MTH 100 — Linear Algebra

Notes for Inner Product

In $\mathbb{R}^2$ and $\mathbb{R}^3$, we can adopt another definition for the inner product:-

Given $\bar{u} = u_1\bar{e}_1 + u_2\bar{e}_2 + u_3\bar{e}_3$

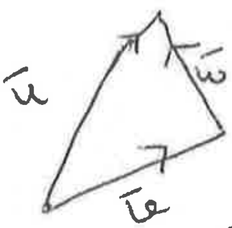
and $\bar{v} = v_1\bar{e}_1 + v_2\bar{e}_2 + v_3\bar{e}_3$,

$$\bar{u}\circ\bar{v} = \bar{u}^T\bar{v} = u_1v_1 + u_2v_2 + u_3v_3 \qquad (1)$$

and $\bar{u}\circ\bar{v} = |\bar{u}||\bar{v}|\cos\theta$, where θ is the angle between the vectors. (2)

Definition 2 is a geometrical definition, based on the notion of vectors as directed line segments. Note that two non-parallel vectors determine a plane, hence the angle θ is ~~well~~ defined; since $\cos(-\theta) = \cos\theta$, the direction in which the angle is measured is immaterial.

To show equivalence of the two definitions:-



Applying the Law of Cosines from Trigonometry:-

$$|\bar{w}|^2 = |\bar{u}|^2 + |\bar{v}|^2 - 2|\bar{u}||\bar{v}|\cos\theta \qquad (1)$$

or $2|\bar{u}||\bar{v}|\cos\theta = |\bar{u}|^2 + |\bar{v}|^2 - |w|^2$

Now, $\bar{w} = \bar{u} - \bar{v} = (u_1 - v_1)\bar{e}_1 + (u_2 - v_2)\bar{e}_2 + (u_3 - u_3)\bar{e}_3$

$$\therefore |\bar{w}|^2 = \left(\sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}\right)^2$$

$$= \left[(u_1^2 + u_2^2 + u_3)^2 + (v_1^2 + v_2^2 + v_3)^2 - 2(u_1v_1 + u_2v_2 + u_3v_3)\right] \qquad (2)$$

Continued from previous page (2)

Since $|\bar{u}|^2 = u_1^2 + u_2^2 + u_3^2$

and $|\bar{v}|^2 = v_1^2 + v_2^2 + v_3^2$,

using (2), we get:

$$2|\bar{u}||\bar{v}|\cos\theta = 2(u_1v_1 + u_2v_2 + u_3v_3)$$

or $\bar{u} \circ \bar{v} = |\bar{u}||\bar{v}|\cos\theta = u_1v_1 + u_2v_2 + u_3v_3$ as desired.

NB 1: ~~the geometrical~~ The equivalence for $\mathbb{R}^2$ is a simplified version of the above derivation for $\mathbb{R}^3$.

NB 2:- The geometrical derivation is meaningful only for $\mathbb{R}^2$ and $\mathbb{R}^3$ - hence, for $\mathbb{R}^n$ in general we adopt the more abstract definition $\bar{u} \circ \bar{v} = \bar{u}^T \bar{v}$

MTH 100 – Lecture Notes for ~~Monday~~

~~Thursday 3-11-2016~~ MONDAY 07-11-2016 etc.

Prop. ~~3~~4.7 (a):- $\bar{v} \in W^{\perp}$ iff $\bar{v}$ is orthogonal to every vector in a spanning set for W.

Proof: [$\Longrightarrow$] is obvious since if $\bar{v} \in W^{\perp}$, then $\bar{v} \perp \bar{w}$ forall $\bar{w} \in W$.

[$\Longleftarrow$] Suppose $\bar{v}$ is orthogonal to every vector in a spanning set K for W.

Suppose $\bar{w} \in W$.

Then, $\bar{w} = c_1\bar{u}_1 + \cdots + c_p\bar{u}_p$ where the $\bar{u}_i \in K$.

$$\therefore \langle \bar{v}, \bar{w} \rangle = \langle \bar{v}, c_1\bar{u}_1 + \cdots + c_p\bar{u}_p \rangle$$
$$= c_1\langle \bar{v}, \bar{u}_1 \rangle + \cdots + c_p\langle \bar{v}, \bar{u}_p \rangle$$
$$= 0, \text{ by hypothesis.}$$

$\therefore$ $\bar{v} \in W^{\perp}$, as reqd.

(b) $W^{\perp}$ is a subspace of V and $W \cap W^{\perp} = \{\bar{0}\}$.

