Math 133A, Weeks 11 & 12: Review of Linear Algebra

Section 1: Linear Algebra

We will soon begin dealing with interdependent systems of first-order differential equations. Before we can begin such consideration, however, it is important to consider the study of linear systems of equations in general, which is the study of **linear algebra**. In particular, we will need to define a few new of objects, called **matrices** and **vectors**, and describe how they interact. Buried inside these objects will be the individual quantities (e.g. variables, functions, derivatives, etc.) which we have been interested in up to this point in the course.

The advantage of briefly introducing linear algebra now is two-fold:

1. It will allow us to condense expressions in a way which will make a first-order system of differential equations analogous to the first-order differential equations studied to date.

2. Theoretical results from linear algebra will be *essential* in deriving solutions to all but the simplest of such systems.

Of course, linear algebra is not a primary topic of **Math 133A**. We will introduce only the topics which are necessary for our study, and for theoretical and computational simplicity, will limit ourselves to two-dimensional systems. (The interested student is encouraged to take **Math 129A** for a fuller introduction to the area.)

Section 2: Matrices & Vectors

The basic objects of study in linear algebra are **matrices** and **vectors**, which we will (usually) consist of elements drawn from the real numbers $\mathbb{R}$.

Definition 1

A **matrix** $A \in \mathbb{R}^{m \times n}$ is defined to be

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

where $a_{ij} \in \mathbb{R}$ for all $i = 1, \ldots, m$ and $j = 1, \ldots, n$. A **vector** $\mathbf{v} \in \mathbb{R}^m$ is defined to be

$$\mathbf{v} = (v_1, v_2, \ldots, v_m)$$

where $v_i \in \mathbb{R}$ for all $i = 1, \ldots, m$.

It is important to note that, for each matrix entry a_{ij}, the first index i corresponds to the *row* while the second index j corresponds to the column. This is standard across all disciplines which use linear algebra (which is a lot!). Vectors will alternatively be denoted by bold-faced (e.g. $\mathbf{v}$) or with an arrow overtop (e.g. $\vec{v}$). For graphical clarity, bold-faced will be favored in the online notes while the arrow notation will be favored in class.

Example 1

A matrix can be thought of as a rectangular *grid* of numbers. For example,

$$A = \begin{bmatrix} 2 & 0 & -1 & 0 \\ 1 & \frac{1}{2} & 6 & -2 \end{bmatrix}$$

and

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

are matrices. A vector can be thought of as a matrix with either one row (a *row vector*) or one column (a *column vector*). For example, if we have

$$\mathbf{v} = \begin{bmatrix} 2 & 3 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

then $\mathbf{v}$ is a row vector, and $\mathbf{w}$ is a column vector.

Since matrices consist of entries which are real numbers, which have well-defined operations $+$, $-$, $\times$, $\div$, it is natural to ask whether matrices have similar operations. We define the following.

Definition 2

Consider two matrices $A, B \in \mathbb{R}^{m\times n}$, and a constant $c \in \mathbb{R}$. Then the matrix $A + B \in \mathbb{R}^{m\times n}$ is the matrix with entries

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

and the matrix $cA \in \mathbb{R}^{m\times n}$ is the matrix with entries

$$[cA]_{ij} = ca_{ij}$$

for all $i = 1, \dots, m$, and $j = 1, \dots, n$.

Example 2

Consider the matrices

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

By the definition of matrix addition, we have

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

Note: It is important to recognize that the operation $A + B$ is only defined for matrices A and B which have the same dimensions. We may not, for instance, add two matrixes with the dimensions 2×3 and 2×5. In such a case, we say the operation is not defined.

Section 3: Matrix Multiplication

We will also need to define multiplication for matrices. This is the first operation which is not as intuitive as applying the standard operation component-wise to the relevant matrices. We have the following definition.

