

# SUPPLEMENT FOR MATH 138

## TABLE OF CONTENTS

A.1 Population Growth Models	2
A.1.1 The exponential growth models	2
A.1.2 Defects	4
A.2 The Logistic Equation	5
A.2.1 Justification and derivation	5
A.2.2 Explicit solution	5
A.2.3 Calculations with the logistic equation	8
B.1 Systems of Linear Equations and Matrices	13
B.2 Algebra of Matrices	22
B.3 Determinant of a Matrix	28
B.4 Properties of the Determinant	35
B.5 Invertible Matrices	39
B.6 Cramer's Rule	42
B.7 Eigenvalues	45
B.8 Eigenvectors	48
B.9 Final Remarks	50

## A.1 Population Growth Models

### A.1.1 The exponential growth model

As preparation for the logistic equation, we shall quickly review the exponential growth model for the number  $Q = Q(t)$  of individuals (people, mice, bacteria) in a population as a function of time  $t$ . The key idea in this model is that the rate of change of  $Q$  is proportional to  $Q$ . This leads to the differential equation

$$dQ/dt = kQ$$

for some relative growth rate  $k$ . ( $k$  is the fraction that the population increases by per unit time.) The solution to this differential equation is

$$Q(t) = Q_0 e^{kt}$$

where  $Q_0 = Q(0)$  is the population at time  $t = 0$ .

Usually  $k$  is of the form  $k = k_b - k_d$ , where  $k_b$  is the birth rate and  $k_d$  is the death rate. If time is measured in years (or days), then  $k_b$  is the average number of offspring produced by one individual per year (or per day). (In animals, this should really be the average number of female offspring produced by one female per year, but if the population is 50% female we would get the same value for  $k_b$ .) One common approximation of the death rate is to set  $k_d = 1/A$ , where  $A$  is the average lifespan.

**Example 1** In a human population with an average lifespan of 70 years and 1.77 children born per year per 100 people, the birthrate and death rate are approximately

$$k_b = \frac{1.77}{100} = 0.0177, \quad k_d = \frac{1}{70} = 0.0143$$

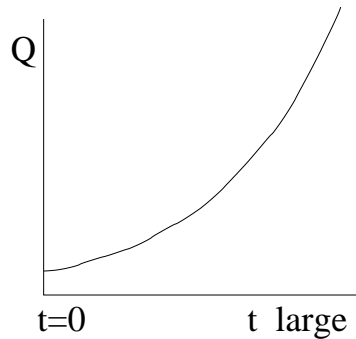
so that  $k = 0.0034 \text{years}^{-1}$ , and the exponential model predicts

$$Q(t) = Q_0 e^{(0.0034)t}, \quad \text{with } t \text{ measured in years.}$$

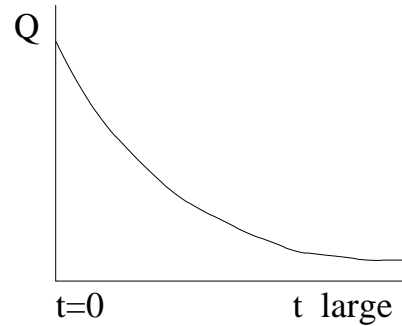
In the limit, for large values of  $t$ , we have:

for $k > 0$	$Q(\infty) = \infty$	population explosion
for $k = 0$	$Q(\infty) = Q_0$	constant population
for $k < 0$	$Q(\infty) = 0$	population dies out.

Here are the graphs of the solutions:



Population explosion ( $k > 0$ )



Species extinction ( $k < 0$ )

The most interesting case is the explosive growth case, when  $k > 0$ . It is common to describe this situation in terms of the time  $T$  that is needed for the population to double, i.e., for  $Q(T) = 2Q(0)$ . From the explicit solution, this is when

$$e^{kT} = 2, \quad \text{or} \quad kT = \ln 2, \quad \text{or:} \quad T = \frac{\ln 2}{k}.$$

Note that the population doubles over any time interval of length  $T$  (say between time  $t$  and  $t + T$ ), because  $Q(t + T) = Q_0 e^{kt} e^{kT} = 2Q(t)$ .

### EXERCISES

1. The U.S. population was 4 million in 1790 and 248.7 million in 1990. Using the exponential model, estimate the population in the years 1620, 1865, 1973, 1998, 2000. Compare these with the actual population of 0.003, 35, 179.3 and 268 million people.
2. The population in Monrovia grows by 10% per year, and was 1 million in 1980. What is the relative growth rate (in new citizens per year per individual)?

### A.1.2 Defects

The most obvious defect of the exponential growth model is the assumption made that the growth rate depends *only* on the size of the population. The following factors are totally neglected:

1. Limits to growth. No population can increase exponentially forever. One limiting factor is the availability of food supply. Another is the population density. In models of parasite population, the size of the host population is a limit to the growth of the parasite. The logistic equation is designed to model these defects.
2. Age and Sex. Not all individuals can reproduce. Among humans, for example, only females in a certain age range can give birth. Moreover, the death rate increases with age. Neither the exponential nor the logistic equation takes this into consideration. You will encounter a population model addressing these points in Keller's supplement *Population Projection*. This model requires the use of matrices and linear algebra.
3. Predators. A population increase encourages the increase of natural enemies, such as predators and infectious diseases. Thus the death rate  $r_d$  is not constant. Most models that consider these factors use linear algebra in a way similar to the way we will handle the age/sex problem.
4. Immigration/Emigration. In modeling the U.S. population, immigration is a very important factor, especially from 1620 to 1720. Currently, about 320,000 people immigrate to the U.S. each year.
5. Technology. Over long periods of time, the birth rate and death rate of human population have changed, due to an increase in the standard of living, the development of birth control, the control of disease (especially in the 19th century), and the rise of medical death-postponing technology.

## A.2 The Logistic Equation

### A.2.1 Justification and derivation

In some populations, such as micro-parasite populations (smallpox, polio, herpes), the birth rate is proportional to both the infected host population  $Q$  and the susceptible host population  $L - Q$  (where  $L$  is the total infectable host population, i.e., the upper limit of  $Q$ ). That is, the birth rate is proportional to the product  $(L - Q)Q$ .

In other populations, the death rate has a component proportional to the *probability of an encounter* between two members of the population, i.e., to  $Q^2$ . A very important example of this is the epidemic transmission of deadly diseases (bubonic plague in 1348, cholera epidemics in the 19th century, AIDS in today's third world), or diseases that weaken and shorten life (the epidemic spread of syphilis in the 16th century, or hookworm in the American south in the 1870's).

Thus a reasonable (yet simple) modification of the exponential growth model is to subtract a quadratic term:

$$dQ/dt = aQ - kQ^2.$$

This is the *logistic equation*. It is useful to set  $L = a/k$  and rewrite this equation as

$$dQ/dt = kQ(L - Q)$$

This form makes it clear that one particular solution of the logistic equation is a constant population  $Q = L$ . Our textbook uses the letter  $B$  instead of  $L$  and calls it the *carrying capacity* but the more traditional term is the *limit to growth*.

The constant  $L$  is an *equilibrium population* in the following sense. If the population is less than  $L$  ( $0 < Q < L$ ) then the rate of change  $dQ/dt$  is positive, so that the population increases. If the population  $Q$  is more than  $L$ , then the rate of change  $dQ/dt$  is negative, so that the population decreases. In fact, the explicit formula for  $Q$  found in the next section will show that the population will always tend toward  $L$  as the time  $t$  tends to  $\infty$ . In symbols:

$$\lim_{t \rightarrow \infty} Q(t) = L.$$

### A.2.2 Explicit solution

In order to solve the logistic equation, we put it in the form

$$k dt = \frac{dQ}{Q(L-Q)}.$$

To integrate this, we multiply by  $L$  and use the “partial fractions” trick that:

$$\frac{L}{Q(L-Q)} = \frac{1}{Q} + \frac{1}{L-Q}.$$

Substituting the right hand side into the equation, we get:

$$L \int k dt = \int \frac{L dQ}{Q(L-Q)} = \int \left[ \frac{1}{Q} + \frac{1}{L-Q} \right] dQ,$$

which (since  $L$  and  $k$  are constant) integrates to:

$$(Lk)t = \ln |Q| - \ln |L-Q| + C.$$

We want to solve this for  $Q$  as a function of  $t$ . Recall that  $a = Lk$ . Taking exponentials of both sides, and setting  $A = e^C$ , we get

$$e^{at} = A \left| \frac{Q}{L-Q} \right| \quad \text{for some constant } A > 0.$$

Suppose first that  $0 < Q < L$ , which is the case in most biological models. In this case, we can ignore the absolute value signs, and calculate:

$$AQ = e^{at}(L-Q) = e^{at}L - e^{at}Q;$$

$$e^{at}Q + AQ = e^{at}L;$$

$$Q = \left[ \frac{e^{at}L}{e^{at} + A} \right] \cdot \left[ \frac{e^{-at}}{e^{-at}} \right];$$

$$\boxed{Q = \frac{L}{1 + Ae^{-at}}, \quad 0 < Q < L \text{ and } A > 0.}$$

Similarly, if we keep track of the absolute value signs, we get

$$Q = \frac{L}{1 - Ae^{-at}}, \quad Q > L \text{ and } A > 0.$$

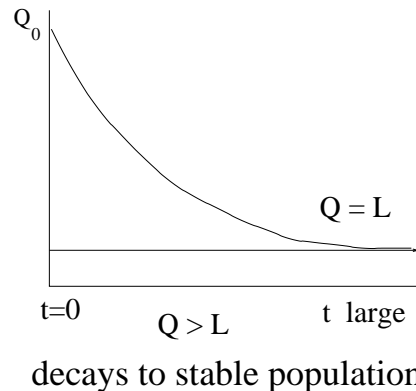
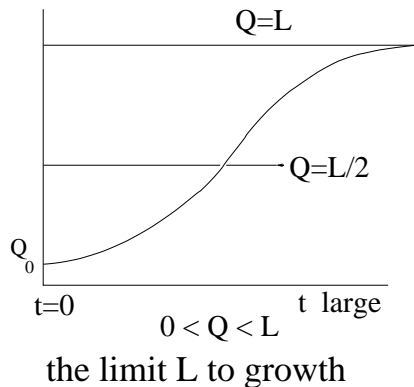
To find the value of the constant  $A$ , set  $t = 0$ ,  $Q = Q(0) = Q_0$ . We get

$$Q_0 = \frac{L}{1 \pm A}.$$

Solving for  $A$  yields

$$A = \left| (L/Q_0) - 1 \right|.$$

Notice that as  $t$  goes to  $\infty$ ,  $e^{-at}$  goes to 0, and the population  $Q$  goes to  $L$  in the limit. This confirms the qualitative assertion we made earlier that  $L$  is the equilibrium population. Here are the graphs of the solutions:



#### EXERCISE

1. Show that the inflection point in the left graph occurs when  $Ae^{-at} = 1$  and  $Q = L/2$ . *Hint:* start with  $dQ/dt = aQ - kQ^2$  and differentiate, using  $L = a/k$ .

The most interesting case of the logistic equation is when  $0 < Q < L$ . When  $Q$  is much less than  $L = a/k$ ,  $kQ$  is much less than  $a$ , and  $kQ^2$  is much smaller than  $aQ$ . The logistic equation tells us then that we have approximately

$$dQ/dt \cong aQ,$$

i.e., exponential growth. When  $Q$  reaches  $L/2$ , though, the growth rate  $dQ/dt$  switches from increasing to decreasing because of the inflection point. The growth rate then slows more and more as the population  $Q$  approaches the limiting population  $Q \approx L$ , finally reaching a stable, equilibrium population.

