y}},$$

and the derivative of the inner layer is

$$\frac{d}{dy}\left(\sqrt{y}\right) = \frac{1}{2\sqrt{y}}.$$

Multiplying the derivatives of both layers together, we get

$$k'(y) = \frac{e^{\sqrt{y}}}{2\sqrt{y}}.$$

(f) $l(z) = \dfrac{\ln(z)}{z}$

This function has a quotient, so let's set $f(z) = \ln(z)$ (high) and $g(z) = z$ (low), so that the quotient rule ("low d-high minus high d-low over low-squared") is

$$l'(z) = \frac{d}{dz}\left(\frac{f(z)}{g(z)}\right) = \frac{g(z)f'(z) - f(z)g'(z)}{g(z)^2}.$$

The derivative of $f(z)$ is

$$f'(z) = \frac{d}{dz}\left(\ln(z)\right) = \frac{1}{z}.$$

The derivative of $g(z)$ is

$$g'(z) = \frac{d}{dz}\left(z\right) = 1.$$

Plugging both functions and their derivatives into the quotient rule, we get

$$l'(z) = \frac{z(\frac{1}{z}) - \ln(z)}{z^2} = \frac{1 - \ln(z)}{z^2}.$$

(g) $m(x) = \dfrac{1}{1+e^{-x}}$

This function has layers, so let's apply the chain rule. First we rewrite the function as

$$m(x) = \frac{1}{A}, \quad A = 1 + e^B, \quad B = -x.$$

The derivative of the outermost layer is

$$\frac{d}{dA}\left(\frac{1}{A}\right) = \frac{-1}{A^2} = \frac{-1}{(1+e^B)^2} = \frac{-1}{(1+e^{-x})^2}.$$

The derivative of the middle layer is

$$\frac{d}{dB}\left(1 + e^B\right) = e^B = e^{-x}.$$

And the derivative of the innermost layer is

$$\frac{d}{dx}(-x) = -1.$$

Multiplying the layers together, we get

$$m'(x) = \frac{-1}{(1+e^{-x})^2} \times e^{-x} \times -1 = \frac{e^{-x}}{(1+e^{-x})^2}.$$

(h) $n(y) = \sqrt{y}e^{\sqrt{y}}$

This derivative requires the product rule. Let $f(y) = \sqrt{y}$ and let $g(y) = e^{\sqrt{y}}$, so that the derivative is given by

$$n'(y) = f'(y)g(y) + g'(y)f(y).$$

First, the derivative of $f(y)$ is

$$\frac{d}{dy}(\sqrt{y}) = \frac{1}{2\sqrt{y}}.$$

Next, the derivative of $g(y)$ requires the chain rule. We rewrite $g(y)$ as

$$g(y) = e^A, \qquad A = \sqrt{P},$$

we take the derivative of the outer layer,

$$\frac{d}{dA}\left(e^A\right) = e^A = e^{\sqrt{y}},$$

we take the derivative of the inner layer,

$$\frac{d}{dy}(\sqrt{y}) = \frac{1}{2\sqrt{y}},$$

and we multiply the layers together

$$g'(y) = \frac{e^{\sqrt{y}}}{2\sqrt{y}}.$$

Now that we have $f'(y)$ and $g'(y)$, we plug these derivatives into the product rule:

$$\begin{aligned} n'(y) &= \left(\frac{1}{2\sqrt{y}}\right)\left(e^{\sqrt{y}}\right) + \left(\frac{e^{\sqrt{y}}}{2\sqrt{y}}\right)\left(\sqrt{y}\right) \\ &= (\sqrt{y} + 1)\frac{e^{\sqrt{y}}}{2\sqrt{y}}. \end{aligned}$$

(i) $p(z) = \dfrac{e^{z^2+4}}{\ln(z)}$

We apply the quotient rule, with $f(z) = e^{z^2+4}$ and $g(z) = \ln(z)$ so that

$$p'(z) = \frac{d}{dz}\left(\frac{f(z)}{g(z)}\right) = \frac{g(z)f'(z) - f(z)g'(z)}{g(z)^2}.$$

The derivative of $f(z)$ requires the chain rule, so we rewrite the function as

$$f(z) = e^A, \qquad A = z^2 + 4.$$

The derivative of the outer layer is

$$\frac{d}{dA}\left(e^A\right) = e^A = e^{z^2+4},$$

and the derivative of the inner layer is

$$\frac{d}{dz}(z^2 + 4) = 2z,$$

so the entire derivative of $f(z)$ is

$$f'(z) = 2ze^{z^2+4}.$$

The derivative of $g(z)$ is

$$\frac{d}{dz}\left(\ln(z)\right) = \frac{1}{z}.$$

Substituting these functions and their derivatives into the quotient rule, we get

$$p'(z) = \frac{\ln(z)2re^{z^2+4} - \frac{e^{z^2+4}}{z}}{\ln(z)^2}.$$

(j) $q(x) = \ln(x^3 + 2x)$

We can rewrite the function as

$$q(x) = \ln(A), \qquad A = x^3 + 2x.$$

The chain rule tells us that

$$\frac{df}{dx} = \frac{df}{dA}\frac{dA}{dx}.$$

$$\frac{df}{dA} = \frac{d}{dA}\left(\ln(A)\right) = \frac{1}{A} = \frac{1}{x^3 + 2x},$$

$$\frac{dA}{dx} = \frac{d}{dx}(x^3 + 2x) = 3x^2 + 2.$$

$$\frac{df}{dx} = \frac{df}{dA}\frac{dA}{dx} = \frac{1}{x^3 + 2x} \times (3x^2 + 2)$$

$$q'(x) = \frac{3x^2 + 2}{x^3 + 2x}.$$

(k) $r(y) = e^{1/(y^2+2y-2)}$

We can rewrite the function with the following layers:

$$r(x) = e^A, \quad A = \tfrac{1}{B}, \quad B = y^2 + 2y - 2.$$

$$\frac{dm}{dA} = \frac{d}{dA}\left(e^A\right) = e^A = e^{1/(y^2+2y-2)}.$$

$$\frac{dA}{dB} = \frac{d}{dB}\left(\frac{1}{B}\right) = \frac{-1}{B^2} = \frac{-1}{(y^2+2y-2)^2}.$$

$$\frac{dB}{dy} = \frac{d}{dy}\left(y^2 + 2y - 2\right) = 2y - 2.$$

$$\frac{dm}{dy} = \frac{dm}{dA}\frac{dA}{dB}\frac{dB}{dy} = e^{1/(y^2+2y-2)} \times \frac{-1}{(y^2+2y-2)^2} \times (2y-2).$$

$$r'(y) = \frac{-(2y-2)e^{1/(y^2+2y-2)}}{(y^2+2y-2)^2}.$$

No need to simplify any further.

(l) $s(z) = \ln(z^3 + 2z)e^{(1/z^2+2z-2)}$

Notice that this is the product of the functions from parts (a) and (b), so we can write

$$s(z) = f(z)g(z).$$

By the product rule of derivatives, we know that

$$s'(z) = f(z)g'(z) + f'(z)g(z).$$

We've already found each of these terms, so we can simply plug them in:

$$s'(z) = \left(\ln(z^3+2z)\right)\left(\frac{-(2z-2)e^{1/(z^2+2z-2)}}{(z^2+2z-2)^2}\right) + \left(\frac{3z^2+2}{z^3+2z}\right)\left(e^{1/(z^2+2z-2)}\right).$$

(m) $t(x) = \dfrac{\sqrt{x^2+3}}{x}$

This function is the quotient of two functions, so we apply the quotient rule of differentiation. In this case,

$$t(x) = \frac{f(x)}{g(x)},$$

where

$$f(x) = \sqrt{x^2+3}, \quad \text{and} \quad g(x) = x.$$

The derivative of $f(x)$ requires the chain rule. Set $f(x) = \sqrt{A}$ and $A = x^2 + 3$. Then

$$f'(x) = \frac{1}{2\sqrt{A}} \times 2x = \frac{2x}{2\sqrt{x^2+3}} = \frac{x}{\sqrt{x^2+3}}.$$

The derivative of $g(x) = x$ is just $g'(x) = 1$. Plugging these terms into the formula for the quotient rule:

$$t'(x) = \frac{\frac{x^2}{\sqrt{x^2+3}} - \sqrt{x^2+3}}{x^2}.$$

(n) $v(y) = \sqrt{\dfrac{(y^4 - 3y^2)\ln(7y-4)}{e^{y^3-2y}}}$

This is an advanced chain rule problem, but the steps are the same as always. First we write the function in terms of its layers:

$$v(y) = \sqrt{A}, \qquad A = \frac{BC}{D},$$

$$B = y^4 - 3y^2, \qquad C = \ln(E),$$

$$D = e^F, \qquad E = 7y - 4,$$

$$F = y^3 - 2y.$$

Next we take the derivative of each layer. The derivative of the outermost layer is

$$v'(y) = \frac{1}{2\sqrt{A}}.$$

The trickiest part here is the derivative of A, which involves *both* the product rule and the quotient rule! Let's consider this derivative first. Note that A is a quotient of functions, so we can first apply the quotient rule:

$$A' = \frac{D(BC)' - (BC)D'}{D^2}.$$

This expression contains the derivative of the product of two functions $(BC)'$, for which we apply the product rule:

$$A' = \frac{D(B'C + C'B) - (BC)D'}{D^2}.$$

Since this step involved the quotient and product rules, we **stop here** and multiply the derivatives together:

$$v'(y) = \frac{1}{2\sqrt{A}} \frac{D(B'C + C'B) - (BC)D'}{D^2}.$$

The problem is now to find B', C' and D' and substitute them into this function. The derivative of B is straightforward:

$$B' = 4y^3 - 6y.$$

The derivative of C itself requires that we use the chain rule:

$$C' = \frac{1}{E}E' = \frac{7}{E}.$$

The derivative of D also requires the chain rule:

$$D' = e^F F' = (3y^2 - 2)e^F.$$

Substituting B', C', and D' into the overall derivative gives us

$$v'(y) = \frac{1}{2\sqrt{A}} \frac{D\Big((4y^3-6y)C + (\frac{7}{E})B\Big) - (BC)(3y^2-2)e^F}{D^2}.$$

Now we substitute back in for the capital letters, starting with A,

$$v'(y) = \frac{1}{2\sqrt{\frac{BC}{D}}} \frac{D\Big((4y^3-6y)C + (\frac{7}{E})B\Big) - (BC)(3y^2-2)e^F}{D^2},$$

then B,

$$v'(y) = \frac{1}{2\sqrt{\frac{(y^4-3y^2)C}{D}}} \frac{D\Big((4y^3-6y)C + (\frac{7}{E})(y^4-3y^2)\Big) - (y^4-3y^2)C(3y^2-2)e^F}{D^2},$$

then C,

$$v'(y) = \frac{1}{2\sqrt{\frac{(y^4-3y^2)\ln(E)}{D}}} \frac{D\Big((4y^3-6y)\ln(E) + (\frac{7}{E})(y^4-3y^2)\Big) - (y^4-3y^2)\ln(E)(3y^2-2)e^F}{D^2},$$

then D,

$$v'(y) = \frac{1}{2\sqrt{\frac{(y^4-3y^2)\ln(E)}{e^F}}} \frac{e^F\Big((4y^3-6y)\ln(E) + (\frac{7}{E})(y^4-3y^2)\Big) - (y^4-3y^2)\ln(E)(3y^2-2)e^F}{e^{2F}},$$

then E,

$$v'(y) = \frac{1}{2\sqrt{\frac{(y^4-3y^2)\ln(7y-4)}{e^F}}} \frac{e^F\Big((4y^3-6y)\ln(7y-4) + (\frac{7}{7y-4})(y^4-3y^2)\Big) - (y^4-3y^2)\ln(7y-4)(3y^2-2)e^F}{e^{2F}},$$

and finally F,

$$v'(y) = \frac{1}{2\sqrt{\frac{(y^4-3y^2)\ln(7y-4)}{e^{y^3-2y}}}} \frac{e^{y^3-2y}\Big((4y^3-6y)\ln(7y-4) + (\frac{7}{7y-4})(y^4-3y^2)\Big) - (y^4-3y^2)\ln(7y-4)(3y^2-2)e^{y^3-2y}}{e^{2y^3-4y}}.$$

This is about as difficult a derivative as anyone will have to take by hand. But we've not shied away from the hard problems and we won't start now. The answer isn't pretty, but if you follow and understand the steps we need to take this derivative, then you have a strong grasp of differentiation.

6. (a) The first derivative requires the product rule:

$$f'(x) = x\frac{d}{dx}\Big(\ln(x)\Big) + \ln(x)\frac{d}{dx}\Big(x\Big) = x\frac{1}{x} + \ln(x) = 1 + \ln(x).$$

The second derivative is

$$f''(x) = \frac{d}{dx}\Big(1 + \ln(x)\Big) = \frac{1}{x}.$$

The third derivative is

$$f'''(x) = \frac{d}{dx}\left(\frac{1}{x}\right) = \frac{-1}{x^2}.$$

The 4th derivative is

$$f^{(4)}(x) = \frac{d}{dx}\left(\frac{-1}{x^2}\right) = \frac{d}{dx}\left(-x^{-2}\right) = 2x^{-3} = \frac{2}{x^3}.$$

The 5th derivative is

$$f^{(5)}(x) = \frac{d}{dx}\left(\frac{2}{x^3}\right) = \frac{d}{dx}\left(2x^{-3}\right) = -6x^{-4} = \frac{-6}{x^4}.$$

The 6th derivative is

$$f^{(6)}(x) = \frac{d}{dx}\left(\frac{-6}{x^4}\right) = \frac{d}{dx}\left(-6x^{-4}\right) = 24x^{-5} = \frac{24}{x^5}.$$

Are you beginning to see the pattern? For the nth derivative, there's a fraction. The numerator is $(n-2)!$ (where the ! mark denotes a factorial, the product of all positive whole numbers from 1 to $n-2$ – for a definition of a factorial see section 3.3.2) and the denominator is x^{n-1}. But also, the fraction is multiplied by -1 for odd derivatives, and 1 by even derivatives, which we can represent as $(-1)^n$. So overall, a formula for the nth derivative is

$$f^{(n)}(x) = (-1)^n \frac{(n-2)!}{x^{n-1}}.$$

(b) The first derivative requires the product rule:

$$g'(x) = x\frac{d}{dx}\left(e^x\right) + e^x\frac{d}{dx}\left(x\right) = xe^x + e^x = (x+1)e^x.$$

The second derivative is

$$g''(x) = (x+1)\frac{d}{dx}\left(e^x\right) + e^x\frac{d}{dx}\left(x+1\right) = (x+1)e^x + e^x = (x+2)e^x.$$

The third derivative is

$$g'''(x) = (x+2)\frac{d}{dx}\left(e^x\right) + e^x\frac{d}{dx}\left(x+2\right) = (x+2)e^x + e^x = (x+3)e^x.$$

So it appears that a formula for the nth derivative is

$$g^{(n)}(x) = (x+n)e^x.$$

7. (a) At this point, you are able to quickly see that the derivative is $f'(x) = 3x^2$. But why is this true? The limit definitions relate to the basic concept that a derivative represents: the instantaneous rate of change of a function. Using the first limit definition, the derivative is

$$\begin{aligned} f'(a) &= \lim_{x\to a} \frac{f(x)-f(a)}{x-a} \\ &= \lim_{x\to a} \frac{x^3-a^3}{x-a}. \end{aligned}$$

The numerator factors using the difference of cubes ("SOAP") formula from section 1.7.2:

$$f'(a) = \lim_{x\to a} \frac{(x-a)(x^2+ax+a^2)}{x-a},$$
$$= \lim_{x\to a} x^2+ax+a^2.$$

Since we've removed $x-a$ from the denominator, it's now safe to plug in a for x:

$$f'(a) = a^2 + a(a) + a^2 = 3a^2.$$

Using the second limit definition, the derivative is

$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$
$$= \lim_{h\to 0} \frac{(x+h)^3 - x^3}{h}$$
$$= \lim_{h\to 0} \frac{x^3+3x^2h+3xh^2+h^3-x^3}{h}$$
$$= \lim_{h\to 0} \frac{3x^2h+3xh^2+h^3}{h}$$
$$= \lim_{h\to 0} \frac{h(3x^2+3xh+h^2)}{h}$$
$$= \lim_{h\to 0} 3x^2+3xh+h^2.$$

Since we've removed h from the denominator, it's now safe to plug in 0 for h:

$$f'(x) = 3x^2 + 3x(0) + (0)^2 = 3x^2.$$

(b) The derivative of $g(y) = y^2+2y+8$ is $g'(y) = 2y+2$. To prove that this function really is the derivative of $g(y)$, we can use the first limit definition of the derivative:

$$g'(a) = \lim_{y\to a} \frac{f(y)-f(a)}{y-a}$$
$$= \lim_{y\to a} \frac{y^2+2y+8-(a^2+2a+8)}{y-a}$$
$$= \lim_{y\to a} \frac{y^2-a^2+2y-2a+8-8}{y-a}$$
$$= \lim_{y\to a} \frac{(y^2-a^2)+(2y-2a)}{y-a}$$
$$= \lim_{y\to a} \frac{(y-a)(y+a)+2(y-a)}{y-a}$$
$$= \lim_{y\to a}(y+a)+2 = (a+a)+2 = 2a+2.$$

Using the second limit definition:

$$g'(y) = \lim_{h\to 0} \frac{g(y+h)-g(y)}{h}$$
$$= \lim_{h\to 0} \frac{(y+h)^2+2(y+h)+8-(y^2+2y+8)}{h}$$

$$
\begin{aligned}
&= \lim_{h\to 0} \frac{y^2 + 2yh + h^2 + 2y + 2h + 8 - y^2 - 2y - 8}{h} \\
&= \lim_{h\to 0} \frac{y^2 - y^2 + 2yh + h^2 + 2y - 2y + 2h + 8 - 8}{h} \\
&= \lim_{h\to 0} \frac{2yh + h^2 + 2h}{h} \\
&= \lim_{h\to 0} \frac{h(2y + h + 2)}{h} \\
&= \lim_{h\to 0} 2y + h + 2 = 2y + 2.
\end{aligned}
$$

8. (a) The first derivative of the normal distribution requires the chain rule. I rewrite the function as

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{A}, \qquad A = -\frac{x^2}{2}.$$

Then the derivative is

$$\frac{df}{dx} = \frac{df}{dA}\frac{dA}{dx} = \left(\frac{1}{\sqrt{2\pi}} e^{A}\right)\left(-x\right) = \frac{-x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

The second derivative requires the product rule. If we let $g(x) = -x$, then the first derivative can be written as

$$f'(x) = \left(-x\right)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right) = g(x)f(x).$$

Note that the second factor in this particular problem is the original function $f(x)$. So the second derivative is

$$
\begin{aligned}
f''(x) &= g'(x)f(x) + g(x)f'(x) \\
&= (-1)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right) + (-x)\left(\frac{-x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right) \\
&= \left(\frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right) - \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right). \\
&= (x^2 - 1)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right).
\end{aligned}
$$

If we plug in 0 into the original function, we get

$$f(0) = \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{0^2}{2}}\right) = \frac{1}{\sqrt{2\pi}}.$$

If we plug 0 into the first derivative, we get

$$f'(0) = \left(\frac{-0}{\sqrt{2\pi}} e^{-\frac{0^2}{2}}\right) = 0.$$

If we plug 0 into the second derivative, we get

$$f''(0) = (0^2 - 1)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{0^2}{2}}\right) = \frac{-1}{\sqrt{2\pi}}.$$

So the second-order Taylor approximation is the polynomial:

$$
\begin{aligned}
f(x) &\approx \frac{-1}{\sqrt{2\pi}}\frac{x^2}{2} + 0x + \frac{1}{\sqrt{2\pi}} \\
&= \frac{-1}{\sqrt{8\pi}} x^2 + \frac{1}{\sqrt{2\pi}}.
\end{aligned}
$$

(b) I graphed the normal distribution and its Taylor approximation using Stata, but any program that is capable of graphing can do the trick. First, type

```
gen x = 6*(_n/_N) - 3
```

The statement `_n` refers to each particular observation number, and the statement `_N` refers to the total number of observations. Type "browse" in the command window to see that this command produces an equally spaced sequence of numbers between -3 and 3. This variable will be used as the independent variable for both the normal distribution and the Taylor approximation. Next, get the y values of the normal distribution by typing

```
gen norm = normalden(x)
```

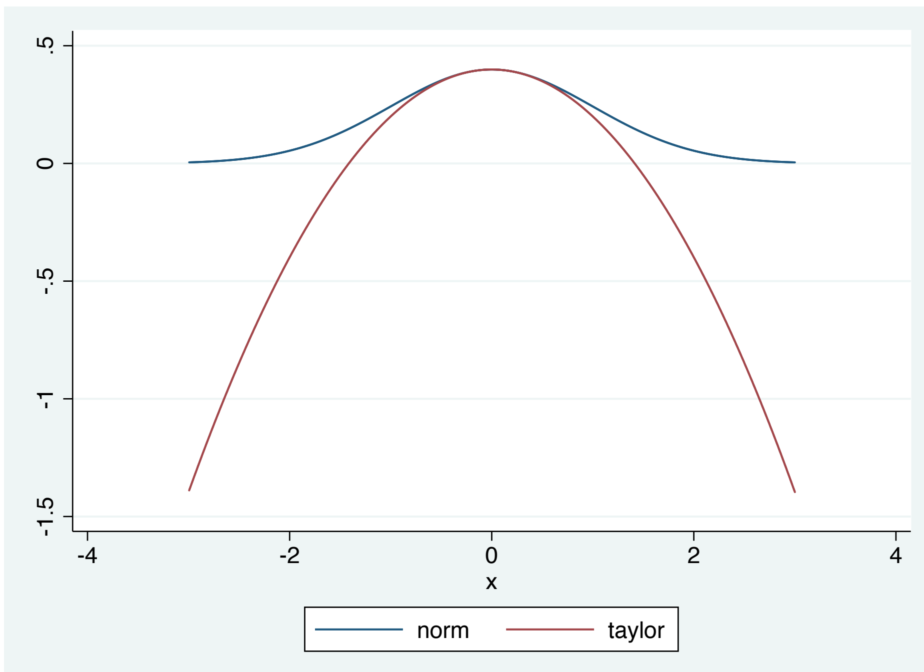
and get the y values of the Taylor approximation by typing

```
gen taylor = (-1/sqrt(8*c(pi)))*x^2 + (1/sqrt(2*c(pi)))
```

where I ignore `x` since it's coefficient is 0. Finally, to produce the graph, type

```
twoway(line norm x)(line taylor x)
```

Here is the graph I obtain:



We approximate the normal distribution well around the centering point 0, and continue to approximate well out to about 1 and -1. The quality of the approximation declines rapidly for points further than 1 away from 0. If we estimated more derivatives though, we could obtain a higher-order Taylor polynomial which would conform to the normal distribution for points farther out.

9. (a) The log-likelihood is the natural logarithm of the likelihood function, which in this case is

$$\ell(p|y_1,\ldots,y_N) = \ln\left(\prod_{i=1}^{N} p^{y_i}(1-p)^{1-y_i}\right).$$

To simplify, we first remember that $\ln(ab) = \ln(a) + \ln(b)$, that a logarithm turns multiplication inside into addition outside. The same rule applies to long-products: a logarithm turns a long-product inside into a summation outside. So the log-likelihood simplifies to

$$\ell(p|y_1,\dots,y_N) = \sum_{i=1}^{N} \ln\left(p^{y_i}(1-p)^{1-y_i}\right).$$

Next the logarithm breaks up the multiplication inside:

$$\ell(p|y_1,\dots,y_N) = \sum_{i=1}^{N} \left[\ln\left(p^{y_i}\right) + \ln\left((1-p)^{1-y_i}\right)\right].$$

Now the exponents inside each logarithm can be brought down as factors:

$$\ell(p|y_1,\dots,y_N) = \sum_{i=1}^{N} \left[y_i \ln(p) + (1-y_i)\ln(1-p)\right].$$

The summation can be broken up over addition,

$$\ell(p|y_1,\dots,y_N) = \sum_{i=1}^{N} y_i \ln(p) + \sum_{i=1}^{N}(1-y_i)\ln(1-p),$$

and any factor that does not contain the index i can be brought outside the summation. In this case, such factors are the logged probabilities:

$$\ell(p|y_1,\dots,y_N) = \ln(p)\sum_{i=1}^{N} y_i + \ln(1-p)\sum_{i=1}^{N}(1-y_i).$$

Finally, the last summation can be broken up over subtraction,

$$\ell(p|y_1,\dots,y_N) = \ln(p)\sum_{i=1}^{N} y_i + \ln(1-p)\left(\sum_{i=1}^{N}(1) - \sum_{i=1}^{N} y_i\right),$$

and the sum of 1s, repeated N times, is just N:

$$\ell(p|y_1,\dots,y_N) = \ln(p)\sum_{i=1}^{N} y_i + \ln(1-p)\left(N - \sum_{i=1}^{N} y_i\right).$$

(b) We are taking the derivative of the log-likelihood function, treating p as the independent variable and the y terms as constants. The problem is

$$\frac{d\ell}{dp} = \frac{d}{dp}\left[\ln(p)\sum_{i=1}^{N} y_i + \ln(1-p)\left(N - \sum_{i=1}^{N} y_i\right)\right].$$

The first step is to remember that a derivative breaks up over addition. Since a summation is just repeated addition, the derivative of a summation is the summation of the derivative of each term:

$$\frac{d\ell}{dp} = \frac{d}{dp}\left[\ln(p)\sum_{i=1}^{N} y_i\right] + \frac{d}{dp}\left[\ln(1-p)\left(N - \sum_{i=1}^{N} y_i\right)\right].$$

Here we treat p as the independent variable and the y_i terms as constants. We next bring the constant factors out of each derivative:

$$\frac{d\ell}{dp} = \left(\sum_{i=1}^{N} y_i\right)\frac{d}{dp}\left(\ln(p)\right) + \left(N - \sum_{i=1}^{N} y_i\right)\frac{d}{dp}\left(\ln(1-p)\right).$$

We've reduced the problem to taking the derivative of two logarithms. The second one requires the chain rule:

$$\frac{d}{dp}\ln(p) = \frac{1}{p},$$

$$\frac{d}{dp}\ln(1-p) = \frac{1}{1-p}(1-p)' = \frac{-1}{1-p}.$$

Substituting these terms into the derivative we get

$$\begin{aligned}\frac{d\ell}{dp} &= \left(\sum_{i=1}^{N} y_i\right)\frac{1}{p} + \left(N - \sum_{i=1}^{N} y_i\right)\frac{-1}{1-p} \\ &= \frac{\sum_{i=1}^{N} y_i}{p} - \frac{N - \sum_{i=1}^{N} y_i}{1-p}.\end{aligned}$$

(c) The problem is to solve

$$\frac{\sum_{i=1}^{N} y_i}{p} - \frac{N - \sum_{i=1}^{N} y_i}{1-p} = 0$$

for p. First we can bring one term over to the other side,

$$\frac{\sum_{i=1}^{N} y_i}{p} = \frac{N - \sum_{i=1}^{N} y_i}{1-p},$$

cross-multiply,

$$(1-p)\sum_{i=1}^{N} y_i = p\left(N - \sum_{i=1}^{N} y_i\right),$$

and distribute,

$$\sum_{i=1}^{N} y_i - p\sum_{i=1}^{N} y_i = pN - p\sum_{i=1}^{N} y_i.$$

We cancel $p\sum y_i$ from both sides,

$$\sum_{i=1}^{N} y{=}pN,$$

and solve for p by dividing by N:

$$p = \frac{\sum_{i=1}^{N} y_i}{N}.$$

Substantively, this answer is the average of all the observed 0s and 1s. It makes perfect sense. If we survey 1000 people, and 600 say they voted for the incumbent, then there is a $\frac{600}{1000} = 0.6$ probability of voting for the incumbent.

(d) All we have to do is plug -3, 0 and 3 into

$$p_i = \frac{1}{1 + e^{-(0.2+0.5x_i)}}.$$

These calculations are

$$p_i(-3) = \frac{1}{1 + e^{-(0.2+0.5(-3))}} = 0.21,$$

$$p_i(0) = \frac{1}{1 + e^{-(0.2+0.5(0))}} = 0.55,$$

$$p_i(3) = \frac{1}{1 + e^{-(0.2+0.5(3))}} = 0.85.$$

So a very liberal voter has a 0.21 probability of voting for the incumbent, a moderate voter has a 0.55 probability of voting for the incumbent (indicating that the incumbent has won the moderate voters), and a very conservative voter has a 0.85 probability of voting for the incumbent.

(e) The problem is to take the derivative of

$$p_i = \frac{1}{1 + e^{-(0.2+0.5x_i)}}.$$

This derivative requires the chain rule. We break the function into layers as follows:

$$p_i = \frac{1}{A}, \quad A = 1 + e^B, \quad B = -(0.2 + 0.5x_i).$$

The derivatives of the layers are

$$p_i' = \frac{-1}{A^2}, \quad A' = e^B, \quad B' = -0.5.$$

No derivative involves the quotient or product rule. So the next step is to multiply the layers' derivatives together:

$$\frac{dp_i}{dx_i} = \frac{0.5e^B}{A^2}.$$

We substitute for A,

$$\frac{dp_i}{dx_i} = \frac{0.5e^B}{(1 + e^B)^2},$$

and we substitute for B:

$$\frac{dp_i}{dx_i} = \frac{0.5e^{-(0.2+0.5x_i)}}{(1 + e^{-(0.2+0.5x_i)})^2}.$$

This formula expresses the marginal change in the probability of voting for the incumbent for any ideology x. So for a moderate voter, the change in probability is

$$\frac{dp_i}{dx_i}(0) = \frac{0.5e^{-(0.2+0.5(0))}}{(1 + e^{-(0.2+0.5(0)i)})^2} = \frac{0.5e^{-0.2}}{(1 + e^{-0.2})^2} = 0.124.$$

(f) Remember that a derivative tells us the *instantaneous slope* of a function at a particular point. So the instantaneous slope of the probability of voting for the incumbent is 0.124 for moderate voters. But what does this really mean? Slope is the change in y for a change in x. We consider the change in x to be 1, as in change in miles per ONE hour. So, when x is 0, a one-unit increase in x is associated with an increase in y of about 0.124.

That's not exactly the same thing as saying that moving from moderate ($x = 0$) to to slightly conservative ($x = 1$) is associated with a 0.124 increase in the probability of voting for the incumbent. It's possible that this instantaneous slope is slightly different at $x = 1$ than it is as $x = 0$, which would alter this difference. But exactly at $x = 0$, the rate of change of the dependent variable for a unit increase in x is 0.124.

5 Optimization

1. (a) The notation $x \in [-4, 4]$ means that x is bounded, and that both boundary points are included in the domain. That means we have to compare the value of the function at the critical points to the value of the function at the boundary points to find the global maximum and global minimum. First we find the critical points by taking the derivative, setting it equal to 0, and solving for x. The derivative is

$$f'(x) = 12x^3 - 12x^2 - 72x.$$

If we set the derivative equal to 0, we can solve for x through factoring:

$$12x^3 - 12x^2 - 72x = 0,$$

$$12(x^3 - x^2 - 6x) = 0,$$
$$12x(x^2 - x - 6) = 0,$$
$$12x(x^2 - x - 6) = 0,$$
$$12x(x - 3)(x + 2) = 0.$$

So the critical points are $x = 0$, $x = 3$, and $x = -2$. Next, we can check whether each critical point c describes a local maximum, a local minimum, or a saddle point using the second derivative test. The second derivative of the function is

$$f''(x) = 36x^2 - 24x - 72.$$

c	$f''(c)$	Result
0	$36(0)^2 - 24(0) - 72 = -72$	Local max
3	$36(3)^2 - 24(3) - 72 = 180$	Local min
-2	$36(-2)^2 - 24(-2) - 72 = 120$	Local min

Finally, we compare the local maximum to the boundary points to find the global maximum, and we compare the local minimums to the boundary points to find the global minimum. At the boundary points, the function is

$$f(-4) = 3(-4)^4 - 4(-4)^3 - 36(-4)^2 = 448, \qquad f(4) = 3(4)^4 - 4(4)^3 - 36(4)^2 = -64,$$

for the local maximum the function is

$$f(0) = 3(0)^4 - 4(0)^3 - 36(0)^2 = 0,$$

and at the local minimums the function is

$$f(3) = 3(3)^4 - 4(3)^3 - 36(3)^2 = -189, \qquad f(-2) = 3(-2)^4 - 4(-2)^3 - 36(-2)^2 = -64.$$

So the global maximum occurs at the boundary point $x = -4$, and the global minimum occurs at $x = 3$.

(b) Since the domain is bounded and since the upper bound 3 is included in the domain (the parenthesis around 0 means that it is not included in the domain) we have to compare the value of the function at $x = 3$ to the value of the function at the critical points. To find the critical points, we take the derivative. First note that the derivative can be broken up across subtraction:

$$g'(x) = \frac{d}{dx}\Big(x \ln(x)\Big) - \frac{d}{dx}(x)$$

$$= \frac{d}{dx}\Big(x\ln(x)\Big) - 1.$$

The product rule applies

$$\begin{aligned} g'(x) &= x\frac{d}{dx}\Big(\ln(x)\Big) + \ln(x)\frac{d}{dx}(x) - 1, \\ &= x\frac{1}{x} + \ln(x) - 1, \\ &= 1 + \ln(x) - 1 \\ &= \ln(x). \end{aligned}$$

Next we set the derivative equal to 0. We know that any logarithm of 1 is equal to 0, so the critical point is $x = 1$. To test whether this point is a local max or a local min, we find the second derivative,

$$g''(x) = \frac{d}{dx}\Big(\ln(x)\Big) = \frac{1}{x},$$

and plug the critical point in:

$$g'(1) = \frac{1}{1} = 1.$$

Since the second derivative is positive at the critical point, $x = 1$ describes a local minimum. Finally we compare the value of the function at the local min to the boundary point:

$$f(1) = (1)\ln(1) - 1 = 1(0) - 1 = -1,$$

$$f(3) = (3)\ln(3) - 1 = 2.3.$$

So $x = 1$ is the location of the global minimum, and $x = 3$ is the location of the global maximum.

2. (a) First we find the critical points by taking the derivative and setting it equal to 0:

$$f'(x) = 3x^2 - 15x + 12 = 0$$

We can factor out a 3:

$$3(x^2 - 5x + 4) = 0.$$

Two numbers that add to -5 and multiply to 4 are -1 and -4, so the derivative factors to

$$3(x - 1)(x - 4) = 0.$$

That implies that the critical points are $x = 1$ and $x = 4$, both of which exist in the stated domain. Next we check whether each one is a local minimum, local maximum, or a saddle point. The second derivative is a pretty simple function in this case, so we use the second derivative test by finding the second derivative, plugging in the critical points, and observing whether the second derivative is positive or negative:

$$f''(x) = 6x - 15.$$

The second derivative at $x = 1$ is

$$f''(1) = 6(1) - 15 = -9,$$

so $x = 1$ represents a local maximum. The second derivative at $x = 5$ is

$$f''(4) = 6(4) - 15 = 9,$$

so $x = 4$ is a local minimum. Finally, we compare the value of the function at the critical points to the value of the function at the boundary points:

$$f(0) = (0)^3 - \frac{15}{2}(0)^2 + 12(0) + 8 = 8,$$

$$f(1) = (1)^3 - \frac{15}{2}(1)^2 + 12(1) + 8 = 13.5,$$
$$f(4) = (4)^3 - \frac{15}{2}(4)^2 + 12(4) + 8 = 0,$$
$$f(6) = (6)^3 - \frac{15}{2}(6)^2 + 12(6) + 8 = 26.$$

So $x = 5$ is the global minimum, but $x = 6$ is the global maximum.

(b) The Newton-Raphson algorithm, beginning at 2, yields the following iterations:

Iteration	x	$f'(x)$	$f''(x)$	$x - \frac{f'(x)}{f''(x)}$
0	2	-6	-3	0
1	0	12	-15	0.8
2	0.8	1.92	-10.2	0.988235294
3	0.988235294	0.106297578	-9.070588235	0.999954223
4	0.999954223	0.000412	-9.000274662	0.999999999
5	0.999999999	6.28643E-09	-9.000000004	1
6	1	0	-9	1

The Newton-Raphson algorithm, beginning at 5, yields the following iterations:

Iteration	x	$f'(x)$	$f''(x)$	$x - \frac{f'(x)}{f''(x)}$
0	5	12	15	4.2
1	4.2	1.92	10.2	4.011764706
2	4.011764706	0.106297578	9.070588235	4.000045777
3	4.000045777	0.000412	9.000274662	4.000000001
4	4.000000001	6.28643E-09	9.000000004	4
5	4	0	9	4

(c) i. The NR-algorithm might stop at a minimum or a maximum, and does not tell you which one it arrives at.

ii. The NR-algorithm might stop at a local extreme point that is not a global extreme point. If the global max or min is on the boundary, then the NR-algorithm will never return that point.

3. To find the global maximum of $U(t)$, we first have to find the critical points. Technically, t is bounded below by 0, so we have to compare the value of the function at the critical points to the value at 0. To find the critical points, we first take the derivative:

$$U'(t) = \frac{d}{dt}\left(30\ln(t+1) - \frac{t^2}{10}\right)$$
$$= 30\frac{d}{dt}\left(\ln(t+1)\right) - \frac{\frac{d}{dt}(t^2)}{10}$$

$$= \frac{30}{t+1} - \frac{t}{5}.$$

Next we set the derivative equal to 0 and we solve for t:

$$\frac{30}{t+1} - \frac{t}{5} = 0,$$

$$\frac{30}{t+1} = \frac{t}{5},$$

$$t(t+1) = 150,$$

$$t^2 + t - 150 = 0.$$

This quadratic equation does not factor neatly, but we can use the quadratic formula to calculate the solutions. The quadratic formula says that for an equation of the form $Ax^2 + Bx + C = 0$, x is

$$x = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}.$$

Let $A = 1$, $B = 1$, and $C = -150$. Then t is

$$t = \frac{-1 \pm \sqrt{1^2 - 4(1)(-150)}}{2(1)}$$

$$= \frac{-1 \pm \sqrt{1 + 600}}{2}$$

$$= -\frac{1}{2} - \frac{\sqrt{601}}{2} \quad \text{and} \quad -\frac{1}{2} + \frac{\sqrt{601}}{2}$$

$$= -12.76 \quad \text{and} \quad 11.76.$$

t is days, so we throw the solution of $t = -12.76$ out because it doesn't make sense. Our critical point is $t = 11.76$ days. To demonstrate that this point describes a local maximum, we use the second derivative test. the second derivative of $U(t)$ is

$$U''(t) = \frac{d}{dt}\left(\frac{30}{t+1} - \frac{t}{5}\right)$$

$$U''(t) = \frac{-30}{(t+1)^2} - \frac{1}{5}.$$

Plugging in $t = 11.76$, we get

$$U''(11.76) = \frac{-30}{(11.76+1)^2} - \frac{1}{5} = -.38.$$

So the second derivative test tells us that $t = 11.76$ is a local maximum. There are no additional critical points to consider, but t is bounded below by $t = 0$. The function at $t = 0$ is

$$U(0) = 30\ln(0+1) - \frac{0^2}{10}$$

$$= 30\ln(1) - 0 = 0.$$

The value of the function at $t = 11.76$ is

$$U(11.76) = 30\ln(11.76+1) - \frac{11.76^2}{10} = 62.6.$$

So $t = 11.76$ is also the global maximum. Therefore, the House Republicans maximize their utility of the shutdown if it lasts 11.76 days.

4. The sum of squared errors is the following function of β:

$$f(\beta) = \sum_{i=1}^{N}(y_i^2 - 2\beta x_i y_i + \beta^2 x_i^2).$$

In order to minimize this function, we first have to take the derivative with respect to β:

$$f'(\beta) = \frac{d}{d\beta}\left(\sum_{i=1}^{N}(y_i^2 - 2\beta x_i y_i + \beta^2 x_i^2)\right).$$

Notice that a summation $\sum$ is the same thing as a sum. Since derivatives break up over addition, we can rewrite the derivative of the sum as the sum of the derivatives of the addends:

$$f'(\beta) = \sum_{i=1}^{N}\frac{d}{d\beta}\left(y_i^2 - 2\beta x_i y_i + \beta^2 x_i^2\right).$$

We treat y_i and x_i as constants, so the derivative is

$$f'(\beta) = \sum_{i=1}^{N}(-2x_i y_i + 2\beta x_i^2).$$

To find the critical point, we set the derivative equal to 0 and solve for β:

$$\sum_{i=1}^{N}(-2x_i y_i + 2\beta x_i^2) = 0.$$

We can rewrite this by taking the sum of each part:

$$\sum_{i=1}^{N}(-2x_i y_i) + \sum_{i=1}^{N}(2\beta x_i^2) = 0.$$

Factors without a subscript i can be brought outside the summations:

$$-2\sum_{i=1}^{N}x_i y_i + 2\beta\sum_{i=1}^{N}x_i^2 = 0.$$

And now we can simply solve for β:

$$\begin{aligned} 2\beta\sum_{i=1}^{N}x_i^2 &= 2\sum_{i=1}^{N}x_i y_i, \\ \beta\sum_{i=1}^{N}x_i^2 &= \sum_{i=1}^{N}x_i y_i, \\ \beta &= \frac{\sum_{i=1}^{N}x_i y_i}{\sum_{i=1}^{N}x_i^2}. \end{aligned}$$

Finally, to demonstrate that this critical point represents a local minimum, we find the second derivative of the sum of squares:

$$\begin{aligned} f''(\beta) &= \frac{d}{d\beta}\left(\sum_{i=1}^{N}(-2x_i y_i + 2\beta x_i^2)\right) \\ &= \sum_{i=1}^{N}\frac{d}{d\beta}\left(-2x_i y_i + 2\beta x_i^2\right) \\ &= \sum_{i=1}^{N}(2x_i^2). \end{aligned}$$

The second derivative does not depend on β, and the since the x_i datapoints are squared, the sum $\sum_{i=1}^{N}(2x_i^2)$ must be positive. Therefore the critical point describes a local minimum.

5. (a) The trick here is to remember that we are taking the derivative with respect to μ. Notice that the first term in the log-likelihood function does not contain μ, so it is a constant and it drops out of the derivative. The derivative is

$$\ell'(\mu) = \frac{d}{d\mu}\Big(-.5\sum_{i=1}^{n}(x_i-\mu)^2\Big).$$

The constant factor -.5 comes in front of the derivative:

$$\ell'(\mu) = -.5\frac{d}{d\mu}\Big(\sum_{i=1}^{n}(x_i-\mu)^2\Big),$$

and the sum comes outside the derivative too since derivatives break up across addition:

$$\ell'(\mu) = -.5\sum_{i=1}^{n}\frac{d}{d\mu}(x_i-\mu)^2.$$

Here x_i is a constant and μ is the variable, so the derivative becomes

$$\ell'(\mu) = -.5\sum_{i=1}^{n}-2(x_i-\mu),$$

$$\ell'(\mu) = \sum_{i=1}^{n}(x_i-\mu).$$

To simplify this function, the summation can be applied to each term in the parentheses:

$$\ell'(\mu) = \sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\mu,$$

and since μ does not have an index, it is added n times. Therefore it can be rewritten as

$$\ell'(\mu) = \sum_{i=1}^{n}x_i - n\mu.$$

(b) We've done almost all the work in part (a). We set the derivative equal to 0,

$$\ell'(\mu) = \sum_{i=1}^{n}x_i - n\mu = 0,$$

and solve for μ:

$$\sum_{i=1}^{n}x_i = n\mu,$$

$$\mu = \frac{\sum_{i=1}^{n}x_i}{n}.$$

(c) The first derivative again is

$$\ell'(\mu) = \sum_{i=1}^{n}x_i - n\mu.$$

So the second derivative is just

$$\ell''(\mu) = -n.$$

That value is negative everywhere, so the value at the critical point we derived in part (b) is negative, and therefore the critical value refers to a local maximum. Since the domain of the normal distribution is unbounded, this value is also the global maximum.

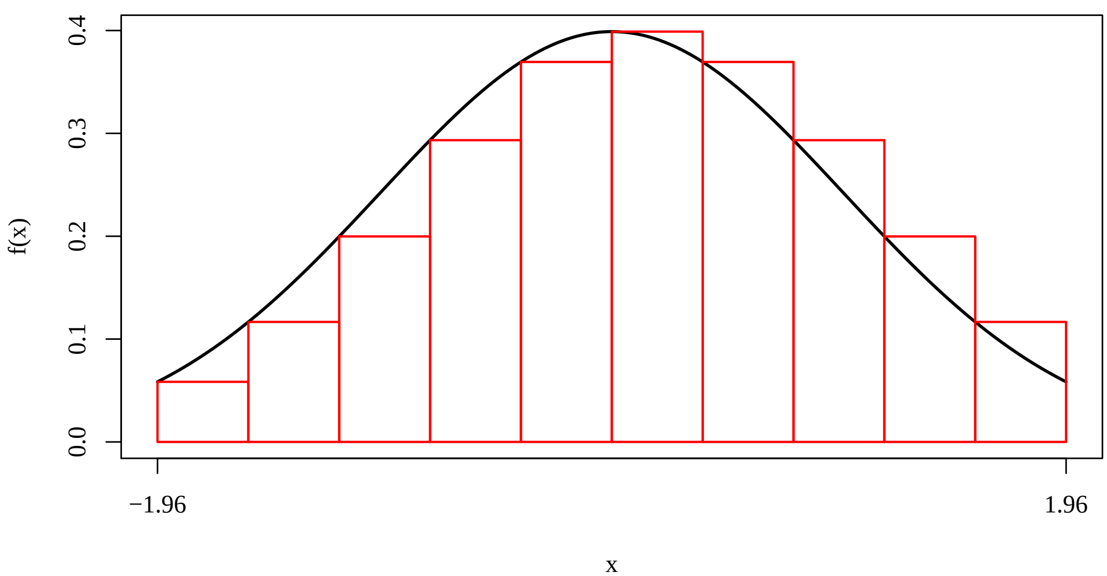
(d) The critical point

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

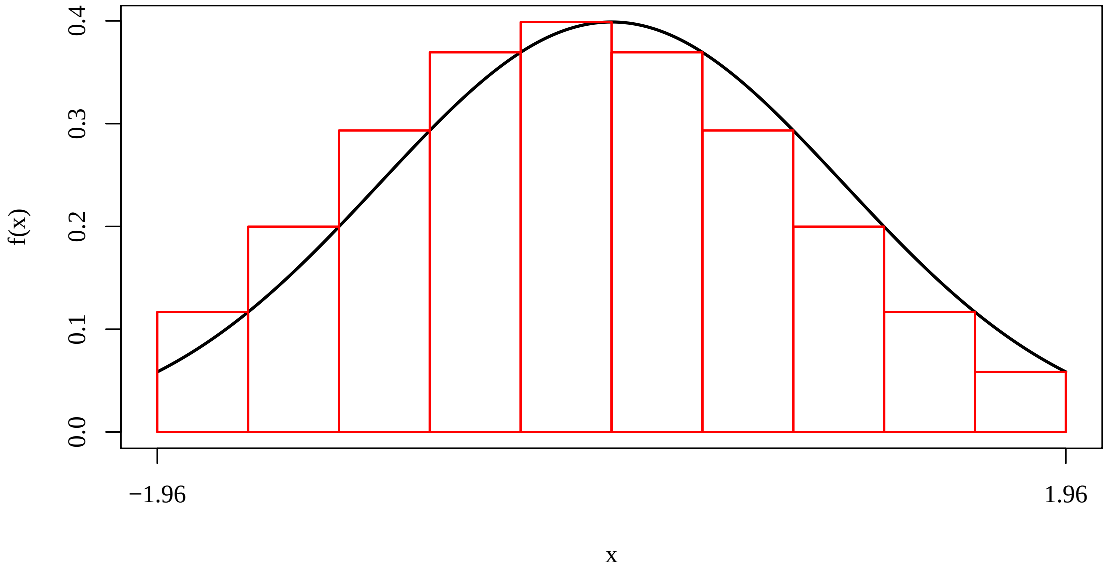
is actually the mean of the sample of x values since it is the sum of the n values of x in the sample, divided by n. So we are estimating the mean of the normal distribution with the mean of the sample. This estimate makes perfect sense.

6 Integration

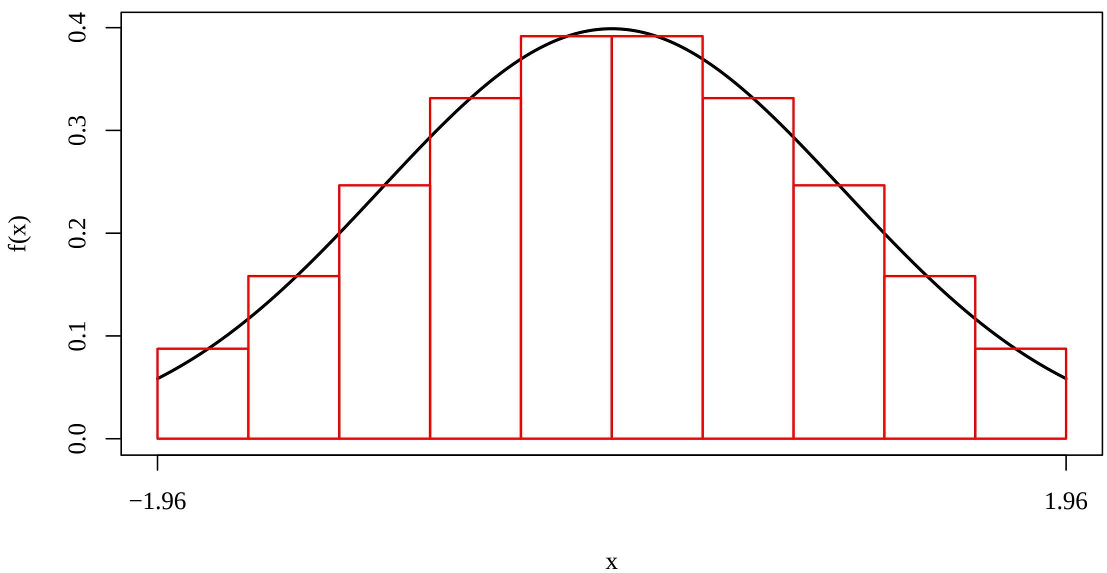
1. (a) The Riemann sums are graphed below. For the left Riemann sum, the top LEFT corner of each rectangle is the corner that touches the graph.