Definition 3

Suppose $A \in \mathbb{R}^{m\times p}$ and $B \in \mathbb{R}^{p\times n}$. Then the matrix $AB \in \mathbb{R}^{m\times n}$ is the matrix with entries

$$[AB]_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}.$$

Alternatively, we may think of matrix multiplication as consisting of multiplying rows of the first matrix component-wise with columns of the second, and then adding the results. The result of this operation is then placed in the row and column of the new matrix corresponding to the row and column which were just multiplied together.

Example 3

For the A and B defined in Example 2, we have

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} (1)(0)+(2)(1) & (1)(-1)+(2)(2) \\ (-3)(0)+(1)(1) & (-3)(-1)+(1)(2) \end{bmatrix} \\ &= \begin{bmatrix} 2 & 3 \\ 1 & 5 \end{bmatrix}. \end{aligned}$$

Note: The operation of matrix multiplication is only defined if the first matrix has *exactly* the same number of columns as the second matrix has rows. Otherwise, the matrix product AB is not defined. It can also be checked that, it is not generally the case that $AB = BA$ (check with the previous example!). In mathematical terminology, we say that matrix multiply is **noncommutative**. In plain English, we say **order matters**. This is a significant distinction with multiplication of real numbers, where $xy = yx$ holds trivially.

Section 4: Inverses and Determinants

A particularly important matrix is the *identity matrix*, which for 2×2 systems is defined as $I \in \mathbb{R}^{2\times 2}$ where

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

This matrix has the property that $AI = A$ and $IA = A$ for any matrix for which the multiplication is defined. Consequently, in the theory of linear algebra, the identity matrix can be thought of as be analogous to the value *one* in standard algebra.

An important question is whether, given a square matrix A, we can find a matrix which multiplies to give the identity matrix. That is to say, can be find a matrix B such that

$$AB = I.$$

The most important application of such a matrix is it allows an analogue to *division* for real numbers to be defined for matrices.

Definition 4

Consider a matrix $A \in \mathbb{R}^{2\times 2}$ defined by

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then the **inverse** matrix $A^{-1} \in \mathbb{R}^{2\times 2}$ is defined by

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

and has the property that $A^{-1}A = AA^{-1} = I$.

Proof

The identities can be checked directly. We have that

$$\begin{aligned} A^{-1}A &= \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} \\ &= \frac{1}{ad-bc}\begin{bmatrix} ad-bc & -ab+ab \\ cd-cd & -bc+ad \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

The other direction can be checked similarly.

The 2×2 inverse matrix A^{-1} has many nice properties. Most notably, it is guaranteed to be an inverse on the right *and* on the left ($A^{-1}A = I$ implies $AA^{-1} = I$, and vice versa), and it is unique (i.e. given a matrix A there exactly one matrix satisfying $A^{-1}A = AA^{-1} = I$).

An immediate consequence of this formula is that not every matrix has an inverse (i.e. not all matrices are *invertible*). This follows from the fact that $ad - bc$ may be zero, in which case the formula tells us to divide by zero. (This could have been anticipated by our interpretation of inverses as the matrix analogue of division!) The quantity telling us to avoid dividing by zero is important enough in the theory of linear algebra that it is given its own name.

Definition 5

The **determinant** of a matrix $A \in \mathbb{R}^{2\times 2}$ is defined by

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Many properties follow from the value of the determinant, but the one which is more readily apparent (and frequently used) is that a matrix is invertible *if and only if* $\det(A) \neq 0$. This is true not just for 2×2 matrices, but for any square matrix of any size. We note, however, that the form for the determinant is more complicated in higher dimensions.

