Linear Algebra Tutorial

Math 2310-360: Differential Equations Spring 2019

This document summarizes some of the key definitions, theorems, and other information we may need from linear algebra.

1. Matrix Algebra

Definition 1.1.

An n ***by*** m ***matrix*** is a rectangular array of elements with n rows and m columns in which not only the value of an element is important, but also its position in the array. Mathematically, this is expressed as

$$A = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \ddots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}$$

Matrices are generally denoted by capital letters like A and an element of A in the ith row and jth column is denoted by a_{ij}.

Definition 1.2.

An $1 \times n$ matrix is referred to as an n***-dimensional row vector***

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{bmatrix}$$

and an $n \times 1$ matrix is referred to as an n***-dimensional column vector***

$$A = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}$$

Note that notationally, vectors are denoted by boldface lower case letters. A subscript will denote individual components of the vector.

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

This may be different than what you have seen in a Calculus III course. In other words, we use $\mathbf{x}$ not $\overrightarrow{x}$.

Matrix Operations

Equality. Two $m \times n$ matrices A and B are equal if all corresponding elements are equal. In other words, $a_{ij} = bij$ for each i and j.

Zero. The symbol **0** is used to denote the vector or matrix for which all elements are zero.

Addition. The sum of two $m \times n$ matrices A and B is defined as the matrix obtained by adding corresponding elements:

$$A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

With this definition, it follows that matrix addition is commutative and associative, so that

$$A + B = B + A, \qquad A + (B + C) = (A + B) + C$$

Subtraction. The difference $A - B$ of two $m \times n$ matrices is defined by

$$A - B = A + (-B)$$

Thus

$$A - B = (a_{ij}) - (b_{ij}) = (a_{ij} - b_{ij}),$$

Scalar Multiplication. The product of a matrix A by a real or complex number α is defined as follows

$$\alpha A = \alpha(a_{ij}) = (\alpha a_{ij})$$

That is, each element of A is multiplied by α. The distributive laws

$$\alpha(A + B) = \alpha A + \alpha B, \qquad (\alpha + \beta)A = \alpha A + \beta A$$

are satisfied for this type of multiplication. In particular, the negative of A, denoted by $-A$, is defined by

$$-A = (-1)A$$

Multiplication. The product AB of two matrices is defined whenever the number of columns in the first factor is the same as the number of rows in the second. If A and B are $m \times n$ and $n \times r$ matrices, respectively, then the product $C = AB$ is an $m \times r$ matrix. The element in the ith row and jth column of C is found by multiplying each element of the ith row of A by the corresponding element of the jth column of B and then adding the resulting products. In symbols,

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

By direct calculation, it can be shown that matrix multiplication satisfies the associative law

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

and the distributive law

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

In general matrix multiplication is not commutative. For both products AB and BA to exist and to be of the same size, it is necessary that A and B be square matrices of the same order. Even in that case the two products are usually unequal, so that, in general,

$$AB \neq BA$$

Transpose. The ***transpose*** of an $n \times m$ matrix $A = (a_{ij})$ is the $m \times n$ matrix $A^\top = (a_{ji})$ where for each i, the ith column of $A^\top$ is the same as the ith row of A

Vector Operations

Vector Addition. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1,\ x_2 + y_2,\ \ldots,\ x_n + y_n)$$

Scalar Multiplication. For $\mathbf{x} \in \mathbb{R}^n$ and $a \in \mathbb{R}$,

$$a\mathbf{x} = (ax_1, ax_2, \ldots, ax_n)$$

Multiplication of Vectors. There are several ways of forming a product of two vectors $\mathbf{x}$ and $\mathbf{y}$, each with n components. One is a direct extension to n dimensions of the familiar dot product from physics and calculus; we denote it by $\mathbf{x}^\top \mathbf{y}$ and write

$$\mathbf{x}^\top \mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

it follows that

- $\mathbf{x}^\top \mathbf{y} = \mathbf{y}^\top \mathbf{x}$
- $\mathbf{x}^\top(\mathbf{y} + \mathbf{z}) = \mathbf{x}^\top \mathbf{y} + \mathbf{x}^\top \mathbf{z}$
- $(\alpha \mathbf{x})^\top \mathbf{y} = \alpha(\mathbf{x}^\top \mathbf{y}) = \mathbf{x}^\top(\alpha \mathbf{y})$

2. Inner Products

Definition 2.1.

