Linear Algebra

Lecture Notes

Adriel Ong

August 13, 2021

Sylvester II Institute

Contents

1.0 Vector Spaces

1.1 Operations on Vector Spaces

Suppose a field K with elements k_i. We define *vectors* as elements $\vec{k}$ of a set $\mathbb{K}$ defined over K. These elements have k_i as their *components* such that for all $\vec{k}$ we have that

$$\vec{k} = (k_1, k_2, \ldots, k_n). \tag{1.1}$$

Elements of $\mathbb{K}$ follow these axioms.

Axiom 1.1. (Addition).

1. For all $\vec{x}, \vec{y} \in \mathbb{K}(\vec{x} + \vec{y} \in \mathbb{K})$ where $\vec{x} + \vec{y} = (x_1 + y_1, \ldots, x_n + y_n)$.
2. There exists $\vec{0} \in \mathbb{K}$ for all $\vec{x} \in \mathbb{K}$ such that $\vec{x} + \vec{0} = \vec{x}$, where $\vec{0} = (0, \ldots, 0)$.
3. For all $\vec{x}$ there exists $\vec{u}$ such that $\vec{x} + \vec{u} = \vec{0}$.
4. For all $\vec{x}, \vec{y} \in \mathbb{K}(\vec{x} + \vec{y} = \vec{y} + \vec{x})$.
5. For all $\vec{x}, \vec{y}, \vec{z} \in \mathbb{K}((\vec{x} + \vec{y}) + \vec{z} = \vec{y} + (\vec{x} + \vec{z}))$.

Axiom 1.2. (Scalar multiplication).

1. For all $\vec{x} \in \mathbb{K}, a \in K(a\vec{x} \in \mathbb{K})$.
2. For all $\vec{x} \in \mathbb{K}, a, b \in K((ab)\vec{x} = a(b\vec{x}))$.
3. For all $\vec{x} \in \mathbb{K}, a, b \in K((a + b)\vec{x} = a\vec{x} + b\vec{x})$.
4. For all $\vec{x} \in \mathbb{K}, 1 \in K(1\vec{x} = \vec{x})$.
5. For all $\vec{x}, \vec{y} \in \mathbb{K}, a \in K(a(\vec{x} + \vec{y}) = a\vec{x} = a\vec{y})$.

These operations lead to the following facts.

Theorem 1.1.1. The zero vector $\vec{0}$ is unique.

Proof. Suppose two zero vectors $\vec{0}_1$ and $\vec{0}_2$ belonging to $\mathbb{K}$. We have that

$$\vec{0}_1 + \vec{0}_2 = \vec{0}_1$$
$$\vec{0}_2 + \vec{0}_1 = \vec{0}_2$$

where by commutativity and the identity element, $\vec{0}_1 = \vec{0}_2$. □

Theorem 1.1.2. Each vector's negative element is unique.

Proof. Suppose that $\vec{x} \in \mathbb{K}$ has two negative elements, y_1 and y_1. Then,

$$\vec{y}_1 = \vec{y}_1 + \vec{0} = \vec{y}_1 + (\vec{x} + \vec{y}_2) = (\vec{y}_1 + \vec{x}) + \vec{y}_2 = \vec{0} + \vec{y}_2 = \vec{y}_2.$$

□

From now, we denote the negative element of $\vec{x}$ by $-\vec{x} = -1\vec{x}$ (prove that this relation holds). The sum denoted by $\vec{x} + (-\vec{y}) \equiv \vec{x} - \vec{y}$ is called the vectors' *difference*.

Theorem 1.1.3. For all $\vec{x} \in \mathbb{K}$, $0\vec{x} = \vec{0}$.