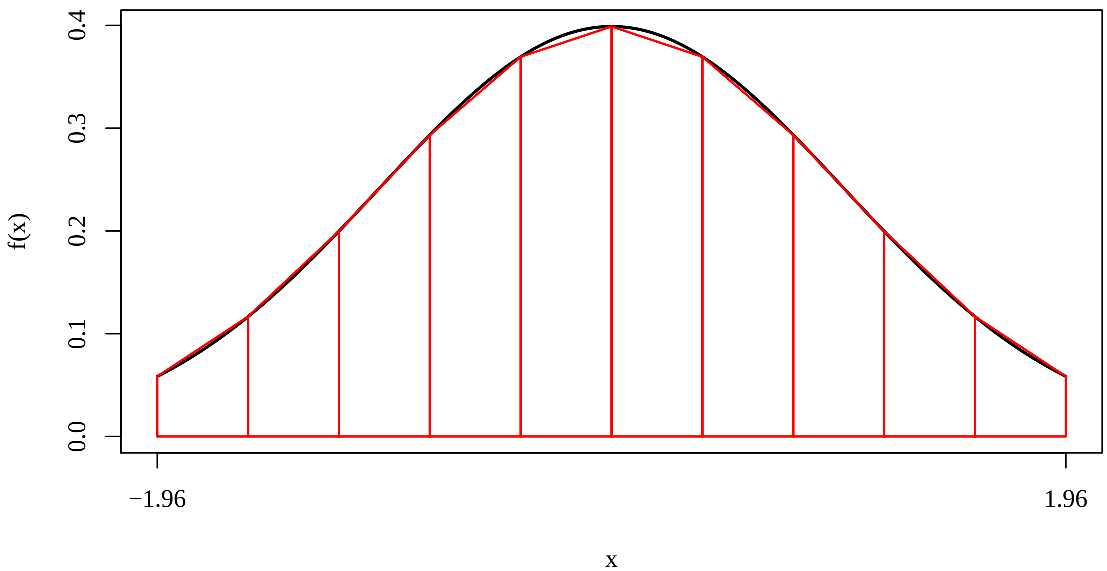
For the right Riemann sum, the top RIGHT corner of each rectangle is the corner that touches the graph.



For the midpoint Riemann sum, midpoint of the top of each rectangle touches the graph.



For the trapezoidal Riemann sum, both the top LEFT and top RIGHT corners of each trapezoid that touch the graph, and we draw a straight line between these two points.



(b) For the left Riemann sum we use the formula

$$A \approx \sum_{i=0}^{N-1} f\left(a + \frac{b-a}{N}i\right)\frac{b-a}{N},$$

and plug in $N = 10$ partitions, from $a = -1.96$ to $b = 1.96$, where f is

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-.5x^2}.$$

We get

$$A \approx \sum_{i=0}^{10-1} f\left(-1.96 + \frac{1.96-(-1.96)}{10}i\right)\frac{1.96-(-1.96)}{10}$$

$$= \sum_{i=0}^{9} f(-1.96 + .392i).392$$

$$= .392\sum_{i=0}^{9} f(-1.96 + .392i)$$

$$= .392\sum_{i=0}^{9} \frac{1}{\sqrt{2\pi}}e^{-.5(-1.96+.392i)^2}$$

$$= \frac{.392}{\sqrt{2\pi}}\sum_{i=0}^{9} e^{-.5(-1.96+.392i)^2}.$$

For the right Riemann sum we use the formula

$$A \approx \sum_{i=1}^{N} f\left(a + \frac{b-a}{N} i\right) \frac{b-a}{N},$$

and plug in $N = 10$ partitions, from $a = -1.96$ to $b = 1.96$, where f is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-.5x^2}.$$

We get

$$\begin{aligned}
A &\approx \sum_{i=1}^{10} f\left(-1.96 + \frac{1.96-(-1.96)}{10} i\right) \frac{1.96-(-1.96)}{10} \\
&= \sum_{i=1}^{10} f(-1.96 + .392i).392 \\
&= .392 \sum_{i=1}^{10} f(-1.96 + .392i) \\
&= .392 \sum_{i=1}^{10} \frac{1}{\sqrt{2\pi}} e^{-.5(-1.96+.392i)^2} \\
&= \frac{.392}{\sqrt{2\pi}} \sum_{i=1}^{10} e^{-.5(-1.96+.392i)^2}.
\end{aligned}$$

For the midpoint Riemann sum we use the formula

$$A \approx \sum_{i=1}^{N} f\left(a + \frac{b-a}{N}(i - .5)\right) \frac{b-a}{N},$$

and plug in $N = 10$ partitions, from $a = -1.96$ to $b = 1.96$, where f is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-.5x^2}.$$

We get

$$\begin{aligned}
A &\approx \sum_{i=1}^{10} f\left(-1.96 + \frac{1.96-(-1.96)}{10}(i - .5)\right) \frac{1.96-(-1.96)}{10} \\
&= \sum_{i=1}^{10} f(-1.96 + .392(i - .5)).392 \\
&= .392 \sum_{i=1}^{10} f(-1.96 + .392(i - .5)) \\
&= .392 \sum_{i=1}^{10} f(-2.16 + .392i) \\
&= .392 \sum_{i=1}^{10} \frac{1}{\sqrt{2\pi}} e^{-.5(-2.16+.392i)^2} \\
&= \frac{.392}{\sqrt{2\pi}} \sum_{i=1}^{10} e^{-.5(-2.16+.392i)^2}.
\end{aligned}$$

And for the trapezoidal Riemann sum we use the formula

$$A \approx \sum_{i=0}^{N-1} \frac{\left[f\left(a + \frac{b-a}{N}i\right) + f\left(a + \frac{b-a}{N}(i+1)\right)\right]}{2} \frac{b-a}{N},$$

and plug in $N = 10$ partitions, from $a = -1.96$ to $b = 1.96$, where f is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-.5x^2}.$$

We get

$$A \approx \sum_{i=0}^{9} \frac{\left[f\left(-1.96 + \frac{1.96-(-1.96)}{10}i\right) + f\left(-1.96 + \frac{1.96-(-1.96)}{10}(i+1)\right)\right]}{2} \frac{1.96-(-1.96)}{10}$$

$$= \sum_{i=0}^{9} \frac{\left[f\left(-1.96 + .392i\right) + f\left(-1.96 + .392(i+1)\right)\right]}{2} .392$$

$$= \frac{.392}{2} \sum_{i=0}^{9} f\left(-1.96 + .392i\right) + f\left(-1.96 + .392(i+1)\right)$$

$$= \frac{.196}{\sqrt{2\pi}} \sum_{i=0}^{9} e^{-.5(-1.96+.392i)^2} + e^{-.5(-1.96+.392(i+1))^2}.$$

(c) We simply write out the sums and evaluate them. For the left Riemann sum:

$$A \approx \frac{.392}{\sqrt{2\pi}} \sum_{i=0}^{9} e^{-.5(-1.96+.392i)^2}$$

$$= \frac{.392}{\sqrt{2\pi}} \Big[e^{-.5(-1.96+.392(0))^2} + e^{-.5(-1.96+.392(1))^2} + e^{-.5(-1.96+.392(2))^2} + e^{-.5(-1.96+.392(3))^2}$$
$$+ e^{-.5(-1.96+.392(4))^2} + e^{-.5(-1.96+.392(5))^2} + e^{-.5(-1.96+.392(6))^2} + e^{-.5(-1.96+.392(7))^2}$$
$$+ e^{-.5(-1.96+.392(8))^2} + e^{-.5(-1.96+.392(9))^2} \Big]$$
$$= 0.947.$$

For the right Riemann sum

$$A \approx \frac{.392}{\sqrt{2\pi}} \sum_{i=1}^{10} e^{-.5(-1.96+.392i)^2}$$

$$= \frac{.392}{\sqrt{2\pi}} \Big[e^{-.5(-1.96+.392(1))^2} + e^{-.5(-1.96+.392(2))^2} + e^{-.5(-1.96+.392(3))^2} + e^{-.5(-1.96+.392(4))^2}$$
$$+ e^{-.5(-1.96+.392(5))^2} + e^{-.5(-1.96+.392(6))^2} + e^{-.5(-1.96+.392(7))^2} + e^{-.5(-1.96+.392(8))^2}$$

$$+e^{-.5(-1.96+.392(9))^2} + e^{-.5(-1.96+.392(10))^2}\Big]$$

$$= 0.947.$$

For the midpoint Riemann sum

$$A \approx \frac{.392}{\sqrt{2\pi}} \sum_{i=1}^{10} e^{-.5(-2.16+.392i)^2}$$

$$= \frac{.392}{\sqrt{2\pi}} \Big[e^{-.5(-2.16+.392(1))^2} + e^{-.5(-2.16+.392(2))^2} + e^{-.5(-2.16+.392(3))^2} + e^{-.5(-2.16+.392(4))^2}$$

$$+e^{-.5(-2.16+.392(5))^2} + e^{-.5(-2.16+.392(6))^2} + e^{-.5(-2.16+.392(7))^2} + e^{-.5(-2.16+.392(8))^2}$$

$$+e^{-.5(-2.16+.392(9))^2} + e^{-.5(-2.16+.392(10))^2}\Big]$$

$$= 0.951.$$

For the trapezoidal Riemann sum

$$A \approx \frac{.196}{\sqrt{2\pi}} \sum_{i=0}^{9} e^{-.5(-1.96+.392i)^2} + e^{-.5(-1.96+.392(i+1))^2}$$

$$= \frac{.196}{\sqrt{2\pi}} \Big[e^{-.5(-1.96+.392(0))^2} + 2e^{-.5(-1.96+.392(1))^2} + 2e^{-.5(-1.96+.392(2))^2} + 2e^{-.5(-1.96+.392(3))^2}$$

$$+2e^{-.5(-1.96+.392(4))^2} + 2e^{-.5(-1.96+.392(5))^2} + 2e^{-.5(-1.96+.392(6))^2} + 2e^{-.5(-1.96+.392(7))^2}$$

$$+2e^{-.5(-1.96+.392(8))^2} + 2e^{-.5(-1.96+.392(9))^2} + e^{-.5(-1.96+.392(10))^2}\Big]$$

Every added term, with the exception of the first and last, has a factor of 2. This factor represents that fact that the term appears twice. For example, the term $e^{-.5(-1.96+.392(5))^2}$ appears once when $i = 4$ and again when $i = 5$. The total of this sum is

$$= 0.947.$$

2. (a) There are no bounds on this integral, so this is an indefinite integral. Our purpose in calculating this integral is to find the anti-derivative of $f(x) = x^{100} + 3e^x - 7(4^x)$, that is, the function $F(x)$ whose derivative is $f(x)$.

First, we break up the integral over addition and subtraction:

$$F(x) = \int x^{100}\, dx + \int 3e^x\, dx - \int 7(4^x)\, dx.$$

Next we bring the constant factors outside the integrals:

$$F(x) = \int x^{100}\, dx + 3\int e^x\, dx - 7\int 4^x\, dx.$$

Finally we apply the rules for integrating power functions and exponential functions of base e and of base 4:

$$F(x) = \frac{x^{101}}{101} + 3e^x - 7\frac{4^x}{\ln(4)} + c,$$

where "+c" represents the constant that drops out of a derivative, and is added to any indefinite integral.

(b) This integral is definite because it has bounds, and our purpose in calculating this integral is to derive the area A between the curve of $f(x) = 5\sqrt{x} + \frac{3}{x^4}$ and the x-axis between $x = 1$ and $x = 9$.

The process for solving a definite integral is nearly the same as the process for solving an indefinite integral. As before, we start by breaking the integral up over addition,

$$A = \int_1^9 5\sqrt{x}\ dx + \int_1^9 \frac{3}{x^4}\ dx,$$

and we bring constant factors outside the integrals:

$$A = 5\int_1^9 \sqrt{x}\ dx + 3\int_1^9 \frac{1}{x^4}\ dx.$$

To simplify each integral, we can rewrite the square root as an exponent of 1/2, and we can rewrite the reciprocated exponent as a negative exponent:

$$A = 5\int_1^9 x^{1/2}\ dx + 3\int_1^9 x^{-4}\ dx.$$

Now we can apply the power rule to each integral and write the bounds to the right of the anti-derivative:

$$\begin{aligned} A &= 5\left(\frac{2}{3}x^{3/2}\right) + 3\left(\frac{-1}{3}x^{-3}\right)\Big|_1^9 \\ &= \frac{10}{3}x^{3/2} - \frac{1}{x^3}\Big|_1^9. \end{aligned}$$

Finally we plug in the upper-bound, plug in the lower-bound, and subtract. When we work with definite integrals, we do not write "+c" in the anti-derivative because it will cancel out when we subtract the anti-derivative at each bound:

$$\begin{aligned} A &= \left(\frac{10}{3}9^{3/2} - \frac{1}{9^3}\right) - \left(\frac{10}{3}1^{3/2} - \frac{1}{1^5}\right) \\ &= \left(\frac{10}{3}27 - \frac{3}{729}\right) - \left(\frac{10}{3} - 1\right) \\ &= \left(90 - \frac{3}{729}\right) - \frac{7}{3} = 87\frac{2}{3}. \end{aligned}$$

(c) This is a definite integral, but because it has an infinite bound, it is also an improper integral. We begin by using a limit to rewrite the integral with bounds that we can work with:

$$\lim_{K\to\infty} \int_2^K \frac{12}{x^2}\ dx.$$

We will solve the definite integral treating K as if it were finite. Then we will solve the limit once we've solved the integral. We can rewrite the integrand with a negative exponent,

$$\lim_{K\to\infty} \int_2^K 12x^{-2}\ dx,$$

bring the 12 outside the integral (and the limit too),

$$12\lim_{K\to\infty} \int_2^K x^{-2}\ dx,$$

and apply the power rule of integration,

$$12 \lim_{K\to\infty} \frac{x^{-1}}{-1}\bigg|_2^K$$

$$= 12 \lim_{K\to\infty} \frac{-1}{x}\bigg|_2^K$$

$$= 12 \lim_{K\to\infty} \frac{-1}{K} + \frac{1}{2}.$$

Finally, note that as K goes to infinity, the fraction $\frac{-1}{K}$ approaches zero. So in the limit we can replace this fraction with 0, leaving us with

$$12\left(\frac{1}{2}\right) = 6.$$

(d) This problem involves the derivative of a definite integral. Since integrals and derivatives are inverse operations, they cancel out – but since the derivative is taken with respect to a variable other than x that exists in the bounds, we apply the first fundamental theorem of calculus for bounds:

$$\frac{d}{dy}\int_a^{g(y)} f(x)\ dx = f\Big(g(y)\Big)g'(y).$$

We don't have to find any anti-derivatives. We simply plug the upper bound y^2 into the integrand, and satisfy the chain rule by multiplying the function by the derivative of y^2. The solution is

$$f\Big(g(y)\Big)g'(y)$$

$$= \sqrt{y^2}(2y) = 2y^2.$$

(e) First we multiply the integral by -1 so that the bounds are reversed:

$$-\frac{d}{dz}\int_{10}^{\sqrt{z}+\ln(z)} e^x\ dx.$$

Now we apply the first fundamental theorem of calculus for bounds:

$$-\frac{d}{dz}\int_a^{g(z)} f(x)\ dx = -f\Big(g(z)\Big)g'(z).$$

We don't have to calculate an anti-derivative, we only need to calculate $f\Big(g(z)\Big)$ and $g'(z)$:

$$-f\Big(g(z)\Big)g'(z)$$

$$= -e^{\sqrt{z}+\ln(z)}\left(\frac{1}{2\sqrt{z}} + \frac{1}{z}\right).$$

3. (a) There are two polynomials in the integrand. The second polynomial, the one taken to the 7th power, is a degree-10 polynomial. When a degree-10 polynomial is differentiated, the result is a degree-9 polynomial. Since the first polynomial in the integrand is degree-9, it makes sense to try to set the second polynomial as u to see if du removes the first polynomial. So,

$$u = 5x^{10} - 25x^4 + 15x,$$

which means that the derivative is

$$\frac{du}{dx} = 50x^9 - 100x^3 + 15,$$
$$\frac{1}{5}du = (10x^9 - 20x^3 + 3)dx,$$

which does in fact contain the first polynomial. Next we substitute u and $\frac{1}{5}du$ into the integral:

$$\int (10x^9 - 20x^3 + 3)(5x^{10} - 25x^4 + 15x)^7 \, dx$$
$$= \int (5x^{10} - 25x^4 + 15x)^7 \left[(10x^9 - 20x^3 + 3) \, dx\right]$$
$$= \int u^7 \left[\frac{1}{5}du\right]$$
$$\frac{1}{5}\int u^7 \, du.$$

This integral can be solved with the power rule of integration:

$$\frac{1}{5}\frac{u^8}{8} + c$$
$$\frac{u^8}{40} + c.$$

Finally, we substitute back for u:

$$\frac{(5x^{10} - 25x^4 + 15x)^8}{40} + c.$$

(b) The trick behind u-substitution is to find part of the integrand whose derivative resembles some factor inside the integrand. In this case, if we set

$$u = x^2 - 2,$$

then

$$\frac{du}{dx} = 2x.$$

Next we multiply both sides by dx, and divide both sides by 2 (to bring all constant factors with du instead of dx):

$$\frac{1}{2}\,du = x\,dx.$$

For a definite integral, we have two options for handling the bounds. First, we can plug each bound into our equation for u, and derive new bounds. If we pursue this option, we won't need to substitute back in for u at the end of the problem. The new lower bound is

$$u(0) = (0)^2 - 2 = -2,$$

and the new upper bound is

$$u(3) = (3)^2 - 2 = 9 - 2 = 7.$$

Now we can substitute u, $\frac{1}{2}du$, and the new bounds into the integral:

$$\int_0^3 xe^{x^2-2}\ dx = \frac{1}{2}\int_{-2}^{7} e^u\ du$$

$$= \frac{1}{2}e^u\Big|_{-2}^{7} = \frac{e^7 - e^{-2}}{2} = 548.25.$$

The second option for dealing with the bounds is to find the anti-derivative of the integral with u, substitute back in for u, and calculate the area under the curve using the *old* bounds. This anti-derivative (ignoring "+c" because this is a definite integral) is

$$e^u.$$

Substituting back in for u and applying the bounds gives us:

$$\frac{1}{2}e^{x^2-2}\Big|_0^3 = \frac{e^{3^2-2} - e^{0^2-2}}{2} = \frac{e^7 - e^{-2}}{2} = 548.25.$$

(c) If we set

$$u = 4x^2 + 6x - 11,$$

then

$$\frac{du}{dx} = 8x + 6, \quad du = 8x + 6\ dx, \quad \frac{1}{2}du = 4x + 3\ dx.$$

So by choosing this particular u, the numerator drops out of the function when we make the substitution. That's exactly the purpose of u-substitution. Substituting for $4x^2 + 6x - 11$ and for $(4x + 3)dx$, the integral now becomes

$$\frac{1}{2}\int \frac{1}{u}\ du,$$

$$= \frac{1}{2}\ln(|u|) + c.$$

Substituting back for u, this becomes

$$\frac{1}{2}\ln(|4x^2 + 6x - 11|) + c.$$

(d) It makes sense to first try setting the expression inside the square root to u,

$$u = x^2 + 1,$$

so that the derivative is

$$\frac{du}{dx} = 2x,$$

$$du = 2x\ dx.$$

The new lower bound is

$$u(0) = (0)^2 + 1 = 1,$$

and the new upper bound is

$$u(2) = (2)^2 + 1 = 5.$$

Substituting u, du, and the new bounds into the integral gives us

$$\int_1^5 \sqrt{u}\ du$$

$$= \int_1^5 u^{\frac{1}{2}}\ du$$

$$= \left.\frac{u^{\frac{3}{2}}}{\frac{3}{2}}\right|_1^5$$

$$= \left.\frac{2}{3}u^{\frac{3}{2}}\right|_1^5$$

$$= \frac{2}{3}\left(5^{\frac{3}{2}} - 1^{\frac{3}{2}}\right)$$

$$= \frac{2}{3}\left(\sqrt{125} - 1\right) = 6.79.$$

4. (a) The integrand is the product of two functions: x and $\sqrt{x+5}$. Of these two functions, x definitely becomes simpler when differentiated, so we denote

$$u = x.$$

We will integrate the other function, so we denote it as

$$dv = \sqrt{x+5}\ dx.$$

Note, it is not necessary to see right away *how* we will integrate the other function at this point. We simply need to choose u and dv and to give integration by parts a chance. It is often the case that one selection of u and dv won't work, so we try a new way to write these functions. Even very experienced mathematicians have to try integration by parts again and again, so don't get frustrated.

The derivative of $u = x$ is

$$\frac{du}{dx} = 1.$$

Solving for du:

$$du = dx.$$

To find v, we take the indefinite integral of the right-hand side of the equation for dv:

$$v = \int \sqrt{x+5}\ dx.$$

To solve this integral, let's apply u-substitution by setting

$$w = x + 5,$$

$$\frac{dw}{dx} = 1, \qquad dw = dx.$$

Substituting w and dw into the integral gives us

$$v = \int \sqrt{w}\ dw$$

$$= \int w^{1/2}\ dw$$
$$= \frac{2}{3}w^{3/2}.$$

Finally we substitute back for w, and we avoid writing "+c" since this problem will eventually involve bounds. Therefore v is

$$v = \frac{2}{3}(x+5)^{3/2}.$$

Plugging these elements as well as the bounds $a = 1$ and $b = 4$ into the above formula gives us

$$\int_1^4 x\sqrt{x+5}\ dx$$
$$= x\left(\frac{2}{3}(x+5)^{3/2}\right)\Bigg|_1^4 - \int_1^4 \frac{2}{3}(x+5)^{3/2}\ dx.$$

We can simplify this equation slightly without solving it:

$$\left(\frac{2}{3}x(x+5)^{3/2}\right)\Bigg|_1^4 - \frac{2}{3}\int_1^4 (x+5)^{3/2}\ dx.$$

We have two steps left to take. First we have to solve the remaining integral, then we have to plug the bounds into the total anti-derivative and calculate the area under the curve. Consider the integral:

$$\int (x+5)^{3/2}\ dx.$$

Since this integral is part of a larger definite integral, we hold off on dealing with the bounds until the very end of the problem. We can solve this integral through u-substitution (using w instead of u to avoid confusion with the u we used earlier in this problem). Set

$$w = (x+5), \quad \text{so that } \tfrac{dw}{dx} = 1, \quad \text{and } dw = dx.$$

Substituting w and dw into the integral gives us:

$$\int w^{3/2}\ dw$$
$$\frac{2}{5}w^{5/2}.$$

Substituting back in for w gives us

$$\frac{2}{5}(x+5)^{5/2},$$

and plugging this anti-derivative into the integration by parts equation gives us

$$\int_1^4 x\sqrt{x+5}\ dx$$
$$= \left(\frac{2}{3}x(x+5)^{3/2}\right) - \frac{2}{3}\left(\frac{2}{5}(x+5)^{5/2}\right)\Bigg|_1^4$$
$$= \frac{2}{3}x(x+5)^{3/2} - \frac{4}{15}(x+5)^{5/2}\Bigg|_1^4$$
$$= \left(\frac{2}{3}(4)(4+5)^{3/2} - \frac{4}{15}(4+5)^{5/2}\right) - \left(\frac{2}{3}(1)(1+5)^{3/2} - \frac{4}{15}(1+5)^{5/2}\right)$$
$$= \left(\frac{2}{3}(4)(9)^{3/2} - \frac{4}{15}(9)^{5/2}\right) - \left(\frac{2}{3}(1)(6)^{3/2} - \frac{4}{15}(6)^{5/2}\right)$$
$$= \left(\frac{2}{3}(4)(27) - \frac{4}{15}(243)\right) - \left(\frac{2}{3}(6)^{3/2} - \frac{4}{15}(6)^{5/2}\right)$$
$$= 20.92.$$

(b) Let

$$u = 3x, \qquad dv = e^x\ dx,$$

which means that

$$du = 3\ dx, \qquad dv = \int e^x\ dx = e^x.$$

Applying the integration by parts formula, we get

$$\begin{aligned}\int u\ dv &= uv - \int v\ du\\ &= 3xe^x - \int e^x (3\ dx)\\ &= 3xe^x - 3\int e^x\ dx\\ &= 3xe^x - 3e^x + c.\end{aligned}$$

(c) When using integration by parts, we have to identify two functions multiplied together, identify the factor that simplifies when we take its derivative and set it to u, and set the other factor to dv. In this case, the integrand is the product of two functions, x and $\ln(x)$. The choice of which function to set to u is tricky. Our first instinct should be to set $u = x$, but then we would have to integrate $dv = \ln(x)\ dx$, which is tricky. Instead, if we set $u = \ln(x)$ then we can easily integrate $dv = x\ dx$. If we make this choice, then

$$du = \frac{1}{x} dx,$$

and

$$v = \int x\ dx = \frac{x^2}{2}.$$

We hold off evaluating the area under the curve between the bounds until the end of the problem. Now that we have u, du, v, and dv, we plug these functions into the integration by parts formula:

$$\begin{aligned}\int u\ dv &= uv - \int v\ du\\ &= \ln(x)\frac{x^2}{2} - \int \frac{x^2}{2}\frac{1}{x}\ dx\\ &= \frac{x^2\ln(x)}{2} - \int \frac{x}{2}\ dx\\ &= \frac{x^2\ln(x)}{2} - \frac{x^2}{4}.\end{aligned}$$

We do not write "$+c$" because we are working with a definite integral. Now that we have evaluated the anti-derivative, we plug in the bounds:

$$\begin{aligned}&\left.\frac{x^2\ln(x)}{2} - \frac{x^2}{4}\right|_1^e\\ &= \frac{e^2\ln(e)}{2} - \frac{e^2}{4} - \left(\frac{1^2\ln(1)}{2} - \frac{1^2}{4}\right)\\ &= \frac{e^2}{2} - \frac{e^2}{4} + \frac{1}{4}\\ &= \frac{e^2+1}{4}.\end{aligned}$$

(d) Note that x^2 becomes simpler if we take the derivative, and e^x is easily integrated. So let's set

$$u = x^2, \qquad dv = e^x\ dx,$$

which implies

$$du = 2x\ dx, \qquad v = \int e^x\ dx = e^x,$$

where we hold off on writing "$+c$" until the end of the problem. Plugging these quantities into the integration by parts formula, we get

$$\begin{aligned}\int u\ dv &= uv - \int v\ du \\ &= x^2e^x - \int e^x(2x\ dx) \\ &= x^2e^x - 2\int xe^x\ dx.\end{aligned}$$

Now we have to contend with the integral $\int xe^x\ dx$, which requires doing integration by parts a second time. This time, let

$$u = x, \qquad dv = e^x\ dx,$$

so that

$$du = dx, \qquad v = \int e^x\ dx = e^x.$$

Then $\int xe^x\ dx$ becomes

$$\begin{aligned}\int u\ dv &= uv - \int v\ du \\ &= xe^x - \int e^x\ dx \\ &= xe^x - e^x.\end{aligned}$$

Plugging this integral in for $\int xe^x\ dx$ above, we get a solution for the entire indefinite integral:

$$\begin{aligned}\int x^2e^x\ dx &= x^2e^x - 2(xe^x - e^x) + c \\ &= x^2e^x - 2xe^x + 2e^x + c.\end{aligned}$$

5. (a) This integral does not require u-substitution or integration by parts because it can be evaluated directly from basic rules of integration. First we can break the integral up over addition:

$$\begin{aligned}\int_{\ln(8)}^{3\sqrt{167}} x^2 + e^x\ dx &= \int_{\ln(8)}^{3\sqrt{167}} x^2\ dx + \int_{\ln(8)}^{3\sqrt{167}} e^x\ dx \\ &= \frac{x^3}{3}\Big|_{\ln(8)}^{3\sqrt{167}} + e^x\Big|_{\ln(8)}^{3\sqrt{167}} \\ &= \left(\frac{(3\sqrt{167})^3}{3} - \frac{\ln(8)^3}{3}\right) + \left(e^{3\sqrt{167}} - e^{\ln(8)}\right) \\ &= \left(\frac{167}{3} - \frac{\ln(8)^3}{3}\right) + \left(e^{3\sqrt{167}} - 8\right) = 291.05.\end{aligned}$$

(b) Observe that the derivative of $\ln(x)$ is $\frac{1}{x}$, which is a factor of the integrand. So let's apply u-substitution where

$$u = \ln(x), \quad \tfrac{du}{dx} = \tfrac{1}{x}, \quad du = \frac{1}{x}dx.$$

We can also recalculate the bounds using this formula for u. The new lower bound is

$$u(e^2) = \ln(e^2) = 2,$$

and the new upper bound is

$$u(e^4) = \ln(e^4) = 4.$$

Substituting for $\ln(x)$, $\frac{1}{x}dx$, and the upper and lower bounds, the definite integral becomes

$$\int_2^4 u\ du$$

$$= \frac{u^2}{2}\bigg|_2^4$$

$$= \frac{4^2}{2} - \frac{2^2}{2} = 8 - 2 = 6.$$

(c) The problem looks a lot like part (b), but note that using u-substitution in which $u = \ln(x)$ won't cancel out the denominator as it did before. It is necessary to solve this problem through integration by parts instead. How were you supposed to know that? You probably had to go through a frustrating process in which you tried various u-substitutions before giving up on that particular tool and trying versions of integration by parts until something worked. That is NOT a problem. In fact, that is what I had to do to solve this problem too. The thing about integration by parts and u-substitution is that they are tools, and integrals are problems in which the correct tools are not always obvious. There is no single process for solving many integrals, and they can be frustrating to even the most experienced mathematicians. If you struggled to find the correct method, you have joined a long and noble tradition.

Anyway, noting that we can differentiate $\ln(x)$ and that we can integrate $\frac{1}{x^5} = x^{-5}$, we set

$$u = \ln(x), \qquad dv = x^{-5}\ dx,$$

which means that

$$du = \frac{1}{x}\ dx, \qquad v = \int x^{-5}\ dx = \frac{x^{-4}}{-4} = \frac{-1}{4x^4}.$$

Plugging these functions into the formula for integration by parts, we get

$$\int u\ dv = uv - \int v\ du$$

$$= \ln(x)\frac{-1}{4x^4} - \int \frac{-1}{4x^4}(\frac{1}{x}\ dx)$$

$$= \frac{-\ln(x)}{4x^4} - \int \frac{-1}{4x^5}\ dx$$

$$= \frac{-\ln(x)}{4x^4} + \frac{1}{4}\int x^{-5}\ dx$$

$$= \frac{-\ln(x)}{4x^4} + \frac{1}{4}\frac{x^{-4}}{-4} + c$$

$$= \frac{-\ln(x)}{4x^4} - \frac{1}{16x^4} + c.$$

(d) First we can break the integral up over addition:

$$\int e^x\ dx + \int \sqrt{x}\ dx$$

Then we can solve each indefinite integral. The first integral evaluates to

$$\int e^x\ dx = e^x + c.$$

The second integral can be rewritten as

$$\int x^{1/2}\ dx,$$

so we apply the power rule for integrals:

$$\int x^{1/2}\ dx = \frac{x^{3/2}}{3/2} + c = \frac{2}{3}x^{3/2} + c.$$

Combining both integrals, we get

$$\int e^x\ dx + \int \sqrt{x}\ dx = e^x + \frac{2}{3}x^{3/2} + c.$$

(e) First we find the anti-derivative as we would while solving an indefinite integral. The integral breaks up over addition and subtraction:

$$\int_0^{10} y^2\ dy - \int_0^{10} 10y\ dy + \int_0^{10} 25\ dy.$$

Leaving the bounds aside for now, the first integral evaluates to

$$\int y^2\ dy = \frac{y^3}{3}.$$

We leave off "+c" since the constant term ultimately would cancel out once we plug in the bounds. The constant factor 10 can be brought outside the second integral

$$10\int y\ dy = 10\frac{y^2}{2} = 5y^2.$$

The third integral evaluates to

$$\int 25\ dy = 25y.$$

So the entire anti-derivative is

$$\frac{y^3}{3} - 5y^2 + 25y.$$

The area underneath the curve from 0 to 10 is

$$\left.\frac{y^3}{3} - 5y^2 + 25y\right|_0^{10}$$

$$= \left(\frac{(10)^3}{3} - 5(10)^2 + 25(10)\right) - \left(\frac{(0)^3}{3} - 5(0)^2 + 25(0)\right)$$

$$= \left(\frac{1000}{3} - 500 + 250\right) - 0 = 83\frac{1}{3}.$$

(f) We can observe that the integrand contains a part whose derivative (without its coefficient) exists elsewhere in the integrand. Specifically, the function underneath the square root, $x^3 + 2$, which becomes a multiple of x^2 when we take its derivative. That indicates that we should use u-substitution. First set u to the term under the square root:

$$u = x^3 + 2.$$

When we take the derivative, we get,

$$\begin{aligned} \frac{du}{dx} &= 3x^2, \\ du &= 3x^2 \; dx, \\ \frac{1}{3}du &= x^2 \; dx. \end{aligned}$$

We can also change the bounds:

$$u(2) = 2^3 + 2 = 10, \quad \text{and} \quad u(4) = 4^3 + 2 = 66.$$

Next, we substitute u for x"$3 + 2$, we substitute $1/3du$ for x^2dx, we substitute 10 for the lower bound, and we substitute 66 for the upper bound. The problem becomes

$$\begin{aligned} &\frac{1}{3}\int_{10}^{66} \sqrt{u} \; du \\ &= \frac{1}{3}\int_{10}^{66} u^{1/2} \; du \\ &= \frac{1}{3}\frac{2}{3}u^{3/2}\Big|_{10}^{66} \\ &= \frac{2}{9}(66^{3/2} - 10^{3/2}) = 112.13. \end{aligned}$$

(g) This integrand is the product of two functions, so we can consider using integration by parts. We set one function as u and take its derivative to find du, and we set the other function as dv and take its anti-derivative to find v, then we plug all of these functions into

$$\int u \; dv = uv - \int v \; du,$$

and with any luck, the remaining integral is easier to solve. Let's set

$$u = \ln(x) \quad \text{and} \quad dv = x^2 \; dx,$$

then du and v are

$$du = \frac{1}{x} \; dx \quad \text{and} \quad v = \int x^2 \; dx = \frac{x^3}{3}.$$

Plugging into the equations by parts formula, this becomes

$$\begin{aligned} &\frac{x^3 \ln(x)}{3} - \frac{1}{3}\int \frac{x^3}{x} \; dx \\ &\frac{x^3 \ln(x)}{3} - \frac{1}{3}\int x^2 \; dx \\ &\frac{x^3 \ln(x)}{3} - \frac{1}{3}\frac{x^3}{3} + c \\ &x^3\left(\frac{\ln(x)}{3} - \frac{1}{9}\right) + c. \end{aligned}$$

(h) Since the integral has an infinite upper bound, this is an improper integral. We can rewrite the integral as

$$\lim_{A\to\infty}\int_1^A \frac{3x^2+2x+1}{(x^3+x^2+x+1)^2}\ dx$$

and solve as if A were finite. We will take the limit only after solving the integral. Note that the derivative of the term inside the square in the denominator is exactly equal to the numerator. So let's use u-substitution by setting

$$u = x^3+x^2+x+1,$$

$$du = 3x^2+2x+1\ dx.$$

Let's not substitute the bounds, instead we will substitute back in for u once we've taken the integral. The integral (without the bounds) now becomes

$$\int \frac{1}{u^2}\ du = \int u^{-2}\ du = \frac{u^{-1}}{-1} = \frac{-1}{u}.$$

Substituting back for u and considering the bounds, this becomes

$$\left.\frac{-1}{x^3+x^2+x+1}\right|_1^A$$

$$\frac{-1}{A^3+A^2+A+1} - \frac{-1}{1^3+1^2+1+1}$$

$$\frac{-1}{A^3+A^2+A+1} + \frac{1}{4}.$$

Finally, we take the limit as $A\to\infty$:

$$\lim_{A\to\infty}\frac{-1}{A^3+A^2+A+1} + \frac{1}{4}.$$

Note that as A gets larger and larger, so does the polynomial A^3+A^2+A+1. The denominator therefore approaches $-\infty$, and the fraction $1/-\infty$ approaches zero. Replacing the limit with 0, the integral is equal to $1/4$.

6. (a) The domain of the function contains values of x between 0 and 1. First we have to demonstrate that the function is never negative in this domain. When $x=0$, the function is

$$f(0) = \frac{3\sqrt{0}}{2} = 0,$$

which isn't positive, but also isn't negative. Every other x in the domain is positive, implying that the function is positive as well. Therefore $f(x)$ is never negative for any x in the domain.

Second, we have to demonstrate that the function integrates to 1 over its domain. That means we have to demonstrate that the following definite integral is equal to 1:

$$\int_0^1 \frac{3\sqrt{x}}{2}\ dx$$

$$= \frac{3}{2}\int_0^1 \sqrt{x}\ dx$$

$$= \frac{3}{2}\int_0^1 x^{1/2}\ dx$$

$$= \frac{3}{2}\left(\frac{2}{3}x^{3/2}\right)\Big|_0^1$$

$$= x^{3/2}\Big|_0^1 = (1)^{3/2} - 0^{3/2} = 1.$$

Therefore $f(x)$ is a PDF.

(b) We already did most of the necessary work in part (a). We simply plug in new bounds.

To calculate the probability that the student gets an A, we find the area under the curve of the PDF from .9 to 1:

$$P(A) = P(.9 < x < 1) = x^{3/2}\Big|_{.9}^{1} = 1^{3/2} - .9^{3/2} = 0.146.$$

To calculate the probability that the student gets a B, we find the area under the curve of the PDF from .8 to .9:

$$P(B) = P(.8 < x < .9) = x^{3/2}\Big|_{.8}^{.9} = .9^{3/2} - .8^{3/2} = 0.138.$$

To calculate the probability that the student gets a C, we find the area under the curve of the PDF from .7 to .8:

$$P(C) = P(.7 < x < .8) = x^{3/2}\Big|_{.7}^{.8} = .8^{3/2} - .7^{3/2} = 0.130.$$

To calculate the probability that the student gets a D, we find the area under the curve of the PDF from .6 to .7:

$$P(D) = P(.6 < x < .7) = x^{3/2}\Big|_{.6}^{.7} = .7^{3/2} - .6^{3/2} = 0.121.$$

Finally, to calculate the probability that the student gets an F, we find the area under the curve of the PDF from 0 to .6:

$$P(F) = P(x < .6) = x^{3/2}\Big|_{0}^{.6} = .6^{3/2} - 0^{3/2} = 0.465.$$

(c) This problem involves two challenges: first we have to plug the relevant information into the formula for an expected value,

$$E(x) = \int_{-\infty}^{\infty} xf(x)\;dx,$$

and second we have to solve that definite integral. The function is $f(x) = \frac{3\sqrt{x}}{2}$, and the domain is $x \in [0,1]$. Plugging this information into the formula gives us

$$E(x) = \int_{-\infty}^{\infty} xf(x)\;dx = \int_0^1 x\frac{3\sqrt{x}}{2}\;dx$$

$$= \frac{3}{2}\int_0^1 x\sqrt{x}\;dx$$

$$= \frac{3}{2}\int_0^1 x(x^{1/2})\;dx$$

$$= \frac{3}{2}\int_0^1 x^{3/2}\ dx$$

$$= \frac{3}{2}\left(\frac{2}{5}x^{5/2}\right)\Big|_0^1$$

$$= \frac{3}{5}\left(x^{5/2}\right)\Big|_0^1 = \frac{3}{5}\left(1^{5/2} - 0^{5/2}\right) = \frac{3}{5} = 0.6.$$

So the average score is 60%. Ouch.

(d) The second part of the variance formula is the square of .6:

$$V(x) = E(x^2) - E(x)^2,$$

$$V(x) = E(x^2) - .36.$$

So the challenge is finding $E(x^2)$, which involves plugging into the given formula and solving the resulting definite integral. The formula is

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\ dx = \int_0^1 x^2 \frac{3\sqrt{x}}{2}\ dx$$

$$= \frac{3}{2}\int_0^1 x^2\sqrt{x}\ dx$$

$$= \frac{3}{2}\int_0^1 x^2(x^{1/2})\ dx$$

$$= \frac{3}{2}\int_0^1 x^{5/2}\ dx$$

$$= \frac{3}{2}\left(\frac{2}{7}x^{7/2}\right)\Big|_0^1$$

$$= \frac{3}{7}\left(x^{7/2}\right)\Big|_0^1 = \frac{3}{7}\left(x^{7/2}\right)\Big|_0^1 = \frac{3}{7}\left(1^{7/2} - 0^{7/2}\right) = \frac{3}{7} = 0.429.$$

Plugging this value back into the formula for variance gives us

$$V(x) = E(x^2) - E(x)^2 = .429 - .36 = 0.069.$$

Finally, the standard deviation is the square root of the variance:

$$SD(x) = \sqrt{V(x)} = \sqrt{0.069} = 0.262.$$

Therefore the mean score is 60%, and the standard deviation is 26.2%.

7. A function is a PDF if it is non-negative everywhere in its domain and if it integrates to 1 over its domain. First, note that $f(x) = \frac{3}{7x^4}$ is positive for all x between $\frac{1}{2}$ and 1, so the first condition is met. Next, we evaluate

$$\int_{1/2}^1 \frac{3}{7x^4}\ dx$$

$$= \frac{3}{7} \int_{1/2}^{1} \frac{1}{x^4}\ dx$$

$$= \frac{3}{7} \int_{1/2}^{1} x^{-4}\ dx$$

$$= \frac{3}{7} \frac{x^{-3}}{-3} \Big|_{1/2}^{1}$$

$$= \frac{-1}{7x^3} \Big|_{1/2}^{1}$$

$$= \frac{-1}{7(1)^3} - \frac{-1}{7(1/2)^3}$$

$$= \frac{-1}{7} + \frac{1}{7(1/8)}$$

$$= \frac{-1}{7} + \frac{1}{7/8}$$

$$= \frac{-1}{7} + \frac{8}{7} = \frac{7}{7} = 1.$$

8. Note that e^{x+1} is positive for all x in the domain $x \in [0, k]$ no matter what k is. Our task here is to choose the value of k that makes the integral evaluate to 1. We solve the definite integral as we always would, and plug k into the anti-derivative in the last step. Then we set this expression equal to 1 and solve for k. The definite integral is:

$$\int_0^k e^{x+1}\ dx.$$

We can solve this integral using u-substitution. Let $u = x + 1$. Then $du = dx$, the lower bound is $u(0) = 0 + 1 = 1$, and the upper bound is $u(k) = k + 1$. Substituting, we get

$$\int_1^{k+1} e^u\ du$$

$$= e^u \Big|_1^{k+1}$$

$$= e^{k+1} - e^1$$

$$= e^{k+1} - e.$$

Now we set this expression equal to 1 and solve for k:

$$e^{k+1} - e = 1,$$

$$e^{k+1} = e + 1,$$

$$k + 1 = \ln(e + 1),$$

$$k = \ln(e + 1) - 1.$$

So $g(x)$ is a PDF on the domain $x \in [0, \ln(e + 1) - 1]$.

9. (a) If $f(x)$ is a PDF, then the formula for a CDF is

$$F(x) = \int_{-\infty}^{x} f(x)\ dx,$$

where the lower bound of $-\infty$ stands in for whatever the lower bound of the distribution happens to be. In this case, the formula for the CDF is

$$\begin{aligned} F(x) &= \int_{0}^{x} \lambda e^{-\lambda x}\ dx \\ &= \lambda \int_{0}^{x} e^{-\lambda x}\ dx. \end{aligned}$$

We can solve this integral using u-substitution, where $u = -\lambda x$, $du = -\lambda\ dx$, and therefore $\frac{-1}{\lambda} du = dx$. The lower bound becomes

$$u(0) = -\lambda(0) = 0,$$

and the upper bound becomes

$$u(x) = -\lambda x.$$

Substituting, the integral becomes

$$\begin{aligned} &\lambda\left(\frac{-1}{\lambda}\right)\int_{0}^{-\lambda x} e^{u}\ du \\ &= -e^{u}\Big|_{0}^{-\lambda x} \\ &= -e^{-\lambda x} - (-e^{0}) \\ &= -e^{-\lambda x} + 1 \\ &= 1 - e^{-\lambda x}, \end{aligned}$$

which is the correct formula for the exponential distribution CDF.

(b) The CDF of a function at a value c tells you the probability that a randomly drawn value from that distribution is less than or equal to c,

$$F(c) = P(x \le c),$$

and the probability that a randomly drawn value is between two values c and d is given by

$$P(d \le x \le c) = P(x \le c) - P(x \le d) = F(c) - F(d).$$

So in this case, the probability that x is between 30 and 50 is

$$\begin{aligned} &F(50) - F(30) \\ &= 1 - e^{-\lambda(50)} - (1 - e^{-\lambda(30)}). \end{aligned}$$

In the case of Claudio Cioffi-Revilla's model, $\lambda = -0.023$, so the probability becomes

$$= 1 - e^{-0.023(50)} - (1 - e^{-0.023(30)}) = 0.185.$$

(c) If $f(x)$ is a PDF, then the formula for the expected value of this PDF is

$$E(x) = \int_{-\infty}^{\infty} x\ f(x)\ dx$$

where the bounds $-\infty$ and ∞ stand in for the bounds of the domain of the PDF. Since the lower bound of the domain is 0, we replace $-\infty$ with 0. Let's replace the upper bound ∞ with a value k for now. At the end of the problem we will plug a very large value in for k to see if the expression simplifies. The integral now becomes

$$\begin{aligned} E(x) &= \int_0^k x(.023e^{-.023x})\ dx \\ &= \int_0^k .023xe^{-.023x}\ dx. \end{aligned}$$

We have to evaluate this integral using integration by parts, and let's not deal with the bounds until the end of the problem. Let

$$u = .023x, \qquad dv = e^{-.023x}\ dx,$$

which means that

$$du = .023\ dx.$$

In order to find v, we have to solve

$$v = \int e^{-.023x}\ dx,$$

which requires u-substitution. Unfortunately, the naming conventions muddle the fact that we have a u from integration by parts and a u from u-substitution. So let's call this method w-substitution in this case. Let

$$w = -.023x, \quad dw = -.023\ dx, \quad \frac{-1}{.023}\ dw = dx.$$

Substituting, this integral becomes

$$v = \frac{-1}{.023}\int e^w\ dw = \frac{-1}{.023}e^w.$$

Substituting back in for w, we get

$$v = \frac{-1}{.023}e^{-.023x}.$$

Now that we have u, v, du, and dv, we plug these functions into the formula for integration by parts:

$$\begin{aligned} \int u\ dv &= uv - \int v\ du \\ &= .023x\left(\frac{-1}{.023}e^{-.023x}\right) - \int \frac{-1}{.023}e^{-.023x}(.023\ dx) \\ &= -xe^{-.023x} + \int e^{-.023x}\ dx. \end{aligned}$$

The remaining integral is the same one we solved above using w-substitution, so the whole expression becomes

$$-xe^{-.023x} - \frac{1}{.023}e^{-.023x}.$$

Since this integral is definite with bounds at 0 and k, we plug in the bounds:

$$\begin{aligned} & -xe^{-.023x} - \frac{1}{.023}e^{-.023x}\Big|_0^k \\ &= \left(-ke^{-.023k} - \frac{1}{.023}e^{-.023k}\right) - \left(-0e^{-.023(0)} - \frac{1}{.023}e^{-.023(0)}\right) \\ &= \left(-ke^{-.023k} - \frac{1}{.023}e^{-.023k}\right) + \frac{1}{.023}. \end{aligned}$$

Now, since k is standing in for ∞, we can get an idea about how this expression evaluates by thinking about plugging very large numbers in for k. When k is very large, the quantity $e^{-.023k}$ is infinitesimally close to 0, which implies that both terms inside the parentheses are 0. Therefore the whole expression evaluates to

$$E(x) = \frac{1}{.023} = 43.5 \text{ weeks}.$$

(d) The variance of a PDF is given by

$$V(x) = E(x^2) - E(x)^2.$$

In part (c) we found that $E(x) = 43.5$, so the variance becomes

$$\begin{aligned} V(x) &= E(x^2) - (43.5)^2 \\ &= E(x^2) - 1890.4. \end{aligned}$$

So we just have to find $E(x^2)$, which is given by the following formula:

$$E(x^2) = \int_{-\infty}^{\infty} x^2\ f(x)\ dx.$$

As in part (c), the integral specifically can be written as

$$\begin{aligned} E(x^2) &= \int_0^k x^2\ (.023e^{-.023x})\ dx \\ &= \int_0^k .023x^2e^{-.023x}\ dx. \end{aligned}$$

To solve this integral, we use integration by parts where

$$u = .023x^2, \qquad dv = e^{-.023x}\ dx,$$

which implies that

$$du = .046x\ dx,$$

and, from the work we did in part (c),

$$v = \int e^{-.023x}\ dx = \frac{-1}{.023}e^{-.023x}.$$

Plugging u, v, du, and dv into the integration by parts formula we get

$$\int u\ dv = uv - \int v\ du$$

$$\begin{aligned} &= .023x^2\left(\frac{-1}{.023}e^{-.023x}\right) - \int \frac{-1}{.023}e^{-.023x}(.046x\ dx) \\ &= -x^2e^{-.023x} - \int \frac{-.046}{.023}xe^{-.023x}\ dx. \\ &= -x^2e^{-.023x}\Big|_0^k + 2\int_0^k xe^{-.023x}\ dx. \end{aligned}$$

Remember that in part (c) we found that

$$E(x) = \int_0^{\infty} .023xe^{-.023x}\ dx = 43.5.$$

That means that we can divide both sides by .023 to find a solution to $\int xe^{-.023x}\ dx$:

$$\int_0^{\infty} xe^{-.023x}\ dx = \frac{43.5}{.023} = 1890.4.$$

Substituting this number into $E(x^2)$ gives us

$$E(x^2) = \lim_{k\to\infty} -x^2e^{-.023x}\Big|_0^k + 2(1890.4).$$

Note that a very large value of k will make the quantity $e^{-.023k}$ approach 0. Therefore the entire first term approaches 0, and

$$E(x^2) = 2(1890.4) = 3780.7.$$

Finally, we plug $E(x^2)$ into the formula for the variance of the distribution:

$$V(x) = E(x^2) - 1890.4$$

$$V(x) = 3780.7 - 1890.4 = 1890.4.$$

The standard deviation is the square root:

$$SD(x) = \sqrt{1890.4} = 43.5 \text{ weeks}.$$

The fact that the mean equals the standard deviation is no accident. That is a property of the exponential distribution. More recent work employs models that do not make such a restrictive assumption about the mean and standard deviation of the dependent variable.

