LECTURE NOTES

MAT1341 Introduction to Linear Algebra
Fall 2019

Based on the textbook: *Vector Spaces First* by Thierry Giordano, Barry Jessup and Monica Nevins.

CHAPTER 2. VECTOR GEOMETRY

1. INTRODUCING R^N AND VECTORS IN R^N

R: the set of real numbers, representing points on a directed line with an origin and a unit, or a 1-dimensional vector.

R^2 = $\{(x_1, x_2) \mid x_1, x_2 \in$ R^2$\}$: **t**he set of ordered pair of real numbers, representing points on a Cartesian plain, or a 2-dimensional vector.

R^3 = $\{(x_1, x_2, x_3) \mid x_1, x_2, x_3 \in$ R^3$\}$: the set of ordered triple of real numbers, representing points in a Cartesian space, or a 3-dimensional vector.

Geometrically, in R, R^2 or R^3, a vector can be represented by an arrow with the *tail* at the origin and the *head* at the point at the space. Sometimes, we may see vectors whose tail is not at the origin. In such cases, this vector equals the vector with tail at the origin that has the same length and direction. For example, a vector in R^2 with head (a_1, b_1) and tail (a_2, b_2) equals the vector $(a_1 - a_2, b_1 - b_2)$.

Generalizing, R^n = $\{(x_1, x_2, \ldots, x_n) \mid x_1, x_2, \ldots, x_n \in$ R$\}$ is the set of ordered n-tuples of real numbers, representing points in an n-dimensional space, or vectors in an n-dimensional space. An ordered n-tuple $\mathbf{v} = (x_1, x_2, \ldots, x_n)$ of R^n is called an (*n-dimensional*) *vector*, and x_i is the *i-th component* of $\mathbf{v}$.

If $x_i = 0$, $i = 1, 2, \ldots, n$, vector $(x_1, x_2, \ldots, x_n)$ is the *zero vector* in R^n, denoted by $\mathbf{0}$ or $\mathbf{0}_n$.

Two vectors $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ are said to be *equal*, denoted by $(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_n)$ if $x_i = y_i$, $i = 1, 2, \ldots, n$.

Since an ordered n-tuple of real numbers can be regarded either as a vector in R^n or a point in R^n, which is the head of the vector, we may **identify** ordered n-tuples, vectors in R^n, or points in R^n.

An n-dimensional vector is denoted by $(x_1, x_2, \dots x_n)$ or $[x_1, x_2, \dots x_n]$ (*row vectors*), or $\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$ or $\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$ (*column vectors*).

With the round bracket, it is called the *vector form*, and, with the square bracket, it is called the *matrix form*. Use the *transpose operator T*, we can convert a row vector to a column vector and vice versa. For instance,

$$(x_1, x_2, \dots x_n)^T = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}^T = [x_1, x_2, \dots x_n].$$

In R^3, $\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, are *unit vectors* in the direction of x-axis, y-axis, and z-axis, respectively. They are usually denoted by column vectors.

The coefficients in an n-variable linear equation $c_1x_1 + c_2x_2 + \dots + c_nx_n = d$ can be regarded as components of a vector $(c_1, c_2, \dots, c_n)$ in R^n. When the constant on the right-hand side of the equation is included, $(c_1, c_2, \dots, c_n, d)$ can be regarded as a vector in R^{n+1}.

2. ADDITION AND SCALAR MULTIPLICATION OF VECTORS

Definitions

The *addition* of two vectors in R^n finds the *sum* of these vectors:

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n).$$

The definition of the sum of two vectors can be justified by the head-tail rule or the parallelogram rule in R^2or R^3.

The *scalar multiplication* of a scalar and a vector in R^n defines the product as

$c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n)$.

If two vectors **u** and **v** have the relation $\mathbf{u} = c\mathbf{v}$, where c is a scalar, $c \neq 0$, **u** and **v** are said to be *parallel* or *collinear*, (i.e., they are in the same or opposite direction).

When $c = -1$, $(-1)\,(x_1, x_2, \ldots, x_n) = (-x_1, -x_2, \ldots, -x_n)$, denoted by $-(x_1, x_2, \ldots, x_n)$, is called the *negation* of $(x_1, x_2, \ldots, x_n)$.

Let **u** and **v** be two (n-dimensional) vectors. Then $\mathbf{u} - \mathbf{v}$ is defined to be $\mathbf{u} + (-\mathbf{v})$. Hence,

$(x_1, x_2, \ldots, x_n) - (y_1, y_2, \ldots, y_n) = (x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n)$.