Proof. By negative element, we have that

$$\vec{0} = \vec{x} - \vec{x} = (1 - 1)\vec{x} = 0\vec{x}.$$

□

1.2 Linear Dependence

Definition 1.2.1. (Linear Combinations). Suppose that we can write $\vec{x} \in \mathbb{K}$ as

$$\vec{x} = \sum_{i=1}^{n} a_i \vec{x}_i$$

for $a_i \in K$ and $\vec{x}_i \in \mathbb{K}$. Then we call $\vec{x}$ a *linear combination* of vectors $\vec{x}_i$.

Definition 1.2.2. (Linear Dependence). Write $\vec{0}$ as

$$\vec{0} = \sum_{i=1}^{n} a_i \vec{x}_i.$$

If there exists $a_i \neq 0$, then we call the vectors $\vec{x}_i$ as *linearly dependent*. If for all $i = 1, 2, \ldots$ we have that $a_i = 0$, then we call the vectors as *linearly independent*.

Lemma 1.2.3. Suppose for n vectors $\vec{x}_i$ we have that

$$\vec{0} = \sum_{i=1}^{n} a_i \vec{x}_i$$

where $k < n$ vectors are linearly dependent. Then, the whole system is linearly dependent.

Proof. Some a_k exists such that $a_k \neq 0$. Thus, whether or not the succeeding $n - k$ vectors are linearly independent, the whole system becomes linearly dependent since one nonzero coefficient exists. □

Lemma 1.2.4. Vectors $\vec{x}_i, i = 1, \ldots, n$ are linearly dependent if and only if we can write one vector as a linear combination of the others.

Proof. Let the vectors $\vec{x}_i$ be linearly dependent. Then by commutativity, we can subtract one vector $\vec{x}_k$ from both sides and divide by its coefficient a_k. Thus,

$$\vec{x}_k = \sum_{i=1}^{n-1} \frac{a_i}{a_k} x_i.$$

For the converse, simply transpose the vector expressed as a linear combination. We note that its coefficient is $1 \neq 0$, making the vectors linearly dependent. □

1.3 Span, Basis, and Dimension

Definition 1.3.1. (Span and Basis). Let $\vec{x} \in \mathbb{K}$ be arbitrary, and suppose that we can write it as a linear combination of vectors $\vec{x}_i, 1 \leq i \leq m$. Then we say that the vectors $\vec{x}_i$ *span* the vector space $\mathbb{K}$. If n linear independent vectors span a vector space, then we call them a *basis* for that space.

Theorem 1.3.2. For all vectors $\vec{x} \in \mathbb{K}$ there exists unique coefficients for their basis representation.

Proof. For arbitrary $\vec{x} \in \mathbb{K}$ suppose two representations

$$\vec{x} = \sum_{i=1}^{n} a_i x_i$$
$$\vec{x} = \sum_{i=1}^{n} b_i x_i.$$

Take the difference

$$\vec{0} = \sum_{i=1}^{n} a_i x_i - \sum_{i=1}^{n} b_i x_i = \sum_{i=1}^{n} (a_i - b_i) \vec{x}_i.$$

By linear independence, the coefficients $(a_i - b_i)$ must be all 0, such that for all $i = 1, \ldots, n$ we have that $a_i = b_i$. $\square$

Definition 1.3.3. (Dimension). Suppose that for a vector space $\mathbb{K}$, only up to n vectors are linearly independent. Then we call n the *dimension* of $\mathbb{K}$.

Theorem 1.3.4. The number of basis vectors of a vector space equals the dimension.

Proof. Suppose n basis vectors $\vec{x}_i$ for a vector space $\mathbb{K}$. If we can find one more vector such that $n+1$ linearly independent vectors exist, then we can find an arbitrary vector such that the $n+2$ vectors are linearly dependent. This $\vec{x}_{n+2}$, however, can be written as a linear combination of the $n+1$ vectors. Since $\vec{x}_{n+2}$ was arbitrary, we find a basis of $n+1$ vectors, contradicting that our basis had n vectors.