Example 4

Determine whether the following matrices are invertible and, if so, find the inverse:

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

Solution: We can simply apply the formula. We first need to compute the determinants. We have

$$\det(A) = 2(-3) - (-5)(1) = -6 + 5 = -1$$

and

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

It follows that A is invertible and B is not. The inverse of A, A^{-1}, can be computed by the formula as

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{-1} \begin{bmatrix} -3 & 5 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ 1 & -2 \end{bmatrix}.$$

We can easily verify this is correct by computing

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

Note: To see how inverses are related to division, consider being asked to solve the algebraic expression $ax = b$ for x. We immediately recognize that we need to divide both sides by a. This is the same as multiplying by the reciprocal of a, which is $(1/a)$. Evaluating step-by-step, we have

$$ax = b \implies \left(\frac{1}{a}\right) ax = \left(\frac{1}{a}\right) b \implies 1 \cdot x = \frac{b}{a} \implies x = \frac{b}{a}.$$

Now consider the analogous matrix expression $A\mathbf{x} = \mathbf{b}$. In order to solve for the vector $\mathbf{x}$, we cannot simply divide by A because matrix division is not formally defined. We can, however, multiply by the inverse A^{-1}. We have

$$A\mathbf{x} = \mathbf{b} \implies A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \implies I\mathbf{x} = A^{-1}\mathbf{b} \implies \mathbf{x} = A^{-1}\mathbf{b}.$$

Section 5: Eigenvalues & Eigenvectors

We now introduce a pair of objects which will be the primary focus of our application of linear algebra to differential equations.

Definition 6

Consider a matrix $A \in \mathbb{R}^{2\times 2}$. Then a vector $\mathbf{v} \in \mathbb{C}^2$ satisfying the equation

$$A\mathbf{v} = \lambda\mathbf{v} \tag{1}$$

for some $\lambda \in \mathbb{C}$ is called an **eigenvector** of A, and the value λ is the corresponding **eigenvalue**.

Note that, for $A \in \mathbb{R}^{2\times 2}$ and $\mathbf{v} \in \mathbb{C}^2$, the operation $A\mathbf{v}$ always produces a 2-dimensional vector. We can therefore think of $A\mathbf{v}$ as an operation on 2-dimensional vectors. That is, it inputs a 2-dimensional vector, and it outputs another 2-dimensional vector. This is example of a **linear transformation**.

What is important to note about equation (1) is that the output vector $A\mathbf{v}$ usually has *nothing* to do with the input vector $\mathbf{v}$. If we picked a 2×2 matrix and a vector $\mathbf{v}$ at random, we would not expect to see much similarity between $\mathbf{v}$ and the vector produced by $A\mathbf{v}$. For example, for

$$A = \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

we compute

$$A\mathbf{v} = \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

As expected, the resulting vector $(2, 3)$ has very little resemblance to the vector we started with, $\mathbf{v} = (1, 1)$.

It is sometimes the case, however, for very carefully selected vectors, computing $A\mathbf{v}$ produces a vector which differs from $\mathbf{v}$ by only a constant value (a scaling in the (x, y)-plane). In some sense, the operation implied by A on the vector $\mathbf{v}$ does not affect the vector. Such a pair is called an eigenvalue/eigenvector pair. For example, if we choose the vector $\mathbf{v} = (1, 2)$ with the above matrix A, we have

$$A\mathbf{v} = \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

We will say that $\mathbf{v} = (1, 2)$ is an eigenvector of A with eigenvalue $\lambda = 2$.

For many applications, the eigenvectors $\mathbf{v}$ and corresponding eigenvalues λ of a matrix A tells us very important information about the what A can do. Differential equations will be no exception, although a more comprehensive treatment is given in **Math 129A**. This recognition does not, however, resolve the question of how we *find* the eigenvalues and eigenvectors of a given matrix. For example, how did we know to choose the vector $\mathbf{v} = (1, 2)$? At this point, it is little more than a lucky guess.

