MA1508E: LINEAR ALGEBRA FOR ENGINEERING

Lecture 5 Notes

References

1. Elementary Linear Algebra: Application Version, Section 3.1-3.2, 3.4, 6.1-6.2
2. Linear Algebra with Application, Section 4.1-4.2

3 Vectors in Euclidean Spaces

3.1 Introduction to Euclidean Spaces

A (real) <u>n-vector</u> (or vector) is a collection of n ordered real numbers,

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \text{ where } v_i \in \mathbb{R} \text{ for } i = 1, ..., n.$$

The collection of all n-vectors is called a <u>vector space</u>, which in this course, is synonymous to the <u>Euclidean n-space</u>. It is denoted as $\mathbb{R}^n$,

$$\mathbb{R}^n = \left\{ \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \middle| v_i \in \mathbb{R} \text{ for } i = 1, ..., n. \right\}$$

Remark.

1. We can similarly define complex n-vectors, denoted as $\mathbb{C}^n$. Later, we will also define vector-valued functions, which are vectors with functions entries.
2. Strictly speaking, the above is called a column vector. A <u>column vector</u> is a $n \times 1$ matrix (column matrix). A <u>row vector</u> is a $1 \times n$ matrix (row matrix), $\mathbf{v} = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}$.
3. The space of all column vectors is isomorphic to the space of row vectors (meaning they have all the same properties we care about, except for only how they look). So in most subjects, they are both denoted as $\mathbb{R}^n$ (in fact any real n-dimensional space is isomorphic to $\mathbb{R}^n$, but we digressed).
4. In this course, if not explicitly mentioned, a vector would mean a column vector.
5. Geometrically, we can think of $\mathbb{R}^2$ as a plane, and $\mathbb{R}^3$ as our physical universe (without considering time).
6. A vector $\mathbf{v}$ can be interpreted as an arrow, with the tail placed at the origin $\mathbf{0}$, and the head of the arrow at $\mathbf{v}$, or it could represent a point in the Euclidean n-space.

3.2 Vectors addition and Scalar Multiplication

Since vectors are (column or row) matrices, the properties of matrix addition and scalar multiplication holds for vectors. For any vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ and scalars $a, b \in \mathbb{R}$,

(i) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$,

(ii) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$,

(iii) $\mathbf{0} + \mathbf{v} = \mathbf{v}$,

(iv) $\mathbf{v} - \mathbf{v} = \mathbf{0}$,

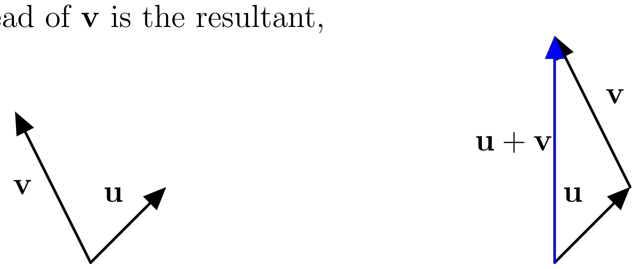
(v) $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$,

(vi) $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$,

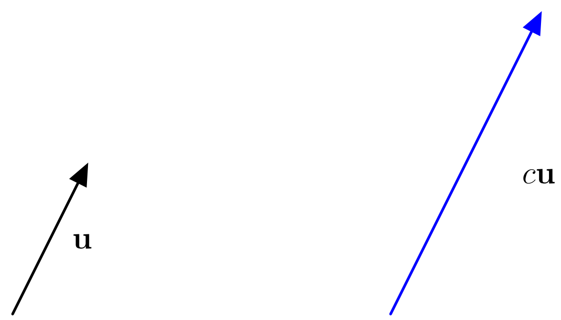
(vii) $(ab)\mathbf{u} = a(b\mathbf{u})$,

(viii) if $a\mathbf{u} = \mathbf{0}$, then $a = 0$ or $\mathbf{u} = 0$.

Geometrically, adding $\mathbf{u}$ to $\mathbf{v}$ can be visualized by joining the tail of $\mathbf{v}$ at the head of $\mathbf{u}$, and the head of $\mathbf{v}$ is the resultant,



and a scalar multiple of a vector is scaling (stretching or compressing) the vector,



3.3 Solutions to linear systems (revisit)

There are 2 ways to express subsets of $\mathbb{R}^n$, implicit and explicit form.

