MATH 725 NOTES
INNER PRODUCT SPACES

MARIO L. GUTIERREZ ABED

Basics of Inner Product Spaces

Definition. *Let V be a vector space over $\mathbb{F}$ (where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$). An **inner product** on V is a function $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{F}$ with the following properties:*

- *For all $v \in V$, the inner product $\langle v, v \rangle$ is real and*
$$\langle v, v \rangle \geq 0 \qquad \text{and} \qquad \langle v, v \rangle = 0 \iff v = 0. \qquad \textit{(Positive definiteness)}$$
- *For $u, v \in V$, we have*
$$\langle u, v \rangle = \overline{\langle v, u \rangle} \qquad \textit{(Conjugate symmetry)}$$
- *For all $u, v, w \in V$ and $\alpha, \beta \in \mathbb{F}$, we have*
$$\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle \qquad \textit{(Linearity in the first coordinate)}$$
- *Combining the linearity in the first coordinate with conjugate symmetry, in the complex case we get*
$$\langle w, \alpha u + \beta v \rangle = \overline{\langle \alpha u + \beta v, w \rangle} = \overline{\alpha}\overline{\langle u, w \rangle} + \overline{\beta}\overline{\langle v, w \rangle} = \overline{\alpha}\langle w, u \rangle + \overline{\beta}\langle w, v \rangle.$$

*A vector space V, together with an inner product, is called an **inner product space**.* ★

<u>*Remark*</u>: The very last property on the definition is referred to as ***conjugate linearity*** in the second coordinate. Thus, a complex inner product is linear in its first coordinate and conjugate linear in its second coordinate. This is often described by saying that the inner product is ***sesquilinear*** (sesqui means "one and a half times"). In the real case when $\mathbb{F} = \mathbb{R}$, the inner product is linear in both coordinates—a property referred to as ***bilinearity***.

Examples. *1) The vector space $\mathbb{R}^n$ is an inner product space under the **standard inner product**, or **dot product**, defined by*
$$\langle (r_1, \ldots, r_n), (s_1, \ldots, s_n) \rangle = r_1 s_1 + \cdots + r_n s_n.$$
*The inner product space $\mathbb{R}^n$ is often called the n-**dimensional Euclidean space**.*

2) The vector space $\mathbb{C}^n$ is an inner product space under the standard inner product defined by
$$\langle (r_1, \ldots, r_n), (s_1, \ldots, s_n) \rangle = r_1 \overline{s_1} + \cdots + r_n \overline{s_n}.$$
*The inner product space is often called the n-**dimensional unitary space**.*

3) The vector space $C[a,b]$ of all continuous complex-valued functions on the closed interval $[a,b]$ is a complex inner product space under the inner product

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx.$$

4) One of the most important inner product spaces is the vector space ℓ^2 of all real (or complex) sequences (s_n) with the property that $\sum |s_n|^2 < \infty$, under the inner product

$$\langle (s_n), (t_n)\rangle = \sum_{n=0}^{\infty} s_n\overline{t_n}.$$

<u>Remark</u>: Of course, for this last inner product to make sense, the sum on the right must converge. To see that this is indeed the case, note that if $(s_n), (t_n) \in \ell^2$, then

$$\begin{aligned} 0 &\leq (|s_n| - |t_n|)^2 \\ &= |s_n|^2 - 2|s_n||t_n| + |t_n|^2 \\ \Longrightarrow\ & 2|s_n t_n| \leq |s_n|^2 + |t_n|^2. \end{aligned}$$

This yields

$$2\left|\sum_{n=0}^{\infty} s_n\overline{t_n}\right| \leq 2\sum_{n=0}^{\infty} |s_n t_n| \leq \sum_{n=0}^{\infty} |s_n|^2 + \sum_{n=0}^{\infty} |t_n|^2 < \infty.$$

Lemma 1. *If V is an inner product space and $\langle u, x\rangle = \langle v, x\rangle$ for all $x \in V$, then $u = v$.*