To answer this question, we consider the eigenvector/eigenvalue equation $A\mathbf{v} = \lambda\mathbf{v}$ in more depth. This can be rewritten as

$$A\mathbf{v} - \lambda I\mathbf{v} = \mathbf{0} \quad \Longrightarrow \quad (A - \lambda I)\mathbf{v} = \mathbf{0}. \tag{2}$$

We are allowed to do this because the identity matrix does not alter a matrix (or vector) upon multiplication (i.e. $\mathbf{v} = I\mathbf{v}$). The matrix $A - \lambda I$ is A with λ subtracted from the diagonal. It is a basic fact of linear algebra (but unfortunately not one we will be able to justify in this course) that a matrix can satisfy the equation $A\mathbf{v} = \mathbf{0}$ for a non-trivial vector $\mathbf{v}$ if and only if its determinant is zero. It follows that (2) can be true if and only if

$$\det(A - \lambda I) = 0. \tag{3}$$

Now consider the general 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

We can now clearly see that

$$A - \lambda I = \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix}.$$

It follows that $\det(A - \lambda I) = 0$ implies

$$(a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc) = 0. \tag{4}$$

This equation is called the **characteristic polynomial** of the matrix A. We can determine the eigenvalues of the matrix A by solving this quadratic for λ. It remains then to find the associated eigenvector $\mathbf{v} = (v_1, v_2)$ to a given eigenvalues λ. For this, we return to (2) and note that we can solve the expression

$$(A - \lambda I)\mathbf{v} = \mathbf{0}$$

for the vector $\mathbf{v}$ by expanding out the components (or row reduction).

Note: The characteristic polynomial (4) can also be solved by the quadratic formula, which gives

$$\lambda = \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad-bc)}}{2}.$$

This expression tells us a great deal about the general nature of eigenvalues and eigenvectors. There are three distinct cases:

1. Two real-valued and distinct eigenvalues (λ_1 and λ_2).
2. One repeated real-valued eigenvalue (λ).
3. A complex-valued conjugate pair of eigenvalues ($\lambda_{1,2} = \alpha \pm \beta i$ where $i = \sqrt{-1}$).

We will consider these cases for the eigenvalues separately, as they will require slightly different techniques for determining the corresponding eigenvectors.

Example 5

Find the eigenvalues and eigenvectors of the matrix

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

Solution: To compute the eigenvalues, we need to solve

$$\begin{aligned}\det(A-\lambda I) &= \begin{vmatrix} 2-\lambda & 0 \\ 2 & 1-\lambda \end{vmatrix} \\ &= (2-\lambda)(1-\lambda)-(2)(0) \\ &= (2-\lambda)(1-\lambda)=0.\end{aligned}$$

It is clear that we have two distinct real-valued roots, $\lambda_1 = 1$ and $\lambda_2 = 2$. Substituting the first value in the expression $(A - I)\mathbf{v} = \mathbf{0}$ to find the corresponding eigenvector gives

$$\begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

where we are trying to solve for v_1 and v_2. An expression like this is called a *linear system of equations.* They are commonly rewritten in the form

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

The question becomes how to solve such systems of equations. It turns (details omitted) that we are allowed perform the following three operations without changing the value of the solution, v_1 and v_2:

1. Multiply rows by scalar values.

2. Add rows.

3. Switch the order of rows.

What we want to do is use these operations to rearrange this matrix expression into a form satisfying the following two objectives:

1. The first non-zero entry in each row is a 1.

2. The first non-zero entry in each successive row is further to the right than the previous row.

Such a matrix is said to be in *echelon form.*

For our example (and the vast majority we will see), the method for obtaining this form will be obvious. In this case, we are almost there, since the first row begins with a 1. The only thing we need to worry about is the 2 in the second row. We need to remove it, because we are not permit to have any entry in that column by condition 2 above. We can remove it by taking the second row and subtracting two of the first row from it. We have

$$\xrightarrow{\tilde{\mathcal{R}}_2=\mathcal{R}_2-2\mathcal{R}_1} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

There are no further operations to perform. We recognize now that this corresponds to the system of equations

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which can be expanded into

$$\begin{aligned} (1)v_1 + (0)v_2 = 0 &\quad\Longrightarrow\quad v_1 = 0 \\ (0)v_1 + (0)v_2 = 0 &\quad\Longrightarrow\quad 0 = 0. \end{aligned}$$

The second equation is trivially satisfied while the first places a restriction on v_1. It follows that the system can be satisfied for any choice of $v_1 = 0$ and $v_2 = t$ where $t \in \mathbb{R}$. That is to say, the vectors we are looking for are

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = t\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Choosing $t = 1$ (a convenient choice, but any value will do), we obtain $\mathbf{v}_1 = (0, 1)$.