- Implicit form: $\left\{ \mathbf{v} \in \mathbb{R}^n \;\middle|\; \mathbf{v} \text{ fulfills some conditions.} \right\}$,
- Explicit form: $\left\{ \mathbf{u} + s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots s_k\mathbf{v}_k \;\middle|\; s_1, s_2, ..., s_k \in \mathbb{R} \right\}$.

Example. 1. Implicit form: $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; x = -y, z = 1 \right\}$ = Explicit form: $\left\{ \begin{pmatrix} s \\ -s \\ 1 \end{pmatrix} \;\middle|\; s \in \mathbb{R} \right\}$

2. Implicit form: $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; x = y = z \right\}$ = Explicit form: $\left\{ s \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; s \in \mathbb{R} \right\}$

Consider the linear system

$$\begin{array}{rcrcrcl} 3x & + & 2y & - & z & = & 1 \\ & & y & - & z & = & 0 \end{array}$$

The implicit form is

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; 3x + 2y - z = 1, y - z = 0 \right\}$$

A general solution is $x = \frac{1}{3}(1 - s)$, $y = s$, $z = s$, $s \in \mathbb{R}$. We may present this in vector form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{3}(1-s) \\ s \\ s \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} -1/3 \\ 1 \\ 1 \end{pmatrix}, s \in \mathbb{R}.$$

The set containing all possible s is then called the <u>solution set</u> of the linear system, and is denoted as

$$\left\{ \tfrac{1}{3}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} -1/3 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; s \in \mathbb{R} \right\} \subseteq \mathbb{R}^3.$$

In order words, from the linear system, we get the implicit form of the solution set, and the explicit form is derived from a general solution.

Remark. Note that even though general solutions are not unique, the solution set is, different expressions of general solutions gives the same solution set for a linear system.

3.4 Dot Product

How do we multiply vectors? Given two (column) vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we cannot multiply them since their size don't match. However, if we transpose one of the vectors, we can multiply,

1. $\mathbf{u}\mathbf{v}^T = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} = \begin{pmatrix} u_1v_1 & u_1v_2 & \cdots & u_1v_n \\ u_2v_1 & u_2v_2 & \cdots & u_2v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_nv_1 & u_nv_2 & \cdots & u_nv_n \end{pmatrix} = (u_iv_j)_n$, or

2. $\mathbf{u}^T\mathbf{v} = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n = \sum_{i=1}^n u_iv_i.$

The first multiplication is known as <u>outer product</u>, denoted as $\mathbf{u} \otimes \mathbf{v}$, and the second is known as <u>inner product</u>, or dot product, denoted as $\mathbf{u} \cdot \mathbf{v}$. In this course, we will only be discussing inner product.

Example. 1. $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 2 + 4 - 2 = 4.$

2. $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 1 + 0 - 1 = 0.$

3. $\begin{pmatrix} 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} = 2 - 6 = -4.$

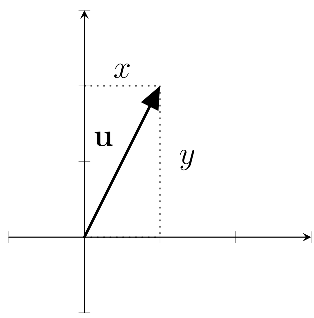
The <u>norm</u> of a vector $\mathbf{u} \in \mathbb{R}^n$ is defined to be the square root of the inner product of $\mathbf{u}$ with itself, and is denoted as $\|\mathbf{u}\|$,

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}.$$

Geometric meaning of norm. The distance between the point $\begin{pmatrix} x \\ y \end{pmatrix}$ and the origin in $\mathbb{R}^2$ is given by

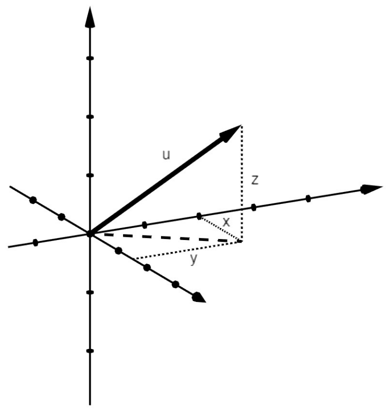
$$\text{distance} = \sqrt{x^2 + y^2} = \left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|.$$

That is, in $\mathbb{R}^2$, the norm of a vector can be interpreted as its distance from the origin.



