lecture 22

MA 353

Linear Algebra II

(Last time: we finished with the Cauchy – Schwarz inequality)

Theorem: If V is a Euclidean vector space, the norm $\|-\|: V \longrightarrow \mathbb{R}$ has the following properties:

(i) $\|x\| \geq 0$ for all x

(ii) $\|x\| = 0 \iff x = 0$.

(iii) $\|cx\| = |c| \cdot \|x\|$ for all $x \in V$, $c \in \mathbb{R}$

(iv) $\|x+y\| \leq \|x\| + \|y\|$ for all $x, y \in V$

Property (iv) of the norm is called the triangle inequality.

Proof:

(i) — (iii) are clear from the definition.

For the triangle inequality we have

$$(\|x\| + \|y\|)^2 = \|x\|^2 + 2\,\|x\|\cdot\|y\| + \|y\|^2$$

$$\geq \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$$

$$= \|x+y\|^2$$

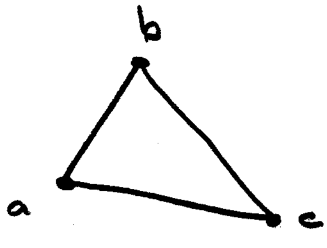
(Here : we used the Cauchy - Schwarz inequality)

$\therefore\ (\|x\| + \|y\|)^2 \geq \|x+y\|^2$

Applying $(-)^{\frac{1}{2}}$ (a strictly increasing function)

$\therefore\ \|x\| + \|y\| \geq \|x+y\|$ ∎

(Picture : (Triangle inequality)



$\|a-c\| \leq \|a-b\| + \|b-c\|$

$x := a-b$

$y := b-c$)

<u>Definition</u> : Two elements x, y of a Euclidean vector space are said to be <u>orthogonal</u> (or perpendicular) to each other (written $x \perp y$) if

$$\langle x, y \rangle = 0 .$$

The definition of $\alpha(x,y)$ for nonzero vectors implies that, if $\langle x,y \rangle = 0$, then the angle between x and y is 90°.

<u>Definition</u> : If M is a subset of the Euclidean vector space V, then

$$M^{\perp} := \{x \in V \mid x \perp y \text{ for all } y \in M\}$$

is called the <u>orthogonal complement</u> of M

Instead of " $x \perp y$ for all $y \in M$ " one can use the shorter notation

$$x \perp M .$$

<u>Proposition</u> : $M^{\perp}$ is a subspace of V.

<u>Proof</u> : $M^{\perp} \neq \emptyset$, since $0 \perp M$,

and if $x, y \in M^\perp$ and $c \in \mathbb{R}$, then $x+y \in M^\perp$ and $cx \in M^\perp$ follows immediately from the linearity of

$$\langle -, u \rangle : V \longrightarrow \mathbb{R} .$$

<u>Definition</u> : An r-tuple $(v_1, \dots, v_r)$ of vectors in a Euclidean vector space is said to be <u>orthonormal</u> (or an <u>orthonormal system</u>) if

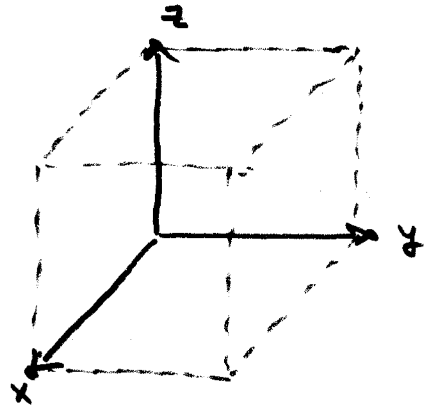
$$(*) \begin{cases} \|v_i\| = 1, & i = 1, \dots, r, \quad \text{and} \\ v_i \perp v_j & \text{for } i \neq j \end{cases}$$

<u>Note</u>: This is sometimes expressed differently:

$$(*)' \quad \langle v_i, v_j \rangle = \delta_{ij}$$

Here, $\delta_{ij} := \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$

Picture :



(An orthonormal system (x, y, z))

Example : If x, y, z are the three "edge vectors" of a cube sitting at the origin of $\mathbb{R}^3$ with edge length equal to 1, then (x, y, z) is an orthonormal system

(i.e., pairwise mutually perpendicular vectors of length 1)

(Note: The pair (x, y) and even (z) alone are also orthonormal systems)

Proposition : An orthonormal system is always linearly independent.

