MAT 1332: Calculus for Life Sciences

A course based on the book
Modeling the dynamics of life
by F.R. Adler

Supplementary material
University of Ottawa

Frithjof Lutscher

February 24, 2009

Complex numbers

Introductory consideration

We can easily solve the equation $x^2 - 4 = 0$. The answer is $x = \pm 2$, in particular, x is a rational number, even an integer. The equation $x^2 - 2 = 0$ is a bit more tricky. The solution $x = \pm\sqrt{2}$ is *not* a rational number. Instead, we have *defined* the square root of a positive number as the real number that gives the original number back when multiplied by itself. But what should we do with the equation $x^2 + 1 = 0$? The answer cannot be a real number. (Why?) Can we do the same as above and *define* a number whose square equals -1? Indeed, this is what mathematicians did in the eighteenth century (it was a daring act and caused a lot of controversy), and they call that number 'i' or *imaginary*. (We will see that complex numbers are hardly more imaginary than $\sqrt{2}$.)

Definition

A *complex number* z is a number of the form

$$z = a + b\mathrm{i}$$

with real numbers a, b and the symbol i that satisfies $\mathrm{i}^2 = -1$. We call $a = \mathrm{Re}(z)$ the *real part* of z and $b = \mathrm{Im}(z)$ the *imaginary part* of z. The real number a can be considered the complex number $a + 0\mathrm{i}$. A complex number of the form $z = b\mathrm{i}$ is called *purely imaginary.*

Addition, subtraction, and multiplication of complex numbers

Complex numbers are easily added, subtracted and multiplied, if we keep the rule $\mathrm{i}^2 = -1$ in mind and use the distributive laws.

$$(a + b\mathrm{i}) \pm (c + d\mathrm{i}) = (a \pm c) + (b \pm d)\mathrm{i}$$

$$(a + b\mathrm{i}) * (c + d\mathrm{i}) = ac + bc\mathrm{i} + ad\mathrm{i} + bd\mathrm{i}^2 = (ac - bd) + (ad + bc)\mathrm{i}$$

Examples

1. $(3 + 5\mathrm{i}) + (2 - 7\mathrm{i}) = 5 - 2\mathrm{i}$
2. $(0.5 + 1.7\mathrm{i}) - (0.8 - 2.6\mathrm{i}) = -0.3 + 4.3\mathrm{i}$
3. $(-3 + 2\mathrm{i}) * (4 - 5\mathrm{i}) = (-12 - (-10)) + (15 + 8)\mathrm{i} = -2 + 23\mathrm{i}$
4. $(2 - 0.5\mathrm{i}) * (3 + 4\mathrm{i}) = (6 - (-2)) + (-1.5 + 8)\mathrm{i} = 8 + 6.5\mathrm{i}$
5. $(9 + 2\mathrm{i}) + 5 = (9 + 2\mathrm{i}) + (5 + 0\mathrm{i}) = 14 + 2\mathrm{i}$
6. $-3\mathrm{i} + (2 + 3\mathrm{i}) = (0 - 3\mathrm{i}) + (2 + 3\mathrm{i}) = 2 + 0\mathrm{i} = 2$
7. $2 * (3 - 5\mathrm{i}) = 6 - 10\mathrm{i}$

8. $3\mathrm{i} * (-1 + 4\mathrm{i}) = -12 - 3\mathrm{i}$

Before we look at inverses and division of complex numbers, we introduce the *complex conjugate* of a complex number.

Definition and observation

The *complex conjugate* of $z = a + b\mathrm{i}$ is $\bar{z} = a - b\mathrm{i}$, i.e., we simply change the sign of the imaginary part. Since the multiplication

$$z\bar{z} = (a + b\mathrm{i})(a - b\mathrm{i}) = a^2 + b^2$$

always produces a non-negative real number, we can take the square root. We define the *modulus* or *absolute value* of $z = a + b\mathrm{i}$ as

$$|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}.$$

From the identity $z\bar{z} = |z|^2$, we find the inverse of z to be

$$\frac{1}{z} = z^{-1} = \bar{z}/|z|^2.$$

Example 1

Start with $z = 3 + 4\mathrm{i}$ and $w = 2 - \mathrm{i}$. The complex conjugates are $\bar{z} = 3 - 4\mathrm{i}$ and $\bar{w} = 2 + \mathrm{i}$. The absolute values are $|z| = 5$ and $|w| = \sqrt{5}$. The inverses are

$$z^{-1} = \frac{1}{25}(3 - 4\mathrm{i}), \qquad w^{-1} = \frac{1}{5}(2 + \mathrm{i}).$$

Finally, we can divide

$$\frac{z}{w} = z\frac{\bar{w}}{|w|^2} = \frac{1}{5}(2 + 11\mathrm{i}), \quad \frac{w}{z} = w\frac{\bar{z}}{|z|^2} = \frac{1}{25}(2 - 11\mathrm{i}).$$

Example 2

Start with $z = 1 - 4\mathrm{i}$ and $w = 0.5 + 3\mathrm{i}$. The complex conjugates are $\bar{z} = 1 + 4\mathrm{i}$ and $\bar{w} = 0.5 - 3\mathrm{i}$. The absolute values are $|z| = \sqrt{17}$ and $|w| = \sqrt{37/4}$. The inverses are

$$z^{-1} = \frac{1}{17}(1 + 4\mathrm{i}), \qquad w^{-1} = \frac{4}{37}(0.5 - 3\mathrm{i}).$$

Division gives

$$\frac{z}{w} = z\frac{\bar{w}}{|w|^2} = \frac{4}{37}(-11.5 - 5\mathrm{i}), \quad \frac{w}{z} = w\frac{\bar{z}}{|z|^2} = \frac{1}{17}(-11.5 + 5\mathrm{i}).$$

Geometric interpretation

It is very helpful to think of a complex number as a point in the plane with the real part as the x-value and the imaginary part as the y-value. Hence, we identify the complex number $z = a + b\mathrm{i}$ with the point (a, b) or with the vector (arrow) from the origin to the point (a, b). (We will talk about vectors in more detail shortly). Then the absolute value of the complex number is simply the distance of the corresponding point from the origin or the length of the vector (arrow), see Figure 3.

With this correspondence, the addition of complex numbers become the addition of vectors as it is known from the physics of forces, see Figure 3.

To interpret multiplication, we take a slightly different point of view. Instead of giving the coordinates of the vector as the endpoint (a, b), we consider its length $r \geq 0$ and the angle ϕ it makes with the x-axis (counterclockwise) as $(r\cos\phi, r\sin\phi)$. This representation is called *polar coordinates*. Then multiplication of two numbers is simply multiplication of the lengths and addition of the angles, see Figure 3. We write

$$z = r(\cos\phi + \mathrm{i}\sin\phi), \quad \text{and} \quad w = s(\cos\psi + \mathrm{i}\sin\psi).$$

Then we multiply, using the trigonometric identities

$$\begin{aligned} zw &= r(\cos\phi + \mathrm{i}\sin\phi) * s(\cos\psi + \mathrm{i}\sin\psi) \\ &= rs[\cos\phi\cos\psi - \sin\phi\sin\psi + \mathrm{i}(\cos\phi\sin\psi + \cos\psi\sin\phi)] \\ &= rs[\cos(\phi+\psi) + \mathrm{i}\sin(\phi+\psi)]. \end{aligned}$$

Observation and definition

Every complex number of the form $z = \cos\phi + \mathrm{i}\sin\phi$ has absolute value one. We introduce the exponential notation (known as Euler's formula)

$$\exp(\mathrm{i}\phi) = e^{\mathrm{i}\phi} = \cos\phi + \mathrm{i}\sin\phi.$$

It might look strange at first, but the same rules as for the real exponential function apply. This has many advantages. First of all, we can write any complex number in polar coordinates as $z = re^{\mathrm{i}\phi}$. And we can easily multiply complex numbers in this form. For example, the calculation above becomes a single step (no need to look up the trig identities)

$$re^{\mathrm{i}\phi} * se^{\mathrm{i}\psi} = rse^{\mathrm{i}(\phi+\psi)}.$$

Examples

1. The complex number $z = 1 + \mathrm{i}$ has modulus $|z| = \sqrt{2}$ and angle $\phi = \pi/4$. Hence $z = 1 + \mathrm{i} = \sqrt{2}e^{\mathrm{i}\pi/4}$.
2. The complex number $w = \sqrt{3} + \mathrm{i}$ has modulus $|w| = 2$ and angle $\phi = \pi/6$. Hence $w = \sqrt{3} + \mathrm{i} = 2e^{\mathrm{i}\pi/6}$.
3. Their product is $zw = (\sqrt{3} - 1) + (\sqrt{3} + 1)\mathrm{i} = 2\sqrt{2}e^{\mathrm{i}5\pi/12}$.

4. In general, if $z = a + b\mathrm{i}$ then $r = |z| = \sqrt{a^2 + b^2}$. The argument ϕ is not uniquely defined. If we restrict it between $-\pi$ and π the we get

$$\begin{cases} \phi = \arctan(b/a) & \text{if} \quad a > 0 \\ \phi = \arctan(b/a) + \pi & \text{if} \quad a < 0, b > 0 \\ \phi = \arctan(b/a) - \pi & \text{if} \quad a < 0, b < 0 \end{cases}$$