Proof: Suppose $\bar{v}_1, \bar{v}_2 \in W^{\perp}$ and $\bar{w} \in W$.

(i) Then: $\langle \bar{v}_1 + \bar{v}_2, \bar{w} \rangle = \langle \bar{v}_1, \bar{w} \rangle + \langle \bar{v}_2, \bar{w} \rangle$

$= \bar{0} + \bar{0} \quad = \bar{0}$

$\Longrightarrow \bar{v}_1 + \bar{v}_2 \in W^{\perp}$

(ii) Again, if $c \in F$, then $\langle c\bar{v}_1, \bar{w} \rangle = c\langle \bar{v}_1, w \rangle$

$= c\,\bar{0} = \bar{0} \Longrightarrow c\bar{v}_1 \in W^{\perp}$.

By (i) and (ii), $W^{\perp}$ is a subspace of V.

Finally, suppose $\bar{w} \in W \cap W^{\perp}$

Then $\langle \bar{w}, \bar{w} \rangle = 0$ and hence by axiom 4 for inner products, $\bar{w} = \bar{0}$.

Proposition ~~8~~ 4 :- If $\{\bar{u}_1, \ldots, \bar{u}_p\}$ is an orthogonal basis for ~~for~~ a subspace W, and if $\bar{y} \in W$, then

$$\bar{y} = \sum_{j=1}^{p} \frac{\langle \bar{y}, \bar{u}_j \rangle}{\langle \bar{u}_j, u_j \rangle} \bar{u}_j \qquad (1)$$

Proof: Suppose $\bar{y} = c_1\bar{u}_1 + \cdots + c_p\bar{u}_p$. (2)

We need to determine the coefficients $c_1, c_2, \ldots, c_p$.

Let us take the inner product with $\bar{u}_j$, where j lies between 1 and p (inclusive).

Then, $\langle \bar{y}, \bar{u}_j \rangle = \langle c_1\bar{u}_1 + \cdots + c_p\bar{u}_p, \bar{u}_j \rangle$ from (2)

$$= c_1\langle \bar{u}_1, \bar{u}_j \rangle + c_2\langle \bar{u}_2, \bar{u}_j \rangle + \cdots + c_p\langle \bar{u}_p, \bar{u}_j \rangle$$

$$= c_j \langle \bar{u}_j, \bar{u}_j \rangle$$ since other terms are 0,

$$\therefore\ c_j = \frac{\langle \bar{y}, \bar{u}_j \rangle}{\langle \bar{u}_j, \bar{u}_j \rangle}$$, which

gives us (1).

An example for Prop. ~~8~~ 4 8:-

→ Given orthogonal basis $\mathcal{B}$:-

$$\bar{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad \bar{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \bar{v}_3 = \begin{bmatrix} -1 \\ 4 \\ -1 \end{bmatrix},$$

find the coordinates of

$\bar{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ w r t this basis

— First check : $\bar{v}_1 \circ \bar{v}_2 = 2 - 2 = 0$

$\bar{v}_1 \circ \bar{v}_3 = -2 + 4 - 2 = 0$

$\bar{v}_2 \circ \bar{v}_3 = -1 + 1 = 0$

So now :

$$c_1 = \frac{\bar{v} \circ \bar{v}_1}{\bar{v}_1 \circ \bar{v}_1} = \frac{2 + 2 + 6}{9} = \frac{10}{9}$$

$$c_2 = \frac{\bar{v} \circ \bar{v}_2}{\bar{v}_2 \circ \bar{v}_2} = \frac{1 - 3}{2} = -\frac{2}{2} = -1$$

$$c_3 = \frac{\bar{v} \circ \bar{v}_3}{\bar{v}_3 \circ \bar{v}_3} = \frac{-1 + 8 - 3}{18} = \frac{2}{9}$$

Check :-

$$\frac{10}{9}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \frac{2}{9}\begin{bmatrix} -1 \\ 4 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{20}{9} - 1 - \frac{2}{9} \\ \frac{10}{9} + \frac{8}{9} \\ \frac{20}{9} + 1 - \frac{2}{9} \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \text{ as reqd.}$$

Proof of Theorem 6 (ODT):-

We will prove the original version, from which the alternative version can be quickly derived.

We assume that any finite-dimensional subspace W of an inner product space has an orthogonal basis.

This assumption is Theorem 7 (Gram-Schmidt Process) which has been ~~cov~~ covered later – however, the proof of Theorem 7 does not require Theorem 6, so the assumption is logically valid.