Conversely, let our dimension for $\mathbb{K}$ be n. For n linearly independent vectors, any other arbitrary vector $\vec{x}$ causes the system to be linearly dependent. We can write $\vec{x}$ as a linear combination of the other vectors. Since $\vec{x}$ was arbitrary, we have a basis of n vectors for $\mathbb{K}$. $\square$

1.4 Linear Subspaces

Definition 1.4.1. (Linear Subspaces). For some subset $\mathbb{L} \subseteq \mathbb{K}$, suppose that

1. $\vec{x}, \vec{y} \in \mathbb{L} \implies \vec{x} + vecy \in \mathbb{L}$.
2. $\vec{x} \in \mathbb{L}, a \in K \implies a\vec{x} \in \mathbb{L}$.

Then we call $\mathbb{L}$ a *linear subspace* of $\mathbb{K}$. One can verify that these two conditions make $\mathbb{L}$ a vector space.

2.0 Determinants

2.1 Number Fields

We suppose a set K of objects called *numbers*, subjected to four arithmetic operations which also give elements of K. These operations have four properties, called the *Field Axioms*.

Axiom 2.1. (Addition). To every $\alpha, \beta \in K$ corresponds a unique *sum* $\alpha + \beta \in K$, where

1. $\forall \alpha, \beta \in K, \alpha + \beta = \beta + \alpha$ (commutativity)
2. $\forall \alpha, \beta, \gamma \in K, (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ (associativity)
3. $\exists 0 \in K$ s.t. $0 + \alpha = \alpha, \forall \alpha \in K$ (identity element)
4. $\forall \alpha \in K \exists \gamma \in K$ s.t. $\alpha + \gamma = 0$ (negative element)

From Property 4, we can define the *difference* $\beta - \alpha$ as $\beta + \gamma$, with γ the solution to $\alpha + \gamma = 0$.

Axiom 2.2. (Multiplication). To every $\alpha, \beta \in K$ corresponds a unique *product* $\alpha\beta \in K$ (sometimes denoted $\alpha \cdot \beta$), where

1. $\forall \alpha, \beta \in K, \alpha\beta = \beta\alpha$ (commutativity)
2. $\forall \alpha, \beta, \gamma \in K, (\alpha\beta)\gamma = \alpha(\beta\gamma)$ (associativity)
3. $\exists (1 \neq 0) \in K$ s.t. $1 \cdot \alpha = \alpha, \forall \alpha \in K$ (identity element)
4. $\forall (\alpha \neq 0) \in K \exists \gamma \in K$ s.t. $\alpha\gamma = 1$ (reciprocal element)

From Property 4, we can define the *quotient* $\frac{\beta}{\alpha}$ as $\beta\gamma$, with γ the solution to $\alpha\gamma = 1$.

Axiom 2.3. (Distributive property).

1. $\forall \alpha, \beta, \gamma \in K, \alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$

The set K with these operations defined on its elementsis called a *field*.

Numbers $1, 1+1=2, 2+1=3$, etc are said to be in the set of *natural numbers* $\mathbb{N}$—none are equal to 0. The *integers* $\mathbb{Z}$ in K are all naturals with 0 and their negative elements. The rationals $\mathbb{Q}$ in K are the set of quotients $\frac{p}{q}, p, q \in \mathbb{Z}$ and $q \neq 0$.

The fields K and K' are *isomorphic* if one can set an injection between them such that one can associate to every sum (or product) of numbers in K another number that is the sum (or product) of corresponding numbers in K'.

We list the most commonly encountered fields:

- Rational numbers $\mathbb{Q}$: of quotients $\frac{p}{q}, (p, q \neq 0) \in \mathbb{Z}$.

  The set of integers Z is not a field since it fails to satisfy Multiplication Property 4. Every field K also has a *subfield*—subset satisfying field axioms—isomorphic to $\mathbb{Q}$.

- Real numbers $\mathbb{R}$: set of all points on the real line in geometry.

  One may build an axiomatic treatment of $\mathbb{R}$ by supplementing the field axioms with order axioms and the axiom of completeness.