Now consider $\lambda_2 = 2$. We have

$$(A - 2I)\mathbf{v} = \mathbf{0} \quad\Longrightarrow\quad \begin{bmatrix} 0 & 0 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We put this in the condensed matrix form and solve to get

$$\left[\begin{array}{cc|c} 0 & 0 & 0 \\ 2 & -1 & 0 \end{array}\right] \xrightarrow{\tilde{\mathcal{R}}_1=(1/2)\mathcal{R}_2} \left[\begin{array}{cc|c} 1 & -(1/2) & 0 \\ 0 & 0 & 0 \end{array}\right].$$

It follows that we have the system

$$\begin{aligned} v_1 - (1/2)v_2 &= 0 \\ 0 &= 0. \end{aligned}$$

The second equation (again) is meaningless. This will be a general property of the eigenvector system we will have to solve. This can be satisfied if and only if $v_1 = (1/2)v_2$ which can be parameterized as $v_1 = (1/2)t$ and $v_2 = t$ for $t \in \mathbb{R}$. It follows that we have

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = t \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}.$$

Choosing the value $t = 2$ (an easy choice), we obtain the vector $\mathbf{v}_2 = (1, 2)$. This corresponds to the earlier eigenvector which we confirmed was associated with the eigenvalue $\lambda = 2$ by direct computation, and we are done!

Example 6

Find the eigenvalues and eigenvectors of the matrix

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

Solution: We have

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} -4-\lambda & -1 \\ 4 & -\lambda \end{vmatrix} \\ &= (-4-\lambda)(-\lambda) + 4 \\ &= \lambda^2 + 4\lambda + 4 = (\lambda + 2)^2 = 0. \end{aligned}$$

We do not even need to use the quadratic formula to determine that $\lambda = -2$. We notice that there is something distinctively different about this example since we have only found one eigenvalue instead of the typical two. Nevertheless, we can proceed as before and compute

$$
\begin{aligned}
(A - 2I) = \mathbf{0} \quad &\Longrightarrow \quad \left[\begin{array}{cc|c} -4-(-2) & -1 & 0 \\ 4 & -(-2) & 0 \end{array}\right] \\
&\longrightarrow \quad \left[\begin{array}{cc|c} -2 & -1 & 0 \\ 4 & 2 & 0 \end{array}\right] \\
&\longrightarrow \quad \left[\begin{array}{cc|c} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right].
\end{aligned}
$$

This corresponds to $v_1 + (1/2)v_2 = 0$. Setting $v_2 = t = 2$ (to clear the denominator) gives us the eigenvector $\mathbf{v} = (-1, 2)$, and since $\lambda = 2$ was the only eigenvalue, we are done.

Something is very different about this example. We have obtained an eigenvector, but one rather than the expected two. For the applications involving differential equations we will *need* to have a full set of eigenvectors—in this case, we will need two of them. Otherwise, we will not be able to solve the systems. The resolution is contained in the following definition.

Definition 7

A vector $\mathbf{w} \in \mathbb{R}^2$ is called a **generalized eigenvector** of $A \in \mathbb{R}^{2\times 2}$ if it satisfies

$$(A - \lambda I)\mathbf{w} = \mathbf{v}$$

where $\mathbf{v} \in \mathbb{R}^2$ is a regular eigenvector.