We first prove uniqueness :- i.e. any vector $\bar{y} \in V$ cannot be expressed in more than one way as a sum of a vector in W and ~~to~~ a vector in $W^\perp$.

Suppose BWOC that

$$\left.\begin{aligned} \bar{y} &= \hat{y} + \bar{z} \\ \text{and} \quad \bar{y} &= \hat{y}_1 + \bar{z}_1 \end{aligned}\right\} \begin{aligned} &\text{where } \hat{y}, \hat{y}_1 \in W \\ &\text{and } \bar{z}, \bar{z}_1 \in W^\perp \end{aligned}$$

Subtracting $\bar{0} = (\hat{y} - \hat{y}_1) + (\bar{z} - \bar{z}_1)$

or $\hat{y} - \hat{y}_1 = -(\bar{z} - \bar{z}_1)$

LHS $\in W$, RHS $\in W^\perp$, i.e. they both are in $W \cap W^\perp = \{\bar{0}\}$, by Prop. 4.7.

$$\left.\begin{aligned} \therefore \quad \hat{y} - \hat{y}_1 = \bar{0} &\implies \hat{y} = \hat{y}_1 \\ \text{and} \quad \bar{z} - \bar{z}_1 = \bar{0} &\implies \bar{z} = \bar{z}_1 \end{aligned}\right| \quad \text{Result follows.}$$

~~5~~ 4

We next prove that a decomposition exists, i.e. that $\bar{y} = \hat{y} + \bar{z}$, where $\hat{y} \in W$ and $\bar{z} \in W^{\perp}$.

In fact, put $\hat{y} = c_1\bar{u}_1 + \cdots + c_p\bar{u}_p$

where $c_j = \dfrac{\langle \bar{y}, \bar{u}_j \rangle}{\langle \bar{u}_j, \bar{u}_j \rangle}$ for $j = 1, 2, \cdots, p$ (1)

So clearly $\hat{y} \in W$.

Now, put $\bar{z} = \bar{y} - \hat{y}$ so obviously

$$\bar{y} = \hat{y} + \bar{z}.$$

It ~~only remains~~ to show that $\bar{z} \in W^{\perp}$.

We use Prop. ~~3~~4.7 (a) for this, so it suffices to show that $\langle \bar{z}, \bar{u}_j \rangle = 0$ for $j = 1, 2, \cdots, p$.

But $\langle \bar{z}, \bar{u}_j \rangle = \langle \bar{y} - \hat{y}, \bar{u}_j \rangle$

$$= \langle \bar{y}, \bar{u}_j \rangle - \langle \hat{y}, u_j \rangle$$

$$= \langle \bar{y}, \bar{u}_j \rangle - [c_1 \langle \bar{u}_1, \bar{u}_j \rangle + \cdots + c_p \langle \bar{u}_1, u_p \rangle]$$

$$= \langle \bar{y}, \bar{u}_j \rangle - c_j \langle \bar{u}_j, \bar{u}_j \rangle$$

$$= \langle \bar{y}, \bar{u}_j \rangle - \frac{\langle \bar{y}, \bar{u}_j \rangle}{\langle \bar{u}_j, \bar{u}_j \rangle} \langle \bar{u}_j, u_j \rangle \quad \text{from (1)}$$

$= 0$, as desired.

5

Example for Gram-Schmidt Process:-

Construct an orthonormal basis for $\mathbb{R}^3$ starting with the basis:

$$\bar{x}_1 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \quad \bar{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \bar{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Put $\bar{v}_1 = \bar{x}_1 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ (1)

Then $\bar{v}_2 = \bar{x}_2 - \dfrac{\bar{v}_1 \circ \bar{x}_2}{\bar{v}_1 \circ \bar{v}_1}\,\bar{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \dfrac{13}{14}\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$

$$= \begin{bmatrix} -\frac{6}{7} \\ \frac{15}{14} \\ \frac{3}{14} \end{bmatrix} \qquad (2)$$

$$\bar{v}_3 = \bar{x}_3 - \frac{\bar{v}_1 \circ \bar{x}_3}{\bar{v}_1 \circ \bar{v}_1}\,\bar{v}_1 - \frac{\bar{v}_2 \circ \bar{x}_3}{\bar{v}_2 \circ \bar{v}_2}\,\bar{v}_2$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{6}{14}\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \frac{-\frac{6}{7} + \frac{15}{14} + \frac{3}{14}}{\frac{12^2 + 15^2 + 3^2}{14^2}} \begin{bmatrix} -\frac{6}{7} \\ \frac{15}{14} \\ \frac{3}{14} \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{7}\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \frac{2}{9}\begin{bmatrix} -\frac{6}{7} \\ \frac{15}{14} \\ \frac{3}{14} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{21-18+4}{21} \\ \frac{21-9-5}{21} \\ \frac{21-27-1}{21} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix} \qquad (3)$$