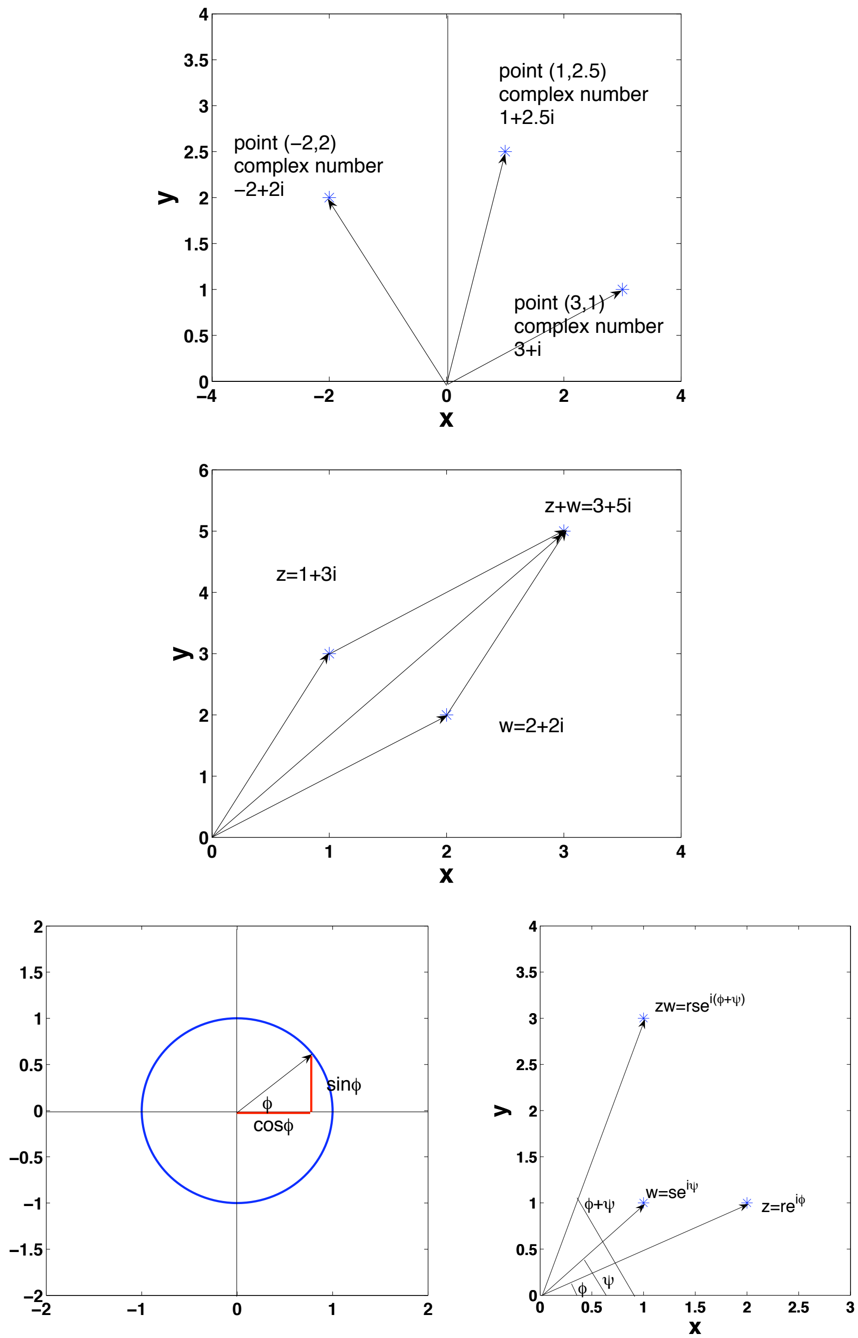

Figure 3: Top panel: correspondence between complex numbers and points in the plane. Middle panel: Addition of complex numbers. Bottom panel: polar coordinates on the unit circle (left) and multiplication of complex numbers using polar coordinates (right)

Linear Algebra I - Linear systems of equations

Introductory example

Suppose there are two types of food. Type I contains 10g of protein and 5g of carbohydrates per 100g, type II contains 8g of protein and 12g of carbohydrates per 100g.

Easy question Suppose I take 50g of type I and 75g of type II, how much protein and carbohydrates do I get? Answer: First we have to choose units, let's say 100g. Then I take 0.5 units of type I and 0.75 units of type II. Now let x_1, x_2 denote the respective units of food type I and II. Then the amount of protein and carbohydrates are given by

$$\begin{aligned} b_1 &= 10x_1 + 8x_2 = 10 * 0.5 + 8 * 0.75 = 11, \\ b_2 &= 5x_1 + 12x_2 = 5 * 0.5 + 12 * 0.75 = 11.5, \end{aligned}$$

grams respectively. NOTE: we are given the x_1, x_2 and want to find the b_1, b_2.

Harder question Suppose I want to take 16g of protein and 20g of carbohydrates. How much of each food type to I have to take? The equations are just as above.

$$\begin{aligned} 10x_1 + 8x_2 &= b_1 = 16, \\ 5x_1 + 12x_2 &= b_2 = 20. \end{aligned}$$

HOWEVER: this time, with the same notation as above, we are given the b_1, b_2 and want to get the x_1, x_2.
SOLUTION: We multiply the second equation by -2:

$$\begin{aligned} 10x_1 + 8x_2 &= 16, \\ -10x_1 - 24x_2 &= -40. \end{aligned}$$

Then we add the two equations:

$$-16x_2 = -24,$$

which gives $x_2 = 3/2$. This we put back into one of the original equations to get $10x_1 + 12 = 16$ or $x_1 = 2/5$.

The example above is a special case of a *linear system of equations.* More generally, we write a *linear system of m equations with n unknowns* as

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

where

1. $x_1, \dots x_n$ are the variables or unknowns,
2. $a_{11}, \dots a_{mn}$ are the coefficients,
3. $b_1, \dots b_m$ are the right hand side.

A *solution* of such a system is a set of numbers $(s_1, \dots, s_n)$ that makes *all* the equations true when we substitute the s_i for the x_i. The equations are called *linear* since each of the variables x_i appears only linearly (as opposed to higher powers or other nonlinear functions). The subject of linear algebra is to study such linear systems of equations.

Goal of this chapter

In this chapter we learn how to solve linear systems of equations. In particular, we answer the three questions:

1. Is there always a solution?
2. Can there be more than one solution?
3. How can we compute all solutions?

Fact: A linear system of equations has either

1. exactly one solution,
2. infinitely many solutions, or
3. no solution.

Cases (1) and (2) are called *consistent* whereas case (3) is called *inconsistent*.

Examples

Consider the three systems of equations:

$$\left\{ \begin{array}{ll} x_1 - x_2 & = 1 \\ x_1 - x_2 & = 0 \end{array} \right\}, \qquad \left\{ \begin{array}{ll} x_1 - x_2 & = 1 \\ x_1 - 2x_2 & = 0 \end{array} \right\}, \qquad \left\{ \begin{array}{ll} x_1 - x_2 & = 1 \\ -x_1 + x_2 & = -1 \end{array} \right\}.$$

The first system has no solution, the second has exactly one solution ($x_1 = 2, x_2 = 1$) and the last system has infinitely many solutions (for all real numbers t, the pair $x_1 = t, x_2 = t - 1$ works). We can see graphically, why there are these three cases. Each equation is the equation of a line in x_1-x_2-space. In the first case, the two lines are parallel and have no point in common. In the second case, there is one point of intersection. In the third case, the lines are identical, and all points are part of both lines, see Figure 4.

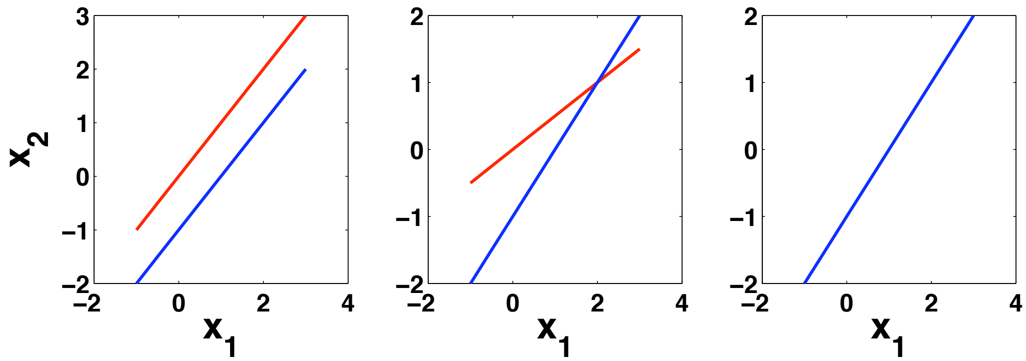

Figure 4: Graphical interpretation of systems of linear equations. Left: no solution, Middle: unique solution, Right: infinitely many solutions.

The Gaussian Elimination Algorithm

Elementary row operations: The solution of a linear system of equations does not change under the following three elementary row operations:

1. multiply a row by a nonzero number,
2. add a multiple of one row to another,
3. exchange the order of two rows.

Example 1

Find the solution of

$$\begin{cases} 2x_1 - x_2 &= 1 \\ 4x_1 + 2x_2 &= 10 \end{cases}$$

Answer:

$$\begin{cases} 2x_1 - x_2 &= 1 \\ 4x_1 + 2x_2 &= 10 \end{cases} \xrightarrow{(-0.5)*R2} \begin{cases} 2x_1 - x_2 &= 1 \\ -2x_1 - x_2 &= -5 \end{cases} \xrightarrow{R2+R1} \begin{cases} 2x_1 - x_2 &= 1 \\ \quad -2x_2 &= -4 \end{cases}$$

$$\xrightarrow{-2*R1+R2} \begin{cases} -4x_1 & &= -6 \\ & -2x_2 &= -4 \end{cases} \xrightarrow{(-1/4)*R1}\xrightarrow{(-1/2)*R2} \begin{cases} x_1 & &= 3/2 \\ & x_2 &= 2 \end{cases}$$

Now, the solution is obvious: $x_1 = 3/2$ and $x_2 = 2$, or $(3/2, 2)$.

Example 2

Find the solution of

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

Answer:

$$\begin{cases} 2x_2+3x_3 &= 1\\ x_1+x_2+x_3 &= 2\\ x_1-x_2 &= -1\end{cases} \xrightarrow{R1\leftrightarrow R2} \begin{cases} x_1+x_2+x_3 &= 2\\ 2x_2+3x_3 &= 1\\ x_1-x_2 &= -1\end{cases} \xrightarrow{(-1)*R3+R1} \begin{cases} x_1+x_2+x_3 &= 2\\ 2x_2+3x_3 &= 1\\ 2x_2+x_3 &= 3\end{cases}$$

$$\xrightarrow{(-1)*R3+R2} \begin{cases} x_1+x_2+x_3 &= 2\\ 2x_2+3x_3 &= 1\\ 2x_3 &= -2\end{cases} \xrightarrow{(1/2)*R3} \overset{(-3)*R3+R2}{\xrightarrow{R1-R3}} \begin{cases} x_1+x_2 &= 3\\ 2x_2 &= 4\\ x_3 &= -1\end{cases}$$

and finally

$$\overset{(1/2)*R2}{\xrightarrow{R1-R2}} \begin{cases} x_1 &= 1\\ x_2 &= 2\\ x_3 &= -1\end{cases}$$

So the solution is $(1,2,-1)$.

Simplify Notation!

Looking at the two examples above, we realize that we do not need to write the variables x_i all the time, provided we agree and stick to a particular order. Similarly, we can do away with the '+' and the '=' signs. The only things that matters are the coefficients in front of the x_i and the right hand side. Hence, we collect these into an array of numbers, which we call *matrix*. We can then simply perform the row operations on the rows of the matrix.