For any two vectors having the same number of components. The product, denoted by (x, y) or $\langle x, y\rangle$, is called the scalar or ***inner product*** and is defined by

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^{n} x_i \bar{y}_i$$

The scalar product is also a real or complex number meaning

$$\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^\top \bar{\mathbf{y}}$$

Thus, if all the elements of $\mathbf{y}$ are real, then the two products are identical.

Definition 2.2.

If V is a vector space, a function $\langle \cdot, \cdot\rangle : V \times V \to \mathbb{R}$ is an ***inner product*** on V if for for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$

(i) $\langle \mathbf{x}, \mathbf{x}\rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x}\rangle = 0$ iff $\mathbf{x} = \mathbf{0}$.

(ii) $\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{y}, \mathbf{x}\rangle$

(iii) $\langle \mathbf{x}, a\mathbf{y} + b\mathbf{z}\rangle = a\langle \mathbf{x}, \mathbf{y}\rangle + b\langle \mathbf{x}, \mathbf{z}\rangle$ for any $a, b \in \mathbb{R}$.

Theorem 2.1 (Properties of Inner Product). We have the following properties for an inner product

- $\langle \mathbf{x}, \mathbf{y}\rangle = \overline{\langle \mathbf{x}, \mathbf{y}\rangle}$
- $\langle \alpha \mathbf{x}, \mathbf{y}\rangle = \alpha\langle \mathbf{x}, \mathbf{y}\rangle$
- $\langle \mathbf{x}, \alpha \mathbf{y}\rangle = \overline{\alpha}\langle \mathbf{x}, \mathbf{y}\rangle$
- $\langle \mathbf{x}, \mathbf{y} + \mathbf{z}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{x}, \mathbf{z}\rangle$

3. Vector Norms

Definition 3.1.

A ***norm*** on a vector space V is a function $\|\cdot\| : V \to \mathbb{R}$ that satisfies the following conditions for any $x, v \in V$ and $a \in \mathbb{R}$:

(i) $\|x\| \geq 0$, and $\|x\| = 0$ iff $x = 0$.

(ii) $\|ax\| = |a| \, \|x\|$

(iii) $\|x + y\| \leq \|x\| + \|y\|$

Definition 3.2.

There are 3 key vector norms for a vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$:

(i) $\|x\|_1 = |x_1| + \cdots + |x_n|$

(ii) $\|x\|_2 = (x_1^2, \ldots, x_n^2)^{1/2}$

(iii) $\|x\|_\infty = \max\limits_{1 \leq i \leq n} |x_i|$

Definition 3.3.

A ***normalized*** vector is a vector is one that has length 1. In other words,

$$\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2} = 1$$

Any vector can be normalized by dividing each element in the vector by its length. In other words, the normalized vector $\mathbf{t}$ is given by

$$\mathbf{t} = \left(\frac{x_1}{\|\mathbf{x}\|}, \frac{x_2}{\|\mathbf{x}\|}, \ldots, \frac{x_n}{\|\mathbf{x}\|} \right)$$

Remark: We note the following facts regarding inner products

(i) If $\mathbf{x}, \mathbf{u} \in \mathbb{R}^n$ then

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \ldots x_n y_n$$

(ii) If $f, g \in C^k[a, b]$ then

$$\langle f, g \rangle = \int_a^b f(x) g(x) \, dx$$

(iii) $\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$

Definition 3.4.

Two vectors $\mathbf{x}, \mathbf{y} \in V$ are ***orthogonal*** if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$

Theorem 3.1. (Cauchy-Schwarz Inequality) If V is an inner product space with inner product $\langle \cdot, \cdot \rangle$, then for any $\mathbf{x}, \mathbf{y} \in V$

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|_2 \, \|\mathbf{y}\|_2$$

4. Solving Systems of Equations

Our goal is to solve a system of linear equations of the form

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}$$

In other words, a system of n equations in n unknowns $x_1, x_2, \ldots, x_n$. Note that a_{ij} and b_{ij} are assumed to be real numbers.

