

# HONORS LINEAR ALGEBRA AND DIFFERENTIAL EQUATIONS NOTES

Bryan Li

Fall 2022

# CONTENTS

0.1	Overview . . . . .	1
0.1.1	What is Linear Algebra? . . . . .	1
0.1.2	What are Ordinary Differential Equations? . . . . .	1
0.1.3	What is the relation between them? . . . . .	1
0.1.4	Grading Policy . . . . .	1
0.1.5	Office Hours . . . . .	1
I	LINEAR ALGEBRA . . . . .	2
1.1	Vectors in the real plane . . . . .	2
1.2	Conic Sections and Quadratic Forms . . . . .	4
1.3	More on Linear Transformations in 2D . . . . .	7
1.3.1	Operations on Linear Transformations . . . . .	7
1.3.2	Linear groups . . . . .	8
1.4	Orthogonal Transformations . . . . .	8
1.4.1	Complex Numbers . . . . .	9
1.5	The Main Theorems of Linear Algebra . . . . .	9
1.6	Matrices in Higher Dimensions . . . . .	10
1.6.1	Determinant . . . . .	11

## O.I OVERVIEW

### O.I.1 WHAT IS LINEAR ALGEBRA?

- Linear: The structure of scalar multiplication and addition.
- Algebra: Study of formal structures in math. Examples: Groups  $(S_n, \text{GL}(n, \mathbb{F}))$ , Rings  $(F[X], \mathbb{Z})$ , and Fields  $(\mathbb{R}, \mathbb{C})$ .

### O.I.2 WHAT ARE ORDINARY DIFFERENTIAL EQUATIONS?

### O.I.3 WHAT IS THE RELATION BETWEEN THEM?

Solutions to ordinary differential equations form a linear vector space which is a subspace of certain function spaces.

### O.I.4 GRADING POLICY

- HW: 0%
- Weekly Quiz: 20%
- Participation: 10%
- Midterms (2): 30%
- Final: 40%

### O.I.5 OFFICE HOURS

Tuesday 12:30–1:30, Thursday 3–4.

# I LINEAR ALGEBRA

Class responses to “What is a vector?”:

1. Something with magnitude and direction.
2. A column of numbers.
3. An element of a vector space.
4. An ordered collection of magnitudes.

## I.1 VECTORS IN THE REAL PLANE

Vectors are commonly defined as arrows on the plane, especially in elementary treatments. The following is a semi-formal treatment of this definition.

In order to define vectors, we must first define what a directed segment is.

**Definition 1.1.1** (Directed Segment). A directed segment is a line segment between two points with a direction, which can be translated around the plane.<sup>1</sup>

If a segment starts on point  $A$  and ends on point  $B$ , then we call that directed segment  $\overrightarrow{AB}$ . Now, we can define vectors using these directed segments.

**Definition 1.1.2** (Vector). Vectors are equivalence classes of directed segments modulo translation. What this means is that if one segment can be translated such that it is equal to another segment, then those segments represent the same vector.

Now, we can define some basic operations and properties on these vectors.

1. Vector addition: Given two vectors  $u$  and  $v$ , we define  $u + v$  by choosing directed segments  $\overrightarrow{AB}$  and  $\overrightarrow{CD}$  that represent vectors  $u$  and  $v$ , respectively, such that  $B = C$ . Then, we can concatenate these segments into  $\overrightarrow{AC}$ , then taking the equivalence class  $[\overrightarrow{AC}]$  to be the sum.
2. Scalar multiplication: Given a vector  $v$  and a scalar  $\lambda$ , we can define scalar multiplication by integers by adding the vector multiple times to itself, and negation as being reflection.

---

<sup>1</sup>This is a very informal definition. Formally, we may view directed segments as ordered pairs of points in  $\mathbb{R}^2$  where the translation operation is an  $\mathbb{R}^2$ -action on the set of directed segments.

Multiplication by rational and irrational numbers are harder to define intuitively, however, it can be defined formally on a vector  $[\overrightarrow{AB}]$  by  $\lambda[\overrightarrow{AB}] = [\overrightarrow{A'B'}]$  where  $A' = \lambda A$  and  $B' = \lambda B$ .