For the example above:

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

We call the matrix above the *augmented matrix*, where we have the coefficients, a_{ij} of the linear system together with the right hand side b_i, separated by the vertical lines. We also call the matrix

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

the *coefficient matrix*, i.e., this matrix contains only the coefficients a_{ij}.

The same steps as in example 2 above, but in matrix notation are:

$$\left[\begin{array}{ccc|c} 0 & 2 & 3 & 1\\ 1 & 1 & 1 & 2\\ 1 & -1 & 0 & -1\end{array}\right] \xrightarrow{R1\leftrightarrow R2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2\\ 0 & 2 & 3 & 1\\ 1 & -1 & 0 & -1\end{array}\right] \xrightarrow{(-1)*R3+R1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2\\ 0 & 2 & 3 & 1\\ 0 & 2 & 1 & 3\end{array}\right]$$

$$\xrightarrow{(-1)*R3+R2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2\\ 0 & 2 & 3 & 1\\ 0 & 0 & 2 & -2\end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 0 & 3\\ 0 & 2 & 0 & 4\\ 0 & 0 & 1 & -1\end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1\\ 0 & 1 & 0 & 2\\ 0 & 0 & 1 & -1\end{array}\right]$$

Now, we remember that the first column corresponds to x_1, the second to x_2, and the third to x_3. Then the solution is $x_1 = 1, x_2 = 2, x_3 = -1$ or simply $(1,2,-1)$.

Reduced row-echelon form

Now that we have seen several examples of 'simple' forms of systems where the solutions could easily be read off, we will formalize this process a bit.

Definition: The *leading entry* of a row in a matrix is the leftmost nonzero coefficient in that row.
A matrix is in *row-echelon form* if the following three rules are true

1. Rows of zeros are below any nonzero row.
2. The leading entry of any row is to the right of any leading entry in any row above it.
3. All entries in the column below a leading entry are zero.

A matrix is in *reduced row-echelon form* if it is in row-echelon form and *in addition*

4. Each leading entry is 1.
5. All entries in the column above a leading entry are zero.

Examples

The first two systems are in row-echelon form but not reduced. The third and fourth are in reduced row-echelon form, the last one is neither.

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

Example 3

Find the solution of

$$\begin{cases} x_1 + x_2 + x_3 & = 5 \\ 2x_1 \qquad - x_3 & = 0 \\ \qquad x_2 + 4x_3 & = 15 \end{cases}$$

Answer: Use the matrix notation to do the row operations.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 2 & 0 & -1 & 0 \\ 0 & 1 & 4 & 15 \end{array}\right] \xrightarrow{(-2)R1+R2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & -2 & -3 & -10 \\ 0 & 1 & 4 & 15 \end{array}\right] \xrightarrow[(-1)*R2]{R2+2R3} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & 2 & 3 & 10 \\ 0 & 0 & 5 & 20 \end{array}\right]$$

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

Hence, we have the solution $x_1 = 2, x_2 = -1, x_3 = 4$ or $(2, -1, 4)$.

Example 4

Find the solution of

$$\begin{cases} 2x_1 + x_2 + 2x_3 &= 1 \\ -4x_1 \qquad - x_3 &= 2 \\ 2x_1 + 5x_2 + 8x_3 &= 11 \end{cases}$$

Answer in matrix notation:

$$\left[\begin{array}{ccc|c} 2 & 1 & 2 & 1 \\ -4 & 0 & -1 & 2 \\ 2 & 5 & 8 & 11 \end{array}\right] \overset{\substack{2*R1+R2 \\ -R1+R3}}{\longrightarrow} \left[\begin{array}{ccc|c} 2 & 1 & 2 & 1 \\ 0 & 2 & 3 & 4 \\ 0 & 4 & 6 & 10 \end{array}\right] \overset{(-1/2)*R3+R2}{\longrightarrow} \left[\begin{array}{ccc|c} 2 & 1 & 2 & 1 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

At this point we pause and look at the last row of zeros. In the original notation with variables x_i, this row reads: $0x_1 + 0x_2 + 0x_3 = 2$. This is clearly impossible. Hence, this system is inconsistent, it has no solution.

Example 5

Find the solution of

$$\begin{cases} 2x_1 - 4x_2 - 8x_3 &= -18 \\ 3x_1 + 3x_2 + 15x_3 &= 18 \end{cases}$$

Answer in matrix notation:

$$\left[\begin{array}{ccc|c} 2 & -4 & -8 & -18 \\ 3 & 3 & 15 & 18 \end{array}\right] \overset{\substack{0.5*R1 \\ 1/3*R2}}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & -2 & -4 & -9 \\ 1 & 1 & 5 & 6 \end{array}\right] \overset{R2-R1}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & -2 & -4 & -9 \\ 0 & 3 & 9 & 15 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & -2 & -4 & -9 \\ 0 & 1 & 3 & 5 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 1 & 3 & 5 \end{array}\right]$$

This time, the last row reads

$$x_2 + 3x_3 = 5.$$

Whatever we choose for x_3, we can always find an x_2 to make the equation true. Hence, this equation has infinitely many solutions. We set $x_3 = t$ as a *free variable*, then we get $x_2 = 5 - 3t$. We plug this into the first equation and solve for x_1 as

$$x_1 + 2t = 1, \qquad \text{or} \qquad x_1 = 1 - 2t.$$

Hence, the infinitely many solutions can be written as the set $\{(1-2t, 5-3t, t) \colon t \in \mathbb{R}\}$.

Example 6

Find the solution of

$$\begin{cases} 2x_1 + 2x_2 + 2x_3 &= 0 \\ -2x_1 + 5x_2 + 2x_3 &= 1 \\ 8x_1 + x_2 + 4x_3 &= -1 \end{cases}$$

Answer in matrix notation:

$$\left[\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ -2 & 5 & 2 & 1 \\ 8 & 1 & 4 & -1 \end{array}\right] \overset{\substack{0.5*R1 \\ R1+R2 \\ -4*R1+R3}}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 7 & 4 & 1 \\ 0 & -7 & -4 & -1 \end{array}\right] \overset{R2+R3}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 7 & 4 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Again, we have a row of zeros at the bottom. However, this time the last row reads: $0x_1 + 0x_2 + 0x_3 = 0$. This equation is satisfied for all values of x_1, x_2, x_3. We do not run into the same problem as in example 4. We simply continue as in example 5.

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 4/7 & 1/7 \\ 0 & 0 & 0 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 3/7 & -1/7 \\ 0 & 1 & 4/7 & 1/7 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The second row

$$x_2 + (4/7)x_3 = 1/7.$$

has again infinitely many solutions. We denote $x_3 = t$ as the *free variable* and compute $x_2 = 1/7 - (4/7)t$. In the first row, we get

$$x_1 + (3/7)t = -1/7 \qquad \text{or} \qquad x_1 = -1/7 - 3t/7.$$

Hence, the solution set is $\{(-\frac{1}{7} - \frac{3t}{7}, \frac{1}{7} - \frac{4t}{7}, t) \colon t \in \mathbb{R}\}$.

Example 7

For which values of h does the system

$$\begin{cases} x_1 + hx_2 & = -3 \\ -2x_1 + 4x_4 & = 6 \end{cases}$$

have (a) a unique solution, (b) infinitely many solutions, and (c) no solution?

Answer:

$$\left[\begin{array}{cc|c} 1 & h & 3 \\ -2 & 4 & 6 \end{array}\right] \overset{2*R1+R2}{\longrightarrow} \left[\begin{array}{cc|c} 1 & h & 3 \\ 0 & 2h-4 & 0 \end{array}\right]$$

The last row gives the equation

$$(2h-4)x_2 = 0.$$

If $2h - 4 = 0$, (i.e., $h = 2$) then $x_2 = t$ is a free variable, and $x_1 = -3 - hx_2 = -3 - 2t$. If, on the other hand, $2h - 4 \neq 0$ then the only way to satisfy the second row is $x_2 = 0$. In this case, the first row gives $x_1 = -3$. Hence, if $h \neq 2$ then there is a unique solution, if $h = 2$ then there are infinitely many solutions. For no value of h is there no solution.

Practice Problems

1. Solve the following systems of equations by bringing them into reduced row-echelon form

$$(a)\quad \begin{cases} x_1 + x_2 + 2x_3 &= 9 \\ 2x_1 + 4x_2 - 3x_3 &= 1 \\ 3x_1 + 6x_2 - 5x_3 &= 0 \end{cases} \qquad (b)\quad \begin{cases} 5x_1 - 2x_2 + 6x_3 &= 0 \\ -2x_1 + x_2 + 3x_3 &= 1 \end{cases}$$

$$(c)\quad \begin{cases} 2x_1 \phantom{{}-x_2} + 2x_3 &= 1 \\ 3x_1 - x_2 + 4x_3 &= 7 \\ 6x_1 + x_2 - x_3 &= 0 \end{cases} \qquad (d)\begin{cases} 7x_1 + 2x_2 + x_3 - 3x_4 &= 5 \\ x_1 + 2x_2 + 4x_3 &= 1 \end{cases}$$

$$(e)\quad \begin{cases} 3x_1 + 2x_2 - x_3 &= -15 \\ 3x_1 + x_2 + 3x_3 &= 11 \\ -6x_1 - 4x_2 + 2x_3 &= 30 \end{cases} \qquad (f)\quad \begin{cases} 2x_1 - x_2 - 3x_3 &= 0 \\ -x_1 + 2x_2 - 3x_3 &= 0 \\ x_1 + x_2 + 4x_3 &= 0 \end{cases}$$

2. Problems with parameters

(a) For which values of a, b does the system

$$\begin{cases} x_1 + ax_2 &= 1 \\ 2x_1 + 3x_2 &= b \end{cases}$$

have (i) a unique solution, (ii) infinitely many solutions, or (iii) no solutions?

(b) Explain why the system

$$\begin{cases} x_1 + x_2 + 2x_3 &= a \\ x_1 \phantom{{}+x_2} + x_3 &= b \\ 2x_1 + x_2 + 3x_3 &= c \end{cases}$$

is consistent if $c = a + b$ but inconsistent in all other cases.

3. Application

Insects of two species are reared on two types of food. Species 1 consumes 5 units of food A and 3 units of food B per day. Species 2 consumes 2 units of A and 4 units of B, respectively. Every day, 900 units of food A and 960 units of food B are provided. How many individuals of each species are reared?