Written in matrix form, we obtain the matrix equation:

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \ddots & \ldots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

This is equivalently expressed as the matrix equation

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

In order to solve the linear system we construct the ***augmented matrix***

$$[A, \mathbf{b}] = \left[\begin{array}{cccc:c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \ddots & \ldots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} & b_n \end{array}\right]$$

Where the goal is to obtain an upper triangular matrix otherwise known as the ***reduced form*** of the matrix through the use of Gaussian Elimination. Recall that the following operations are permitted when completing Gaussian Elimination:

Definition 4.1.
The ***elementary row operations*** on a system are defined as the following where $\mathcal{E}_i$ is the ith equation in the system.

(1) Interchange of two rows: $\mathcal{E}_i \leftrightarrow \mathcal{E}_j$

(2) Scalar (nonzero) Multiplication: $\alpha\mathcal{E}_i \leftrightarrow \mathcal{E}_i$

(3) Addition of any multiple of one row to another row: $\mathcal{E}_i + \alpha\mathcal{E}_j \leftrightarrow \mathcal{E}_i$.

Theorem 4.1. If one system of equations is obtained from another system of equations by a finite sequence of elementary row operations, then the two systems of equations are equivalent.

Definition 4.2.

Suppose $A\mathbf{x} = \mathbf{b}$ and $B\mathbf{x} = \mathbf{d}$ are both systems of n equations with n unknowns. If the two systems have the same solution $\mathbf{x}$ then they are ***equivalent systems***.

To find a solution to a system of equations, the goal is to transform our augmented matrix $[A, \mathbf{b}]$ into a form in which we can use one of the two direct approaches:

- **Forward Substitution** which implies we are solving a ***lower triangular matrix***.

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

- **Backward Substitution** which implies we are solving an ***upper triangular matrix***.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

We can reorder equations to get one of these forms and then solve. For example

Example 4.1.

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & 0 & 0 \end{bmatrix} \Longrightarrow A = \begin{bmatrix} a_{31} & 0 & 0 \\ a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

5. Determinants

The determinant of a square matrix A is a value that indicates whether A is invertible. The easiest determinant to calculate is for a 2×2 matrix.

Definition 5.1.

The ***determinant*** of a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is given by $\det(A) = ad - cb$

Definition 5.2.

The (ij)th ***minor*** of $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$ is the determinant of the matrix we get by eliminating row i and column j, denoted M_{ij}.

Definition 5.3.

The (ij)th ***cofactor*** of A is

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

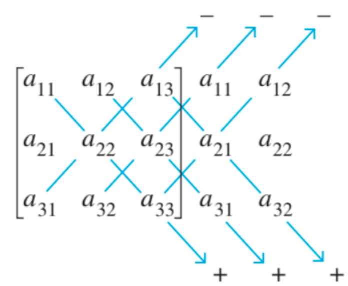
is the determinant of the matrix we get by eliminating row i and column j, denoted M_{ij}.

Then the ***determinant*** of a square matrix A is

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{nn}$$

There is also a shortcut for calculating a 3×3 determinant as shown in the following figure.

1. Re-copy the first two columns of the matrix to the right of the matrix.
2. Multiply entries in diagonal with 3 entries.
3. Add or subtract the products of these entries according to the pattern in the figure.



Theorem 5.1. If $C = AB$ then $\det(C) = \det(A)\det(B)$

6. Inverse of a Matrix

The inverse is essential when trying to solve a system of linear equations since

$$A\mathbf{x} = \mathbf{b} \qquad \Longrightarrow \qquad \mathbf{x} = A^{-1}\mathbf{b}$$

Naturally, this means we need to know when A^{-1} will exist.

Definition 6.1.
The multiplicative identity, or simply the ***identity matrix*** I, is given by

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

From the definition of matrix multiplication, we have

$$AI = IA = A$$

for any (square) matrix A. Hence the commutative law does hold for square matrices if one of the matrices is the identity.