Definition. *If V is an inner product space, the **norm** of $v \in V$ is defined by*

$$\|v\| = \sqrt{\langle v, v\rangle}. \qquad \bigstar$$

Theorem 1. *Here are the basic properties of the norm:*

- *i) $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$.*
- *ii) $\|\alpha v\| = |\alpha|\|v\|$ for all $\alpha \in \mathbb{F}$ and $v \in V$.*
- *iii)* **(Cauchy-Schwarz Inequality)** *For all $u, v \in V$, we have*

  $$|\langle u, v\rangle| \leq \|u\|\|v\|,$$

  with equality if and only if one of u and v is a scalar multiple of the other.
- *iv)* **(Triangle Inequality)** *For all $u, v \in V$, we have*

  $$\|u + v\| \leq \|u\| + \|v\|.$$

  with equality if and only if one of u and v is a scalar multiple of the other.
- *v) For all $u, v, x \in V$, we have*

  $$\|u - v\| \leq \|u - x\| + \|x - v\|.$$
- *vi) For all $u, v \in V$, we have*

  $$|\|u\| - \|v\|| \leq \|u - v\|.$$

vii) **(The Parallelogram Law)** *For all $u, v \in V$, we have*

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Proof. We prove only Cauchy-Schwarz and the triangle inequality. For Cauchy-Schwarz, if either u or v is zero the result follows, so assume that $u, v \neq 0$. Then, for any scalar $\alpha \in \mathbb{F}$,

$$\begin{aligned} 0 &\leq \|u - \alpha v\|^2 \\ &= \langle u - \alpha v, u - \alpha v\rangle \\ &= \langle u, u\rangle - \bar{\alpha}\langle u, v\rangle - \alpha[\langle v, u\rangle - \bar{\alpha}\langle v, v\rangle]. \end{aligned}$$

Choosing $\bar{\alpha} = \langle v, u\rangle / \langle v, v\rangle$ makes the value in the square brackets equal to 0 and so

$$0 \leq \langle u, u\rangle - \frac{\langle v, u\rangle\langle u, v\rangle}{\langle v, v\rangle} = \|u\|^2 - \frac{|\langle u, v\rangle|^2}{\|v\|^2},$$

which is equivalent to the Cauchy-Schwarz inequality. Furthermore, equality holds if and only if $\|u - \alpha v\|^2 = 0$, that is, if and only if $u - \alpha v = 0$, which is equivalent to u and v being scalar multiples of one another.

Now to prove the triangle inequality, the Cauchy-Schwarz inequality gives

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2, \end{aligned}$$

from which the triangle inequality follows. The proof of the statement concerning equality is left to the reader. □

Definition. *A vector space V, together with a function $\|\cdot\| \colon V \to \mathbb{R}$ that satisfies the first two properties of the above theorem, along with the triangle inequality property, is called a **normed linear space**. Thus, any inner product space is a normed linear space, under the norm given by the inner product.* ★

It is interesting to observe that the inner product on V can be recovered from the norm, as we show in the following theorem:

Theorem 2 (**Polarization Identity**). *If V is a complex inner product space, then*

$$\langle u, v\rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right) + \frac{1}{4}i\left(\|u+iv\|^2 - \|u-iv\|^2\right).$$

Notice that if V is instead a real inner product space, then we are left with the real part of the right hand side on the equality above. That is,

$$\langle u, v\rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right).$$

Definition. *Let V and W be inner product spaces and let $\tau \in \mathcal{L}(V, W)$. Then*

- *τ is called an **isometry** if it preserves the inner product, that is, if*
$$\langle \tau(u), \tau(v)\rangle = \langle u, v\rangle \qquad \forall\ u, v \in V.$$
- *A bijective isometry is called an **isometric isomorphism**. When $\tau\colon V \to W$ is a bijective isometry, we say that V and W are **isometrically isomorphic**.* ★

Remark: It is not hard to show that an isometry is injective and so it is an isometric isomorphism provided it is also surjective. Moreover, if $\dim V = \dim W < \infty$, then injectivity implies surjectivity and so the concepts of isometry and isometric isomorphism are equivalent. However, as the following example shows, this is not the case for infinite-dimensional inner product spaces:

Example: Let $\tau\colon \ell^2 \to \ell^2$ be defined by
$$\tau(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots).$$
Then τ is an isometry (known as the *right shift operator*), but it is clearly not surjective.

