QM-03 Lecture Notes *

Vector Spaces With Inner Product

Inner product and orthogonality

A. K. Kapoor
kapoor.proofs@gmail.com
akkhcu@gmail.com

January 31, 2016

Contents

* qm-lnu-03001.tex; 16.x   Proofs   Jan 29, 2016

§1 Inner Product in Vector Spaces

From now onwards all the vector spaces we deal are complex vector spaces of finite dimension unless mentioned otherwise.

Definition 1 Norm *of a vector f in a vector space V is a real number $\|f\|$ satisfying the following properties.*

$*1)$ $\|f\| \geq 0$, and, $\|f\| = 0$ if and only if $f = 0$.

$*2)$ $\|\alpha f\| = |\alpha| \|f\|$

$*3)$ $\|f + g\| \leq \|f\| + \|g\|$ (Triangle Inequality)

Q:*Is norm a linear functional !? WHY ?*

Definition 2 *An* **inner product** *(or scalar product), denoted by (f, g), in a complex vector space $\mathcal{V}$ is a complex valued function of the ordered pair of vectors $f, g \in \mathcal{V}$ such that*

$\star 1)$ $(f, f) \geq 0, \text{and } (f, f) = 0 \quad \text{iff } f = 0$

$\star 2)$ $(f, g) = (g, f)^*$

$\star 3)$ $(f, \alpha_1 g_1 + \alpha_2 g_2) = \alpha_1 (f, g_2) + \alpha_2 (f, g_2)$

$\star 4)$ $(\alpha_1 f_1 + \alpha_2 f_2, g) = \alpha_1^* (f_1, g) + \alpha_2^* (f, g)$

We shall not discuss real vector spaces with inner product.

Examples Of Properties Of Inner Product

The property $\star 4)$ can be proved from $\star 2)$ and $\star 3)$. Thus we have

$$\begin{aligned} (\alpha_1 f_1 + \alpha_2 f_2, g) &= [(g, \alpha_1 f_1) + (g, \alpha_2 f_2)]^* && (1) \\ &= [\alpha_1 (g, f_1) + \alpha_2 (g, f_2)]^* && (2) \\ &= \alpha_1^* (g, f_1) + \alpha_2^* (g, f_2) && (3) \end{aligned}$$

Using the property $\star 2)$ once again we get the desired result:

$$(\alpha_1 f_1 + \alpha_2 f_2, g) = \alpha_1 (f_1, g) + \alpha_2 (f_:2, g)$$

We shall now prove two important identities.

Parallelogram Identity

$$\|f + g\|^2 + \|f - g\|^2 = 2(\|f\|^2 + \|g\|^2) \tag{4}$$

Polarization Identity

$$4(f,g) = \|f+g\|^2 - \|f-g\|^2 + i\|f-ig\|^2 - i\|f+ig\|^2 \tag{5}$$

PROOF :

$$\begin{aligned}
\|f+g\|^2 &= (f+g, f+g) = (f,f) + (f,g) + (g,f) + (g,g) && (6)\\
\|f-g\|^2 &= (f-g, f-g) = (f,f) - (f,g) - (g,f) + (g,g) && (7)\\
\|f-ig\|^2 &= (f-ig, f-ig) = (f,f) - i(f,g) + i(g,f) + (g,g) && (8)\\
\|f+ig\|^2 &= (f+ig, f+ig) = (f,f) + i(f,g) - i(g,f) + (g,g) && (9)
\end{aligned}$$

Adding Eq.(6) and Eq.(7) gives the parallelogram identity. In a similar fashion taking Eq.(6) - Eq.(7) + $i \otimes$ Eq.(8) - $i\otimes$ Eq.(9) gives the polarization identity.

Relating Norm and Inner Product In a vector space with an inner product if we define

$$\|f\| = \sqrt{(f,f)},$$

then $\|f\|$ has all the properties of the norm. The two properties (1) and (2) of the norm are automatically satisfied. The third property, viz., the triangle inequality will be proved below after the proof of Cauchy Schwarz inequality.

Conversely, if a norm is defined in a complex vector space we ask: "can we introduce a norm such that the relation is maintained?" The answer is YES if and only if the norm satisfies the parallelogram identity. The right hand side of the polarization identity can then be taken as the definition of inner product. The result will satisfy all the axioms for the inner product.

§2 Orthogonality and Gram Schmidt Procedure

Definition 3 *We say that two vectors f and g are* **orthogonal** *if* $(f, g) = 0$

LEMMA : If $g \neq 0$ then the vector

$$x = f - \frac{(g, f)}{(g, g)} g$$

is orthogonal to g.

Proof :

Consider

$$(g, x) = (g, f - \frac{(g, f)}{(g, g)} g) = (g, f) - \frac{(g, f)}{(g, g)} (g, g) \tag{10}$$

$$= (g, f) - (g, f) = 0 \tag{11}$$

Therefore, g is orthogonal to $x = f - \dfrac{(g, f)}{(g, g)} g$.