Definition 6.2.
The matrix $A \in \mathbb{R}^{n\times n}$ is ***nonsingular*** if for any $b \in \mathbb{R}^n$ there exists a unique vector $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$. Otherwise A is ***singular***.

Definition 6.3.
The square matrix A is said to be ***invertible*** if there is a unique matrix $A^{-1} \in \mathbb{R}^{n\times n}$ such that

$$AA^{-1} = A^{-1}A = I$$

The easiest inverse to compute is that of a 2×2 matrix.

Definition 6.4 (Inverse of a 2×2 Matrix).
The inverse of the 2 by 2 matrix

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

is found by the following calculation

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

To compute the inverse of a matrix A, it exists, is to perform Gaussian Elimination to transform the augmented matrix $[A, I]$ to the identity matrix on the right hand side. Any nonsingular matrix A can be transformed into the identity I by a systematic sequence of row operations, often denoted by E_i. It is possible to show that if the same sequence of operations is then performed on I, it is transformed into A^{-1}.

Example 6.1.
To find the inverse A^{-1} of the matrix A

$$\left[\begin{array}{rrr:rrr} 1 & -1 & -1 & 1 & 0 & 0 \\ 3 & -1 & 2 & 0 & 1 & 0 \\ 2 & 2 & 2 & 0 & 0 & 1 \end{array}\right]$$

We construct the augmented matrix $[A, I]$

$$\left[\begin{array}{rrr:rrr} 1 & -1 & -1 & 1 & 0 & 0 \\ 3 & -1 & 2 & 0 & 1 & 0 \\ 2 & 2 & 2 & 0 & 0 & 1 \end{array}\right] \underset{\substack{\text{Row}\\ \text{Reduce}}}{\rightarrow} \left[\begin{array}{rrr:rrr} 1 & 0 & 0 & \frac{7}{10} & -\frac{1}{10} & \frac{3}{10} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{4}{5} & \frac{2}{5} & -\frac{1}{5} \end{array}\right]$$

and so we obtain the inverse matrix A^{-1}

$$A^{-1} = \begin{bmatrix} \frac{7}{10} & -\frac{1}{10} & \frac{3}{10} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{4}{5} & \frac{2}{5} & -\frac{1}{5} \end{bmatrix}$$

Definition 6.5.
If A, B are matrices where $AB = I$ then B is a ***right inverse*** of A and A is a ***left inverse*** of B.

Theorem 6.1. A square matrix can possess at most one right inverse.

Theorem 6.2. If A and B are square matrices such that $AB = I$, then $BA = I$.

Proposition 6.1. If a matrix is invertible a sequence of elementary row operations can be applied to A reducing it to I. In other words,

$$E_m E_{m-1} \dots E_2 E_1 A = I$$

where the sequence of elementary row operations applied to A is denoted by $E_m E_{m-1} \dots E_2 E_1$. It also follows that

$$A^{-1} = E_m E_{m-1} \dots E_2 E_1$$

Theorem 6.3. For an $n \times n$ matrix A the following are equivalent

(1) The inverse of A exists (A is nonsingular).

(2) The determinant of A is not zero: $\det(A) \neq 0$.

(3) The rows of A form a basis for $\mathbb{R}^n$.

(4) The columns of A form a basis for $\mathbb{R}^n$.

(5) The map $A : \mathbb{R}^n \to \mathbb{R}^n$ is injective (one to one).

(6) The map $A : \mathbb{R}^n \to \mathbb{R}^n$ is surjective (onto).

(7) The equation $A\mathbf{x} = 0$ implies that $\mathbf{x} = \mathbf{0}$.

(8) For each $\mathbf{b} \in \mathbb{R}^n$ there exists exactly one $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$.

(9) A is a product of elementary matrices.

(10) Zero is not an eigenvalue of A.

7. Matrix Functions

We also may need to consider vectors or matrices whose elements are functions. For example,

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} \qquad \text{and} \qquad A(t) = (a_{ij}) = \begin{bmatrix} a_{11}(t) & a_{12}(t) & \dots & a_{1m}(t) \\ a_{21}(t) & a_{22}(t) & \dots & a_{2m}(t) \\ \vdots & \ddots & \dots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \dots & a_{nm}(t) \end{bmatrix}$$

The matrix $A(t)$ is continuous at a point or on an interval if each element of A is continuous at that point or on that interval.