- Complex Numbers $\mathbb{C}$: of form $a + ib, (a, b \in \mathbb{R})$, and i defined as $i^2 = -1$. The field $\mathbb{C}$ has the defined operations:

  - Addition: $(a_1 + ib_1) + (a_2 + ib_2) = (a_1 + a_2) + i(b_1 + b_2)$.

  - Multiplication: $(a_1 + ib_1)(a_2 + ib_2) = (a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1)$.

  Numbers of form $(a + i \cdot 0)$ reduce to operations on real numbers. Numbers of form $0 + ib$ are purely imaginary: $i^2 = i \cdot i = (0 + i1)(0 + i1) = -1$. The complex numbers also have subfields isomorphic to $\mathbb{R}$.

Remark. By the *Fundamental Theorem of Algebra,* any polynomial of form

$$z^n + a_1 z^{n-1} + \ldots + a_n = 0$$

has solutions in $\mathbb{C}$—not necessarily in $\mathbb{R}$. For example, $z^2 + 1 = 0$ has no real solution—the imaginary unit i solves it.
From now on, we will use K to denote any arbitrary field. Anything that holds for K also holds for $\mathbb{R}$ and $\mathbb{C}$.

2.2 Problems in the Theory of Linear Functions

Definition 2.2.1. A *system of equations* takes the following form:

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \ldots + a_{2,n}x_n &= b_2 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \ldots + a_{m,n}x_n &= b_m \end{aligned} \tag{2.1}$$

The values x_i denote *unknowns*, or elements of K to be determined. Quantities a_i denote *coefficients* of system 2.1. Quantities b_i denote *constants* in the system. We assume both coefficients and constants to be known.

Definition 2.2.2. A *solution* to system 2.1 denotes the set $c_1, c_2, \ldots, c_n \in K$ which, when substituted for the unknowns, turns the equations into identities.

Not all systems of the above form have solutions. For example, take the system

$$\begin{aligned} 2x_1 + 3x_2 &= 5 \\ 2x_1 + 3x_2 &= 6 \end{aligned}$$

No solution exists at all.

Definition 2.2.3. Systems of form 2.1 with solutions are called *compatible*. Otherwise, they are *incompatible*.

Compatible systems may have one or more solutions. We can distinguish multiple solutions with superscripts: $c_1^{(1)}, c_2^{(1)}, \ldots, c_n^{(1)}, c_1^{(2)}, c_2^{(2)}, \ldots, c_n^{(2)}$.

If multiple solutions have it so that $c_i^{(1)} \neq c_i^{(2)}$ for any $i = 1, 2, \ldots, n$, then we call them *distinct*.

For example, the system

$$\begin{aligned} 2x_1 + 3x_2 &= 0 \\ 4x_1 + 6x_2 &= 0 \end{aligned}$$

has distinct solutions $c_1^{(1)} = c_2^{(1)} = 0$ and $c_1^{(2)} = 3, c_2^{(2)} = -2$.

Definition 2.2.4. We call systems of form 2.1 with a unique solution *determiante*. Systems with multiple solutions are *indeterminate*.

We have the following basic problems in studying a system of linear equations:

1. Ascertain whether the system is compatible or incompatible
2. If compatible, whether it is determinate
3. If determinate, find the unique solution
4. If indeterminate, describe the set of solutions

The basic tool of doing such is the theory of determinants.

2.3 Determinants of Order n

We shall primarily work with square matrices.

Definition 2.3.1. A *square matrix* is an array of n^2 numbers $(a_{ij}), i, j = 1, 2, \dots, n$, and $a_{ij} \in K, \forall i, j$.

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots, & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix} \tag{2.2}$$

We call its number of rows and columns the matrix *order*. The numbers a_{ij} are its *elements*. The first index i denotes row, the second index j column. Elements $a_{ij}, i = j$ form a matrix's *principal diagonal*.