10. (a) The expected value of income is

$$\text{E(Income)} = \text{E}\Big(\alpha{+}\beta_1\text{Gender} + \beta_2\text{Education} + \beta_3(\text{Gender}{\times}\text{Education}) + \varepsilon\Big).$$

We can break this expected value up over addition,

$$\text{E}\Big(\alpha{+}\beta_1\text{Gender} + \beta_2\text{Education} + \beta_3(\text{Gender}{\times}\text{Education})\Big) + \text{E}(\varepsilon).$$

The problem tells us to treat the coefficients and variables as if they are constant. That means that the whole expression $\alpha{+}\beta_1\text{Gender} + \beta_2\text{Education} + \beta_3(\text{Gender}{\times}\text{Education})$ is a constant, and therefore that the first expected value is the expected value of a constant. Since the expected value of a constant is just the constant itself, the whole expression becomes

$$\alpha{+}\beta_1\text{Gender} + \beta_2\text{Education} + \beta_3(\text{Gender}{\times}\text{Education}) + \text{E}(\varepsilon).$$

For the second expected value, we know that a regression error term is a random variable with a mean of zero. Since expected value is another name for the mean, we can replace this term with zero. The expected value of income is

$$\text{E(Income)} = \alpha{+}\beta_1\text{Gender} + \beta_2\text{Education} + \beta_3(\text{Gender}{\times}\text{Education}).$$

(b) The derivative of E(Income) with respect to gender is

$$\tfrac{dE(\text{Income})}{d\text{Gender}} = \tfrac{d}{d\text{Gender}}\Big(\alpha{+}\beta_1\text{Gender} + \beta_2\text{Education} + \beta_3(\text{Gender}{\times}\text{Education})\Big).$$

It might be difficult to connect this equation to the derivatives we took in chapter 4. If might help to replace every occurrence of the word gender with x, every occurrence of the word income with y, and every occurrence of the word education with another symbol like z. Then the equation is

$$\frac{dE(y)}{dx} = \frac{d}{dx}(\alpha + \beta_1 x + \beta_2 z + \beta_3 xz).$$

The derivative is

$$\frac{dE(y)}{dx} = \beta_1 + \beta_3 z,$$

and plugging words back in for the symbols gives us

$$\frac{dE(\text{Income})}{d\text{Gender}} = \beta_1 + \beta_3\text{Education}.$$

(c) If $\alpha = 10000$, $\beta_1 = 15000$, $\beta_2 = 5000$, and $\beta_3 = -500$ then the derivative of E(Income) with respect to gender is

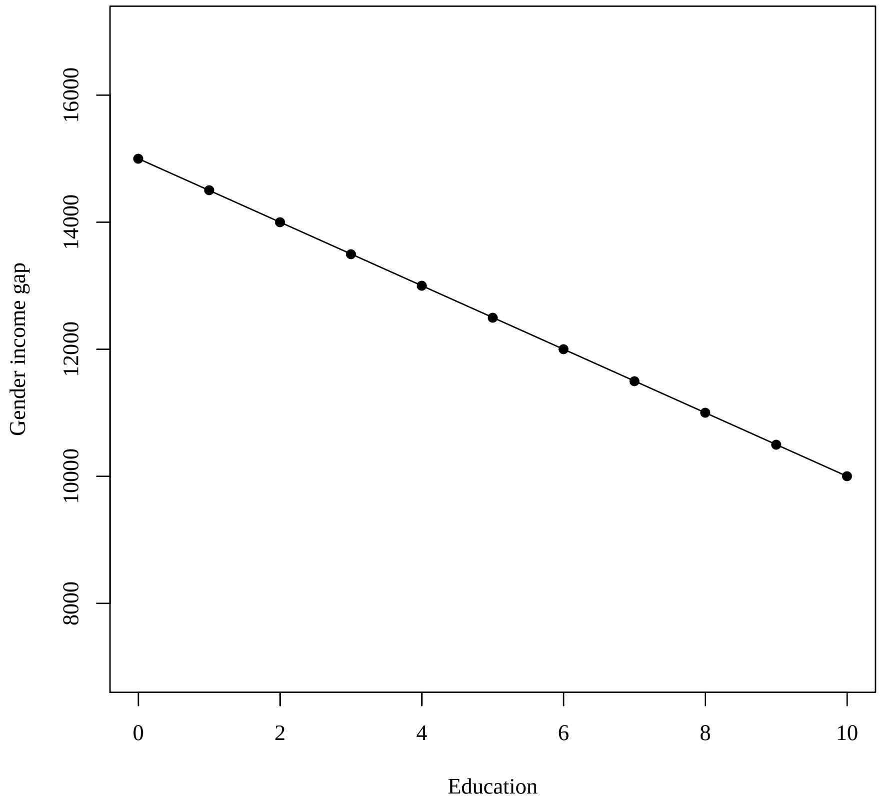
$$\frac{dE(\text{Income})}{d\text{Gender}} = 15000 \text{ - } 500\text{Education}.$$

The derivative of E(Income) with respect to gender is one way to measure the gender income gap. We simply plug in the level of education and compute $\frac{dE(\text{Income})}{d\text{Gender}}$ using the above equation. These values are listed in the table below:

At education level	The gender income gap is
0	15000-500(0)=15000
1	15000-500(1)=14500
2	15000-500(2)=14000
3	15000-500(3)=13500
4	15000-500(4)=13000
5	15000-500(5)=12500
6	15000-500(6)=12000
7	15000-500(7)=11500
8	15000-500(8)=11000
9	15000-500(9)=10500
10	15000-500(10)=10000

The table tells us that while the gender income gap is decreasing with higher levels of education, the gap exists everywhere. Even among people with the highest education level men make $10,000 more than women, on average.

(d) The graph of these values is below:



(e) The variance of the gender gap in expected income is

$$V\left(\frac{dE(\text{Income})}{d\text{Gender}}\right) = V(\beta_1 + \beta_3\text{Education}).$$

Here we are treating β_1 and β_3 as the random variables and education as constant. The rule regarding the variance of sums of random variables is

$$V(aX + bY) = a^2V(x) + b^2V(y) + 2abCov(x, y).$$

Substituting 1 for a, β_1 for x, education for b, and β_3 for y, the variance is

$$V\left(\frac{dE(\text{Income})}{d\text{Gender}}\right) = V(\beta_1) + \left(\text{Education}^2 \times V(\beta_3)\right) + \left(2 \times \text{Education} \times Cov(\beta_1, \beta_3)\right).$$

(f) Assuming that $V(\beta_1) = 60000$, $V(\beta_3) = 20000$, and $Cov(\beta_1, \beta_3) = 5000$, the variance we derived in part (e) becomes

$$V\left(\frac{dE(\text{Income})}{d\text{Gender}}\right) = 60000 + 20000\,\text{Education}^2 + 10000\,\text{Education}.$$

The standard errors are given by the square root:

$$SE\left(\frac{dE(\text{Income})}{d\text{Gender}}\right) = \sqrt{60000 + 20000\,\text{Education}^2 + 10000\,\text{Education}}.$$

The standard errors at each education level are listed in the following table

At education level	Standard Error
0	$\sqrt{60000 + 20000(0)^2 + 10000(0)} = 173.21$
1	$\sqrt{60000 + 20000(1)^2 + 10000(1)} = 244.95$
2	$\sqrt{60000 + 20000(2)^2 + 10000(2)} = 360.56$
3	$\sqrt{60000 + 20000(3)^2 + 10000(3)} = 489.90$
4	$\sqrt{60000 + 20000(4)^2 + 10000(4)} = 624.50$
5	$\sqrt{60000 + 20000(5)^2 + 10000(5)} = 761.58$
6	$\sqrt{60000 + 20000(6)^2 + 10000(6)} = 900.00$
7	$\sqrt{60000 + 20000(7)^2 + 10000(7)} = 1039.23$
8	$\sqrt{60000 + 20000(8)^2 + 10000(8)} = 1178.98$
9	$\sqrt{60000 + 20000(9)^2 + 10000(9)} = 1319.09$
10	$\sqrt{60000 + 20000(10)^2 + 10000(10)} = 1459.45$

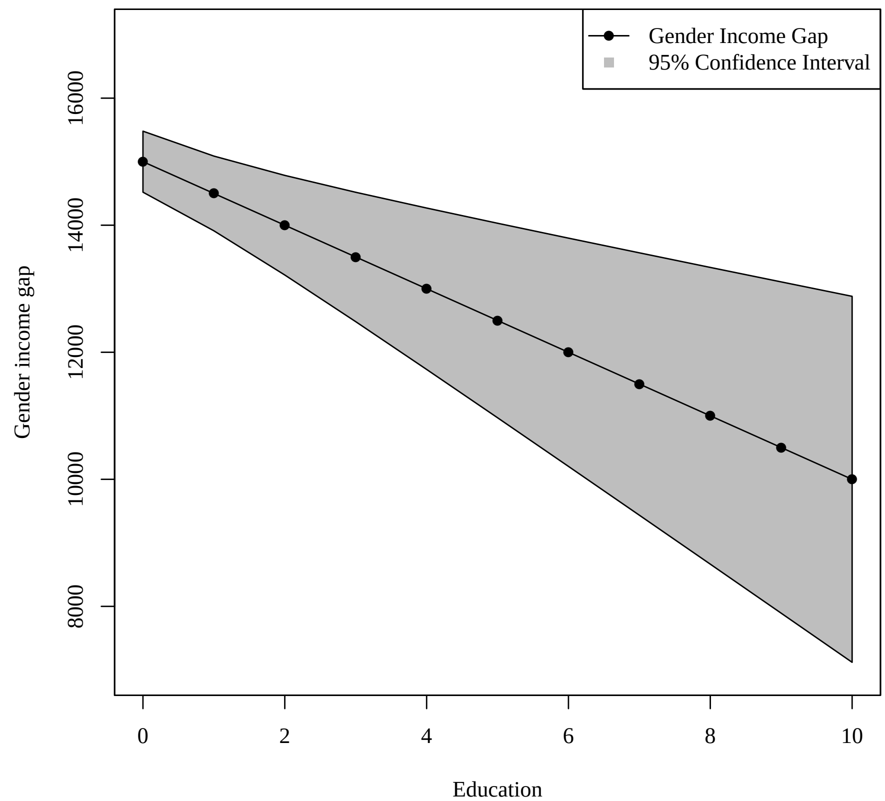
The 95% confidence interval's lower and upper bounds, given by

Lower bound = effect - 1.96 × standard error
Upper bound = effect + 1.96 × standard error,

are:

At education level	Lower bound	Upper bound
0	15000-1.96(173.21)=**14661**	15000+1.96(173.21)=**15339**
1	14500-1.96(244.95)=**14020**	14500+1.96(244.95)=**14980**
2	14000-1.96(360.56)=**13293**	14000+1.96(360.56)=**14707**
3	13500-1.96(489.90)=**12540**	13500+1.96(489.90)=**14460**
4	13000-1.96(624.50)=**11776**	13000+1.96(624.50)=**14224**
5	12500-1.96(761.58)=**11007**	12500+1.96(761.58)=**13993**
6	12000-1.96(900.00)=**10236**	12000+1.96(900.00)=**13764**
7	11500-1.96(1039.23)=**9463**	11500+1.96(1039.23)=**13537**
8	11000-1.96(1178.98)=**8689**	11000+1.96(1178.98)=**13311**
9	10500-1.96(1319.09)=**7915**	10500+1.96(1319.09)=**13085**
10	10000-1.96(1459.45)=**7139**	10000+1.96(1459.45)=**12861**

(g) The graph of the gender income gap against education is again drawn below. This time the lower and upper bounds are drawn on the graph as well, and the entire 95% confidence interval is shaded in grey.



To read this graph: first choose a level of education on the x-axis. Then trace a vertical line up to the middle line. The y-value of the middle line at this point is the average gender income gap among people with the education level you've selected. On the same vertical line, move down to the bottom of the grey area and up to the top of the grey area. The y-values of these two points are the lower and upper bounds of the 95% confidence interval – in statistics we are never certain about any prediction, and here we state that we are 95% confident that the true gender income gap at the given education level is between these two bounds.

7 Multivariate Calculus

1. (a) A function with a domain of R^6 and a range of R takes inputs which are ordered sextuplets (6 dimensional numbers) and returns outputs which are unidimensional ("regular" real) numbers. An example of a function of this type is

$$f(x_1, x_2, x_3, x_4, x_5, x_6) = x_1 + x_2 + x_3 + x_4 + x_5 + x_6.$$

This function would map the ordered sextuplet $(1, 2, 3, 4, 5, 6)$ to the real number $1+2+3+4+5+6 = 21$. Incidentally, any linear regression model with one dependent variable and six independent variables is a function that maps $R^6 \to R$, such as

$$f(x_1, x_2, x_3, x_4, x_5, x_6) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6.$$

(b) A function with a domain of R^3 and a range of R^3 takes inputs which are ordered triplets and returns outputs which are also ordered triplets. An example of a function of this type is

$$f(x_1, x_2, x_3) = \left(x_1 + x_2 + x_3, x_1x_2x_3, \frac{x_1 + x_2}{x_3}\right).$$

This function would map the ordered triplet $(1, 2, 3)$ to the ordered triplet $\left(1+2+3, 1 \times 2 \times 3, \frac{1+2}{3}\right) = (6, 6, 1)$.

(c) A function with a domain of R and a range of R^4 takes inputs which are unidimensional real numbers and returns outputs which are ordered quadruplets. An example of a function of this type is

$$f(x) = (2x, x^2, x - 7, e^x).$$

This function would map 2 to the ordered quadruplet $\left(2(2), 2^2, 2 - 7, e^2\right) = (4, 4, -5, 7.39)$.

2. The first step to solving this problem is simply translating the set-builder notation into English. The domain is

$$\{(x, y) \in \mathbf{R}^2 \,|\, x + y \geq 0\},$$

which translates to the set of all ordered pairs in the set of real-numbered ordered pairs such that the sum of the two numbers in the ordered pair add to at least zero. In other words, we need to find a function that is undefined if $x + y$ is negative.

The range is

$$\{(x, y) \in \mathbf{R} \,|\, f(x, y) \geq 0\},$$

which is the set of (unidimensional) real numbers that are greater than or equal to 0. That means that the function must only be able to output numbers that are zero or positive.

A example of such a function is

$$f(x, y) = \sqrt{x + y}.$$

Since we are taking the square root of $x + y$, that sum must be either positive or zero. And the square root function only returns values that are positive or zero. Therefore this function has the domain and range specified in this problem.

3. (a) First, observe that we cannot simply plug in $x = 3$ and $y = 1$ as that would make the denominator equal zero. But we can use a trick: the numerator,

$$x^2 - xy + 6y^2,$$

is quadratic and can be factored. Factoring a quadratic expression is conceptually trickier when it contains two variables instead of one, but the math is basically the same. Pretend that $y = 1$. Then the expression becomes

$$x^2 - x + 6,$$

and two numbers that add to -1 and multiply to 6 are -3 and 2, so this expression factors to $(x-3)(x+2)$. The more general expression factors to $(x-3y)(x+2y)$. So the limit can be rewritten as

$$\lim_{(x,y)\to(3,1)} \frac{(x-3y)(x+2y)}{x-3y}$$
$$= \lim_{(x,y)\to(3,1)} (x+2y).$$

Since we've cancelled out the denominator, it is now safe to plug in (3,1). The limit equals $3 + 2(1) = 5$.

(b) Since both x and y are approaching infinity, a good strategy is to rewrite the limit so that as many terms as possible become fractions with constants in the numerator and variables in the denominator. We saw in chapter 4 that the limit of such a fraction is 0. This strategy is easier to implement here because the function inside the limit is itself a fraction, and we can multiply or divide the top and bottom of a fraction by the same thing without changing its value. With that in mind, let's divide the top and bottom by x^2y^2. The limit becomes

$$\lim_{(x,y)\to(\infty,\infty)} \frac{2x^2y^2 + 3y^2 - 7x - 15}{7x^2y^2 - 5x^2 + 2y + 10} \frac{\frac{1}{x^2y^2}}{\frac{1}{x^2y^2}}$$
$$= \lim_{(x,y)\to(\infty,\infty)} \frac{2 + \frac{3}{x^2} - \frac{7}{xy^2} - \frac{15}{x^2y^2}}{7 - \frac{5}{y^2} + \frac{2}{x^2y} + \frac{10}{x^2y^2}}$$

Each of these newly created fractions has a constant in the numerator and variables in the denominator. Since both variables go to infinity, these fractions all individually go to zero. The limit is equal to $\frac{2}{7}$.

4. (a) $f(x,y) = (x+y)\sqrt{x-y}$

To find the gradient, we take the first partial derivative of the function with respect to each independent variable, and then arrange these partial derivatives in a vector.

The first partial derivative with respect to x,

$$f_x(x,y) = \frac{\partial}{\partial x}\Big((x+y)\sqrt{x-y}\Big),$$

requires the product rule:

$$f_x(x,y) = \frac{\partial}{\partial x}\Big((x+y)\Big)\sqrt{x-y} + (x+y)\frac{\partial}{\partial x}\Big(\sqrt{x-y}\Big).$$

This expression contains two partial derivatives. The first is

$$\frac{\partial}{\partial x}\Big((x+y)\Big) = 1.$$

Remember that we are taking the partial derivative with respect to x, meaning that we treat x as the variable and y as a constant. Here the derivative of x is 1 and the derivative of y, treated as constant, is 0. Next, the second partial derivative is

$$\frac{\partial}{\partial x}\Big(\sqrt{x-y}\Big) = \frac{1}{2\sqrt{x-y}},$$

where again y does not enter into the calculation because it is being treated as a constant. Substituting these two partial derivatives into the overall partial derivative gives us

$$f_x(x,y) = \sqrt{x-y} + \frac{x+y}{2\sqrt{x-y}}.$$

The first partial derivative with respect to y,

$$f_y(x,y) = \frac{\partial}{\partial y}\Big((x+y)\sqrt{x-y}\Big),$$

requires nearly the same calculation as the partial derivative with respect to x. The only difference is that the chain rule implies that we multiply the derivative of the square root by -1:

$$f_y(x,y) = \sqrt{x-y} - \frac{x+y}{2\sqrt{x-y}}.$$

Therefore the gradient is

$$\nabla f(x,y) = \begin{bmatrix} \sqrt{x-y} + \frac{x+y}{2\sqrt{x-y}} \\ \\ \sqrt{x-y} - \frac{x+y}{2\sqrt{x-y}} \end{bmatrix}.$$

(b) $g(x,y) = e^{x^2+y^2-2x+5y+7}$

Remember from chapter 4 that the derivative of e^x is simply e^x again, and that the chain rule tells us that if $f(x) = e^{g(x)}$, then

$$f'(x) = e^{g(x)}g'(x).$$

With that in mind, the partial derivative with respect to x is

$$g_x(x,y) = e^{x^2+y^2-2x+5y+7}\left(\frac{\partial}{\partial x}(x^2+y^2-2x+5y+7)\right),$$

and the partial derivative with respect to y is

$$g_y(x,y) = e^{x^2+y^2-2x+5y+7}\left(\frac{\partial}{\partial y}(x^2+y^2-2x+5y+7)\right).$$

All we have to do to find the gradient is to find

$$\frac{\partial}{\partial x}(x^2+y^2-2x+5y+7) \quad \text{and} \quad \frac{\partial}{\partial y}(x^2+y^2-2x+5y+7)$$

and substitute them into the overall partial derivatives. Treating x as a variable and y as a constant shows us that

$$\frac{\partial}{\partial x}(x^2+y^2-2x+5y+7) = 2x-2.$$

Treating y as a variable and x as a constant shows us that

$$\frac{\partial}{\partial y}(x^2+y^2-2x+5y+7) = 2y+5.$$

Therefore the gradient is

$$\nabla g(x,y) = \begin{bmatrix} (2x-2)e^{x^2+y^2-2x+5y+7} \\ \\ (2y+5)e^{x^2+y^2-2x+5y+7} \end{bmatrix}.$$

(c) $h(x, y) = \ln(x + \sqrt{y})$

Remember from chapter 4 that the derivative of $\ln(x)$ is $\frac{1}{x}$ again, and that the chain rule tells us that if $f(x) = \ln\Big(g(x)\Big)$, then

$$f'(x) = \frac{1}{g(x)} g'(x) = \frac{g'(x)}{g(x)}.$$

With that in mind, the partial derivative with respect to x is

$$h_x(x, y) = \frac{1}{\ln(x + \sqrt{y})} \left(\frac{\partial}{\partial x}(x + \sqrt{y})\right),$$

and the partial derivative with respect to y is

$$h_y(x, y) = \frac{1}{\ln(x + \sqrt{y})} \left(\frac{\partial}{\partial y}(x + \sqrt{y})\right),$$

All we have to do to find the gradient is to find

$$\frac{\partial}{\partial x}(x + \sqrt{y}) \quad \text{and} \quad \frac{\partial}{\partial y}(x + \sqrt{y})$$

and substitute them into the overall partial derivatives. Treating x as a variable and y as a constant shows us that

$$\frac{\partial}{\partial x}(x + \sqrt{y}) = 1.$$

Treating y as a variable and x as a constant shows us that

$$\frac{\partial}{\partial y}(x + \sqrt{y}) = \frac{1}{2\sqrt{y}}.$$

Therefore the gradient is

$$\nabla h(x, y) = \begin{bmatrix} \frac{1}{x+\sqrt{y}} \\ \\ \frac{1}{2\sqrt{y}(x+\sqrt{y})} \end{bmatrix}.$$

(d) $j(x, y) = \dfrac{x^2 + y^2}{x^3 - 4xy - y^2}$

Both partial derivatives require the quotient rule. The partial derivative with respect to x is

$$j_x(x, y) = \frac{(x^3 - 4xy - y^2)\frac{\partial}{\partial x}(x^2 + y^2) - (x^2 + y^2)\frac{\partial}{\partial x}(x^3 - 4xy - y^2)}{(x^3 - 4xy - y^2)^2},$$

and the partial derivative with respect to y is

$$j_y(x, y) = \frac{(x^3 - 4xy - y^2)\frac{\partial}{\partial y}(x^2 + y^2) - (x^2 + y^2)\frac{\partial}{\partial y}(x^3 - 4xy - y^2)}{(x^3 - 4xy - y^2)^2}.$$

All we have to do to find these partial derivatives is calculate the smaller derivatives inside these expressions. These derivatives are

$$\frac{\partial}{\partial x}(x^2 + y^2) = 2x \quad , \quad \frac{\partial}{\partial x}(x^3 - 4xy - y^2) = 3x^2 - 4y,$$

$$\frac{\partial}{\partial y}(x^2+y^2) = 2y, \quad \text{and} \quad \frac{\partial}{\partial y}(x^3-4xy-y^2) = 2y-4x.$$

Substituting into the overall partial derivatives, the gradient is

$$\nabla j(x,y) = \begin{bmatrix} \frac{2x(x^3-4xy-y^2)-(3x^2-4y)(x^2+y^2)}{(x^3-4xy-y^2)^2} \\ \\ \frac{2y(x^3-4xy-y^2)-(2y-4x)(x^2+y^2)}{(x^3-4xy-y^2)^2} \end{bmatrix}.$$

(e) $k(x,y,z) = -4x^5y^3z^2 - 3y^2z^4 + 7xz^3 + 10y - 9z + 9$

Since there are three independent variables, we have to find three partial derivatives: one with respect to x, y, and z.

To find the partial derivative with respect to x, we treat y and z as constant. Then finding the derivative is straightforward:

$$k_x(x,y,z) = \frac{\partial}{\partial x}\Big(-4x^5y^3z^2 - 3y^2z^4 + 7xz^3 + 10y - 9z + 9\Big) = -20x^4y^3z^2 + 7z^3.$$

Similarly, the partial derivative with respect to y is

$$k_y(x,y,z) = \frac{\partial}{\partial y}\Big(-4x^5y^3z^2 - 3y^2z^4 + 7xz^3 + 10y - 9z + 9\Big) = -12x^5y^2z^2 - 6yz^4 + 10,$$

and the partial derivative with respect to z is

$$k_z(x,y,z) = \frac{\partial}{\partial z}\Big(-4x^5y^3z^2 - 3y^2z^4 + 7xz^3 + 10y - 9z + 9\Big) = -8x^5y^3z - 12y^2z^3 + 21xz^2 - 9.$$

The gradient is

$$\nabla k(x,y,z) = \begin{bmatrix} -20x^4y^3z^2 + 7z^3 \\ \\ -12x^5y^2z^2 - 6yz^4 + 10 \\ \\ -8x^5y^3z - 12y^2z^3 + 21xz^2 - 9 \end{bmatrix}.$$

(f) $l(x,y,z) = \dfrac{x^2-y^2+z^2}{\ln(x)}$

The partial derivative with respect to x requires the quotient rule since x appears in the numerator and the denominator:

$$\begin{aligned} l_x(x,y,z) = \frac{\partial}{\partial x}\left(\frac{x^2-y^2+z^2}{\ln(x)}\right) &= \frac{\ln(x)\frac{\partial}{\partial x}(x^2-y^2+z^2) - (x^2-y^2+z^2)\frac{\partial}{\partial x}(\ln(x))}{\ln(x)^2} \\ &= \frac{2x\ln(x) - \frac{x^2-y^2+z^2}{x}}{\ln(x)^2} \\ &= \frac{2x\ln(x)}{\ln(x)^2} - \frac{\frac{x^2-y^2+z^2}{x}}{\ln(x)^2} \end{aligned}$$

$$= \frac{2x}{\ln(x)} - \frac{x^2 - y^2 + z^2}{x\ln(x)^2}.$$

The partial derivatives of y and z *do not* require the quotient rule because neither y nor z appear in the denominator. We can treat the denominator as a constant factor that can be brought outside of these derivatives. The partial derivative with respect to y is

$$l_y(x, y, z) = \frac{\partial}{\partial y}\left(\frac{x^2 - y^2 + z^2}{\ln(x)}\right)$$
$$= \frac{1}{\ln(x)}\frac{\partial}{\partial y}(x^2 - y^2 + z^2)$$
$$= \frac{-2y}{\ln(x)},$$

and the partial derivative with respect to z is

$$l_z(x, y, z) = \frac{\partial}{\partial z}\left(\frac{x^2 - y^2 + z^2}{\ln(x)}\right)$$
$$= \frac{1}{\ln(x)}\frac{\partial}{\partial z}(x^2 - y^2 + z^2)$$
$$= \frac{2z}{\ln(x)}.$$

The gradient is

$$\nabla l(x, y, z) = \begin{bmatrix} \frac{2x}{\ln(x)} - \frac{x^2-y^2+z^2}{x\ln(x)^2} \\ \\ \frac{-2y}{\ln(x)} \\ \\ \frac{2z}{\ln(x)} \end{bmatrix}.$$

5. A gradient is a vector of the first partial derivatives. The partial derivative of $f(x, y)$ with respect to x is

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\Big((x^2 - y^2)\ln(x + y)\Big)$$
$$= (x^2 - y^2)\frac{\partial}{\partial x}\ln(x + y) + \ln(x + y)\frac{\partial}{\partial x}(x^2 - y^2)$$
$$= \frac{x^2 - y^2}{x + y} + 2x\ln(x + y)$$
$$= \frac{(x + y)(x - y)}{x + y} + 2x\ln(x + y)$$
$$= x - y + 2x\ln(x + y).$$

The partial derivative of $f(x, y)$ with respect to y is

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\Big((x^2 - y^2)\ln(x + y)\Big)$$
$$= (x^2 - y^2)\frac{\partial}{\partial y}\ln(x + y) + \ln(x + y)\frac{\partial}{\partial y}(x^2 - y^2)$$
$$= \frac{x^2 - y^2}{x + y} - 2y\ln(x + y)$$

$$
\begin{aligned}
&= \frac{(x+y)(x-y)}{x+y} - 2y\ln(x+y) \\
&= x - y - 2y\ln(x+y).
\end{aligned}
$$

Therefore the gradient of $f(x, y)$ is

$$
\nabla f(x,y) = \begin{bmatrix} x - y + 2x\ln(x+y) \\ \\ x - y - 2y\ln(x+y) \end{bmatrix}.
$$

The Hessian is the matrix of second partial derivatives. We have to calculate the partial derivative with respect to x and x again, the partial derivative with respect to y and y again, and the partial derivative with respect to x and then y which must be equal to the partial derivative with respect to y then x. First, we calculate the partial derivative with respect to x and x again:

$$
\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial^2}{\partial x^2}\Big((x^2 - y^2)\ln(x+y)\Big) \\
&= \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\Big((x^2 - y^2)\ln(x+y)\Big)\right) \\
&= \frac{\partial}{\partial x}\Big(x - y + 2x\ln(x+y)\Big) \\
&= 1 + 2\ln(x+y) + \frac{2x}{x+y} \\
&= \frac{-(x+y)}{x+y} + 2\ln(x+y) + \frac{2x}{x+y} \\
&= \frac{-x-y+2x}{x+y} - +2\ln(x+y) \\
&= \frac{x-y}{x+y} + 2\ln(x+y).
\end{aligned}
$$

Next, we calculate the partial derivative with respect to y and y again:

$$
\begin{aligned}
\frac{\partial^2 f}{\partial y^2} &= \frac{\partial^2}{\partial y^2}\Big((x^2 - y^2)\ln(x+y)\Big) \\
&= \frac{\partial}{\partial y}\left(\frac{\partial}{\partial y}\Big((x^2 - y^2)\ln(x+y)\Big)\right) \\
&= \frac{\partial}{\partial y}\Big(x - y - 2y\ln(x+y)\Big) \\
&= -1 - 2\ln(x+y) - \frac{2y}{x+y} \\
&= \frac{-(x+y)}{x+y} - 2\ln(x+y) + \frac{2x}{x+y} \\
&= \frac{-x-y+2x}{x+y} - 2\ln(x+y) \\
&= \frac{x-y}{x+y} - 2\ln(x+y).
\end{aligned}
$$

Finally, the partial derivative with respect to x then y is

$$
\begin{aligned}
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial^2}{\partial x \partial y}\Big((x^2 - y^2)\ln(x+y)\Big) \\
&= \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}\Big((x^2 - y^2)\ln(x+y)\Big)\right)
\end{aligned}
$$

$$= \frac{\partial}{\partial y}\Big(x - y + 2x\ln(x+y)\Big)$$
$$= -1 + \frac{2x}{x+y}$$
$$= \frac{-(x+y)}{x+y} + \frac{2x}{x+y}$$
$$= \frac{-x-y+2x}{x+y}$$
$$= \frac{x-y}{x+y}.$$

So the Hessian of $f(x,y)$ is

$$H\Big(f(x,y)\Big) = \begin{bmatrix} \frac{x-y}{x+y} + 2\ln(x+y) & \frac{x-y}{x+y} \\ \frac{x-y}{x+y} & \frac{x-y}{x+y} - 2\ln(x+y) \end{bmatrix}.$$

6. (a) The first partial derivative of the function with respect to x is

$$\frac{\partial f(x,y)}{\partial x} = \frac{\partial}{\partial x}\Big(-x^2 + xy - y^2 + 2x + y\Big) = -2x + y + 2.$$

The first partial derivative with respect to y is

$$\frac{\partial f(x,y)}{\partial y} = \frac{\partial}{\partial y}\Big(-x^2 + xy - y^2 + 2x + y\Big) = x - 2y + 1.$$

So the gradient of $f(x,y)$ is

$$\nabla f(x,y) = \begin{bmatrix} -2x + y + 2 \\ x - 2y + 1 \end{bmatrix}.$$

(b) We now have to solve the following system of equations:

$$\begin{cases} -2x + y + 2 = 0, \\ x - 2y + 1 = 0. \end{cases}$$

We can solve this system by solving the top equation for y,

$$y = 2x - 2,$$

and plugging in for y in the second equation:

$$x - 2(2x-2) + 1 = 0,$$
$$x - 4x + 4 + 1 = 0,$$
$$-3x + 5 = 0,$$
$$3x = 5,$$

$$x = \frac{5}{3}.$$

Then plugging the solution for x back into the first equation solved for y, we get

$$y = 2(5/3) - 2 = 10/3 - 6/3 = \frac{4}{3}.$$

So the ordered pair $(x, y) = \left(\frac{5}{3}, \frac{4}{3}\right)$ is a critical point for the function. Since we obtained only one possible value of x, and this value implied only one value of y, that means that $(x, y) = \left(\frac{5}{3}, \frac{4}{3}\right)$ is the only critical point.

(c) First, the second partial derivative with respect to x and x again is

$$\frac{\partial f(x,y)}{\partial x} = \frac{\partial}{\partial x}\left(-2x + y + 2\right) = -2$$

The second partial derivative with respect to y and y again is

$$\frac{\partial f(x,y)}{\partial y} = \frac{\partial}{\partial y}\left(x - 2y + 1\right) = -2.$$

Finally, the second partial derivative with respect to x then y (or y then x) is

$$\frac{\partial f(x,y)}{\partial y} = \frac{\partial}{\partial y}\left(-2x + y + 2\right) = 1.$$

So the Hessian is

$$H\left(f(x,y)\right) = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}.$$

(d) First we check whether

$$\begin{gathered} f_{xx}(-1,-4)f_{yy}(-1,-4) - f_{xy}(-1,-4)^2 > 0, \\ (-2)(-2) - (1)^2 > 0, \\ 4 - 1 > 0, \\ 3 > 0, \end{gathered}$$

so the first condition is met. Next we check the sign of

$$f_{xx}(-1,-4) + f_{yy}(-1,-4) = -2 + -2 = -4 < 0.$$

So the critical point represents a local maximum.

7. The problem is to maximize the function

$$f(x,y) = 150x^{1/3}y^{2/3}$$

subject to the constraint that

$$300x + 500y = 100000.$$

In order to apply the method of Lagrange multipliers, we first find the gradient of $f(x, y)$. The partial derivative with respect to x is

$$\begin{aligned} f_x(x,y) &= \frac{\partial}{\partial x}(150x^{1/3}y^{2/3}) \\ &= 150y^{2/3}\frac{\partial}{\partial x}(x^{1/3}) \\ &= 150y^{2/3}\left(\frac{1}{3}x^{-2/3}\right) \\ &= \frac{50y^{2/3}}{x^{2/3}}. \end{aligned}$$

The partial derivative with respect to y is

$$\begin{aligned} f_y(x,y) &= \frac{\partial}{\partial y}(150x^{1/3}y^{2/3}) \\ &= 150x^{1/3}\frac{\partial}{\partial y}(y^{2/3}) \\ &= 150x^{1/3}\left(\frac{2}{3}y^{-1/3}\right) \\ &= \frac{100x^{1/3}}{y^{1/3}}. \end{aligned}$$

The gradient is

$$\nabla f(x,y) = \begin{bmatrix} \frac{50y^{2/3}}{x^{2/3}} \\ \\ \frac{100x^{1/3}}{y^{1/3}} \end{bmatrix}.$$

Note that there are no critical points for this function. The point (0,0) does not make the gradient's elements equal 0 because it places 0 in the denominator of each partial derivative.

Next, we write the constraint function, replacing the constant 100,000 with a general function name $g(x, y)$,

$$g(x,y) = 300x + 500y.$$

The partial derivative of $g(x, y)$ with respect to x is 300, and the partial derivative of $g(x, y)$ with respect to y is 500. So the gradient of $g(x, y)$ is

$$\nabla g(x,y) = \begin{bmatrix} 300 \\ 500 \end{bmatrix}.$$

Next we create a system of three equations by setting $f(x, y) = \lambda g(x, y)$, and including the constraint as well:

$$\begin{cases} \frac{50y^{2/3}}{x^{2/3}} = 300\lambda, \\ \\ \frac{100x^{1/3}}{y^{1/3}} = 500\lambda, \\ \\ 300x + 500y = 100000. \end{cases}$$

To solve this system, let's first solve the first two equations for λ then set them equal to each other. Dividing both sides by 300, the first equation can be rewritten as

$$\frac{y^{2/3}}{6x^{2/3}} = \lambda,$$

and dividing both sides by 500, the second equation can be written as

$$\frac{x^{1/3}}{5y^{1/3}} = \lambda.$$

These two equations imply that

$$\begin{aligned}\frac{y^{2/3}}{6x^{2/3}} &= \frac{x^{1/3}}{5y^{1/3}},\\ (y^{2/3})(5y^{1/3}) &= (x^{1/3})(6x^{2/3}),\\ 5y &= 6x,\\ y &= \frac{6}{5}x.\end{aligned}$$

Finally we substitute for y in the last equation:

$$\begin{aligned}300x + 500\left(\frac{6}{5}x\right) &= 100000,\\ 300x + 600x &= 100000,\\ 900x &= 100000,\\ x = \frac{100000}{900} &= 111.11.\end{aligned}$$

Given this value of x, we can calculate y:

$$y = \frac{6}{5}(111.11) = 133.33.$$

We have no critical points to compare the point (111.11, 133.33) to, and it is our only optimum on the line $300x + 500y = 100000$. The value of the Cobb-Douglas production function at this point is

$$f(111.11, 133.33) = 150(111.11)^{1/3}(133.33)^{2/3} = 18820.34.$$

Any other point on this line produces a lower value of $f(x, y)$ than $f(111.11, 133.33)$. Consider, for example, the case in which $x = 50$. According to the constraint, the value of y must be

$$\begin{aligned}300(50) + 500y &= 100000,\\ 15000 + 500y &= 100000,\\ 500y &= 85000,\\ y &= 170.\end{aligned}$$

The value of the Cobb-Douglas production function at the point (50,170) is

$$f(50, 170) = 150(50)^{1/3}(170)^{2/3} = 16958.23.$$

Therefore the firm maximizes its profits by spending its resources on 111.11 units of labor and 133.33 units of capital.

8. (a) The problem asks us to solve the definite double integral

$$\int_0^3 \int_1^5 3x^2 - 3y^2 - 2xy \; dy \; dx.$$

Because neither set of bounds depends on a variable, these bounds can be considered in either order. This integral is equal to the one in which the order of x and y are reversed:

$$\int_1^5 \int_0^3 3x^2 - 3y^2 - 2xy\ dx\ dy.$$

To solve this double integral, we start by solving the innermost integral:

$$\int_0^3 \left(\int_1^5 3x^2 - 3y^2 - 2xy\ dy \right) dx$$

$$= \int_0^3 \left(3x^2y - y^3 - xy^2 \Big|_1^5 \right) dx$$

$$= \int_0^3 \left([3x^2(5) - (5)^3 - x(5)^2] - [3x^2(1) - (1)^3 - x(1)^2] \right) dx$$

$$= \int_0^3 \left([15x^2 - 125 - 25x] - [3x^2 - 1 - x] \right) dx$$

$$= \int_0^3 12x^2 - 124 - 24x\ dx.$$

We now have a single definite integral to solve:

$$\int_0^3 12x^2 - 124 - 24x\ dx = 4x^3 - 124x - 12x^2 \Big|_0^3$$

$$= 4(3)^3 - 124(3) - 12(3)^2 - (4(0)^3 - 124(0) - 12(0)^2) = -372.$$

(b) The problem asks us to solve

$$\int_2^4 \int_1^{e^x} \frac{x}{y}\ dy\ dx.$$

Note that, unlike part (a), here we have no choice in the order of the bounds. We must work with y before x because the bounds of y themselves contain a function of x. We solve the inner-integral first:

$$\int_2^4 \left(\int_1^{e^x} \frac{x}{y}\ dy \right) dx$$

In that inner-integral, we treat y as the variable and x as a constant. Since x is a constant factor, we can bring it outside that integral:

$$\int_2^4 x \left(\int_1^{e^x} \frac{1}{y}\ dy \right) dx$$

$$= \int_2^4 x \left(\ln(y) \Big|_1^{e^x} \right) dx^{1}$$

$$= \int_2^4 x \left(\ln(e^x) - \ln(1) \right) dx$$

$$= \int_2^4 x \left(x - 0 \right) dx$$

$$= \int_2^4 x^2\ dx = \frac{x^3}{3} \Big|_2^4$$

$$\frac{4^3}{3} - \frac{2^3}{3} = \frac{64 - 8}{3} = \frac{56}{3}.$$

[1] Technically, the anti-derivative of $\frac{1}{y}$ is $\ln(|y|)$. Here we can omit the absolute value because all of the numbers within the range of the inner-integral are positive.

(c) The problem asks us to solve

$$\int_0^1 \int_0^{3x^2} \sqrt{x^2+y}\ dy\ dx.$$

Note that, like part (b), we must work with y before x because the bounds of y themselves contain a function of x. We solve the inner-integral first:

$$\int_0^1 \left(\int_0^{3x^2} \sqrt{x^2+y}\ dy \right)\ dx.$$

$$= \int_0^1 \left(\int_0^{3x^2} (x^2+y)^{1/2}\ dy \right)\ dx.$$

We need to employ u-substitution to solve the inner-integral. Let $u = x^2 + y$. Then $\frac{du}{dy} = 1$, and so $du = dy$. We also change the bounds:

$$u(0) = x^2 + 0 = x^2, \quad \text{and} \quad u(x^2) = x^2 + (3x^2) = 4x^2.$$

Substituting, the integral becomes:

$$\int_0^1 \left(\int_{x^2}^{4x^2} u^{1/2}\ du \right)\ dx$$

$$= \int_0^1 \left(\frac{2}{3} u^{3/2} \Big|_{x^2}^{4x^2} \right)\ dx$$

$$= \frac{2}{3} \int_0^1 \left((4x^2)^{3/2} - (x^2)^{3/2} \right)\ dx$$

$$= \frac{2}{3} \int_0^1 \left((4^{3/2}(x^2)^{3/2} - (x^2)^{3/2} \right)\ dx$$

$$= \frac{2}{3} \int_0^1 \left((4^{3/2} - 1)(x^2)^{3/2} \right)\ dx$$

$$= \frac{2}{3} \int_0^1 7x^3\ dx$$

$$= \frac{14}{3} \int_0^1 x^3\ dx = \frac{14}{3} \left(\frac{x^4}{4} \Big|_0^1 \right)$$

$$= \frac{14}{3} \left(\frac{1}{4} - \frac{0}{4} \right) = \frac{7}{6}.$$

(d) The problem asks us to solve

$$\int_1^{e^4} \int_1^{e^5} \frac{\ln(x) + \ln(y)}{xy}\ dx\ dy.$$

Like part (a), neither set of bounds depends on the other variable, so we can consider these bounds in any order. Let's consider the bounds of x first. Note that the y in the denominator can be considered to be a constant factor that can be brought outside of this integral:

$$\int_1^{e^4} \frac{1}{y} \left(\int_1^{e^5} \frac{\ln(x) + \ln(y)}{x}\ dx \right)\ dy.$$

We need to use u-substitution to solve the inner-integral. Let $u = \ln(x) + \ln(y)$. Remember that in this inner-integral, we are treating x as the variable and y as constant, so we take the partial derivative of u with respect to x,

$$\frac{du}{dx} = \frac{1}{x}, \qquad du = \frac{1}{x}\, dx.$$

Notice that the factor $\frac{1}{x}$ will cancel out of the integrand. We also change the bounds:

$$u(1) = \ln(1) + \ln(y) = \ln(y), \quad \text{and} \quad u(e^5) = \ln(e^5) + \ln(y) = 5 + \ln(y).$$

Substituting, the integral becomes

$$\begin{aligned}
&\int_1^{e^4} \frac{1}{y}\left(\int_{\ln(y)}^{5+\ln(y)} u\, du\right) dy \\
&= \int_1^{e^4} \frac{1}{y}\left(\frac{u^2}{2}\Big|_{\ln(y)}^{5+\ln(y)}\right) dy \\
&= \int_1^{e^4} \frac{1}{2y}\Big((5+\ln(y))^2 - \ln(y)^2\Big)\, dy \\
&= \int_1^{e^4} \frac{1}{2y}\Big(25 + 10\ln(y) + \ln(y)^2 - \ln(y)^2\Big)\, dy \\
&= \int_1^{e^4} \frac{1}{2y}\Big(25 + 10\ln(y)\Big)\, dy \\
&= \frac{1}{2}\int_1^{e^4} \frac{25 + 10\ln(y)}{y}\, dy.
\end{aligned}$$

Now we have to employ u-substitution again to solve this remaining integral, which treats y as the variable. Let $u = 25 + 10\ln(y)$. Then

$$\frac{du}{dy} = \frac{10}{y}, \qquad \frac{1}{10}du = \frac{1}{y}\, dy.$$

Now the factor $\frac{1}{y}$ cancels out of the integrand. We replace the bounds:

$$u(1) = 25 + \ln(1) = 25, \quad \text{and} \quad u(e^4) = 25 + \ln(e^4) = 29.$$

Substituting, the integral becomes

$$\begin{aligned}
&\frac{1}{2}\int_{25}^{29} u\left(\frac{1}{10}du\right) \\
&= \frac{1}{20}\frac{u^2}{2}\Big|_{25}^{29} \\
&= \frac{29^2 - 25^2}{40} = 5.4.
\end{aligned}$$

9. (a) In order to demonstrate that a multivariate function is a joint PDF, we have to show that the function is nonnegative for all ordered pairs in the domain, and that the multiple integral over the domain evaluates to 1. First, note that the domain only contains nonnegative values of x and y, so

$$f(x,y) = \frac{6}{7}\left(x^2 + \frac{xy}{2}\right)$$

must also be nonnegative since it contains no subtraction or multiplication by negative numbers. Next we have to demonstrate that the integral

$$\int_0^1 \int_0^2 \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \, dy \, dx = 1.$$

The integral is solved using the following steps:

$$= \frac{6}{7} \int_0^1 \int_0^2 x^2 + \frac{xy}{2} \, dy \, dx$$

$$= \frac{6}{7} \int_0^1 \left(\int_0^2 x^2 + \frac{xy}{2} \, dy \right) dx$$

$$= \frac{6}{7} \int_0^1 \left(x^2 y + \frac{xy^2}{4} \Big|_0^2 \right) dx$$

$$= \frac{6}{7} \int_0^1 \left(2x^2 + \frac{4x}{4} \right) - \left(0x^2 + \frac{0x}{4} \right) dx$$

$$= \frac{6}{7} \int_0^1 2x^2 + x \, dx$$

$$= \frac{6}{7} \left(\frac{2}{3}x^3 + \frac{x^2}{2} \Big|_0^1 \right)$$

$$= \frac{6}{7} \left(\left(\frac{2}{3}1^3 + \frac{1^2}{2} \right) - \left(\frac{4}{3}0^3 + \frac{0^2}{2} \right) \right)$$

$$= \frac{6}{7} \left(\frac{2}{3} + \frac{1}{2} \right)$$

$$= \frac{6}{7} \left(\frac{4}{6} + \frac{3}{6} \right)$$

$$= \frac{6}{7} \left(\frac{7}{6} \right) = 1.$$

So the function is indeed a joint PDF.

(b) Now that we've demonstrated that $f(x, y)$ is a joint PDF, we can calculate the probability that x and y fall within certain bounds by plugging these new bounds into the multivariate definite integral and finding the volume under the curve. In this case, the problem is to solve

$$P(0 < x < 0.5, 0 < y < x) = \int_0^{.5} \int_0^{x} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \, dy \, dx.$$

The integral is solved using the following steps:

$$= \frac{6}{7} \int_0^{.5} \int_0^{x} x^2 + \frac{xy}{2} \, dy \, dx$$

$$= \frac{6}{7} \int_0^{.5} \left(\int_0^{x} x^2 + \frac{xy}{2} \, dy \right) dx$$

$$= \frac{6}{7} \int_0^{.5} \left(x^2 y + \frac{xy^2}{4} \Big|_0^{x} \right) dx$$

$$= \frac{6}{7}\int_0^{.5} \left(x^3 + \frac{x^3}{4}\right) - \left(0x^2 + \frac{0x}{4}\right)\, dx$$

$$= \frac{6}{7}\int_0^{.5} \left(\frac{5}{4}x^3\right)\, dx$$

$$= \frac{15}{14}\int_0^{.5} x^3\, dx$$

$$= \frac{15}{14}\frac{x^4}{4}\Big|_0^{.5}$$

$$= \frac{15}{14}\left(\frac{.5^4}{4} - \frac{0^4}{4}\right)$$

$$= \frac{15}{14} \times \frac{1}{64} = \frac{15}{896} = 0.017.$$

(c) To find the marginal distribution of x, we integrate the joint PDF over the domain of y:

$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

$$= \int_0^2 \frac{6}{7}\left(x^2 + \frac{xy}{2}\right)\, dy$$

$$= \frac{6}{7}\int_0^2 x^2 + \frac{xy}{2}\, dy$$

$$= \frac{6}{7}\left(x^2y + \frac{xy^2}{4}\Big|_0^2\right)$$

$$= \frac{6}{7}\left(x^2(2) + \frac{x(2)^2}{4} - x^2(0) - \frac{x(0)^2}{4}\Big|_0^2\right)$$

$$= \frac{6}{7}\left(2x^2 + x\right)$$

$$= \frac{12x^2 + 6x}{7}.$$

(d) We repeat the calculations we conducted for part (c), but this time we integrate the joint PDF over the domain of x:

$$f_y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$$

$$= \int_0^1 \frac{6}{7}\left(x^2 + \frac{xy}{2}\right)\, dx$$

$$= \frac{6}{7}\left(\frac{x^3}{3} + \frac{x^2y}{4}\Big|_0^1\right)$$

$$= \frac{6}{7}\left(\frac{(1)^3}{3} + \frac{(1)^2y}{4} - \frac{(0)^3}{3} - \frac{(0)^2y}{4}\right)$$

$$= \frac{6}{7}\left(\frac{1}{3} + \frac{y}{4}\right)$$
$$= \frac{4+3y}{14}.$$

(e) x and y are independent if and only if the product of their marginal distributions equals the joint PDF. The product of the marginal distributions is

$$f_x(x)f_y(y)$$
$$= \left(\frac{12x^2+6x}{7}\right)\left(\frac{4+3y}{14}\right)$$
$$= \frac{(12x^2+6x)(4+3y)}{98}$$
$$= \frac{48x^2+24x+36x^2y+18xy}{98},$$

which does not equal the joint PDF $\frac{6}{7}\left(x^2 + \frac{xy}{2}\right)$. To see this explicitly, consider the point (1,0) which is in the domain of the joint PDF. At this point, the product of the marginal distributions is

$$f_x(1)f_y(0) = \frac{48(1)^2+24(1)+36(1)^2(0)+18(1)(0)}{98} = \frac{48+24}{98} = 0.735.$$

And the joint PDF at (1,0) is

$$f(1,0) = \frac{6}{7}\left((1)^2 + \frac{(1)(0)}{2}\right) = \frac{6}{7} = 0.857.$$

Since there are instances in which the product of marginal distributions does not equal the joint PDF, these variables are not independent.