Differentiation of Vector & Matrix Functions

The matrix $A(t)$ is differentiable if each of it's elements $a_{ij}(t)$ is differentiable.

Definition 7.1.
The derivative of a matrix function is denoted $\frac{dA}{dt}$ and is defined by

$$\frac{dA}{dt} = \left(\frac{da_{ij}}{dt}\right) = \left(\frac{d}{dt} a_{ij}(t)\right)$$

i.e. to find the derivative of the matrix A, find the derivative of each element of A.

We have some of the same basic rules for derivatives as we do in Calculus.

Theorem 7.1. For matrices A and B and the scalar c we have

$$\frac{d}{dt}(cA) = c\frac{dA}{dt} \qquad \text{where } c \text{ can also be a constant matrix}$$

$$\frac{d}{dt}(A+B) = \frac{dA}{dt} + \frac{dB}{dt} \qquad \text{Sum Rule}$$

$$\frac{d}{dt}(AB) = \frac{dA}{dt}B + A\frac{dB}{dt} \qquad \text{Product Rule}$$

Integration of Vector & Matrix Functions

The matrix $A(t)$ is integrable if each of it's elements $a_{ij}(t)$ is integrable.

Definition 7.2.
The integral of a matrix function is defined as

$$\int_a^b A(t)\,dt = \left(\int_a^b a_{ij}(t)\,dt\right)$$

i.e. to find the integral of the matrix A, find the integral of each element of A.

8. Vector Spaces

Definition 8.1.
We define the following operations for a set V.

- An ***addition*** on a set V is a function that assigns an element $u+v \in V$ to each pair of elements $u, v \in V$.
- A ***scalar multiplication*** on V is a function that assigns an element $a \in V$ to each $a \in \mathbb{R}$ and each $v \in V$.

Definition 8.2.
A ***vector space*** v is a set V along with addition on V and scalar multiplication on V such that the following properties hold

1. **commutativity:**
$$u+v = v+u, \qquad \text{for all } u, v \in V$$
2. **associativity:**
$$(u+v)+w = u+(v+w) \quad \text{and} \quad (ab)v = a(bv) \quad \text{for all} u, v, w \in V \text{and all } a, b \in F$$
3. **additive identity:** there exists an element $0 \in V$ such that $v+0 = v$ for all $v \in V$
4. **additive inverse:** for every $v \in V$, there exists $w \in V$ such that $v+w = 0$

5. **multiplicative identity:** $1v = v$ for all $v \in V$

6. **distributive properties:**

$$a(u+v) = au + av \quad \text{and} \quad (a+b)u = au + bu \quad \text{for all } a, b \in F \text{ and all } u, v \in V$$

Definition 8.3 (Linear Combination).
If V is a real vector space, a ***linear combination*** of the vectors $x_1, \ldots, x_n \in V$ is a vector of the form

$$c_1 x_1 + \cdots + c_n x_n$$

where $c_1, \ldots, c_n \in \mathbb{R}$.

Definition 8.4.
If $S \subset V$ the ***span*** of S denoted Span (S) is the set of all linear combinations of vectors belonging to S. If $U =$ Span (S) then S ***spans*** U.

Definition 8.5.
A list $(v_1, \ldots, v_m)$ of vectors in V is called ***linearly independent*** if the only choice of $a_1, \ldots, a_m \in F$ that makes

$$a_1 v_1 + \cdots + a_m v_m = 0$$

is $a_1 = \ldots = a_m = 0$. Otherwise, they are ***linearly dependent***

Definition 8.6.
A subset S of a vector space V is a ***basis*** for V if S is linearly independent and Span $(S) = V$

Example 8.1.
The set of vectors

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a basis for $\mathbb{R}^3$.