Consider any product of n elements appearing in different rows and columns of matrix 2.2:

$$a_{\alpha_1,1} a_{\alpha_2,2}, \dots, a_{\alpha_n,n} \tag{2.3}$$

We can always choose the element in the first column of matrix 2.2 for our first factor a_{α_1}. Here we denote by α_1 the element's row. We get by hypothesis that factors $a_{\alpha_i,i}$ appear in different rows of matrix 2.2. The row indices α_i then represent some permutation of $1, 2, \dots, n$.

Definition 2.3.2. We mean by an *inversion* of the sequence (α_i) an arrangement of two indices such that larger indices come before smaller ones.
We count the total number of inversions by $N(\alpha_1, \alpha_2, \dots, \alpha_n)$.

Take the sequence $(2, 1, 4, 3)$. Two inversions exist: $(2, 1)$, and $(4, 3)$. We then have $N(2, 1, 4, 3) = 2$.

Consider also the sequence $(4, 3, 1, 2)$. We count five inversions: $(4, 3)$, $(4, 1)$, $(4, 2)$, $(3, 1)$, and $(3, 2)$, so that $N(4, 3, 1, 2) = 5$.

We go back now to the product of n elements 2.3. We multiply the product by $(-1)^{N(\alpha_1,\ldots,\alpha_n)}$, so that even inversions put a plus sign and odd inversions put a minus.

The total number of products from elements of order n matrices is equal to the number of permutations for numbers $1, 2, \ldots, n$, which is $n!$. We now arrive to the definiition of a determinant:

Definition 2.3.3. The *determinant* D of a matrix with form 2.2 is the algebraic sum of $n!$ products of form 2.3. Each product is multiplied by a rule-determined sign:

$$D = \sum (-1)^{N(\alpha_1,\ldots,\alpha_n)} a_{\alpha_1,1} a_{\alpha_2,2} \ldots a_{\alpha_n,n}$$

For order 2 determinants,

$$\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{2,1}a_{1,2}$$

Order 3 determinants,

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1}a_{2,2}a_{3,3} - a_{2,1}a_{1,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3}$$
$$- a_{3,1}a_{2,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} - a_{1,1}a_{3,2}a_{2,3}$$

Determinants become useful in solving systems like 2.2

$$a_{1,1}x_1 + a_{1,2}x_2 = b_1$$
$$a_{2,1}x_1 + a_{2,2}x_2 = b_2$$

by obtaining the formulas

$$x_1 = \frac{b_1a_{2,2} - b_2a_{1,2}}{a_{1,1}a_{2,2} - a_{2,1}a_{1,2}}, x_2 = \frac{b_2a_{1,1} - b_1a_{2,1}}{a_{1,1}a_{2,2} - a_{2,1}a_{1,2}}$$

Assuming a nonzero denominator, we can represent the numerators and denominator as

$$b_1a_{2,2} - b_2a_{1,2} = \begin{vmatrix} b_1 & a_{1,2} \\ b_2 & a_{2,2} \end{vmatrix}, a_{2,1}b_2 - a_{2,1}b_1 = \begin{vmatrix} a_{1,1} & b_1 \\ a_{2,1} & b_2 \end{vmatrix}$$
$$a_{1,1}a_{2,2} - a_{2,1}a_{1,2} = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix}$$

Similar formulas hold for solutions of systems with arbitrary number of unknowns.

2.4 Properties of Determinants

Definition 2.4.1. The *transposition operation* $(a_{ij})^T = (a_{ji})$ obtained from a matrix of form 2.2 comes from interchanging rows and columns with identical indices. One may see that transposes have identical determinants.

Taking the transpose of a determinant is geometrically equivalent to a $180°$ rotation in space about the principal diagonal. All inversions remain, so term signs also remain.

Since this property establishes equivalence between rows and columns of a matrix, further theorems will be proved for columns only.

Proposition 2.4.2. (Antisymmetry). Interchanging two adjacent columns of a determinants changes its sign.