Solutions to practice problems

1.

$$(a)\quad (1,2,3) \qquad (b)\quad \{(2-12t, 5-27t, t)\colon t\in\mathbb{R}\} \qquad (c)\quad \left(\frac{11}{12}, -\frac{71}{12}, -\frac{5}{12}\right)$$

$$(d)\quad \left\{\left(\frac{4}{3}+\frac{1}{2}s+\frac{1}{2}t, -\frac{1}{6}-\frac{27}{12}s-\frac{1}{4}t, s, t\right) : s,t\in\mathbb{R}\right\} \qquad (e)\quad \left\{\left(\frac{37}{3}-\frac{7}{3}t, 4t-26, t\right) : t\in\mathbb{R}\right\}$$

$$(f)\quad (0,0,0)$$

2.

(a) Unique solution if $a \neq 3/2$. infinitely many solutions if $a = 3/2$ and $b = 2$. No solution if $a = 3/2$ and $b \neq 2$.

(b) Add the first and second row.

3.

Let x_i be the number of individuals of species i. Then the system is

$$5x_1 + 2x_2 = 900, \qquad 3x_1 + 4x_2 = 960,$$

and the solution is $(x_1, x_2) = (120, 150)$.

Linear Algebra II - Vectors and matrices

In the last section, we introduced matrices to simplify our life, as a short hand notation for linear systems of equations. In this section, we study matrices as objects in their own right. We learn when it is possible to add and multiply them and how to do it. While there is no direct biological application in this section, the content presented here is the foundation of everything to come. It is like learning the grammar of a language so that one can speak it properly later.

Definition: An $m \times n$-matrix A is a rectangular array of numbers with m rows and n columns, i.e.,

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

The numbers a_{ij} are called *entries*.
If $m = n$ then A is a *square matrix*.
A $1 \times n$-matrix is called a row vector: $[c_1, c_2, \dots, c_n]$.
An $m \times 1$-matrix is called a column vector: $\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$.
Two matrices $A = [a_{ij}], B = [b_{ij}]$ are said to be *equal* if they have the same dimension and if for all i, j we have $a_{ij} = b_{ij}$.

We note two special (classes of) matrices, the *zero matrix*

$$0 = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix},$$

and the identity matrix

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

The number of rows and columns of these two matrices is usually clear from the context.
If A is a square matrix, then the elements a_{ii} are called the *diagonal elements* and their sum of called the *trace* of A, i.e.,

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

has the diagonal elements 5, 1, -4 and the trace is $\mathrm{tr}(A) = 5 + 1 - 4 = 2$.

Matrix addition and scalar multiplication

If two matrices, $A = [a_{ij}]$ and $B = [b_{ij}]$ are both of the format $m \times n$, then we can form the sum of the two by entrywise addition to obtain a matrix of the same format

$$C = A + B = [c_{ij}] = [a_{ij} + b_{ij}].$$

If k is a number and $A = [a_{ij}]$ an $m \times n$-matrix, then we define the entrywise product

$$kA = [ka_{ij}],$$

which is again an $m \times n$-matrix.

Examples

Consider the two 2×3-matrices

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

Then we have

$$A + B = \begin{bmatrix} 2 & 2 & 4 \\ -3 & -1 & -1 \end{bmatrix}, \qquad 5A = \begin{bmatrix} 5 & 10 & 15 \\ -15 & -10 & -5 \end{bmatrix},$$

$$3A + 2B = \begin{bmatrix} 3 & 6 & 9 \\ -9 & -6 & -3 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 6 & 11 \\ -9 & -4 & -3 \end{bmatrix}.$$

The transpose of a matrix

The *transpose* of a matrix A is obtained from A by interchanging rows and columns, or, loosely speaking, by flipping the matrix along its diagonal. More formally, if $A = [a_{ij}]$, then the transpose is

$$A^T = [a_{ji}].$$

If A is of the format $m \times n$ then A^T is of the format $n \times m$. In particular, the transpose of a column vector is a row vector and vice versa.

Examples

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

$$v = \begin{bmatrix} -4 \\ 7 \\ 0 \end{bmatrix} \qquad v^T = \begin{bmatrix} -4 & 7 & 0 \end{bmatrix}.$$

Matrix-vector multiplication

If the matrix A has n columns and the column vector x has n rows, then we can define the product Ax as the following column vector:

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

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

To remember this definition, simply think about linear systems of equations!
Note that the resulting vector has the same number of rows as the matrix. In general, the dimensions work as follows:

$$(m \times n) * (n \times 1) = (m \times 1).$$

Examples

1.

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

2.

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

3.

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

4.

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

5.

$$\begin{bmatrix} -3 & 2 \\ -7 & 4 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \end{bmatrix} = \begin{bmatrix} -15-6 \\ -35-12 \\ 5-18 \end{bmatrix} \begin{bmatrix} -21 \\ -47 \\ -13 \end{bmatrix}$$

6.

$$\begin{bmatrix} 30 & 25 \end{bmatrix} \begin{bmatrix} 2 \\ 9 \end{bmatrix} = \begin{bmatrix} 60 + 225 \end{bmatrix} = 285$$

Matrix-matrix multiplication

We use the definition of the matrix-vector product to define a product of two matrices. Consider a matrix A with n columns and a matrix B with n rows. We may think of each column of B as a column vector of length n. We know how to multiply each of these with the matrix A and then we put the resulting vectors into one matrix. More formally, the product of

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & \dots & b_{1k} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nk} \end{bmatrix} = \left[\begin{array}{c|c|c|c} B_1 & B_2 & \dots & B_k \end{array}\right]$$

is given by

$$AB = \left[\begin{array}{c|c|c|c} AB_1 & AB_2 & \dots & AB_k \end{array}\right].$$

Note that the dimensions multiply as follows:

$$(m \times n) * (n \times k) = (m \times k).$$

Examples

Take the following matrices:

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

The product AB is defined and it is

$$AB = \begin{bmatrix} -1 & 2 & 9 \\ -1 & 6 & 19 \end{bmatrix}.$$

The product BA is not defined since the number of rows of B does not equal the number of columns of A. However, if we transpose B first, then we can multiply $B^T A$:

$$\begin{bmatrix} -2 & -2 \\ 2 & 4 \\ 13 & 18 \end{bmatrix}.$$

The product CA is defined, but not vice versa (check the format!):

$$CA = \begin{bmatrix} -2 & -2 \\ 10 & 16 \\ 9 & 12 \end{bmatrix}.$$

Again, if we transpose C, then we can multiply AC^T

$$AC^T = \begin{bmatrix} -1 & 8 & 6 \\ -1 & 20 & 12 \end{bmatrix}.$$

Since A is a square matrix, we can form the power

$$A^2 = AA = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}.$$

Since A and D are square of the same size, we can multiply them both ways:

$$AD = \begin{bmatrix} -1 & 0 \\ -3 & -2 \end{bmatrix}, \quad DA = \begin{bmatrix} -7 & -10 \\ 3 & 4 \end{bmatrix}.$$

NOTE: The order of the product matters! Matrix multiplication is NOT commutative, even if both products are defined.

Multiplication of the identity matrix and the zero matrix are just as easy as multiplying the real numbers zero and one:

$$AI = IA = A, \qquad A0 = 0A = 0.$$

As a final, not so obvious example for matrix multiplication, we note (check the formats!)

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

Note: We have seen that many operations with real numbers (addition, multiplication) also work with (some) matrices. Two of several differences are

1. The commutative law does not hold.
2. The cancellation law does not hold.

Example

The matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is an example where $A^2 = 0$ but $A \neq 0$.

Practice Problems

Consider the following matrices:

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

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

1. Compute the following if possible, if not, explain why

(a) $D+E$	(b) $D-E$	(c) $5A$	(d) $-7C$
(e) $2B-C$	(f) $4E-2D$	(g) $-3(D+2E)$	(h) $A-A$
(i) $\mathrm{tr}(D)$	(j) $\mathrm{tr}(D-3E)$	(k) $\mathrm{tr}(A)$	(l) $\mathrm{tr}(B)$
(m) $2A^T+C$	(n) D^T-E^T	(o) $(D-E)^T$	(p) $B-B^T$

2. Compute the following if possible, if not, explain why

(a) AB	(b) BA	(c) $3ED$	(d) $(AB)C$	(e) $A(BC)$
(f) CC^T	(g) $(DA)^T$	(h) $(C^TB)A^T$	(i) $\mathrm{tr}(DD^T)$	(j) B^3

Solutions to Practice Problems

1.

(a) $D+E=\begin{bmatrix} 7 & 6 & 5 \\ -2 & 1 & 3 \\ 7 & 3 & 7 \end{bmatrix}$ (b) $D-E=\begin{bmatrix} -5 & 4 & -1 \\ 0 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix}$ (c) $5A=\begin{bmatrix} 15 & 0 \\ -5 & 10 \\ 5 & 5 \end{bmatrix}$,

(d) $-7C=\begin{bmatrix} -7 & -28 & -14 \\ -21 & -7 & -35 \end{bmatrix}$, (e) $2B-C$ not defined (f) $4E-2D=\begin{bmatrix} 22 & -6 & 8 \\ -2 & 4 & 6 \\ 10 & 0 & 4 \end{bmatrix}$,

(g) $-3(D+2E)=\begin{bmatrix} -39 & -21 & -24 \\ 9 & -6 & -15 \\ -33 & -12 & -30 \end{bmatrix}$, (h) $A-A=0,$ the zero matrix (i) $\mathrm{tr}(D)=1+0+4=5$,

(j) $\mathrm{tr}(D-3E)=-17-3-5=-25$, (k) $\mathrm{tr}(A)$ not defined (l) $\mathrm{tr}(B)=4+2=6$.

(m) $2AT+C=\begin{bmatrix} 7 & 2 & 4 \\ 3 & 5 & 7 \end{bmatrix}$, (n) $D^T-E^T=\begin{bmatrix} -5 & 0 & -1 \\ 4 & -1 & 1 \\ -1 & -1 & 1 \end{bmatrix}$ (o) $(D_E)^T=D^T-E^T$

(p) $B-B^T=\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

2.

(a) $AB=\begin{bmatrix} 12 & -3 \\ -4 & 5 \\ 4 & 1 \end{bmatrix}$, (b) BA not defined, (c) $3ED=\begin{bmatrix} 42 & 108 & 75 \\ 12 & -3 & 21 \\ 36 & 78 & 63 \end{bmatrix}$,

(d) $(AB)C=\begin{bmatrix} 3 & 45 & 9 \\ 11 & -11 & 17 \\ 7 & 17 & 13 \end{bmatrix}$, (e) $A(BC)=(AB)C$, (f) $CC^T=\begin{bmatrix} 21 & 17 \\ 17 & 35 \end{bmatrix}$,

(g) $(DA)^T\begin{bmatrix} 0 & -2 & 11 \\ 12 & 1 & 8 \end{bmatrix}$, (h) $(C^TB)A^T=\begin{bmatrix} 12 & 6 & 9 \\ 48 & -20 & 14 \\ 24 & 8 & 16 \end{bmatrix}$, (i) $\mathrm{tr}(DD^T)=30+2+29=61$,

(j) $B^3=\begin{bmatrix} 64 & -28 \\ 0 & 8 \end{bmatrix}$.