You may encounter the logistic equation in an alternative form, where we choose  $t = t_0$  to be the time when  $Q = L/2$ . This form is obtained by setting  $A = e^{at_0}$ :

$$Q = \frac{L}{1 + e^{a(t_0-t)}} .$$

### A.2.3 Calculations with the logistic equation

Suppose that we have a number of observations of a population, e.g. bacteria growing in a medium. How can we test whether the growth fits a logistic equation? Because the equation contains transcendental functions, this may appear difficult. However, it is quite easy to calculate with populations at *equally spaced* time intervals,  $t = nT$ ,  $n = 0, 1, 2, \dots$  (Each interval has length  $T$ .)

To do this, we turn the solution upside down, and get

$$\frac{1}{Q} = \frac{1}{L} + \frac{A}{L}e^{-at} .$$

Setting  $Q_0 = Q(0)$ ,  $Q_1 = Q(T)$  and  $Q_2 = Q(2T)$  yields the equations

$$\frac{1}{Q_0} = \frac{1}{L} + \frac{A}{L}, \quad \text{and} \quad \frac{1}{Q_1} = \frac{1}{L} + \frac{A}{L}e^{-aT}, \quad \text{and} \quad \frac{1}{Q_2} = \frac{1}{L} + \frac{A}{L}e^{-2T} \quad (1)$$

$$\frac{1}{Q_0} - \frac{1}{Q_1} = \left[ \frac{A}{L} \right] \cdot [1 - e^{-aT}] \quad (2)$$

$$\frac{1}{Q_1} - \frac{1}{Q_2} = \left[ \frac{A}{L} \right] \cdot [e^{-aT} - e^{-2aT}] = e^{-aT} \cdot \left[ \frac{A}{L} \right] \cdot [1 - e^{-aT}] . \quad (3)$$

The ratio of (3) to (2) is

$$\boxed{\frac{\left[ \frac{1}{Q_1} - \frac{1}{Q_2} \right]}{\left[ \frac{1}{Q_0} - \frac{1}{Q_1} \right]} = e^{-aT}}$$



and this can be used to find the constant  $a$ . (Take the logarithm of both sides.) Plugging this value of  $e^{-aT}$  back into equation (2) lets us find  $A/L$  :

$$\boxed{\left[ \frac{A}{L} \right] = \frac{\left[ \frac{1}{Q_0} - \frac{1}{Q_1} \right]}{\left[ 1 - e^{-aT} \right]}} . \quad (4)$$

But from (1) we get  $L$ , since its reciprocal  $L^{-1}$  is

$$\frac{1}{L} = \frac{1}{Q_0} - \frac{A}{L} .$$

We now know (but don't need)  $k = a/L$ . Finally,  $A = L \cdot (A/L)$ . We have now found the entire logistic equation.

To test whether the growth fits the logistic equation, measure  $Q_3 = Q(3T)$  and compare it to the formula we just found.

**Example 2** Three successive observations of a bacterial population gave results:

<u>time <math>t</math></u>	<u>population <math>Q(t)</math></u>
0	$5.00 \times 10^5$
1 day	$1.60 \times 10^6$
2 days	$2.50 \times 10^6$

Assume that  $Q(t)$  obeys the logistic equation, and use this model to predict the equilibrium population  $L$ , as well as the bacterial population  $Q(3)$  after 3 days.

To solve this problem, it is convenient to measure  $Q$  in millions to get rid of the  $10^6$  factor. We choose  $T = 1$  day for the time intervals and form the following tableau:

<u><math>t</math></u>	<u><math>Q</math></u>	<u><math>1/Q</math></u>	<u>difference</u>	<u>ratio = <math>e^{-a}</math></u>
0	0.500	2.0000		
			1.3750	
1	1.600	0.6250		0.163636    so $a = -\ln(.163636) = 1.181001$
			0.2250	
2	2.500	0.4000		

Equation (4) gives  $A/L = 1.6440$ . From this we get  $L = 2.8092$  and  $A = 4.6756$ . Thus our population model uses the logistic equation:

$$Q = \frac{2.8092}{1 - 4.6756 \cdot e^{-1.81011t}} \quad \text{million bacteria after } t \text{ days.}$$

## EXERCISES

1. The world population (in billions) was

3.049 in 1960

3.721 in 1970

4.473 in 1980 (source: U.S. Bureau of census).

Assume that the world population obeys the logistic equation.

- (a) Predict the population in 1990 (estimated to be 5.25 billion).
- (b) Predict the equilibrium world population (Answer: 12.77 billion).
- (c) To test whether the growth fits the logistic equation, predict the population in 1950, which was 2.490 billion. Why is it not a good idea to test the equation with the 1800 population of 0.910 billion?

2. The U.S. population (in millions) was

0.905 in 1740 (British Colonial census)

3.929 in 1790 (first U.S. census)

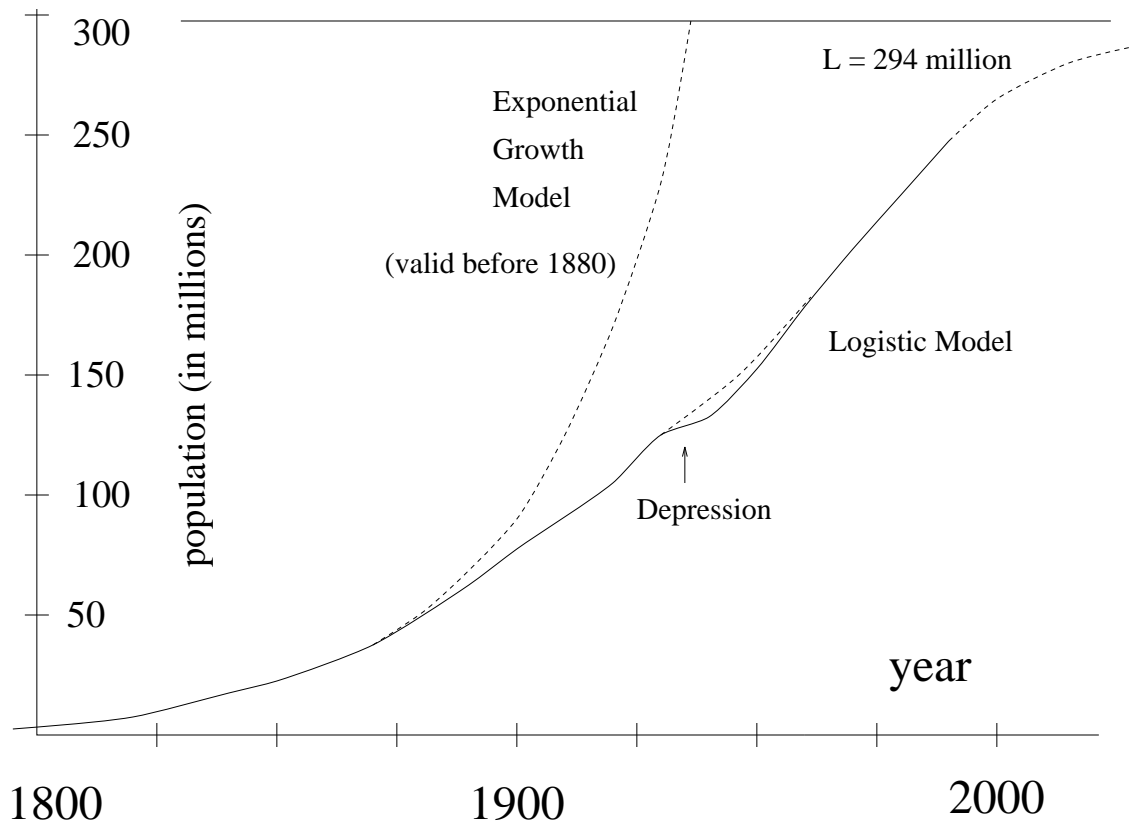
17.069 in 1840

62.979 in 1890

132.164 in 1940

- (a) Fit the U.S. populations of 1790, 1840 and 1890 to a logistic equation, and find the equilibrium population  $L$  of that model. (Answer: 250.7 million.)
- (b) Test your answer against the populations in 1740 and 1940. The birth rate during the Depression (1930's) was very low; can you give reasons for this?

- (c) How does the (legal) immigration rate of 800,000 people per year affect the differential equation  $\frac{dQ}{dt} = aQ - kQ^2$ ? (Until about 1990 this rate was about 300,000 people per year.)
- (d) The Census Bureau made two population projections recently: 1) that the U.S. population will peak at 302 million sometime around the year 2050, and 2) that it will continue to grow, reaching 393 million by the year 2050. Which do you believe, and why?
- (See <http://www.census.gov> for more information.)



U.S. Population 1790--2000

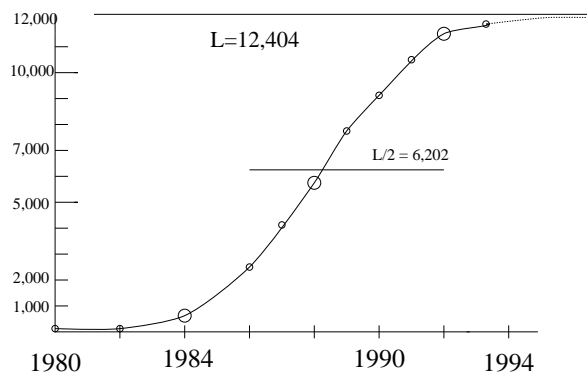
3. The following statistics were released by the State Health Department in 1994.  
(The right column updates these numbers; reporting methods changed in 1993.)

N.J. AIDS Cases — Cumulative Known Dead (By Year of Diagnosis)

1980	9	1985	1,337	1989	7,600	1995	20,000
1981	33	1986	2,465	1990	9,267	2000	25,500
1982	100	1987	4,078	1991	10,696	2005	32,000
1984	622	1988	5,722	1992	11,571	2010	38,000

- (a) Fit the data for 1984, 1988, 1992 to a logistic equation, and find the equilibrium number  $L$  of deaths by AIDS in New Jersey, predicted by this model.

(Answer:  $e^{-a} = .498$ ,  $L = 12,404$ .)



Cumulative no. Deaths by AIDS in N.J. (1994 data)

- (b) Using the logistic model predict the cumulative number of deaths by AIDS by the end of 1993. (Answer: 11,974, or 203 deaths during 1993.) There were 173 reported deaths from Jan. 1 – Oct. 1 in 1993.

- (c) Give three reasons why the logistic model is inappropriate for this statistic. Then explain why the logistic model fits the data as well as it does.

Hint: People die many years after they contract AIDS. In 1993 there were 7,000 reported cases of people with AIDS in New Jersey. Anti-retroviral therapy became available in 1995. In 2010 there were 35,000 people living with HIV or AIDS in New Jersey.

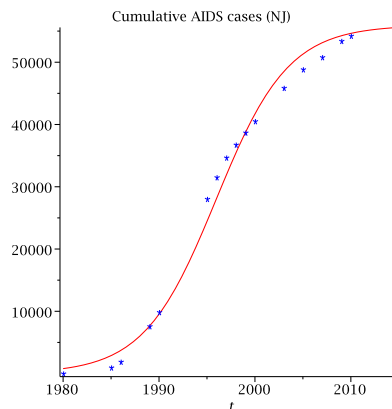
4. Better reporting methods were adopted by New Jersey in 1994, and anti-retroviral therapy became available in 1995. After that date, the NJ Health Department focussed upon cumulative cases with AIDS, and also HIV cases which had not yet developed into AIDS.