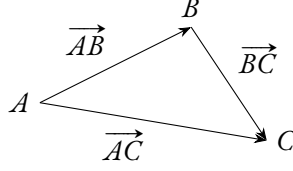


Figure 1.1: Concatenation of Directed Segments

3. Vector multiplication? In general, there is no good way to multiply vectors. However, in this example, we can identify  $\mathbb{R}^2$  with  $\mathbb{C}$  and define multiplication in that fashion, although this method does not generalize. Vector spaces with a notion of multiplication are called *algebras*.
4. Linear combinations: If  $u, v$  are vectors and  $a, b \in \mathbb{R}$ , then  $au + bv$  is a linear combination of  $u$  and  $v$ .
5. Inner product<sup>2</sup>: A way to “multiply” two vectors to output a scalar, we can define  $\langle v|u \rangle \in \mathbb{R}$  as  $|u||v| \cos(\theta)$ , where  $|u|$  is the length of a directed segment that represents  $u$ , and  $\theta$  being the angle between representatives of  $u$  and  $v$  whose starting points are equal. This satisfies bilinearity (linear in both inputs), which means  $\langle au|w \rangle = a\langle u|w \rangle$ , and  $\langle u + v|w \rangle = \langle u|w \rangle + \langle v|w \rangle$ , with analogous identities for the second input.
6. Bases: The standard basis  $e_1$  and  $e_2$  are represented the unit directed segments pointing to the right and up. Any vector can be written as a linear combination of these two basis vectors. This basis is also orthonormal, i.e. their inner product is 0 and their lengths are 1.  
  
In general, a basis in  $\mathbb{R}^2$  is a set of two vectors that satisfies the property where any vector in  $\mathbb{R}^2$  can be written as a linear combination of those two vectors. In addition, one can prove that as long as two vectors aren’t pointing in the exact same or opposite directions, they form a basis.
7. Coordinates: In order to give a vector coordinates (numerical representation), we must specify a basis. After we have specified a basis  $e_1, e_2$ , vector  $v$  has coordinates  $(a_1, a_2)$  if  $v = a_1 e_1 + a_2 e_2$ .

<sup>2</sup>Note: Not all vector spaces have an inner product. Also note that in general, the inner product is defined first, then the norm (magnitude) of the vector is defined as  $\sqrt{\langle v|v \rangle}$

A special fact about orthonormal bases is that the coordinates of a vector with respect to such a basis corresponds with the inner product of the vector with elements of the basis. This makes it straightforward to compute coordinates as long as you have a formula for the inner product.

**Theorem 1.1.1.** For an orthonormal basis  $e_1, e_2$ , we have  $v = \langle v|e_1\rangle e_1 + \langle v|e_2\rangle e_2$ .

*Proof.*  $v$  can be decomposed into  $ae_1 + be_2$ , and thus

$$\begin{aligned} \langle ae_1 + be_2|e_1\rangle e_1 + \langle ae_1 + be_2|e_2\rangle e_2 &= \langle ae_1|e_1\rangle e_1 + \langle be_2|e_1\rangle e_1 + \langle ae_1|e_2\rangle e_2 + \langle be_2|e_2\rangle e_2 \\ &= a\langle e_1|e_1\rangle e_1 + b\langle e_2|e_1\rangle e_1 + a\langle e_1|e_2\rangle e_2 + b\langle e_2|e_2\rangle e_2 \\ &= (a \cdot 1)e_1 + (b \cdot 0)e_1 + (a \cdot 0)e_2 + (b \cdot 1)e_2 \\ &= ae_1 + be_2. \end{aligned}$$

□

## 1.2 CONIC SECTIONS AND QUADRATIC FORMS

Conic sections are the intersections of a infinite double-cone by a plane. The “shape” generated is dependent on how tilted this plane is.

- Flat plane: Circle
- Slightly tilted: Ellipse
- Tilted until parallel to slope of cone: Parabola
- Tilted past slope of cone: Hyperbola

Ellipses on the plane can be defined by fixing two focal points (foci) and specifying a total length, then taking the set of all points such that the distances to the two foci sum to the total length specified.

What’s more interesting is that the intersection of a plane to the cone that generates an ellipse allows for two spheres that are tangent to both the plane and the cone, one below and one above the plane. The intersection between those spheres and the plane are the foci of the ellipse. Furthermore, the distance between the two circles of tangency along the cone is the total length.