Definition 4 *Two vectors f and g are* **orthogonal** *if* $(f, g) = 0$.

Definition 5 *A set of vectors $\mathcal{X}$ is an* **orthogonal set** *if $\forall$ pair $x, y \in \mathcal{X}$, we have $(x, y) = 0$.*

Definition 6 *A set of vectors $\mathcal{X}$ is called* **orthonormal set** *if*

(a) *for every pair $x, y \in \mathcal{X}$ we have $(x, y) = 0$ and*

(b) *for every $x \in \mathcal{X}$ we have $\|x\| = 1$.*

Definition 7 *A set $\{x_1, x_2, ..., x_r\}$ is an* **orthonormal set** *iff $(x_i, x_j) = \delta_{ij}$.*

Definition 8 *An orthonormal set is called a* **complete orthonormal set** *if it is not contained in any larger orthonormal set.*

Theorem 1 *An orthogonal set $\mathcal{X} = \{x_1, x_2, ..x_r\}$ of non-zero vectors is linearly independent.*

Proof : Consider

$$\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_r x_r = 0 \tag{12}$$

Taking scalar product with x_1 gives zero for all terms except the first one. Thus

$$\alpha_1 (x_1, x_1) = 0 \Rightarrow \alpha_1 = 0 \tag{13}$$

$$(\because x_1 \neq 0 \Rightarrow (x_1, x_1) \neq 0). \tag{14}$$

Remark : Earlier we have seen that the vector $h = f - \lambda g$ is orthogonal to the vector g if λ is taken to be $(g, f)/(g, g)$. The following theorem generalizes this result to orthogonal sets.

Theorem 2 *If $\mathcal{U} = u_1, u_2, ..., u_n$ is any finite orthogonal set containing nonzero vectors of an inner product space and if $\lambda_k = (u_k, x)/(u_k, u_k)$, then the vector h defined by*

$$h = f - \lambda_1 u_1 - \lambda_2 u_2 - ... - \lambda_k u_k$$

is orthogonal to every element u_k in the set $\mathcal{U}$

The result follows easily by taking the scalar products (h, u_k) for different k.

Grahm Schmidt Orthogonalization Procedure

Let $\mathcal{X} = \{x_1, x_2, \ldots, x_r\}$ be a linearly independent set. Then one can construct a set of vectors $\mathcal{E} = \{e_1, e_2, ...e_r\}$ such that the vectors e_k are linear combinations of the vectors in $\mathcal{X}$ and the set $\mathcal{E}$ is an orthonormal set.

Proof:

Define

$$\begin{aligned} u_1 &= x_1, & e_1 &= u_1/\|u_1\| \\ u_2 &= x_2 - (e_1, x_2)e_2, & e_2 &= u_2/\|u_2\| \\ u_3 &= x_3 - (e_1, x_3)e_3 - (e_2, x_3)e_2, & e_3 &= u_3/\|u_3\| \\ u_r &= x_r - \sum_{k=1}^{r-1} (e_k, x_r)e_k, & e_r &= u_r/\|u_r\| \end{aligned}$$

It is easily verified that $\{e_1, e_2, ...\}$ is an o.n. set.

Bessel's Inequality

If $\mathcal{U} = u_1, u_2, ..., u_r$ is any finite orthonormal set in an inner product space then for all $x \in \mathcal{V}$ we have

$$\sum_k |(u_k, x)|^2 \leq \|x\|^2 \qquad \text{(Bessel Inequality)} \tag{15}$$

Proof :

For every vector y , we have $(y, y) \geq 0$. Therefore, taking y to be

$$y = x - \sum_k \lambda_k u_k \qquad \text{with } u_k = (u_k, x).$$

we get

$$(y,y) = (x - \sum_k \lambda_k u_k, x - \sum_j \lambda_j u_j) \tag{16}$$
$$= (x,x) - \sum_k \lambda_k^*(u_k, x) - \sum_j \lambda_j (x, u_j) + \sum_j \sum_k \lambda_k^* \lambda_j (u_j, u_k) \tag{17}$$
$$= (x,x) - \sum_k \lambda_k^*(u_k, x) - \sum_j \lambda_j (x, u_j) + \sum_k \lambda_k^* \lambda_k \tag{18}$$

One of two the summations in the last term has been done using $(u_j, u_k) = \delta_{jk}$. Substituting $\lambda_j = (u_j, x)$ we get

$$(y,y) = (x,x) - \sum (x, u_k)(u_k, x) - \sum (u_j, x)(x, u_j) + \sum (x, u_j)(u_j, x) \tag{19}$$
$$= (x,x) - \sum_k (x, u_k)(u_k, x) \tag{20}$$
$$= (x,x) - \sum_k |(u_k, x)|^2 \tag{21}$$

Using $(y,y) \geq 0$ we get the desired Bessel's inequality.

$$\sum_k |(u_k, x)|^2 \leq \|x\|^2 \tag{22}$$

§3 Cauchy Schwarz and Triangle Inequalities

Cauchy Schwarz Inequality

As a preparation we first prove an intermediate result.