The following statistics were released by the State Health Department in 2010.

N.J. AIDS Cases — Cumulative AIDS cases in New Jersey

1985	972	1990	9,867	1995	28,047	1996	31,523
1997	34,687	1998	36,761	1999	38,700	2000	40,531
2003	45,881	2005	48,858	2007	50,810	2010	53,420

- (a) Fit the data for 1990, 2000, 2010 to a logistic equation, and find the equilibrium number  $L$  of deaths by AIDS in New Jersey, predicted by this model. (Answer:  $e^{-10a} = 0.088$ ,  $L = 55,120$ .)
- (b) Using the logistic model predict the cumulative number of AIDS cases by the end of 2005. (Answer: 49,791)
- (c) The logistics model predicts only 25,000 AIDS cases by 1995. Explain why the cumulative number of AIDS cases diagnosed is more than the logistics model for 1992–1995.
- (d) Explain why it is unreasonable to expect that there will be at most  $L$  cases of AIDS by the end of 2020. What is missing from the model?



## B.1 Systems of Linear Equations and Matrices

A solution to a system of simultaneous equations in  $n$  unknowns is an ordered  $n$ -tuple of numbers which satisfies all the equations of the system. For example, consider the following system of linear equations in the unknowns  $x$ ,  $y$  and  $z$ :

$$\begin{aligned}x - 5y + 2z &= -5, \\3x - 14y + 3z &= -8, \\4x - 18y + 3z &= -8.\end{aligned}\tag{1}$$

To solve this system, we need to find a 3-tuple  $(x, y, z)$  of numbers which satisfies all three equations. Here is one way to go about it. If we subtract 4 times the first equation from the third equation and leave the first two equations alone, we obtain

$$\begin{aligned}x - 5y + 2z &= -5, \\3x - 14y + 3z &= -8, \\2y - 5z &= 12.\end{aligned}\tag{2}$$

By subtracting 3 times the first equation from the second equation in (2), we get

$$\begin{aligned}x - 5y + 2z &= -5, \\y - 3z &= 7, \\2y - 5z &= 12.\end{aligned}\tag{3}$$

We now subtract 2 times the second equation from the third equation in (3) to obtain

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

If we add 3 times the third equation to the second, and subtract 2 times the third equation from the first, we have

$$\begin{aligned}x - 5y &= -1, \\y &= 1, \\z &= -2.\end{aligned}\tag{5}$$

Finally, by adding 5 times the second equation to the first, we get

$$\begin{aligned}x &= 4, \\y &= 1, \\z &= -2.\end{aligned}\tag{6}$$

Thus the only possible solution of (1), the original system of equations, is  $(4, 1, -2)$ , or  $x = 4, y = 1, z = -2$ .

We can verify by substitution that  $(4, 1, -2)$  is a solution of the original system (1). It is also instructive to verify that  $(4, 1, -2)$  is also a solution of the systems (2), (3), (4), and (5). These systems, (1) – (5), are *equivalent*:

**Definition 1** *Two systems are said to be equivalent if every solution of each system is also a solution of the other.*

Here are three operations we can do to a system of linear equations which will result in an equivalent system:

- i. changing the order of the equations;
- ii. multiplying all members of an equation by a nonzero constant;
- iii. adding a multiple of one equation to another equation.

The letters used to represent the unknowns in a system of linear equations are not important. The solution of the system (1) is exactly the same as the solution of the system

$$\begin{aligned}u - 5v + 2w &= -5, \\3u - 14v + 3w &= -8, \\4u - 18v + 3w &= -8,\end{aligned}$$

namely  $(4, 1, -2)$ . Since only the coefficients and the constant terms affect the solution, we might omit the unknowns and the equality signs and simply write the numbers:

$$\begin{bmatrix} 1 & -5 & 2 & -5 \\ 3 & -14 & 3 & -8 \\ 4 & -18 & 3 & -8 \end{bmatrix}.$$

Such a rectangular array of numbers is called a *matrix*. Its rows run horizontally and its columns run vertically. This matrix has three rows and four columns. Notice that the first column contains the coefficients of the first unknown variable, and so on for the next two columns, and the last column contains the constant terms. Since this matrix has 3 rows and 4 columns, it's *dimension* is  $3 \times 4$  (three by four). In the dimension, the number of rows comes before the number of columns.

Let's repeat our process for solving the system (1) and see what happens to the matrix at each step. We start with

$$\begin{array}{rclcl} x & - & 5y & + & 2z & = & -5, & & \\ 3x & - & 14y & + & 3z & = & -8, & & \\ 4x & - & 18y & + & 3z & = & -8. & & \end{array} \quad \begin{bmatrix} 1 & -5 & 2 & \left| & -5 \\ 3 & -14 & 3 & \left| & -8 \\ 4 & -18 & 3 & \left| & -8 \end{bmatrix} \quad (1')$$

The vertical bar is not part of the matrix; it was inserted to remind us that the numbers to the left of the bar are the coefficients of the unknowns, and the column to the right of it shows the right hand side. This is called the *augmented matrix* of the system.

As before, we first subtract 4 times the first equation from the third equation (and 4 times the first row of the matrix from the second row of the matrix).

$$\begin{array}{rclcl} x & - & 5y & + & 2z & = & -5, & & \\ 3x & - & 14y & + & 3z & = & -8, & & \\ 0x & + & 2y & - & 5z & = & 12. & & \end{array} \quad \begin{bmatrix} 1 & -5 & 2 & \left| & -5 \\ 3 & -14 & 3 & \left| & -8 \\ 0 & 2 & 5 & \left| & 12 \end{bmatrix} \quad (2')$$

We next subtract 4 times the first equation (row) from the second equation (row).

$$\begin{array}{rclcl} x & - & 5y & + & 2z & = & -5, & & \\ 0x & + & y & - & 3z & = & 7, & & \\ 0x & + & 2y & - & 5z & = & 12. & & \end{array} \quad \begin{bmatrix} 1 & -5 & 2 & \left| & -5 \\ 0 & 1 & -3 & \left| & 7 \\ 0 & 2 & 5 & \left| & 12 \end{bmatrix} \quad (3')$$

Now we subtract 2 times the second equation (row) from the third.



$$\begin{array}{rcl}
x & - & 5y + 2z = -5, \\
0x & + & y - 3z = 7, \\
0x & + & 0y + z = -2.
\end{array}
\quad \left[ \begin{array}{ccc|c}
1 & -5 & 2 & -5 \\
0 & 1 & -3 & 7 \\
0 & 0 & 1 & -2
\end{array} \right] \quad (4')$$

We add 3 times the third equation (row) to the second, and subtract 2 times the third equation (row) from the first.

$$\begin{array}{rcl}
x & - & 5y + 0z = -1, \\
0x & + & y - 0z = 1, \\
0x & + & 0y + z = -2.
\end{array}
\quad \left[ \begin{array}{ccc|c}
1 & -5 & 0 & -1 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & -2
\end{array} \right] \quad (5')$$

Finally, we add 5 times the second equation (row) to the first to get:

$$\begin{array}{rcl}
x & - & 0y + 0z = 4, \\
0x & + & y - 0z = 1, \\
0x & + & 0y + z = -2.
\end{array}
\quad \left[ \begin{array}{ccc|c}
1 & 0 & 0 & 4 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & -2
\end{array} \right]. \quad (6')$$

The solution  $(4, 1, -2)$  of the original system of equations  $(1')$  can now be read from the fourth column of the matrix in  $(6')$ .

The operations on the augmented matrix of the system of equations in steps  $(2')$ – $(6')$  are called *elementary row operations*. They are

- i. changing the order of the rows;
- ii. multiplying a row by a nonzero constant;
- iii. adding a multiple of one row to another row.

If a matrix can be obtained from another matrix by elementary row operations, the matrices are said to be *row equivalent*. Since the corresponding operations on a system of linear equations always result in an equivalent system, row equivalent matrices must belong to equivalent systems of linear equations.

The first three columns in the matrix of the system of equations  $(1)$ ,

$$\left[ \begin{array}{ccc}
1 & -5 & 2 \\
3 & -14 & 3 \\
4 & -18 & 3
\end{array} \right]$$

form the *coefficient matrix*. Since this matrix has three rows and three columns, it is a square matrix. The diagonal of elements from the upper left-hand corner to the lower right-hand corner of a square matrix is called its *principal diagonal*.

We found the solution of the given system of equations (1) by reducing its augmented matrix above to the matrix

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

by means of elementary row operations. Notice that the final coefficient matrix has 1's in its principal diagonal and zeros elsewhere. This is the method which is applied to any system of  $n$  linear equations in  $n$  unknowns.

**Example 3** By elementary row operations on its augmented matrix, solve the system

$$5x - 4y = -22,$$

$$3x + 7y = 15.$$

*Solution.* The augmented matrix of this system of linear equations is

$$\left[ \begin{array}{cc|c} 5 & -4 & -22 \\ 3 & 7 & 15 \end{array} \right].$$

Our objective is to reduce its coefficient matrix to

$$\left[ \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right]$$

by elementary row operations on the augmented matrix.

At each stage, let  $R_1$  denote the first row and  $R_2$  denote the second row. We begin with the augmented matrix

$$\left[ \begin{array}{cc|c} 5 & -4 & -22 \\ 3 & 7 & 15 \end{array} \right].$$

Combining two operations, we multiply the second row by 5 and add  $(-3R_1)$  to  $(5R_2)$ .

This replaces the second row with  $(-3R_1) + 5R_2$  yielding

$$\left[ \begin{array}{cc|c} 5 & -4 & -22 \\ 0 & 47 & 141 \end{array} \right].$$

Now we multiply the first row by  $\frac{1}{5}$  and the second row by  $\frac{1}{47}$  to get

$$\left[ \begin{array}{cc|c} 1 & \frac{-4}{5} & \frac{-22}{5} \\ 0 & 1 & 3 \end{array} \right].$$

Finally, we replace the first row with  $R_1 + \frac{4}{3}R_2$  with the result

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

The solution of the given system is then  $(-2, 3)$ . You should verify that our solution  $x = -2$ ,  $y = 3$  does indeed satisfy both the given equations of the system.

Both the examples of linear systems of equations we have looked at so far have one and only one solution. But some systems of  $n$  linear equations in  $n$  unknowns have infinitely many solutions and others have no solutions.

To understand how each case may occur, we consider systems of three equations and three unknowns. In three-dimensional Cartesian space the solutions of each equation form a plane. If the three equations correspond to three planes which intersect in just one point, the system of equations has a unique solution, the coordinates of the single common point. If the three planes have a line in common, their system of equations has an infinity of solutions since the coordinates of every point on their common line satisfy all three equations. If the line of intersection of two planes is parallel to the third plane, their system of equations has no solution.

The matrix method of solution we have been discussing may be applied to all these cases and will reveal which case occurs.

**Example 4** Solve the system

$$\begin{aligned} x + 2y - z &= 6, \\ -x + 4y - z &= 8, \\ 2x + y - z &= 5. \end{aligned}$$

by elementary row operations on its matrix.