9. Eigenvalues and Eigenvectors

Definition 9.1.
A number $\lambda \in \mathbb{C}$ is an ***eigenvalue*** of $A \in \mathbb{R}^{n \times n}$ if there is a nonzero vector $\mathbf{x} \in \mathbb{C}^n$ for which

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

Any such vector $\mathbf{x}$ is an ***eigenvector*** of A associated with λ

Definition 9.2.
The equation in (1) is equivalent to

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

and has a nonzero solution if and only if λ is found such that

$$\det(A - \lambda I) = 0 \tag{2}$$

This is a polynomial equation of degree n in λ and is called the ***characteristic equation*** of the matrix A.

Definition 9.3.
If A is a square matrix, the ***characteristic polynomial*** is the polynomial defined by

$$p(\lambda) = \det(\lambda I - A)$$

and has a degree n in λ.

Since eigenvalues are defined as the zeros of the characteristic polynomial we have the following definition.

Definition 9.4.
If an eigenvalue λ is repeated m times in the characteristic polynomial then its ***algebraic multiplicity*** is m.

Every eigenvalue will have at least one associated eigenvector. It is possible for an eigenvalue to have more than one associated eigenvector.

Definition 9.5.
An eigenvalue λ of multiplicity m may have q eigenvectors. This value q is defined as the ***geometric multiplicity*** and can be any value in the interval

$$1 \leq q \leq m$$

Definition 9.6.
An eigenvalue is ***simple*** if it has algebraic multiplicity of 1

Theorem 9.1. If each eigenvalue of A is simple, then each eigenvalue also has geometric multiplicity of 1

Definition 9.7.
The collection λ_A of all eigenvalues of A is the ***spectrum*** of A and the ***spectral radius*** is the number

$$\rho(A) = \max |\lambda|$$

Theorem 9.2. Let $A \in \mathbb{R}^{n\times n}$. Then

(i) A is singular iff 0 is an eigenvalue of A.

(ii) If A is upper or lower triangular, then its eigenvalues are its diagonal entries.

(iii) if A is symmetric, then all of its eigenvalues are real numbers.

(iv) If A is symmetric *and* non-negative, that is $\mathbf{x}^T A\mathbf{x} \geq 0$ for every $\mathbf{x} \in \mathbb{R}^n$, then all eigenvalues of A are non-negative.

(v) If A is symmetric *and* positive definite, then all of its eigenvalues are positive.

(vi) If A is symmetric, then there exists an orthonormal basis for $\mathbb{R}^n$, each of whose elements is an eigenvector of A.

Remark: Note the following

- Eigenvalues are the factors by which A stretches its eigenvectors.
- $A\mathbf{x} = \lambda\mathbf{x}$ with $\mathbf{x} \neq 0$ implies that the matrix defined by $\lambda I - A$ is singular, thus any eigenvalue λ of A is a zero of the characteristic polynomial.

Definition 9.8.
A set of vectors $\{v_1, \ldots, v_n\}$ is ***orthonormal*** if the vectors in it are pairwise orthogonal and each vector has norm 1.

Example 9.1.
Find Eigenvalues of the matrix

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

Solution. We find the eigenvalues of A by finding the set of λ_i that satisfy $\det(A - \lambda I) = 0$.

$$\det(A - \lambda I) = \begin{vmatrix} -1-\lambda & 0 & 0 \\ 0 & 2-\lambda & -4 \\ 1 & 4 & 2-\lambda \end{vmatrix} = -\lambda^3 + 3\lambda^2 - 16\lambda - 20 = -(\lambda+1)\left(\lambda^2 - 4\lambda + 20\right) = 0$$

So our three eigenvalues are $\lambda_1 = -1$, and $\lambda_{2,3} = 2 \pm 4i$. Next we find the corresponding eigenvectors by finding the solution to

$$(A - \lambda_k I)\xi^{(k)} = 0$$

For $\lambda = -1$ we have

$$(A + I) = \begin{bmatrix} -1-(-1) & 0 & 0 \\ 0 & 2-(-1) & -4 \\ 1 & 4 & 2-(-1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & -4 \\ 1 & 4 & 3 \end{bmatrix} \underset{\substack{\text{Row} \\ \text{Reduce}}}{=} \begin{bmatrix} 1 & 0 & \frac{25}{3} \\ 0 & 1 & \frac{4}{3} \\ 0 & 0 & 0 \end{bmatrix}$$