Proof : Let $(v_1, \dots, v_r)$ be orthonormal and suppose

$$c_1 v_1 + \cdots + c_r v_r = 0 . \qquad (*)$$

(**We want** : to verify that $c_1 = 0$, $\vdots$, $c_r = 0$)

Applying $\langle -, v_i \rangle$ to $(*)$ gives

$$\begin{aligned} \langle c_1 v_1 + \cdots + c_r v_r, v_i \rangle &= \langle 0, v_i \rangle = 0 \\ \| \\ c_1 \langle v_1, v_i \rangle + \cdots + c_r \langle v_r, v_i \rangle \\ \| \\ c_i \langle v_i, v_i \rangle &= c_i \qquad \text{for } i = 1, \dots, r. \end{aligned}$$

$\therefore \; c_1 = \cdots = c_r = 0 .$ ∎

Proposition:

If $(v_1, \ldots, v_n)$ is an orthonormal basis of V, then for each $x \in V$ we have the "expansion formula"

$$x = \sum_{i=1}^{n} \langle v, v_i \rangle v_i$$

Proof: We know that x is expressible as

$$x = c_1 v_1 + \cdots + c_n v_n \qquad (**)$$

since $(v_1, \ldots, v_n)$ is a basis.

Applying $\langle -, v_i \rangle$ to $(**)$ gives

$$\langle x, v_i \rangle = c_i \langle v_i, v_i \rangle$$

$$= c_i \qquad \text{for } i = 1, \ldots, r. \qquad \blacksquare$$

We have always assumed that the orthonormal system is given.

(**Q**. How does one obtain orthonormal systems ?)

Theorem : Let V be a finite-dimensional Euclidean vector space. Then there exists an orthonormal basis for V.

Proof : Start with an arbitrary basis

$$\underline{\underline{B}} := (v_1, \dots, v_n)$$

of V. We will describe a method called the <u>Gram – Schmidt procedure</u> which successively changes an arbitrary linearly independent r-tuple of vectors

$$(v_1, \dots, v_r)$$

into an orthonormal system

$$(w_1, \ldots, w_r)$$

and in such a way that

$$\left(\begin{array}{c} \mathrm{Span}(v_1, \ldots, v_k) = \mathrm{Span}(w_1, \ldots, w_k) \\ \text{for } k = 1, \ldots, r. \end{array} \right)$$

<u>Step</u>. Normalize v_1, so that $\langle v_1, v_1 \rangle = 1$.
To do this, we note that

$$\langle cv, cv \rangle = c^2 \langle v, v \rangle$$

Since $\langle v_1, v_1 \rangle > 0$ (by positive def.)
we set $c := \langle v_1, v_1 \rangle^{-\frac{1}{2}} = \frac{1}{\|v_1\|}$

and replace v_1 by

$$\boxed{w_1 := c v_1 = \frac{v_1}{\|v_1\|}}$$

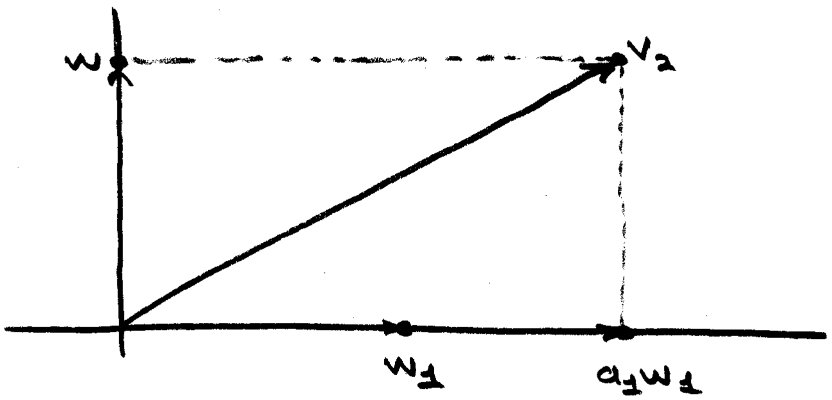
(<u>Note:</u> $\mathrm{span}(v_1) = \mathrm{Span}(w_1)$)