Now, we define these objects algebraically:

**Definition 1.2.1** (Double Cone). The double-cone is defined in  $\mathbb{R}^3$  as the set of points satisfying  $x_1^2 + x_2^2 = x_3^2$ .

**Definition 1.2.2** (Plane). A plane in  $\mathbb{R}^3$  is the set of points satisfying  $ax_1 + bx_2 + cx_3 = d$ .

Now, to consider the intersection, we consider the solution set to the *system of equations*

$$\begin{cases} x_1^2 + x_2^2 = x_3^2 \\ ax_1 + bx_2 + cx_3 = d. \end{cases}$$

How do we solve this system? Substitution. Assuming  $c \neq 0$ , we can substitute  $x_3 = -\frac{1}{c}(ax_1 + bx_2 - d)$  into the first equation and get a quadratic equation in two variables. We then get an equation of the form  $Ax_1^2 + Bx_1x_2 + Cx_2^2 + Dx_1 + Ex_2 + F = 0$ . This leads us to the following definition:

**Definition 1.2.3** (Conic curve). A conic curve in  $\mathbb{R}^2$  is the solution set to  $Ax_1^2 + Bx_1x_2 + Cx_2^2 + Dx_1 + Ex_2 + F = 0$ .

**Example 1.2.1** (Circle). The equation  $x_1^2 + x_2^2 - 1 = 0$  defines the unit circle and  $(x_1/r)^2 + (x_2/r)^2 - 1 = 0 = x_1^2 + x_2^2 - r^2 = 0$  defines a circle of radius  $r$ .

**Example 1.2.2** (Ellipse). The equation  $(x_1/r_1)^2 + (x_2/r_2)^2 - 1 = 0$  defines an ellipse with radii  $r_1, r_2$ .

**Example 1.2.3** (Hyperbola). The standard hyperbola is defined as  $x_1^2 - x_2^2 - 1 = 0$ . Other hyperbolas include  $x_1x_2 = 1$ , or  $x_2^2 - x_1^2 - 1 = 0$ .

**Example 1.2.4** (Parabola).  $x_1^2 = x_2$ .

**Theorem 1.2.1** (Classification of conic sections). From three basic examples (the unit circle, standard hyperbola, and  $x_2 = x_1^2$ ), we can obtain ALL conic sections by affine linear transformations.

#### A SHORT PRIMER ON MATRIX MULTIPLICATION

If  $A, B$  are matrices of dimensions  $n \times \ell$  and  $\ell \times m$ , then  $AB$  is a  $n \times m$  matrix, with  $(AB)_{ij} = \sum_{k=1}^{\ell} A_{ik}B_{kj}$ , where  $M_{ij}$  is the element in the  $i$ th row and  $j$ th column of  $M$ .

First, we must define what an affine linear transformation is, which requires defining what a regular linear transformation is.

**Definition 1.2.4** (Linear transformation).  $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is a linear transformation if it can be expressed as a matrix where

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

That's kind of a non-intuitive definition. Here's another, more abstract and conceptual definition. This improved definition can also be generalized to any vector space in the domain, and any (not necessarily the same) vector space in the codomain.

**Definition 1.2.5** (Linear Transformation).  $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is a linear transformation if

1.  $T(cv) = cT(v)$ .
2.  $T(v + w) = T(v) + T(w)$ .

In other words,  $T$  preserves the linear structure of the space.

**Proposition 1.2.2.** These definitions are equivalent.

*Proof.* It is easy to verify that matrices produce linear transformations through simple computations. To prove that all linear transformations correspond with a matrix, we consider the action on basis vectors. If we fix a basis  $v_1, v_2$ , since  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 v_1 + x_2 v_2$ , we have

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = T(x_1 v_1 + x_2 v_2) = T(x_1 v_1) + T(x_2 v_2) = x_1 T(v_1) + x_2 T(v_2).$$

In this case, if  $T(v_1) = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$  and  $T(v_2) = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ , we then have

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = T\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

□

Now, we define affine transformations. But what does “affine” mean? Affine spaces like vector spaces with no origin, so you can't add or multiply, but only subtract.

You can make an affine space from the vector space by “forgetting the origin”, and a vector space from an affine space by setting an origin. Subtraction,  $-: \mathbb{A}^2 \times \mathbb{A}^2 \rightarrow \mathbb{R}^2$  gives a vector.