(f) The conditional distribution of x given y is given by the quotient of the joint PDF and the marginal distribution of y:

$$f_{x|y}(x,y) = \frac{f(x,y)}{f_y(x,y)}.$$

Since we know the joint PDF, and we've already derived the marginal distribution of y in part (d), we do not need to take any more integrals – we simply plug these two known functions into the above formula:

$$f_{x|y}(x,y) = \frac{\frac{6}{7}\left(x^2 + \frac{xy}{2}\right)}{\frac{4+3y}{14}}$$
$$= \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \times \frac{14}{4+3y}$$
$$= \frac{12\left(x^2 + \frac{xy}{2}\right)}{4+3y}$$
$$= \frac{12x^2+6xy}{4+3y}.$$

(g) The conditional distribution of y given x is given by the quotient of the joint PDF and the marginal distribution of x:

$$f_{y|x}(x,y) = \frac{f(x,y)}{f_x(x,y)}.$$

Since we know the joint PDF, and we've already derived the marginal distribution of x in part (c), we do not need to take any more integrals – we simply plug these two known functions into the above formula:

$$\begin{aligned}
f_{y|x}(x,y) &= \frac{\frac{6}{7}\left(x^2 + \frac{xy}{2}\right)}{\frac{12x^2+6x}{7}} \\
&= \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \times \frac{7}{12x^2+6x} \\
&= \frac{6\left(x^2 + \frac{xy}{2}\right)}{12x^2+6x} \\
&= \frac{x^2 + \frac{xy}{2}}{2x^2+x} \\
&= \frac{2x^2+xy}{4x^2+2x} \\
&= \frac{2x+y}{4x+2}.
\end{aligned}$$

(h) The expected value of x is given by the following formula,

$$E(x) = \int_{-\infty}^{\infty} x\, f_x(x)\ dx,$$

where the infinite bounds stand in for whatever the domain of x happens to be, and $f_x(x)$ is the marginal distribution of x. We found the marginal distribution of x in part (c), so in this case the expected value is

$$\begin{aligned}
E(x) &= \int_0^1 x\left(\frac{12x^2+6x}{7}\right)\ dx \\
&= \frac{6}{7}\int_0^1 2x^3 + x^2\ dx \\
&= \frac{6}{7}\left(\frac{x^4}{2} + \frac{x^3}{3}\Big|_0^1\right) \\
&= \frac{6}{7}\left(\frac{1^4}{2} + \frac{1^3}{3} - \frac{0^4}{2} - \frac{0^3}{3}\right) \\
&= \frac{6}{7}\left(\frac{1}{2} + \frac{1}{3}\right) \\
&= \frac{6}{7} \times \frac{5}{6} = \frac{5}{7} = .714.
\end{aligned}$$

(i) The expected value of y is given by

$$E(y) = \int_{-\infty}^{\infty} y\, f_y(y)\ dy,$$

where the infinite bounds stand in for whatever the domain of y happens to be, and $f_y(y)$ is the marginal distribution of y. We found the marginal distribution of y in part (d), so in this case the expected value is

$$\begin{aligned}
E(y) &= \int_0^2 y\left(\frac{4+3y}{14}\right)\ dy \\
&= \frac{1}{14}\int_0^2 4y + 3y^2\ dy \\
&= \frac{1}{14}\left(2y^2 + y^3\Big|_0^2\right) \\
&= \frac{1}{14}\left(2(2)^2 + (2)^3 - 2(0)^2 - (0)^3\right) \\
&= \frac{1}{14}(8+8) = \frac{16}{14} = \frac{8}{7} = 1.14.
\end{aligned}$$

(j) The variance of x is given by

$$V(x) = \int_{-\infty}^{\infty} (x-c)^2\, f_x(x)\ dx,$$

where $c = E(y)$. We found the marginal distribution of x in part (c) and the expected value of x in part (h), so in this case the variance is

$$V(x) = \int_0^1 (x-c)^2\left(\frac{12x^2+6x}{7}\right)\ dx.$$

To keep the calculation neater, let's leave the expected value as c for now, and we will plug in .714 at the end. The calculation proceeds as follows:

$$\begin{aligned}
V(x) &= \frac{6}{7}\int_0^1 (x^2 - 2cx + c^2)(2x^2 + x)\ dx \\
&= \frac{6}{7}\int_0^1 2x^4 - 4cx^3 + 2c^2x^2 + x^3 - 2cx^2 + c^2x\ dx \\
&= \frac{6}{7}\int_0^1 2x^4 - (4c-1)x^3 + 2c(c-1)x^2 + c^2x\ dx \\
&= \frac{6}{7}\left(\frac{2}{5}x^5 - \frac{4c-1}{4}x^4 + \frac{2c(c-1)}{3}x^3 + \frac{c^2}{2}x^2\Big|_0^1\right) \\
&= \frac{6}{7}\left(\frac{2}{5} - \frac{4c-1}{4} + \frac{2c(c-1)}{3} + \frac{c^2}{2}\right) \\
&= \frac{6}{7}\left(\frac{2}{5} - \frac{4(.714)-1}{4} + \frac{2(.714)(.714-1)}{3} + \frac{.714^2}{2}\right) = 0.047.
\end{aligned}$$

Finally, the standard deviation of x is simply the square root of the variance:

$$SD(x) = \sqrt{V(x)} = \sqrt{0.047} = 0.217.$$

(k) The variance of y is given by

$$V(y) = \int_{-\infty}^{\infty} (y-d)^2 \, f_y(y) \; dy,$$

where $d = E(y)$. We found the marginal distribution of y in part (d) and the expected value of y in part (i), so in this case the variance is

$$V(y) = \int_0^2 (y-d)^2 \left(\frac{4+3y}{14}\right) \; dy.$$

Again, to keep the calculation neater, let's leave the expected value as d for now, and we will plug in 1.14 at the end. The calculation proceeds as follows:

$$\begin{aligned}
V(y) &= \frac{1}{14}\int_0^2 (y^2 - 2dy + d^2)(4+3y) \; dy \\
&= \frac{1}{14}\int_0^2 4y^2 - 8dy + 4d^2 + 3y^3 - 6dy^2 + 3d^2y \; dy \\
&= \frac{1}{14}\int_0^2 3y^3 + (4-6d)y^2 + (3d^2 - 8d)y + 4d^2 \; dy \\
&= \frac{1}{14}\left(\frac{3}{4}y^4 + \frac{4-6d}{3}y^3 + \frac{3d^2-8d}{2}y^2 + 4d^2y\Big|_0^2\right) \\
&= \frac{1}{14}\left(\frac{3}{4}(2)^4 + \frac{4-6d}{3}(2)^3 + \frac{3d^2-8d}{2}(2)^2 + 4d^2(2)\right) \\
&= \frac{1}{14}\left(\frac{3}{4}(16) + \frac{4-6d}{3}(8) + \frac{3d^2-8d}{2}(4) + 8d^2\right) \\
&= \frac{1}{14}\left(12 + \frac{8(4-6d)}{3} + 2(3d^2 - 8d) + 8d^2\right) \\
&= \frac{1}{14}\left(12 + \frac{8[4-6(1.14)]}{3} + 2[3(1.14)^2 - 8(1.14)] + 8(1.14)^2\right) = 0.313.
\end{aligned}$$

Finally, the standard deviation of y is simply the square root of the variance:

$$SD(y) = \sqrt{V(y)} = \sqrt{0.313} = 0.559.$$

(l) The covariance between x and y is given by

$$\text{Cov}(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-c)(y-d) \, f(x,y) \; dy \; dx,$$

where $c = E(x)$ and $d = E(y)$, which we will plug in only at the end of the calculation.[2] Substituting the bounds of x and y and their joint distribution gives us

$$\begin{aligned}
\text{Cov}(x,y) &= \int_0^1\int_0^2 (x-c)(y-d)\left[\frac{6}{7}\left(x^2 + \frac{xy}{2}\right)\right] \; dy \; dx \\
&= \frac{6}{7}\int_0^1\int_0^2 (xy - dx - cy + cd)\left(x^2 + \frac{xy}{2}\right) \; dy \; dx
\end{aligned}$$

[2]Please forgive the notational confusion here: d is used to denote both the expected value of y and the integration differential. The term dx appears inside the integrand as x times the expected value of y, and also as always at the end of the integral. Please note that every instance of dx inside the integrand refers to the former, not the latter.

$$= \frac{6}{7}\int_0^1 \int_0^2 x^3y - dx^3 - cx^2y + cdx^2 + \frac{x^2y^2}{2} - \frac{dx^2y}{2} - \frac{cxy^2}{2} + \frac{cdxy}{2}\ dy\ dx$$

$$= \frac{6}{7}\int_0^1 \left(\frac{x^3y^2}{2} - dx^3y - \frac{cx^2y^2}{2} + cdx^2y + \frac{x^2y^3}{6} - \frac{dx^2y^2}{4} - \frac{cxy^3}{6} + \frac{cdxy^2}{4}\Big|_0^2 \right) dx$$

$$= \frac{6}{7}\int_0^1 \left(\frac{x^3(2)^2}{2} - dx^3(2) - \frac{cx^2(2)^2}{2} + cdx^2(2) + \frac{x^2(2)^3}{6} - \frac{dx^2(2)^2}{4} - \frac{cx(2)^3}{6} + \frac{cdx(2)^2}{4} \right) dx$$

$$= \frac{6}{7}\int_0^1 2x^3 - 2dx^3 - 2cx^2 + 2cdx^2 + \frac{4x^2}{3} - dx^2 - \frac{4cx}{3} + cdx\ dx$$

$$= \frac{6}{7}\left(\frac{x^4}{2} - \frac{dx^4}{2} - \frac{2cx^3}{3} + \frac{2cdx^3}{3} + \frac{4x^3}{9} - \frac{dx^3}{3} - \frac{2cx^2}{3} + \frac{cdx^2}{2}\Big|_0^1 \right)$$

$$= \frac{6}{7}\left(\frac{(1)^4}{2} - \frac{d(1)^4}{2} - \frac{2c(1)^3}{3} + \frac{2cd(1)^3}{3} + \frac{4(1)^3}{9} - \frac{d(1)^3}{3} - \frac{2c(1)^2}{3} + \frac{cd(1)^2}{2} \right)$$

$$= \frac{6}{7}\left(\frac{1}{2} - \frac{d}{2} - \frac{2c}{3} + \frac{2cd}{3} + \frac{4}{9} - \frac{d}{3} - \frac{2c}{3} + \frac{cd}{2} \right)$$

$$= \frac{6}{7}\left(\frac{1}{2} - \frac{1.14}{2} - \frac{2(.714)}{3} + \frac{2(.714)(1.14)}{3} + \frac{4}{9} - \frac{1.14}{3} - \frac{2(.714)}{3} + \frac{(.714)(1.14)}{2} \right)$$

$$= -0.007.$$

(m) After all that work to calculate the marginal distributions, the expected values, the variances and standard deviations, and the covariance, all we have left to do to calculate the correlation between x and y is divide the covariance we calculated in part (l) by the standard deviations we calculated in parts (j) and (k):

$$\text{Corr}(x, y) = \frac{\text{Cov}(x, y)}{SD(x)SD(y)}$$

$$= \frac{-0.007}{0.217 \times 0.559} = -0.056.$$

10. (a) The correlation between two random variables is the covariance of the two variables divided by the product of their standard derivations. The standard deviations are the square roots of the variances:

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{SD(X)SD(Y)}$$

$$= \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}.$$

We know that $\text{Cov}(X, Y) = 5$, $V(X) = 4$, and $V(Y) = 9$ from the given information in this problem. So we simply plug these numbers in:

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{5}{\sqrt{4}\sqrt{9}} = \frac{5}{2 \times 3} = \frac{5}{6} = .833.$$

(b) We know from the problem that $A = 3X - 3Y + 7$. The variance of A is

$$V(A) = V(3X - 3Y + 7).$$

We can apply the rules of the variance of a sum from section 6.7 to find the value of this variance. First, adding a constant to a random variable does not change the variance of that random variable. So the variance becomes

$$V(A) = V(3X - 3Y).$$

Next, the variance of a weighted sum of two random variables is given by

$$V(aX + bY) = a^2V(X) + b^2V(Y) - 2ab\text{Cov}(X, Y),$$

which in this case is

$$\begin{aligned} V(3X - 3Y) &= (3)^2V(X) + (-3)^2V(Y) - 2(3)(-3)\text{Cov}(X, Y), \\ &= 9(4) + 9(9) + 18(5) = 207. \end{aligned}$$

(c) The problem tells us that $B = 5 - 2X + Y$. The goal is to calculate

$$V(B) = V(5 - 2X + Y).$$

We again use the rules for variances listed in section 6.7. The constant addend does not change the variance, so we can remove it,

$$V(B) = V(-2X + Y),$$

and the remaining expression is a weighted sum of random variables, subject to the rule

$$V(aX + bY) = a^2V(X) + b^2V(Y) - 2ab\text{Cov}(X, Y),$$

which in this case is

$$\begin{aligned} V(-2X + Y) &= (-2)^2V(X) + (1)^2V(Y) - 2(-2)(1)\text{Cov}(X, Y) \\ &= 4(4) + 9 + 4(5) = 45. \end{aligned}$$

(d) This problem asks us to calculate

$$\text{Cov}(A, B) = \text{Cov}\Big(3X - 3Y + 7,\ 5 - 2X + Y\Big)$$

Here we can apply the rules of covariances from section 7.4.4. First, adding a constant to either term of the covariance does not change the covariance, so we can remove both constant addends:

$$\text{Cov}(A, B) = \text{Cov}\Big(3X - 3Y,\ -2X + Y\Big).$$

Next we apply the rule for handling covariances of weighted sums of random variables,

$$\text{Cov}(aX + bY,\ cW + dZ) = ac\text{Cov}(X, W) + ad\text{Cov}(X, Z) + bc\text{Cov}(Y, W) + bd\text{Cov}(Y, Z),$$

which in this case is

$$\text{Cov}\Big(3X - 3Y,\ -2X + Y\Big) = (3)(-2)\text{Cov}(X, X) + (3)(1)\text{Cov}(X, Y) + (-3)(-2)\text{Cov}(Y, X) + (-3)(1)\text{Cov}(Y, Y).$$

$$= -6\text{Cov}(X, X) + 3\text{Cov}(X, Y) + 6\text{Cov}(Y, X) - 3\text{Cov}(Y, Y).$$

We also know that changing the order of the terms does not change the covariance, so we can combine the two middle covariances:

$$-6\text{Cov}(X, X) + 3\text{Cov}(X, Y) + 6\text{Cov}(X, Y) - 3\text{Cov}(Y, Y)$$
$$= -6\text{Cov}(X, X) + 9\text{Cov}(X, Y) - 3\text{Cov}(Y, Y).$$

Finally, we know that the covariance of a variable with itself is the variance. So our covariance becomes

$$-6V(X) + 9\text{Cov}(X, Y) - 3V(Y)$$
$$= -6(4) + 9(5) - 3(9) = -6.$$

(e) Now that we know the variance of A, the variance of B, and the covariance between them, we can plug these quantities into the formula for the correlation:

$$\text{Corr}(A, B) = \frac{\text{Cov}(A, B)}{SD(A)SD(B)}$$
$$= \frac{\text{Cov}(A, B)}{\sqrt{V(A)}\sqrt{V(B)}}$$
$$= \frac{-6}{\sqrt{207}\sqrt{45}} = -0.06.$$

Notice that the two transformations we applied, to generate A and B from X and Y, completely wiped out most of the correlation between these two variables, even though X and Y are themselves highly correlated.

11. (a) The first partial derivative of $E(y_i)$ with respect to x_{i1} is

$$\frac{\partial E(y_i)}{\partial x_{i1}} = \frac{\partial}{\partial x_{i1}}\left(\alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 x_{i3} + \beta_6 x_{i4} + \beta_7 x_{i3} x_{i4}\right) = \beta_1.$$

(b) When the variable under consideration is assumed to have a strictly linear effect – no curvilinear terms or interactions – then the coefficient IS the first partial derivative. That means that the classic interpretation of a regression coefficient,
"a one-unit increase in x_{i1} is associated with a β_1 change in y_i, on average, after controlling for the other x variables in the model,"
is the interpretation of the first-partial derivative. To break this down further:

Classic regression interpretation	How it relates to partial derivatives
"A one-unit increase in x_{i1}"	This part comes from the fact that a derivative is a slope. When you calculate slope, you divide the rise over the run. The resulting slope is a change in y for a 1-unit change in x, simply because you can write any slope as a fraction of the slope over 1.
"is associated with"	"Associated" simply means we haven't taken steps to ensure this model is giving us true causation.
"a β_1 change in y_i,"	This also comes from the basic definition of a slope. Change in y over change in x.
"on average,"	This part refers to the fact that we are taking the derivative of the expected value of y_i instead of y_i itself.
"after controlling for the other x variables in the model."	This phrase refers to the fact that we are taking a *partial* derivative instead of a regular derivative. We are only considering the slope in one direction: the direction that refers to x_{i1}.

It is often useful to approach the interpretation of regression models by thinking about derivatives instead of coefficients.

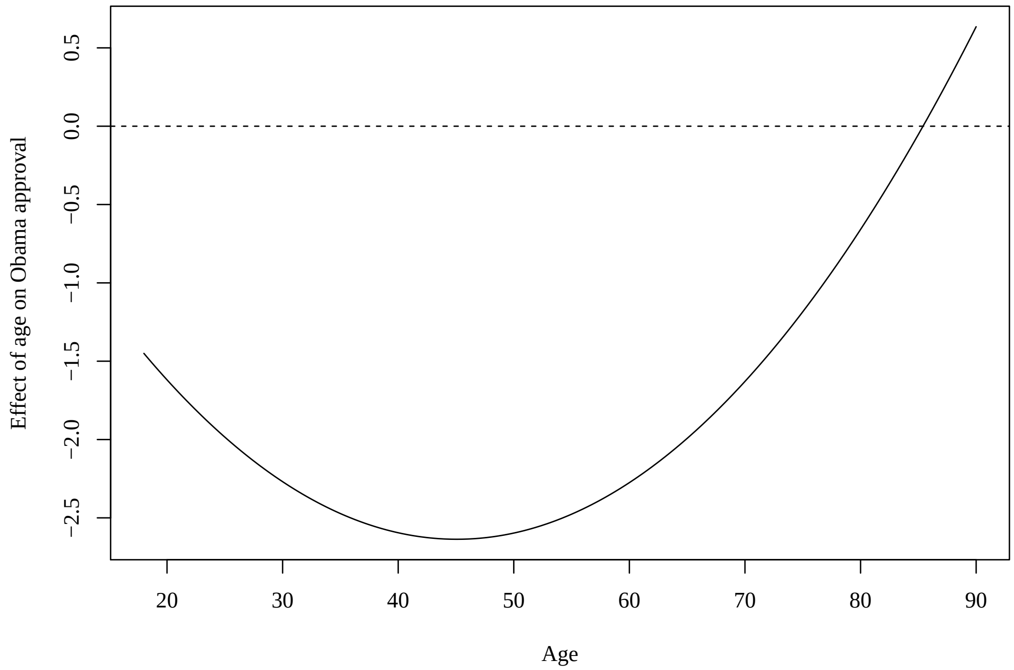
(c) The first partial derivative of $E(y_i)$ with respect to x_{i2} is

$$\frac{\partial E(y_i)}{\partial x_{i2}} = \frac{\partial}{\partial x_{i2}}\left(\alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 x_{i3} + \beta_6 x_{i4} + \beta_7 x_{i3} x_{i4}\right) = \beta_2 + 2\beta_3 x_{i2} + 3\beta_4 x_{i2}^2.$$

(d) If $\beta_2 = 0.653, \beta_3 = -0.073$, and $\beta_4 = 0.00054$, then we can plug these values into the partial derivative:

$$\begin{aligned}\frac{\partial E(y_i)}{\partial x_{i2}} &= 0.653 + 2(-0.073)x_{i2} + 3(0.00054)x_{i2}^2 \\ &= 0.00162x_{i2}^2 - 0.146x_{i2} + 0.653.\end{aligned}$$

Remember that x_{i2} is age, so it is reasonable to plot this function over a domain of x_{i2} from 18 to 90. The graph is:



The graph says that age generally has a negative effect on a person's approval of Obama: that is, older people approve of Obama less. But this effect is most highly pronounced for middle-aged people, with the largest negative effect estimated for people who are about 45 years old. There is a smaller negative effect of age for younger people and for older people. People over 85 years old actually have a positive effect, although this effect is unlikely to be significantly different than zero. This graph would support a theory that says that people's political preferences change most in their 40s. In order to really draw these inferences, however, we would need to graph the 95% confidence interval around this partial derivative for every x.

(e) The first partial derivative of $E(y_i)$ with respect to x_{i4} is

$$\frac{\partial E(y_i)}{\partial x_{i4}} = \frac{\partial}{\partial x_{i4}}\Big(\alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 x_{i3} + \beta_6 x_{i4} + \beta_7 x_{i3} x_{i4}\Big) = \beta_6 + \beta_7 x_{i3}.$$

(f) If we plug in $\beta_6 = 1$ and $\beta_7 = 3$ then this partial derivative becomes

$$\frac{\partial E(y_i)}{\partial x_{i4}} = 1 + 3x_{i3}.$$

Since x_{i3} is binary, we can break down this partial derivative into two cases. For men the partial derivative is $1 + 3(0) = 1$, and for women the partial derivative is $1 + 3(1) = 4$.

(g) That means that, in general, people who are more pro-choice are also more approving of Obama. For every unit more pro-choice a man is, he is on average 1 unit more favorable towards Obama. But for every unit more pro-choice a woman is, she is on average 4 units more favorable towards Obama. So this issue has a more dramatic effect on Obama approval for women than for men.

12. (a) We know from the problem that

$$\theta = \beta_1 x_1 + \beta_2 x_2 + \delta.$$

That means that the variance of θ is

$$V(\theta) = V(\beta_1 x_1 + \beta_2 x_2 + \delta),$$

where the we will treat the β coefficients as constants and the other terms as random variables. The formula for the variance of a weighted sum of three independent variables is

$$V(aX + bY + cZ) = a^2 V(X) + b^2 V(Y) + c^2 V(Z) + 2ab\text{Cov}(X, Y) + 2ac\text{Cov}(X, Z) + 2bc\text{Cov}(Y, Z).$$

We apply this formula to $V(\theta)$ to get

$$V(\beta_1 x_1 + \beta_2 x_2 + \delta) = \beta_1^2 V(x_1) + \beta_2^2 V(x_2) + V(\delta) + 2\beta_1\beta_2\text{Cov}(x_1, x_2) + 2\beta_1\text{Cov}(x_1, \delta) + 2\beta_2\text{Cov}(x_2, \delta).$$

We can directly measure $V(x_1)$, $V(x_2)$, and $\text{Cov}(x_1, x_2)$ from the data, and we assume that $\text{Cov}(x_1, \delta)$=0, $\text{Cov}(x_2, \delta)$=0, and $V(\delta) = 1$, so that this formula becomes

$$V(\beta_1 x_1 + \beta_2 x_2 + \delta) = \beta_1^2 V(x_1) + \beta_2^2 V(x_2) + 2\beta_1\beta_2\text{Cov}(x_1, x_2) + 1.$$

(b) We know that y_1 has the following regression equation:

$$y_1 = \delta_1\theta + \varepsilon_1.$$

To find the variance of y_1, we apply the formula for the variance of a weighted sum, again treating the coefficient λ as a constant:

$$\begin{aligned} V(y_1) &= V(\delta_1\theta + \varepsilon_1) \\ &= \delta_1^2 V(\theta) + V(\varepsilon_1) + 2\delta_1\text{Cov}(\theta, \varepsilon_1). \end{aligned}$$

We continue to assume that no error has a covariance with any other variable, so we can replace the covariance in this formula with zero,

$$V(y_1) = \delta_1^2 V(\theta) + V(\varepsilon_1),$$

we also assume that every error has a variance equal to 1,

$$V(y_1) = \delta_1^2 V(\theta) + 1.$$

We now replace $V(\theta)$ with the formula we derived in part (a):

$$V(y_1) = \delta_1^2 \left(\beta_1^2 V(x_1) + \beta_2^2 V(x_2) + 2\beta_1\beta_2\text{Cov}(x_1, x_2) + 1 \right) + 1.$$

(c) We can understand the covariance of x_1 and y_1 by substituting y_1's linear equation for y_1,

$$\mathrm{Cov}(x_1, y_1) = \mathrm{Cov}(x_1, \lambda_1\theta + \varepsilon_1),$$

and applying the rule of the covariance of a weighted sum,

$$\mathrm{Cov}(aX + bY,\, cW + dZ) = ac\mathrm{Cov}(X, W) + ad\mathrm{Cov}(X, Z) + bc\mathrm{Cov}(Y, W) + bd\mathrm{Cov}(Y, Z),$$

$$\mathrm{Cov}(x_1, \lambda_1\theta + \varepsilon_1) = \lambda_1\mathrm{Cov}(x_1, \theta) + \mathrm{Cov}(x_1, \varepsilon_1).$$

Once again, we assume that the errors are independent from all other variables, so that the covariance $\mathrm{Cov}(x_1, \varepsilon_1) = 0$, leaving us with

$$\lambda_1\mathrm{Cov}(x_1, \theta).$$

The problem now requires us to find this remaining covariance. First, we replace θ with its linear model,

$$\lambda_1\mathrm{Cov}(x_1, \beta_1 x_1 + \beta_2 x_2 + \delta),$$

and then we apply the rule regarding the covariances of weighted sums:

$$\lambda_1\mathrm{Cov}(x_1, \beta_1 x_1 + \beta_2 x_2 + \delta) = \lambda_1\Big(\beta_1\mathrm{Cov}(x_1, x_1) + \beta_2\mathrm{Cov}(x_1, x_2) + \mathrm{Cov}(x_1, \delta)\Big).$$

We can replace $\mathrm{Cov}(x_1, x_1)$ with $V(x_1)$, and as in part (a), we can directly measure $V(x_1)$, $V(x_2)$, and $\mathrm{Cov}(x_1, x_2)$ from the data, and we assume that $\mathrm{Cov}(x_1, \delta)$=0. The covariance is

$$\mathrm{Cov}(x_1, y_1) = \lambda_1\beta_1 V(x_1) + \lambda_1\beta_2\mathrm{Cov}(x_1, x_2).$$

The point of this exercise is to derive the formulas that underlie structural equation modeling. Most practitioners will consider these formulas to be too complex to try to understand. But now you know that all you need to derive these formulas is an understanding of covariance.

8 Matrix Notation and Arithmetic

1. (a) Let the 10 rows represent the 10 respondents, let the columns respectively represent each respondents' union membership, marital status, and religious service attendance. Let 1 indicate that the respondent belongs to a union or is married, and 0 represent that the respondent is not in a union or is not married. Let the last column be represented by the number of services per week the individual attends. The dataset can be written in a table as

Obs.	Union	Married	Religious Services
1	1	0	0
2	0	1	1
3	0	1	1
4	1	1	3
5	1	0	0
6	0	1	5
7	0	0	1
8	1	0	0
9	0	1	1
10	0	1	0

This table can be written as a (10×3) matrix X:

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \\ 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Note that it is no longer necessary to include labels for the rows and columns, as it is with the table. For the matrix, the row and column numbers describe the meaning of each particular datapoint.

(b) These data consist of the distance between each pair of cities within a group of these four: Washington, Boston, New York, and Chicago. Let's create a matrix with 4 rows and 4 columns in each the first row and first column each refer to Washington, the second row and column refer to Boston, the third row and column refer to New York, and the fourth row and column refer to Chicago. The elements of the matrix will be the distance between the city represented by the row and the city represented by the column. Note that there will be 0s on the diagonal because a city is 0 miles away from itself. Also note that the matrix will be symmetric – that is, every element will be equal to the element with the row and column numbers switched – because the distance from New York to Chicago is equal to the distance from Chicago to New York. The (4×4) matrix is

$$X = \begin{bmatrix} 0 & 440 & 229 & 695 \\ 440 & 0 & 213 & 983 \\ 229 & 213 & 0 & 840 \\ 695 & 983 & 840 & 0 \end{bmatrix}.$$

(c) We have data on three countries and 8 years, so let's create a table with 8 rows to represent the years and 3 columns to represent the countries. We only have one economic indicator to consider. For the UK in 2008 we start with 4.8. From the problem we know that the next 7 values are (4.6, 4.4, 4.2, 4.5, 4.8, 5.1, 5.4). For France in 2008 we start with 4.1 and the next seven values are (3.6, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9). For Germany we start with 5.1, and the next seven values are (4.8, 4.5, 4.2, 4.7, 5.2, 5.7, 6.2). Arranging these values in the (8×3) matrix gives us

$$X = \begin{bmatrix} 4.8 & 4.1 & 5.1 \\ 4.6 & 3.6 & 4.8 \\ 4.4 & 3.4 & 4.5 \\ 4.2 & 3.5 & 4.2 \\ 4.5 & 3.6 & 4.7 \\ 4.8 & 3.7 & 5.2 \\ 5.1 & 3.8 & 5.7 \\ 5.4 & 3.9 & 6.2 \end{bmatrix}.$$

(d) There are three patients, and three variables: the times at which each patient is observed, the amount of medication each patient receives, and whether the patient leaves the study or not. Note that patient 3 hasn't yet left the study even at week 7. There are a few ways to arrange these data. But one way is to create a table that has two nested ID variables, a patient identifier and a time identifier. For each combination of patient and time, we can record the dosage and whether or not the patient leaves the study (1 if the patient leaves, 0 if not). We can create a table:

Patient	Week	Dosage	Left Study
1	1	1	0
1	2	1.5	0
1	3	2	1
2	1	1	0
2	2	1.5	0
2	3	2	0
2	4	2.5	0
2	5	3	1
3	1	1	0
3	2	1.5	0
3	3	2	0
3	4	2.5	0
3	5	3	0
3	6	3.5	0
3	7	4	0

We can place this table in a (15×3) matrix:

$$X = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1.5 & 0 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \\ 2 & 1.5 & 0 \\ 3 & 2 & 0 \\ 4 & 2.5 & 0 \\ 5 & 3 & 1 \\ 1 & 1 & 0 \\ 2 & 1.5 & 0 \\ 3 & 2 & 0 \\ 4 & 2.5 & 0 \\ 5 & 3 & 0 \\ 6 & 3.5 & 0 \\ 7 & 4 & 0 \end{bmatrix}.$$

These data are a basic example of a dataset that can be analyzed using survival models, also known as duration or event history models. These models can be used to understand medical prognoses, regime survival, the length of marriages, the duration of wars, and other social phenomena.

(e) There are two voters each making a choice between 3 parties. There are two variables measured: the voters' favorability rating of each party and the voters' choice among the parties. Let's create a data table with 6 rows – one for each unique combination of voter and party – and two columns for the favorability and vote variables. Let's denote favorability on a 1 to 5 scale, where 1 means "very unfavorable," 2 means "slightly unfavorable," 3 means "neutral," 4 means "slightly favorable," and 5 means "very favorable." Let's denote the vote as a binary outcome where 1 indicates a vote for the party and 0 indicates that the voter did not cote for the party. The table is

Voter	Party	Favorability	Vote
1	Labour	5	1
1	Conservative	1	0
1	Liberal	2	0
2	Labour	2	0
2	Conservative	5	1
2	Liberal	4	0

We can place this table in a (6×2) matrix:

$$X = \begin{bmatrix} 5 & 1 \\ 1 & 0 \\ 2 & 0 \\ 2 & 0 \\ 5 & 1 \\ 4 & 0 \end{bmatrix}.$$

2. (a) 12 is just a number. But in matrix terms, we can think of it as a matrix with only one row and one column. That makes 12 a scalar.

   (b) This is a matrix with 3 rows and 1 column. That makes this object a column vector.

   (c) This is a matrix with 1 row and 3 columns. That makes this object a row vector.

   (d) This matrix has two rows and two columns. It is a matrix, and that its elements are not numbers does not change that fact. Remember that a matrix is a two-dimensional array of objects that are often, *but not necessarily*, numbers. Since this matrix has the same number of rows and columns, it is a square matrix.

   (e) This matrix has three rows and three columns. It is a matrix, and as with part (d) that its elements are not numbers does not change that fact. Since this matrix has the same number of rows and columns, it is a square matrix. Also, remember that a matrix is a two-dimensional array of objects in which the row and column positions have meaning. In this case the first row and column each represent zombies, the second row and column represent ninjas, and the third row and column represent pirates. Each element is the combination of the categories represented by the row and column. Since the combination of zombies and ninjas produce the same result as the combination of ninjas and zombies (the dreaded "zombie ninjas"), each element is equal to the one in which the row and column positions are switched. Therefore this matrix is a symmetric matrix.

   (f) This matrix has the same number of rows and columns, so it is square. The dots indicate that the upper triangular elements are equal to the elements with the row and column numbers switched. Therefore this matrix is symmetric.

   (g) This matrix has the same number of rows and columns, so it is square. All of the elements above the diagonal are zero, so this is a lower-triangular matrix.

   (h) This matrix has the same number of rows and columns, so it is square. Each element is equal to the one in which the row and column positions are switched, so this matrix is a symmetric matrix. All of the elements above and below the diagonal are zero, so this is a diagonal matrix. Furthermore, all of the elements on the diagonal are equal, so this is a scalar matrix. Finally, since all of the diagonal elements are equal to 1, this is an identity matrix. Since there are 9 rows and columns, the name of this identity matrix is I_9.

   (i) This matrix has the same number of rows and columns, so it is square, but it is not symmetric or an upper or lower triangular matrix. Instead, observe that the matrix can be broken into partitions as follows:

$$\left[\begin{array}{ccc|cc} 8 & 7 & -9 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & -7 & 8 \\ 0 & 0 & 0 & 12 & 2 \\ 0 & 0 & 0 & -9 & 7 \end{array}\right].$$

So this matrix is a partitioned matrix. We can rewrite this matrix as

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where

$$A = \begin{bmatrix} 8 & 7 & -9 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 6 \\ -7 & 8 \\ 12 & 2 \\ -9 & 7 \end{bmatrix}.$$

Since matrices B and C are off-diagonal in the partitioned matrix and since they consist of all zeroes, this matrix is a block-diagonal matrix.

3. (a) First, we apply scalar multiplication to find $3A$ and $2B$:

$$3A = 3\begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \times -1 \\ 3 \times 3 \\ 3 \times 4 \end{bmatrix} = \begin{bmatrix} -3 \\ 9 \\ 12 \end{bmatrix}.$$

$$2B = 2\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \times 3 \\ 2 \times -2 \\ 2 \times 1 \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 2 \end{bmatrix}.$$

Addition and subtraction of two matrices or vectors is conformable (that is, possible) only if the dimensions of the two matrices/vectors are exactly the same. In this case, $3A$ and $2B$ are both (3×1) vectors, so subtraction is conformable, and the difference is a (3×1) matrix whose elements are differences of the corresponding elements of the two vectors:

$$3A - 2B = \begin{bmatrix} -3 \\ 9 \\ 12 \end{bmatrix} - \begin{bmatrix} 6 \\ -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -3-6 \\ 9-(-4) \\ 12-2 \end{bmatrix} = \begin{bmatrix} -9 \\ 13 \\ 10 \end{bmatrix}.$$

(b) $A \cdot B$ is the inner-product (or dot-product) of two vectors, and to find it we multiply the corresponding elements of each vector together, and sum these products:

$$A \cdot B = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = (-1 \times 3) + (3 \times -2) + (4 \times 1) = -3 - 6 + 4 = -5.$$

(c) $A \times B$ is the outer-product (or cross-product) of two vectors, and to find it we multiply each element of A by each element of B, and we align these products in a matrix whose row number corresponds to the row of the element from A and whose column number corresponds to the row of the element from B:

$$A \times B = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} \times \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \times 3 & -1 \times -2 & -1 \times 1 \\ 3 \times 3 & 3 \times -2 & 3 \times 1 \\ 4 \times 3 & 4 \times -2 & 4 \times 1 \end{bmatrix} = \begin{bmatrix} -3 & 2 & -1 \\ 9 & -6 & 3 \\ 12 & -8 & 4 \end{bmatrix}.$$

(d) CA is an example of matrix multiplication, which is conformable only if the number of columns of the left factor equal the number of rows in the right factor. If multiplication is conformable, then the product is a matrix with the same number of rows as the left factor and the same number of columns as the right factor. In this case, C is a (2×3) matrix and A is a (3×1) vector. Therefore multiplication is conformable, and the product is a (2×1) vector.

In general, the (i, j) element of the product is the inner-product of the ith row of the left factor and the jth column of the right factor. in this case the product is

$$CA = \begin{bmatrix} 3 & 2 & -4 \\ -8 & 0 & 6 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} (3 \times -1) + (2 \times 3) + (-4 \times 4) \\ (-8 \times -1) + (0 \times 3) + (6 \times 4) \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \end{bmatrix}.$$

(e) B is a (3×1) vector and D is a (3×2) matrix. In this case the number of columns of the left factor does not equal the number of columns of the right factor. So matrix multiplication is not conformable.

(f) $B \otimes D$ is the Kronecker product of matrices B and D. To find the Kronecker product, we scalar multiply the entire right factor by every element of the left factor, and expand the number of rows and columns of the product as necessary. In this case the Kronecker product is

$$B \otimes D = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix}$$

$$= \begin{bmatrix} 3 \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix} \\ \\ -2 \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix} \\ \\ 1 \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 18 & -6 \\ -3 & 9 \\ -9 & 24 \\ -12 & 4 \\ 2 & -6 \\ 6 & -16 \\ 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix}.$$

(g) C is a (2×3) matrix and D is a (3×2) matrix. Since the number of columns of the left factor equal the number of rows of the right factor, matrix multiplication is conformable and the product is (2×2).

Each element of the product is the inner-product of the corresponding row in the left factor and the corresponding column in the right factor. In this case the product is

$$CD = \begin{bmatrix} 3 & 2 & -4 \\ -8 & 0 & 6 \end{bmatrix} \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix}$$

$$= \begin{bmatrix} (3\times 6)+(2\times -1)+(-4\times -3) & (3\times -2)+(2\times 3)+(-4\times 8) \\ \\ (-8\times 6)+(0\times -1)+(6\times -3) & (-8\times -2)+(0\times 3)+(6\times 8) \end{bmatrix} = \begin{bmatrix} 28 & -32 \\ -66 & 64 \end{bmatrix}.$$

(h) D is a (3×2) matrix and C is a (2×3) matrix. Since the number of columns of the left factor equal the number of rows of the right factor, matrix multiplication is conformable and the product is (3×3).

Each element of the product is the inner-product of the corresponding row in the left factor and the corresponding column in the right factor. In this case the product is

$$DC = \begin{bmatrix} 6 & -2 \\ -1 & 3 \\ -3 & 8 \end{bmatrix} \begin{bmatrix} 3 & 2 & -4 \\ -8 & 0 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} (6\times 3)+(-2\times -8) & (6\times 2)+(-2\times 0) & (6\times -4)+(-2\times 6) \\ \\ (-1\times 3)+(3\times -8) & (-1\times 2)+(3\times 0) & (-1\times -4)+(3\times 6) \\ \\ (-3\times 3)+(8\times -8) & (-3\times 2)+(8\times 0) & (-3\times -4)+(8\times 6) \end{bmatrix} = \begin{bmatrix} 34 & 12 & -36 \\ -27 & -2 & 22 \\ -73 & -6 & 60 \end{bmatrix}.$$

The fact that our calculations for CD and DC differ demonstrates that the order of matrix multiplication can change the answer.

(i) The transpose of matrix C contains the elements on C, but the row and column position of each element is switched:

$$C' = \begin{bmatrix} 3 & -8 \\ 2 & 0 \\ -4 & 6 \end{bmatrix}.$$

C' is a (3×2) matrix and C is a (2×3) matrix. Since the number of columns of the left factor equal the number of rows of the right factor, matrix multiplication is conformable and the product is (3×3).

Each element of the product is the inner-product of the corresponding row in the left factor and the corresponding column in the right factor. In this case the product is

$$C'C = \begin{bmatrix} 3 & -8 \\ 2 & 0 \\ -4 & 6 \end{bmatrix} \begin{bmatrix} 3 & 2 & -4 \\ -8 & 0 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} (3\times 3)+(-8\times -8) & (3\times 2)+(-8\times 0) & (3\times -4)+(-8\times 6) \\ \\ (2\times 3)+(0\times -8) & (2\times 2)+(0\times 0) & (2\times -4)+(0\times 6) \\ \\ (-4\times 3)+(6\times -8) & (-4\times 2)+(6\times 0) & (-4\times -4)+(6\times 6) \end{bmatrix} = \begin{bmatrix} 73 & 6 & -60 \\ 6 & 4 & -8 \\ -60 & -8 & 52 \end{bmatrix}.$$

4. A symmetric matrix is a square matrix that has the property that any element is equal to the element with switched row and column numbers. The answer in question 3, part (i) is symmetric because it is square and

the (1,2), (1,3), and (2,3) elements are equal to the (2,1), (3,1), and (3,2) elements respectively.

Any matrix that is left-multiplied by its transpose is symmetric because the ith row in the transpose is equal to the ith column in the original matrix. Consider a general element in the (i, j) position, where $i \neq j$. This element is equal to the inner-product of the ith row of the transpose and the jth column of the original matrix. The symmetric element in the (j, i) position is equal to the inner-product of the jth row of the transpose and the ith column of the original matrix. But since we are multiplying by the transpose, these two inner-products must be exactly equal. Therefore every (i, j) element is equal to the (j, i) element and the matrix is symmetric.

The same argument applies when a matrix is right-multiplied by its transpose. So CC' must also be a symmetric matrix.

5. The transpose of the X matrix, X', switches the row and column position of every element in X and is therefore a (2×10) matrix:

$$X' = \begin{bmatrix} 2 & 4 & -12 & -5 & 9 & 6 & -7 & 7 & -13 & 2 \\ -7 & 3 & 9 & -8 & 0 & -1 & 5 & 10 & -2 & -4 \end{bmatrix},$$

The product $X'X$ multiplies a (2×10) by a (10×2) matrix, which is conformable to matrix multiplication and results in a (2×2) matrix. The (1,1) element is the dot product of the first row of X' and the first column of X:

$$(2\times2)+(4\times4)+(-12\times-12)+(-5\times-5)+(9\times9)+(6\times6)+(-7\times-7)+(7\times7)+(-13\times-13)+(2\times2) = 577.$$

The (1,2) element is the dot product of the first row of X' and the second column of X:

$$(2\times-7)+(4\times3)+(-12\times9)+(-5\times-8)+(9\times0)+(6\times-1)+(-7\times5)+(7\times10)+(-13\times-2)+(2\times-4) = -23.$$

Note that the calculation for the (2,1) element – the dot product of the second row of X' and the first column of X – results in exactly the same calculation we encountered for the (1,2) element:

$$(-7\times2)+(3\times4)+(9\times-12)+(-8\times-5)+(0\times9)+(-1\times6)+(5\times-7)+(10\times7)+(-2\times-13)+(-4\times2) = -23.$$

The (2,2) element is the dot product of the second row of X' and the second column of X:

$$(-7\times-7)+(3\times3)+(9\times9)+(-8\times-8)+(0\times0)+(-1\times-1)+(5\times5)+(10\times10)+(-2\times-2)+(-4\times-4) = 349.$$

Put together, the whole product is

$$X'X = \begin{bmatrix} 577 & -23 \\ -23 & 349 \end{bmatrix}.$$

Multiplying this matrix by $\frac{1}{10}$ is scalar multiplication, which involves multiplying every element of the matrix by the scalar. Applying this multiplication, we find that

$$\frac{1}{10}X'X = \begin{bmatrix} 57.7 & -2.3 \\ -2.3 & 34.9 \end{bmatrix}.$$

6. (a) First let's compute the product AB. This matrix is the product of a (3×2) and (2×3) matrix, so it is conformable and (3×3). The product is

$$AB = \begin{bmatrix} (10\times-4)+(-2\times2) & (10\times9)+(-2\times1) & (10\times-12)+(-2\times6) \\ (-1\times-4)+(6\times2) & (-1\times9)+(6\times1) & (-1\times-12)+(6\times6) \\ (8\times-4)+(3\times2) & (8\times9)+(3\times1) & (8\times-12)+(3\times6) \end{bmatrix} = \begin{bmatrix} -44 & 88 & -132 \\ 16 & -3 & 48 \\ -26 & 75 & -78 \end{bmatrix}.$$

The trace of AB is the sum of the diagonal elements:

$$tr(AB) = -44 - 3 - 78 = -125.$$

Next let's compute the product BA. This matrix is the product of a (2×3) and (3×2) matrix, so it is conformable and (2×2). The product is

$$BA = \begin{bmatrix} (-4 \times 10) + (9 \times -1) + (-12 \times 8) & (-4 \times -2) + (9 \times 6) + (-12 \times 3) \\ (2 \times 10) + (1 \times -1) + (6 \times 8) & (2 \times -2) + (1 \times 6) + (6 \times 3) \end{bmatrix} = \begin{bmatrix} -145 & 26 \\ 67 & 20 \end{bmatrix}.$$

The trace of BA is the sum of the diagonal elements:

$$tr(BA) = -145 + 20 = -125.$$

We have confirmed that $tr(AB) = tr(BA)$ in this case. This property is true in general beyond this special case.

(b) First let's compute the product CD. This matrix is the product of a (2×3) and (3×2) matrix, so it is conformable and (2×2). The product is

$$CD = \begin{bmatrix} (6 \times -1) + (1 \times 4) + (-2 \times 5) & (6 \times 3) + (1 \times -3) + (-2 \times 2) \\ (4 \times -1) + (3 \times 4) + (7 \times 5) & (4 \times 3) + (3 \times -3) + (7 \times 2) \end{bmatrix} = \begin{bmatrix} -12 & 11 \\ 43 & 17 \end{bmatrix}.$$

The transpose of CD switches the row and column number of every element of CD:

$$(CD)' = \begin{bmatrix} -12 & 43 \\ 11 & 17 \end{bmatrix}.$$

Likewise, the transposes of C and D individually switch the row and column numbers of C and D,

$$C' = \begin{bmatrix} 6 & 4 \\ 1 & 3 \\ -2 & 7 \end{bmatrix}, \qquad D' = \begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix}$$

and the product $D'C'$ is the product of a (2×3) and (3×2) matrix, so it is conformable and (2×2). The product is

$$D'C' = \begin{bmatrix} (-1 \times 6) + (4 \times 1) + (5 \times -2) & (-1 \times 4) + (4 \times 3) + (5 \times 7) \\ (3 \times 6) + (-3 \times 1) + (2 \times -2) & (3 \times 4) + (-3 \times 3) + (2 \times 7) \end{bmatrix} = \begin{bmatrix} -12 & 43 \\ 11 & 17 \end{bmatrix}.$$

We've confirmed that $(CD)' = D'C'$. This property is also true in general situations.

(c) A and D are both (3×2) matrices, and since they have the same dimensions they are conformable to addition. The sum of these matrices is

$$A + D = \begin{bmatrix} 10 - 1 & -2 + 3 \\ -1 + 4 & 6 - 3 \\ 8 + 5 & 3 + 2 \end{bmatrix} = \begin{bmatrix} 9 & 1 \\ 3 & 3 \\ 13 & 5 \end{bmatrix}.$$

The transpose of $A + D$ switches the row and column numbers of every element of $A + D$:

$$(A + D)' = \begin{bmatrix} 9 & 3 & 13 \\ 1 & 3 & 5 \end{bmatrix}.$$

Likewise the transposes of A and D individually are

$$A' = \begin{bmatrix} 10 & -1 & 8 \\ -2 & 6 & 3 \end{bmatrix}, \qquad D' = \begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix},$$

and the sum of A' and D' is

$$A' + D' = \begin{bmatrix} 10-1 & -1+4 & 8+5 \\ -2+3 & 6-3 & 3+2 \end{bmatrix} = \begin{bmatrix} 9 & 3 & 13 \\ 1 & 3 & 5 \end{bmatrix}.$$

We've demonstrated that $(A+D)' = A' + D'$ in this case, and this property is true in general as well.

(d) The scalar product $5B$ is equal to

$$5B = 5\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} = \begin{bmatrix} -20 & 45 & -60 \\ 10 & 5 & 30 \end{bmatrix}.$$

The transpose of $5B$ is

$$(5B)' = \begin{bmatrix} -20 & 10 \\ 45 & 5 \\ -60 & 30 \end{bmatrix}.$$

The scalar product $5B'$ is equal to

$$5B' = 5\begin{bmatrix} -4 & 2 \\ 9 & 1 \\ -12 & 6 \end{bmatrix} = \begin{bmatrix} -20 & 10 \\ 45 & 5 \\ -60 & 30 \end{bmatrix}.$$

We've demonstrated that $(5B)' = 5B'$ in this case, and this property is true in general as well.

(e) The sum $B+C$ equals

$$B + C = \begin{bmatrix} -4+6 & 9+1 & -12-2 \\ 2+4 & 1+3 & 6+7 \end{bmatrix} = \begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix}.$$

The Kronecker product $A \otimes (B+C)$ equals

$$A \otimes (B+C) = \begin{bmatrix} 10 & -2 \\ -1 & 6 \\ 8 & 3 \end{bmatrix} \otimes \begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix}$$

$$= \begin{bmatrix} 10\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} & -2\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} \\ -1\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} & 6\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} \\ 8\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} & 3\begin{bmatrix} 2 & 10 & -14 \\ 6 & 4 & 13 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 20 & 100 & -140 & -4 & -20 & 28 \\ 60 & 40 & 130 & -12 & -8 & -26 \\ -2 & -10 & 14 & 12 & 60 & -84 \\ -6 & -4 & -13 & 36 & 24 & 78 \\ 16 & 80 & -112 & 6 & 30 & -42 \\ 48 & 32 & 104 & 18 & 12 & 39 \end{bmatrix}.$$

The Kronecker product $A \otimes B$ is

$$A \otimes B = \begin{bmatrix} 10 & -2 \\ -1 & 6 \\ 8 & 3 \end{bmatrix} \otimes \begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} 10\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} & -2\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} \\ -1\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} & 6\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} \\ 8\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} & 3\begin{bmatrix} -4 & 9 & -12 \\ 2 & 1 & 6 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} -40 & 90 & -120 & 8 & -18 & 24 \\ 20 & 10 & 60 & -4 & -2 & -12 \\ 4 & -9 & 12 & -24 & 54 & -72 \\ -2 & -1 & -6 & 12 & 6 & 36 \\ -32 & 78 & -106 & -12 & 27 & -36 \\ 16 & 8 & 48 & 6 & 3 & 18 \end{bmatrix},$$

the Kronecker product $A \otimes C$ is

$$A \otimes B = \begin{bmatrix} 10 & -2 \\ -1 & 6 \\ 8 & 3 \end{bmatrix} \otimes \begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix}$$

$$= \begin{bmatrix} 10\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} & -2\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} \\ -1\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} & 6\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} \\ 8\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} & 3\begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 60 & 10 & -20 & -12 & -2 & 4 \\ 40 & 30 & 70 & -8 & -6 & -14 \\ -6 & -1 & 2 & 36 & 6 & -12 \\ -4 & -3 & -7 & 24 & 18 & 42 \\ 48 & 8 & -16 & 18 & 3 & -6 \\ 32 & 24 & 56 & 12 & 9 & 21 \end{bmatrix},$$

and the sum of these two Kronecker products $(A \otimes B) + (A \otimes C)$ is

$$\begin{bmatrix} -40 & 90 & -120 & 8 & -18 & 24 \\ 20 & 10 & 60 & -4 & -2 & -12 \\ 4 & -9 & 12 & -24 & 54 & -72 \\ -2 & -1 & -6 & 12 & 6 & 36 \\ -32 & 78 & -106 & -12 & 27 & -36 \\ 16 & 8 & 48 & 6 & 3 & 18 \end{bmatrix} + \begin{bmatrix} 60 & 10 & -20 & -12 & -2 & 4 \\ 40 & 30 & 70 & -8 & -6 & -14 \\ -6 & -1 & 2 & 36 & 6 & -12 \\ -4 & -3 & -7 & 24 & 18 & 42 \\ 48 & 8 & -16 & 18 & 3 & -6 \\ 32 & 24 & 56 & 12 & 9 & 21 \end{bmatrix}$$

$$= \begin{bmatrix} 20 & 100 & -140 & -4 & -20 & 28 \\ 60 & 40 & 130 & -12 & -8 & -26 \\ -2 & -10 & 14 & 12 & 60 & -84 \\ -6 & -4 & -13 & 36 & 24 & 78 \\ 16 & 80 & -112 & 6 & 30 & -42 \\ 48 & 32 & 104 & 18 & 12 & 39 \end{bmatrix}.$$

We've demonstrated that $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ in this case, and this property is true in general as well.