and so

$$\begin{bmatrix} 1 & 0 & \frac{25}{3} \\ 0 & 1 & \frac{4}{3} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \implies \begin{cases} \xi_1 + \dfrac{25}{3}\xi_3 = 0, \\ \xi_2 - \dfrac{4}{3}\xi_3 = 0 \end{cases}$$

$$\implies \xi_3 = -\frac{3}{25}\xi_1$$

$$\xi_2 = \frac{4}{3}\xi_3 = \left(\frac{4}{3}\right)\left(-\frac{3}{25}\xi_1\right) = -\frac{4}{25}\xi_1$$

So our corresponding eigenvector is

$$\xi^{(1)} = \begin{bmatrix} -25 \\ 4 \\ 3 \end{bmatrix}$$

To find the corresponding eigenvector for $\lambda_{2,3} = 2 \pm 4i$ we use $\lambda_2 = 2 + 4i$. We have

$$(A - (2+4i)I) = \begin{bmatrix} -1-(2+4i) & 0 & 0 \\ 0 & 2-(2+4i) & -4 \\ 1 & 4 & 2-(2+4i) \end{bmatrix} = \begin{bmatrix} -4i-3 & 0 & 0 \\ 0 & -4i & -4 \\ 1 & 4 & -4i \end{bmatrix}$$

Since the two eigenvalues are complex conjugates, then the eigenvectors must also be complex conjugates and so the corresponding eigenvectors are

$$\xi^{(2)} = \begin{bmatrix} 0 \\ i \\ 1 \end{bmatrix}, \qquad \xi^{(3)} = \begin{bmatrix} 0 \\ -i \\ 1 \end{bmatrix} \implies \xi^{(2,3)} = \begin{bmatrix} 0 \\ \pm i \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{u}} \pm i \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{w}}$$

10. Special Matrices

Definition 10.1.
For a matrix A the ***conjugate matrix***, denoted $\bar{A}$ is found by replacing each element of A, denoted (a_{ij}) with its complex conjugate, denoted $(\bar{a}_{ij})$.

Definition 10.2.
The transpose of the conjugate matrix A, in other words $\bar{A}^\top$ is called the ***adjoint***, of A. This is denoted as A^*.

Definition 10.3.
A matrix A is ***self-adjoint*** (or ***Hermitian***) if $A^* = A$. In other words, if $(a_{ij}) = (\bar{a}_{ij})$.

Definition 10.4.
A matrix A is ***symmetric*** if $A^T = A$.

Hermitian matrices have useful properties which are outlined in the following theorem.

Theorem 10.1. If a matrix A is Hermitian then,

1. All eigenvalues are real.
2. There always exists a full set of n linearly independent eigenvectors, regardless of the algebraic multiplicities of the eigenvalues.
3. If $x^{(1)}$ and $x^{(2)}$ are eigenvectors that correspond to different eigenvalues, then $\langle x^{(1)}, x^{(2)} \rangle = 0$. Thus, if all eigenvalues are simple, then the associated eigenvectors form an orthogonal set of vectors.
4. Corresponding to an eigenvalue of algebraic multiplicity m, it is possible to choose m eigenvectors that are mutually orthogonal. Thus the full set of n eigenvectors can always be chosen to be orthogonal as well as linearly independent.

Definition 10.5.

A ***tridiagonal matrix*** has the form

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & 0 \\ \vdots & & \ddots & \ddots & \ddots\ a_{n-1,n} \\ 0 & 0 & 0 & a_{n,n-1} & a_{nn} \end{bmatrix}$$

Definition 10.6.

The matrix $A \in \mathbb{R}^{n \times n}$ is ***strictly diagonally dominant*** if

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|$$

Theorem 10.2. When A is symmetric,

$$\|A\|_2 = \sqrt{\rho(A^T A)} = \sqrt{\rho(A^2)} = \rho(A)$$

Definition 10.7.

A matrix is ***positive definite*** if $x \neq 0$ and

$$x^T A x = \langle Ax, x \rangle > 0$$

and ***symmetric positive definite*** if $A^T = A$