*Solution.* The augmented matrix of the given system is

$$\left[ \begin{array}{ccc|c} 1 & 2 & -1 & 6 \\ -1 & 4 & -1 & 8 \\ 2 & 1 & -1 & 5 \end{array} \right].$$

Again at each stage, we refer to the first row by  $R_1$ , the second row by  $R_2$  and the third row by  $R_3$ .

First, we replace  $R_3$  by  $(-2R_1) + R_3$

$$\left[ \begin{array}{ccc|c} 1 & 2 & -1 & 6 \\ -1 & 4 & -1 & 8 \\ 0 & -3 & 1 & -7 \end{array} \right].$$

Now we replace  $R_2$  by  $R_1 + R_2$

$$\left[ \begin{array}{ccc|c} 1 & 2 & -1 & 6 \\ 0 & 6 & -2 & 14 \\ 0 & -3 & 1 & -7 \end{array} \right].$$

Next, we replace  $R_2$  by  $R_2 + 2R_3$ , multiply  $R_3$  by  $-1$  and switch  $R_2$  and  $R_3$

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

So the original system of three equations is equivalent to the system of two equations:

$$\begin{aligned} x + 2y - z &= 6, \\ 3y - z &= 7. \end{aligned}$$

This system has infinitely many solutions. If  $y$  is assigned an arbitrary value  $t$ , we use the equation  $3y - z = 7$  to get that  $z = 3t - 7$ , then we use the equation  $x + 2y - z = 6$  and our values for  $y$  and  $z$  in terms of  $t$  to get  $x = t - 1$ . Therefore, every solution of the given system has the form  $(t - 1, t, 3t - 7)$ , where  $t$  is a real number. If we plot the solutions of the given system, we get the line in 3-space whose parametric equations are

$$x = t - 1, \quad y = t, \quad z = 3t - 7, \quad -\infty < t < +\infty.$$

**Example 5** Investigate the solutions of the system

$$\begin{aligned}x + 2y - z &= 6, \\-x + 4y - z &= 8, \\2x + y - z &= 4\end{aligned}$$

by elementary row operations on its matrix.

*Solution.* Notice that this system differs from the system of Example 2 only in the right hand side of the third equation. The matrix of this system is

$$\left[ \begin{array}{ccc|c} 1 & 2 & -1 & 6 \\ -1 & 4 & -1 & 8 \\ 2 & 1 & -1 & 4 \end{array} \right].$$

If we perform exactly the same elementary row operations on this matrix as we did on the augmented matrix of Example 2 we will find that it reduces to the row equivalent matrix

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

The third row of this matrix is equivalent to the equation  $0 = -2$ , which is false for all  $x, y, z$ . Therefore the given system has no solution.

## EXERCISES

1. Describe all solutions of a linear system whose corresponding augmented matrix can be row reduced to the matrix below. Also give a specific solution with  $x_3 = 2$ .

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

2. In (a) and (b) below, find all solutions of the given linear system by using elementary row operations on its augmented matrix to reduce it to the form

$$\left[ \begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & b \end{array} \right].$$

$$(a) \quad \begin{array}{rcl} 2x & - & y = 8 \\ 6x & - & 5y = 32 \end{array}$$

$$(b) \quad \begin{array}{rcl} 7x_1 & - & 2x_2 = 7 \\ 3x_1 & + & 13x_2 = 3 \end{array}$$

3. In (a) and (b) below, find all solutions of the given linear system by using elementary row operations on its augmented matrix to reduce it to the form

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & a \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & c \end{array} \right].$$

$$(a) \quad \begin{array}{rcl} x_1 + 3x_2 - 2x_3 & = & -7 \\ 2x_1 - x_2 + x_3 & = & -9 \\ 4x_1 - 2x_2 - 3x_3 & = & -23 \end{array}$$

$$(b) \quad \begin{array}{rcl} 7x - y - z & = & 7 \\ 4x + 3y - 5z & = & -4 \\ -2x + 6y - 11z & = & -19 \end{array}$$

4. Use elementary row operations on the augmented matrix of the system of linear equations below to show that the system has no solutions.

$$\begin{array}{rcl} x + 4y - 2z & = & 4 \\ 2x + 7y - z & = & -2 \\ 2x + 9y - 7z & = & 1 \end{array}$$

5. Find all solutions of the given linear system below by using elementary row operations on its augmented matrix.

$$\begin{array}{rcl} -2x_1 - 3x_2 & = & 0 \\ 2x_1 + 3x_2 & = & 0 \end{array}$$

## B.2 Algebra of Matrices

Now that we have seen what matrices are good for, we need to know how to manipulate them. Let's begin with some vocabulary.

**Definition 2** Let  $m$  and  $n$  be positive integers. An  $m \times n$  ( $m$  by  $n$ ) matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

is a rectangular array of numbers  $a_{ij}$ , arranged in  $m$  rows and  $n$  columns (so that the first subscript  $i$  goes from 1 to  $m$ , and the second subscript  $j$  goes from 1 to  $n$ ). We say that  $A$  has dimension  $m \times n$ .

The numbers  $a_{ij}$  are the *entries* of the matrix; the first subscript indicates the row in which the entry appears and the second subscript indicates the column. So for example,  $a_{25}$  is the entry in the second row and fifth column.

It is often convenient to indicate an  $m \times n$  matrix briefly by  $[a_{ij}]_{mn}$ .

Two  $m \times n$  matrices,  $A = [a_{ij}]_{mn}$  and  $B = [b_{ij}]_{mn}$ , are said to be *equal* if and only if

$$a_{ij} = b_{ij}$$

for  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ . Thus two matrices are equal if and only if they have the same dimension, and all their corresponding entries are equal.

**Definition 3** If  $A = [a_{ij}]_{mn}$  and  $B = [b_{ij}]_{mn}$  are  $m \times n$  matrices, their sum is the  $m \times n$  matrix

$$A + B = [a_{ij} + b_{ij}]_{mn}.$$

In other words, if two matrices have the same dimension, they may be added by adding corresponding entries.

For example, if

$$A = \begin{bmatrix} 2 & -7 \\ -3 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} -5 & 0 \\ 1 & 6 \end{bmatrix}$$

then

$$A + B = \begin{bmatrix} 2 + (-5) & -7 + 0 \\ -3 + 1 & 4 + 6 \end{bmatrix} = \begin{bmatrix} -3 & -7 \\ -2 & 10 \end{bmatrix}.$$

Addition of matrices, like equality of matrices, is defined only for matrices having the same dimension. Note that addition of matrices is commutative and associative. That is, if  $A$ ,  $B$ , and  $C$  are matrices having the same dimension, then

$$A + B = B + A \quad \text{and} \quad A + (B + C) = (A + B) + C.$$

**Definition 4** *The product of a number  $k$  and an  $m \times n$  matrix  $A = [a_{ij}]_{mn}$  is the  $m \times n$  matrix*

$$kA = [ka_{ij}]_{mn}.$$

As an example,

$$6 \begin{bmatrix} -1 & 0 & 4 \\ 5 & 2 & -7 \end{bmatrix} = \begin{bmatrix} 6(-1) & 6(0) & 6(4) \\ 6(5) & 6(2) & 6(-7) \end{bmatrix} = \begin{bmatrix} -6 & 0 & 24 \\ 30 & 12 & -42 \end{bmatrix}.$$

If all the entries of a matrix are zeros, it is called a *zero matrix* and is written as  $\mathbf{0}$ .

A matrix of 1 row and  $n$  columns, that is, a  $1 \times n$  matrix

$$[a_1 \ a_2 \ \cdots \ a_n],$$

is called an  $n$ -dimensional *row vector*. A matrix of  $n$  rows and 1 column, that is, an  $n \times 1$  matrix

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$



is called an  $n$ -dimensional *column vector*. The *dot product* of an  $n$ -dimensional row vector and an  $n$ -dimensional column vector is the number

$$[a_1 a_2 \cdots a_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Using this, we can define the product of two matrices whose sizes match up.

**Definition 5** *If  $A$  is an  $m \times n$  matrix*

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

*and if  $B$  is an  $n \times r$  matrix*

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1r} \\ b_{21} & b_{22} & b_{23} & \cdots & b_{2r} \\ b_{31} & b_{32} & b_{33} & \cdots & b_{3r} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ b_{n1} & b_{n2} & b_{n3} & \cdots & b_{nr} \end{bmatrix}$$

*the product  $AB$  is the  $m \times r$  matrix  $P = [p_{ij}]_{mr}$  where  $p_{ij}$  is the dot product of the  $i$ th row vector of matrix  $A$  and the  $j$ th column vector of matrix  $B$ . That is to say,*

$$p_{ij} = [a_{i1} a_{i2} \cdots a_{in}] \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj},$$

$$p_{ij} = \sum_{t=1}^n a_{it} b_{jt}.$$

For the definition of the product  $AB$  to make sense, it is necessary that the number of columns in  $A$  be the same as the number of rows in  $B$ . That is, the second number  $n$  in the dimension  $m \times n$  of  $A$  must be the same as the first number  $n$  in the dimension  $n \times r$  of  $B$ . The numbers  $m$  and  $r$  can be any positive integers.

**Example 6** Compute  $AB$  if

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}.$$

*Solution.* The dimension of  $A$  is  $2 \times 3$  and the dimension of  $B$  is  $3 \times 1$ , so  $AB$  is a  $2 \times 1$  matrix. Using the definition, we find that

$$AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \end{bmatrix}$$

**Example 7** Compute  $AB$  where

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

*Solution.* Here  $A$  is a  $3 \times 2$  matrix and  $B$  is a  $2 \times 4$  matrix.

$$\begin{aligned} AB &= \begin{bmatrix} 2 & 3 \\ -1 & 4 \\ 5 & -2 \end{bmatrix} \begin{bmatrix} 3 & 1 & 4 & -5 \\ -2 & 0 & 3 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 2(3) + 3(-2) & 2(1) + 3(0) & 2(4) + 3(3) & 2(-5) + 3(4) \\ -1(3) + 4(-2) & -1(1) + 4(0) & -1(4) + 4(3) & -1(-5) + 4(4) \\ 5(3) - 2(-2) & 5(1) - 2(0) & 5(4) - 2(3) & 5(-5) - 2(4) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 2 & 17 & 2 \\ -11 & -1 & 8 & 21 \\ 19 & 5 & 14 & -33 \end{bmatrix}. \end{aligned}$$

Here are some basic facts about matrix multiplication:

- i. Associative law: If  $A$ ,  $B$ , and  $C$  are matrices of dimension  $m \times n$ ,  $n \times p$  and  $p \times q$ , respectively, then

$$(AB)C = A(BC).$$

- ii. Distributive law: If  $A$  is an  $m \times n$  matrix, if  $B$  and  $C$  are  $n \times p$  matrices, and if  $D$  is a  $p \times q$  matrix, then

$$A(B + C) = AB + AC$$

and

$$(B + C)D = BD + CD.$$

However, there is no commutative law. Let  $A$  be an  $m \times n$  matrix and  $B$  an  $n \times p$  matrix. Then the product  $AB$  is defined, but the product  $BA$  has no meaning unless  $m = p$ . For *both*  $AB$  AND  $BA$  to exist,  $A$  must have dimension  $m \times n$  and  $B$  must have dimension  $n \times m$ . In such an event,  $AB$  will be an  $m \times m$  matrix and  $BA$  will be an  $n \times n$  matrix. So in order for there to even be a chance that  $AB$  and  $BA$  are equal,  $A$  and  $B$  must be square matrices (a square matrix has the same number of rows as columns) of the same dimension. Even then the commutative law does not hold for matrix multiplication as this next example shows.