Similarly, in $\mathbb{R}^3$, the distance of a vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ to the origin is

$$\text{distance} = \sqrt{x^2 + y^2 + z^2} = \left\| \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\|.$$



We may thus generalize and define the distance between a vector $\mathbf{v}$ and the origin in $\mathbb{R}^n$ is its norm, $\|\mathbf{v}\|$.

Observe that the distance between two vector $\mathbf{v} = (v_i)$ and $\mathbf{u} = (u_i)$ is

$$d(\mathbf{u}, \mathbf{v}) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2} = \|\mathbf{u} - \mathbf{v}\|.$$

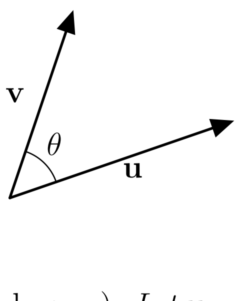
Example. 1. $\left\| \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \right\| = \sqrt{1^2 + 2^2 + (-1)^2} = \sqrt{6}.$

2. $d\left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \end{pmatrix} \right) = \left\| \begin{pmatrix} 1-0 \\ 3-5 \end{pmatrix} \right\| = \sqrt{1^2 + (-2)^2} = \sqrt{5}.$

The angle between two nonzero vectors, $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$ is the number θ with $0 \leq \theta \leq \pi$ such that

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

This is a natural definition because once again, in $\mathbb{R}^2$, this is indeed the definition of the trigonometric function cosine.



Theorem (Properties of dot product and norm)**.** *Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ be vectors and $a, b, c \in \mathbb{R}$ be scalars.*

(i) (Symmetric) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$.

(ii) (Scalar multiplication) $c\mathbf{u} \cdot \mathbf{v} = (c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v})$.

(iii) (Distribution) $\mathbf{u} \cdot (a\mathbf{v} + b\mathbf{w}) = a\mathbf{u} \cdot \mathbf{v} + b\mathbf{u} \cdot \mathbf{w}$.

(iv) (Positive definite) $\mathbf{u} \cdot \mathbf{u} \geq 0$ *with equality if and only if* $\mathbf{u} = \mathbf{0}$.

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

Proof. Let $\mathbf{u} = (u_i)$, $\mathbf{v} = (v_i)$, and $\mathbf{w} = (w_i)$.

(i) $\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^n u_i v_i = \sum_{i=1}^n v_i u_i = \mathbf{v} \cdot \mathbf{u}$. Alternatively, since $\mathbf{u}^T\mathbf{v}$ is a 1×1 matrix, it is symmetric. So $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = (\mathbf{u}^T\mathbf{v})^T = \mathbf{v}^T\mathbf{u} = \mathbf{v} \cdot \mathbf{u}$.

(ii) $c\sum_{i=1}^n u_i v_i = \sum_{i=1}^n (cu_i)v_i = \sum_{i=1}^n u_i(cv_i)$.

(iii) $\sum_{i=1}^n u_i(av_i + bw_i) = \sum_{i=1}^n (au_iv_i + bu_iw_i) = a\sum_{i=1}^n u_iv_i + b\sum_{i=1}^n u_iw_i$.

(iv) $\mathbf{u} \cdot \mathbf{u} = \sum_{i=1}^n u_i^2 \geq 0$ since $u_i \in \mathbb{R}$ are real numbers. It is clear that a sum of square of real numbers is equal to 0 if and only if all the numbers are 0.

(v) $\|c\mathbf{u}\| = \sqrt{\sum_{i=1}^n (cu_i^2)} = \sqrt{c^2}\sqrt{\sum_{i=1}^n u_i^2} = |c|\|\mathbf{u}\|$.