Theorem 3. *A linear transformation $\tau \in \mathcal{L}(V, W)$ is an isometry if and only if it preserves the norm, that is, if and only if*
$$\|\tau(v)\| = \|v\|.$$

Hilbert Spaces

Definition. *A set $\mathcal{H}$ is a **Hilbert space** if it satisfies the following:*

i) *$\mathcal{H}$ is a vector space over $\mathbb{C}$ (or $\mathbb{R}$).*

ii) *$\mathcal{H}$ is equipped with an inner product $\langle\cdot,\cdot\rangle$, so that*
  - *$f \mapsto \langle f, g\rangle$ is linear on $\mathcal{H}$ for every $g \in \mathcal{H}$.*
  - *$\langle f, g\rangle = \overline{\langle g, f\rangle}$.*
  - *$\langle f, f\rangle \geq 0$ for all $f \in \mathcal{H}$.*

iii) *We let $\|f\| = \sqrt{\langle f, f\rangle}$. Then $\|f\| = 0$ if and only if $f = 0$.*

iv) *$\mathcal{H}$ is complete in the metric $d(f, g) = \|f - g\| = \sqrt{\langle f - g, f - g\rangle}$.*

Remark 1: Notice that the Cauchy-Schwarz and triangle inequalities
$$|\langle f, g\rangle| \leq \|f\|\|g\| \qquad \text{and} \qquad \|f + g\| \leq \|f\| + \|g\|$$
are in fact easy consequences of assumptions *i)* and *ii)* of our definition.

Remark 2: Notice that saying that $\mathcal{H}$ a *Hilbert space* is the same as saying that $\mathcal{H}$ is a *Banach space* (i.e., a *complete* normed linear space), with the norm induced by an inner product $\langle\cdot,\cdot\rangle$.

Definition. *Let $\mathcal{H}$ and $\mathcal{H}'$ be Hilbert spaces with respective inner products $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and $\langle \cdot, \cdot \rangle_{\mathcal{H}'}$ and the corresponding norms $\| \cdot \|_{\mathcal{H}}$ and $\| \cdot \|_{\mathcal{H}'}$. A mapping $U \colon \mathcal{H} \to \mathcal{H}'$ between these spaces is called* ***unitary*** *if:*

- *U is linear, i.e., $U(\alpha f + \beta g) = \alpha U(f) + \beta U(g)$.*
- *U is a bijection.*
- *$\|Uf\|_{\mathcal{H}'} = \|f\|_{\mathcal{H}}$ for all $f \in \mathcal{H}$.* ★

Definition. *A nonempty set $\mathcal{O} = \{u_i \mid i \in \Lambda\}$ of vectors in an inner product space is said to be an* ***orthogonal set*** *if $u_i \perp u_j$ for all $i \neq j \in \Lambda$. If, in addition, each vector u_i is a unit vector, then the set $\mathcal{O}$ is an* ***orthonormal set****. Thus, a set is orthonormal if*

$$\langle u_i, u_j \rangle = \delta_{i,j} \qquad \forall\, i, j \in \Lambda.$$

Note that an orthonormal set $\mathcal{O}$ is ***maximal*** *if and only if $\mathcal{O}^{\perp} = \{0\}$.* ★

Definition. *A maximal orthonormal set in an inner product space V is called a* ***Hilbert basis*** *for V.* ★

<u>*Remark 1*</u>: In general, a Hilbert space $\mathcal{H}$ has a Hilbert basis $\{e_i\}$ if the e_i form an orthonormal basis and every element $v \in \mathcal{H}$ can be written as

$$v = \sum_{i=1}^{\infty} a_i e_i$$

for some a_i with $\sum |a_i|^2 < \infty$.

<u>*Remark 2*</u>: Extreme care must be taken here not to confuse the concepts of a basis for a vector space and a Hilbert basis for an inner product space. To avoid confusion, a vector space basis, that is, a maximal linearly independent set of vectors, is referred to as a ***Hamel basis***.

The following example shows that, in general, the two concepts of bases discussed above are not the same:

<u>*Example*</u>:
• Let $V = \ell^2$ and let M be the set of all vectors of the form

$$e_i = (0, \dots, 0, 1, 0, \dots),$$

where e_i has a 1 in the i^{th} coordinate and 0's elsewhere. Clearly, M is an orthonormal set. Moreover, it is maximal. For if $v = (x_n) \in \ell^2$ has the property that $v \perp M$, then

$$x_i = \langle v, e_i \rangle = 0$$

for all i and so $v = 0$. Hence, no nonzero vector $v \notin M$ is orthogonal to M. This shows that M is a Hilbert basis for the inner product space ℓ^2.