**Example 8** Let

$$A = \begin{bmatrix} 4 & -3 \\ 2 & 5 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 6 \\ -7 & -2 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 4(1) + (-3)(-7) & 4(6) + (-3)(-2) \\ 2(1) + 5(-7) & 2(6) + 5(-2) \end{bmatrix} = \begin{bmatrix} 25 & 30 \\ -33 & 2 \end{bmatrix},$$

and

$$BA = \begin{bmatrix} 1(4) + 6(2) & 1(-3) + 6(5) \\ -7(4) + (-2)(2) & (-7)(-3) + (-2)(5) \end{bmatrix} = \begin{bmatrix} 16 & 27 \\ -32 & 11 \end{bmatrix}.$$

Thus  $AB \neq BA$ .

We can express a system of  $m$  linear equations in  $n$  unknowns as one matrix equation. As an example, consider the system of three linear equations in four unknowns,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 &= b_3 \end{aligned}$$

Let

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

The matrix  $A$  is called the *coefficient matrix*, the column vector  $\mathbf{x}$  encodes the unknowns, and the column vector  $\mathbf{b}$  encodes the constants on the right hand side of the equations. It is easy to verify by matrix multiplication that this system can be written

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

or

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

## EXERCISES

In exercises 1-10, let

$$A = \begin{bmatrix} -2 & 1 & 3 \\ 4 & 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 4 & 1 & -2 \\ 5 & -1 & 3 \end{bmatrix}, C = \begin{bmatrix} 2 & -1 \\ 0 & 6 \\ -3 & 2 \end{bmatrix}, D = \begin{bmatrix} -4 & 2 \\ 3 & 5 \\ -1 & -3 \end{bmatrix}.$$

Compute the indicated quantity when it exists. If it does not exist, tell why.

1.  $3A$
2.  $0B$
3.  $A + B$
4.  $B + C$
5.  $2C - D$
6.  $AB$
7.  $(2A)(5C)$
8.  $(2A - B)D$
9.  $BC$  and  $CB$
10.  $ADB$

11. Let  $A = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix}$ . Find  $A^2$  and  $A^5$ .

In exercises 12 and 13, write the given system of linear equations in the form  $A\mathbf{x} = \mathbf{b}$  where  $A$  is the coefficient matrix of the system,  $\mathbf{x}$  is the column vector of unknowns and  $\mathbf{b}$  is the column vector of constant terms.

$$\begin{array}{l} 2x + 9y = 12 \\ 12. \quad 3x - 5y = -19 \\ \quad -x + 4y = 11 \end{array} \qquad \begin{array}{l} 13. \quad 3x_1 + 5x_2 - 2x_3 = -7 \\ \quad -x_1 + 4x_2 + 8x_3 = 24 \end{array}$$

### B.3 Determinant of a matrix

Attached to each square matrix  $A$  is a number  $\det A$  called the *determinant* of  $A$ . If

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

then the determinant of  $A$  is symbolized by the same square array, enclosed between vertical bars:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

If  $A$  is an  $n \times n$  square matrix, (ie, if  $A$  is of dimension  $n \times n$ ) then we say  $A$  has order  $n$ . We need a few preliminary definitions before we can define the determinant.

**Definition 6** A submatrix of a matrix  $A$  is any matrix obtained from  $A$  by deleting some of its rows or columns or both.

**Definition 7** If  $A$  is an order  $n$  matrix and  $a_{ij}$  is the entry in the  $i$ th row and  $j$ th column of matrix  $A$ , then we define the minor of  $a_{ij}$  to be the determinant of the submatrix of  $A$  gotten by deleting the  $i$ th row and the  $j$ th column of  $A$ . We write  $M_{ij}$  for the minor of  $a_{ij}$ . The cofactor of  $a_{ij}$ , written  $A_{ij}$ , is defined by

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

**Example 9** Find  $M_{32}$  and  $A_{32}$  if

$$A = \begin{bmatrix} 1 & -5 & 2 \\ 3 & 4 & -8 \\ 4 & -7 & 3 \end{bmatrix}.$$

*Solution.* To find  $M_{32}$ , we need to find the submatrix of  $A$  gotten by deleting the third row and the second column. We get the submatrix

$$\begin{bmatrix} 1 & 2 \\ 3 & -8 \end{bmatrix}.$$

Then the minor  $M_{32}$  is the determinant

$$M_{32} = \begin{vmatrix} 1 & 2 \\ 3 & -8 \end{vmatrix} = -14,$$

and the cofactor  $A_{32}$  is the signed minor

$$A_{32} = (-1)^{3+2}M_{32} = - \begin{vmatrix} 1 & 2 \\ 3 & -8 \end{vmatrix} = +14.$$

We define the determinant of a square matrix by induction on the order  $n$ .

**Definition 8** Let  $A$  be a square matrix of order  $n$ .

If  $n = 1$ ,  $A = [a_{11}]$  and  $\det A$  is defined to be the number  $a_{11}$ .

If  $n = 2$ ,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ and } \det A \text{ is defined by } \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

If  $n > 2$ ,  $\det A$  is the number defined by

$$\det A = a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}$$

where  $A_{1j}$  is the cofactor of the element  $a_{1j}$ .

**Example 10** Find the determinant of the matrix

$$A = \begin{bmatrix} 1 & 5 & 6 \\ -3 & 4 & 8 \\ 2 & -7 & 3 \end{bmatrix}.$$

*Solution.* According to the definition above,

$$\det A = \begin{vmatrix} 1 & 5 & 6 \\ -3 & 4 & 8 \\ 2 & -7 & 3 \end{vmatrix} = (1)A_{11} + (5)A_{12} + (6)A_{13}.$$

We compute

$$A_{11} = (-1)^{1+1} \begin{vmatrix} 4 & 8 \\ -7 & 3 \end{vmatrix} = (-1)^2[(4)(3) - (8)(-7)] = 68,$$

$$A_{12} = (-1)^{1+2} \begin{vmatrix} -3 & 8 \\ 2 & 3 \end{vmatrix} = (-1)^3[(-3)(3) - (8)(2)] = 25,$$

$$A_{13} = (-1)^{1+3} \begin{vmatrix} -3 & 4 \\ 2 & -7 \end{vmatrix} = (-1)^4[(-3)(-7) - (4)(2)] = 13.$$

Thus,

$$\det A = (1)(68) + (5)(25) + (6)(13) = 271.$$

Recall that for an  $n \times n$  matrix  $A$ , the cofactor  $A_{ij}$  is defined to be  $(-1)^{i+j}M_{ij}$ , where  $M_{ij}$  is the determinant of the submatrix gotten by eliminating the  $i$ th row and the  $j$ th column from the original matrix  $A$ . We can easily determine the sign  $(-1)^{i+j}$  attached to the minor  $M_{ij}$  in each individual example by noticing the “checkerboard” pattern the signs make in the entries of the matrix.

To illustrate this, suppose we want to find the cofactor  $A_{23}$  in the  $4 \times 4$  matrix

$$A = \begin{bmatrix} a_{11} & \cdots & a_{14} \\ \vdots & \ddots & \vdots \\ a_{41} & \cdots & a_{44} \end{bmatrix}.$$

Then putting the sign  $(-1)^{i+j}$  in the  $(i, j)$  position yields the checkerboard pattern

$$\begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix}$$

To get the sign for  $A_{23}$  we look in the  $(2, 3)$  position (the second row and third column) and find a negative sign so

$$A_{23} = -M_{23}.$$



Then also we have

$$\det A = a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13} - a_{14}M_{14}.$$

**Example 11** Find the value of the determinant of the following matrix:

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

*Solution.* We have

$$\begin{aligned} \det A &= (5)A_{11} + (-2)A_{12} + (0)A_{13} + (-1)A_{14}, \\ &= (5)M_{11} - (-2)M_{12} + (0)M_{13} - (-1)M_{14}. \end{aligned}$$

So we need to compute  $M_{11}$ ,  $M_{12}$  and  $M_{14}$ .

$$\begin{aligned} M_{11} &= \begin{vmatrix} 1 & 5 & 2 \\ 2 & 0 & 1 \\ 1 & -1 & 1 \end{vmatrix} \\ &= (1)(-1)^{1+1} \begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} + (5)(-1)^{1+2} \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} + (2)(-1)^{1+3} \begin{vmatrix} 2 & 0 \\ 1 & -1 \end{vmatrix} \\ &= (1)[(0)(1) - (1)(-1)] - 5[(2)(1) - (1)(1)] + 2[(2)(-1) - (0)(1)] = -8. \end{aligned}$$

Similarly,  $M_{12} = -22$  and  $M_{14} = 36$ . So,

$$\det A = (5)(-8) - (-2)(-22) + 0 - (-1)(36) = -48.$$

The computation of  $\det A$  using our definition is called *expansion by minors on the first row*. It turns out that  $\det A$  can be obtained by using an expansion by minors on any row or on any column:

**Theorem 1** Let  $A$  be the  $n \times n$  matrix  $[a_{ij}]_{nn}$ , and let  $r$  and  $c$  be any selections from the list of numbers  $1, 2, \dots, n$ . Then

i.  $\det A = a_{r1}A_{r1} + a_{r2}A_{r2} + \cdots + a_{rn}A_{rn}$ , and

ii.  $\det A = a_{1c}A_{1c} + a_{2c}A_{2c} + \cdots + a_{nc}A_{nc}$ .

**Example 12** Find the determinant of the matrix

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

*Solution.* We could simply use the definition of the determinant to compute  $\det B$  but, as we have seen in the previous example, this requires a lot of work. However, using the above theorem, we can also find  $\det B$  by expanding by minors on the third row. This will greatly eliminate the work necessary to find the determinant:

$$\det B = (0)M_{31} - (3)M_{32} + (0)M_{33} - (0)M_{34} = (-3)M_{32}.$$

By computing  $\det B$  in this way we need only find the value of one cofactor,  $M_{32}$ .

$$\begin{aligned} M_{32} &= \begin{vmatrix} 1 & 2 & -5 \\ 3 & 1 & 0 \\ 2 & 7 & -3 \end{vmatrix} \\ &= (1) \begin{vmatrix} 1 & 0 \\ 7 & -3 \end{vmatrix} - (2) \begin{vmatrix} 3 & 0 \\ 2 & -3 \end{vmatrix} + (-5) \begin{vmatrix} 3 & 1 \\ 2 & 7 \end{vmatrix} = -80 \end{aligned}$$

So we get that  $\det B = (-3)(-80) = 240$ .

It is not much harder to expand by minors on the fourth column. In this case we get

$$\det B = -(-5)M_{14} + (-3)M_{44} = (5)(57) - (3)(15) = 240.$$

## EXERCISES

1. Find the determinant of the given matrices.

$$A = \begin{bmatrix} -8 & 3 \\ 7 & -2 \end{bmatrix}$$

$$B = \begin{bmatrix} 2 & 5 \\ -4 & 9 \end{bmatrix}$$

$$C = \begin{bmatrix} -13 & 0 \\ 36 & 0 \end{bmatrix}$$

2. Find the determinant of the matrix  $\begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \\ 3 & -1 & 5 \end{bmatrix}$ .