Theorem 3 *If f is a given vector and $g \neq 0$ be any vector $\|f - \lambda g\|$ is minimum when $\lambda = \lambda_0$ where*

$$\lambda_0 = \frac{(f,g)^*}{\|g\|^2} = \frac{(g,f)}{(g,g)}$$

and the minimum value of $\|f - \lambda g\|$ is given by

$$\|f - \lambda g\|_{min} = \|f\|^2 - |(f,g)|^2/\|g\|^2$$

Proof:
Let $F(\lambda) = \|f - \lambda g\|^2$. We compute $F(\lambda)$, write it as function of the real and imaginary parts of $\lambda(\equiv \alpha + i\beta)$ and minimize $F(\lambda)$ w.r.t. α and β.

$$\begin{aligned} F(\lambda) &= \|f - \lambda g\|^2 && (23)\\ &= (f - \lambda g, f - \lambda g) && (24)\\ &= (f,f) - \lambda(f,g) - \lambda^*(g,f) + \|\lambda\|^2(g,g) && (25) \end{aligned}$$

Substituting $\lambda = \alpha + i\beta$ we get

$$F(\lambda) = (f,f) - \alpha[(f,g) + (g,f)] + i\beta[(g,f) - (f,g)] + (\alpha^2 + \beta^2)(g,g)$$

Note that the right hand side has to be real. WHY ?! Setting

$$\frac{\partial F}{\partial \alpha} = 0, \qquad \text{and} \qquad \frac{\partial F}{\partial \beta} = 0$$

we get

$$\begin{aligned} -(f,g) - (g,f) + 2\alpha(g,g) = 0 && (26)\\ i(g,f) - i(f,g) + 2\beta(g,g) = 0 && (27)\\ && (28) \end{aligned}$$

hence

$$\begin{aligned} \alpha = [(f,g) + (g,f)]/2(g,g) && (29)\\ \beta = i[(f,g) - (g,f)]/2(g.g) && (30)\\ && (31) \end{aligned}$$

This gives the desired value λ_0 corresponding to the minimum of $F(\lambda)$ as

$$\lambda_0 = \alpha + i\beta = \frac{(g,f)}{(g,g)} = \frac{(f,g)^*}{(g,g)}$$

and the minimum value of $F(\lambda_0)$ is then computed to be

$$F(\lambda)|_{\min} = (f,f) - |(f,g)|^2/(g,g)$$

Theorem 4 (Cauchy Schwarz Inequality) *Let $f, g \in \mathcal{V}$. Then*

$$|(f,g)| \leq \|f\| \, \|g\|$$

The equality holds if and only if f and g are linearly dependent.

Proof :If $f = 0$ or $g = 0$, the equality holds trivially and there is nothing to prove because both sides are zero. Therefore, we assume $g \neq 0$. Consider $x = f - \lambda g$. Then we have $\|x\| \geq 0$ for all values of λ. We find the minimum of $\|x\|$ and set it ≥ 0

$$\min \|x\|^2 \geq 0$$

Using the previous result $\|x\|^2 = \|f - \lambda g\|^2$ is minimum when λ is equal to $\frac{(g,f)}{(g,g)}(\equiv \lambda_0)$ and minimum value of $\|x\|^2 = \|f - \lambda g\|^2$ is given by

$$\min \|x\|^2 = (f,f) - \frac{|(f,g)|^2}{(g,g)}$$

Thus we get

$$(f,f) - \frac{|(f,g)|^2}{(g,g)} \geq 0$$

or

$$(f,f)(g,g) \geq |(f,g)|^2$$

which is just the desired Cauchy Schwarz inequality

$$|(f,g)| \leq \|f\| \, \|g\|.$$

Note that when the Cauchy Schwarz inequality becomes equality $\min\|x\|^2 = 0$. This is possible if and only if $x = 0$ for $\lambda = \lambda_0$. This gives $f - \lambda_0 g = 0$ which means that f and g are linearly dependent.

Triangle Inequality

We are now in a position to prove the triangle inequalities

$$\|f + g\| \leq \|f\| + \|g\|$$

$$\|f - g\| \leq \|f\| + \|g\|$$

Proof : Consider

$$\begin{aligned}
\|f+g\|^2 &= (f+g, f+g) && (32)\\
&= (f,f) + (g,g) + (f,g) + (g,f) && (33)\\
&= (f,f) + (g,g) + 2Re(f,g) && (34)\\
&\leq \|f\|^2 + \|g\|^2 + 2|(f,g)| [\because Re z \leq |z|] && (35)
\end{aligned}$$

Using the Cauchy Schwarz inequality we get

$$\begin{aligned}
\|f+g\|^2 &= \|f\|^2 + \|g\|^2 + 2\|f\|\|g\| && (36)\\
&= [\|f\| + \|g\|]^2 && (37)
\end{aligned}$$

$\therefore$ We get the desired inequality

$$\|f + g\| \; \leq \|f\| \; + \|g\|$$