NB: We can check that $\bar{v}_1, \bar{v}_2, \bar{v}_3$ are orthogonal (whereas original basis vectors were not.)

$$\bar{v}_1 \circ \bar{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \circ \begin{bmatrix} -\frac{6}{7} \\ \frac{15}{14} \\ \frac{3}{14} \end{bmatrix} = -\frac{12}{7} + \frac{15}{14} + \frac{9}{14}$$

$$= 0$$

$$\bar{v}_1 \circ \bar{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \circ \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix} = \frac{2}{3} + \frac{1}{3} - \frac{3}{3}$$

$$= 0$$

$$\bar{v}_2 \circ \bar{v}_3 = \begin{bmatrix} -\frac{6}{7} \\ \frac{15}{14} \\ \frac{3}{14} \end{bmatrix} \circ \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix} = -\frac{6}{21} + \frac{15}{42} - \frac{3}{42}$$

$$= 0$$

NB2: If we desire an orthonormal basis, we each vector $\bar{v}_i$ by its length, to get:

$$\bar{v}_1' = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

$$\bar{v}_2' = \frac{1}{\sqrt{42}} \begin{bmatrix} -4 \\ 5 \\ 1 \end{bmatrix}$$

$$\bar{v}_3' = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Proof of Other Results

PROP. 49 (Pythagorean Theorem) $\bar{u}$ and $\bar{v}$ are orthogonal to each other iff $\|\bar{u}+\bar{v}\|^2 = \|\bar{u}\|^2 + \|\bar{v}\|^2$.

Proof: $\|\bar{u}+\bar{v}\|^2 = \langle \bar{u}+\bar{v}, \bar{u}+\bar{v} \rangle = \underset{\|\bar{u}\|^2}{\langle \bar{u}, \bar{u} \rangle} + \underset{\|\bar{v}\|^2}{\langle \bar{v}, \bar{v} \rangle} + 2\langle \bar{u}, \bar{v} \rangle$ 1

$\therefore \|\bar{u}+\bar{v}\|^2 = \|\bar{u}\|^2 + \|\bar{v}\|^2$

$\iff 2\langle \bar{u}, \bar{v} \rangle = 0 \iff \bar{u} \perp \bar{v}$

PROP 50 (Best Approx. Theorem).

Proof: Let $\bar{v} \in W$.

Then $\|\bar{y}-\bar{v}\|^2 = \langle \bar{y}-\bar{v}, \bar{y}-\bar{v} \rangle$

$= \langle (\bar{y}-\hat{y}) + (\hat{y}-\bar{v}), (\bar{y}-\hat{y}) + (\hat{y}-\bar{v}) \rangle$

$= \langle \bar{y}-\hat{y}, \bar{y}-\hat{y} \rangle + \langle \hat{y}-\bar{v}, \hat{y}-\bar{v} \rangle + 2\langle \bar{y}-\hat{y}, \hat{y}-\bar{v} \rangle$ 1

Now, $\bar{y}-\hat{y} \in W^{\perp}$ whereas $\hat{y}-\bar{v} \in W$

Hence, the 3rd term on RHS of 1 is 0.

$\therefore \|\bar{y}-\bar{v}\|^2 = \|\bar{y}-\hat{y}\|^2 + \|\hat{y}-\bar{v}\|^2$ 2

Now, if $\bar{y} = \bar{v}$, then $\bar{y} \in W \Rightarrow \bar{y} = \hat{y} = \bar{v}$, which is not allowed.

Hence, $\|\bar{y}-\bar{v}\|^2 > 0$ whence

$\|\bar{y}-\bar{v}\| > \|\bar{y}-\hat{y}\|$ from 2

Proof of Corollary 50.1 : We have that
$\bar{v} = \text{Proj}_W \bar{v} + \bar{z}$, where $\bar{z} \in W^\perp$

$\therefore$ Applying Pythagorean Theorem,

$$\|\bar{v}\|^2 = \|\text{Proj}_W \bar{v} + \bar{z}\|^2$$

$$= \|\text{Proj}_W \bar{v}\|^2 + \|\bar{z}\|^2.$$

Since $\|\bar{z}\|^2 \geq 0$, the result follows.