3. Find the determinant of the following matrix by expanding by minors on the third row, and also by expanding by minors on the second column

$$\begin{bmatrix} 1 & 0 & 7 \\ 17 & 1 & -23 \\ 0 & 0 & 1 \end{bmatrix}.$$

4. Find the determinant of the following matrix by expanding by minors on the fourth column

$$\begin{bmatrix} -2 & 1 & 3 & -5 \\ 8 & 0 & -5 & 0 \\ 0 & -7 & 0 & 13 \\ 5 & 0 & 12 & 0 \end{bmatrix}.$$

5. Find  $\det \begin{bmatrix} -13 & 2 & -5 \\ 0 & 0 & 0 \\ 12 & 4 & 3 \end{bmatrix}$  by expanding by minors on the second row.

6. Show that the determinant of the  $4 \times 4$  matrix

$$M = \begin{bmatrix} d_1 & 1 & 2 & 3 \\ 0 & d_2 & 4 & 5 \\ 0 & 0 & d_3 & 6 \\ 0 & 0 & 0 & d_4 \end{bmatrix}$$

is the product of the entries along the principal diagonal. A square matrix is *upper triangular* if all entries below the principal diagonal are zero. What is the determinant of any upper triangular matrix?

7. Find the determinant of the following matrix

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

where  $\lambda$  is any number.

## B.4 Properties of the Determinant

**Definition 9** The transpose of an  $m \times n$  matrix  $A$  is the  $n \times m$  matrix  $A^T$  formed by interchanging the rows and columns of  $A$ . The  $i$ th row of  $A$  is the  $i$ th column of  $A^T$ .

For example, if

$$A = \begin{bmatrix} 2 & -5 & 1 & 7 \\ 3 & 4 & -3 & -8 \end{bmatrix},$$

then

$$A^T = \begin{bmatrix} 2 & 3 \\ -5 & 4 \\ 1 & -3 \\ 7 & -8 \end{bmatrix}.$$

For the rest of this section we will always consider  $A$  to be a square  $n \times n$  matrix unless otherwise stated.

**Property 1**  $\det A = \det A^T$ .

**Property 2** If two consecutive rows (or two consecutive columns) of matrix  $A$  are interchanged, the determinant of the resulting matrix is  $-\det A$ .

*Proof of Property 2:* We will prove this for interchanging two consecutive rows. The column case follows from Property 1. Let  $A$  be an  $n \times n$  matrix and let  $B$  be the matrix  $A$  with row  $i$  and row  $i + 1$  interchanged. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{(i+1)1} & a_{(i+1)2} & \cdots & a_{(i+1)n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Then

$$B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{(i+1)1} & a_{(i+1)2} & \cdots & a_{(i+1)n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

We write the determinant of matrix  $A$  using the expansion by minors along the  $i$ th row:

$$\det A = a_{i1}(-1)^{i+1}A_{i1} + a_{i2}(-1)^{i+2} + \cdots + a_{in}(-1)^{i+n}A_{in}.$$

Notice that the cofactor  $B_{(i+1)j}$  of position  $(i + 1, j)$  of  $B$  is the same as the cofactor  $A_{ij}$  of position  $(i, j)$  of  $A$ . We write the determinant of matrix  $B$  using the expansion by minors along the  $(i + 1)$ st row:

$$\begin{aligned} \det B &= a_{i1}(-1)^{(i+1)+1}B_{(i+1)1} + a_{i2}(-1)^{(i+1)+2}B_{(i+1)2} + \cdots + a_{in}(-1)^{(i+1)+n}B_{(i+1)n} \\ &= a_{i1}(-1)^{(i+1)+1}A_{i1} + a_{i2}(-1)^{(i+1)+2} + \cdots + a_{in}(-1)^{(i+1)+n}A_{in} \\ &= (-1)a_{i1}(-1)^{i+1}A_{i1} + a_{i2}(-1)^{i+2} + \cdots + a_{in}(-1)^{i+n}A_{in} \\ &= -\det A. \end{aligned}$$

**Property 3** If all the entries in a row (or column) of matrix  $A$  are zero, the value of the determinant is zero.

**Property 4** If two rows (or columns) are proportional, the value of the determinant is zero.

**Property 5** If a single row (or column) of matrix  $A$  is multiplied by a constant  $k$ , the determinant of the resulting matrix is  $k \cdot \det A$ .

*Proof of Property 5:* Let  $B$  be the matrix obtained from  $A$  by replacing the  $i$ th row  $[a_{i1} \ a_{i2} \ \cdots \ a_{in}]$  of  $A$  by  $[ka_{i1} \ ka_{i2} \ \cdots \ ka_{in}]$ . Since the rows of  $B$  are equal to the rows of  $A$  except possibly for the  $i$ th row, the cofactors  $A_{ij}$  and  $B_{ij}$  are the same for each  $j$ . If we compute  $\det B$  by expanding by minors on the  $i$ th row, we have

$$\begin{aligned} \det B &= ka_{i1}A_{i1} + ka_{i2}A_{i2} + \cdots + ka_{in}A_{in} \\ &= k(a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}) \\ &= k \cdot \det A. \end{aligned}$$

The proof for columns is similar.

**Property 6** If we add a constant times one row of a matrix  $A$  to a different row of  $A$  the determinant of the resulting matrix is the same as  $\det A$ .

*Proof of Property 6:* Let  $R_i = [a_{i1} \ a_{i2} \ \cdots \ a_{in}]$  be the  $i$ th row of  $A$ . Suppose that  $k \cdot R_i$  is added to the  $p$ th row  $R_p = [a_{p1} \ a_{p2} \ \cdots \ a_{pn}]$ , where  $k$  is any constant and  $i \neq p$ . In this way we obtain a matrix  $B$  whose rows are the same as the rows of  $A$  except possibly for the  $p$ th row which is  $[ka_{i1} + a_{p1} \ ka_{i2} + a_{p2} \ \cdots \ ka_{in} + a_{pn}]$ . The cofactors  $A_{pj}$  and  $B_{pj}$  are the same for all  $j$ , so

$$\begin{aligned} \det B &= (ka_{i1} + a_{p1})B_{p1} + (ka_{i2} + a_{p2})B_{p2} + \cdots + (ka_{in} + a_{pn})B_{pn} \\ &= (ka_{i1}B_{p1} + ka_{i2}B_{p2} + \cdots + ka_{in}B_{pn}) + (a_{p1}B_{p1} + a_{p2}B_{p2} + \cdots + a_{pn}B_{pn}) \\ &= k \cdot \det M + \det A, \end{aligned}$$

where  $M$  is the matrix obtained from  $A$  by replacing the  $p$ th row of  $A$  with the  $i$ th row of  $A$ . Now, since  $M$  is a matrix with two equal rows, its determinant is zero. So,  $\det B = \det A$ .

**Property 7** If  $A$  and  $B$  are  $n \times n$  matrices, then  $\det AB = (\det A)(\det B)$ .

**Example 13** Use the properties of determinants in evaluating the determinant of the following matrix:

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

*Solution.* At each stage, let  $R_i$  denote the  $i$ th row of the matrix being considered. Notice that if we formed a new matrix,  $B$  by replacing  $R_1$  of matrix  $A$  with  $R_1 + R_3$  then the first row of  $B$  will have three zeros. By Property 6,  $\det A = \det B$ , so we will compute  $\det A$  by computing  $\det B$  expanding by minors on the 1st row.

$$B = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 1 & 5 & 2 \\ 1 & 2 & 0 & 1 \\ -3 & 1 & -1 & 1 \end{bmatrix}$$

$$\det B = 6 \begin{vmatrix} 1 & 5 & 2 \\ 2 & 0 & 1 \\ 1 & -1 & 1 \end{vmatrix} = 6 \det D$$

Now we will form a new matrix  $F$  by replacing  $R_3$  of matrix  $D$  with  $-R_1 + R_3$  and  $R_2$  of matrix  $D$  with  $-2R_1 + R_2$  to get

$$F = \begin{bmatrix} 1 & 5 & 2 \\ 0 & -10 & -3 \\ 0 & -6 & -1 \end{bmatrix}.$$

Since  $\det D = \det F$ ,

$$\det A = \det B = (6) \det D = (6) \det F = 6[(1)((-10)(-1) - (-3)(-6)) - 0 + 0] = -48.$$

Notice that this is the same matrix as in Example 9.

## EXERCISES

1. Show that  $\det A = \det A^T$  for the matrix  $A = \begin{bmatrix} 1 & -4 & 5 \\ 3 & 2 & -1 \\ -2 & 6 & 7 \end{bmatrix}$ .

2. Find the determinant of each matrix given below. Note that the idea here is to try to reduce each matrix to a matrix having two zeros in some row or column and then expanding by minors along that row or column.

$$A = \begin{bmatrix} 6 & -4 & 3 \\ 5 & 0 & 10 \\ -2 & -2 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & 3 & -6 \\ -1 & 2 & -5 \\ 7 & -1 & 1 \end{bmatrix}$$

## B.5 Invertible Matrices

**Definition 10** The identity matrix of order  $n$ , written  $I_n$ , or simply  $I$  is the square matrix  $[a_{ij}]_{nn}$  such that  $a_{ii} = 1$  and  $a_{ij} = 0$  for  $i \neq j$ . In other words, the identity matrix is the matrix with ones along the principal diagonal and zeros elsewhere.

For example, the identity matrix of order 4 is the following matrix:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

If  $A$  is any  $m \times n$  matrix and  $B$  is any  $n \times s$  matrix we have that

$$A \cdot I = A \text{ and } I \cdot B = B.$$

**Definition 11** A matrix  $A$  is called invertible if there is another matrix  $B$  so that the product matrices  $AB$  and  $BA$  are both the identity matrix. The matrix  $B$  is called the inverse matrix of  $A$ , and is usually written  $A^{-1}$ .



It turns out that  $A$  is not invertible unless  $A$  is a square matrix whose determinant is nonzero. If  $A$  is a square matrix and  $\det A \neq 0$ , there is a formula for the entries of  $A^{-1}$ . For example, for  $2 \times 2$  matrices:

$$\text{if } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ then } A^{-1} = \begin{bmatrix} d/\det A & -b/\det A \\ -c/\det A & a/\det A \end{bmatrix}.$$

Here is the explicit formula for the inverse matrix of larger matrices: the  $(i, j)$  entry of  $A^{-1}$  is  $A_{ji}/\det A$ , where  $A_{ji}$  is the cofactor of the  $(j, i)$  entry of  $A$ . *Beware the switch in the order of  $i$  and  $j$ !* For example, in the  $(1, 2)$  entry of the  $2 \times 2$  example above,  $-b$  is the cofactor  $A_{21}$  of the  $(2, 1)$  entry of  $A$ .