Linear Algebra III - Inverses and Determinants

We know that for every nonzero real number, x, there exists an inverse, $x^{-1} = 1/x$ such that the product $xx^{-1} = 1$. In the last section, we learned how to multiply matrices (and when this is possible). Here we ask whether inverses exist for matrices as well, and how we can compute them.

Introductory example

The multiplication of two square matrices can give the identity matrix:

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

But of course, not every (square) matrix has an inverse. The zero matrix, for example, is not invertible.

Definition: A square matrix A is *invertible* if there exists another square matrix B of the same dimension, such that

$$AB = I = BA.$$

We then write $A^{-1} = B$.

Note: If a matrix is not square, then it cannot be invertible.

How can we find out whether a given matrix has an inverse and what the inverse is? Let's go back to the example above with

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

We want to find a matrix

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix},$$

such that

$$AB = \left[A \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} \middle| A \begin{bmatrix} b_{12} \\ b_{22} \end{bmatrix} \right] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Hence, we have to solve the two systems

$$A \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad A \begin{bmatrix} b_{12} \\ b_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

We can do this simultaneously, using the three allowed elementary row operations of multiplication, addition and interchange.

$$\left[\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right] \xrightarrow{R1-2R2} \left[\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 0 & -1 & 1 & -2 \end{array}\right] \xrightarrow{R1+5R2} \left[\begin{array}{cc|cc} 2 & 0 & 6 & -10 \\ 0 & -1 & 1 & -2 \end{array}\right]$$

$$\xrightarrow{0.5*R1,-R2} \left[\begin{array}{cc|cc} 1 & 0 & 3 & -5 \\ 0 & 1 & -1 & 2 \end{array}\right]$$

Once we have reduced the left hand side to the identity matrix, we have the inverse of the original matrix on the right hand side.

Algorithm to compute an inverse matrix: Given a square matrix A, we take the identity matrix I and write the system

$$[\, A \mid I \,].$$

Then we use the three row operations to reduce the left hand side to the identity matrix. If this is possible, we end up with

$$[\, I \mid B \,] = [\, I \mid A^{-1} \,].$$

If we cannot reduce the left hand side to the identity matrix, then the original matrix A is not invertible.

Example 1

Find the inverse of $A = \begin{bmatrix} 3 & 7 \\ 1 & 3 \end{bmatrix}$. We write

$$\left[\begin{array}{cc|cc} 3 & 7 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 3 & 7 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 0 & -2 & 1 & -3 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{cc|cc} 1 & 3 & 0 & 1 \\ 0 & 1 & -1/2 & 3/2 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 0 & 3/2 & -7/2 \\ 0 & 1 & -1/2 & 3/2 \end{array}\right]$$

Hence, A is invertible and its inverse is $A^{-1} = \begin{bmatrix} 3/2 & -7/2 \\ -1/2 & 3/2 \end{bmatrix}$.

Example 2

Find the inverse of $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. We write

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

Because we have a row of zeros on the left hand side, we cannot transform the left hand side into the identity matrix by elementary row operations. Therefore, the matrix A is not invertible.

Example 3

Find the inverse of $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 0 & 2 & 4 \end{bmatrix}$. We write

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

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

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

Hence, the matrix A is invertible and the inverse is given by the right side of the last step above.

Example 4

Find the inverse of $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 0 & 2 & 2 \end{bmatrix}$. We write

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

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

Since the left half of the above matrix is the zero matrix, we cannot transform it into the identity matrix, using row operations. Therefore, A is NOT invertible.

When is a matrix invertible?

Sometimes we are not interested in the exact inverse of a matrix but only in the question of whether or not the matrix is invertible. Let's go back to the case of 2×2-matrices. When is the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

invertible? We take the potential inverse $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, and write out the equations $AB = I$ explicitly. We get the two systems of equations

$$\begin{cases} a_{11}b_{11} + a_{12}b_{21} &= 1 \\ a_{21}b_{11} + a_{22}b_{21} &= 0 \end{cases} \qquad \begin{cases} a_{11}b_{12} + a_{12}b_{22} &= 0 \\ a_{21}b_{12} + a_{22}b_{22} &= 0 \end{cases}$$

For the system on the left, we multiply the first row by $-a_{21}$ and the second by a_{11} and add the two in order to eliminate b_{11} from the second equation.

$$\begin{cases} a_{11}b_{11} + a_{12}b_{21} &= 1 \\ -a_{21}a_{12}b_{21} + a_{11}a_{22}b_{21} &= -a_{21} \end{cases}$$

If the expression $a_{11}a_{22} - a_{12}a_{21}$ is nonzero, we can solve for b_{21} as

$$b_{21} = \frac{-a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

The system on the right for b_{12}, b_{22} is similar: If $a_{11}a_{22} - a_{12}a_{21} \neq 0$ then we can solve for

$$b_{22} = \frac{a_{11}}{a_{11}a_{22} - a_{12}a_{21}}.$$

Since the expression $a_{11}a_{22} - a_{12}a_{21}$ is so important, we give it a special name.

Definition and Result: For a 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

the *determinant* of A is given by

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}.$$

If $\det(A) \neq 0$ then A is invertible and

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

If $\det(A) = 0$ then A is not invertible.

Example 1, revisited

The determinant of the matrix $A = \begin{bmatrix} 3 & 7 \\ 1 & 3 \end{bmatrix}$ is

$$\det(A) = 3 \cdot 3 - 1 \cdot 7 = 9 - 7 = 2.$$

The inverse according to our new formula is therefore $A^{-1} = \frac{1}{2} \begin{bmatrix} 3 & -7 \\ -1 & 3 \end{bmatrix}$, just as we had computed earlier.

Example 2, revisited

The determinant of the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ is

$$\det(A) = 1 \cdot 4 - 2 \cdot 2 = 4 - 4 = 0.$$

Hence, the matrix is not invertible, as we have already seen before.

Our theory so far only applies to matrices of size 2×2, therefore we cannot revisit the examples 3 and 4. We do that later. For now, we look at some more examples of matrices of size 2×2.

Example 5

The determinant of the matrix $A = \begin{bmatrix} 3 & 7 \\ 1 & 2 \end{bmatrix}$ is

$$\det(A) = 3 \cdot 2 - 1 \cdot 7 = 6 - 7 = -1.$$

Therefore, the inverse of A exists and is given by $A^{-1} = -\begin{bmatrix} 2 & -7 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} -2 & 7 \\ 1 & -3 \end{bmatrix}$.

Example 6

The determinant of the matrix $A = \begin{bmatrix} 3 & -1 \\ 2 & 1/2 \end{bmatrix}$ is

$$\det(A) = 3 \cdot 0.5 - 2 \cdot (-1) = 3/2 + 2 = 7/2.$$

Therefore, the inverse of A exists and is given by $A^{-1} = \frac{2}{7}\begin{bmatrix} 1/2 & 1 \\ -2 & 3 \end{bmatrix}$.

Example 7

The determinant of the matrix $A = \begin{bmatrix} -7 & 2 \\ 4 & -8/7 \end{bmatrix}$ is

$$\det(A) = (-7) \cdot (-8/7) - 2 \cdot 4 = 8 - 8 = 0.$$

Hence, the matrix is not invertible.

Determinants for matrices of size 3×3

Determinants can be defined for square matrices of all sizes, and it is always true that if $\det(A) \neq 0$ then the matrix A is invertible. We will only consider the case of 3×3-matrices here, since it has a fairly simple form. Determinants of bigger matrices can be computed, but it takes time.

The determinant of the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

is given by

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}.$$

Unfortunately, there is no really simple formula for the inverse of a 3×3 matrix (or larger) analogous to the 2×2 case. To find the inverse, we still have to use the row-reduction algorithm.

This determinant formula can be remembered more easily by attaching the first two columns of the matrix A as columns 4 and 5 of a larger matrix and then taking products along the diagonal down with plus signs and products along the diagonal up with minus signs, i.e.,

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix}$$

Example 3, revisited

The determinant of $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 0 & 2 & 4 \end{bmatrix}$ is

$$\det(A) = 1 \cdot 3 \cdot 4 + 2 \cdot 4 \cdot 0 + 3 \cdot 1 \cdot 2 - 0 \cdot 3 \cdot 3 - 2 \cdot 4 \cdot 1 - 4 \cdot 1 \cdot 2 = 18 - 14 = 4.$$

Hence, the matrix is invertible, as we already know.

Example 4, revisited

The determinant of $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 0 & 2 & 2 \end{bmatrix}$ is

$$\det(A) = 1 \cdot 3 \cdot 2 + 2 \cdot 4 \cdot 0 + 3 \cdot 1 \cdot 2 - 0 \cdot 3 \cdot 3 - 2 \cdot 4 \cdot 1 - 2 \cdot 1 \cdot 2 = 12 - 12 = 0.$$

Hence, the matrix is not invertible, confirming our previous result.

Important Observation: If A is an invertible square matrix then the unique solution of the system of linear equations

$$Ax = b$$

is given by

$$x = A^{-1}b.$$

Example 1, revisited

The solution of the system

$$\begin{cases} 3x_1 + 7x_2 & = 4 \\ x_1 + 3x_2 & = 6 \end{cases}$$

is given by

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 3 & -7 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} -15 \\ 7 \end{bmatrix}$$

Practice Problems

1. Compute the determinant and find the inverse if it exists.

$$\text{(a)}\quad A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \qquad \text{(b)}\quad A = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\text{(c)}\quad A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \qquad \text{(d)}\quad A = \begin{bmatrix} 3 & -2 \\ -6 & 4 \end{bmatrix}$$

$$\text{(e)}\quad A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix} \qquad \text{(f)}\quad A = \begin{bmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{bmatrix}$$

2. Solve the linear system of equations

For each of the matrices above solve the system $Ax = b$ where $b = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ in cases (a)-(d) and $b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in cases (e) and (f).

Solutions to Practice Problems

1.

(a) $A^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}$ (b) $A^{-1} = \begin{bmatrix} 1/4 & 0 \\ 0 & -1 \end{bmatrix}$ (c) $A^{-1} = \begin{bmatrix} 1 & -2/3 \\ 0 & 1/3 \end{bmatrix}$

(d) and (f) are not invertible

$$(e) \quad A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$

2.