It is not immediately obvious how this new vector helps us (and the rigorous justification is beyond the scope of the course), but it is known that a generalized eigenvector only exists when there is a *repeated* eigenvalue $\lambda \in \mathbb{R}$ corresponding to only a single eigenvector $\mathbf{v}$ of a matrix $A \in \mathbb{R}^{2\times 2}$.

Example 7

Find a generalized eigenvector of the matrix

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

Solution: We want to find a generalized eigenvector $\mathbf{w}$ by using the equation $(A - \lambda I)\mathbf{w} = \mathbf{v}$ where $\mathbf{v} = (-1, 2)$. We have

$$\begin{aligned}(A - 2I)\mathbf{w} = \mathbf{v} \quad &\Longrightarrow \left[\begin{array}{cc|c} -2 & -1 & -1 \\ 4 & 2 & 2 \end{array}\right] \\ &\Longrightarrow \left[\begin{array}{cc|c} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 \end{array}\right].\end{aligned}$$

This corresponds to the system $w_1 + (1/2)w_2 = (1/2)$. Setting $w_2 = t$ gives $w_1 = (1/2) - (1/2)t$. Picking $t = 1$ gives $\mathbf{w} = (w_1, w_2) = (0, 1)$. So, corresponding to the repeated eigenvalue $\lambda = 2$, we have the regular eigenvector $\mathbf{v} = (-1, 2)$ and a corresponding generalized eigenvector $\mathbf{w} = (0, 1)$.

Note: As with regular eigenvalues, the choice of t in the selection of a generalized eigenvalue is arbitrary, but can be chosen for algebraic simplicity. This flexibility can, however, lead to vectors which seem wildly different than each other. For instance, choosing $t = 2$ gives the generalized eigenvector $\mathbf{w} = (-1/2, 2)$.

Example 8

Find the eigenvalues and eigenvectors of the matrix

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

Solution: We need to compute

$$\begin{aligned}\det(A - \lambda I) &= \begin{vmatrix} 3 - \lambda & -1 \\ 5 & -1 - \lambda \end{vmatrix} \\ &= (3 - \lambda)(-1 - \lambda) + 5 \\ &= \lambda^2 - 2\lambda + 2 = 0\end{aligned}$$

$$\Longrightarrow \quad \lambda = \frac{2 \pm \sqrt{4 - 8}}{2} = 1 \pm i.$$

We might be tempted to give up on the equation, but it turns our **complex-valued** eigenvalues and eigenvectors are extremely important in the study of linear algebra, and in solutions of real-valued differential equations. We need to find complex-valued eigenvectors $\mathbf{v} \in \mathbb{C}^2$.

There is some good new and some bad new with regards to solve $(A - \lambda I)\mathbf{v} = \mathbf{0}$ for $\mathbf{v} \in \mathbb{C}^2$:

1. The algebra is undoubtedly more challenging. Practice will help significantly in recognizing the general tricks.

2. We only have to perform the computation once. (It is a fact that if $\lambda = \alpha + \beta i$ corresponds to eigenvector $\mathbf{v} = \mathbf{a} + \mathbf{b}i$, then $\lambda = \alpha - \beta i$ corresponds to eigenvector $\mathbf{v} = \mathbf{a} - \mathbf{b}i$.)

We will start with $\lambda = 1 + i$. We have

$$\begin{aligned}
(A - (1+i)\,I)\,\mathbf{v} = \mathbf{0} \quad &\Longrightarrow \left[\begin{array}{cc|c} 3-(1+i) & -1 & 0 \\ 5 & -1-(1+i) & 0 \end{array}\right] \\
&\Longrightarrow \left[\begin{array}{cc|c} 2-i & -1 & 0 \\ 5 & -2-i & 0 \end{array}\right] \\
&\Longrightarrow \left[\begin{array}{cc|c} (2-i)\,(2+i) & -(2+i) & 0 \\ 5 & -2-i & 0 \end{array}\right] \\
&\Longrightarrow \left[\begin{array}{cc|c} 5 & -2-i & 0 \\ 5 & -2-i & 0 \end{array}\right] \\
&\Longrightarrow \left[\begin{array}{cc|c} 1 & -\frac{2}{5}-\frac{1}{5}i & 0 \\ 0 & 0 & 0 \end{array}\right]
\end{aligned}$$