**Example 14** Find the inverse of the following matrix:

$$A = \begin{bmatrix} 1 & 5 & 6 \\ -3 & 4 & 8 \\ 2 & -7 & 3 \end{bmatrix}.$$

*Solution.* We have seen in Example 8 that  $\det A = 271$ . We need to compute each cofactor:

$$A_{11} = \begin{vmatrix} 4 & 8 \\ -7 & 3 \end{vmatrix} = 68 \quad A_{12} = - \begin{vmatrix} -3 & 8 \\ 2 & 3 \end{vmatrix} = 25 \quad A_{13} = \begin{vmatrix} -3 & 4 \\ 2 & -7 \end{vmatrix} = 13$$

$$A_{21} = - \begin{vmatrix} 5 & 6 \\ -7 & 3 \end{vmatrix} = -57 \quad A_{22} = \begin{vmatrix} 1 & 6 \\ 2 & 3 \end{vmatrix} = -9 \quad A_{23} = - \begin{vmatrix} 1 & 5 \\ 2 & -7 \end{vmatrix} = 17$$

$$A_{31} = \begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix} = 16 \quad A_{32} = - \begin{vmatrix} 1 & 6 \\ -3 & 8 \end{vmatrix} = -26 \quad A_{33} = \begin{vmatrix} 1 & 5 \\ -3 & 4 \end{vmatrix} = 19$$

We use the formula to get that the inverse of  $A$  is then

$$A^{-1} = \begin{bmatrix} \frac{68}{271} & \frac{-57}{271} & \frac{16}{271} \\ \frac{25}{271} & \frac{-9}{271} & \frac{-26}{271} \\ \frac{13}{271} & \frac{17}{271} & \frac{19}{271} \end{bmatrix}.$$

The most common use of the inverse matrix is in solving equations. A system of  $n$  equations in  $n$  unknown variables can be written as  $A \cdot \mathbf{x} = \mathbf{b}$ , where  $A$  is an  $n \times n$

matrix and both  $\mathbf{x}$  and  $\mathbf{b}$  are  $(n \times 1)$  column vectors. Multiplying both sides on the left by  $A^{-1}$  gives the solution:

$$\mathbf{x} = (A^{-1} \cdot A) \cdot \mathbf{x} = A^{-1} \cdot (A \cdot \mathbf{x}) = A^{-1} \cdot \mathbf{b}$$

For example, to find the solution to the equations

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

we form

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}, A = \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix}, \text{ and } A^{-1} = \begin{bmatrix} 2 & -2.5 \\ -1 & 1.5 \end{bmatrix}.$$

We then get the solution  $x = 1.5, y = 0.5$  from

$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1} \cdot \mathbf{b} = \begin{bmatrix} 2 & -2.5 \\ -1 & 1.5 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ 5 \end{bmatrix} = \begin{bmatrix} 14 - 12.5 \\ -7 + 7.5 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix}.$$

## EXERCISES

1. Find the inverse of the matrix  $A = \begin{bmatrix} -1 & 4 & 5 \\ 3 & 6 & 2 \\ 4 & -3 & 0 \end{bmatrix}$ .

2. (a) Show that the matrix  $A = \begin{bmatrix} 2 & -3 \\ 5 & -7 \end{bmatrix}$  is invertible and find its inverse.

(b) Use the result in (a) to find the solution of the system of equations

$$\begin{aligned} 2x_1 - 3x_2 &= 4 \\ 5x_1 - 7x_2 &= -3 \end{aligned}$$

(c) Use the result in (a) to find the solution of the system of equations

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

3. Find all numbers  $r$  such that the matrix  $A = \begin{bmatrix} 2 & 4 & 2 \\ 1 & r & 3 \\ 1 & 2 & 1 \end{bmatrix}$  is invertible.

4. Find all numbers  $r$  such that the matrix  $A = \begin{bmatrix} 2 & 4 & 2 \\ 1 & r & 3 \\ 1 & 1 & 1 \end{bmatrix}$  is invertible.

## B.6 Cramer's Rule

In this section we exhibit formulas in terms of determinants for the entries in the solution vector of a square linear system  $A \cdot \mathbf{x} = \mathbf{b}$ , where  $A$  is an invertible matrix (in particular,  $\det A \neq 0$ ). The formulas are contained in the following theorem:

**Theorem 2** *Cramer's Rule.* Consider the linear system  $A \cdot \mathbf{x} = \mathbf{b}$ , where  $A = [a_{ij}]_{nn}$  is an  $n \times n$  invertible matrix,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \text{ and } \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Let  $B_p$  be the  $n \times n$  matrix formed from  $A$  by replacing the  $p$ th column vector of  $A$  by the column vector  $\mathbf{b}$ . Then the linear system has the unique solution given by

$$x_1 = \frac{\det B_1}{\det A}, x_2 = \frac{\det B_2}{\det A}, \dots, x_n = \frac{\det B_n}{\det A}.$$

*Proof of Cramer's Rule:* If  $(y_1, y_2, \dots, y_n)$  is a solution of the given linear system, then

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n &= b_1, \\ a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n &= b_2, \\ \vdots & \\ a_{n1}y_1 + a_{n2}y_2 + \dots + a_{nn}y_n &= b_n. \end{aligned} \tag{1}$$

We multiply the first equation by  $A_{11}$ , the second equation by  $A_{21}$ ,  $\dots$ , the  $n$ th equation by  $A_{n1}$  and add the resulting equations to obtain the equation

$$(a_{11}A_{11} + a_{21}A_{21} + \dots + a_{n1}A_{n1})y_1 + (a_{12}A_{11} + a_{22}A_{21} + \dots + a_{n2}A_{n1})y_2 + \dots + (a_{1n}A_{11} + a_{2n}A_{21} + \dots + a_{nn}A_{n1})y_n = b_1A_{11} + b_2A_{21} + \dots + b_nA_{n1}.$$

Notice that the coefficients of  $y_j$  for  $j \neq 1$  are zero (why?), so this equation can be written

$$(\det A)y_1 = \det B_1. \tag{2}$$

In an analogous manner, we also deduce from (1) that

$$(\det A)y_p = \det B_p \text{ for each } 2 \leq p \leq n. \tag{3}$$

From (2) and (3), and because  $\det A \neq 0$  we get

$$y_p = \frac{\det B_p}{\det A} \quad \text{for each } 1 \leq p \leq n. \tag{4}$$

This proves that *if*  $(y_1, y_2, \dots, y_n)$  is a solution of the given linear system, then the values of  $y_p$  must be the unique numbers shown in (4).

To complete the proof, we must show that (4) *is* a solution of the given system; that is, we must show that (4) implies (1). But this follows from the fact that all the above steps are reversible.

**Example 15** Solve the system

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= 11, \\ 4x_1 - 2x_2 + x_3 &= -15, \\ 3x_1 + 4x_2 - x_3 &= 3 \end{aligned}$$

by Cramer's rule.

*Solution.* Using the notation in the above theorem we find that

$$\det A = \begin{vmatrix} 1 & 3 & -2 \\ 4 & -2 & 1 \\ 3 & 4 & -1 \end{vmatrix} = -25, \quad \det B_1 = \begin{vmatrix} 11 & 3 & -2 \\ -15 & -2 & 1 \\ 3 & 4 & -1 \end{vmatrix} = 50,$$

$$\det B_2 = \begin{vmatrix} 1 & 11 & -2 \\ 4 & -15 & 1 \\ 3 & 3 & -1 \end{vmatrix} = -25, \quad \det B_3 = \begin{vmatrix} 11 & 3 & 11 \\ -15 & -2 & -15 \\ 3 & 4 & 3 \end{vmatrix} = 125.$$

Hence

$$x_1 = \frac{50}{-25} = -2,$$

$$x_2 = \frac{-25}{-25} = 1,$$

$$x_3 = \frac{125}{-25} = -5.$$

## EXERCISES

1. Solve the given system of linear equations by Cramer's rule.

$$x_1 + 2x_2 = -2$$

$$x_1 + 5x_2 = -2$$

2. Solve the given system of linear equations by Cramer's rule. Notice that the coefficient matrix is from a previous exercise.

$$x_1 + 2x_2 - x_3 = -2$$

$$2x_1 + x_2 + x_3 = 1$$

$$3x_1 - x_2 + 5x_3 = 1$$

3. Solve the given system of linear equations by Cramer's rule.

$$3x_1 + 2x_2 - x_3 = 1$$

$$x_1 - 4x_2 + x_3 = -2$$

$$5x_1 + 2x_2 = 1$$

4. Tell why Cramer's rule does not apply to the following system of linear equations. Then solve the given system by elementary row operations on its augmented matrix.

$$11x_1 + 5x_2 - 5x_3 = -8$$

$$2x_1 + x_2 - x_3 = -2$$

$$7x_1 + 3x_2 - 3x_3 = -4$$

## B.7 Eigenvalues

Let us define the eigenvalues and eigenvectors of an  $n \times n$  matrix  $A$ . Let  $\mathbf{x}$  be an  $n \times 1$  column matrix, (also called a column vector), and  $\lambda$  a number. Then if

$$A \cdot \mathbf{x} = \lambda \mathbf{x}, \quad \mathbf{x} \neq \mathbf{0},$$

$\mathbf{x}$  is called an *eigenvector* of  $A$  for the *eigenvalue*  $\lambda$ . In other words,  $\mathbf{x}$  is an eigenvector of  $A$  if and only if the result of  $A$  acting on  $\mathbf{x}$  is to multiply by a constant — and the constant is the eigenvalue. For example,

$$\begin{bmatrix} 1 & -3 \\ 2 & 6 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 9 \\ -6 \end{bmatrix} = 3 \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

So  $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$  is an eigenvector with eigenvalue 3.

We may rewrite the eigenvalue equation  $A \cdot \mathbf{x} = \lambda \mathbf{x}$  in the form

$$(A - \lambda I) \cdot \mathbf{x} = \mathbf{0},$$

where the  $\mathbf{0}$  on the right hand side stands for a column vector with all entries 0. It is not hard to see that if  $\mathbf{x} \neq \mathbf{0}$  then  $(A - \lambda I)$  is not invertible. Indeed, if it were invertible we would have the contradiction

$$\mathbf{x} = (A - \lambda I)^{-1} \cdot \mathbf{0} = \mathbf{0}.$$

We conclude: if  $\lambda$  is an eigenvalue of  $A$ , then we must have

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

$\det(A - \lambda I)$  is called the *characteristic polynomial* of the matrix  $A$ . When the determinant is multiplied out, the characteristic polynomial is seen to be a polynomial of degree  $n$  in  $\lambda$ .

This leads to the following way to find the eigenvalues of a matrix  $A$ : they are the roots of the characteristic polynomial. For example,

$$A = \begin{bmatrix} 1 & -3 \\ 2 & 6 \end{bmatrix} \text{ has } \det(A - \lambda I) = \det \begin{bmatrix} 1 - \lambda & -3 \\ 2 & 6 - \lambda \end{bmatrix} = \lambda^2 - 7\lambda + 12.$$

Factoring the characteristic polynomial as

$$\lambda^2 - 7\lambda + 12 = (\lambda - 3)(\lambda - 4) = 0,$$

we see that the eigenvalues of  $A$  are  $\lambda = 3$  and  $\lambda = 4$ . In fact, we have already seen that  $\lambda = 3$  is an eigenvalue, and  $\lambda = 4$  is an eigenvalue because

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

We will see later how to find eigenvectors for the eigenvalues of  $A$ . The point is that we don't have to find any eigenvectors in order to find the eigenvalues of a matrix.

In general, any polynomial of degree  $n$  has at most  $n$  roots. The characteristic polynomial is no exception. This means that an  $n \times n$  matrix can have at most  $n$  eigenvalues.

Sometimes we have a *double* eigenvalue. For example, if

$$A = \begin{bmatrix} 5 & 4 \\ -1 & 1 \end{bmatrix} \quad \text{then} \quad \det(A - \lambda I) = \det \begin{bmatrix} 5 - \lambda & 4 \\ -1 & 1 - \lambda \end{bmatrix}$$

and the characteristic polynomial is

$$\lambda^2 - 6\lambda + 9 = (\lambda - 3)^2.$$

Here  $\lambda = 3$  is a double eigenvalue (and in fact is the only eigenvalue of  $A$ ).