Proof of Prop. 51:- ~~Proof of Theorem~~ (Cauchy-Schwarz Inequality) :-

$$|\langle \bar{u}, \bar{v}\rangle| \leq \|\bar{u}\| \, \|\bar{v}\|.$$

Proof: Clearly result holds if either $\bar{u}$ or $\bar{v} = 0$.
So assume both $\bar{u}$ and $\bar{v}$ non-zero, and
apply Cor 50.1 above, taking $W = \text{span}\{\bar{v}\}$.

$$\therefore \|\text{Proj}_W \bar{u}\| \leq \|\bar{u}\| \quad (1)$$

In (1), LHS $= \|\text{Proj}_W \bar{u}\| = \left\| \frac{\langle \bar{u}, \bar{v}\rangle}{\langle \bar{v}, \bar{v}\rangle} \bar{v} \right\|$

$$= \frac{|\langle \bar{u}, \bar{v}\rangle| \, \|\bar{v}\|}{\|\bar{v}\|^2} \quad (2)$$

From (1) and (2), $|\langle \bar{u}, \bar{v}\rangle| \leq \|\bar{u}\| \|\bar{v}\|$, as reqd.

~~Pro~~ Prop. 52 (Triangle Inequality): $\|\bar{u} + \bar{v}\| \leq \|\bar{u}\| + \|\bar{v}\|$

Proof: We have $\|\bar{u} + \bar{v}\|^2 = \langle \bar{u} + \bar{v}, \bar{u} + \bar{v}\rangle$

$$= \|\bar{u}\|^2 + \|\bar{v}\|^2 + 2\langle \bar{u}, \bar{v}\rangle$$

$$\leq \|\bar{u}\|^2 + \|\bar{v}\|^2 + 2|\langle \bar{u}, \bar{v}\rangle|$$

$$\leq \|\bar{u}\|^2 + \|\bar{v}\|^2 + 2\|\bar{u}\| \|\bar{v}\|, \text{ using C-S inequality}$$

$$= (\|\bar{u}\| + \|\bar{v}\|)^2.$$

Result follows.

Inner Products

- **Definition**: An **inner product** on a (real) vector space V is a function, that to each pair of vectors u and v in V associates a scalar (real number) ⟨u,v⟩ and satisfies the following axioms for all vectors u, v, w in V and all scalars c:

1. ⟨u,v⟩ = ⟨v,u⟩
2. ⟨u + v, w⟩ = ⟨u,w⟩ + ⟨v,w⟩
3. ⟨cu,v⟩ = c ⟨u,v⟩
4. ⟨u,u⟩ ≥ 0 and ⟨u,u⟩ = 0 if and only if **u = 0**

A vector space with an inner product is called an **inner product space.**

Note: The above definition holds for real inner products. For complex inner products, the first axiom above becomes:
⟨u,v⟩ = $\overline{\langle v,u\rangle}$ (in other words, the complex conjugate)

Standard Example of an Inner Product

- **Definition:** If we regard vectors **u**, **v** in R^n as $n\times1$ matrices, then the transpose $\mathbf{u}^T$ is a $1\times n$ matrix. Thus the matrix product $\mathbf{u}^T\mathbf{v}$ is a 1×1 matrix which we regard as a real number. This real number is called the **inner product** or **dot product**, written **u·v.**
- If $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$,

then $\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = u_1v_1 + u_2v_2 + \ldots + u_nv_n$

Another Example of an Inner Product Space

- The space $R_n[t]$ of all polynomials of degree less than or equal to n can be made into an inner product space in the following way. Let $t_0, t_1, t_2, \ldots, t_n$ be **distinct** real numbers (nb: there are n + 1 numbers). For any two polynomials p and q in $R_n[t]$, we define:
$\langle p,q\rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \ldots\ldots + p(t_n)q(t_n)$
- **Remark:** It can be verified that the four axioms for an inner product hold with the above definition (*exercise !*).
- The above inner product for polynomials is used when the values at specific points are important (interpolation problems).

Yet Another Example of an Inner Product Space

- The space C[a,b] of all continuous functions on the closed interval [a,b] can be made into an inner product space with the following definition:

$$\langle f,g\rangle = \int_a^b f(t)g(t)dt$$

- **Remark:** Again, it can be verified that the four axioms for an inner product hold with the above definition (*exercise !*).
- The above inner product plays a very important role in the study of continuous functions and their applications in signals and systems.