Geometrically, the left hand side of this equality is a vector from a point $\mathbf{p}_2 = (y_1, y_2, \ldots, y_n)$ to a point $\mathbf{p}_1 = (x_1, x_2, \ldots, x_n)$ and the right-hand side is a vector from the origin to the point $\mathbf{p} = (x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n)$. These two vectors have the same direction and the same length.

Properties

Suppose **u** and **v** are vectors in R^n, and c and d are scalars.

(i) $\mathbf{u} + \mathbf{0} = \mathbf{u}$, $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$, $1\mathbf{u} = \mathbf{u}$,

(ii) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$, (the commutative law)

(iii) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$, $c(d\mathbf{u}) = (cd)\mathbf{u}$, (the associative law)

(iv) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$, $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$, (the distributive law)

3. LINEAR COMBINATION

Definition

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ be k vectors, in R^n and $c_1, c_2, \ldots c_k$ are scalars, then

$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k$

is called a *linear combination* of these vectors. A linear combination of k vectors in R^n is a vector in R^n.

Examples

1. $(-3)(1, 2, 0) + 2(2, 1, 3) = (1, -4, 6)$.

2. The vector (1, 0, 0) is not a linear combination of (1, 2, 0) and (2, 1, 3) because, assume to the contrary, we have $a(1, 2, 0) + b(2, 1, 3) = (1, 0, 0)$. Then $a + 2b = 1$, $2a + b = 0$, $3b = 0$. From the last equation, $b = 0$, then the first equation gives $a = 1$, and the second equation gives $a = 0$, a contradiction.

4. THE DOT PRODUCT

Definition

The *dot product* (also called the *inner product* or the *scalar product*) of two vectors $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ in R^n is defined to be

$$(x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = x_1y_1 + x_2y_2 + \ldots + x_ny_n.$$

Note that the result is a scalar (a number)!

$\sqrt{\mathbf{u}\cdot\mathbf{u}} = \| \mathbf{u} \|$ is called the *norm* of vector **u**. Let $\mathbf{u} = (x_1, x_2, \ldots, x_n)$. Then

$\| \mathbf{u} \| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$. In R^2 or R^3, this is the *length* of the vector **u**.

Examples

1. $(1, 2, 3) \cdot (-1, 2, 1) = -1 + 4 + 3 = 6$.

2. $(0, 1, 2, -1) \cdot (2, 3, -1, 1) = 0 + 3 - 2 - 1 = 0$.

3. $\| (1, -4, 8) \| = \sqrt{1+16+64} = \sqrt{81} = 9$.

Properties

Let **u** and **v** be two vectors in R^n, and let c be a scalar.

(i) If **u** or **v** is a zero vector, then $\mathbf{u} \cdot \mathbf{v} = 0$.

(ii) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$. (the commutative law)

(iii) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$. (the distributive law)

(iv) $\mathbf{u} \cdot (c\mathbf{v}) = c(\mathbf{u} \cdot \mathbf{v})$.

(v) $\| \mathbf{u} \| = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

(vi) $\| c\mathbf{u} \| = | c | \, \| \mathbf{u} \|$.

(vi) Cauchy-Schwarz Inequality $| \mathbf{u} \cdot \mathbf{v} | \le \| \mathbf{u} \| \, \| \mathbf{v} \|$. (See textbook for a proof.)

(vii) Triangle inequality $\| \mathbf{u} + \mathbf{v} \| \le \| \mathbf{u} \| + \| \mathbf{v} \|$.

$\| \mathbf{u} + \mathbf{v} \|^2 = (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \| \mathbf{u} \|^2 + 2(\mathbf{u} \cdot \mathbf{v}) + \| \mathbf{v} \|^2 \le \| \mathbf{u} \|^2 + 2\| \mathbf{u} \| \| \mathbf{v} \| + \| \mathbf{v} \|^2$
$= (\| \mathbf{u} \| + \| \mathbf{v} \|)^2$.

Geometric Interpretation

In R^2 or R^3, the norm of a vector **u** is the length of **u**. If **u** and **v** are regarded as two points, (i.e., the heads of vectors **u** and **v**), then $\| \mathbf{u} - \mathbf{v} \|$ is the distance between these two points.

In R^2 or R^3, $\mathbf{u} \cdot \mathbf{v} = \| \mathbf{u} \| \, \| \mathbf{v} \| \cos \theta$, where θ is the angle between **u** and **v**.

Proof. Let **u** and **v** be two vectors. Suppose the angle between these two vectors is θ. Then the vector from the tail of **v** to the tail of **u** is $\mathbf{w} = \mathbf{u} - \mathbf{v}$.

By the cosine law, $\| \mathbf{u} - \mathbf{v} \|^2 = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 - 2\| \mathbf{u} \| \, \| \mathbf{v} \| \cos \theta$.