(f) The Kronecker product $C \otimes D$ is

$$C \otimes D = \begin{bmatrix} 6 & 1 & -2 \\ 4 & 3 & 7 \end{bmatrix} \otimes \begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 6\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} & 1\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} & -2\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} \\ 4\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} & 3\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} & 7\begin{bmatrix} -1 & 3 \\ 4 & -3 \\ 5 & 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} -6 & 18 & -1 & 3 & 2 & -6 \\ 24 & -18 & 4 & -3 & -8 & 6 \\ 30 & 12 & 5 & 2 & -10 & -4 \\ -4 & 12 & -3 & 9 & -7 & 21 \\ 16 & -12 & 12 & -9 & 28 & -21 \\ 20 & 8 & 15 & 6 & 35 & 14 \end{bmatrix},$$

and the transpose of this matrix is

$$(C \otimes D)' = \begin{bmatrix} -6 & 24 & 30 & -4 & 16 & 20 \\ 18 & -18 & 12 & 12 & -12 & 8 \\ -1 & 4 & 5 & -3 & 12 & 15 \\ 3 & -3 & 2 & 9 & -9 & 6 \\ 2 & -8 & -10 & -7 & 28 & 35 \\ -6 & 6 & -4 & 21 & -21 & 14 \end{bmatrix}.$$

Individually, the transposes of matrices C and D are

$$C' = \begin{bmatrix} 6 & 4 \\ 1 & 3 \\ -2 & 7 \end{bmatrix}, \qquad D' = \begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix},$$

and the Kronecker product $C' \otimes D'$ is

$$C' \otimes D' = \begin{bmatrix} 6 & 4 \\ 1 & 3 \\ -2 & 7 \end{bmatrix} \otimes \begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 6\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} & 4\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} \\ 1\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} & 3\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} \\ -2\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} & 7\begin{bmatrix} -1 & 4 & 5 \\ 3 & -3 & 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} -6 & 24 & 30 & -4 & 16 & 20 \\ 18 & -18 & 12 & 12 & -12 & 8 \\ -1 & 4 & 5 & -3 & 12 & 15 \\ 3 & -3 & 2 & 9 & -9 & 6 \\ 2 & -8 & -10 & -7 & 28 & 35 \\ -6 & 6 & -4 & 21 & -21 & 14 \end{bmatrix}.$$

We've demonstrated that $(C \otimes D)' = C' \otimes D'$ in this case, and this property is true in general as well.

(g) We previously calculated $A + D$ in part (c):

$$A + D = \begin{bmatrix} 9 & 1 \\ 3 & 3 \\ 13 & 5 \end{bmatrix}.$$

The product $B(A+D)$ multiplies a (2×3) matrix by a (3×2) matrix. The product, which is conformable and (2×2), is

$$B(A+D) = \begin{bmatrix} (-4\times9)+(9\times3)+(-12\times13) & (-4\times1)+(9\times3)+(-12\times5) \\ (2\times9)+(1\times3)+(6\times13) & (2\times1)+(1\times3)+(6\times5) \end{bmatrix} = \begin{bmatrix} -165 & -37 \\ 99 & 35 \end{bmatrix}.$$

We previously calculated the product BA in part (a):

$$BA = \begin{bmatrix} -145 & 26 \\ 67 & 20 \end{bmatrix},$$

and the product BD multiplies a (2×3) matrix by a (3×2) matrix. The product, which is conformable and (2×2), is

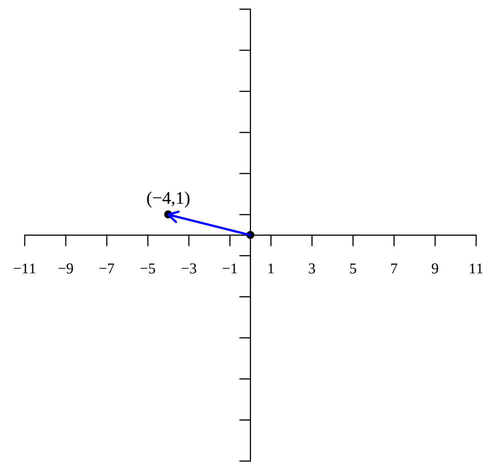
$$BD = \begin{bmatrix} (-4\times-1)+(9\times4)+(-12\times5) & (-4\times3)+(9\times-3)+(-12\times2) \\ (2\times-1)+(1\times4)+(6\times5) & (2\times3)+(1\times-3)+(6\times2) \end{bmatrix} = \begin{bmatrix} -20 & -63 \\ 32 & 15 \end{bmatrix}.$$

Finally, the sum of BA and BD is

$$BA + BD = \begin{bmatrix} -145 & 26 \\ 67 & 20 \end{bmatrix} + \begin{bmatrix} -20 & -63 \\ 32 & 15 \end{bmatrix} = \begin{bmatrix} -165 & -37 \\ 99 & 35 \end{bmatrix}.$$

We've demonstrated that $B(A+D) = BA+BD$ in this case, and this property is true in general as well.

7. (a) A plot of this vector is



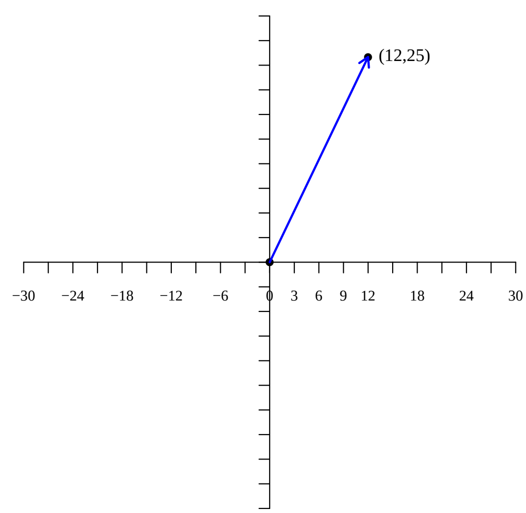
The magnitude of the vector is

$$\left\| \begin{bmatrix} -4 \\ 1 \end{bmatrix} \right\| = \sqrt{(-4)^2 + (1)^2} = \sqrt{17} = 4.12.$$

A unit vector in the same direction is

$$\frac{1}{4.12} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.97 \\ 0.24 \end{bmatrix}.$$

(b) A plot of this vector is



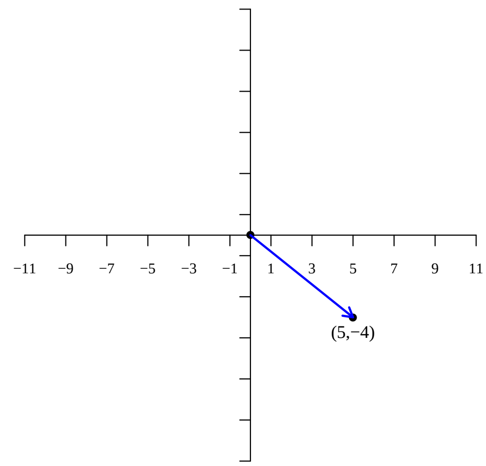
The magnitude of the vector is

$$\left\| \begin{bmatrix} 12 \\ 25 \end{bmatrix} \right\| = \sqrt{(12)^2 + (25)^2} = \sqrt{769} = 27.73.$$

A unit vector in the same direction is

$$\frac{1}{27.73} \begin{bmatrix} 12 \\ 25 \end{bmatrix} = \begin{bmatrix} 0.43 \\ 0.90 \end{bmatrix}.$$

(c) A plot of this vector is



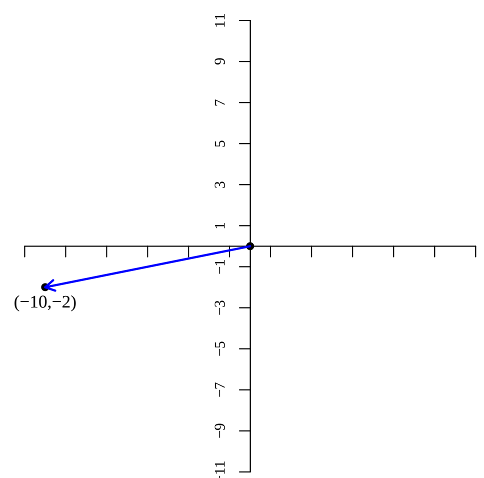
The magnitude of the vector is

$$\left| \left[\begin{array}{r} 5 \\ -4 \end{array} \right] \right| = \sqrt{(5)^2 + (-4)^2} = \sqrt{41} = 6.4.$$

A unit vector in the same direction is

$$\frac{1}{6.4} \left[\begin{array}{r} 5 \\ -4 \end{array} \right] = \left[\begin{array}{r} 0.78 \\ -0.62 \end{array} \right].$$

(d) A plot of this vector is



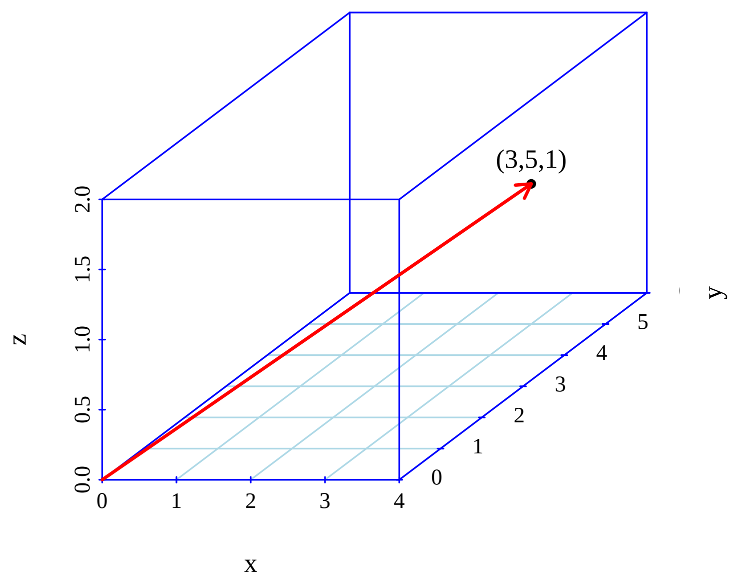
The magnitude of the vector is

$$\left| \left[\begin{array}{r} -10 \\ -2 \end{array} \right] \right| = \sqrt{(-10)^2 + (-2)^2} = \sqrt{104} = 6.4.$$

A unit vector in the same direction is

$$\frac{1}{6.4} \left[\begin{array}{r} -10 \\ -2 \end{array} \right] = \left[\begin{array}{r} -0.98 \\ -0.20 \end{array} \right].$$

(e) A plot of this vector is



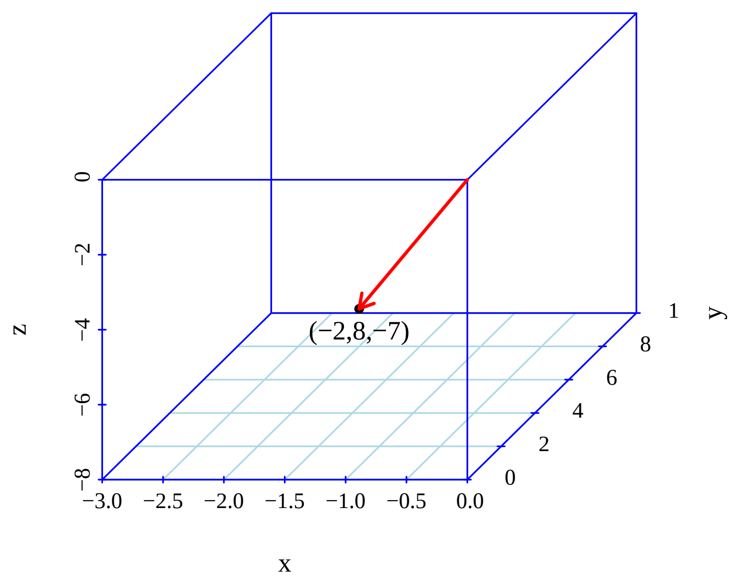
The magnitude of the vector is

$$\left|\begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix}\right| = \sqrt{(3)^2 + (5)^2 + (1)^2} = \sqrt{35} = 5.92.$$

A unit vector in the same direction is

$$\frac{1}{5.92}\begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.51 \\ 0.84 \\ 0.17 \end{bmatrix}.$$

(f) A plot of this vector is



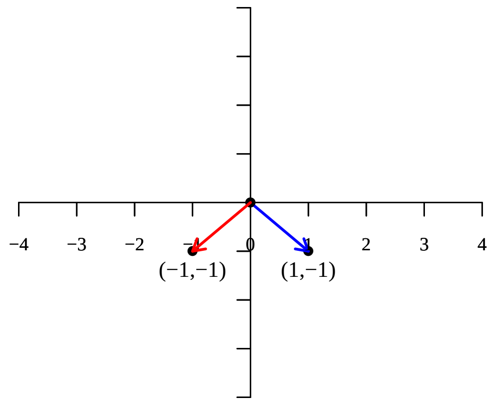
The magnitude of the vector is

$$\left| \left[\begin{array}{r} -2 \\ 8 \\ -7 \end{array} \right] \right| = \sqrt{(-2)^2 + (8)^2 + (-7)^2} = \sqrt{117} = 10.82.$$

A unit vector in the same direction is

$$\frac{1}{10.82} \left[\begin{array}{r} -2 \\ 8 \\ -7 \end{array} \right] = \left[\begin{array}{r} -0.18 \\ 0.74 \\ -0.65 \end{array} \right].$$

8. The following plot indicates that the vector that is turned 90 degrees clockwise is $\left[\begin{array}{r} -1 \\ -1 \end{array} \right]$:



The trick is finding a matrix that when left-multiplied by the vector yields a product of $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$. Consider a general (2×2) matrix as a left-factor:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix},$$

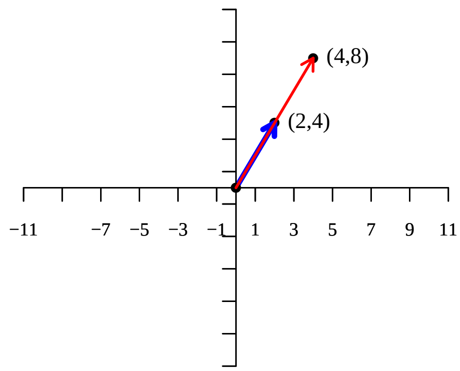
which will be true only if

$$\begin{cases} a - b = -1, \\ c - d = -1. \end{cases}$$

There are many matrices that perform this transformation, but one example is the (2×2) matrix where $a = 2$, $b = 3$, $c = 4$, and $d = 5$. To see that this works, we compute the product:

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

9. Any multiple of $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ is in the same direction. Consider, for example, $\begin{bmatrix} 4 \\ 8 \end{bmatrix}$. Both vectors are graphed below:



The trick is finding a matrix that when left-multiplied by the vector yields a product of $\begin{bmatrix} 4 \\ 8 \end{bmatrix}$. Consider a general (2×2) matrix as a left-factor:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix},$$

which will be true only if

$$\begin{cases} 2a + 4b = 4, \\ 2c + 4d = 8. \end{cases}$$

There are many matrices that perform this transformation, but one example is the (2×2) matrix where $a = 0$, $b = 1$, $c = 2$, and $d = 1$. To see that this works, we compute the product:

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

10. (a) We begin with

$$\begin{bmatrix} 2 & 1 & 6 \\ 8 & 4 & 3 \\ 12 & 1 & 4 \end{bmatrix}.$$

We multiply the first row by -4 and add it to the second row,

$$\begin{bmatrix} 2 & 1 & 6 \\ 0 & 0 & -21 \\ 12 & 1 & 4 \end{bmatrix},$$

then multiply the first row by -6 and add it to the third row,

$$\begin{bmatrix} 2 & 1 & 6 \\ 0 & 0 & -21 \\ 0 & -5 & -32 \end{bmatrix}.$$

We interchange second and third rows,

$$\begin{bmatrix} 2 & 1 & 6 \\ 0 & -5 & -32 \\ 0 & 0 & -21 \end{bmatrix},$$

divide the third row by -21,

$$\begin{bmatrix} 2 & 1 & 6 \\ 0 & -5 & -32 \\ 0 & 0 & 1 \end{bmatrix},$$

multiply the third row by 32 and add it to the second row,

$$\begin{bmatrix} 2 & 1 & 6 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and multiply it by -6 and add it to the first row,

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

We divide the second row by -5,

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

multiply it by -1, and add it to the first row,

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

Finally we divide the first row by 2,

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

(b) We begin with

$$\begin{bmatrix} 1 & -3 & 4 \\ 3 & 7 & -4 \\ 5 & 9 & 12 \end{bmatrix}.$$

We multiply the first row by -3 and add it to the second row,

$$\begin{bmatrix} 1 & -3 & 4 \\ 0 & 16 & -16 \\ 5 & 9 & 12 \end{bmatrix},$$

and multiply it by -5 and add it to the third row,

$$\begin{bmatrix} 1 & -3 & 4 \\ 0 & 16 & -16 \\ 0 & 24 & -8 \end{bmatrix}.$$

We divide the second row by 16 and the third row by 8,

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

We then multiply the second row by -3 and add it to the third row,

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

and divide the third row by 2,

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

Now we add the third row to the second,

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

we multiply the third row by -4 and add it to the first row,

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

and we multiply the second row by 3 and add it to the first row,

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

(c) We begin with

$$\begin{bmatrix} 3 & 4 & 7 \\ 9 & -3 & 1 \\ -3 & 11 & 13 \end{bmatrix}.$$

We multiply the first row by -3 and add it to the second row,

$$\begin{bmatrix} 3 & 4 & 7 \\ 0 & -15 & -20 \\ -3 & 11 & 13 \end{bmatrix},$$

and add the first row to the third row,

$$\begin{bmatrix} 3 & 4 & 7 \\ 0 & -15 & -20 \\ 0 & 15 & 20 \end{bmatrix}.$$

Next we add the second row to the third,

$$\begin{bmatrix} 3 & 4 & 7 \\ 0 & -15 & -20 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since we now have a row of all zeroes, it will not be possible to reduce this matrix to an identity matrix. Let's continue reducing the matrix as much as possible. Next we divide the second row by -5,

$$\begin{bmatrix} 3 & 4 & 7 \\ 0 & 3 & 4 \\ 0 & 0 & 0 \end{bmatrix},$$

and in order to avoid fractions, we multiply the first row by 3 and the second row by 4,

$$\begin{bmatrix} 9 & 12 & 21 \\ 0 & 12 & 16 \\ 0 & 0 & 0 \end{bmatrix}.$$

We multiply the second row by -1 and add it to the first,

$$\begin{bmatrix} 9 & 0 & 5 \\ 0 & 12 & 16 \\ 0 & 0 & 0 \end{bmatrix}.$$

Finally – there's no avoiding fractions any longer – we divide the first row by 9 and the second row by 12,

$$\begin{bmatrix} 1 & 0 & \frac{5}{9} \\ 0 & 1 & \frac{3}{4} \\ 0 & 0 & 0 \end{bmatrix}.$$

This matrix is reduced as much as possible because the first two rows and columns contain the I_2 identity matrix. There are two rows that do not consist entirely of zeroes.

(d) We begin with

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 6 & 6 & -8 & 3 \\ -4 & 0 & 2 & 9 \\ 10 & 0 & 6 & -2 \end{bmatrix}.$$

We multiply the first row by -3 and add it to the second,

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ -4 & 0 & 2 & 9 \\ 10 & 0 & 6 & -2 \end{bmatrix},$$

then multiply it by 2 and add it to the third row (we don't worry about erasing the zero in the (3,2) spot even though eventually we want this element to be 0 – we will come back to fix this element eventually),

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 6 & 4 & -5 \\ 10 & 0 & 6 & -2 \end{bmatrix},$$

and then multiply it by -5 and add it to the fourth row (we will also fix the (4,2) element later),

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 6 & 4 & -5 \\ 0 & -15 & 1 & 33 \end{bmatrix}.$$

Next we multiply the second row by 2 and add it to the third row (see?),

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 0 & -18 & 43 \\ 0 & -15 & 1 & 33 \end{bmatrix},$$

and we multiply it by 5 and add it to the fourth row,

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 0 & -18 & 43 \\ 0 & 0 & -54 & 177 \end{bmatrix}.$$

We multiply the third row by -3 and add it to the fourth row,

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 0 & -18 & 43 \\ 0 & 0 & 0 & 48 \end{bmatrix},$$

and divide the fourth row by 48,

$$\begin{bmatrix} 2 & 3 & 1 & -7 \\ 0 & -3 & -11 & 24 \\ 0 & 0 & -18 & 43 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

We can take a series of easy and useful steps now. We multiply the fourth row by -43 and add it to the third row, then we multiply it by -24 and add it to the second row, and multiply it by 7 and add it to the first row,

$$\begin{bmatrix} 2 & 3 & 1 & 0 \\ 0 & -3 & -11 & 0 \\ 0 & 0 & -18 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Next we divide the third row by -18,

$$\begin{bmatrix} 2 & 3 & 1 & 0 \\ 0 & -3 & -11 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

then multiply it by 11 and add it to the second row, and multiply it by -1 and add it to the first row,

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

We add the second row to the first,

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

and divide the first row by 2 and the second row by -3,

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

(e) We begin with

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 3 & 6 & 2 & -2 \\ 4 & 8 & -6 & 1 \\ -2 & -4 & -10 & 5 \end{bmatrix}.$$

We multiply the first row by -3 and add it to the second row,

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 0 & 0 & 26 & -11 \\ 4 & 8 & -6 & 1 \\ -2 & -4 & -10 & 5 \end{bmatrix},$$

then we multiply it by -4 and add it to the third row,

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 0 & 0 & 26 & -11 \\ 0 & 0 & 26 & -11 \\ -2 & -4 & -10 & 5 \end{bmatrix},$$

and then we multiply it by 2 and add it to the fourth row,

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 0 & 0 & 26 & -11 \\ 0 & 0 & 26 & -11 \\ 0 & 0 & -26 & 11 \end{bmatrix}.$$

Next we multiply the second row by -1 and add it to the third row,

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 0 & 0 & 26 & -11 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -26 & 11 \end{bmatrix},$$

and add it to the fourth row as is,

$$\begin{bmatrix} 1 & 2 & -8 & 3 \\ 0 & 0 & 26 & -11 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We now have two rows that consist entirely of zeroes, so it will not be possible to reduce this matrix to an identity matrix. Let's reduce it as much as possible. Here's where it gets ugly. To avoid fractions, multiply the first row by 13 and the second row by 4,

$$\begin{bmatrix} 13 & 26 & -104 & 36 \\ 0 & 0 & 104 & -44 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and add the second row to the first,

$$\begin{bmatrix} 13 & 26 & 0 & -8 \\ 0 & 0 & 104 & -44 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Finally, divide the first row by 13 and the second row by 104,

$$\begin{bmatrix} 1 & 2 & 0 & -\frac{8}{13} \\ 0 & 0 & 1 & -\frac{11}{26} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We cannot reduce the fraction any further, and there are two rows remaining that do not consist entirely of zeroes.

11. Let's create a (10×1) column vector named Y that contains the values of the dependent variable. Likewise we can create a (10×1) column vector of regression errors named ϵ and a (3×1) column vector called B that contains the constant and the two β coefficients:

$$Y = \begin{bmatrix} 9 \\ 5 \\ 4 \\ 7 \\ 4 \\ 6 \\ 5 \\ 7 \\ 2 \\ 5 \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \end{bmatrix}, \quad B = \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{bmatrix}.$$

The trick is finding a matrix for the X variables that works within a matrix equation that is equivalent to the regression model. The best approach is to set X to be the following (10×3) matrix,

$$X = \begin{bmatrix} 1 & 0 & 8 \\ 1 & 0 & 6 \\ 1 & 1 & 7 \\ 1 & 1 & 8 \\ 1 & 1 & 3 \\ 1 & 0 & 4 \\ 1 & 1 & 6 \\ 1 & 0 & 3 \\ 1 & 1 & 5 \\ 1 & 0 & 4 \end{bmatrix},$$

where the column of all 1s is designed to work with the constant, as we will see below. How were you supposed to know that? Hopefully you gave this problem a lot of thought, and realized that accounting for the constant in the regression equation is the tricky part. If you tried several different approaches, then you are thinking about this problem correctly. Moreover, you are thinking about designing matrices for specific data applications, which is the very best way for a social scientist to think about matrices.

Now that we have defined vectors and matrices for every component of the regression, we can write the regression model as

$$Y = XB + \epsilon.$$

To check whether this equation is conformable, note that X is (10×3) and B is (3×1), so their product is conformable and (10×1). This product is itself conformable to add to the (10×1) vector of errors, and the sum is (10×1), which corresponds to the vector Y which is also (10×1). If we replace these terms with the actual data, the equation becomes

$$\begin{bmatrix} 9 \\ 5 \\ 4 \\ 7 \\ 4 \\ 6 \\ 5 \\ 7 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 8 \\ 1 & 0 & 6 \\ 1 & 1 & 7 \\ 1 & 1 & 8 \\ 1 & 1 & 3 \\ 1 & 0 & 4 \\ 1 & 1 & 6 \\ 1 & 0 & 3 \\ 1 & 1 & 5 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \end{bmatrix},$$

which when multiplied out is the following system of equations:

$$\begin{cases} 9 = 1\alpha + 0\beta_1 + 8\beta_2 + \varepsilon_1, \\ 5 = 1\alpha + 0\beta_1 + 6\beta_2 + \varepsilon_2, \\ 4 = 1\alpha + 1\beta_1 + 7\beta_2 + \varepsilon_3, \\ 7 = 1\alpha + 1\beta_1 + 8\beta_2 + \varepsilon_4, \\ 4 = 1\alpha + 1\beta_1 + 3\beta_2 + \varepsilon_5, \\ 6 = 1\alpha + 0\beta_1 + 4\beta_2 + \varepsilon_6, \\ 5 = 1\alpha + 1\beta_1 + 6\beta_2 + \varepsilon_7, \\ 7 = 1\alpha + 0\beta_1 + 3\beta_2 + \varepsilon_8, \\ 2 = 1\alpha + 1\beta_1 + 5\beta_2 + \varepsilon_9, \\ 5 = 1\alpha + 0\beta_1 + 4\beta_2 + \varepsilon_{10}. \end{cases}$$

Each one of these equations is another realization of the familiar linear regression equation. When we use matrices instead, we can work with all of them simultaneously with the simple equation listed above.

9 Matrix Inverses, Singularity, and Rank

1. There's really no way around tedious work to solve this problem. Either we have to take the inverse of a (4×4) matrix, or we can do slightly less work by showing that

$$\begin{bmatrix} 2 & 10 & 2 & -3 \\ 8 & -4 & 5 & -8 \\ 0 & 5 & -1 & 6 \\ -8 & 4 & -4 & 3 \end{bmatrix} \frac{1}{11} \begin{bmatrix} -44.5 & 138 & 99 & 125.5 \\ -4 & 15 & 11 & 14 \\ 100 & -309 & -220 & -284 \\ 20 & -64 & -44 & -59 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and also that

$$\frac{1}{11} \begin{bmatrix} -44.5 & 138 & 99 & 125.5 \\ -4 & 15 & 11 & 14 \\ 100 & -309 & -220 & -284 \\ 20 & -64 & -44 & -59 \end{bmatrix} \begin{bmatrix} 2 & 10 & 2 & -3 \\ 8 & -4 & 5 & -8 \\ 0 & 5 & -1 & 6 \\ -8 & 4 & -4 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Let's try this latter approach. First consider the product

$$\begin{bmatrix} 2 & 10 & 2 & -3 \\ 8 & -4 & 5 & -8 \\ 0 & 5 & -1 & 6 \\ -8 & 4 & -4 & 3 \end{bmatrix} \frac{1}{11} \begin{bmatrix} -44.5 & 138 & 99 & 125.5 \\ -4 & 15 & 11 & 14 \\ 100 & -309 & -220 & -284 \\ 20 & -64 & -44 & -59 \end{bmatrix}.$$

For matrices, the order in which the matrices are multiplied matters, but for scalars the order still doesn't matter. So we can bring the factor $\frac{1}{11}$ to the front,

$$\frac{1}{11} \begin{bmatrix} 2 & 10 & 2 & -3 \\ 8 & -4 & 5 & -8 \\ 0 & 5 & -1 & 6 \\ -8 & 4 & -4 & 3 \end{bmatrix} \begin{bmatrix} -44.5 & 138 & 99 & 125.5 \\ -4 & 15 & 11 & 14 \\ 100 & -309 & -220 & -284 \\ 20 & -64 & -44 & -59 \end{bmatrix}.$$

This product multiplies a (4×4) matrix by another (4×4) matrix and results in a (4×4) matrix. We also have to divide every element of the product by 11. The elements of this product are

$$\begin{aligned}
(1,1) \text{ element} &= \tfrac{1}{11}\Big((2 \times -44.5) + (10 \times -4) + (2 \times 100) + (-3 \times 20)\Big) &= 1 \\
(1,2) \text{ element} &= \tfrac{1}{11}\Big((2 \times 138) + (10 \times 15) + (2 \times -309) + (-3 \times -64)\Big) &= 0 \\
(1,3) \text{ element} &= \tfrac{1}{11}\Big((2 \times 99) + (10 \times 11) + (2 \times -220) + (-3 \times -44)\Big) &= 0 \\
(1,4) \text{ element} &= \tfrac{1}{11}\Big((2 \times 125.5) + (10 \times 14) + (2 \times -284) + (-3 \times -59)\Big) &= 0 \\
(2,1) \text{ element} &= \tfrac{1}{11}\Big((8 \times -44,5) + (-4 \times -4) + (5 \times 100) + (-8 \times 20)\Big) &= 0 \\
(2,2) \text{ element} &= \tfrac{1}{11}\Big((8 \times 138) + (-4 \times 15) + (5 \times -309) + (-8 \times -64)\Big) &= 1 \\
(2,3) \text{ element} &= \tfrac{1}{11}\Big((8 \times 99) + (-4 \times 11) + (5 \times -220) + (-8 \times -44)\Big) &= 0 \\
(2,4) \text{ element} &= \tfrac{1}{11}\Big((8 \times 125.5) + (-4 \times 14) + (5 \times -284) + (-8 \times -59)\Big) &= 0 \\
(3,1) \text{ element} &= \tfrac{1}{11}\Big((0 \times -44.5) + (5 \times -4) + (-1 \times 100) + (6 \times 20)\Big) &= 0 \\
(3,2) \text{ element} &= \tfrac{1}{11}\Big((0 \times 138) + (5 \times 15) + (-1 \times -309) + (6 \times -64)\Big) &= 0 \\
(3,3) \text{ element} &= \tfrac{1}{11}\Big((0 \times 99) + (5 \times 11) + (-1 \times -220) + (6 \times -44)\Big) &= 1 \\
(3,4) \text{ element} &= \tfrac{1}{11}\Big((0 \times 125.5) + (5 \times 14) + (-1 \times -284) + (6 \times -59)\Big) &= 0 \\
(4,1) \text{ element} &= \tfrac{1}{11}\Big((-8 \times -44.5) + (4 \times -4) + (-4 \times 100) + (3 \times 20)\Big) &= 0 \\
(4,2) \text{ element} &= \tfrac{1}{11}\Big((-8 \times 138) + (4 \times 15) + (-4 \times -309) + (3 \times -64)\Big) &= 0 \\
(4,3) \text{ element} &= \tfrac{1}{11}\Big((-8 \times 99) + (4 \times 11) + (-4 \times -220) + (3 \times -44)\Big) &= 0 \\
(4,4) \text{ element} &= \tfrac{1}{11}\Big((-8 \times 125.5) + (4 \times 14) + (-4 \times -284) + (3 \times -59)\Big) &= 1
\end{aligned}$$

Therefore the product is equal to

$$\begin{bmatrix} 2 & 10 & 2 & -3 \\ 8 & -4 & 5 & -8 \\ 0 & 5 & -1 & 6 \\ -8 & 4 & -4 & 3 \end{bmatrix} \frac{1}{11} \begin{bmatrix} -44.5 & 138 & 99 & 125.5 \\ -4 & 15 & 11 & 14 \\ 100 & -309 & -220 & -284 \\ 20 & -64 & -44 & -59 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and the first matrix the the left-inverse of the second. But to prove that these two matrices are full inverses of each other, we also have to compute

$$\frac{1}{11} \begin{bmatrix} -44.5 & 138 & 99 & 125.5 \\ -4 & 15 & 11 & 14 \\ 100 & -309 & -220 & -284 \\ 20 & -64 & -44 & -59 \end{bmatrix} \begin{bmatrix} 2 & 10 & 2 & -3 \\ 8 & -4 & 5 & -8 \\ 0 & 5 & -1 & 6 \\ -8 & 4 & -4 & 3 \end{bmatrix}.$$

Again, this product multiplies a (4×4) matrix by another (4×4) matrix and results in a (4×4) matrix. We also have to divide every element of the product by 11. The elements of this product are

$$
\begin{aligned}
(1,1) \text{ element} &= \tfrac{1}{11}\Big((-44.5 \times 2) + (138 \times 8) + (99 \times 0) + (125.5 \times -8)\Big) &= 1 \\
(1,2) \text{ element} &= \tfrac{1}{11}\Big((-44.5 \times 10) + (138 \times -4) + (99 \times 5) + (125.5 \times 4)\Big) &= 0 \\
(1,3) \text{ element} &= \tfrac{1}{11}\Big((-44.5 \times 2) + (138 \times 5) + (99 \times -1) + (125.5 \times -4)\Big) &= 0 \\
(1,4) \text{ element} &= \tfrac{1}{11}\Big((-44.5 \times -3) + (138 \times -8) + (99 \times 6) + (125.5 \times 3)\Big) &= 0 \\
(2,1) \text{ element} &= \tfrac{1}{11}\Big((-4 \times 2) + (15 \times 8) + (11 \times 0) + (14 \times -8)\Big) &= 0 \\
(2,2) \text{ element} &= \tfrac{1}{11}\Big((-4 \times 10) + (15 \times -4) + (11 \times 5) + (14 \times 4)\Big) &= 1 \\
(2,3) \text{ element} &= \tfrac{1}{11}\Big((-4 \times 2) + (15 \times 5) + (11 \times -1) + (14 \times -4)\Big) &= 0 \\
(2,4) \text{ element} &= \tfrac{1}{11}\Big((-4 \times -3) + (15 \times -8) + (11 \times 6) + (14 \times 3)\Big) &= 0 \\
(3,1) \text{ element} &= \tfrac{1}{11}\Big((100 \times 2) + (-309 \times 8) + (-220 \times 0) + (-284 \times -8)\Big) &= 0 \\
(3,2) \text{ element} &= \tfrac{1}{11}\Big((100 \times 10) + (-309 \times -4) + (-220 \times 5) + (-284 \times 4)\Big) &= 0 \\
(3,3) \text{ element} &= \tfrac{1}{11}\Big((100 \times 2) + (-309 \times 5) + (-220 \times -1) + (-284 \times -4)\Big) &= 1 \\
(3,4) \text{ element} &= \tfrac{1}{11}\Big((100 \times -3) + (-309 \times -8) + (-220 \times 6) + (-284 \times 3)\Big) &= 0 \\
(4,1) \text{ element} &= \tfrac{1}{11}\Big((20 \times 2) + (-64 \times 8) + (-44 \times 0) + (-59 \times -8)\Big) &= 0 \\
(4,2) \text{ element} &= \tfrac{1}{11}\Big((20 \times 10) + (-64 \times -4) + (-44 \times 5) + (-59 \times 4)\Big) &= 0 \\
(4,3) \text{ element} &= \tfrac{1}{11}\Big((20 \times 2) + (-64 \times 5) + (-44 \times -1) + (-59 \times -4)\Big) &= 0 \\
(4,4) \text{ element} &= \tfrac{1}{11}\Big((20 \times -3) + (-64 \times -8) + (-44 \times 6) + (-59 \times 3)\Big) &= 1
\end{aligned}
$$

So this product is also equal to the identity matrix,

$$
\frac{1}{11}\begin{bmatrix} -44.5 & 138 & 99 & 125.5 \\ -4 & 15 & 11 & 14 \\ 100 & -309 & -220 & -284 \\ 20 & -64 & -44 & -59 \end{bmatrix}\begin{bmatrix} 2 & 10 & 2 & -3 \\ 8 & -4 & 5 & -8 \\ 0 & 5 & -1 & 6 \\ -8 & 4 & -4 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},
$$

and the two matrices are full inverses of each other.

2. We find the inverse of each of the following (2×2) matrices with elements denoted

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

by plugging their elements into the formula for a (2×2) inverse,

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

The inverse doesn't exist when $ad - bc = 0$.

(a) $A^{-1} = \frac{1}{(2\times 4)-(7\times 1)}\begin{bmatrix} 4 & -7 \\ -1 & 2 \end{bmatrix} = \frac{1}{1}\begin{bmatrix} 4 & -7 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & -7 \\ -1 & 2 \end{bmatrix}.$

(b) $A^{-1} = \frac{1}{(4\times 5)-(5\times 3)}\begin{bmatrix} 5 & -3 \\ -5 & 4 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 5 & -3 \\ -5 & 4 \end{bmatrix} = \begin{bmatrix} 1 & -\frac{3}{5} \\ -1 & \frac{4}{5} \end{bmatrix}.$

(c) $A^{-1} = \frac{1}{(4\times -5)-(10\times -2)}\begin{bmatrix} -5 & -10 \\ 2 & 4 \end{bmatrix} = \frac{1}{\mathbf{0}}\begin{bmatrix} -5 & -10 \\ 2 & 4 \end{bmatrix}$. This inverse does not exist.

(d) $A^{-1} = \frac{1}{(-10\times -1)-(2\times 8)}\begin{bmatrix} -1 & -2 \\ -8 & -10 \end{bmatrix} = -\frac{1}{6}\begin{bmatrix} -1 & -2 \\ -8 & -10 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} \\ \frac{4}{3} & \frac{5}{3} \end{bmatrix}.$

(e) $A^{-1} = \frac{1}{(5\times 6)-(15\times 2)}\begin{bmatrix} 6 & -2 \\ -15 & 5 \end{bmatrix} = \frac{1}{\mathbf{0}}\begin{bmatrix} 6 & -2 \\ -15 & 5 \end{bmatrix}$. This inverse does not exist.

(f) $A^{-1} = \frac{1}{(-2\times -6)-(6\times -7)}\begin{bmatrix} -6 & -6 \\ 7 & -2 \end{bmatrix} = \frac{1}{54}\begin{bmatrix} -6 & -6 \\ 7 & -2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{9} & -\frac{1}{9} \\ \frac{7}{54} & -\frac{1}{27} \end{bmatrix}.$

(g) $A^{-1} = \frac{1}{(9\times 8)-(7\times 2)}\begin{bmatrix} 8 & -7 \\ -2 & 9 \end{bmatrix} = \frac{1}{58}\begin{bmatrix} 8 & -7 \\ -2 & 9 \end{bmatrix} = \begin{bmatrix} \frac{4}{29} & -\frac{7}{58} \\ -\frac{1}{29} & \frac{9}{58} \end{bmatrix}.$

(h) $A^{-1} = \dfrac{1}{(-5 \times 1) - (0 \times -7)} \begin{bmatrix} 1 & 0 \\ 7 & -5 \end{bmatrix} = -\dfrac{1}{5} \begin{bmatrix} 1 & 0 \\ 7 & -5 \end{bmatrix} = \begin{bmatrix} -\frac{1}{5} & 0 \\ -\frac{7}{5} & 1 \end{bmatrix}.$

3. (a) We first compute the minor elements for each position in the matrix by removing the indicated row and column and taking the determinant of the remaining (2×2) matrix. In this case, the minor elements are

$$\begin{aligned}
M_{11} &= \left| \begin{bmatrix} 4 & 6 \\ 0 & 5 \end{bmatrix} \right| &&= (4 \times 5) - (6 \times 0) &&= 20 \\
M_{12} &= \left| \begin{bmatrix} -3 & 6 \\ -9 & 5 \end{bmatrix} \right| &&= (-3 \times 5) - (6 \times -9) &&= 39 \\
M_{13} &= \left| \begin{bmatrix} -3 & 4 \\ -9 & 0 \end{bmatrix} \right| &&= (-3 \times 0) - (4 \times -9) &&= 36 \\
M_{21} &= \left| \begin{bmatrix} 6 & 6 \\ 0 & 5 \end{bmatrix} \right| &&= (6 \times 5) - (6 \times 0) &&= 30 \\
M_{22} &= \left| \begin{bmatrix} 3 & 6 \\ -9 & 5 \end{bmatrix} \right| &&= (3 \times 5) - (6 \times -9) &&= 69 \\
M_{23} &= \left| \begin{bmatrix} 3 & 6 \\ -9 & 0 \end{bmatrix} \right| &&= (3 \times 0) - (6 \times -9) &&= 54 \\
M_{31} &= \left| \begin{bmatrix} 6 & 6 \\ 4 & 6 \end{bmatrix} \right| &&= (6 \times 6) - (6 \times 4) &&= 12 \\
M_{32} &= \left| \begin{bmatrix} 3 & 6 \\ -3 & 6 \end{bmatrix} \right| &&= (3 \times 6) - (6 \times -3) &&= 36 \\
M_{33} &= \left| \begin{bmatrix} 3 & 6 \\ -3 & 4 \end{bmatrix} \right| &&= (3 \times 4) - (6 \times -3) &&= 30,
\end{aligned}$$

the cofactors are

$$\begin{aligned}
C_{11} = (-1)^{1+1} M_{11} = 20, \quad & C_{12} = (-1)^{1+2} M_{12} = -39, \quad & C_{13} = (-1)^{3+1} M_{31} = 36, \\
C_{21} = (-1)^{2+1} M_{21} = -30, \quad & C_{22} = (-1)^{2+2} M_{22} = 69, \quad & C_{23} = (-1)^{3+2} M_{32} = -54, \\
C_{31} = (-1)^{3+1} M_{31} = 12, \quad & C_{32} = (-1)^{3+2} M_{32} = -36, \quad & C_{33} = (-1)^{3+3} M_{33} = 30,
\end{aligned}$$

and the cofactor matrix is therefore

$$C = \begin{bmatrix} 20 & -39 & 36 \\ -30 & 69 & -54 \\ 12 & -36 & 30 \end{bmatrix}.$$

Then the adjoint matrix is simply the transpose of the cofactor matrix,

$$C' = \begin{bmatrix} 20 & -30 & 12 \\ -39 & 69 & -36 \\ 36 & -54 & 30 \end{bmatrix}.$$

To find the determinant, we choose one row or column of the original matrix, multiply every element by its corresponding cofactor, and add. Using the first row gives us a determinant of

$$(3 \times 20) + (6 \times -39) + (6 \times 36) = 42.$$

The inverse is the adjoint matrix scalar multiplied by the reciprocal of the determinant,

$$\frac{1}{42}\begin{bmatrix} 20 & -30 & 12 \\ -39 & 69 & -36 \\ 36 & -54 & 30 \end{bmatrix} = \begin{bmatrix} 0.48 & -0.71 & 0.29 \\ -0.93 & 1.64 & -0.86 \\ 0.86 & -1.29 & 0.71 \end{bmatrix}.$$

(b) In this case, the minor elements are

$$\begin{aligned}
M_{11} &= \left|\begin{bmatrix} -8 & 1 \\ -2 & 3 \end{bmatrix}\right| &&= (-8 \times 3) - (1 \times -2) &&= -22 \\
M_{12} &= \left|\begin{bmatrix} 9 & 1 \\ 0 & 3 \end{bmatrix}\right| &&= (9 \times 3) - (1 \times 0) &&= 27 \\
M_{13} &= \left|\begin{bmatrix} 9 & -8 \\ 0 & -2 \end{bmatrix}\right| &&= (9 \times -2) - (-8 \times 0) &&= -18 \\
M_{21} &= \left|\begin{bmatrix} 10 & -5 \\ -2 & 3 \end{bmatrix}\right| &&= (10 \times 3) - (-5 \times -2) &&= 20 \\
M_{22} &= \left|\begin{bmatrix} -2 & -5 \\ 0 & 3 \end{bmatrix}\right| &&= (-2 \times 3) - (-5 \times 0) &&= -6 \\
M_{23} &= \left|\begin{bmatrix} -2 & 10 \\ 0 & -2 \end{bmatrix}\right| &&= (-2 \times -2) - (10 \times 0) &&= 4 \\
M_{31} &= \left|\begin{bmatrix} 10 & -5 \\ -8 & 1 \end{bmatrix}\right| &&= (10 \times 1) - (-5 \times -8) &&= -30 \\
M_{32} &= \left|\begin{bmatrix} -2 & -5 \\ 9 & 1 \end{bmatrix}\right| &&= (-2 \times 1) - (-5 \times 9) &&= 43 \\
M_{33} &= \left|\begin{bmatrix} -2 & 10 \\ 9 & -8 \end{bmatrix}\right| &&= (-2 \times -8) - (10 \times 9) &&= -74,
\end{aligned}$$

the cofactors are

$$C_{11} = (-1)^{1+1}M_{11} = -22, \quad C_{12} = (-1)^{1+2}M_{12} = -27, \quad C_{13} = (-1)^{3+1}M_{31} = -18,$$

$$C_{21} = (-1)^{2+1}M_{21} = -20, \quad C_{22} = (-1)^{2+2}M_{22} = -6, \quad C_{23} = (-1)^{3+2}M_{32} = -4,$$

$$C_{31} = (-1)^{3+1}M_{31} = -30, \quad C_{32} = (-1)^{3+2}M_{32} = -43, \quad C_{33} = (-1)^{3+3}M_{33} = -74,$$

and the cofactor matrix is therefore

$$C = \begin{bmatrix} -22 & -27 & -18 \\ -20 & -6 & -4 \\ -30 & -43 & -74 \end{bmatrix}.$$

Then the adjoint matrix is simply the transpose of the cofactor matrix,

$$C' = \begin{bmatrix} -22 & -20 & -30 \\ -27 & -6 & -43 \\ -18 & -4 & -74 \end{bmatrix}.$$

To find the determinant, we choose one row or column of the original matrix, multiply every element by its corresponding cofactor, and add. Using the first row gives us a determinant of

$$(-2 \times -22) + (10 \times -27) + (-5 \times -18) = -136.$$

The inverse is the adjoint matrix scalar multiplied by the reciprocal of the determinant,

$$-\frac{1}{136} \begin{bmatrix} -22 & -20 & -30 \\ -27 & -6 & -43 \\ -18 & -4 & -74 \end{bmatrix} = \begin{bmatrix} 0.16 & 0.15 & 0.22 \\ 0.20 & 0.04 & 0.32 \\ 0.13 & 0.03 & 0.54 \end{bmatrix}.$$

(c) In this case, the minor elements are

$$\begin{aligned}
M_{11} &= \left| \begin{bmatrix} 4 & 5 \\ 0 & -5 \end{bmatrix} \right| &&= (4 \times -5) - (5 \times 0) &&= -20 \\
M_{12} &= \left| \begin{bmatrix} -10 & 5 \\ 4 & -5 \end{bmatrix} \right| &&= (-10 \times -5) - (5 \times 4) &&= 30 \\
M_{13} &= \left| \begin{bmatrix} -10 & 4 \\ 4 & 0 \end{bmatrix} \right| &&= (-10 \times 0) - (4 \times 4) &&= -16 \\
M_{21} &= \left| \begin{bmatrix} 9 & 0 \\ 0 & -5 \end{bmatrix} \right| &&= (9 \times -5) - (0 \times 0) &&= -45 \\
M_{22} &= \left| \begin{bmatrix} 7 & 0 \\ 4 & -5 \end{bmatrix} \right| &&= (7 \times -5) - (0 \times 4) &&= -35 \\
M_{23} &= \left| \begin{bmatrix} 7 & 9 \\ 4 & 0 \end{bmatrix} \right| &&= (7 \times 0) - (9 \times 4) &&= -36 \\
M_{31} &= \left| \begin{bmatrix} 9 & 0 \\ 4 & 5 \end{bmatrix} \right| &&= (9 \times 5) - (0 \times 4) &&= 45 \\
M_{32} &= \left| \begin{bmatrix} 7 & 0 \\ -10 & 5 \end{bmatrix} \right| &&= (7 \times 5) - (0 \times -10) &&= 35 \\
M_{33} &= \left| \begin{bmatrix} 7 & 9 \\ -10 & 4 \end{bmatrix} \right| &&= (7 \times 4) - (9 \times -10) &&= 118,
\end{aligned}$$

the cofactors are

$$C_{11} = (-1)^{1+1} M_{11} = -20, \quad C_{12} = (-1)^{1+2} M_{12} = -30, \quad C_{13} = (-1)^{3+1} M_{31} = -16,$$

$$C_{21} = (-1)^{2+1} M_{21} = 45, \quad C_{22} = (-1)^{2+2} M_{22} = -35, \quad C_{23} = (-1)^{3+2} M_{32} = -36,$$

$$C_{31} = (-1)^{3+1} M_{31} = 45, \quad C_{32} = (-1)^{3+2} M_{32} = -35, \quad C_{33} = (-1)^{3+3} M_{33} = 118,$$

and the cofactor matrix is therefore

$$C = \begin{bmatrix} -20 & -30 & -16 \\ 45 & -35 & -36 \\ 45 & -35 & 118 \end{bmatrix}.$$

Then the adjoint matrix is simply the transpose of the cofactor matrix,

$$C' = \begin{bmatrix} -20 & 45 & 45 \\ -30 & -35 & -35 \\ -16 & -36 & 118 \end{bmatrix}.$$

To find the determinant, we choose one row or column of the original matrix, multiply every element by its corresponding cofactor, and add. Using the first row gives us a determinant of

$$(7 \times -20) + (9 \times -30) + (0 \times -16) = -410.$$

The inverse is the adjoint matrix scalar multiplied by the reciprocal of the determinant,

$$-\frac{1}{410}\begin{bmatrix} -20 & 45 & 45 \\ -30 & -35 & -35 \\ -16 & -36 & 118 \end{bmatrix} = \begin{bmatrix} 0.05 & -0.11 & -0.11 \\ 0.07 & 0.09 & 0.09 \\ 0.04 & -0.09 & -0.29 \end{bmatrix}.$$