(a) $x = \frac{1}{3}\begin{bmatrix} -1 \\ -4 \end{bmatrix}$, (b) $x = \begin{bmatrix} 1/4 \\ 2 \end{bmatrix}$, (c) $x = \begin{bmatrix} 7/3 \\ -2/3 \end{bmatrix}$,

In case (d) the solution is $\left\{ \begin{bmatrix} (1+2t)/3 \\ t \end{bmatrix} : t \in \mathbb{R} \right\}$.

$$(e) \quad x = \begin{bmatrix} -15 \\ 5 \\ 2 \end{bmatrix}.$$

In case (f), there is no solution.
NOTE: Even if the matrix is not invertible, there can still be a solution (see (d)), but it is not unique in that case.

Linear Algebra IV: Eigenvalues and Eigenvectors

Observation and introductory example

Consider the matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ so that $A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$. For example

$$A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad A \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

On the other hand, we also have

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = - \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

In these two cases, the vectors $A \begin{bmatrix} x \\ y \end{bmatrix}$ are simply multiples of the vectors $\begin{bmatrix} x \\ y \end{bmatrix}$, whereas in the examples above, there was no obvious relationship between $A \begin{bmatrix} x \\ y \end{bmatrix}$ and $\begin{bmatrix} x \\ y \end{bmatrix}$. The goal of this section is to find and study these special cases where a vector multiplied by a matrix is simply a multiple of the vector.

Definition: Let A be a square matrix. A vector v that is not the zero vector and a number λ are called *eigenvector* and *eigenvalue,* respectively if

$$Av = \lambda v.$$

Example 1

The vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of the matrix $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$ corresponding to the eigenvalue $\lambda = 3$, since

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

Example 2

The vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector of the matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ corresponding to the eigenvalue $\lambda = 0$, since

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

NOTE: $\lambda = 0$ is allowed, but the eigenvector cannot be the zero vector.

Example 3

The vector $\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$ corresponding to the eigenvalue $\lambda = 1$ since

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

How to compute eigenvalues and eigenvectors?

We are looking for a nonzero solution of the equation $Av = \lambda v$, which is the same as saying that we are looking for a nonzero solution of the homogeneous equation

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

If the matrix $(A - \lambda I)$ is invertible, then the only solution to the equation is the zero solution, which we do not want. Hence, we need to make sure that the matrix $(A - \lambda I)$ is *not* invertible. In the previous section we saw that a matrix is not invertible precisely if its determinant is zero. Therefore, we need to find values of λ that make the determinant of $(A - \lambda I)$ to zero.

Result: The number λ is an eigenvalue of the square matrix A if and only if it satisfies the equation

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

If λ is an eigenvalue of the square matrix A then we find the corresponding eigenvector(s) by solving the linear system of equations

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

Note: Any scalar multiple of an eigenvector is again an eigenvector.

Example 1

Find the eigenvalues and eigenvectors of $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

First we form the matrix

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

Then we calculate its determinant as

$$\det \begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 - 1.$$

Setting the determinant to zero, we get the two eigenvalues $\lambda = \pm 1$. For $\lambda_1 = 1$ we solve the system

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

The second row is satisfied, the first row reads $-x_1 + x_2 = 0$. Hence, $x_2 = t$ is a free variable, and $x_1 = t$ is the resulting solution, so that $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is the eigenvector corresponding to $\lambda_1 = 1$. Note that any multiple of v is also an eigenvector to the same eigenvalue.

For $\lambda_2 = -1$ we solve the system

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

The second row is satisfied, the first row reads $x_1 + x_2 = 0$. Hence, $x_2 = t$ is a free variable, and $x_1 = -t$ is the resulting solution, so that $v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is the eigenvector corresponding to $\lambda_2 = -1$. Note that any multiple of v is also an eigenvector to the same eigenvalue.

Example 2

Find the eigenvalues and eigenvectors of $A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

First we form the matrix

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

Then we calculate its determinant as

$$\det(A - \lambda I) = -\lambda(2-\lambda)(1-\lambda) - (2-\lambda) = (2-\lambda)(\lambda^2 - 1).$$

Hence the eigenvalues are $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = 2$. For $\lambda_1 = 1$ we solve the system (writing only the coefficient matrix and suppressing the right hand column of zeros)

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

We have $x_3 = t$ as the free variable and solve for $x_2 = -2t$, and $x_1 = t$ to get the eigenvector $v_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$.

For $\lambda_2 = -1$ we solve the system (writing only the coefficient matrix and suppressing the right hand column of zeros)

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

The second row gives the equation $3x_2 = 0$, which means $x_2 = 0$. We have $x_3 = t$ the free variable and solve for $x_1 = -t$ to get the eigenvector $v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

For $\lambda_3 = 2$ we solve the system (writing only the coefficient matrix and suppressing the right hand column of zeros)

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

The second row gives the equation $0x_2 = 0$, which means $x_2 = t$ is free. The first equation says $-2x_1 + x_3 = 0$, while the third equation says $x_1 - 2x_3 = 0$. The only way to make both of them true is to set $x_1 = x_3 = 0$. Hence we get the eigenvector $v_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

Example 3

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

First we form the matrix

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

Then we calculate its determinant as

$$\det(A - \lambda I) = -\lambda(2-\lambda)(3-\lambda) - (2-\lambda)(-2) = (2-\lambda)(\lambda-2)(\lambda-1).$$

Hence the eigenvalues are $\lambda_1 = 2, \lambda_2 = 2, \lambda_3 = 1$. For the double eigenvalue $\lambda_1 = \lambda_2 = 2$ we solve the system

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

The first equation is $x_1 + x_3 = 0$. The variables $x_2 = s$ and $x_3 = t$ are free. The solution is $x_1 = -x_3 = -t$. Hence, we get essentially two eigenvectors. For one, we set $t = 0$ and $s = 1$. For the other, we set $t = 1$ and $s = 0$. Then

$$v_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

For $\lambda_3 = 1$, we solve the system (suppressing the column of zeros)

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

In this case, we get $x_3 = t$ to be the free variable. The resulting solution then is $x_2 = x_3 = t$ and $x_1 = -2x_3 = -2t$. The eigenvector therefore is $v_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$.

Example 4

Find the eigenvalues and eigenvectors of $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.
First we form the matrix

$$A - \lambda I = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} -\lambda & 0 \\ 0 & -\lambda \end{bmatrix} = \begin{bmatrix} 1-\lambda & -11 \\ 1 & 1-\lambda \end{bmatrix}.$$

Then we calculate its determinant as

$$\det \begin{bmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{bmatrix} = (1-\lambda)^2 + 1 = \lambda^2 - 2\lambda + 2.$$

Setting the determinant to zero, we get

$$\lambda = \frac{1}{2}\left(2 \pm \sqrt{4-8}\right) = 1 \pm \sqrt{-1} = 1 \pm \mathrm{i}.$$

Hence, we get complex numbers as eigenvalues. Just to be sure, we check that $\lambda_1 = 1 + \mathrm{i}$ solves the original equation:

$$(1+\mathrm{i})^2 - 2(1+\mathrm{i}) + 2 = 1 + 2\mathrm{i} - 1 - 2 - 2\mathrm{i} + 2 = 0.$$

We can compute the corresponding eigenvectors in the same way as above, but the calculation involves complex numbers. For $\lambda_1 = 1 + \mathrm{i}$ we get

$$\begin{bmatrix} -\mathrm{i} & -1 \\ 1 & -\mathrm{i} \end{bmatrix} \longrightarrow (-\mathrm{i} * R1 + R2) \longrightarrow \begin{bmatrix} -\mathrm{i} & -1 \\ 0 & 0 \end{bmatrix}.$$

Then $x_2 = t$ is a free variable and $x_1 = \mathrm{i}t$, so that the eigenvector is $v = \begin{bmatrix} \mathrm{i} \\ 1 \end{bmatrix}$.
For $\lambda_2 = 1 - \mathrm{i}$ we get

$$\begin{bmatrix} \mathrm{i} & -1 \\ 1 & \mathrm{i} \end{bmatrix} \longrightarrow (\mathrm{i} * R1 + R2) \longrightarrow \begin{bmatrix} \mathrm{i} & -1 \\ 0 & 0 \end{bmatrix}.$$

Then $x_2 = t$ is a free variable and $x_1 = -\mathrm{i}t$, so that the eigenvector is $v = \begin{bmatrix} -\mathrm{i} \\ 1 \end{bmatrix}$.

Practice Problems

Find the eigenvalues and eigenvectors of the following matrices.

$$(a) \quad A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 3 & 1 \\ 2 & 0 & 0 \end{bmatrix}, \qquad (b) \quad A = \begin{bmatrix} -5 & 2 & -3 \\ 0 & 1 & -2 \\ 0 & 1 & 4 \end{bmatrix},$$

$$(c) \quad A = \begin{bmatrix} 3 & 2 & -6 \\ 0 & -2 & 1 \\ 0 & 4 & 1 \end{bmatrix}, \qquad (d) \quad A = \begin{bmatrix} 4 & 2 & -5 \\ 0 & 1 & 2 \\ 0 & 6 & -3 \end{bmatrix},$$

$$(e) \quad A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}, \qquad (f) \quad A = \begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}, \qquad (g) \quad A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

$$(h) \quad A = \begin{bmatrix} -1 & 0 & 1 \\ -3 & 4 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \qquad (i) \quad A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}.$$

Solutions

For problem (a):

$\det(\lambda I - A) = \det\begin{bmatrix} \lambda & 0 & -2 \\ -1 & \lambda-3 & -1 \\ -2 & 0 & \lambda \end{bmatrix} = \lambda^2(\lambda-3) - 4(\lambda-3) = (\lambda-3)(\lambda^2-4) = 0$

So the eigenvalues are $\lambda_1 = 3$, $\lambda_2 = 2$ and $\lambda_3 = -2$. To find the eigenvectors corresponding to an eigenvalue λ we have to solve the system $(\lambda I - A)x = 0$ where $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

For $\lambda_1 = 3$ we have to solve

$$3x_1 - 2x_3 = 0$$
$$-x_1 - x_3 = 0$$
$$-2x_1 + 3x_3 = 0$$

$\left[\begin{array}{ccc|c} 3 & 0 & -2 & 0 \\ -1 & 0 & -1 & 0 \\ -2 & 0 & 3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 3 & 0 & -2 & 0 \\ 0 & 0 & -7 & 0 \\ 0 & 0 & 5 & 0 \end{array}\right]$ Hence $x_3 = 0$ and from the first equation also $x_1 = 0$. The eigenvectors are $\begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} = t\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$,with $t \in R$

For $\lambda_2 = 2$

$$2x_1 - 2x_3 = 0$$
$$-x_1 - x_2 - x_3 = 0$$
$$-2x_1 + 2x_3 = 0$$

$\left[\begin{array}{ccc|c} 2 & 0 & -2 & 0 \\ -1 & -1 & -1 & 0 \\ -2 & 0 & 2 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 2 & 0 & -2 & 0 \\ 0 & -2 & -4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$ Then $x_3 = t$ is free, $x_2 = -2t$ and $x_1 = t$. The eigenvectors are $\begin{bmatrix} t \\ -2t \\ t \end{bmatrix} = t\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ with $t \in R$.