• On the other hand, the vector space $\operatorname{span}(M)$ is the subspace S of all sequences in ℓ^2 that have finite support, and since $\operatorname{span}(M) = S \neq \ell^2$, we see that M is not a Hamel basis for the vector space ℓ^2.

Proposition 1. *A Hilbert space $\mathcal{H}$ is separable*[1] *if and only if $\mathcal{H}$ has a countable Hilbert basis. Moreover, if $\mathcal{H}$ is separable, all Hilbert bases are countable.*

Theorem 4. *Let V be an inner product space.*

i) **(Gram-Schmidt Orthogonalization)** *If $\mathcal{B} = (v_1, v_2, \dots)$ is a linearly independent sequence in V, then there is an orthogonal sequence $\mathcal{O} = (u_1, u_2, \dots)$ in V for which*

$$\operatorname{span}(u_1, \dots, u_n) = \operatorname{span}(v_1, \dots, v_n) \qquad \forall\, n > 0.$$

ii) If $\dim V = n$ is finite, then V has a Hilbert basis of size n and all Hilbert bases for V have size n.

iii) If V has a finite Hilbert basis of size n, then $\dim V = n$.

Theorem 5 (**Gram-Schmidt Orthogonalization Process**)**.** *If $\mathcal{B} = (v_1, v_2, \dots)$ is a sequence of linearly independent vectors in an inner product space V, then the sequence $\mathcal{O} = (u_1, u_2, \dots)$ defined by*

$$u_k = v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, u_i \rangle}{\langle u_i, u_i \rangle} u_i$$

is an orthogonal sequence in V with the property that

$$\operatorname{span}(u_1, \dots, u_k) = \operatorname{span}(v_1, \dots, v_k) \qquad \forall\, k > 0.$$

Of course, from the orthogonal sequence (u_i), we get the orthonormal sequence (w_i), where $w_i = u_i / \|u_i\|$.

Remark: Orthonormal bases have a great advantage over arbitrary bases. From a computational point of view, if $\mathcal{B} = \{v_1, \dots, v_n\}$ is a basis for V, then each $v \in V$ has the form

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

In general, however, determining the coordinates α_i requires solving a system of linear equations of size $n \times n$. On the other hand, if $\mathcal{O} = \{u_1, \dots, u_n\}$ is an orthonormal basis for V and

$$v = \alpha_1 u_1 + \cdots + \alpha_n u_n,$$

then the coefficients are quite easily computed:

$$\langle v, u_i \rangle = \langle \alpha_1 u_1 + \cdots + \alpha_n u_n, u_i \rangle = \alpha_i \langle u_i, u_i \rangle = \alpha_i.$$

Definition. *If S is a subspace of a Hilbert space $\mathcal{H}$, we define the* ***orthogonal complement*** *of S by*

$$S^\perp = \{ f \in \mathcal{H} \mid \langle f, g \rangle = 0 \quad \forall\, g \in S \}. \qquad \bigstar$$

[1] Recall that a space is said to be ***separable*** if it contains a countable dense subset (e.g. $\mathbb{R}$ is separable because it contains $\mathbb{Q}$, which is countable and is also dense in $\mathbb{R}$).

Remark 1: Clearly, $S^{\perp}$ is also a subspace of $\mathcal{H}$, and moreover $S \cap S^{\perp} = \{0\}$. To see this, note that if $f \in S \cap S^{\perp}$, then f must be orthogonal to itself; thus $0 = \langle f, f \rangle = \|f\|$, and therefore $f = 0$. Moreover, $S^{\perp}$ is itself a closed subspace. Indeed, if $f_n \to f$, then $\langle f_n, g \rangle \to \langle f, g \rangle$ for every g by the Cauchy-Schwarz inequality. Hence if $\langle f_n, g \rangle = 0$ for all $g \in S$ and all n, then $\langle f, g \rangle = 0$ for all those g.