**Definition 1.2.6** (Affine Linear Transformation). An affine linear transformation is a transformation that preserves the affine structure.

Affine linear transformations are precisely linear transformations composed of a translation by a vector.

**Theorem 1.2.3.** Any quadratic form  $Q$  can be written in one of the following forms with a correct choice of coordinates

1.  $x_1^2 + x_2^2$



2.  $x_1^2 - x_2^2$
3.  $x_1^2$
4. 0

*Proof.* If  $Q(x_1, x_2) = ax_1^2 + bx_1 + cx_2^2$ , then assume  $a \neq 0$ . Then, we have

$$Q(x_1, x_2) = a\left(x_1 + \frac{b}{2a}x_2\right)^2 + \left(c - \frac{b^2}{4a}\right)x_2^2.$$

From this form, it is easy to convert into one of the preceding forms. □

As a corollary, you can prove the classification of conic sections theorem.

### 1.3 MORE ON LINEAR TRANSFORMATIONS IN 2D

As a reminder, to find a matrix for a transformation  $T$  with respect to a basis  $e_1, e_2$ , we have that  $(Te_1 \ Te_2) = (e_1 \ e_2)M$ , where  $M$  is the matrix representing  $T$  in this choice of basis. We then compute

$$T(ae_1 + be_2) = T\begin{pmatrix} e_1 & e_2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} e_1 & e_2 \end{pmatrix} M \begin{pmatrix} a \\ b \end{pmatrix}$$

Examples of linear transformations include

1. Stretching: For example  $T(e_1) = 2e_1, T(e_2) = \frac{1}{3}e_2$ .
2. Projection:  $T(e_1) = 0, T(e_2) = 1$ .

#### 1.3.1 OPERATIONS ON LINEAR TRANSFORMATIONS

**Definition 1.3.1** (Linear combination of transformations). If  $T_1$  and  $T_2$  are linear maps  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ , then  $c_1T_1 + c_2T_2$  is defined as the transformation such that

$$(c_1T_1 + c_2T_2)v = c_1T_1(v) + c_2T_2(v).$$

We can see that the matrix representation of  $c_1T_1 + c_2T_2$  behaves exactly how we'd expect it to—that is, if  $T_1, T_2$  are represented by  $M_1$  and  $M_2$ , then  $c_1T_1 + c_2T_2$  is represented by  $c_1M_1 + c_2M_2$ .

**Definition 1.3.2** (Composition). If  $T_1, T_2$  are transformations then their composition  $T_2 \circ T_1$  is the transformation making this diagram commutative

$$\begin{array}{ccccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & & \searrow & \nearrow & \\ & & & & \end{array}$$

In other words, it is the transformation defined as  $(T_2 \circ T_1)(v) = T_2(T_1(v))$ .

The matrix representing the composition is the matrix product.

**Definition 1.3.3** (Inverse). A transformation is invertible if there exists an inverse. Concretely, if  $T$  is a transformation, it is invertible if there exists a transformation  $S$  such that  $ST = TS = \text{id}$ . This transformation  $S$  is denoted  $T^{-1}$ .

How do we determine whether a transformation is invertible? And how do we compute the matrix of its inverse?

**Definition 1.3.4** (Determinant). The determinant of a matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , is defined as  $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$ .

**Proposition 1.3.1.** A transformation  $T$  is invertible iff  $\det([T]) \neq 0$ .

To compute the inverse of a matrix, we have

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

### 1.3.2 LINEAR GROUPS

Invertible transformations/matrices, along with the composition operation, satisfy the properties of a group. Composition is associative, and there exists an identity transformation/matrix represented as  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . This group is denoted as  $\text{GL}(\mathbb{R}^2)$  if we consider transformations and  $\text{GL}(2, \mathbb{R})$  for matrices.

An important subgroup of these groups is the orthogonal group  $\text{O}(2)$ . This consists of orthogonal transformations.

**Definition 1.3.5** (Orthogonal transformation). A transformation  $O$  is orthogonal if  $\langle v|w \rangle = \langle Ov|Ow \rangle$  for all vectors  $v, w$ . In other words, it preserves the inner product.

**Proposition 1.3.2.** A transformation is orthogonal iff  $O^T O = I$ .

### 1.4 ORTHOGONAL TRANSFORMATIONS

The set of orthogonal transformations on  $\mathbb{R}^2$  is denoted  $\text{O}(2)$ .