For $\lambda_3 = -2$

$$-2x_1 - 2x_3 = 0$$
$$-x_1 - 5x_2 - x_3 = 0$$
$$-2x_1 - 2x_3 = 0$$

$\left[\begin{array}{ccc|c} -2 & 0 & -2 & 0 \\ -1 & -5 & -1 & 0 \\ -2 & 0 & -2 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} -2 & 0 & -2 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$. Then $x_3 = t$ is free, $x_1 = -t$ and $x_2 = 0$.

The eigenvectors are $\begin{bmatrix} -t \\ 0 \\ t \end{bmatrix} = t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ with $t \in R$.

For problem (b):
We have

$$\det(A - \lambda I) = \det \begin{bmatrix} -5-\lambda & 2 & -3 \\ 0 & 1-\lambda & -2 \\ 0 & 1 & 4-\lambda \end{bmatrix}$$
$$= (-5-\lambda)(1-\lambda)(4-\lambda) - (1)(-2)(-5-\lambda) = (-5-\lambda)\left[(1-\lambda)(4-\lambda)+2\right]$$
$$= -(5+\lambda)(\lambda^2 - 5\lambda + 6) = -(5+\lambda)(\lambda-2)(\lambda-3)$$

Thus, the eigenvalues of A are $\lambda_1 = -5$, $\lambda_2 = 2$ and $\lambda_3 = 3$.

The eigenvector v_1 associated with the eigenvalue $\lambda_1 = -5$ is the solution of the system

$$(A - \lambda_1 I)v_1 = \begin{bmatrix} 0 & 2 & -3 \\ 0 & 6 & -2 \\ 0 & 1 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The augmented matrix is

$$\begin{bmatrix} 0 & 2 & -3 & | & 0 \\ 0 & 6 & -2 & | & 0 \\ 0 & 1 & 9 & | & 0 \end{bmatrix}$$

$R_1 - 2R_3 \rightarrow R_1$ and $R_2 - 6R_3 \rightarrow R_2$ gives

$$\begin{bmatrix} 0 & 0 & -21 & | & 0 \\ 0 & 0 & -56 & | & 0 \\ 0 & 1 & 9 & | & 0 \end{bmatrix}$$

$(-1/21)R_1 \rightarrow R_1$ gives

$$\begin{bmatrix} 0 & 0 & 1 & | & 0 \\ 0 & 0 & -56 & | & 0 \\ 0 & 1 & 9 & | & 0 \end{bmatrix}$$

$R_2 + 56R_1 \rightarrow R_2$ and $R_3 - 9R_1 \rightarrow R_3$ gives

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

Finally, $R_3 \leftrightarrow R_1$ and $R_2 \leftrightarrow R_3$ gives

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

The solution is $y = z = 0$ and $x = t$. Thus, the eigenvectors are $v_1 = s \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ where t is any real number except 0.

Next, for $\lambda_2 = 2$ we have

$$(A - \lambda_2 I)v_2 = \begin{pmatrix} -7 & 2 & -3 \\ 0 & -1 & -2 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Thus

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

$R_3 + R_2 \to R_3$ and $R_1 + 2R_2 \to R_1$ gives

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

$(-1/7)R_1 \to R_1$ and $-R_2 \to R_2$ gives

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

The solution is thus $x = -z$ and $y = -2z$. With $z = s$, we get $v_2 = s \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$ where s is any real number except 0.

Finally, for $\lambda_3 = 3$ we have

$$(A - \lambda_3 I)v_3 = \begin{pmatrix} -8 & 2 & -3 \\ 0 & -2 & -2 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The augmented matrix is

$$\begin{bmatrix} -8 & 2 & -3 & | & 0 \\ 0 & -2 & -2 & | & 0 \\ 0 & 1 & 1 & | & 0 \end{bmatrix}$$

$R_2 + 2R_3 \to R_2$ and $R_1 - 2R_3 \to R_1$ gives

$$\left[\begin{array}{ccc|c} -8 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right]$$

$(-1/8)R_1 \to R_1$ and $R_2 \leftrightarrow R_3$ gives

$$\left[\begin{array}{ccc|c} 1 & 0 & 5/8 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The solution is $x = -5z/8$ and $y = -z$. With $z = r$, we get $v_3 = r \begin{bmatrix} -5/8 \\ -1 \\ 1 \end{bmatrix}$ where r is any real number except 0.

For problem (c):

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 3-\lambda & 2 & -6 \\ 0 & -2-\lambda & 1 \\ 0 & 4 & 1-\lambda \end{bmatrix} \\ &= (3-\lambda)(-2-\lambda)(1-\lambda) + 0 + 0 - 0 - 0 - (3-\lambda)(4) \\ &= (3-\lambda)[-2-\lambda+\lambda^2-4] \\ &= (3-\lambda)[\lambda^2-\lambda-6] \\ &= (3-\lambda)(\lambda+3)(\lambda-2) \end{aligned}$$

Thus $\lambda_1 = 3, \lambda_2 = -3, \lambda_3 = 2$.

Eigenvectors are

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 4/3 \\ -1 \\ 1 \end{bmatrix}, \qquad v_3 = \begin{bmatrix} 11/2 \\ 1/4 \\ 1 \end{bmatrix}.$$

For problem (d):

$$\begin{aligned} \det(A - \lambda I) = \det &\begin{pmatrix} 4-\lambda & 2 & -5 \\ 0 & 1-\lambda & 2 \\ 0 & 6 & -3-\lambda \end{pmatrix} \\ &= (4-\lambda)\,[(1-\lambda)(-3-\lambda) - 12] \\ &= (4-\lambda)(\lambda^2 + 2\,\lambda - 15) = (4-\lambda)(\lambda-3)(\lambda+5) \end{aligned}$$

Thus, the eigenvalues of A are $\lambda_1 = -5$, $\lambda_2 = 3$ and $\lambda_3 = 4$.

Eigenvectors are

$$v_1 = \begin{bmatrix} 17/27 \\ -1/3 \\ 1 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \qquad v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

For problem (e):
Eigenvalues are

$$\lambda_1 = 5, \lambda_2 = -1.$$

Eigenvectors are

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

For problem (f):
Eigenvalues are

$$\lambda_1 = 5, \lambda_2 = -5.$$

Eigenvectors are

$$v_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

For problem (g):
Eigenvalues are

$$\lambda_1 = -\mathrm{i}, \lambda_2 = \mathrm{i}.$$

Eigenvectors are

$$v_1 = \begin{bmatrix} \mathrm{i} \\ 1 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} -\mathrm{i} \\ 1 \end{bmatrix}.$$

For problem (h):
Eigenvalues are

$$\lambda_1 = 2, \lambda_2 = -1, \lambda_3 = 4.$$

Eigenvectors are

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 5/3 \\ 1 \\ 0 \end{bmatrix}, \qquad v_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

For problem (i):
Eigenvalues are

$$\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 2.$$

Eigenvectors are

$$v_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad v_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

LinearAlgebra V - Markov chains

Example 1

We consider two ponds, a smaller and a larger one, where a population of a total of 300 ducks lives. We observe that a duck on the smaller pond switches to the larger pond the next day with a probability of 40% and stays on the smaller pond with a probability of 60%. A duck on the large pond stays with a probability of 80% and moves on to the smaller pond with a probability of 20%. Suppose that on the first day of observation, there are 150 ducks on each pond. How many ducks are there on each pond the next day, the third day, in the long run?

Let's denote x_n as the number of ducks on the small pond and $y^{(n)}$ the number on the large pond on day n. Then at time $t = 0$ we have $x^{(0)} = 150, y^{(0)} = 150$. To compute the number of ducks on day 1, we simply add the ducks that stay to the ones that newly arrive, i.e.,

$$x^{(1)} = 0.6x_0 + 0.2y_0 = 90 + 30 = 120, \quad y^{(1)} = 0.4x_0 + 0.8y_0 = 60 + 120 = 180.$$

We can write these equations in matrix form as

$$\begin{bmatrix} x^{(1)} \\ y^{(1)} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} \begin{bmatrix} x^{(0)} \\ y^{(0)} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} \begin{bmatrix} 150 \\ 150 \end{bmatrix} = \begin{bmatrix} 120 \\ 180 \end{bmatrix}.$$

Now it is easy to continue. For the numbers on day 2, we simply apply the matrix again

$$\begin{bmatrix} x^{(2)} \\ y^{(2)} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ y^{(1)} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} \begin{bmatrix} 120 \\ 180 \end{bmatrix} = \begin{bmatrix} 108 \\ 192 \end{bmatrix}.$$

We can continue from here. But we cal also ask if there will be a state where the arrivals equal the departures on both ponds so that the number of ducks stays the same day after day. If the numbers stay the same, say x^*, y^*, then we can find them from the equation

$$\begin{bmatrix} x^* \\ y^* \end{bmatrix} = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} \begin{bmatrix} x^* \\ y^* \end{bmatrix},$$

or

$$\left(\begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x^* \\ y^* \end{bmatrix} = \begin{bmatrix} -0.4 & 0.2 \\ 0.4 & -0.2 \end{bmatrix} \begin{bmatrix} x^* \\ y^* \end{bmatrix} = 0.$$

We see that this is the same as finding an eigenvector of the matrix with eigenvalue one. The solution is

$$\begin{bmatrix} x^* \\ y^* \end{bmatrix} = r \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

for any real number $r \neq 0$. But we know that the total number of ducks is 300, hence $x^* + y^* = (1 + 2)r = 300$, which makes $r = 100, x^* = 100, y^* = 200$.

There are two alternative approaches and interpretations of the same problem.