Remark 2: Having seen the result stated on the above remark, we may ask ourselves whether or not the orthogonal complement $S^{\perp}$ is a vector space complement of S, that is, whether or not $V = S \oplus S^{\perp}$.

It turns out that if S is a finite-dimensional subspace of V, then the answer to this question is yes, but for infinite-dimensional subspaces, S must have the topological property of being complete[2] (or closed[3]).

In the next example we show that, in general, $V \neq S \oplus S^{\perp}$:

Example: As in a previous example, let $V = \ell^2$ and let S be the subspace of all sequences of finite support, that is, S is spanned by the vectors

$$e_i = (0, \ldots, 0, 1, 0, \ldots)$$

where e_i has a 1 in the i^{th} coordinate and 0's elsewhere. If $x = (x_n) \in S^{\perp}$, then

$$\begin{aligned} x_i = \langle x, e_i \rangle &= 0 \\ \implies x &= 0 \\ \implies S^{\perp} &= \{0\} \\ \implies S \oplus S^{\perp} &= S \neq \ell^2. \end{aligned}$$

The next lemma shows that closed subspaces enjoy an important characteristic property of Euclidean geometry:

Lemma 2. *Suppose S is a closed subspace of $\mathcal{H}$ and $f \in \mathcal{H}$. Then*

- *There exists a (unique) element $g_0 \in S$ which is closest to f, in the sense that*
$$\|f - g_0\| = \inf_{g \in S} \|f - g\|.$$
- *The element $f - g_0$ is perpendicular to S, that is,*
$$\langle f - g_0, g \rangle = 0 \qquad \forall\, g \in S.$$

Remark: The situation described in the above lemma can be visualized in Figure 1 below:

[2] Recall that a ***complete*** space is one in which every Cauchy sequence converges.

[3] By a ***closed*** subspace $S \subset \mathcal{H}$, we mean a subspace S in which for every sequence $\{f_n\} \subset S$ that converges to some $f \in \mathcal{H}$, we have that $f \in S$.

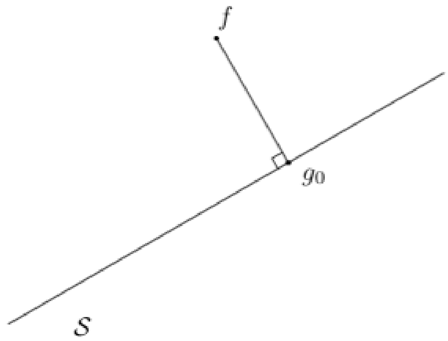

FIGURE 1. Nearest element to f in S.

As the next theorem shows, in the finite-dimensional case, orthogonal complements are also vector space complements. This theorem is often called the *projection theorem*, for reasons that will become apparent when we discuss projection operators:

Theorem 6 (**The Projection Theorem**). *If S is a finite-dimensional subspace of an inner product space V (which need not be finite-dimensional), then*

$$V = S \oplus S^{\perp}.$$

Now, here's the more generalized version of the theorem, in which we do not assume that S is finite:

Theorem 7. *If S is a closed subspace of a Hilbert space $\mathcal{H}$, then*

$$\mathcal{H} = S \oplus S^{\perp}.$$

<u>*Remark*</u>: The notation in the above theorem means that every $f \in \mathcal{H}$ can be written uniquely as $f = g + h$, where $g \in S$ and $h \in S^{\perp}$; we then say that $\mathcal{H}$ is the ***direct sum*** of S and $S^{\perp}$. This is equivalent to saying that any $f \in \mathcal{H}$ is the sum of two elements, one in S, the other in $S^{\perp}$, and that $S \cap S^{\perp}$ contains only 0.

With the decomposition $\mathcal{H} = S \oplus S^{\perp}$ one has the natural projection onto S defined by

$$\pi_S(f) = g, \qquad \text{where} \quad f = g + h \quad \text{and} \quad g \in S,\ h \in S^{\perp}.$$

The mapping π_S is called the ***orthogonal projection*** onto S and satisfies the following simple properties:

- $f \mapsto \pi_S(f)$ is linear.
- $\pi_S(f) = f$ whenever $f \in S$.
- $\pi_S(f) = 0$ whenever $f \in S^{\perp}$.
- $\|\pi_S(f)\| \leq \|f\|$ for all $f \in \mathcal{H}$.