□

A vector $\mathbf{u}$ is a unit vector if $\|\mathbf{u}\| = 1$. We can normalize every nonzero vector by multiplying it by the reciprocal of its norm,

$$\mathbf{u} \longrightarrow \frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

Then

$$\left(\frac{\mathbf{u}}{\|\mathbf{u}\|}\right) \cdot \left(\frac{\mathbf{u}}{\|\mathbf{u}\|}\right) = \frac{\mathbf{u} \cdot \mathbf{u}}{\|\mathbf{u}\|^2} = 1.$$

Example. Let $\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$. We have computed that $\|\mathbf{u}\| = \sqrt{6}$ and so

$$\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

is a unit vector.

We say that vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are orthogonal if

$$\mathbf{u} \cdot \mathbf{v} = 0.$$

Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are orthogonal.

- Case 1: Either $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$.
- Case 2: Otherwise,

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = 0$$

  tells us that $\theta = \frac{\pi}{2}$, that is, $\mathbf{u}$ and $\mathbf{v}$ are perpendicular.

That is, $\mathbf{u}, \mathbf{v}$ are orthogonal if and only if either one of them is the zero vector or they are perpendicular to each other.

Example.
$$\begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\0\\-1\end{pmatrix}=0,\quad \begin{pmatrix}1\\0\\0\end{pmatrix}\cdot\begin{pmatrix}0\\1\\0\end{pmatrix}=0,\quad \begin{pmatrix}1\\1\\0\end{pmatrix}\cdot\begin{pmatrix}0\\0\\1\end{pmatrix}=0.$$

Exercise: Suppose $\mathbf{u}, \mathbf{v}$ are orthogonal. Show that for any $s, t \in \mathbb{R}$ scalars, $s\mathbf{u}, t\mathbf{v}$ are also orthogonal.

A set $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\} \subseteq \mathbb{R}^n$ of vectors is <u>orthogonal</u> if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for every $i \neq j$, that is, vectors in S are pairwise orthogonal. A set $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\} \subseteq \mathbb{R}^n$ of vectors is <u>orthonormal</u> if
$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases}$$
That is, S is orthogonal, and all the vectors are unit vectors.

Remark. Every orthogonal set of nonzero vectors can be normalized to an orthonormal set.

Example.

1. $S = \left\{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \right\}$ is an orthonormal set.

2. $S = \left\{ \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\-1\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix} \right\}$ is not an orthogonal set since $\begin{pmatrix}0\\1\\-1\end{pmatrix} \cdot \begin{pmatrix}-1\\0\\1\end{pmatrix} = -1 \neq 0$.

3. $S = \left\{ \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\-1\end{pmatrix}, \begin{pmatrix}1\\-2\\1\end{pmatrix} \right\}$ is an orthogonal but not orthonormal set. It can be normalized to an orthonormal set
$$\left\{ \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix}1\\-2\\1\end{pmatrix} \right\}.$$

4. $S = \left\{ \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}, \begin{pmatrix}-1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \right\}$ is an orthonormal set.

5. $S = \left\{ \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix} \right\}$ is an orthogonal set but it cannot be normalized to an orthonormal set since it contains the zero vector.

Appendix to Lecture 5

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\} \subseteq \mathbb{R}^n$ be a subset. Form a $n \times k$ matrix $\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix}$, where the columns of $\mathbf{A}$ are the vectors in S. Consider the product

$$
\begin{aligned}
\mathbf{A}^T\mathbf{A} &= \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \cdots \\ \mathbf{u}_k^T \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \cdots & \mathbf{u}_1^T\mathbf{u}_k \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \cdots & \mathbf{u}_2^T\mathbf{u}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_k^T\mathbf{u}_1 & \mathbf{u}_k^T\mathbf{u}_2 & \cdots & \mathbf{u}_k^T\mathbf{u}_k \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{u}_k \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{u}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_k \cdot \mathbf{u}_1 & \mathbf{u}_k \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_k \cdot \mathbf{u}_k \end{pmatrix}.
\end{aligned}
$$

Then S is orthogonal if and only if the product $\mathbf{A}^T\mathbf{A}$ is a diagonal matrix, and S is orthonormal if and only if $\mathbf{A}^T\mathbf{A}$ is the identity matrix of order k.