On the other hand, $\| \mathbf{u} - \mathbf{v} \|^2 = (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) = \mathbf{u} \cdot \mathbf{u} - 2\mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v} = \| \mathbf{u} \|^2 - 2\mathbf{u} \cdot \mathbf{v} + \| \mathbf{v} \|^2$.

Therefore, $\| \mathbf{u} \| \, \| \mathbf{v} \| \cos \theta = \mathbf{u} \cdot \mathbf{v}$.

Use this relation, we can find the angle between two vectors.

Example

Let $\mathbf{u} = (\sqrt{3}, 1)$, and $\mathbf{v} = (\sqrt{3} - 1, \sqrt{3} + 1)$. Then $\mathbf{u} \cdot \mathbf{v} = (3 - \sqrt{3}) + (\sqrt{3} + 1) = 4$.
$\| \mathbf{u} \| = 2$, $\| \mathbf{v} \| = \sqrt{8}$. Then $\cos \theta = \dfrac{4}{2\sqrt{8}} = \dfrac{1}{\sqrt{2}}$. $\theta = \dfrac{\pi}{4}$.

Generalizing this idea, in R^n we define angle θ to be the *angle* between vectors **u** and **v** in R^n if

$\cos\theta = \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$ and $0 \le \theta \le \pi$.

When $\mathbf{u} \cdot \mathbf{v} = 0$, the angle between $\mathbf{u}$ and $\mathbf{v}$ is $\dfrac{\pi}{2}$. We said that $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal* (i.e., *perpendicular* to each other).

Example. Since $(0, 1, 2, -1) \cdot (2, 3, -1, 1) = 0$, vectors $(0, 1, 2, -1)$ and $(2, 3, -1, 1)$ are orthogonal.

Orthogonal Projection

Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors in R^n. Vector $\mathbf{w} = \|\mathbf{v}\|\cos\theta\dfrac{\mathbf{u}}{\|\mathbf{u}\|} = \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|}\dfrac{\mathbf{u}}{\|\mathbf{u}\|} = \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|^2}\mathbf{u}$ is the the (*orthogonal*) *projection* of $\mathbf{v}$ onto $\mathbf{u}$, denoted by $\mathrm{proj}_{\mathbf{u}}\,\mathbf{v}$. The length of $\mathbf{w}$ is $\|\,\mathrm{proj}_{\mathbf{u}}\,\mathbf{v}\,\| = |\,\|\mathbf{v}\|\cos\theta| = \dfrac{|\mathbf{u}\cdot\mathbf{v}|}{\|\mathbf{u}\|}$. If $\mathbf{u} \cdot \mathbf{v} > \mathbf{0}$, w and $\mathbf{u}$ are in the same direction; if $\mathbf{u} \cdot \mathbf{v} < \mathbf{0}$, w and $\mathbf{u}$ are in the opposite direction.

Let $\mathbf{v}_1 = \mathbf{v} - \mathrm{proj}_{\mathbf{u}}\,\mathbf{v}$. Then $\mathbf{v}_1 \cdot \mathbf{u} = \left(\mathbf{v} - \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|^2}\mathbf{u}\right)\cdot\mathbf{u} = \mathbf{v}\cdot\mathbf{u} - \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|^2}(\mathbf{u}\cdot\mathbf{u}) = 0$.

Hence, $\mathbf{v}_1$ is orthogonal to $\mathbf{u}$. In this way, we can *decompose* vector $\mathbf{v}$ into two components $\mathbf{v}_1$ and $\mathbf{v}_2$ (i.e., express $\mathbf{v}$ as a sum of two vectors, $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$) such that $\mathbf{u}$ and $\mathbf{v}_1$ are orthogonal, and $\mathbf{u}$ and $\mathbf{v}_2 = \mathrm{proj}_{\mathbf{u}}\,\mathbf{v}$ are collinear.

Example. Let $\mathbf{u} = (1, 2, -2)$, $\mathbf{v} = (3, -6, 0)$. Decompose $\mathbf{v}$ into two vectors so that one is orthogonal to $\mathbf{u}$, and the other is collinear with $\mathbf{u}$.

Since $\mathbf{u} \cdot \mathbf{v} = -9$ and $\|\mathbf{u}\|^2 = 9$, $\mathbf{v}_1 = \mathrm{proj}_{\mathbf{u}}\,\mathbf{v} = -\mathbf{u} = (-1, -2, 2)$ is collinear with $\mathbf{u}$, and $\mathbf{v}_2 = \mathbf{v} - \mathbf{v}_1 = (4, -4, -2) =$ is orthogonal to $\mathbf{u}$.