If we know the eigenvalues of a matrix  $A$ , we can determine the eigenvalues of all powers of  $A$ , i.e., the eigenvalues of  $A^2$ ,  $A^3$ ,  $\dots$ . Let  $\mathbf{x}$  be an eigenvector of  $A$ , with eigenvalue  $\lambda$ . Then

$$A^2 \cdot \mathbf{x} = A \cdot (A \cdot \mathbf{x}) = A \cdot (\lambda \mathbf{x}) = \lambda(A \cdot \mathbf{x}) = \lambda^2 \mathbf{x}.$$

Similarly, for any positive integer  $k$ ,

$$A^k \cdot \mathbf{x} = \lambda^k \mathbf{x}.$$

This important result tells us that the eigenvectors of  $A$  are also the eigenvectors of  $A^k$ , but the eigenvalues for  $A^k$  are  $\lambda^k$  instead of  $\lambda$ .

## EXERCISES

1. Show that the characteristic polynomial of the following matrix

$$A = \begin{bmatrix} 5 & 1 & 1 \\ 2 & 4 & 1 \\ -6 & -3 & 0 \end{bmatrix}$$

is  $\lambda^3 - 9\lambda^2 + 27\lambda - 27 = (\lambda - 3)^3$ . ( $\lambda = 3$  is a triple eigenvalue of  $A$ .)

2. Let  $A = \begin{bmatrix} 7 & 5 \\ -10 & -8 \end{bmatrix}$ .

- (a) Find the eigenvalues of  $A$ .
- (b) Find the eigenvalues of  $A^2$  by doing the matrix product  $B = A \cdot A$  and finding the eigenvalues of  $B$ .
- (c) Find the eigenvalues of  $A^5$ .

3. Find the eigenvalues of the matrix  $\begin{bmatrix} -2 & 0 & 0 \\ -5 & -2 & -5 \\ 5 & 0 & 3 \end{bmatrix}$ .



## B.8 Eigenvectors

Suppose that we have found out that  $\lambda$  is an eigenvalue of the matrix  $A$ . How do we obtain the eigenvectors that go with the eigenvalue  $\lambda$ ? The equation satisfied by the eigenvectors is

$$(A - \lambda I) \cdot \mathbf{x} = 0.$$

We can think of this equation as a system of  $n$  equations in  $n$  unknowns. To solve this system of linear equations, we saw in section 1 that it is easiest to use elementary row operations. In fact, that is how computer programs do it! For the matrix  $A = \begin{bmatrix} 1 & -3 \\ 2 & 6 \end{bmatrix}$  we have already found that the eigenvalues are  $\lambda = 3$  and  $\lambda = 4$ . The computation of the eigenvectors proceeds as shown below:

For  $\lambda = 3$  we have

$$A - \lambda I = \begin{bmatrix} 1 & -3 \\ 2 & 6 \end{bmatrix} - 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -2 & -3 \\ 2 & 3 \end{bmatrix}.$$

So we are looking for the vector  $\mathbf{x}$  which satisfies the equation

$$\begin{bmatrix} -2 & -3 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

By row reduction we get

$$\begin{bmatrix} -2 & -3 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

so  $\mathbf{x} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$  is an eigenvector for  $\lambda = 3$ . Similarly, for  $\lambda = 4$  we have

$$\begin{bmatrix} -3 & -3 \\ 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ or}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ so that } \mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ is an eigenvector for } \lambda = 4.$$

Note that the eigenvectors are defined only up to a *non-zero* constant.

(N.B.: the zero vector does *not* count as an eigenvector of anything).

A basic fact about the number of eigenvectors is that a  $k$ -fold eigenvalue has associated with it at most  $k$  “independent” eigenvectors. (Every eigenvalue has at least one eigenvector, because an eigenvector is need to *define* an eigenvalue). However, it can have fewer than  $k$  “independent” eigenvectors.

It turns out that a  $2 \times 2$  matrix with a *double* eigenvalue either has only one eigenvector, or else it is a multiple of the identity,  $A = \lambda I$ . In the latter case, *every* vector (except 0) is an eigenvector. When we count *two* eigenvectors, we appeal to the fact that any vector can be written as a linear combination of *two* independent vectors. For example,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

## EXERCISES

1. In the last section, the matrix  $\begin{bmatrix} 5 & 4 \\ -1 & 1 \end{bmatrix}$  was found to have a double eigenvalue

$$\lambda = 3. \text{ Show that it has only one eigenvector, } \mathbf{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

2. Find the eigenvalues and eigenvectors of the matrices  $\begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$  and  $\begin{bmatrix} 1 & 1 \\ 6 & 2 \end{bmatrix}$ .

3. Find all the eigenvectors for the triple eigenvalue  $\lambda = 3$  of the matrix

$$A = \begin{bmatrix} 5 & 1 & 1 \\ 2 & 4 & 1 \\ -6 & -3 & 0 \end{bmatrix}.$$

4. Find the eigenvalues and eigenvectors of the the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -8 & 4 & -6 \\ 8 & 1 & 9 \end{bmatrix}.$$

5. Find the characteristic polynomial, the eigenvalues, and the corresponding eigenvectors of the matrices  $A$ ,  $A^2$  and  $A^5$  where  $A = \begin{bmatrix} -1 & -2 \\ 4 & 5 \end{bmatrix}$ .

## B.9 Final Remarks

Let us look at some of the other results used in Keller's *Population Projection*.

- (1) When an  $n \times n$  matrix  $A$  has  $n$  eigenvectors (say  $\mathbf{x}_1, \dots, \mathbf{x}_n$  with associated eigenvalues  $\lambda_1, \dots, \lambda_n$ ), any vector  $\mathbf{y}$  can be written as a linear combination of those eigenvectors:

$$\mathbf{y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_n\mathbf{x}_n.$$

Now let  $A$  act on  $\mathbf{y}$   $k$  times:  $A^k\mathbf{y} = A \cdot A \cdots A \cdot \mathbf{y}$ . Then we get

$$A^k \cdot \mathbf{y} = (\lambda_1^k a_1)\mathbf{x}_1 + (\lambda_2^k a_2)\mathbf{x}_2 + \dots + (\lambda_n^k a_n)\mathbf{x}_n.$$

This gives us a way to calculate the effect of  $A^k$  on the vector  $\mathbf{y}$ , and avoids the need for calculating the matrix  $A^k$ , as is done in the UMAP supplement, p.5.

- (2) When there is one eigenvalue  $\lambda_1$  that is *larger in absolute value* than any other ( $|\lambda_1| > |\lambda_i|$  for all  $i \neq 1$ ), and  $k$  large enough,  $A^k \cdot \mathbf{y}$  looks like a multiple of the eigenvector  $\mathbf{x}_1$  of  $\lambda_1$ . (Recall that eigenvalues may be positive, negative, or complex, so that you could have two eigenvalues that are equal in absolute value without being numerically equal).

To see this, suppose that  $a_1 \neq 0$  in (1) above. As  $k$  is allowed to get larger, the term in  $\mathbf{x}_1$  gets larger *relative to the others* in absolute value, and  $A^k \cdot \mathbf{y}$  looks more and more like a multiple of  $\mathbf{x}_1$ :

$$A^k \cdot \mathbf{y} \cong \lambda^k a_1 \mathbf{x}_1.$$

You will encounter this type of behavior in calculations with the Leslie matrix, which has a positive  $\lambda_1$ . It justifies the following assertions on p.10 of the UMAP:

- i. The age distribution vectors  $A^k \cdot \mathbf{y}$  eventually behave like a multiple of a fixed vector (the eigenvector  $\mathbf{x}_1$ ). Given the initial population  $\mathbf{y}$ , we can estimate the population after  $k$  generations, because it is approximately  $\lambda_1^k$  times  $a_1 \mathbf{x}_1$ , where  $y = a_1 \mathbf{x}_1 + \dots + a_n \mathbf{x}_n$ .
- ii. The population eventually tends to grow at a rate proportional to itself, being multiplied by  $\lambda_1$  every generation. That is, *over long time intervals* the population will approximately satisfy the differential equation  $d\mathbf{y}/dt = \lambda_1 \mathbf{y}$ , and the population will increase exponentially.

(3) If an  $n \times n$  matrix has  $n$  different eigenvectors, then it can be *diagonalised*. This has the following meaning. Let us order the eigenvectors by labeling them  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , and the corresponding eigenvalues  $\lambda_1, \dots, \lambda_n$  (it is possible for some of the  $\lambda_i$  to be equal). Make up a matrix  $P$  such that the first column is  $\mathbf{x}_1$ , the second column is  $\mathbf{x}_2$ , ... and the  $n$ th column is  $\mathbf{x}_n$ . Make up the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

Then it is true that  $A \cdot P = P \cdot D$ . This matrix equation is nothing more than all the equations

$$A \cdot \mathbf{x}_i = \lambda_i \mathbf{x}_i$$

combined into a single equation. Multiplying on the left by  $P^{-1}$  gives

$$P^{-1}AP = D \quad \text{and} \quad A = PDP^{-1}.$$

We can do this because  $P^{-1}$  exists whenever  $A$  has  $n$  eigenvectors. (There is a generalisation for matrices with fewer than  $n$  eigenvectors, but the matrix that replaces  $D$  is not quite diagonal).

From  $A = PDP^{-1}$  we have by  $k$ -fold multiplication, using  $P^{-1}P = I$ ,

$$A^k = (PDP^{-1})(PDP^{-1}) \dots (PDP^{-1}) = PD(P^{-1}P)D(P^{-1}P)D \dots DP^{-1},$$

and by cancelling we get  $A^k = PD^kP^{-1}$ . The matrix  $D^k$  is calculated easily: it is the diagonal matrix with entries  $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ . This gives us an easy way to calculate high powers of  $A$ .

Now suppose that one eigenvalue  $\lambda_1$  is larger in absolute value than all the others. For large  $k$ , this means that  $\lambda_i^k/\lambda_1^k \approx 0$  for all the other eigenvalues  $\lambda_i$ . Therefore

$$D^k = \begin{bmatrix} \lambda_1^k & & & 0 \\ & \lambda_2^k & & \\ & & \ddots & \\ 0 & & & \lambda_n^k \end{bmatrix} \approx \lambda_1^k \begin{bmatrix} 1 & & & 0 \\ & \lambda_2^k/\lambda_1^k & & \\ & & \ddots & \\ 0 & & & \lambda_n^k/\lambda_1^k \end{bmatrix} \approx \begin{bmatrix} 1 & & & 0 \\ & 0 & & \\ & & 0 & \\ & & & \ddots \\ 0 & & & & 0 \end{bmatrix}$$

This means that  $A^k = P \cdot D^k \cdot P^{-1}$  is approximately  $\lambda_1^k$  times the matrix

$$M = P \cdot \begin{bmatrix} 1 & & & 0 \\ & 0 & & \\ & & 0 & \\ & & & \ddots \\ 0 & & & & 0 \end{bmatrix} \cdot P^{-1}.$$

This proves the assertion on p.16 of the UMAP supplement:

$$\lim_{k \rightarrow \infty} A^k/\lambda_1^k = M.$$

If  $u$  is the first column of  $P$  (an  $n \times 1$  matrix) and  $v$  is the first row of  $P^{-1}$  (a  $1 \times n$  matrix), then the matrix product  $uv$  is an  $n \times n$  matrix, and in fact is the matrix  $M$ . The argument presented on pp.15-17 of the UMAP supplement is not much different from the one we just gave, but provides more discussion of  $u$  and  $v$ .