This process may seem a little mysterious at first glance, but it will become very familiar after practice. What we want to recognize is that:

1. The eigenvector equation guarantees that we *absolutely must* be able to row reduce the condensed matrix so that there is a row of zeroes. That means there is *necessarily* some combination of the two rows which adds to zero. (That combination, however, involves complex numbers.)

2. Multiplying a complex number by its conjugate (e.g. for $\lambda = \alpha + \beta i$, the value $\bar{\lambda} = \alpha - \beta i$ is the conjugate) always produces a real number.

It follows that we just need to match up the real and complex numbers in the columns of the condensed matrix, which is what we have done above.

The final equation is equivalent to $v_1 - ((2/5) + (1/5)i)v_2 = 0$. Setting $v_2 = t$, we have $v_1 = ((2/5) + (1/5)i)t$. Writing in vector form, and then taking $t = 5$, it follows that we have the complex eigenvector

$$\mathbf{v} = \begin{bmatrix} 2+i \\ 5 \end{bmatrix} = \left(\begin{bmatrix} 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot i \right).$$

The importance of what we have done is that we have obtained two vectors, corresponding to the real and imagine parts of $\mathbf{v}$. We have $\mathbf{a} = (2, 5)$ and $\mathbf{b} = (1, 0)$. These are the two distinct vectors we will need when we return to considering differential equations.

It can be checked that choosing the eigenvalue $\lambda = 1 - i$ would have given the complex eigenvector

$$\mathbf{v} = \left(\begin{bmatrix} 2 \\ 5 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot i \right).$$

In particular, notice that we have the same real-valued vectors we identified before—$\mathbf{a} = (2, 5)$ and $\mathbf{b} = (1, 0)$.

Suggested Problems

1. For the following matrix, compute $A + B$, $B - 3C$, and $A - 2B + C$:

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

2. If possible, compute AB, CB, and ABC for the following matrices:

$$A = \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 8 & 1 & 5 \\ -1 & -3 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

3. Determine whether the following matrices are invertible. If so, determine the inverse:

(i) $\begin{bmatrix} 5 & -1 \\ 3 & 3 \end{bmatrix}$

(ii) $\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$

(iii) $\begin{bmatrix} -2 & -3 \\ 3 & 2 \end{bmatrix}$

(iv) $\begin{bmatrix} e & 1 \\ 1 & e^{-1} \end{bmatrix}$

(v) $\begin{bmatrix} 5 & 3 \\ -1 & 0 \end{bmatrix}$

(vi) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

4. Determine the eigenvalues, eigenvectors, and generalized eigenvectors (if applicable) of the following matrices:

(a) $\begin{bmatrix} -5 & 3 \\ 0 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 5 \\ -1 & -3 \end{bmatrix}$

(c) $\begin{bmatrix} -5 & 2 \\ -10 & 4 \end{bmatrix}$

(d) $\begin{bmatrix} 0 & -1 \\ 4 & -4 \end{bmatrix}$

(e) $\begin{bmatrix} 8 & 1 \\ -9 & 2 \end{bmatrix}$

(f) $\begin{bmatrix} -2.5 & 2 \\ -0.5 & 2.5 \end{bmatrix}$

(g) $\begin{bmatrix} \sqrt{2} & 1 \\ 4 & \sqrt{2} \end{bmatrix}$

(h) $\begin{bmatrix} -11 & -4 \\ 10 & 1 \end{bmatrix}$

(i) $\begin{bmatrix} -3 & -9 \\ 1 & 3 \end{bmatrix}$