Proof. Interchanging two adjacent columns changes the number of determinant inversions by one, changing whether it is even or odd. The sign of the determinant then changes. □

Remark. Interchanging nonadjacent columns with m columns in between causes $m + (m+1)$ sign changes in the determinant. The resulting sign will always be opposite the original, since $m + (m+1) = 2m+1$, an odd number.

Corollary 2.4.3. A determinant with two identical columns vanishes.

Proof. A determinant with two identical columns has at least one pair of identical terms. Interchanging two identical columns doesn't change the determinant D since two identical terms exist. However, the determinant's sign must change with column interchanges. Since $D = -D$, it must be that $D = 0$. □

Definition 2.4.4. A matrix column a_{ij} is a *linear combination* of two columns if for all elements of the column,

$$a_{ij} = \lambda b_i + \mu c_i, i = 1, 2, \ldots, n$$

for fixed numbers λ and μ.

Theorem 2.4.5. (Linear Property of Determinants). If a matrix column a_{ij} is a linear combination $a_{ij} = \lambda b_i + \mu c_i, i = 1, 2, \ldots, n$ of two columns, then the determinant D is a linear combination of two determinants:

$$D = \lambda D_1 + \mu D_2 \tag{2.4}$$

with D_1 and D_2 having the same columns as D except for the j-th; determinant D_1 has elements of b_i while D_2 has elements of c_i for the j-th column.

Proof. We can write each term of the determinant D as

$$\begin{aligned} a_{\alpha_1,1}a_{\alpha_2,1}\ldots a_{\alpha_j,j}\ldots a_{\alpha_n,n} &= a_{\alpha_1,1}a_{\alpha_2,1}\ldots(\lambda b_{\alpha_j} + \mu c_{\alpha_j})\ldots a_{\alpha_n,n} \\ &= \lambda b_{\alpha_j} a_{\alpha_1,1}a_{\alpha_2,1}\ldots a_{\alpha_n,n} + \mu c_{\alpha_j} a_{\alpha_1,1}a_{\alpha_2,1}\ldots a_{\alpha_n,n} \end{aligned}$$

All the first terms of the determinant add to λD_1, while the second terms add to μD_2. □

We can write the determinant formula more conveniently. Let D be a fixed determinant, and $D_i(p_i)$ the determinant obtained by replacing the j-th column of D by $p_i, i = 1, \ldots, n$. The determinant 2.4 takes the form

$$D_j(\lambda b_i + \mu c_i) = \lambda D_i(b_i) + \mu D_i(c_i).$$

One may easily extend the linear property of determinants to cases where every element of the j-th column is a linear combination of any other term

$$a_{ij} = \lambda b_i + \mu c_i + \ldots + \tau f_i$$

so that we obtain the determinant

$$\begin{aligned} D_j a_{ij} &= D_j(\lambda b_i + \mu c_i + \ldots + \tau f_i) \\ &= \lambda D_j(b_i) + \mu D_j(c_i) + \ldots + \tau D_j(f_i). \end{aligned}$$

Corollary 2.4.6. Any common factor of a column of a determinant can be factored out of the determinant.

Proof. If $a_{i,j} = \lambda b_{i,j}$, then we have that $D_j(a_{i,j}) = D_j(\lambda b_{i,j}) = \lambda D_j(b_{i,j})$ □

Corollary 2.4.7. If a determinant column consists entirely of 0s, then the determinant vanishes.

Proof. We can factor the 0 out of the determinant column

$$D_j(0b_{i,j}) = 0D_j(b_{i,j}) = 0.$$

□

Theorem 2.4.8. A determinant value is not changed by adding elements of a column multiplied by an arbitrary number to elements of another column.

Proof. Suppose that we multiply the kth column by λ and add it to the jth column. By Linear Property of Determinants, we can write D_j as

$$D_j(\lambda a_{ik} + a_{ij}) = \lambda D_j(a_{ik}) + D_j(a_{ij})$$

The column j consists of elements a_{ik} and is identical to the kth column, causing the determinant to vanish. □