1. If we do not know the total number of ducks to begin with, but only the percentage of ducks on each pond, we can do the same calculation with percentages. For example, initially, both ponds have the same number of ducks, i.e., $x^{(0)} = y^{(0)} = 50\% = 0.5$. Then

$$\begin{bmatrix} x^{(1)} \\ y^{(1)} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} \begin{bmatrix} x^{(0)} \\ y^{(0)} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}.$$

   The final distribution of ducks in the long run would then be $x^* + y^* = 3r = 1$, and hence $r = 1/3$ or $x^* = 1/3$ and $y^* = 2/3$.

2. If we don't look at the whole population but only at a single duck, we could be interested in the probability that the duck is on the small pond versus on the big pond. Then x_n and y_n denote the probability that the duck is on the small or large pond on day n. We could assume that the duck is initially on either pond with equal probability as to get $x_0 = y_0 = 50\% = 0.5$. Then the rest is as above, and we compute that in the long run, the duck is twice as likely to be on the big pond than on the small pond.

A *Markov chain* is a process where the state of a system at a given time can be predicted by knowing just the state of the system at the previous time. A Markov chain with k states is described by the *transition matrix*

$$P = [p_{ij}], \qquad 1 \le i, j \le k,$$

where p_{ij} is the probability that the system will be in state i after is was in state j at the previous time. For each j, we have

$$p_{1j} + p_{2j} + \cdots + p_{kj} = 1,$$

since the system has to go to *some* state. A *state vector* is a vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}$$

where x_i is the probability that the system is in state i. In particular

$$x_1 + x_2 + \cdots + x_k = 1.$$

If $x^{(n)}$ is the state vector at time n, then the state vector at the next time, $x^{(n+1)}$ is given by

$$x^{(n+1)} = Px^{(n)}.$$

The steady state distribution is given by the equation

$$x^* = Px^*.$$

Following up on the previous example: a duck can be in one of two states: on the small pond or on the big pond. If it is on the small pond then the transition probability to the big pond is 0.4 whereas the transition probability to the small pond (i.e., the duck remains) is 0.6. Vice versa, a duck on the big pond has a transition probability of 0.2 to the small pond and of 0.8 to the big pond (i.e., remaining). The transition matrix is therefore

$$P = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix}.$$

Example 2

Consider the two soft drinks Choke and Popsi. Market research found that in a given year, 70% of Choke drinkers switch to Popsi and 30% stick with Choke, while 50% of Popsi drinkers switch to Choke and 50% keep drinking Popsi. Assuming that that a system behaves like a finite Markov chain, find the transition matrix P of probabilities for switching from one brand to another, and determine the percentage of Choke and Popsi drinkers in the long run.

Solution: Denote by x the fraction of Choke drinkers and by y the fraction of Popsi drinkers. Then the transition matrix is given by

$$P = \begin{bmatrix} 0.3 & 0.5 \\ 0.7 & 0.5 \end{bmatrix}.$$

The steady state satisfies

$$\left(\begin{bmatrix} 0.3 & 0.5 \\ 0.7 & 0.5 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -0.7 & 0.5 \\ 0.7 & -0.5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0.$$

The solution is $x = 5/7t, y = t$. Since the sum has to equal one, $5/7t + t = 12/7t = 1$, we get $t = 7/12$ and $x = 5/12, y = 7/12$. Hence, in the long run, approximately 42% drink Choke and 58% drink Popsi.

Example 3: Forest succession (Horn, 1980)

Suppose that a forest contains three types of trees: Gray Birch (GB), Red Maple (RM) and Beech (BE). By counting saplings under each tree, we can guess the probability that a certain tree is replaced by another tree. For example, under a GB, we find 2 GB, 15 RM and 3 BE saplings for a total of 20. We conclude that the probability of the GB to be replaced by GB is 0.1, to be replaced by RM is 0.75, and to be replaced by BE is 0.15. Under a RM, we find 2 GB, 3RM, and 5 BE saplings for a total of 10. Hence, the probability of RM to be replaced by GB is 0.1, to be replaced by RM is 0.3 and to be replaced by BE is 0.5.

We choose the state vector to give the probability to be GB, RM, or BE at any given location, i.e.,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \text{Prob. of GB} \\ \text{Prob. of RM} \\ \text{Prob. of BE} \end{bmatrix}.$$

Then the transition matrix is

$$P = \begin{bmatrix} 0.1 & 0.2 & 0 \\ 0.75 & 0.3 & 0.25 \\ 0.15 & 0.5 & 0.75 \end{bmatrix} = \frac{1}{20}\begin{bmatrix} 2 & 4 & 0 \\ 15 & 6 & 5 \\ 3 & 10 & 15 \end{bmatrix}.$$

If all tree species are equally present in the first year, then the distribution in the second year is

$$\begin{bmatrix} x^{(1)} \\ y^{(1)} \\ z^{(1)} \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 & 0 \\ 0.75 & 0.3 & 0.25 \\ 0.15 & 0.5 & 0.75 \end{bmatrix}\begin{bmatrix} x^{(0)} \\ y^{(0)} \\ z^{(0)} \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 & 0 \\ 0.75 & 0.3 & 0.25 \\ 0.15 & 0.5 & 0.75 \end{bmatrix}\begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} =$$

$$\frac{1}{60}\begin{bmatrix} 2 & 4 & 0 \\ 15 & 6 & 5 \\ 3 & 10 & 15 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{60}\begin{bmatrix} 6 \\ 26 \\ 28 \end{bmatrix} = \begin{bmatrix} 1/10 \\ 13/30 \\ 14/30 \end{bmatrix}$$

The steady state distribution is given by the solution of

$$\frac{1}{20}\begin{bmatrix} 2 & 4 & 0 \\ 15 & 6 & 5 \\ 3 & 10 & 15 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Equivalently, we solve

$$\frac{1}{20}\begin{bmatrix} 2-20 & 4 & 0 \\ 15 & 6-20 & 5 \\ 3 & 10 & 15-20 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0.$$

Row reduction gives (we drop the right hand column of zeros)

$$\begin{bmatrix} -18 & 4 & 0 \\ 15 & -14 & 5 \\ 3 & 10 & -5 \end{bmatrix} \longrightarrow \begin{bmatrix} 3 & 10 & -5 \\ 0 & 64 & -30 \\ 0 & -64 & 30 \end{bmatrix} \longrightarrow \begin{bmatrix} 3 & 10 & -5 \\ 0 & 32 & -15 \\ 0 & 0 & 0 \end{bmatrix}$$

The second equation is

$$32y - 15z = 0.$$

We set $z = t$, the free variable and solve for $y = 15t/32$. The first equation is

$$3x + 10y - 5z = 0, \text{ or } 3x + \frac{150t}{32} - 5t = 0,$$

which we solve to get

$$x = \frac{5}{3}t - \frac{1}{3}\frac{150t}{32} = \frac{10t}{96}.$$

Hence the solution is

$$\begin{bmatrix} x^* \\ y^* \\ z^* \end{bmatrix} = t\begin{bmatrix} 10/96 \\ 15/32 \\ 1 \end{bmatrix}.$$

With the condition the $x + y + z = 1$, we get

$$\left(\frac{10}{96} + \frac{15}{32} + 1\right) t = \frac{151}{96} t = 1,$$

so that $t = 96/151$ and

$$\begin{bmatrix} x^* \\ y^* \\ z^* \end{bmatrix} = \begin{bmatrix} 10/151 \\ 45/151 \\ 96/151 \end{bmatrix} \approx \begin{bmatrix} 6\% \\ 30\% \\ 64\% \end{bmatrix}.$$

Practice Problems

1. There are two weekly Farmers Markets in Ottawa: one on Bank and Heron (B), the other in the Glebe (G). People who shop at B one week return there the next week with a probability of 80% and go to G in 20% of the cases. People who go to G return there with a probability of 70% and shop at B the next week with probability 30%. Write down the matrix that describes this Markov chain. (You may assume that the total number of people who get groceries at a Farmers Market is constant.) Show that the matrix has one eigenvalue equal to one. Find the proportion of people who shop at B and G in the long run, respectively.

2. In a lab experiment, there are three chambers: A, B, C, with connecting corridors between A and B, as well as B and C, but not between A and C. A rat is placed into chamber A, and its location is recorded every hour. The rat stays within a chamber with probability 0.5 and leaves the chamber with probability 0.5. If it leaves chamber A or C, it goes to B (it has to, there are no other corridors). If it leaves chamber B, it chooses to go to A or C with equal probability. Write down the matrix that describes this Markov chain. What is the probability to find the rat in chamber A, B, C after the first hour? After the second hour? In the long run, what is the probability to find the rat in A, B, or C, respectively?

Solutions to Practice Problems

1. We denote the percent of people at B by x and the percent of people at G by y, and consider the vector $\begin{bmatrix} x \\ y \end{bmatrix}$. Then the matrix has the form

$$P = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}.$$

Note that the entries in a column add up to one. The matrix has eigenvalue 1, if the system

$$P \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

has a nontrivial solution. We solve

$$(P - I) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -0.2 & 0.3 \\ 0.2 & -0.3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0.$$

The solution is

$$y = t, \quad \text{free variable} \quad x = \frac{0.3}{0.2} t = \frac{3}{2} t.$$

Since we are dealing with proportions, we need $x + y = 1$, which gives $1.5t + t = 2.5t = 1$ and hence $t = 2/5$. In the long run the percentage of people shopping at B is $x = 3/5 = 60\%$ and at G is $y = 2/5 = 40\%$.

2. We denote by x, y, z the probability that the rat is in chamber A, B, C, respectively, and write the vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$. Then the matrix is

$$P = \begin{bmatrix} 0.5 & 0.25 & 0 \\ 0.5 & 0.5 & 0.5 \\ 0 & 0.25 & 0.5 \end{bmatrix}.$$

Note again that the columns add up to one. If the rat starts at A, then the corresponding vector is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. After one hour, the probabilities are

$$P \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.25 & 0 \\ 0.5 & 0.5 & 0.5 \\ 0 & 0.25 & 0.5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix}.$$

Hence, the rat is in A or B with equal probability, but not in C. After the second hour the probabilities are

$$\begin{bmatrix} 0.5 & 0.25 & 0 \\ 0.5 & 0.5 & 0.5 \\ 0 & 0.25 & 0.5 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 3/8 \\ 1/2 \\ 1/8 \end{bmatrix}.$$

To see what is happening in the long run, we have to solve the system

$$P \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

or

$$\begin{bmatrix} -0.5 & 0.25 & 0 \\ 0.5 & -0.5 & 0.5 \\ 0 & 0.25 & -0.5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0.$$

The solution is $t \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$. Together with the condition that $x + y + z = 1$, we get $t = 1/4$ and the respective probabilities of $x = 0.25, y = 0.5, z = 0.25$.