(d) In this case, the minor elements are

$$\begin{aligned}
M_{11} &= \left|\begin{bmatrix} -7 & 2 \\ 7 & 2 \end{bmatrix}\right| &&= (-7 \times 2) - (2 \times 7) &&= -28 \\
M_{12} &= \left|\begin{bmatrix} 3 & 2 \\ 7 & 2 \end{bmatrix}\right| &&= (3 \times 2) - (2 \times 7) &&= -8 \\
M_{13} &= \left|\begin{bmatrix} 3 & -7 \\ 7 & 7 \end{bmatrix}\right| &&= (3 \times 7) - (-7 \times 7) &&= 70 \\
M_{21} &= \left|\begin{bmatrix} 1 & -1 \\ 7 & 2 \end{bmatrix}\right| &&= (1 \times 2) - (-1 \times 7) &&= 9 \\
M_{22} &= \left|\begin{bmatrix} 9 & -1 \\ 7 & 2 \end{bmatrix}\right| &&= (9 \times 2) - (-1 \times 7) &&= 25 \\
M_{23} &= \left|\begin{bmatrix} 9 & 1 \\ 7 & 7 \end{bmatrix}\right| &&= (9 \times 7) - (1 \times 7) &&= 56 \\
M_{31} &= \left|\begin{bmatrix} 1 & -1 \\ -7 & 2 \end{bmatrix}\right| &&= (1 \times 2) - (-1 \times -7) &&= -5 \\
M_{32} &= \left|\begin{bmatrix} 9 & -1 \\ 3 & 2 \end{bmatrix}\right| &&= (9 \times 2) - (-1 \times 3) &&= 21 \\
M_{33} &= \left|\begin{bmatrix} 9 & 1 \\ 3 & -7 \end{bmatrix}\right| &&= (9 \times -7) - (1 \times 3) &&= -66,
\end{aligned}$$

the cofactors are

$$C_{11} = (-1)^{1+1}M_{11} = -28, \qquad C_{12} = (-1)^{1+2}M_{12} = 8, \qquad C_{13} = (-1)^{3+1}M_{31} = 70,$$

$$C_{21} = (-1)^{2+1}M_{21} = -9, \qquad C_{22} = (-1)^{2+2}M_{22} = 25, \quad C_{23} = (-1)^{3+2}M_{32} = -56,$$

$$C_{31} = (-1)^{3+1}M_{31} = -5, \quad C_{32} = (-1)^{3+2}M_{32} = -21, \quad C_{33} = (-1)^{3+3}M_{33} = -66,$$

and the cofactor matrix is therefore

$$C = \begin{bmatrix} -28 & 8 & 70 \\ -9 & 25 & -56 \\ -5 & -21 & -66 \end{bmatrix}.$$

Then the adjoint matrix is simply the transpose of the cofactor matrix,

$$C' = \begin{bmatrix} -28 & -9 & -5 \\ 8 & 25 & -21 \\ 70 & -56 & -66 \end{bmatrix}.$$

To find the determinant, we choose one row or column of the original matrix, multiply every element by its corresponding cofactor, and add. Using the first row gives us a determinant of

$$(9 \times -28) + (1 \times 8) + (-1 \times 70) = -314.$$

The inverse is the adjoint matrix scalar multiplied by the reciprocal of the determinant,

$$-\frac{1}{314}\begin{bmatrix} -28 & -9 & -5 \\ 8 & 25 & -21 \\ 70 & -56 & -66 \end{bmatrix} = \begin{bmatrix} 0.09 & 0.03 & 0.02 \\ -0.03 & -0.08 & 0.07 \\ -0.22 & 0.18 & 0.21 \end{bmatrix}.$$

(e) In this case, the minor elements are

$$\begin{aligned}
M_{11} &= \left|\begin{bmatrix} -2 & 4 \\ 2 & 6 \end{bmatrix}\right| &&= (-2 \times 6) - (4 \times 2) &&= -20 \\
M_{12} &= \left|\begin{bmatrix} -3 & 4 \\ 2 & 6 \end{bmatrix}\right| &&= (-3 \times 6) - (4 \times 2) &&= -26 \\
M_{13} &= \left|\begin{bmatrix} -3 & -2 \\ 2 & 2 \end{bmatrix}\right| &&= (-3 \times 2) - (-2 \times 2) &&= -2 \\
M_{21} &= \left|\begin{bmatrix} 6 & 10 \\ 2 & 6 \end{bmatrix}\right| &&= (6 \times 6) - (10 \times 2) &&= 16 \\
M_{22} &= \left|\begin{bmatrix} -5 & 10 \\ 2 & 6 \end{bmatrix}\right| &&= (-5 \times 6) - (10 \times 2) &&= -50 \\
M_{23} &= \left|\begin{bmatrix} -5 & 6 \\ 2 & 2 \end{bmatrix}\right| &&= (-5 \times 2) - (6 \times 2) &&= -22 \\
M_{31} &= \left|\begin{bmatrix} 6 & 10 \\ -2 & 4 \end{bmatrix}\right| &&= (6 \times 4) - (10 \times -2) &&= 44 \\
M_{32} &= \left|\begin{bmatrix} -5 & 10 \\ -3 & 4 \end{bmatrix}\right| &&= (-5 \times 4) - (10 \times -3) &&= 10 \\
M_{33} &= \left|\begin{bmatrix} -5 & 6 \\ -3 & -2 \end{bmatrix}\right| &&= (-5 \times -2) - (6 \times -3) &&= 28,
\end{aligned}$$

the cofactors are

$$C_{11} = (-1)^{1+1}M_{11} = -20, \quad C_{12} = (-1)^{1+2}M_{12} = 26, \quad C_{13} = (-1)^{3+1}M_{31} = -2,$$

$$C_{21} = (-1)^{2+1}M_{21} = -16, \quad C_{22} = (-1)^{2+2}M_{22} = -50, \quad C_{23} = (-1)^{3+2}M_{32} = 22,$$

$$C_{31} = (-1)^{3+1}M_{31} = 44, \quad C_{32} = (-1)^{3+2}M_{32} = -10, \quad C_{33} = (-1)^{3+3}M_{33} = 28,$$

and the cofactor matrix is therefore

$$C = \begin{bmatrix} -20 & 26 & -2 \\ -16 & -50 & 22 \\ 44 & -10 & 28 \end{bmatrix}.$$

Then the adjoint matrix is simply the transpose of the cofactor matrix,

$$C' = \begin{bmatrix} -20 & -16 & 44 \\ 26 & -50 & -10 \\ -2 & 22 & 28 \end{bmatrix}.$$

To find the determinant, we choose one row or column of the original matrix, multiply every element by its corresponding cofactor, and add. Using the first row gives us a determinant of

$$(-5 \times -20) + (6 \times 26) + (10 \times -2) = 236.$$

The inverse is the adjoint matrix scalar multiplied by the reciprocal of the determinant,

$$-\frac{1}{236}\begin{bmatrix} -20 & -16 & 44 \\ 26 & -50 & -10 \\ -2 & 22 & 28 \end{bmatrix} = \begin{bmatrix} -0.08 & -0.07 & 0.19 \\ 0.11 & -0.21 & -0.04 \\ -0.01 & 0.09 & 0.12 \end{bmatrix}.$$

(f) In this case, the minor elements are

$$
\begin{aligned}
M_{11} &= \left|\begin{bmatrix} -3 & -5 \\ 0 & 4 \end{bmatrix}\right| &= (-3 \times 4) - (-5 \times 0) &= -12 \\
M_{12} &= \left|\begin{bmatrix} 0 & -5 \\ -9 & 4 \end{bmatrix}\right| &= (0 \times 4) - (-5 \times -9) &= -45 \\
M_{13} &= \left|\begin{bmatrix} 0 & -3 \\ -9 & 0 \end{bmatrix}\right| &= (0 \times 0) - (-3 \times -9) &= -27 \\
M_{21} &= \left|\begin{bmatrix} 8 & 6 \\ 0 & 4 \end{bmatrix}\right| &= (8 \times 4) - (6 \times 0) &= 32 \\
M_{22} &= \left|\begin{bmatrix} 3 & 6 \\ -9 & 4 \end{bmatrix}\right| &= (3 \times 4) - (6 \times -9) &= 66 \\
M_{23} &= \left|\begin{bmatrix} 3 & 8 \\ -9 & 0 \end{bmatrix}\right| &= (3 \times 0) - (8 \times -9) &= 72 \\
M_{31} &= \left|\begin{bmatrix} 8 & 6 \\ -3 & -5 \end{bmatrix}\right| &= (8 \times -5) - (6 \times -3) &= -22 \\
M_{32} &= \left|\begin{bmatrix} 3 & 6 \\ 0 & -5 \end{bmatrix}\right| &= (3 \times -5) - (6 \times 0) &= -15 \\
M_{33} &= \left|\begin{bmatrix} 3 & 8 \\ 0 & -3 \end{bmatrix}\right| &= (3 \times -3) - (8 \times 0) &= -9
\end{aligned}
$$

the cofactors are

$$
\begin{aligned}
C_{11} = (-1)^{1+1}M_{11} = -12, \quad & C_{12} = (-1)^{1+2}M_{12} = 45, \quad & C_{13} = (-1)^{3+1}M_{31} = -27, \\
C_{21} = (-1)^{2+1}M_{21} = 32, \quad & C_{22} = (-1)^{2+2}M_{22} = 66, \quad & C_{23} = (-1)^{3+2}M_{32} = -72, \\
C_{31} = (-1)^{3+1}M_{31} = -22, \quad & C_{32} = (-1)^{3+2}M_{32} = 15, \quad & C_{33} = (-1)^{3+3}M_{33} = -9,
\end{aligned}
$$

and the cofactor matrix is therefore

$$
C = \begin{bmatrix} -12 & 45 & -27 \\ 32 & 66 & -72 \\ -22 & 15 & -9 \end{bmatrix}.
$$

Then the adjoint matrix is simply the transpose of the cofactor matrix,

$$
C' = \begin{bmatrix} -12 & 32 & -22 \\ 45 & 66 & -72 \\ -27 & -72 & -9 \end{bmatrix}.
$$

To find the determinant, we choose one row or column of the original matrix, multiply every element by its corresponding cofactor, and add. Using the first row gives us a determinant of

$$
(3 \times -12) + (8 \times 45) + (6 \times -27) = 162.
$$

The inverse is the adjoint matrix scalar multiplied by the reciprocal of the determinant,

$$
-\frac{1}{162}\begin{bmatrix} -12 & 32 & -22 \\ 45 & 66 & -72 \\ -27 & -72 & -9 \end{bmatrix} = \begin{bmatrix} -0.07 & -0.20 & -0.14 \\ 0.28 & 0.41 & 0.09 \\ -0.17 & -0.44 & -0.06 \end{bmatrix}.
$$

4. (a) To perform Sarrus' rule, write the matrix twice, side by side. Then for every element in the first row of the left matrix, draw an oval moving down and to the right that contains three elements. Multiply the elements in each oval together. In this case, these products are

$$\begin{aligned}(6 \times 4 \times 6) &= 162,\\ (-9 \times 10 \times -9) &= 144,\\ (-8 \times -5 \times 8) &= 320.\end{aligned}$$

For every element in the first row of the right matrix, draw an oval moving down and to the left that contains three elements. In this case, these products are

$$\begin{aligned}(6 \times 10 \times 8) &= 480,\\ (-9 \times -5 \times 6) &= 270,\\ (-8 \times 4 \times -9) &= 288.\end{aligned}$$

Finally, we add the three products from the ovals moving down and to the right, and subtract the products from the ovals moving down and to the left. The determinant is:

$$(144 + 810 + 320) - (480 + 270 + 288) = 236.$$

(b) We write the matrix twice, side by side. Then for every element in the first row of the left matrix, we draw an oval moving down and to the right that contains three elements and multiply the elements in each oval together. In this case, these products are

$$\begin{aligned}(-9 \times -3 \times 6) &= 162,\\ (8 \times 4 \times 15) &= 480,\\ (-9 \times 1 \times -7) &= 63.\end{aligned}$$

For every element in the first row of the right matrix, we draw an oval moving down and to the left that contains three elements. In this case, these products are

$$\begin{aligned}(-9 \times 4 \times -7) &= 252,\\ (8 \times 1 \times 6) &= 48,\\ (-9 \times -3 \times 15) &= 405.\end{aligned}$$

Finally, we add the three products from the ovals moving down and to the right, and subtract the products from the ovals moving down and to the left. The determinant is:

$$(162 + 480 + 63) - (252 + 48 + 405) = 0.$$

(c) We write the matrix twice, side by side. Then for every element in the first row of the left matrix, we draw an oval moving down and to the right that contains three elements and multiply the elements in each oval together. In this case, these products are

$$\begin{aligned}(3 \times -6 \times -8) &= 144,\\ (0 \times -6 \times -2) &= 0,\\ (-2 \times 5 \times -4) &= 40.\end{aligned}$$

For every element in the first row of the right matrix, we draw an oval moving down and to the left that contains three elements. In this case, these products are

$$\begin{aligned}(3 \times -6 \times -4) &= 72,\\ (0 \times 5 \times -8) &= 0,\\ (-2 \times -6 \times -2) &= -24.\end{aligned}$$

Finally, we add the three products from the ovals moving down and to the right, and subtract the products from the ovals moving down and to the left. The determinant is:

$$(144 + 0 + 40) - (72 + 0 + -24) = 136.$$

(d) We write the matrix twice, side by side. Then for every element in the first row of the left matrix, we draw an oval moving down and to the right that contains three elements and multiply the elements in each oval together. In this case, these products are

$$\begin{aligned}(-8 \times -8 \times 10) &= 640,\\ (-3 \times 3 \times 0) &= 0,\\ (-8 \times -4 \times -1) &= -32.\end{aligned}$$

For every element in the first row of the right matrix, we draw an oval moving down and to the left that contains three elements. In this case, these products are

$$\begin{aligned}(-8 \times -8 \times 10) &= 640,\\ (-3 \times 3 \times 0) &= 0,\\ (-8 \times -4 \times -1) &= -32.\end{aligned}$$

Finally, we add the three products from the ovals moving down and to the right, and subtract the products from the ovals moving down and to the left. The determinant is:

$$(640 + 0 + -32) - (24 + 120 + 0) = 464.$$

5. A (4×4) matrix is much more unwieldy than a (3×3) matrix. Fortunately, we only have to find the determinant, not the full inverse matrix. The steps we will take are as follows:

   - We have to choose one row or column to work with, so let's work with the first row.
   - We will calculate the minor elements that correspond to each of the four elements on the first row. To find a minor element, we remove the indicated row and column from the (4×4) matrix and find the determinant of the remaining (3×3) matrix. Let's use Sarrus' rule to find those determinants.

- We will find the cofactors.
- We will multiply each element on the first row of the original matrix by its cofactor and add. The sum is the determinant of the (4×4) matrix.

First, we find the four minor elements for the first row. For the (1,1) minor element, we remove the first row and first column of the (4×4) matrix and take the determinant of what remains,

$$M_{11} = \left| \begin{bmatrix} 3 & 0 & 1 \\ -6 & -2 & -4 \\ -6 & 6 & 3 \end{bmatrix} \right|,$$

by applying Sarrus' rule,

$$\begin{aligned} (3 \times -2 \times 3) &= -18, \\ (0 \times -4 \times -6) &= 0, \\ (1 \times -6 \times 6) &= -36, \\ (3 \times -4 \times 6) &= -72, \\ (0 \times -6 \times 3) &= 0, \\ (1 \times -2 \times -6) &= 12, \end{aligned}$$

$$M_{11} = (-18 + 0 + -36) - (-72 + 0 + 12) = 6.$$

For the (1,2) minor element, we remove the first row and second column of the (4×4) matrix and take the determinant of what remains,

$$M_{12} = \left| \begin{bmatrix} -5 & 0 & 1 \\ 7 & -2 & -4 \\ 8 & 6 & 3 \end{bmatrix} \right|,$$

by applying Sarrus' rule,

$$\begin{aligned} (-5 \times -2 \times 3) &= 30, \\ (0 \times -4 \times 8) &= 0, \\ (1 \times 7 \times 6) &= 42, \\ (-5 \times -4 \times 6) &= 120, \\ (0 \times 7 \times 3) &= 0, \\ (1 \times -2 \times 8) &= -16, \end{aligned}$$

$$M_{12} = (30 + 0 + 42) - (120 + 0 + -16) = -32.$$

For the (1,3) minor element, we remove the first row and third column of the (4×4) matrix and take the determinant of what remains,

$$M_{13} = \left| \begin{bmatrix} -5 & 3 & 1 \\ 7 & -6 & -4 \\ 8 & -6 & 3 \end{bmatrix} \right|,$$

by applying Sarrus' rule,

$$\begin{aligned} (-5 \times -6 \times 3) &= 90, \\ (3 \times -4 \times 8) &= -96, \\ (1 \times 7 \times -6) &= -42, \\ (-5 \times -4 \times -6) &= -120, \\ (3 \times 7 \times 3) &= 63, \\ (1 \times -6 \times 8) &= -48, \end{aligned}$$

$$M_{13} = (90 + -96 + -42) - (-120 + 63 + -48) = 57.$$

Finally, for the (1,4) minor element, we remove the first row and third column of the (4×4) matrix and take the determinant of what remains,

$$M_{14} = \left| \begin{bmatrix} -5 & 3 & 0 \\ 7 & -6 & -2 \\ 8 & -6 & 6 \end{bmatrix} \right|,$$

by applying Sarrus' rule,

$$\begin{aligned}
(-5 \times -6 \times 6) &= 180, \\
(3 \times -2 \times 8) &= -48, \\
(0 \times 7 \times -6) &= 0, \\
(-5 \times -2 \times -6) &= -60, \\
(3 \times 7 \times 6) &= 126, \\
(0 \times -6 \times 8) &= 0,
\end{aligned}$$

$$M_{14} = (180 + -48 + 0) - (-60 + 126 + 0) = 66.$$

The cofactors are

$$C_{11} = (-1)^{1+1} M_{11} = 6, \qquad C_{12} = (-1)^{1+2} M_{12} = 32,$$

$$C_{13} = (-1)^{1+3} M_{13} = 57, \quad C_{14} = (-1)^{1+4} M_{14} = -66.$$

And the last step is to multiply the elements on the first row of the original matrix by their cofactors and add. The sum is the determinant of the (4×4) matrix:

$$(-3 \times 6) + (-3 \times 32) + (-1 \times 57) + (-9 \times -66) = 423.$$

6. (a) As we saw in the previous problem, in order to find the determinant of a (4×4) matrix, it is necessary to take the find the determinant of 4 (3×3) matrices. Putting aside the time it takes to identify these matrices, calculate the cofactors, and add the products of the elements on a row and the cofactors, the time it should take a student to take this determinant will be *at least* 4 minutes.

   For a (5×5) matrix, the student can expand along the first row which requires the student to find 5 determinants of (4×4) matrices, each of which takes at least 4 minutes to find. So the calculation will take at least $5 \times 4 = 20$ minutes.

   For a (6×6) matrix, the student can expand along the first row which requires the student to find 6 determinants of (5×5) matrices, each of which takes at least 20 minutes to find. So the calculation will take at least $6 \times 20 = 120$ minutes (2 hours).

   And for a (7×7) matrix, the student can expand along the first row which requires the student to find 7 determinants of (6×6) matrices, each of which takes at least 120 minutes to find. So the calculation will take at least $7 \times 120 = 840$ minutes (14 hours!).

   In contrast, the computer can take a (4×4) determinant in about $4(.1) = .4$ seconds, a (5×5) determinant in about $5(.4) = 2$ seconds, a (6×6) determinant in about $6(2) = 12$ seconds, and a (7×7) determinant in about $7(12) = 84$ seconds.

   (b) The quote by Michael Sand refers to the fact that amount of tedious work involved in performing these matrix arithmetic computations increases exponentially with the size of the matrices. For anything larger than (4×4) the computation is a real project, and anything with 6 or 7 or more dimensions becomes quickly untenable. These computations are not difficult, they are tedious. And humans have severe constraints on the amount of tedious work they are able to perform quickly and accurately. In contrast, a computer is designed to excel at monotonous tasks. This advantage of computers is the primary reason why modern applied statistics in the social sciences relies so heavily on computers.

7. A matrix is singular only when its determinant is zero. Our strategy here will be to find the (2×2) determinants in terms of λ and then to find the value(s) of λ that make the determinant equal 0.

   (a)
$$\left|\begin{bmatrix} 8-\lambda & 7 \\ 7 & 8-\lambda \end{bmatrix}\right| = 0,$$

$$(8-\lambda)^2 - 49 = 0,$$

$$(64 - 16\lambda + \lambda^2) - 49 = 0,$$

$$\lambda^2 - 16\lambda + 15 = 0,$$

$$(\lambda - 15)(\lambda - 1) = 0,$$

$$\lambda = 1, \lambda = 15.$$

   (b)
$$\left|\begin{bmatrix} 9-\lambda & -9 \\ 7 & -7-\lambda \end{bmatrix}\right| = 0,$$

$$(-9-\lambda)(-7-\lambda) - 63 = 0,$$

$$(63 + 16\lambda + \lambda^2) - 63 = 0,$$

$$\lambda^2 + 16\lambda = 0,$$

$$\lambda(\lambda + 16) = 0,$$

$$\lambda = 0, \lambda = -16.$$

(c)

$$\left| \begin{bmatrix} 2-\lambda & -8 \\ -10 & 10-\lambda \end{bmatrix} \right| = 0,$$

$$(2-\lambda)(10-\lambda) - 80 = 0,$$

$$(20 - 12\lambda + \lambda^2) - 80 = 0,$$

$$\lambda^2 - 12\lambda - 60 = 0,$$

$$\lambda = \frac{12 \pm \sqrt{144 - 4(1)(-60)}}{2},$$

$$\lambda = \frac{12 \pm \sqrt{144 + 240}}{2},$$

$$\lambda = \frac{12 \pm \sqrt{384}}{2},$$

$$\lambda = \frac{12 \pm 8\sqrt{6}}{2},$$

$$\lambda = 6 \pm 4\sqrt{6},$$

$$\lambda = 6 - 4\sqrt{6} = -3.80,$$
$$\lambda = 6 + 4\sqrt{6} = 15.80.$$

8. We can reduce the matrix as much as possible using the elementary row operations. If the matrix reduces to an identity matrix, then its rank is full and equal to the number of columns. If it contains a column of all zeroes then it's rank is equal to the number of columns that do not consist entirely of zeroes.

   (a) With the goal of reducing this matrix to the identity matrix, let's start by interchanging the first and second rows,

   $$\begin{bmatrix} 1 & 7 & -1 \\ 3 & -3 & 0 \\ -4 & 4 & 6 \end{bmatrix},$$

   multiplying the first row by -3 and adding it to the second row,

   $$\begin{bmatrix} 1 & 7 & -1 \\ 0 & -24 & -3 \\ -4 & 4 & 6 \end{bmatrix},$$

   and multiplying the first row by 4 and adding it to the third row,

   $$\begin{bmatrix} 1 & 7 & -1 \\ 0 & -24 & -3 \\ 0 & 32 & 10 \end{bmatrix}.$$

Next, we divide the second row by -3 and the third row by 2,

$$\begin{bmatrix} 1 & 7 & -1 \\ 0 & 8 & 1 \\ 0 & 16 & 5 \end{bmatrix},$$

multiply the second row by -2 and add it to the third row,

$$\begin{bmatrix} 1 & 7 & -1 \\ 0 & 8 & 1 \\ 0 & 0 & 3 \end{bmatrix},$$

and divide the third row by 3,

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

We now add the third row to the first, and we multiply the third row by -1 and add it to the second,

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

divide the second row by 8,

$$\begin{bmatrix} 1 & 7 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and add -7 times the second row to the first row,

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

Since the matrix breaks down to an identity matrix, it is of full rank – the rank is 3.

(b) Let's begin by interchanging the first and second rows,

$$\begin{bmatrix} 1 & 6 & 0 & 2 \\ -2 & 4 & -1 & 3 \\ 3 & 5 & -2 & -1 \\ 4 & 14 & -5 & 1 \end{bmatrix}.$$

We add 2 times the first row to the second, -3 times the first row to the third, and -4 times the first row to the fourth,

$$\begin{bmatrix} 1 & 6 & 0 & 2 \\ 0 & 16 & -1 & 7 \\ 0 & -13 & -2 & -7 \\ 0 & -10 & -5 & -7 \end{bmatrix}.$$

To avoid fractions in the next step, we multiply the first, second and third rows by 10, and interchange the second and fourth rows

$$\begin{bmatrix} 10 & 60 & 0 & 20 \\ 0 & -10 & -5 & -7 \\ 0 & 160 & -10 & 70 \\ 0 & -130 & -20 & -70 \end{bmatrix},$$

we then add 16 times the second row to the third,

$$\begin{bmatrix} 10 & 60 & 0 & 20 \\ 0 & -10 & -5 & -7 \\ 0 & 0 & -90 & -42 \\ 0 & -130 & -20 & -70 \end{bmatrix},$$

and we add -13 times the second row to the fourth row,

$$\begin{bmatrix} 10 & 60 & 0 & 20 \\ 0 & -10 & -5 & -7 \\ 0 & 0 & -90 & -42 \\ 0 & 0 & 45 & 21 \end{bmatrix}.$$

We divide the third row by 2,

$$\begin{bmatrix} 10 & 60 & 0 & 20 \\ 0 & -10 & -5 & -7 \\ 0 & 0 & -45 & -21 \\ 0 & 0 & 45 & 21 \end{bmatrix},$$

and add it to the fourth row

$$\begin{bmatrix} 10 & 60 & 0 & 20 \\ 0 & -10 & -5 & -7 \\ 0 & 0 & -45 & -21 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We next divide the third row by -3,

$$\begin{bmatrix} 10 & 60 & 0 & 20 \\ 0 & -10 & -5 & -7 \\ 0 & 0 & 15 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

add 6 times the second row to the first,

$$\begin{bmatrix} 10 & 0 & -30 & -22 \\ 0 & -10 & -5 & -7 \\ 0 & 0 & 15 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and add 2 times the third row to the first,

$$\begin{bmatrix} 10 & 0 & 0 & -8 \\ 0 & -10 & -5 & -7 \\ 0 & 0 & 15 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

To continue avoiding fractions, we multiply the second row by 3,

$$\begin{bmatrix} 10 & 0 & 0 & -8 \\ 0 & -30 & -15 & -21 \\ 0 & 0 & 15 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and add the third row to the second,

$$\begin{bmatrix} 10 & 0 & 0 & -8 \\ 0 & -30 & 0 & -14 \\ 0 & 0 & 15 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Finally, we divide the first row by 10, the second row by -30, and the third row by 15,

$$\begin{bmatrix} 1 & 0 & 0 & -\frac{4}{5} \\ 0 & 1 & 0 & \frac{7}{15} \\ 0 & 0 & 1 & \frac{7}{15} \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since the first three rows and columns contain an identity matrix, we've reduced the matrix as much as we can. There are 3 rows that do not consist entirely of zeroes, so the rank of the matrix is 3. Since this matrix is not of full column rank, we also know that it is singular, that there is a linear dependency in its rows and columns, that its determinant is zero, and that it has no inverse.

(c) Let's start by interchanging the first and third rows,

$$\begin{bmatrix} 1 & 4 & 11 & 1 \\ -4 & 6 & 8 & 1 \\ 5 & -2 & 3 & 0 \\ 9 & -8 & -5 & -1 \end{bmatrix},$$

then multiplying the first row by 4 and adding it to the second row, multiplying it by -5 and adding it to the third row, and multiplying it by -9 and adding it to the third row,

$$\begin{bmatrix} 1 & 4 & 11 & 1 \\ 0 & 22 & 49 & 5 \\ 0 & -22 & -52 & -5 \\ 0 & -44 & -104 & -10 \end{bmatrix}.$$

Next we add the second row to the third, and add 2 times the second row to the fourth row,

$$\begin{bmatrix} 1 & 4 & 11 & 1 \\ 0 & 22 & 49 & 5 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & -6 & 0 \end{bmatrix}.$$

We add 2 times the third row to the fourth row,

$$\begin{bmatrix} 1 & 4 & 11 & 1 \\ 0 & 22 & 49 & 5 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

divide the third row by -3,

$$\begin{bmatrix} 1 & 4 & 11 & 1 \\ 0 & 22 & 49 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and add -49 times the third row to the second row and -11 times the third row to the first row,

$$\begin{bmatrix} 1 & 4 & 0 & 1 \\ 0 & 22 & 0 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We multiply the first row by 11,

$$\begin{bmatrix} 11 & 44 & 0 & 11 \\ 0 & 22 & 0 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and add -2 times the second row to the first,

$$\begin{bmatrix} 11 & 0 & 0 & 1 \\ 0 & 22 & 0 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Finally we divide the first row by 11 and the second row by 22,

$$\begin{bmatrix} 1 & 0 & 0 & \frac{1}{11} \\ 0 & 1 & 0 & \frac{5}{22} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since the first three rows and columns contain an identity matrix, we've reduced the matrix as much as we can. There are 3 rows that do not consist entirely of zeroes, so the rank of the matrix is 3. Since this matrix is not of full column rank, we also know that it is singular, that there is a linear dependency in its rows and columns, that its determinant is zero, and that it has no inverse.

(d) Note that this matrix,

$$\begin{bmatrix} 1 & -5 & 8 \\ 2 & -3 & 4 \\ 6 & 0 & 3 \\ 9 & 1 & 3 \\ -4 & 4 & 0 \end{bmatrix},$$

is not square. Since it has more rows than columns, it cannot be full row rank but it might be full column rank. To reduce, start by multiplying the first row by -2 and adding it to the third row, then multiply it by -6 and add it to the third row, then multiply it by -9 and add it to the fourth row, and multiply it by 4 and add it to the fifth row,

$$\begin{bmatrix} 1 & -5 & 8 \\ 0 & 7 & -12 \\ 0 & 30 & -45 \\ 0 & -44 & -69 \\ 0 & -16 & 32 \end{bmatrix}.$$

Next, interchange the second and fifth rows,

$$\begin{bmatrix} 1 & -5 & 8 \\ 0 & -16 & 32 \\ 0 & 30 & -45 \\ 0 & -44 & -69 \\ 0 & 7 & -12 \end{bmatrix},$$

divide the second row by -16, divide the third row by 15, and divide the fourth row by -3,

$$\begin{bmatrix} 1 & -5 & 8 \\ 0 & 1 & -2 \\ 0 & 2 & -3 \\ 0 & 18 & -23 \\ 0 & 7 & -12 \end{bmatrix}.$$

Multiply the second row by-2 and add it to the third row, then multiply the second row by -18 and add it to the fourth row, and multiply the second row by -7 and add it to the fifth row,

$$\begin{bmatrix} 1 & -5 & 8 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \\ 0 & 0 & 13 \\ 0 & 0 & 2 \end{bmatrix},$$

and multiply the third row by -1,

$$\begin{bmatrix} 1 & -5 & 8 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \\ 0 & 0 & 13 \\ 0 & 0 & 2 \end{bmatrix}.$$

Multiply the third row by -8, 2, -13, and -2 and add it to the first, second, fourth, and fifth rows respectively,

$$\begin{bmatrix} 1 & -5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and add 5 times the second row to the first row,

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

Since we've reduced the matrix so that there is an identity matrix in the first three rows and columns, we've reduced as much as we can. There are three rows that do not consist entirely of zeroes, so the rank is 3, and the matrix is of full column rank.

9. First we calculate the matrix product

$$X'X = \begin{bmatrix} 6 & 4 & 7 & 2 & 6 & 1 & 3 & 3 & 2 & 1 \\ 6 & 3 & 5 & 2 & 6 & 1 & 4 & 3 & 2 & 1 \\ 12 & 7 & 12 & 4 & 12 & 2 & 7 & 6 & 4 & 2 \end{bmatrix} \begin{bmatrix} 6 & 6 & 12 \\ 4 & 3 & 7 \\ 7 & 5 & 12 \\ 2 & 2 & 4 \\ 6 & 6 & 12 \\ 1 & 1 & 2 \\ 3 & 4 & 7 \\ 3 & 3 & 6 \\ 2 & 2 & 4 \\ 1 & 1 & 2 \end{bmatrix},$$

This product multiplies a (3×10) matrix by a (10×3) matrix, which is conformable and results in a (3×3) matrix. As we saw in chapter 8 exercise 4, any matrix that is left-multiplied by its transpose will be symmetric. Therefore it will only be necessary to calculate the 6 unique elements. The (1,1) element is

$$(6 \times 6) + (4 \times 4) + (7 \times 7) + (2 \times 2) + (6 \times 6) + (1 \times 1) + (3 \times 3) + (3 \times 3) + (2 \times 2) + (1 \times 1) = 165.$$

The (1,2) and (2,1) elements are

$$(6 \times 6) + (4 \times 3) + (7 \times 5) + (2 \times 2) + (6 \times 6) + (1 \times 1) + (3 \times 4) + (3 \times 3) + (2 \times 2) + (1 \times 1) = 150.$$

The (1,3) and (3,1) elements are

$$(6 \times 12) + (4 \times 7) + (7 \times 12) + (2 \times 4) + (6 \times 12) + (1 \times 2) + (3 \times 7) + (3 \times 6) + (2 \times 4) + (1 \times 2) = 315.$$

The (2,2) element is

$$(6 \times 6) + (3 \times 3) + (5 \times 5) + (2 \times 2) + (6 \times 6) + (1 \times 1) + (4 \times 4) + (3 \times 3) + (2 \times 2) + (1 \times 1) = 141.$$

The (2,3) and (3,2) elements are

$$(6 \times 12) + (3 \times 7) + (5 \times 12) + (2 \times 4) + (6 \times 12) + (1 \times 2) + (4 \times 7) + (3 \times 6) + (2 \times 4) + (1 \times 2) = 291.$$

And the (3,3) element is

$$(12 \times 12) + (7 \times 7) + (12 \times 12) + (4 \times 4) + (12 \times 12) + (2 \times 2) + (7 \times 7) + (6 \times 6) + (4 \times 4) + (2 \times 2) = 606.$$

The entire product is

$$X'X = \begin{bmatrix} 165 & 150 & 315 \\ 150 & 141 & 291 \\ 315 & 291 & 606 \end{bmatrix}.$$

To find the determinant, we apply Sarrus' rule. We write the matrix twice, adjacent to itself. Then we draw six ovals. Three ovals start in the first row of the left matrix, and move diagonally down and to the right, encompassing three elements. Three ovals start in the first row of the right matrix, and move diagonally down and to the left, also encompassing three elements. We multiply the elements within each oval together, add the products for the ovals that move down and right, and subtract the products of the ovals that move down and left.

The products from the ovals that move down and right are

$$\begin{aligned} (165 \times 141 \times 606) &= 14,098,590, \\ (150 \times 291 \times 315) &= 13,749,750, \\ (315 \times 150 \times 291) &= 13,749,750. \end{aligned}$$

The products from the ovals that move down and left are

$$\begin{aligned} (165 \times 291 \times 291) &= 13,972,365, \\ (150 \times 150 \times 606) &= 13,635,000, \\ (315 \times 141 \times 315) &= 13,990,725. \end{aligned}$$

The determinant is

$$|X'X| = (14,098,590 + 13,749,750 + 13,749,750) - (13,972,365 + 13,635,000 + 13,990,725) = 0.$$

10. (a) We are trying to find the product

$$X'Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 0 & 1 & 3 & 11 & 0 & 5 & 0 \\ 5 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 12 & 2 \end{bmatrix} \begin{bmatrix} 10 \\ 7 \\ 2 \\ 8 \\ 10 \\ 6 \\ 2 \\ 10 \\ 8 \\ 5 \end{bmatrix}$$

$$= \begin{bmatrix} (1\times 10)+(1\times 7)+(1\times 2)+(1\times 8)+(1\times 10)+(1\times 6)+(1\times 2)+(1\times 10)+(1\times 8)+(1\times 5) \\ (2\times 10)+(1\times 7)+(2\times 2)+(0\times 8)+(1\times 10)+(3\times 6)+(11\times 2)+(0\times 10)+(5\times 8)+(0\times 5) \\ (5\times 10)+(1\times 7)+(0\times 2)+(2\times 8)+(0\times 10)+(0\times 6)+(0\times 2)+(0\times 10)+(12\times 8)+(2\times 5) \end{bmatrix}$$

$$= \begin{bmatrix} 68 \\ 121 \\ 179 \end{bmatrix}.$$

(b) We are computing the product

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 0 & 1 & 3 & 11 & 0 & 5 & 0 \\ 5 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 12 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 5 \\ 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 3 & 0 \\ 1 & 11 & 0 \\ 1 & 0 & 0 \\ 1 & 5 & 12 \\ 1 & 0 & 2 \end{bmatrix},$$

which will be a symmetric (3×3) matrix. We have to compute the (1,1), (1,2), (1,3), (2,2), (2,3), and (3,3) elements.

The (1,1) element is

$$(1\times 1)+(1\times 1)+(1\times 1)+(1\times 1)+(1\times 1)+(1\times 1)+(1\times 1)+(1\times 1)+(1\times 1)+(1\times 1) = 10.$$

The (1,2) and (2,1) elements are

$$(1\times 2)+(1\times 1)+(1\times 2)+(1\times 0)+(1\times 1)+(1\times 3)+(1\times 11)+(1\times 0)+(1\times 5)+(1\times 0) = 25.$$

The (1,3) and (3,1) elements are

$$(1\times 5)+(1\times 1)+(1\times 0)+(1\times 2)+(1\times 0)+(1\times 0)+(1\times 0)+(1\times 0)+(1\times 12)+(1\times 2) = 22.$$

The (2,2) element is

$$(2\times 2)+(1\times 1)+(2\times 2)+(0\times 0)+(1\times 1)+(3\times 3)+(11\times 11)+(0\times 0)+(5\times 5)+(0\times 0) = 165.$$

The (2,3) and (3,2) elements are

$$(2\times 5)+(1\times 1)+(2\times 0)+(0\times 2)+(1\times 0)+(3\times 0)+(11\times 0)+(0\times 0)+(5\times 12)+(0\times 2) = 71.$$

Finally, the (3,3) element is

$$(5\times 5)+(1\times 1)+(0\times 0)+(2\times 2)+(0\times 0)+(0\times 0)+(0\times 0)+(0\times 0)+(12\times 12)+(2\times 2)=178.$$

So the entire matrix is

$$X'X=\begin{bmatrix}10 & 25 & 22\\ 25 & 165 & 71\\ 22 & 71 & 178\end{bmatrix}.$$

(c) In order to find the inverse of $X'X$, we will apply the formula

$$(X'X)^{-1}=\frac{1}{|X'X|}\text{adj}(X'X),$$

which involves finding the adjoint matrix, then the determinant. To find the adjoint matrix, we first determine the minor elements of $X'X$, which for the (i,j)th element involves removing the ith row and the jth column of $X'X$ and calculating the determinant of the remaining (2×2) matrix. In general, the determinant of a (2×2) matrix is given by

$$\left|\begin{bmatrix}a & b\\ c & d\end{bmatrix}\right|=ad-bc.$$

The matrix of minor elements is symmetric because $X'X$ is symmetric, and therefore removing the ith row and jth column is the same as removing the jth row and ith column. As a result, we only need to calculate 6 minor elements.

To find the $(1,1)$ minor element, we remove the first row and first column from $X'X$ and find the determinant:

$$M_{11}=\left|\begin{bmatrix}165 & 71\\ 71 & 178\end{bmatrix}\right|=(165\times 178)-(71\times 71)=24,329.$$

To find the $(1,2)$ and $(2,1)$ minor elements:

$$M_{12}=M_{21}=\left|\begin{bmatrix}25 & 71\\ 22 & 178\end{bmatrix}\right|=(25\times 178)-(22\times 71)=2,888.$$

To find the $(1,3)$ and $(3,1)$ minor elements:

$$M_{13}=M_{31}=\left|\begin{bmatrix}25 & 165\\ 22 & 71\end{bmatrix}\right|=(25\times 71)-(22\times 165)=-1,855.$$

To find the $(2,2)$ minor element:

$$M_{22}=\left|\begin{bmatrix}10 & 22\\ 22 & 178\end{bmatrix}\right|=(10\times 178)-(22\times 22)=1,296$$

To find the $(2,3)$ and $(3,2)$ minor elements:

$$M_{23}=M_{32}=\left|\begin{bmatrix}10 & 25\\ 22 & 71\end{bmatrix}\right|=(10\times 71)-(22\times 25)=160$$

Finally, to find the $(3,3)$ minor element:

$$M_{33}=\left|\begin{bmatrix}10 & 25\\ 25 & 165\end{bmatrix}\right|=(10\times 165)-(25\times 25)=1,025.$$

The entire matrix of minor elements is therefore

$$M = \begin{bmatrix} 24,329 & 2,888 & -1,855 \\ 2,888 & 1,296 & 160 \\ -1,855 & 160 & 1,025 \end{bmatrix}.$$

Next we find the cofactor matrix by multiplying the elements of the matrix of minor elements by -1 if the row and column numbers of the element add to an odd number. The cofactor matrix is

$$C = \begin{bmatrix} 24,329 & -2,888 & -1,855 \\ -2,888 & 1,296 & -160 \\ -1,855 & -160 & 1,025 \end{bmatrix}.$$

Next, the adjoint matrix is the transpose of the cofactor matrix. But since $X'X$, M, and C are all symmetric in this case, the adjoint matrix is equal to the cofactor matrix:

$$\mathrm{adj}(X'X) = \begin{bmatrix} 24,329 & -2,888 & -1,855 \\ -2,888 & 1,296 & -160 \\ -1,855 & -160 & 1,025 \end{bmatrix}.$$

To find the determinant of the (3×3) matrix, we choose one row or column of the matrix, multiply each element by its corresponding cofactor element, and sum these products. Let's choose the first row of $X'X$:

$$\begin{bmatrix} 10 & 25 & 22 \end{bmatrix}$$

The corresponding cofactor elements are

$$\begin{bmatrix} 24,329 & -2,888 & -1,855 \end{bmatrix}.$$

Multiplying the corresponding elements and adding gives us:

$$|X'X| = (10 \times 24,329) + (25 \times -2,888) + (22 \times -1,855) = 130,280.$$

Finally, plugging the adjoint matrix and determinant into

$$(X'X)^{-1} = \frac{1}{|X'X|}\mathrm{adj}(X'X),$$

gives us

$$(X'X)^{-1} = \frac{1}{130,280}\begin{bmatrix} 24,329 & -2,888 & -1,855 \\ -2,888 & 1,296 & -160 \\ -1,855 & -160 & 1,025 \end{bmatrix} = \begin{bmatrix} 0.187 & -0.022 & -0.014 \\ -0.022 & 0.010 & -0.001 \\ -0.014 & -0.001 & 0.008 \end{bmatrix}.$$

(d) Having already gone through all of the work to calculate $(X'X)^{-1}$ and $X'Y$, all we have left to do is multiply these two matrices:

$$\hat{\beta} = (X'X)^{-1}(X'Y) = \begin{bmatrix} 0.187 & -0.022 & -0.014 \\ -0.022 & 0.010 & -0.001 \\ -0.014 & -0.001 & 0.008 \end{bmatrix}\begin{bmatrix} 68 \\ 121 \\ 179 \end{bmatrix}.$$

The product multiplies a (3×3) matrix by a (3×1) matrix, so multiplication is possible and the product is the following (3×1) vector:

$$\hat{\beta} = (X'X)^{-1}(X'Y) = \begin{bmatrix} (0.187 \times 68) + (-0.022 \times 121) + (-0.014 \times 179) \\ (-0.022 \times 68) + (0.010 \times 121) + (-0.001 \times 179) \\ (-0.014 \times 68) + (-0.001 \times 121) + (0.008 \times 179) \end{bmatrix} = \begin{bmatrix} 7.47 \\ -0.52 \\ 0.29 \end{bmatrix}.$$

So the OLS estimate for α is 7.47, for β_1 is -0.52, and for β_2 is 0.29. A computer can perform these calculations in a fraction of a second. But now, you have mastered the mathematics that the computer uses. Hopefully, this process demystifies the process of OLS regression, and you can look at a computer as a useful tool, but one no more intelligent than a handheld calculator.