We first prove the following proposition:

**Proposition 1.4.1.** If a transformation preserves norms in an inner product space, then it preserves the inner product.

*Proof.*  $2\langle u|v\rangle = \|u+v\|^2 - \|u\|^2 - \|v\|^2$ , so the left hand side is preserved if the right hand side is preserved.  $\square$

A few more properties about orthogonal transformations are listed below:

**Proposition 1.4.2.** The matrix of an orthogonal transformation  $M$  with respect to an orthonormal basis satisfies the property  $M^T M = I$ .

*Proof.* We have that  $\langle Mv|Mw\rangle = \langle v|M^T Mw\rangle = \langle v|w\rangle$ . Now, by using projections onto the orthonormal basis vectors, we can prove that  $M^T M = I$ .  $\square$

A special subset of the orthogonal group is the special orthogonal group  $SO(2)$ , which induces the additional constraint that  $\det(O) = 1$ , and consists of plane rotations about the origin.

#### 1.4.1 COMPLEX NUMBERS

The imaginary unit  $i$  is defined as a solution of  $x^2 + 1 = 0$ , and the complex numbers are defined as numbers of the form  $a + bi$ .

Multiplication is then easily defined as  $(a + bi)(c + di) = ac - bd + (ad + bc)i$ .

Complex numbers have a length and an angle, and can be expressed in terms of the length and angle using the formula  $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ .

We can see that the group  $SO(2)$  is isomorphic to the unit complex numbers under multiplication.

### 1.5 THE MAIN THEOREMS OF LINEAR ALGEBRA

In linear algebra, we have four main theorems:

- Rank Theorem
- Inertia Theorem
- Orthogonal Diagonalization Theorem
- Jordan Normal Form

**Definition 1.5.1** (Classification Theorem). Classification theorems in linear algebra seek to classify and find “nice” representatives of the orbits of group actions.

**Theorem 1.5.1** (Rank theorem). Any linear equation of the form  $y = Ax$  can be simplified by a change of variables into the form  $\tilde{y} = M\tilde{x}$ , where  $M$  is a block matrix of the form

$$\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}.$$

**Theorem 1.5.2** (Inertia theorem). Any quadratic form can be rewritten after a change of variables into the form  $x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_q^2$ .

**Theorem 1.5.3.**

**Theorem 1.5.4** (Jordan Normal Form). Any matrix can be rewritten as a change of variables into a diagonal block matrix consisting of Jordan blocks, which are of the form

$$\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ 0 & & & \lambda \end{pmatrix}$$

## 1.6 MATRICES IN HIGHER DIMENSIONS

If  $A$  is a matrix, we use  $A_{ij}$  to refer to the  $i$ th row and  $j$ th column of the matrix.

Then, we are able to define operations on matrices:

**Definition 1.6.1** (Transpose). The transpose of a  $m \times n$  matrix  $A$ , denoted by  $A^T$ , is the  $n \times m$  matrix defined by  $(A^T)_{ij} = A_{ji}$ .

**Definition 1.6.2** (Multiplication). If  $A$  is a  $n \times m$  matrix and  $B$  is a  $m \times \ell$  matrix, then  $AB$  is a  $n \times \ell$  matrix with  $(AB)_{ij} = A_{i1}B_{1j} + \dots + A_{im}B_{mj}$ . Using summation, we have that

$$(AB)_{ij} = A_{ik}B_{kj} = \sum_{k=1}^m A_{ik}B_{kj}.$$

Why do we define multiplication of matrices like this? Multiplication of matrices correspond to compositions of linear transformations. An example of an application is a neural network, where “neurons” activity is based off matrices, and these are composed to get a final result.

## 1.6.1 DETERMINANT

**Definition 1.6.3** (Determinant). If  $A$  is an  $n \times n$  matrix, then the determinant of  $A$ , is a number defined as follows:

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^n A_{i\sigma(i)}.$$

$S_n$  denotes the symmetric group on  $n$  elements, or the permutations of  $n$  elements, which has size  $n!$ .  $\text{sgn}(\sigma)$  is the “sign” of the permutation, which is 1 if the permutation swaps the order of an even number of elements and  $-1$  otherwise.