Step : Replace v_2 by its component

$$w \in \mathrm{Span}(w_1)^{\perp}$$

perpendicular to $\mathrm{Span}(w_1)$.

Picture :



(The vector $a_1 w_1$ is the orthogonal projection of v_2 onto the subspace (the line) spanned by w_1)

To do this, we look for a linear combination of w_1 and v_2 which is orthogonal to w_1.

The required linear combination is (see picture)

$$\begin{cases} w := v_2 - a_1 w_1 \quad , \text{ where} \\ a_1 := \langle v_2, w_1 \rangle \end{cases}$$

Note: $\langle w, w_1 \rangle = \langle v_2, w_1 \rangle - a_1 \langle w_1, w_1 \rangle$

$= \langle v_2, w_1 \rangle - a_1$

$= a_1 - a_1$

$= 0$

We normalize this vector w to length 1, obtaining a vector w_2 which we substitute for v_2 :

$$\boxed{w_2 := \frac{w}{\|w\|}}$$

(Note: $\mathrm{Span}(v_1, v_2) = \mathrm{Span}(w_1, w_2)$ and (w_1, w_2) is orthonormal system.)

Step : Replace v_3 by its component

$$w \in \mathrm{Span}(w_1, w_2)^{\perp}$$

perpendicular to $\mathrm{Span}(w_1, w_2)$.

We adjust v_3 as follows: we let

$$\begin{cases} w := v_3 - a_1 w_1 - a_2 w_2 \\ a_1 := \langle v_3, w_1 \rangle \\ a_2 := \langle v_3, w_2 \rangle \end{cases}$$

(Then it follows that

$$\langle w, w_1 \rangle = 0$$
$$\langle w, w_2 \rangle = 0 \quad)$$

We normalize this vector w to length 1, obtaining a vector w_3 which we substitute for v_3:

$$\boxed{w_3 := \frac{w}{\|w\|}}$$

(<u>Note</u>: $\text{Span}(v_1, v_2, v_3) = \text{Span}(w_1, w_2, w_3)$ and (w_1, w_2, w_3) is orthonormal system.)

Step : Suppose that the $k-1$ vectors

$$w_1, \dots, w_{k-1}$$

are orthonormal and that

$$(w_1, \dots, w_{k-1}, v_k, \dots, v_n)$$

is a basis. We replace v_k by its component

$$w \in \mathrm{Span}(w_1, \dots, w_{k-1})^{\perp}$$

and normalize. We adjust v_k as follows: We let

$$\begin{cases} w := v_k - a_1 w_1 - \cdots - a_{k-1} w_{k-1} \\ a_1 := \langle v_k, w_1 \rangle \\ \vdots \\ a_{k-1} := \langle v_k, w_{k-1} \rangle \end{cases}$$

Then it follows that $w \perp w_i \quad (i = 1, \dots, k-1)$

We normalize this vector w to length 1

$$\boxed{w_k := \frac{w}{\|w\|}}$$

obtaining a vector w_k which we substitute for v_k.

(**Note:** $\text{span}(v_1, \dots, v_k) = \text{Span}(w_1, \dots, w_k)$ and $(w_1, \dots, w_k)$ is orthonormal system.)

Since v_k is in the span of

$$(w_1, \dots, w_k, v_{k+1}, \dots, v_n)$$

this is a basis of V. Carrying on in this way finishes the proof (i.e., induction on k finishes the proof). $\blacksquare$