10 Linear Systems of Equations and Eigenvalues

1. (a) We can rewrite this system of equations in terms of matrices:

$$\begin{bmatrix} -3 & 5 & 5 \\ 1 & -4 & -2 \\ 3 & 0 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -43 \\ 31 \\ 7 \end{bmatrix}.$$

We can solve this system by left-multiplying both sides of this equation by the inverse of the (3×3) matrix. To find this matrix, we first determine the minor elements:

$$\begin{aligned}
M_{11} &= \left|\begin{bmatrix} -4 & -2 \\ 0 & -4 \end{bmatrix}\right| &&= (-4 \times -4) - (-2 \times 0) &&= 16 \\
M_{12} &= \left|\begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}\right| &&= (1 \times -4) - (-2 \times 3) &&= 2 \\
M_{13} &= \left|\begin{bmatrix} 1 & -4 \\ 3 & 0 \end{bmatrix}\right| &&= (1 \times 0) - (-4 \times 3) &&= 12 \\
M_{21} &= \left|\begin{bmatrix} 5 & 5 \\ 0 & -4 \end{bmatrix}\right| &&= (5 \times -4) - (5 \times 0) &&= -20 \\
M_{22} &= \left|\begin{bmatrix} -3 & 5 \\ 3 & -4 \end{bmatrix}\right| &&= (-3 \times -4) - (5 \times 3) &&= -3 \\
M_{23} &= \left|\begin{bmatrix} -3 & 5 \\ 3 & 0 \end{bmatrix}\right| &&= (-3 \times 0) - (5 \times 3) &&= -15 \\
M_{31} &= \left|\begin{bmatrix} 5 & 5 \\ -4 & -2 \end{bmatrix}\right| &&= (5 \times -2) - (5 \times -4) &&= 10 \\
M_{32} &= \left|\begin{bmatrix} -3 & 5 \\ 1 & -2 \end{bmatrix}\right| &&= (-3 \times -2) - (5 \times 1) &&= 1 \\
M_{33} &= \left|\begin{bmatrix} -3 & 5 \\ 1 & -4 \end{bmatrix}\right| &&= (-3 \times -4) - (5 \times 1) &&= 7
\end{aligned}$$

Next we find the cofactors,

$$C_{11} = -1^{(1+1)}M_{11} = 16, \quad C_{12} = -1^{(1+2)}M_{12} = -2, \quad C_{13} = -1^{(1+3)}M_{13} = 12,$$

$$C_{21} = -1^{(2+1)}M_{21} = 20, \quad C_{22} = -1^{(2+2)}M_{22} = -3, \quad C_{23} = -1^{(2+3)}M_{23} = 15,$$

$$C_{31} = -1^{(3+1)}M_{31} = 10, \quad C_{32} = -1^{(3+2)}M_{32} = -1, \quad C_{33} = -1^{(3+3)}M_{33} = 7,$$

and take the transpose of the cofactor matrix to find the adjoint matrix,

$$\mathrm{adj}\left(\begin{bmatrix} -3 & 5 & 5 \\ 1 & -4 & -2 \\ 3 & 0 & -4 \end{bmatrix}\right) = \begin{bmatrix} 16 & 20 & 10 \\ -2 & -3 & -1 \\ 12 & 15 & 7 \end{bmatrix}.$$

To find the determinant, we choose one row or column (the first row in this case), multiply the elements by their cofactors, and add the products:

$$\left|\begin{bmatrix} -3 & 5 & 5 \\ 1 & -4 & -2 \\ 3 & 0 & -4 \end{bmatrix}\right| = (-3 \times 16) + (5 \times -2) + (5 \times 12) = 2.$$

We then plug the determinant and the adjoint matrix into the formula for a matrix inverse:

$$\left(\begin{bmatrix} 1 & 3 & 1 \\ 2 & 1 & 5 \\ 2 & 2 & -3 \end{bmatrix}\right)^{-1} = \frac{1}{2}\begin{bmatrix} 16 & 20 & 10 \\ -2 & -3 & -1 \\ 12 & 15 & 7 \end{bmatrix} = \begin{bmatrix} 8 & 10 & 5 \\ -1 & -1.5 & -0.5 \\ 6 & 7.5 & 3.5 \end{bmatrix}.$$

The solution to the system of equations is therefore

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 & 10 & 5 \\ -1 & -1.5 & -0.5 \\ 6 & 7.5 & 3.5 \end{bmatrix}\begin{bmatrix} -43 \\ 31 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \\ -7 \\ -1 \end{bmatrix}.$$

(b) We can rewrite this system of equations in terms of matrices:

$$\begin{bmatrix} -2 & 3 & 0 \\ -4 & 1 & 3 \\ 0 & 5 & -5 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ -5 \\ -15 \end{bmatrix}.$$

We can solve this system by left-multiplying both sides of this equation by the inverse of the (3×3) matrix. To find this matrix, we first determine the minor elements:

$$\begin{aligned}
M_{11} &= \left|\begin{bmatrix} 1 & 3 \\ 5 & -5 \end{bmatrix}\right| &&= (1 \times -5) - (3 \times 5) &&= -20 \\
M_{12} &= \left|\begin{bmatrix} -4 & 3 \\ 0 & -5 \end{bmatrix}\right| &&= (-4 \times -5) - (3 \times 0) &&= 20 \\
M_{13} &= \left|\begin{bmatrix} -4 & 1 \\ 0 & 5 \end{bmatrix}\right| &&= (-4 \times 5) - (1 \times 0) &&= -20 \\
M_{21} &= \left|\begin{bmatrix} 3 & 0 \\ 5 & -5 \end{bmatrix}\right| &&= (3 \times -5) - (0 \times 5) &&= -15 \\
M_{22} &= \left|\begin{bmatrix} -2 & 0 \\ 0 & -5 \end{bmatrix}\right| &&= (-2 \times -5) - (0 \times 0) &&= 10 \\
M_{23} &= \left|\begin{bmatrix} -2 & 3 \\ 0 & 5 \end{bmatrix}\right| &&= (-2 \times 5) - (3 \times 0) &&= -10 \\
M_{31} &= \left|\begin{bmatrix} 3 & 0 \\ 1 & 3 \end{bmatrix}\right| &&= (3 \times 3) - (0 \times 1) &&= 9 \\
M_{32} &= \left|\begin{bmatrix} -2 & 0 \\ -4 & 3 \end{bmatrix}\right| &&= (-2 \times 3) - (0 \times -4) &&= -6 \\
M_{33} &= \left|\begin{bmatrix} -2 & 3 \\ -4 & 1 \end{bmatrix}\right| &&= (-2 \times 1) - (3 \times -4) &&= 10
\end{aligned}$$

Next we find the cofactors,

$$C_{11} = -1^{(1+1)}M_{11} = -20, \quad C_{12} = -1^{(1+2)}M_{12} = -20, \quad C_{13} = -1^{(1+3)}M_{13} = -20,$$

$$C_{21} = -1^{(2+1)}M_{21} = 15, \quad C_{22} = -1^{(2+2)}M_{22} = 10, \quad C_{23} = -1^{(2+3)}M_{23} = 10,$$

$$C_{31} = -1^{(3+1)}M_{31} = 9, \quad C_{32} = -1^{(3+2)}M_{32} = 6, \quad C_{33} = -1^{(3+3)}M_{33} = 10,$$

and take the transpose of the cofactor matrix to find the adjoint matrix,

$$\operatorname{adj}\left(\begin{bmatrix} -2 & 3 & 0 \\ -4 & 1 & 3 \\ 0 & 5 & -5 \end{bmatrix}\right) = \begin{bmatrix} -20 & 15 & 9 \\ -20 & 10 & 6 \\ -20 & 10 & 10 \end{bmatrix}.$$

To find the determinant, we choose one row or column (the first row in this case), multiply the elements by their cofactors, and add the products:

$$\left|\begin{bmatrix} -2 & 3 & 0 \\ -4 & 1 & 3 \\ 0 & 5 & -5 \end{bmatrix}\right| = (-2 \times -20) + (3 \times -20) + (0 \times -20) = -20.$$

We then plug the determinant and the adjoint matrix into the formula for a matrix inverse:

$$\left(\begin{bmatrix} 1 & 3 & -4 \\ 2 & 1 & 3 \\ 2 & 2 & -2 \end{bmatrix}\right)^{-1} = \frac{1}{-20}\begin{bmatrix} -20 & 15 & 9 \\ -20 & 10 & 6 \\ -20 & 10 & 10 \end{bmatrix} = \begin{bmatrix} 1 & -0.75 & -0.45 \\ 1 & -0.5 & -0.3 \\ 1 & -0.5 & -0.5 \end{bmatrix}.$$

The solution to the system of equations is therefore

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & -0.75 & -0.45 \\ 1 & -0.5 & -0.3 \\ 1 & -0.5 & -0.5 \end{bmatrix}\begin{bmatrix} 5 \\ -5 \\ -15 \end{bmatrix} = \begin{bmatrix} 15.5 \\ 12 \\ 15 \end{bmatrix}.$$

(c) We can rewrite this system of equations in terms of matrices:

$$\begin{bmatrix} -2 & 2 & -1 \\ 5 & -3 & -4 \\ 3 & 0 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -11 \\ -12 \end{bmatrix}.$$

We can solve this system by left-multiplying both sides of this equation by the inverse of the (3×3)

matrix. To find this matrix, we first determine the minor elements:

$$
\begin{aligned}
M_{11} &= \left|\begin{bmatrix} -3 & -4 \\ 0 & -3 \end{bmatrix}\right| &&= (-3\times -3)-(-4\times 0) &&= 9 \\
M_{12} &= \left|\begin{bmatrix} 5 & -4 \\ 3 & -3 \end{bmatrix}\right| &&= (5\times -3)-(-4\times 3) &&= -3 \\
M_{13} &= \left|\begin{bmatrix} 5 & -3 \\ 3 & 0 \end{bmatrix}\right| &&= (5\times 0)-(-3\times 3) &&= 9 \\
M_{21} &= \left|\begin{bmatrix} 2 & -1 \\ 0 & -3 \end{bmatrix}\right| &&= (2\times -3)-(-1\times 0) &&= -6 \\
M_{22} &= \left|\begin{bmatrix} -2 & -1 \\ 3 & -3 \end{bmatrix}\right| &&= (-2\times -3)-(-1\times 3) &&= 9 \\
M_{23} &= \left|\begin{bmatrix} -2 & 2 \\ 3 & 0 \end{bmatrix}\right| &&= (-2\times 0)-(2\times 3) &&= -6 \\
M_{31} &= \left|\begin{bmatrix} 2 & -1 \\ -3 & -4 \end{bmatrix}\right| &&= (2\times -4)-(-1\times -3) &&= -11 \\
M_{32} &= \left|\begin{bmatrix} -2 & -1 \\ 5 & -4 \end{bmatrix}\right| &&= (-2\times -4)-(-1\times 5) &&= 13 \\
M_{33} &= \left|\begin{bmatrix} -2 & 2 \\ 5 & -3 \end{bmatrix}\right| &&= (-2\times -3)-(2\times 5) &&= -4
\end{aligned}
$$

Next we find the cofactors,

$$
\begin{aligned}
&C_{11} = -1^{(1+1)}M_{11} = 9, && C_{12} = -1^{(1+2)}M_{12} = 3, && C_{13} = -1^{(1+3)}M_{13} = 9, \\
&C_{21} = -1^{(2+1)}M_{21} = 6, && C_{22} = -1^{(2+2)}M_{22} = 9, && C_{23} = -1^{(2+3)}M_{23} = 6, \\
&C_{31} = -1^{(3+1)}M_{31} = -11, && C_{32} = -1^{(3+2)}M_{32} = -13, && C_{33} = -1^{(3+3)}M_{33} = -4,
\end{aligned}
$$

and take the transpose of the cofactor matrix to find the adjoint matrix,

$$
\mathrm{adj}\left(\begin{bmatrix} -2 & 2 & -1 \\ 5 & -3 & -4 \\ 3 & 0 & -3 \end{bmatrix}\right) = \begin{bmatrix} 9 & 6 & -11 \\ 3 & 9 & -13 \\ 9 & 6 & -4 \end{bmatrix}.
$$

To find the determinant, we choose one row or column (the first row in this case), multiply the elements by their cofactors, and add the products:

$$
\left|\begin{bmatrix} -2 & 2 & -1 \\ 5 & -3 & -4 \\ 3 & 0 & -3 \end{bmatrix}\right| = (-2\times 9)+(2\times 3)+(-1\times 9) = -21.
$$

We then plug the determinant and the adjoint matrix into the formula for a matrix inverse:

$$
\left(\begin{bmatrix} 1 & 3 & 5 \\ 2 & 1 & 2 \\ 2 & 2 & -2 \end{bmatrix}\right)^{-1} = \frac{1}{-21}\begin{bmatrix} 9 & 6 & -11 \\ 3 & 9 & -13 \\ 9 & 6 & -4 \end{bmatrix} = \begin{bmatrix} -0.43 & -0.29 & 0.52 \\ -0.14 & -0.43 & 0.62 \\ -0.43 & -0.29 & 0.19 \end{bmatrix}.
$$

The solution to the system of equations is therefore

$$
\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -0.43 & -0.29 & 0.52 \\ -0.14 & -0.43 & 0.62 \\ -0.43 & -0.29 & 0.19 \end{bmatrix}\begin{bmatrix} 2 \\ -11 \\ -12 \end{bmatrix} = \begin{bmatrix} -4 \\ -3 \\ 0 \end{bmatrix}.
$$

(d) We can rewrite this system of equations in terms of matrices:

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

We can solve this system by left-multiplying both sides of this equation by the inverse of the (3×3) matrix. To find this matrix, we first determine the minor elements:

$$\begin{aligned}
M_{11} &= \left|\begin{bmatrix} 4 & 5 \\ 3 & -1 \end{bmatrix}\right| &&= (4 \times -1) - (5 \times 3) &&= -19 \\
M_{12} &= \left|\begin{bmatrix} -2 & 5 \\ -3 & -1 \end{bmatrix}\right| &&= (-2 \times -1) - (5 \times -3) &&= 17 \\
M_{13} &= \left|\begin{bmatrix} -2 & 4 \\ -3 & 3 \end{bmatrix}\right| &&= (-2 \times 3) - (4 \times -3) &&= 6 \\
M_{21} &= \left|\begin{bmatrix} -1 & 1 \\ 3 & -1 \end{bmatrix}\right| &&= (-1 \times -1) - (1 \times 3) &&= -2 \\
M_{22} &= \left|\begin{bmatrix} 2 & 1 \\ -3 & -1 \end{bmatrix}\right| &&= (2 \times -1) - (1 \times -3) &&= 1 \\
M_{23} &= \left|\begin{bmatrix} 2 & -1 \\ -3 & 3 \end{bmatrix}\right| &&= (2 \times 3) - (-1 \times -3) &&= 3 \\
M_{31} &= \left|\begin{bmatrix} -1 & 1 \\ 4 & 5 \end{bmatrix}\right| &&= (-1 \times 5) - (1 \times 4) &&= -9 \\
M_{32} &= \left|\begin{bmatrix} 2 & 1 \\ -2 & 5 \end{bmatrix}\right| &&= (2 \times 5) - (1 \times -2) &&= 12 \\
M_{33} &= \left|\begin{bmatrix} 2 & -1 \\ -2 & 4 \end{bmatrix}\right| &&= (2 \times 4) - (-1 \times -2) &&= 6
\end{aligned}$$

Next we find the cofactors,

$$\begin{aligned}
&C_{11} = -1^{(1+1)} M_{11} = -19, && C_{12} = -1^{(1+2)} M_{12} = -17, && C_{13} = -1^{(1+3)} M_{13} = 6, \\
&C_{21} = -1^{(2+1)} M_{21} = 2, && C_{22} = -1^{(2+2)} M_{22} = 1, && C_{23} = -1^{(2+3)} M_{23} = -3, \\
&C_{31} = -1^{(3+1)} M_{31} = -9, && C_{32} = -1^{(3+2)} M_{32} = -12, && C_{33} = -1^{(3+3)} M_{33} = 6,
\end{aligned}$$

and take the transpose of the cofactor matrix to find the adjoint matrix,

$$\operatorname{adj}\left(\begin{bmatrix} 2 & -1 & 1 \\ -2 & 4 & 5 \\ -3 & 3 & -1 \end{bmatrix}\right) = \begin{bmatrix} -19 & 2 & -9 \\ -17 & 1 & -12 \\ 6 & -3 & 6 \end{bmatrix}.$$

To find the determinant, we choose one row or column (the first row in this case), multiply the elements by their cofactors, and add the products:

$$\left|\begin{bmatrix} 2 & -1 & 1 \\ -2 & 4 & 5 \\ -3 & 3 & -1 \end{bmatrix}\right| = (2 \times -19) + (-1 \times -17) + (1 \times 6) = -15.$$

We then plug the determinant and the adjoint matrix into the formula for a matrix inverse:

$$\left(\begin{bmatrix} 1 & 3 & -2 \\ 2 & 1 & -1 \\ 2 & 2 & 2 \end{bmatrix}\right)^{-1} = \frac{1}{-15} \begin{bmatrix} -19 & 2 & -9 \\ -17 & 1 & -12 \\ 6 & -3 & 6 \end{bmatrix} = \begin{bmatrix} 1.27 & -0.13 & 0.6 \\ 1.13 & -0.07 & 0.8 \\ -0.4 & 0.2 & -0.4 \end{bmatrix}.$$

The solution to the system of equations is therefore

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1.27 & -0.13 & 0.6 \\ 1.13 & -0.07 & 0.8 \\ -0.4 & 0.2 & -0.4 \end{bmatrix} \begin{bmatrix} 4 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ -1 \end{bmatrix}.$$

2. (a) The augmented matrix that corresponds to this system is

$$\left[\begin{array}{ccc|c} -5 & 5 & 6 & 2 \\ -3 & 3 & -4 & -14 \\ 9 & 6 & 5 & -17 \end{array}\right].$$

We break the first three columns down to an identity matrix using the elementary row operations. Each time we perform an operation, we also apply it to the last column. Once we've created an identity matrix in the first three columns, then the values of the elements in the last column are the (x, y, z) solution to the system.

First we multiply the first row by 3 (to make it easier to eliminate the first element), and interchange the rows to that the first is third, the second is first, and the third is second,

$$\left[\begin{array}{ccc|c} -3 & 3 & -4 & -14 \\ 9 & 6 & 5 & -17 \\ -15 & 15 & 18 & 6 \end{array}\right].$$

Next we multiply the first row by 3 and add it to the second row, and we multiply the first row by -5 and add it to the third row,

$$\left[\begin{array}{ccc|c} -3 & 3 & -4 & -14 \\ 0 & 15 & -7 & -59 \\ 0 & 0 & 38 & 76 \end{array}\right].$$

We divide row 3 by 2,

$$\left[\begin{array}{ccc|c} -3 & 3 & -4 & -14 \\ 0 & 15 & -7 & -59 \\ 0 & 0 & 1 & 2 \end{array}\right],$$

and add 7 times the third row to the second row, and 4 times the third row to the first row,

$$\left[\begin{array}{ccc|c} -3 & 3 & 0 & -6 \\ 0 & 15 & 0 & -45 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

Next we divide the second row by 15,

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

and multiply it by -3 and add it to the first row,

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

Finally we the first row by -3,

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

The reduction implies that the system's solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \\ 2 \end{bmatrix}.$$

(b) The augmented matrix that corresponds to this system is

$$\left[\begin{array}{rrr|r} -6 & -4 & 1 & 10 \\ 1 & 4 & -4 & 16 \\ -3 & 5 & 4 & 19 \end{array}\right].$$

We break the first three columns down to an identity matrix using the elementary row operations. Each time we perform an operation, we also apply it to the last column. Once we've created an identity matrix in the first three columns, then the values of the elements in the last column are the (x, y, z) solution to the system.

First we interchange the rows,

$$\left[\begin{array}{rrr|r} 1 & 4 & -4 & 16 \\ -3 & 5 & 4 & 19 \\ -6 & -4 & 1 & 10 \end{array}\right],$$

then add 3 times the first row to the second row, and 6 times the first row to the third,

$$\left[\begin{array}{rrr|r} 1 & 4 & -4 & 16 \\ 0 & 17 & -8 & 67 \\ 0 & 20 & -23 & 106 \end{array}\right].$$

This next part is very tricky. 17 and 20 share no factors, so to avoid fractions we multiply the second row by 20 and the third row by 17,

$$\left[\begin{array}{rrr|r} 1 & 4 & -4 & 16 \\ 0 & 340 & -160 & 1340 \\ 0 & 340 & -391 & 1802 \end{array}\right],$$

we add -1 times the second row to the third,

$$\left[\begin{array}{rrr|r} 1 & 4 & -4 & 16 \\ 0 & 340 & -160 & 1340 \\ 0 & 0 & -231 & 462 \end{array}\right],$$

and we divide the last row by -231,

$$\left[\begin{array}{rrr|r} 1 & 4 & -4 & 16 \\ 0 & 340 & -160 & 1340 \\ 0 & 0 & 1 & -2 \end{array}\right].$$

Next we multiply the third row by 160 and add it to the second row,

$$\left[\begin{array}{rrr|r} 1 & 4 & -4 & 16 \\ 0 & 340 & 0 & 1020 \\ 0 & 0 & 1 & -2 \end{array}\right],$$

and we multiply the third row by 4 and add it to the first row,

$$\left[\begin{array}{rrr|r} 1 & 4 & 0 & 8 \\ 0 & 340 & 0 & 1020 \\ 0 & 0 & 1 & -2 \end{array}\right].$$

We divide the second row by 340,

$$\left[\begin{array}{ccc|c} 1 & 4 & 0 & 8 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{array}\right],$$

and multiply it by -4 and add it to the first row,

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

The reduction implies that the system's solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ -2 \end{bmatrix}.$$

(c) The augmented matrix that corresponds to this system is

$$\left[\begin{array}{ccc|c} 5 & -1 & -5 & -3 \\ -5 & -5 & 6 & -18 \\ 0 & -2 & -6 & 12 \end{array}\right].$$

We break the first three columns down to an identity matrix using the elementary row operations. Each time we perform an operation, we also apply it to the last column. Once we've created an identity matrix in the first three columns, then the values of the elements in the last column are the (x, y, z) solution to the system.

We start by adding the first row to the second,

$$\left[\begin{array}{ccc|c} 5 & -1 & -5 & -3 \\ 0 & -6 & 1 & -21 \\ 0 & -2 & -6 & 12 \end{array}\right],$$

interchanging the second and third rows,

$$\left[\begin{array}{ccc|c} 5 & -1 & -5 & -3 \\ 0 & -2 & -6 & 12 \\ 0 & -6 & 1 & -21 \end{array}\right],$$

dividing the second row by -2,

$$\left[\begin{array}{ccc|c} 5 & -1 & -5 & -3 \\ 0 & 1 & 3 & -6 \\ 0 & -6 & 1 & -21 \end{array}\right],$$

and adding 6 times the second row to the third,

$$\left[\begin{array}{ccc|c} 5 & -1 & -5 & -3 \\ 0 & 1 & 3 & -6 \\ 0 & 0 & 19 & -57 \end{array}\right].$$

Next we divide the third row by 19,

$$\left[\begin{array}{ccc|c} 5 & -1 & -5 & -3 \\ 0 & 1 & 3 & -6 \\ 0 & 0 & 1 & -3 \end{array}\right],$$

add -3 times the third row to the second,

$$\left[\begin{array}{rrr|r} 5 & -1 & -5 & -3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -3 \end{array}\right],$$

and add 5 times the third row to the first,

$$\left[\begin{array}{rrr|r} 5 & -1 & 0 & -18 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -3 \end{array}\right].$$

Finally we add the second row to the first,

$$\left[\begin{array}{rrr|r} 5 & 0 & 0 & -15 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -3 \end{array}\right],$$

and divide the first row by 5,

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

The reduction implies that the system's solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -3 \\ 3 \\ -3 \end{bmatrix}.$$

(d) The augmented matrix that corresponds to this system is

$$\left[\begin{array}{rrrr|r} -1 & 3 & -6 & -3 & 0 \\ 1 & -1 & 1 & 4 & -11 \\ 5 & -5 & -1 & -2 & -1 \\ 2 & 3 & -3 & 0 & -3 \end{array}\right].$$

We break the first four columns down to an identity matrix using the elementary row operations. Each time we perform an operation, we also apply it to the last column. Once we've created an identity matrix in the first three columns, then the values of the elements in the last column are the (w, x, y, z) solution to the system.

We start by interchanging the first two rows,

$$\left[\begin{array}{rrrr|r} 1 & -1 & 1 & 4 & -11 \\ -1 & 3 & -6 & -3 & 0 \\ 5 & -5 & -1 & -2 & -1 \\ 2 & 3 & -3 & 0 & -3 \end{array}\right],$$

and adding the first row to the second, -5 times the first row to the third, and -2 times the first row to the fourth,

$$\left[\begin{array}{rrrr|r} 1 & -1 & 1 & 4 & -11 \\ 0 & 2 & -5 & 1 & -11 \\ 0 & 0 & -6 & -22 & 54 \\ 0 & 5 & -5 & -8 & 19 \end{array}\right].$$

Next we multiply the fourth row by 2,

$$\left[\begin{array}{rrrr|r} 1 & -1 & 1 & 4 & -11 \\ 0 & 2 & -5 & 1 & -11 \\ 0 & 0 & -6 & -22 & 54 \\ 0 & 10 & -10 & -16 & 38 \end{array}\right],$$

and add -5 times the second row to the fourth row,

$$\left[\begin{array}{rrrr|r} 1 & -1 & 1 & 4 & -11 \\ 0 & 2 & -5 & 1 & -11 \\ 0 & 0 & -6 & -22 & 54 \\ 0 & 0 & 15 & -21 & 93 \end{array}\right].$$

We divide the third row by -2,

$$\left[\begin{array}{rrrr|r} 1 & -1 & 1 & 4 & -11 \\ 0 & 2 & -5 & 1 & -11 \\ 0 & 0 & 3 & 11 & -27 \\ 0 & 0 & 15 & -21 & 93 \end{array}\right],$$

add -5 times the third row to the fourth,

$$\left[\begin{array}{rrrr|r} 1 & -1 & 1 & 4 & -11 \\ 0 & 2 & -5 & 1 & -11 \\ 0 & 0 & 3 & 11 & -27 \\ 0 & 0 & 0 & -76 & 228 \end{array}\right],$$

and divide the fourth row by -76,

$$\left[\begin{array}{rrrr|r} 1 & -1 & 1 & 4 & -11 \\ 0 & 2 & -5 & 1 & -11 \\ 0 & 0 & 3 & 11 & -27 \\ 0 & 0 & 0 & 1 & -3 \end{array}\right].$$

Now we add -11 times the fourth row to the third row, -1 times the fourth row to the second row, and -4 times the fourth row to the first row,

$$\left[\begin{array}{rrrr|r} 1 & -1 & 1 & 0 & 1 \\ 0 & 2 & -5 & 0 & -8 \\ 0 & 0 & 3 & 0 & 6 \\ 0 & 0 & 0 & 1 & -3 \end{array}\right],$$

divide the third row by 3,

$$\left[\begin{array}{rrrr|r} 1 & -1 & 1 & 0 & 1 \\ 0 & 2 & -5 & 0 & -8 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -3 \end{array}\right],$$

add 5 times the third row to the second row and -1 times the third row to the first row,

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

divide the second row by 2,

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

and add the second row to the first,

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

The reduction implies that the system's solution is

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ -3 \end{bmatrix}.$$

3. (a) The augmented matrix that corresponds to this system is

$$\left[\begin{array}{ccc|c} 1 & -2 & -1 & 15 \\ -1 & -1 & 1 & -6 \\ 1 & -6 & -1 & -43 \end{array}\right].$$

We start by adding the first row to the second, and adding -1 times the first row to the third,

$$\left[\begin{array}{ccc|c} 1 & -2 & -1 & 15 \\ 0 & -3 & 0 & 9 \\ 0 & -4 & 0 & -58 \end{array}\right].$$

We divide the second row by -3 and the third row by -4,

$$\left[\begin{array}{ccc|c} 1 & -2 & -1 & 15 \\ 0 & 1 & 0 & -3 \\ 0 & 1 & 0 & 14.5 \end{array}\right],$$

and add -2 times the second row to the first row, and -1 times the second row to the third row,

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 21 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 0 & 20.5 \end{array}\right].$$

Remember that the rows of the augmented matrix refer to equations in the system, that the columns refer to the variables x, y, and z respectively, the elements are coefficients, and the vertical line is an equal sign. At this point we've reduced the system of equations to

$$\begin{cases} x - z = 21, \\ y = -3, \\ 0 = 20.5. \end{cases}$$

The third equation is an obviously untrue statement. Therefore this system of equations has no solution.

(b) The augmented matrix that corresponds to this system is

$$\left[\begin{array}{ccc|c} -4 & -1 & -2 & 15 \\ 1 & -2 & 2 & 18 \\ 0 & -6 & 4 & 20 \end{array}\right].$$

We start by moving the first row to the third position and moving the other two rows up,

$$\left[\begin{array}{rrr|r} 1 & -2 & 2 & 18 \\ 0 & -6 & 4 & 20 \\ -4 & -1 & -2 & 15 \end{array}\right].$$

Next we add 4 times the first row to the third row,

$$\left[\begin{array}{rrr|r} 1 & -2 & 2 & 18 \\ 0 & -6 & 4 & 20 \\ 0 & -9 & 6 & 87 \end{array}\right].$$

We divide the second row by 2,

$$\left[\begin{array}{rrr|r} 1 & -2 & 2 & 18 \\ 0 & -3 & 2 & 10 \\ 0 & -9 & 6 & 87 \end{array}\right],$$

then multiply it by -3 and add it to the third row,

$$\left[\begin{array}{rrr|r} 1 & -2 & 2 & 18 \\ 0 & -3 & 2 & 10 \\ 0 & 0 & 0 & 57 \end{array}\right].$$

The third equation is now 0=57, which is obviously untrue. Therefore the system has no solution.

(c) The augmented matrix that corresponds to this system is

$$\left[\begin{array}{rrr|r} -6 & -2 & 5 & -29 \\ 2 & -5 & 1 & -4 \\ 4 & 5 & -5 & 28 \end{array}\right].$$

Let's start by rearranging the rows,

$$\left[\begin{array}{rrr|r} 2 & -5 & 1 & -4 \\ 4 & 5 & -5 & 28 \\ -6 & -2 & 5 & -29 \end{array}\right],$$

and adding -2 times the first row to the second, and 3 times the first row to the third,

$$\left[\begin{array}{rrr|r} 2 & -5 & 1 & -4 \\ 0 & 15 & -7 & 36 \\ 0 & -17 & 8 & -41 \end{array}\right].$$

Next we multiply the second row by 17 and the third row by 15,

$$\left[\begin{array}{rrr|r} 2 & -5 & 1 & -4 \\ 0 & 255 & -119 & 612 \\ 0 & -255 & 120 & -615 \end{array}\right],$$

add the second row to the third,

$$\left[\begin{array}{rrr|r} 2 & -5 & 1 & -4 \\ 0 & 255 & -119 & 612 \\ 0 & 0 & 1 & -3 \end{array}\right],$$

then add 199 times the third row to the second,

$$\left[\begin{array}{rrr|r} 2 & -5 & 1 & -4 \\ 0 & 255 & 0 & 255 \\ 0 & 0 & 1 & -3 \end{array}\right],$$

and divide the second row by 255,
$$\left[\begin{array}{rrr|r} 2 & -5 & 1 & -4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -3 \end{array}\right].$$
Finally, add -1 times the third row to the first row,
$$\left[\begin{array}{rrr|r} 2 & -5 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -3 \end{array}\right],$$
add 5 times the second row to the first row,
$$\left[\begin{array}{rrr|r} 2 & 0 & 0 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -3 \end{array}\right],$$
and divide the first row by 2,
$$\left[\begin{array}{rrr|r} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -3 \end{array}\right].$$
The reduction implies that the system has one unique solution, which is
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}.$$

(d) The augmented matrix that corresponds to this system is
$$\left[\begin{array}{rrr|r} -1 & 3 & -1 & 9 \\ 1 & -1 & 0 & -8 \\ -5 & 3 & 1 & 39 \end{array}\right].$$
We start by interchanging the first and second rows,
$$\left[\begin{array}{rrr|r} 1 & -1 & 0 & -8 \\ -1 & 3 & -1 & 9 \\ -5 & 3 & 1 & 39 \end{array}\right],$$
adding the first row to the second, and adding 5 times the first row to the third,
$$\left[\begin{array}{rrr|r} 1 & -1 & 0 & -8 \\ 0 & 2 & -1 & 1 \\ 0 & -2 & 1 & -1 \end{array}\right].$$
Next we add the second row to the third,
$$\left[\begin{array}{rrr|r} 1 & -1 & 0 & -8 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right].$$
The third equation in the system is now 0=0, which is an obviously true statement. Therefore this system has infinitely many solutions.

4. (a) The augmented matrix that corresponds to this system is

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 11 \\ -1 & 5 & 2 & 4 \\ 4 & 4 & 1 & 29 \end{array}\right].$$

Let's start by adding the first row to the second, and adding -4 times the first row to the third,

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 11 \\ 0 & 8 & 3 & 15 \\ 0 & -8 & -3 & -15 \end{array}\right].$$

Then we can add the second row to the third,

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 11 \\ 0 & 8 & 3 & 15 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

We can see that this system has infinitely many solutions. Our task is now to derive a formula for these solutions, and to relate the specific solution. Let's continue reducing. We multiply the first row by 8,

$$\left[\begin{array}{ccc|c} 8 & 24 & 8 & 88 \\ 0 & 8 & 3 & 15 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

and add -3 times the second row to the first row,

$$\left[\begin{array}{ccc|c} 8 & 0 & -1 & 43 \\ 0 & 8 & 3 & 15 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Finally, we divide the first and second row by 8,

$$\left[\begin{array}{ccc|c} 1 & 0 & -\frac{1}{8} & \frac{43}{8} \\ 0 & 1 & \frac{3}{8} & \frac{15}{8} \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x - \frac{1}{8}z = \frac{43}{8}, \\ y + \frac{3}{8}z = \frac{15}{8}, \\ z = z, \end{cases}$$

where we add $z = z$ into the system as a general representation of the trivial statement 0=0. That is, by adding or subtracting the same thing to both sides of 0=0, we can turn this statement into $z = z$ for any value of z. Solving the system for x and y in terms of z gives us

$$\begin{cases} x = \frac{1}{8}z + \frac{43}{8}, \\ y = -\frac{3}{8}z + \frac{15}{8}, \\ z = z, \end{cases}$$

which can be written in matrix form as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{8} \\ -\frac{3}{8} \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{43}{8} \\ \frac{15}{8} \\ 1 \end{bmatrix} z.$$

z is a free variable, meaning that we can choose any value for z we want. But for any value of z there is only one solution to this system that can be found by plugging that value of z into the above equation. The specific solution is the name for the solution in which $z = 0$, which in this case is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{8} \\ -\frac{3}{8} \\ 0 \end{bmatrix}.$$

(b) The augmented matrix that corresponds to this system is

$$\left[\begin{array}{ccc|c} 5 & 1 & 4 & -4 \\ -4 & 1 & -3 & -7 \\ -3 & 3 & -2 & -18 \end{array}\right].$$

Let's start by multiplying the second and third rows by 5,

$$\left[\begin{array}{ccc|c} 5 & 1 & 4 & -4 \\ -20 & 5 & -15 & -35 \\ -15 & 15 & -10 & -90 \end{array}\right].$$

Now we can add 4 times the first row to the second row and 3 times the first row to the third row,

$$\left[\begin{array}{ccc|c} 5 & 1 & 4 & -4 \\ 0 & 9 & 1 & -51 \\ 0 & 18 & 2 & -102 \end{array}\right],$$

and add -2 times the second row to the third row,

$$\left[\begin{array}{ccc|c} 5 & 1 & 4 & -4 \\ 0 & 9 & 1 & -51 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The system has infinitely many solutions. Let's continue reducing by multiplying the first row by 9,

$$\left[\begin{array}{ccc|c} 45 & 9 & 36 & -36 \\ 0 & 9 & 1 & -51 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

and adding -1 times the second row to the first,

$$\left[\begin{array}{ccc|c} 45 & 0 & 35 & 15 \\ 0 & 9 & 1 & -51 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Finally, divide the first row by 45 and divide the second row by 9,

$$\left[\begin{array}{ccc|c} 1 & 0 & \frac{7}{9} & \frac{1}{3} \\ 0 & 1 & \frac{1}{9} & -\frac{51}{9} \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x + \frac{7}{9}z = \frac{1}{3}, \\ y + \frac{1}{9}z = -\frac{51}{9}, \\ z = z. \end{cases}$$

Solving the system for x and y in terms of z gives us

$$\begin{cases} x = \frac{1}{3} - \frac{7}{9}z, \\ y = -\frac{51}{9} - \frac{1}{9}z, \\ z = z, \end{cases}$$

which can be written in matrix form as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{51}{9} \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{7}{9} \\ \frac{1}{9} \\ 1 \end{bmatrix} z.$$

The specific solution in which $z = 0$ is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{51}{9} \\ 0 \end{bmatrix}.$$

(c) The augmented matrix that corresponds to this system is

$$\left[\begin{array}{rrr|r} 1 & 6 & -1 & -18 \\ 3 & 2 & 1 & -22 \\ -5 & 6 & -4 & 18 \end{array}\right].$$

Let's start by adding -3 times the first row to the second, and 5 times the first row to the third,

$$\left[\begin{array}{rrr|r} 1 & 6 & -1 & -18 \\ 0 & -16 & 4 & 32 \\ 0 & 36 & -9 & -72 \end{array}\right].$$

We divide the second row by 4 and the third row by 9,

$$\left[\begin{array}{rrr|r} 1 & 6 & -1 & -18 \\ 0 & -4 & 1 & 8 \\ 0 & 4 & -1 & -8 \end{array}\right],$$

and add the second row to the third,

$$\left[\begin{array}{rrr|r} 1 & 6 & -1 & -18 \\ 0 & -4 & 1 & 8 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

This system has infinitely many solutions. To continue reducing, multiply the first row by 2,

$$\left[\begin{array}{rrr|r} 2 & 12 & -2 & -36 \\ 0 & -4 & 1 & 8 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

add 3 times the second row to the first,

$$\left[\begin{array}{rrr|r} 2 & 0 & 1 & -4 \\ 0 & -4 & 1 & 8 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

divide the first row by 2 and divide the second row by -4,

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

The system of equations is now

$$\begin{cases} x + \frac{1}{2}z = -2, \\ y - \frac{1}{4}z = -2, \\ z = z. \end{cases}$$

Solving the system for x and y in terms of z gives us

$$\begin{cases} x = -\frac{1}{2}z - 2, \\ y = -\frac{1}{4}z - 2, \\ z = z, \end{cases}$$

which can be written in matrix form as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \\ 1 \end{bmatrix} z.$$

The specific solution in which $z = 0$ is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \\ 0 \end{bmatrix}.$$

(d) The augmented matrix that corresponds to this system is

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

Note that this system is a homogenous system, which means that it either has infinitely many solutions, or just the trivial (all zero) solution. Let's begin by rearranging the rows,

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

then by adding -1 times the first row to the second, and -2 times the first row to the third,

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -7 & -1 & 0 \\ 0 & -7 & -1 & 0 \end{array}\right].$$

Now add -1 times the second row to the third,

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -7 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

This system has infinitely many solutions. To continue reducing, multiply the top row by 7,

$$\left[\begin{array}{ccc|c} 7 & 7 & -7 & 0 \\ 0 & -7 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

add the second row to the first row,

$$\left[\begin{array}{ccc|c} 7 & 0 & -8 & 0 \\ 0 & -7 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

divide the first row by 7, and divide the second row by -7

$$\left[\begin{array}{ccc|c} 1 & 0 & -\frac{8}{7} & 0 \\ 0 & 1 & \frac{1}{7} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x - \frac{8}{7}z = 0, \\ y + \frac{1}{7}z = 0, \\ z = z. \end{cases}$$

Solving the system for x and y in terms of z gives us

$$\begin{cases} x = \frac{8}{7}z, \\ y = -\frac{1}{7}z, \\ z = z, \end{cases}$$

which can be written in matrix form as

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

The specific solution in which $z = 0$ is the trivial solution,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Another solution in which $z = 1$ is

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

5. Eigenvalues are the constants λ for a matrix A such that

$$Ax = \lambda x,$$

which means that

$$Ax = \lambda I x,$$
$$Ax - \lambda I x = 0,$$
$$(A - \lambda I)x = 0.$$

This equation implies a homogenous system of equations, which only has a non-trivial solution when the determinant is zero,

$$|A - \lambda I| = 0.$$

To find the eigenvalues, we will

- plug A into $(A - \lambda I)$ and write this matrix out,
- then write out the determinant,
- and solve for the values of λ that make the determinant zero.

A matrix is positive-definite if all of its eigenvalues are positive and negative-definite if all of its eigenvalues are negative. To find the unit eigenvectors associated with each matrix, we will

- plug each value of λ back in to $(A - \lambda I)x = 0$,
- write out the augmented matrix for this homogenous system,
- find a non-trivial solution,
- compute the magnitude of this solution vector,
- and divide the vector by its magnitude. This last step ensures that the eigenvector is a unit eigenvector.

(a) $\begin{bmatrix} -4 & 4 \\ 0 & -8 \end{bmatrix}$

$$A - \lambda I = \begin{bmatrix} -4 & 4 \\ 0 & -8 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -4-\lambda & 4 \\ 0 & -8-\lambda \end{bmatrix}$$

$$\begin{aligned} |A - \lambda I| &= (-4-\lambda)(-8-\lambda) - (4 \times 0) \\ &= (\lambda+8)(\lambda+4) = 0, \\ \lambda &= -8, -4. \end{aligned}$$

Since this matrix has only negative eigenvalues, it is negative-definite. The system of equations implied by $(A - \lambda I)x = 0$ is

$$\begin{cases} (-4-\lambda)x + 4y = 0, \\ -8 - \lambda y = 0 \end{cases}$$

with the augmented matrix

$$\left[\begin{array}{cc|c} -4-\lambda & 4 & 0 \\ 0 & -8-\lambda & 0 \end{array}\right].$$

Plugging in the first eigenvalue $\lambda = -8$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} 4 & 4 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

which we reduce with elementary row operations by dividing the first row by 4,

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

The system of equations is now

$$\begin{cases} x + y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = -y, \\ y = y, \end{cases}$$

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 1$ is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, which has magnitude

$$\left| \left| \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right| \right| = \sqrt{(-1)^2 + 1^2} = \sqrt{2}.$$

To find the unit eigenvector associated with $\lambda = -8$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

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

Plugging in the second eigenvalue $\lambda = -4$ turns the augmented matrix into

$$\left[\begin{array}{cc|c} 0 & 4 & 0 \\ 0 & -4 & 0 \end{array}\right],$$

which we reduce with elementary row operations by adding the second row to the first,

$$\left[\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & -4 & 0 \end{array}\right],$$

then dividing the second row by -4,

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

Since the first row corresponds to x, and it reduces to the obviously true statement 0=0, we think of x to be a free variable in which $x = x$. The system of equations is now

$$\begin{cases} x = x, \\ y = 0, \end{cases}$$

The system can be rewritten in matrix form as

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

One eigenvector that arises when $x = 1$ is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, which has magnitude

$$\left| \left| \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right| \right| = \sqrt{1^2 + 0^2} = 1,$$

so it is the unit eigenvector associated with $\lambda = -4$.

(b) $\begin{bmatrix} -3 & 0 \\ 2 & 7 \end{bmatrix}$

$$A - \lambda I = \begin{bmatrix} -3 & 0 \\ 2 & 7 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -3-\lambda & 0 \\ 2 & 7-\lambda \end{bmatrix}$$

$$\begin{aligned} |A - \lambda I| &= (-3-\lambda)(7-\lambda) - (0 \times 2) \\ &= (\lambda+3)(\lambda-7) = 0, \\ \lambda &= -3,\, 7. \end{aligned}$$

Since this matrix has one positive and one negative eigenvalue, it is neither positive-definite nor negative-definite. The system of equations implied by $(A - \lambda I)x = 0$ is

$$\begin{cases} (-3-\lambda)x = 0, \\ 2x + (7-\lambda)y = 0 \end{cases}$$

with the augmented matrix

$$\left[\begin{array}{cc|c} -3-\lambda & 0 & 0 \\ 2 & 7-\lambda & 0 \end{array}\right].$$

Plugging in the first eigenvalue $\lambda = -3$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} 0 & 0 & 0 \\ 2 & 10 & 0 \end{array}\right],$$

which we reduce with elementary row operations by interchanging the first and second rows,

$$\left[\begin{array}{cc|c} 2 & 10 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

then dividing the first row by 2,

$$\left[\begin{array}{cc|c} 1 & 5 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x + 5y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = -5y, \\ y = y, \end{cases}$$

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -5 \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 1$ is $\begin{bmatrix} -5 \\ 1 \end{bmatrix}$, which has magnitude

$$\left\| \begin{bmatrix} -5 \\ 1 \end{bmatrix} \right\| = \sqrt{(-5)^2 + 1^2} = \sqrt{26}.$$

To find the unit eigenvector associated with $\lambda = -3$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\begin{bmatrix} -\frac{5}{\sqrt{26}} \\ \\ \frac{1}{\sqrt{26}} \end{bmatrix}.$$

Plugging in the second eigenvalue $\lambda = 7$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} -10 & 0 & 0 \\ 2 & 0 & 0 \end{array}\right],$$

which we reduce with elementary row operations by dividing the first row by -10,

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 2 & 0 & 0 \end{array}\right],$$

then adding -2 times the first row to the second row,

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

The system of equations is now

$$\begin{cases} x = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten in matrix form as

$$\left[\begin{array}{c} x \\ y \end{array}\right] = y\left[\begin{array}{c} 0 \\ 1 \end{array}\right].$$

One eigenvector that arises when $y = 1$ is $\left[\begin{array}{c} 0 \\ 1 \end{array}\right]$, which has magnitude

$$\left|\left|\left[\begin{array}{c} 0 \\ 1 \end{array}\right]\right|\right| = \sqrt{0^2 + 1^2} = 1,$$

so this vector is the unit eigenvector associated with $\lambda = 7$.

(c) $\left[\begin{array}{cc} 6 & 7 \\ -5 & -6 \end{array}\right]$

$$A - \lambda I = \left[\begin{array}{cc} 6 & 7 \\ -5 & -6 \end{array}\right] - \lambda\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] = \left[\begin{array}{cc} 6-\lambda & 7 \\ -5 & -6-\lambda \end{array}\right]$$

$$\begin{aligned} |A - \lambda I| &= (6-\lambda)(-6-\lambda) - (7 \times -5) \\ &= (\lambda - 6)(\lambda + 6) + 35 \\ &= \lambda^2 - 36 + 35 \\ &= \lambda^2 - 1 \\ &= (\lambda - 1)(\lambda + 1) = 0, \\ \lambda &= -1, 1. \end{aligned}$$

Since this matrix has one positive and one negative eigenvalue, it is neither positive-definite nor negative-definite. The system of equations implied by $(A - \lambda I)x = 0$ is

$$\begin{cases} (6-\lambda)x + 7y = 0, \\ -5x + (-6-\lambda)y = 0 \end{cases}$$

with the augmented matrix

$$\left[\begin{array}{cc|c} 6-\lambda & 7 & 0 \\ -5 & -6-\lambda & 0 \end{array}\right].$$

Plugging in the first eigenvalue $\lambda = -1$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} 7 & 7 & 0 \\ -5 & -5 & 0 \end{array}\right],$$

which we reduce with elementary row operations by dividing the first row by 7,

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ -5 & -5 & 0 \end{array}\right],$$

and adding 5 times the first row to the second,

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

The system of equations is now

$$\begin{cases} x + y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = -y, \\ y = y, \end{cases}$$

and in matrix form as

$$\left[\begin{array}{c} x \\ y \end{array}\right] = y \left[\begin{array}{r} -1 \\ 1 \end{array}\right].$$

One eigenvector that arises when $y = 1$ is $\left[\begin{array}{r} -1 \\ 1 \end{array}\right]$, which has magnitude

$$\left|\left[\begin{array}{r} -1 \\ 1 \end{array}\right]\right| = \sqrt{(-1)^2 + 1^2} = \sqrt{2}.$$

To find the unit eigenvector associated with $\lambda = -1$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\left[\begin{array}{c} -\frac{1}{\sqrt{2}} \\ \\ \frac{1}{\sqrt{2}} \end{array}\right].$$

Plugging in the first eigenvalue $\lambda = 1$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} 5 & 7 & 0 \\ -5 & -7 & 0 \end{array}\right],$$

which we reduce with elementary row operations by adding the first row to the second,

$$\left[\begin{array}{cc|c} 5 & 7 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

and dividing the first row by 5,

$$\left[\begin{array}{cc|c} 1 & \frac{7}{5} & 0 \\ \\ 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x + \frac{7}{5}y = 0, \\ \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = -\frac{7}{5}y, \\ \\ y = y, \end{cases}$$

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -\frac{7}{5} \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 5$ is $\begin{bmatrix} -7 \\ 5 \end{bmatrix}$, which has magnitude

$$\left\| \begin{bmatrix} -7 \\ 5 \end{bmatrix} \right\| = \sqrt{(-7)^2 + 5^2} = \sqrt{74}.$$

To find the unit eigenvector associated with $\lambda = -1$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\begin{bmatrix} -\frac{7}{\sqrt{74}} \\ \frac{5}{\sqrt{74}} \end{bmatrix}.$$

(d) $\begin{bmatrix} 9 & 3 \\ 0 & -1 \end{bmatrix}$

$$A - \lambda I = \begin{bmatrix} 9 & 3 \\ 0 & -1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 9-\lambda & 3 \\ 0 & -1-\lambda \end{bmatrix}$$

$$\begin{aligned} |A - \lambda I| &= (9-\lambda)(-1-\lambda) - (3 \times 0) \\ &= (\lambda - 9)(\lambda + 1) = 0, \\ \lambda &= -1, 9. \end{aligned}$$

Since this matrix has one positive and one negative eigenvalue, it is neither positive-definite nor negative-definite. The system of equations implied by $(A - \lambda I)x = 0$ is

$$\begin{cases} (9-\lambda)x + 3y = 0, \\ (-1-\lambda)y = 0 \end{cases}$$

with the augmented matrix

$$\left[\begin{array}{cc|c} 9-\lambda & 3 & 0 \\ 0 & -1-\lambda & 0 \end{array}\right].$$

Plugging in the first eigenvalue $\lambda = -1$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} 10 & 3 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

which we reduce with elementary row operations by dividing the first row by 10,

$$\left[\begin{array}{cc|c} 1 & \frac{3}{10} & 0 \\ 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x + \frac{3}{10}y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = -\frac{3}{10}y, \\ y = y, \end{cases}$$

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -\frac{3}{10} \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 10$ is $\begin{bmatrix} -3 \\ 10 \end{bmatrix}$, which has magnitude

$$\left| \left| \begin{bmatrix} -3 \\ 10 \end{bmatrix} \right| \right| = \sqrt{(-3)^2 + 10^2} = \sqrt{109}.$$

To find the unit eigenvector associated with $\lambda = -1$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\begin{bmatrix} -\frac{3}{\sqrt{109}} \\ \\ \frac{10}{\sqrt{109}} \end{bmatrix}.$$

Plugging in the second eigenvalue $\lambda = 9$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} 0 & 3 & 0 \\ 0 & -10 & 0 \end{array}\right],$$

which we reduce with elementary row operations by dividing the first row by 3 and the second row by -10,

$$\left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 1 & 0 \end{array}\right],$$

and adding -1 times the second row to the first,

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

The system of equations is now

$$\begin{cases} x = x, \\ y = 0, \end{cases}$$

where x is the free variable. The system can be rewritten in matrix form as

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

One eigenvector that arises when $x = 1$ is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, which has magnitude

$$\left| \left| \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right| \right| = \sqrt{1^2 + 0^2} = 1,$$

so this vector is the unit eigenvector associated with $\lambda = 9$.

(e) $\begin{bmatrix} -6 & -8 \\ -8 & -6 \end{bmatrix}$

$$A - \lambda I = \begin{bmatrix} -6 & -8 \\ -8 & -6 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -6-\lambda & -8 \\ -8 & -6-\lambda \end{bmatrix}$$

$$\begin{aligned} |A - \lambda I| &= (-6-\lambda)^2 - (-8)^2 \\ &= (\lambda + 6)^2 - 64 \\ &= (\lambda^2 + 12\lambda + 36) - 64 \\ &= \lambda^2 + 12\lambda - 28 \\ &= (\lambda + 14)(\lambda - 2) = 0, \\ \lambda &= -14,\ 2. \end{aligned}$$

Since this matrix has one positive and one negative eigenvalue, it is neither positive-definite nor negative-definite. The system of equations implied by $(A - \lambda I)x = 0$ is

$$\begin{cases} (-6-\lambda)x - 8y = 0, \\ -8x + (-6-\lambda)y = 0 \end{cases}$$

with the augmented matrix

$$\left[\begin{array}{rr|r} -6-\lambda & -8 & 0 \\ -8 & -6-\lambda & 0 \end{array}\right].$$

Plugging in the first eigenvalue $\lambda = -14$ turns this augmented matrix into

$$\left[\begin{array}{rr|r} 8 & -8 & 0 \\ -8 & 8 & 0 \end{array}\right],$$

which we reduce with elementary row operations by adding the first row to the second,

$$\left[\begin{array}{rr|r} 8 & -8 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

and dividing the first row by 8,

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

The system of equations is now

$$\begin{cases} x - y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

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

and in matrix form as

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

One eigenvector that arises when $y = 1$ is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, which has magnitude

$$\left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\| = \sqrt{1^2 + 1^2} = \sqrt{2}.$$

To find the unit eigenvector associated with $\lambda = -14$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Plugging in the second eigenvalue $\lambda = 2$ turns this augmented matrix into

$$\left[\begin{array}{rr|r} -8 & -8 & 0 \\ -8 & -8 & 0 \end{array}\right],$$

which we reduce with elementary row operations by adding the first row to the second,

$$\left[\begin{array}{rr|r} -8 & -8 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

and dividing the first row by -8,

$$\left[\begin{array}{rr|r} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x + y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = -y, \\ y = y, \end{cases}$$

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 1$ is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, which has magnitude

$$\left|\left| \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right|\right| = \sqrt{1^2 + 1^2} = \sqrt{2}.$$

To find the unit eigenvector associated with $\lambda = 2$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

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

(f) $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$

$$A - \lambda I = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1-\lambda & 3 \\ 2 & 6-\lambda \end{bmatrix}$$

$$\begin{aligned} |A - \lambda I| &= (1-\lambda)(6-\lambda) - (3 \times 2) \\ &= (\lambda^2 - 7\lambda + 6) - 6 \\ &= \lambda^2 - 7\lambda \\ &= \lambda(\lambda - 7) = 0, \\ \lambda &= 0, 7. \end{aligned}$$

This matrix is not negative-definite because it has a negative eigenvalue. But it's also not positive-definite because 0 is not, strictly speaking, a positive number. We can say that this matrix is positive-*semi*definite. The system of equations implied by $(A - \lambda I)x = 0$ is

$$\begin{cases} (1-\lambda)x + 3y = 0, \\ 2x + (6-\lambda)y = 0 \end{cases}$$

with the augmented matrix

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

Plugging in the first eigenvalue $\lambda = 0$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} 1 & 3 & 0 \\ 2 & 6 & 0 \end{array}\right],$$

which we reduce with elementary row operations by adding -2 times the first row to the second,

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

The system of equations is now

$$\begin{cases} x + 3y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

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

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -3 \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 1$ is $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$, which has magnitude

$$\left| \left| \begin{bmatrix} -3 \\ 1 \end{bmatrix} \right| \right| = \sqrt{(-3)^2 + 1^2} = \sqrt{10}.$$

To find the unit eigenvector associated with $\lambda = 0$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\begin{bmatrix} -\frac{3}{\sqrt{10}} \\ \\ \frac{1}{\sqrt{10}} \end{bmatrix}.$$

In other words, the matrix $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$, when left-multiplied by any vector that is a multiple of $\begin{bmatrix} -\frac{3}{\sqrt{10}} \\ \\ \frac{1}{\sqrt{10}} \end{bmatrix}$, always outputs a vector of all zeroes.

Plugging in the second eigenvalue $\lambda = 7$ turns this augmented matrix into

$$\left[\begin{array}{rr|r} -6 & 3 & 0 \\ 2 & -1 & 0 \end{array}\right],$$

which we reduce with elementary row operations by dividing the first row by -3,

$$\left[\begin{array}{rr|r} 2 & -1 & 0 \\ 2 & -1 & 0 \end{array}\right],$$

then multiplying the first row by -1 and adding it to the second row,

$$\left[\begin{array}{rr|r} 2 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

and finally dividing the first row by 2,

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

The system of equations is now

$$\begin{cases} x - \frac{1}{2}y = 0, \\ \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = \frac{1}{2}y, \\ \\ y = y, \end{cases}$$

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 2$ is $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, which has magnitude

$$\left\| \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\| = \sqrt{1^2 + 2^2} = \sqrt{5}.$$

To find the unit eigenvector associated with $\lambda = 7$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}.$$

(g) $\begin{bmatrix} 8 & -10 \\ 0 & -9 \end{bmatrix}$

$$A - \lambda I = \begin{bmatrix} 8 & -10 \\ 0 & -9 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 8-\lambda & -10 \\ 0 & -9-\lambda \end{bmatrix}$$

$$\begin{aligned} |A - \lambda I| &= (8-\lambda)(-9-\lambda) - (-10 \times 0) \\ &= (\lambda - 8)(\lambda + 9) = 0, \\ \lambda &= -9, 8. \end{aligned}$$

Since this matrix has one positive and one negative eigenvalue, it is neither positive-definite nor negative-definite. The system of equations implied by $(A - \lambda I)x = 0$ is

$$\begin{cases} (8-\lambda)x - 10y = 0, \\ (-9-\lambda)y = 0 \end{cases}$$

with the augmented matrix

$$\left[\begin{array}{cc|c} 8-\lambda & -10 & 0 \\ 0 & -9-\lambda & 0 \end{array}\right].$$

Plugging in the first eigenvalue $\lambda = -9$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} 17 & -10 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

which we reduce with elementary row operations by dividing the first row by 17,

$$\left[\begin{array}{cc|c} 1 & -\frac{10}{17} & 0 \\ 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x - \frac{10}{17}y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = \frac{10}{17}y, \\ y = y, \end{cases}$$

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} \frac{10}{17} \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 17$ is $\begin{bmatrix} 10 \\ 17 \end{bmatrix}$, which has magnitude

$$\left|\left| \begin{bmatrix} 10 \\ 17 \end{bmatrix} \right|\right| = \sqrt{10^2 + 17^2} = \sqrt{389}.$$

To find the unit eigenvector associated with $\lambda = -9$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\begin{bmatrix} \frac{10}{\sqrt{389}} \\ \frac{17}{\sqrt{389}} \end{bmatrix}.$$

Plugging in the second eigenvalue $\lambda = 8$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} 0 & -10 & 0 \\ 0 & -17 & 0 \end{array}\right],$$

which we reduce with elementary row operations by dividing the second row by -17,

$$\left[\begin{array}{cc|c} 0 & -10 & 0 \\ 0 & 1 & 0 \end{array}\right],$$

and adding 10 times the second row to the first row,

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

The system of equations is now

$$\begin{cases} x = x, \\ y = 0, \end{cases}$$

where x is the free variable. The system can be rewritten in matrix form as

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

One eigenvector that arises when $x = 1$ is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, which has magnitude

$$\left|\left| \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right|\right| = \sqrt{1^2 + 0^2} = 1,$$

so this vector is the unit eigenvector associated with $\lambda = 8$.

(h) $\begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix}$

$$A - \lambda I = \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7-\lambda & 0 \\ 0 & 7-\lambda \end{bmatrix}$$

$$\begin{aligned} |A - \lambda I| &= (7-\lambda)^2 - (0 \times 0) \\ &= (7-\lambda)^2 = 0, \\ \lambda &= 7. \end{aligned}$$

Since this matrix has only positive eigenvalues, it is positive-definite. The system of equations implied by $(A - \lambda I)x = 0$ is

$$\begin{cases} (7-\lambda)x = 0, \\ (7-\lambda)y = 0 \end{cases}$$

with the augmented matrix

$$\left[\begin{array}{rr|r} 7-\lambda & 0 & 0 \\ 0 & 7-\lambda & 0 \end{array}\right].$$

Plugging in the only eigenvalue $\lambda = 7$ turns this augmented matrix into

$$\left[\begin{array}{rr|r} 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

which cannot be reduced. The system of equations is now

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

where x and y are both free variables. The system can be rewritten in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix},$$

which means that **any** two-dimensional, real-numbered vector is an eigenvector of this matrix associated with the eigenvalue $\lambda = 7$. To find the unit eigenvectors, we write out the magnitude:

$$\left|\left|\begin{bmatrix} x \\ y \end{bmatrix}\right|\right| = \sqrt{x^2 + y^2}.$$

The unit eigenvalues are all of the form

$$\begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} \\ \\ \frac{y}{\sqrt{x^2+y^2}} \end{bmatrix},$$

which is defined for any values of x and y other than $(0, 0)$ since that would divide each element by 0. For example, the unit eigenvector we derive when $x = 2$ and $y = 5$ is

$$\begin{bmatrix} \frac{2}{\sqrt{2^2+5^2}} \\ \\ \frac{5}{\sqrt{2^2+5^2}} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{29}} \\ \\ \frac{5}{\sqrt{29}} \end{bmatrix}.$$

(i) $\begin{bmatrix} 0 & -3 \\ -6 & -4 \end{bmatrix}$

$$A - \lambda I = \begin{bmatrix} 0 & -3 \\ -6 & -4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0-\lambda & -3 \\ -6 & -4-\lambda \end{bmatrix}$$

$$\begin{aligned} |A - \lambda I| &= (-\lambda)(-4-\lambda) - (-3 \times -6) \\ &= 4\lambda + \lambda^2 - 18 \\ &= \lambda^2 + 4\lambda - 18 = 0. \end{aligned}$$

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-4 \pm \sqrt{16 - 4(1)(-18)}}{2}$$

$$= \frac{-4 \pm \sqrt{88}}{2} = \frac{-4 \pm 2\sqrt{22}}{2} = -2 \pm \sqrt{22},$$

$$\lambda = -2 - \sqrt{22} = -6.69, \qquad \lambda = -2 + \sqrt{22} = 2.69.$$

Since this matrix has one positive and one negative eigenvalue, it is neither positive-definite nor negative-definite. The system of equations implied by $(A - \lambda I)x = 0$ is

$$\begin{cases} -\lambda x - 3y = 0, \\ -6x + (-4 - \lambda)y = 0 \end{cases}$$

with the augmented matrix

$$\left[\begin{array}{cc|c} -\lambda & -3 & 0 \\ -6 & -4-\lambda & 0 \end{array}\right].$$

Plugging in the first eigenvalue $\lambda = -6.69$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} 6.69 & -3 & 0 \\ -6 & 2.69 & 0 \end{array}\right],$$

which we reduce with elementary row operations by diving the first row by 6.69,

$$\left[\begin{array}{cc|c} 1 & -.45 & 0 \\ -6 & 2.69 & 0 \end{array}\right],$$

dividing the second row by -6,

$$\left[\begin{array}{cc|c} 1 & -.45 & 0 \\ 1 & -.45 & 0 \end{array}\right],$$

and adding -1 times the first row to the second row,

$$\left[\begin{array}{cc|c} 1 & -.45 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x - .45y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = .45y, \\ y = y, \end{cases}$$

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} .45 \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 1$ is $\begin{bmatrix} .45 \\ 1 \end{bmatrix}$, which has magnitude

$$\left| \left| \begin{bmatrix} .45 \\ 1 \end{bmatrix} \right| \right| = \sqrt{.45^2 + 1^2} = \sqrt{1.2} = 1.095.$$

To find the unit eigenvector associated with $\lambda = -6.69$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\begin{bmatrix} \frac{.45}{1.095} \\ \\ \frac{1}{1.095} \end{bmatrix} = \begin{bmatrix} .41 \\ .91 \end{bmatrix}.$$

Plugging in the second eigenvalue $\lambda = 2.69$ turns this augmented matrix into

$$\left[\begin{array}{cc|c} -2.69 & -3 & 0 \\ -6 & -6.69 & 0 \end{array}\right],$$

which we reduce with elementary row operations by diving the first row by -2.69,

$$\left[\begin{array}{rr|r} 1 & 1.12 & 0 \\ -6 & -5.69 & 0 \end{array}\right],$$

dividing the second row by -6,

$$\left[\begin{array}{rr|r} 1 & 1.12 & 0 \\ 1 & 1.12 & 0 \end{array}\right],$$

and adding -1 times the first row to the second row,

$$\left[\begin{array}{rr|r} 1 & 1.12 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x + 1.12y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = -1.12y, \\ y = y, \end{cases}$$

and in matrix form as

$$\left[\begin{array}{c} x \\ y \end{array}\right] = y \left[\begin{array}{r} -1.12 \\ 1 \end{array}\right].$$

One eigenvector that arises when $y = 1$ is $\left[\begin{array}{r} -1.12 \\ 1 \end{array}\right]$, which has magnitude

$$\left|\left| \left[\begin{array}{r} -1.12 \\ 1 \end{array}\right] \right|\right| = \sqrt{(-1.12)^2 + 1^2} = \sqrt{2.25} = 1.5.$$

To find the unit eigenvector associated with $\lambda = 2.69$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\left[\begin{array}{c} -\frac{1.12}{1.5} \\ \\ \frac{1}{1.5} \end{array}\right] = \left[\begin{array}{r} -.75 \\ .67 \end{array}\right].$$

(j) $\left[\begin{array}{rr} -5 & 2 \\ 2 & 6 \end{array}\right]$

$$A - \lambda I = \left[\begin{array}{rr} -5 & 2 \\ 2 & 6 \end{array}\right] - \lambda \left[\begin{array}{rr} 1 & 0 \\ 0 & 1 \end{array}\right] = \left[\begin{array}{rr} -5-\lambda & 2 \\ 2 & 6-\lambda \end{array}\right]$$

$$\begin{aligned} |A - \lambda I| &= (-5-\lambda)(6-\lambda) - (2 \times 2) \\ &= (\lambda + 5)(\lambda - 6) - 4 \\ &= (\lambda^2 - \lambda - 30) + 4 \\ &= \lambda^2 - \lambda - 26 = 0. \end{aligned}$$

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{1 - 4(1)(-26)}}{2} = \frac{1 \pm \sqrt{105}}{2}.$$

$$\lambda = \frac{1 - \sqrt{105}}{2} = -5.35, \qquad \lambda = \frac{1 + \sqrt{105}}{2} = 6.35.$$

Since this matrix has one positive and one negative eigenvalue, it is neither positive-definite nor negative-definite. The system of equations implied by $(A - \lambda I)x = 0$ is

$$\begin{cases} (-5-\lambda)x + 2y = 0, \\ 2x + (6-\lambda)y = 0 \end{cases}$$

with the augmented matrix

$$\left[\begin{array}{rr|r} -5-\lambda & 2 & 0 \\ 2 & 6-\lambda & 0 \end{array}\right].$$

Plugging in the first eigenvalue $\lambda = -5.35$ turns this augmented matrix into

$$\left[\begin{array}{rr|r} .35 & 2 & 0 \\ 2 & 11.35 & 0 \end{array}\right],$$

which we reduce with elementary row operations by diving the first row by .35,

$$\left[\begin{array}{rr|r} 1 & 5.7 & 0 \\ 2 & 11.35 & 0 \end{array}\right],$$

dividing the second row by 2,

$$\left[\begin{array}{rr|r} 1 & 5.7 & 0 \\ 1 & 5.7 & 0 \end{array}\right],$$

and adding -1 times the first row to the second row,

$$\left[\begin{array}{rr|r} 1 & 5.7 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x + 5.7y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = -5.7y, \\ y = y, \end{cases}$$

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -5.7 \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 1$ is $\begin{bmatrix} -5.7 \\ 1 \end{bmatrix}$, which has magnitude

$$\left|\left| \begin{bmatrix} -5.7 \\ 1 \end{bmatrix} \right|\right| = \sqrt{(-5.7)^2 + 1^2} = \sqrt{33.49} = 5.79.$$

To find the unit eigenvector associated with $\lambda = -5.35$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\begin{bmatrix} -\frac{5.7}{5.79} \\ \frac{1}{5.79} \end{bmatrix} = \begin{bmatrix} -.98 \\ .17 \end{bmatrix}.$$

Plugging in the second eigenvalue $\lambda = 6.35$ turns this augmented matrix into

$$\left[\begin{array}{rr|r} -11.35 & 2 & 0 \\ 2 & -.35 & 0 \end{array}\right],$$

which we reduce with elementary row operations by diving the first row by -11.35,

$$\left[\begin{array}{rr|r} 1 & -.18 & 0 \\ 2 & -.35 & 0 \end{array}\right],$$

dividing the second row by 2,

$$\left[\begin{array}{cc|c} 1 & -.18 & 0 \\ 1 & -.18 & 0 \end{array}\right],$$

and adding -1 times the first row to the second row,

$$\left[\begin{array}{cc|c} 1 & -.18 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

The system of equations is now

$$\begin{cases} x - .18y = 0, \\ y = y, \end{cases}$$

where y is a free variable. The system can be rewritten as

$$\begin{cases} x = .18y, \\ y = y, \end{cases}$$

and in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} .18 \\ 1 \end{bmatrix}.$$

One eigenvector that arises when $y = 1$ is $\begin{bmatrix} .18 \\ 1 \end{bmatrix}$, which has magnitude

$$\left|\left| \begin{bmatrix} .18 \\ 1 \end{bmatrix} \right|\right| = \sqrt{(.18)^2 + 1^2} = \sqrt{1.03} = 1.02.$$

To find the unit eigenvector associated with $\lambda = 6.35$, we divide an eigenvector by its magnitude. In this case the unit eigenvector is

$$\begin{bmatrix} \frac{.18}{1.02} \\ \\ \frac{1}{1.02} \end{bmatrix} = \begin{bmatrix} .17 \\ .98 \end{bmatrix}.$$

6. (a) To find the gradient, we first take the partial derivative with respect to x,

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial}{\partial x}\left(xy + 2x + y + x^2 + 2y^3\right) \\ &= y + 2 + 2x, \end{aligned}$$

and the partial derivative with respect to y,

$$\begin{aligned} \frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}\left(xy + 2x + y + x^2 + 2y^3\right) \\ &= x + 1 + 6y^2. \end{aligned}$$

The gradient is therefore

$$\nabla f(x, y) = \begin{bmatrix} y + 2 + 2x \\ x + 1 + 6y^2 \end{bmatrix}.$$

(b) The Hessian is given by the (2×2) matrix

$$H\Big(f(x, y)\Big) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}.$$

To find the (1,1) element, the second partial derivative with respect to x and then x again, we take the partial derivative of $\frac{\partial f}{\partial x}$ with respect to x:

$$\frac{\partial f}{\partial x}\left(y+2+2x\right) = 2.$$

To find the (1,2) and (2,1) elements, the second partial derivative with respect to x and then y, we either take the partial derivative of $\frac{\partial f}{\partial x}$ with respect to y, or the partial derivative of $\frac{\partial f}{\partial y}$ with respect to x. These two derivatives are both equal to:

$$\frac{\partial f}{\partial y}\left(y+2+2x\right) = 1.$$

Finally, to find the (2,2) element, the second partial derivative with respect to y and then y again, we take the partial derivative of $\frac{\partial f}{\partial y}$ with respect to y:

$$\frac{\partial f}{\partial y}\left(x+1+6y^2\right) = 12y.$$

So the Hessian is

$$H\left(f(x,y)\right) = \begin{bmatrix} 2 & 1 \\ 1 & 12y \end{bmatrix}.$$

(c) To find the critical points, we set each element of the gradient simultaneously equal to 0, and solve for the values of x and y that make this true. The system of equations is

$$\begin{cases} y+2+2x = 0, \\ x+1+6y^2 = 0. \end{cases}$$

There are many ways to solve this system, but here's one approach. We start by solving the first equation for y,

$$y = -2x-2,$$

and substituting for y in the second equation,

$$x+1+6(-2x-2)^2 = 0,$$

$$x+1+6(4x^2+8x+4) = 0,$$
$$x+1+24x^2+48x+24 = 0,$$
$$24x^2+49x+25 = 0,$$

We can factor this quadratic expression according to the steps outlined in section 1.7.2. This quadratic expression is ax^2+bx+c where $a=24$, $b=49$, and $c=25$. First we multiply a and c together

$$a \times c = 600,$$

find all pairs of integer factors that also multiply to this product:

$$\begin{matrix} 1\times 600, & 2\times 300, & 3\times 200, & 4\times 150, & 5\times 120, & 6\times 100, \\ 8\times 75, & 10\times 60, & 12\times 50, & 15\times 40, & 20\times 30, & 24\times 25, \end{matrix}$$

and look for a pair that adds to $b=49$. In this case such a pair is 24 and 25. Then we break the middle term of the quadratic expression into two addends equal to these two factors,

$$24x^2+(24x+25x)+25 = 0,$$

place parentheses around the first two and last two terms,

$$(24x^2 + 24x) + (25x + 25) = 0,$$

and pull all common factors outside each set of parentheses,

$$24x(x + 1) + 25(x + 1) = 0.$$

Finally we pull the common parenthetical factor out of both terms,

$$(x + 1)(24x + 25) = 0.$$

The values of x that solve this equation are -1 and $-\frac{25}{24}$. When $x = -1$, then y equals

$$y = -2(-1) - 2 = 0,$$

and when $x = -\frac{25}{24}$, y equals

$$y = -2\left(-\frac{25}{24}\right) - 2 = \frac{25}{12} - \frac{24}{12} = \frac{1}{12}.$$

Therefore the points $(x, y) = (-1, 0)$ and $(x, y) = \left(\frac{25}{12}, \frac{1}{12}\right)$ are critical points for the function.

(d) A critical point represents a local maximum if the Hessian is negative-definite after plugging in the critical point, and the critical point represents a local minimum if the Hessian is positive-definite after plugging in the critical point. If the Hessian is neither negative-definite nor positive-definite, then the critical point represents a saddle point. A matrix is negative-definite when all of its eigenvalues are negative, and positive-definite when all of its eigenvalues are positive. So here we have to first plug each critical point we found in part (c) into Hessian we found in part (b), then find the eigenvalues of the resulting matrix, and check whether the eigenvalues are all negative or all positive.
Plugging the critical point (-1,0) into the Hessian gives us

$$H\Big(f(-1, 0)\Big) = \begin{bmatrix} 2 & 1 \\ 1 & 12(0) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}.$$

To find the eigenvalues, we write

$$A - \lambda I = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2-\lambda & 1 \\ 1 & -\lambda \end{bmatrix},$$

and find the values of λ that make the determinant of this matrix equal to 0,

$$\begin{aligned} |A - \lambda I| &= (2-\lambda)(-\lambda) - (1 \times 1) \\ &= \lambda^2 - 2\lambda - 1 = 0. \end{aligned}$$

In this case we have to use the quadratic formula to solve for λ,

$$\begin{aligned} \lambda = \frac{2 \pm \sqrt{4 - 4(1)(-1)}}{2}, \\ = \frac{2 \pm \sqrt{8}}{2}, \\ = \frac{2 \pm 2\sqrt{2}}{2}, \end{aligned}$$

$$= 1 \pm \sqrt{2}.$$

$$\lambda = 1 - \sqrt{2} = -.42, \qquad \lambda = 1 + \sqrt{2} = 2.42.$$

Since this matrix has one positive and one negative eigenvalue, it is neither negative-definite nor positive-definite and the critical point (-1,0) represents a saddle point.

Plugging the critical point $\left(\frac{25}{12}, \frac{1}{12}\right)$ into the Hessian gives us

$$H\Big(f(-1,0)\Big) = \begin{bmatrix} 2 & 1 \\ 1 & 12\left(\frac{1}{12}\right) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

To find the eigenvalues, we write

$$A - \lambda I = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix},$$

and find the values of λ that make the determinant of this matrix equal to 0,

$$\begin{aligned} |A - \lambda I| &= (2-\lambda)(1-\lambda) - (1 \times 1) \\ &= (\lambda^2 - 3\lambda + 2) - 1 \\ &= \lambda^2 - 3\lambda + 1 = 0. \end{aligned}$$

In this case we again have to use the quadratic formula to solve for λ,

$$\lambda = \frac{3 \pm \sqrt{9 - 4(1)(1)}}{2},$$

$$= \frac{3 \pm \sqrt{5}}{2}.$$

$$\lambda = \frac{3 - \sqrt{5}}{2} = .38, \qquad \lambda = \frac{3 + \sqrt{5}}{2} = 2.62.$$

Since all of the eigenvalues of this matrix are positive, it is negative-definite and the critical point $\left(\frac{25}{12}, \frac{1}{12}\right)$ represents a local minimum.

7. (a) First, we find the trace, determinant, and eigenvalues of each of the 6 matrices:

- $\begin{bmatrix} 3 & 0 \\ 6 & 4 \end{bmatrix}$

  The trace is 3+4=7.
  The determinant is $(3 \times 4) - (0 \times 6)$=12.
  The eigenvalues are the values λ that make the determinant of $\begin{bmatrix} 3-\lambda & 0 \\ 6 & 4-\lambda \end{bmatrix}$ equal 0. This characteristic equation is:

$$(3-\lambda)(4-\lambda) - (0 \times 6) = 0,$$

$$(\lambda - 3)(\lambda - 4) = 0,$$

$$\lambda = 3,\ 4.$$

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

  The trace is 5+2=7.
  The determinant is $(5 \times 2) - (-4 \times -7)$=-18.
  The eigenvalues are the values λ that make the determinant of $\begin{bmatrix} 5-\lambda & -4 \\ -7 & 2-\lambda \end{bmatrix}$ equal 0. This characteristic equation is:

$$(5-\lambda)(2-\lambda) - (-4 \times -7) = 0,$$
$$(\lambda-5)(\lambda-2) - 28 = 0,$$
$$(\lambda^2 - 7\lambda + 10) - 28 = 0,$$
$$\lambda^2 - 7\lambda - 18 = 0,$$
$$(\lambda-9)(\lambda+2) = 0,$$
$$\lambda = 9,\ -2.$$

- $\begin{bmatrix} -2 & 2 \\ 9 & 5 \end{bmatrix}$

  The trace is -2+5=3.
  The determinant is $(-2 \times 5) - (2 \times 9)$=-28.
  The eigenvalues are the values λ that make the determinant of $\begin{bmatrix} -2-\lambda & 2 \\ 9 & 5-\lambda \end{bmatrix}$ equal 0. This characteristic equation is:

$$(-2-\lambda)(5-\lambda) - (2 \times 9) = 0,$$
$$(\lambda+2)(\lambda-5) - 18 = 0,$$
$$(\lambda^2 - 3\lambda - 10) - 18 = 0,$$
$$\lambda^2 - 3\lambda - 28 = 0,$$
$$(\lambda-7)(\lambda+4) = 0,$$
$$\lambda = -4,\ 7.$$

- $\begin{bmatrix} 2 & -5 \\ -4 & 2 \end{bmatrix}$

  The trace is 2+2=4.
  The determinant is $(2 \times 2) - (-4 \times -5)$=-16.
  The eigenvalues are the values λ that make the determinant of $\begin{bmatrix} 2-\lambda & -5 \\ -4 & 2-\lambda \end{bmatrix}$ equal 0. This characteristic equation is:

$$(2-\lambda)(2-\lambda) - (-4 \times -5) = 0,$$
$$(\lambda-2)^2 - 20 = 0,$$
$$(\lambda^2 - 4\lambda + 4) - 20 = 0,$$
$$\lambda^2 - 4\lambda - 16 = 0,$$
$$\lambda = \frac{4 \pm \sqrt{16 - 4(1)(-16)}}{2},$$

$$
\begin{aligned}
&= \frac{4 \pm \sqrt{16+64}}{2}, \\
&= \frac{4 \pm \sqrt{80}}{2}, \\
&= \frac{4 \pm 4\sqrt{5}}{2}, \\
\lambda &= (2-2\sqrt{5}),\ (2+2\sqrt{5}).
\end{aligned}
$$

- $\begin{bmatrix} -10 & 9 \\ 6 & 2 \end{bmatrix}$

  The trace is -10+2=-8.
  The determinant is $(-10 \times 2) - (6 \times 9)$=-74.
  The eigenvalues are the values λ that make the determinant of $\begin{bmatrix} -10-\lambda & 9 \\ 6 & 2-\lambda \end{bmatrix}$ equal 0. This characteristic equation is:

$$
\begin{aligned}
(-10-\lambda)(2-\lambda) - (6 \times 9) &= 0, \\
(\lambda+10)(\lambda-2) - 54 &= 0, \\
(\lambda^2 + 8\lambda - 20) - 54 &= 0, \\
\lambda^2 + 8\lambda - 74 &= 0, \\
\lambda &= \frac{-8 \pm \sqrt{64 - 4(1)(-74)}}{2}, \\
&= \frac{-8 \pm \sqrt{64+296}}{2}, \\
&= \frac{-8 \pm 6\sqrt{10}}{2}, \\
\lambda &= (-4-3\sqrt{10}),\ (-4+3\sqrt{10}).
\end{aligned}
$$

- $\begin{bmatrix} 5 & 4 \\ -2 & -6 \end{bmatrix}$

  The trace is 5-6=-1.
  The determinant is $(5 \times -6) - (-2 \times 4)$=-22.
  The eigenvalues are the values λ that make the determinant of $\begin{bmatrix} 5-\lambda & 4 \\ -2 & -6-\lambda \end{bmatrix}$ equal 0. This characteristic equation is:

$$
\begin{aligned}
(5-\lambda)(-6-\lambda) - (-2 \times 4) &= 0, \\
(\lambda-5)(\lambda+6) + 8 &= 0, \\
(\lambda^2 + \lambda - 30) + 8 &= 0, \\
\lambda^2 + \lambda - 22 &= 0, \\
\lambda &= \frac{-1 \pm \sqrt{1 - 4(1)(-22)}}{2},
\end{aligned}
$$

$$= \frac{-1 \pm \sqrt{1+88}}{2},$$

$$= \frac{-1 \pm \sqrt{89}}{2},$$

$$\lambda = \left(-\frac{1}{2} - \frac{\sqrt{89}}{2}\right), \left(-\frac{1}{2} + \frac{\sqrt{89}}{2}\right).$$

(b) The first three matrices show a clear pattern: the trace is the **sum of the eigenvalues** and the determinant is the **product of the eigenvalues**. With a little work we can show that these properties hold for the last three matrices as well, when the eigenvalues are not integers. Consider the fourth matrix. The sum of the eigenvalues is

$$(2 - 2\sqrt{5}) + (2 + 2\sqrt{5}) = 4,$$

which is the trace, and the product of the eigenvalues is[3]

$$(2 - 2\sqrt{5})(2 + 2\sqrt{5}) = 4 - (2\sqrt{5})^2 = 4 - 4(5) = -16,$$

which is the determinant. Next consider the fifth matrix. The sum of the eigenvalues is

$$(-4 - 3\sqrt{10}) + (-4 + 3\sqrt{10}) = -8,$$

which is the trace, and the product of the eigenvalues is

$$(-4 - 3\sqrt{10})(-4 + 3\sqrt{10}) = 16 - 9(10) = -74,$$

which is the determinant. Finally, consider the last matrix. The sum of the eigenvalues is

$$\left(-\frac{1}{2} - \frac{\sqrt{89}}{2}\right) + \left(-\frac{1}{2} + \frac{\sqrt{89}}{2}\right) = -1,$$

which is the trace, and the product of the eigenvalues is

$$\left(-\frac{1}{2} - \frac{\sqrt{89}}{2}\right)\left(-\frac{1}{2} + \frac{\sqrt{89}}{2}\right) = \frac{1}{4} - \frac{89}{4} = -\frac{88}{4} = -22,$$

which is again the determinant.

8. This problem asks us to demonstrate that the product QBQ^{-1}, where

$$Q = \begin{bmatrix} 0.91 & 0.35 & -0.38 & 0.26 \\ 0.28 & -0.44 & 0.00 & 0.64 \\ -0.27 & 0.83 & -0.44 & 0.03 \\ 0.11 & 0.03 & 0.81 & -0.72 \end{bmatrix}, \qquad B = \begin{bmatrix} 10.47 & 0 & 0 & 0 \\ 0 & -9.21 & 0 & 0 \\ 0 & 0 & -7.65 & 0 \\ 0 & 0 & 0 & -3.60 \end{bmatrix},$$

and

$$Q^{-1} = \begin{bmatrix} 0.97 & -0.18 & -0.52 & 0.18 \\ 0.09 & 0.92 & 1.62 & 0.94 \\ -0.45 & 2.00 & 1.19 & 1.68 \\ -0.35 & 2.27 & 1.33 & 0.57 \end{bmatrix},$$

[3]It's easiest to use the difference of squares formula to evaluate this product: $(a+b)(a-b) = a^2 - b^2$.

is equal to

$$A = \begin{bmatrix} 8 & -1 & -8 & 3 \\ 4 & -2 & 2 & 3 \\ -5 & 0 & -7 & -2 \\ 3 & -7 & -5 & -9 \end{bmatrix}.$$

All we have to do is compute the matrix product QBQ^{-1}. First, consider the two left factors QB:

$$QB = \begin{bmatrix} 0.91 & 0.35 & -0.38 & 0.26 \\ 0.28 & -0.44 & 0.00 & 0.64 \\ -0.27 & 0.83 & -0.44 & 0.03 \\ 0.11 & 0.03 & 0.81 & -0.72 \end{bmatrix} \begin{bmatrix} 10.47 & 0 & 0 & 0 \\ 0 & -9.21 & 0 & 0 \\ 0 & 0 & -7.65 & 0 \\ 0 & 0 & 0 & -3.60 \end{bmatrix}.$$

This product multiplies a $(4{\times}4)$ matrix by another one, so multiplication is conformable and the product is also (4×4). The elements are inner-products of the corresponding row of Q and the corresponding column of B:

$$\begin{aligned}
&(1,1) \text{ element}: && (0.91 \times 10.47) + (0.35 \times 0) + (-0.38 \times 0) + (0.26 \times 0) && = 9.53, \\
&(1,2) \text{ element}: && (0.91 \times 0) + (0.35 \times -9.21) + (-0.38 \times 0) + (0.26 \times 0) && = -3.22, \\
&(1,3) \text{ element}: && (0.91 \times 0) + (0.35 \times 0) + (-0.38 \times -7.65) + (0.26 \times 0) && = 2.91, \\
&(1,4) \text{ element}: && (0.91 \times 0) + (0.35 \times 0) + (-0.38 \times 0) + (0.26 \times -3.60) && = -0.94, \\
&(2,1) \text{ element}: && (0.28 \times 10.47) + (-0.44 \times 0) + (0 \times 0) + (0.64 \times 0) && = 2.93, \\
&(2,2) \text{ element}: && (0.28 \times 0) + (-0.44 \times -9.21) + (0 \times 0) + (0.64 \times 0) && = 4.05, \\
&(2,3) \text{ element}: && (0.28 \times 0) + (-0.44 \times 0) + (0 \times -7.65) + (0.64 \times 0) && = 0, \\
&(2,4) \text{ element}: && (0.28 \times 0) + (-0.44 \times 0) + (0 \times 0) + (0.64 \times -3.60) && = -2.30, \\
&(3,1) \text{ element}: && (-0.27 \times 10.47) + (0.83 \times 0) + (-0.44 \times 0) + (0.03 \times 0) && = -2.83, \\
&(3,2) \text{ element}: && (-0.27 \times 0) + (0.83 \times -9.21) + (-0.44 \times 0) + (0.03 \times 0) && = -7.64, \\
&(3,3) \text{ element}: && (-0.27 \times 0) + (0.83 \times 0) + (-0.44 \times -7.65) + (0.03 \times 0) && = 3.37, \\
&(3,4) \text{ element}: && (-0.27 \times 0) + (0.83 \times 0) + (-0.44 \times 0) + (0.03 \times -3.60) && = -0.11, \\
&(4,1) \text{ element}: && (0.11 \times 10.47) + (0.03 \times 0) + (0.81 \times 0) + (-0.72 \times 0) && = 1.15, \\
&(4,2) \text{ element}: && (0.11 \times 0) + (0.03 \times -9.21) + (0.81 \times 0) + (-0.72 \times 0) && = -0.28, \\
&(4,3) \text{ element}: && (0.11 \times 0) + (0.03 \times 0) + (0.81 \times -7.65) + (-0.72 \times 0) && = -6.20, \\
&(4,4) \text{ element}: && (0.11 \times 0) + (0.03 \times 0) + (0.81 \times 0) + (-0.72 \times -3.60) && = 2.59.
\end{aligned}$$

Next we left-multiply this matrix by the inverse of Q:

$$(QB)Q^{-1} = \begin{bmatrix} 9.53 & -3.22 & 2.91 & -0.94 \\ 2.93 & 4.05 & 0 & -2.30 \\ -2.83 & -7.64 & 3.37 & -0.11 \\ 1.15 & -0.28 & -6.20 & 2.59 \end{bmatrix} \begin{bmatrix} 0.97 & -0.18 & -0.52 & 0.18 \\ 0.09 & 0.92 & 1.62 & 0.94 \\ -0.45 & 2.00 & 1.19 & 1.68 \\ -0.35 & 2.27 & 1.33 & 0.57 \end{bmatrix}.$$

Again, this product multiplies a $(4{\times}4)$ matrix by another one, so multiplication is conformable and the product is also $(4{\times}4)$. The elements are inner-products of the corresponding row of QB and the corresponding column of Q^{-1}:

$$\begin{array}{lcl}
(1,1) \text{ element}: & (9.53 \times 0.97) + (-3.22 \times 0.09) + (2.91 \times -0.45) + (-0.94 \times -0.35) & = 8, \\
(1,2) \text{ element}: & (9.53 \times -0.18) + (-3.22 \times 0.92) + (2.91 \times 2.00) + (-0.94 \times 2.27) & = -1, \\
(1,3) \text{ element}: & (9.53 \times -0.52) + (-3.22 \times 1.62) + (2.91 \times 1.19) + (-0.94 \times 1.33) & = -8, \\
(1,4) \text{ element}: & (9.53 \times 0.18) + (-3.22 \times 0.94) + (2.91 \times 1.68) + (-0.94 \times 0.57) & = 3, \\
(2,1) \text{ element}: & (2.93 \times 0.97) + (4.05 \times 0.09) + (0 \times -0.45) + (-2.30 \times -0.35) & = 4, \\
(2,2) \text{ element}: & (2.93 \times -0.18) + (4.05 \times 0.92) + (0 \times 2.00) + (-2.30 \times 2.27) & = -2, \\
(2,3) \text{ element}: & (2.93 \times -0.52) + (4.05 \times 1.62) + (0 \times 1.19) + (-2.30 \times 1.33) & = 2, \\
(2,4) \text{ element}: & (2.93 \times 0.18) + (4.05 \times 0.94) + (0 \times 1.68) + (-2.30 \times 0.57) & = 3, \\
(3,1) \text{ element}: & (-2.83 \times 0.97) + (-7.64 \times 0.09) + (3.37 \times -0.45) + (-0.11 \times -0.35) & = -5, \\
(3,2) \text{ element}: & (-2.83 \times -0.18) + (-7.64 \times 0.92) + (3.37 \times 2.00) + (-0.11 \times 2.27) & = -0, \\
(3,3) \text{ element}: & (-2.83 \times -0.52) + (-7.64 \times 1.62) + (3.37 \times 1.19) + (-0.11 \times 1.33) & = -7, \\
(3,4) \text{ element}: & (-2.83 \times 0.18) + (-7.64 \times 0.94) + (3.37 \times 1.68) + (-0.11 \times 0.57) & = -2, \\
(4,1) \text{ element}: & (1.15 \times 0.97) + (-0.28 \times 0.09) + (-6.20 \times -0.45) + (2.59 \times -0.35) & = 3, \\
(4,2) \text{ element}: & (1.15 \times -0.18) + (-0.28 \times 0.92) + (-6.20 \times 2.00) + (2.59 \times 2.27) & = -7, \\
(4,3) \text{ element}: & (1.15 \times -0.52) + (-0.28 \times 1.62) + (-6.20 \times 1.19) + (2.59 \times 1.33) & = -5, \\
(4,4) \text{ element}: & (1.15 \times 0.18) + (-0.28 \times 0.94) + (-6.20 \times 1.68) + (2.59 \times 0.57) & = -9,
\end{array}$$

so the total product is

$$QBQ^{-1} = \begin{bmatrix} 8 & -1 & -8 & 3 \\ 4 & -2 & 2 & 3 \\ -5 & 0 & -7 & -2 \\ 3 & -7 & -5 & -9 \end{bmatrix},$$

and we've demonstrated that $QBQ^{-1} = A$.

9. (a) The concept we are trying to measure is social capital. We characterize this concept to involve an individual's belief that other people can be trusted, and the individual's inclination to follow societal rules even when there will be no consequence for breaking them. An individual with a large amount of social capital will both trust others and will avoid making decisions that can harm others indirectly. Some of the concepts that we do not want to include in this characterization are: conforming to rules when sanctions for failing to do so are present, the size of an individual's social network, an individual's societal standing, prestige, or class. In this example, we are representing the concept of social capital with the survey question on trust and with the battery of questions regarding civic cooperation. In practice, we would probably want to add additional variables to represent the concept, but this is the representation for this simple example. We will measure the concept with principle components analysis.

   (b) In order to find the eigenvalues of the covariance matrix, we find the determinant of

$$\begin{bmatrix} 6.32 - \lambda & 4.47 \\ 4.47 & 7.51 - \lambda \end{bmatrix},$$

   and solve for λ such that the determinant is 0. The characteristic equation is

$$(6.32 - \lambda)(7.51 - \lambda) - 4.47^2 = 0,$$

$$\begin{aligned}
(\lambda - 6.32)(\lambda - 7.51) - 19.98 &= 0, \\
(\lambda^2 - 13.83\lambda + 47.46) - 19.98 &= 0, \\
\lambda^2 - 13.83\lambda + 27.48 &= 0, \\
\lambda &= \frac{13.83 \pm \sqrt{(-13.83)^2 - 4(1)(27.48)}}{2}, \\
&= \frac{13.83 \pm \sqrt{191.27 - 109.92}}{2}, \\
&= \frac{13.83 \pm \sqrt{81.35}}{2}, \\
&= \frac{13.83 \pm 9.02}{2},
\end{aligned}$$

$$\lambda = \frac{13.83 - 9.02}{2} = 2.41, \qquad \lambda = \frac{13.83 + 9.02}{2} = 11.43.$$

To find the eigenvectors associated with $\lambda = 11.43$, we plug this eigenvalue into the above matrix,

$$\begin{bmatrix} 6.32 - 11.43 & 4.47 \\ 4.47 & 7.51 - 11.43 \end{bmatrix} = \begin{bmatrix} -5.11 & 4.47 \\ 4.47 & -3.92 \end{bmatrix},$$

and reduce as much as possible using elementary row operations. This task is trickier because we have no choice but to deal with the decimals. First we multiply the first row by 4.47/5.11=.875 and add it to the second row:

$$\begin{bmatrix} -5.11 & 4.47 \\ 0 & 0 \end{bmatrix}.$$

This reduction implies the following system of equations,

$$\begin{cases} -5.11x + 4.47y = 0, \\ y = y, \end{cases}$$

solving the top equation for x,

$$\begin{cases} x = .875y, \\ y = y, \end{cases}$$

so in matrix notation the eigenvectors associated with $\lambda = 11.43$ are

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} .875 \\ 1 \end{bmatrix}.$$

A particular eigenvector that arises when $y = 1$ is $\begin{bmatrix} .875 \\ 1 \end{bmatrix}$. To find the unit eigenvector, we compute the length of this vector and divide the vector by it's length. The length is

$$\left\| \begin{bmatrix} .875 \\ 1 \end{bmatrix} \right\| = \sqrt{.875^2 + 1^2} = 1.33.$$

Therefore the unit eigenvector associated with $\lambda = 11.43$ is

$$\frac{1}{1.33} \begin{bmatrix} .875 \\ 1 \end{bmatrix} = \begin{bmatrix} .658 \\ .752 \end{bmatrix}.$$

To find the eigenvectors associated with $\lambda = 2.41$, we plug this eigenvalue into the above matrix,

$$\begin{bmatrix} 6.32 - 2.41 & 4.47 \\ 4.47 & 7.51 - 2.41 \end{bmatrix} = \begin{bmatrix} 3.91 & 4.47 \\ 4.47 & 5.1 \end{bmatrix},$$

and reduce as much as possible using elementary row operations. First we multiply the first row by -4.47/3.91=-1.14 and add it to the second row:

$$\begin{bmatrix} 3.91 & 4.47 \\ 0 & 0 \end{bmatrix}.$$

This reduction implies the following system of equations,

$$\begin{cases} 3.91x + 4.47y = 0, \\ y = y, \end{cases}$$

solving the top equation for x,

$$\begin{cases} x = -1.14y, \\ y = y, \end{cases}$$

so in matrix notation the eigenvectors associated with $\lambda = 2.41$ are

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -1.14 \\ 1 \end{bmatrix}.$$

A particular eigenvector that arises when $y = 1$ is $\begin{bmatrix} -1.14 \\ 1 \end{bmatrix}$. To find the unit eigenvector, we compute the length of this vector and divide the vector by it's length. The length is

$$\left|\left| \begin{bmatrix} -1.14 \\ 1 \end{bmatrix} \right|\right| = \sqrt{(-1.14)^2 + 1^2} = 1.52.$$

Therefore the unit eigenvector associated with $\lambda = 2.41$ is

$$\frac{1}{1.52} \begin{bmatrix} -1.14 \\ 1 \end{bmatrix} = \begin{bmatrix} -.75 \\ .658 \end{bmatrix}.$$

(c) The variance explained by the larger eigenvalue is

$$\frac{11.43}{11.43 + 2.41} = 82.6 \text{ percent}$$

(d) To create an index for social capital, we consider the largest eigenvalue and its unit eigenvector, and use the elements of this unit eigenvector to weight the observed variables that generated the covariance matrix. In part (b) we found that the largest eigenvalue is $\lambda = 11.43$ and that its unit eigenvector is

$$\begin{bmatrix} .658 \\ .752 \end{bmatrix}.$$

Therefore the equation for the social capital index is

$$\text{Social capital} = .658 \times \text{Trust} + .752 \times \text{Civic Cooperation}.$$

Including this index in the data gives us the following table:

Obs.	Trust	Civic Cooperation	Social Capital
1	4	2	.658(4)+.752(2)=4.1
2	7	10	.658(7)+.752(10)=12.1
3	3	6	.658(3)+.752(6)=6.5
4	8	9	.658(8)+.752(9)=12.0
5	9	7	.658(9)+.752(7)=11.2
6	2	3	.658(2)+.752(3)=3.6
7	9	6	.658(9)+.752(6)=10.4
8	5	3	.658(5)+.752(3)=5.5
9	6	4	.658(6)+.752(4)=7.0
10	8	8	.658(8)+.752(8)=11.3

10. This is a hard problem. Usually, researchers quickly turn to computers to process the tedious calculus involved with a problem like this. But it is worthwhile to struggle through one of these problems by hand, to get a strong sense of what the computer does and why. Be patient and careful as you work through the steps. Generally speaking, because we are rounding here to the third decimal place, our results by hand will be slightly different from the results we would get by using a computer, so don't be surprised to see small discrepancies if you compare these results to computer output.

   (a) The cross-tabulation can be written as

$$M = \begin{bmatrix} 452 & 174 \\ 82 & 292 \end{bmatrix}.$$

   To calculate P, we find the sum of all elements in the cross-tab, $452 + 174 + 82 + 292 = 1000$, and divide every element of M by this sum,

$$P = \begin{bmatrix} .452 & .174 \\ .082 & .292 \end{bmatrix}.$$

   (b) R is a vector that contains the row sums of P,

$$R = \begin{bmatrix} 0.626 \\ 0.374 \end{bmatrix},$$

   and C is a vector that contains the column sums of P,

$$C = \begin{bmatrix} 0.534 \\ 0.466 \end{bmatrix}.$$

   The matrix D_r is square with as many rows as R, has zeroes for the off-diagonal elements, and contains the square roots of the elements of R on its diagonal:

$$D_r = \begin{bmatrix} \sqrt{0.626} & 0 \\ 0 & \sqrt{0.374} \end{bmatrix} = \begin{bmatrix} 0.719 & 0 \\ 0 & 0.611 \end{bmatrix}.$$

   Likewise, the matrix D_c is square with as many rows as C, has zeroes for the off-diagonal elements, and contains the square roots of the elements of C on its diagonal:

$$D_c = \begin{bmatrix} \sqrt{0.534} & 0 \\ 0 & \sqrt{0.466} \end{bmatrix} = \begin{bmatrix} 0.731 & 0 \\ 0 & 0.682 \end{bmatrix}.$$

(c) To calculate the matrix $S = D_r(P - RC')D_c$ all we have to do is plug in these matrices and perform the calculations. Let's work step by step starting with RC':

$$RC' = \begin{bmatrix} 0.626 \\ 0.374 \end{bmatrix} \begin{bmatrix} 0.534 & 0.466 \end{bmatrix} = \begin{bmatrix} 0.334 & 0.292 \\ 0.200 & 0.174 \end{bmatrix}.$$

Next let's calculate $P - RC'$,

$$P - RC' = \begin{bmatrix} .452 & .174 \\ .082 & .292 \end{bmatrix} - \begin{bmatrix} 0.334 & 0.292 \\ 0.200 & 0.174 \end{bmatrix} = \begin{bmatrix} 0.118 & -0.118 \\ -0.118 & 0.118 \end{bmatrix}.$$

Next we calculate $D_r(P - Rc')$,

$$D_r(P - RC') = \begin{bmatrix} 0.719 & 0 \\ 0 & 0.611 \end{bmatrix} \begin{bmatrix} 0.118 & -0.118 \\ -0.118 & 0.118 \end{bmatrix} = \begin{bmatrix} 0.093 & -0.093 \\ -0.072 & 0.072 \end{bmatrix}.$$

Finally, we right-multiply this product by D_c,

$$S = D_r(P - RC')D_c = \begin{bmatrix} 0.093 & -0.093 \\ -0.072 & 0.072 \end{bmatrix} \begin{bmatrix} 0.731 & 0 \\ 0 & 0.682 \end{bmatrix} = \begin{bmatrix} 0.068 & -0.064 \\ -0.053 & 0.049 \end{bmatrix}.$$

(d) To find SS', we simply multiply

$$SS' = \begin{bmatrix} 0.068 & -0.064 \\ -0.053 & 0.049 \end{bmatrix} \begin{bmatrix} 0.068 & -0.053 \\ -0.064 & 0.049 \end{bmatrix} = \begin{bmatrix} 0.009 & -0.007 \\ -0.007 & 0.005 \end{bmatrix}.$$

Likewise, to find $S'S$, we multiply

$$S'S = \begin{bmatrix} 0.068 & -0.053 \\ -0.064 & 0.049 \end{bmatrix} \begin{bmatrix} 0.068 & -0.064 \\ -0.053 & 0.049 \end{bmatrix} = \begin{bmatrix} 0.007 & -0.007 \\ -0.007 & 0.007 \end{bmatrix}.$$

(e) First consider SS'. To find the eigenvalues of this matrix, we have to solve the characteristic equation

$$\left| \begin{bmatrix} 0.009 - \lambda & -0.007 \\ -0.007 & 0.005 - \lambda \end{bmatrix} \right| = 0,$$

$$(.009 - \lambda)(.005 - \lambda) - .007^2 = 0,$$

$$(\lambda - .009)(\lambda - .005) - .000049 = 0,$$

$$\lambda^2 - .014\lambda + .000045 - .000049 = 0,$$

$$\lambda^2 - .014\lambda - .000004 = 0.$$

The third term is close enough to 0 for us to round it to 0 (and anyway, if we wanted to perform arithmetic at the sixth decimal place we'd be using computers to do so), and so the eigenvalues of SS' are

$$\lambda^2 - .014\lambda = 0$$

$$\lambda(\lambda - .014) = 0,$$

$$\lambda = 0, \ \lambda = .014.$$

(f) In this step, we only have to rearrange the results we calculated in the previous step. U is the matrix of unit eigenvectors of SS', so

$$U = \begin{bmatrix} -0.791 & -0.612 \\ 0.612 & -0.791 \end{bmatrix}.$$

Σ is the diagonal matrix that contains the square roots of the eigenvalues on the diagonal:

$$\Sigma = \begin{bmatrix} \sqrt{0.014} & 0 \\ 0 & \sqrt{0} \end{bmatrix} = \begin{bmatrix} 0.118 & 0 \\ 0 & 0 \end{bmatrix}.$$

Finally, V is the matrix of unit eigenvectors of $S'S$, so

$$V = \begin{bmatrix} -0.731 & -0.683 \\ 0.683 & -0.731 \end{bmatrix}.$$

Recall that

$$S = \begin{bmatrix} 0.068 & -0.064 \\ -0.053 & 0.049 \end{bmatrix}.$$

If U, Σ, and V comprise the singular value decomposition of S, then

$$S = U\Sigma V'.$$

We have to compute $U\Sigma V'$ and show that it is equal to S. First, let's compute $U\Sigma$:

$$U\Sigma = \begin{bmatrix} -0.791 & -0.612 \\ 0.612 & -0.791 \end{bmatrix} \begin{bmatrix} 0.118 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -0.093 & 0 \\ 0.072 & 0 \end{bmatrix}.$$

Next we compute $U\Sigma V'$:

$$U\Sigma V' = \begin{bmatrix} -0.093 & 0 \\ 0.072 & 0 \end{bmatrix} \begin{bmatrix} -0.731 & 0.683 \\ -0.683 & -0.731 \end{bmatrix} = \begin{bmatrix} 0.068 & -0.064 \\ -0.053 & 0.049 \end{bmatrix}.$$

So we have confirmed this singular value decomposition.

(g) To find the coordinates for the categories that comprise the rows of the cross-tabulation, we calculate $D_r U\Sigma$. In the previous step we calculated $U\Sigma$. To complete the product we multiply

$$D_r U\Sigma = \begin{bmatrix} 0.719 & 0 \\ 0 & 0.611 \end{bmatrix} \begin{bmatrix} -0.093 & 0 \\ 0.072 & 0 \end{bmatrix} = \begin{bmatrix} -0.074 & 0 \\ 0.044 & 0 \end{bmatrix}.$$

To find the coordinates for the categories that comprise the columns, we calculate $D_c V\Sigma$:

$$D_c V\Sigma = \begin{bmatrix} 0.731 & 0 \\ 0 & 0.682 \end{bmatrix} \begin{bmatrix} -0.731 & -0.683 \\ 0.683 & -0.731 \end{bmatrix} \begin{bmatrix} 0.118 & 0 \\ 0 & 0 \end{bmatrix}$$

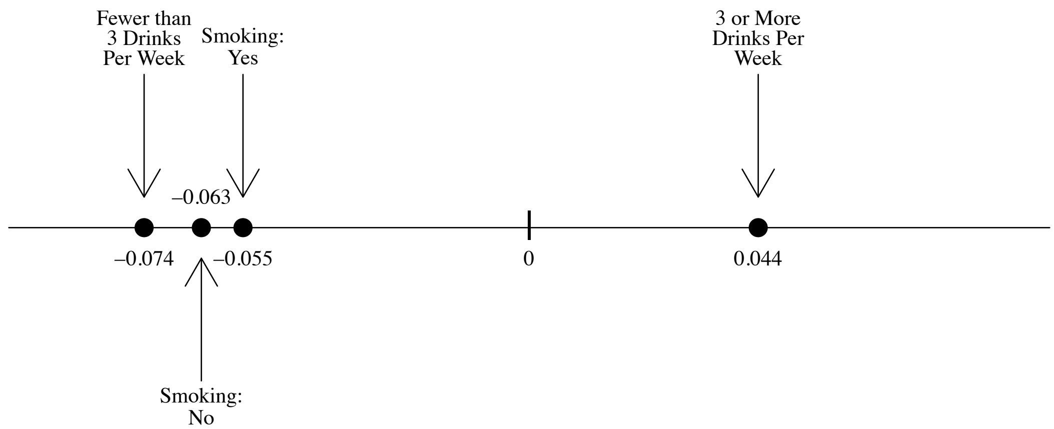
$$= \begin{bmatrix} -0.534 & 0.499 \\ -0.466 & -0.499 \end{bmatrix} \begin{bmatrix} 0.118 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -0.063 & 0 \\ -0.055 & 0 \end{bmatrix}.$$

The coordinates for the categories that appear in the cross-tabulation are as follows:

Category	First latent variable	Second latent variable
Fewer than 3 drinks per week	-0.074	0
3 or more drinks per week	0.044	0
Regular smoker: No	-0.063	0
Regular smoker: Yes	-0.055	0

All of the coordinates for the second latent variable turn out to be zero – we do not have enough data to make a statement about a second dimension that underlies the data.

(h) The plot of these points (on a number line since there's no second dimension) is below:



The first latent variable indicates the similarity between categories. The standout is the "3 or more drinks per week" category.