MATH 325 Q1: LINEAR ALGEBRA III, PART 8

Inner product spaces

Definition. *Let F be the field of real or complex numbers and denote by $z \mapsto \bar{z}$ complex conjugation, given by*

$$a + b \cdot i \to \overline{a + b \cdot i} = a - b \cdot i.$$

An inner product on the F-vector space V is a map

$$\langle -, - \rangle \, : \, V \times V \longrightarrow F \, , \; (v, w) \longmapsto \langle v, w \rangle \, ,$$

subject to the following axioms:

- **(P)** $\langle v, v \rangle \geq 0$ *and* $\langle v, v \rangle = 0$ *if and only if* $v = 0$. *(Note that this in particular means that* $\langle v, v \rangle \in \mathbb{R}$.*)*
- **(L1)** $\langle \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2, w \rangle = \alpha_1 \cdot \langle v_1, w \rangle + \alpha_2 \cdot \langle v_2, w \rangle$ *for all* $\alpha_1, \alpha_2 \in F$ *and all* $v_1, v_2, w \in V$*; and*
- **(H)** $\langle v, w \rangle = \overline{\langle w, v \rangle}$ *for all* $v, w \in V$.

Definition. *The vector space V together with the inner product $\langle -, - \rangle$ is then called an inner product space. We call also the pair $(V, \langle -, - \rangle)$ an inner product space.*

Remark. *Note that if $W \subseteq V$ is a subspace then the restriction of the inner product $\langle -, - \rangle$ to W is an inner product on W and so W is also an inner product space.*

Before we give examples we mention the following consequences of the axioms. Let $\beta_1, \beta_2 \in F$ and $v, w, w_1, w_2 \in V$. Then:

(i)

$$\begin{aligned}
\langle v, \beta_1 \cdot w_1 + \beta_2 \cdot w_2 \rangle &= \overline{\langle \beta_1 \cdot w_1 + \beta_2 \cdot w_2, v \rangle} && \text{by } \textbf{(H)} \\
&= \overline{\beta_1 \cdot \langle w_1, v \rangle + \beta_2 \cdot \langle w_2, v \rangle} && \text{by } \textbf{(L1)} \\
&= \overline{\beta_1} \cdot \overline{\langle w_1, v \rangle} + \overline{\beta_2} \cdot \overline{\langle w_2, v \rangle} \\
&= \overline{\beta_1} \cdot \langle v, w_1 \rangle + \overline{\beta_2} \cdot \langle v, w_2 \rangle && \text{by } \textbf{(H)}.
\end{aligned}$$

We will call this "axiom" **(L2)**.

(ii) If $F = \mathbb{R}$ then **(H)** melts down to $\langle v, w \rangle = \langle w, v \rangle$.

(iii) We have $\langle 0, w\rangle = 0$ since $\langle 0, w\rangle = \langle 0+0, w\rangle = \langle 0, w\rangle + \langle 0, w\rangle$ by **(L1)** and so $\langle 0, w\rangle = 0$. Since $\langle v, 0\rangle = \overline{\langle 0, v\rangle}$ this implies also $\langle v, 0\rangle = 0$.

Examples.

(i) Let $F = \mathbb{R}$ and $V = \mathbb{R}^n$. Then the scalar product

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle := \sum_{i=1}^{n} x_i \cdot y_i$$

is an inner product on the n-dimensional vector space $\mathbb{R}^n$.

(ii) let $F = \mathbb{C}$. Then the scalar product

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle := \sum_{i=1}^{n} x_i \cdot \bar{y}_i$$

is an inner product on the n-dimensional vector space $\mathbb{C}^n$.

Exercise. Verify three axioms (P), (L1) and (H).

Definition. *The above two products are called the euclidean (inner) products on $\mathbb{R}^n$ and $\mathbb{C}^n$, respectively.*

Orthogonal vectors

Definition. *Let $(V, \langle -, - \rangle)$ be an inner product space, and $v \in V$. We say that the vector $w \in V$ is orthogonal to v if $\langle v, w\rangle = 0$.*

Remark. *Note that since $\langle v, w\rangle = \overline{\langle w, v\rangle}$ by* **(H)** *this property is symmetric,* i.e. *w is orthogonal to v if and only if v is orthogonal to w.*

Notation: $v \perp w$.

Notation. Let $S \subseteq V$ be a subset (not necessary a vector subspace). Then we denote by $S^\perp$ the set

$$S^\perp := \{ w \in V \mid w \perp s \text{ for all } s \in S \}$$

Lemma. *The set*

$$S^\perp := \{ w \in V \mid w \perp s \textit{ for all } s \in S \}$$

is a vector subspace of V.

Proof. Let $w_1, w_2 \in S^\perp$. Then

$$\langle s, w_1 + w_2 \rangle = \langle s, w_1 \rangle + \langle s, w_2 \rangle = 0$$

by **(L2)** for all $s \in S$, and this implies $w_1 + w_2 \in S^\perp$. Also, for all $\alpha \in F$ one has

$$\langle s, \alpha \cdot w_1 \rangle = \bar{\alpha} \cdot \langle s, w_1 \rangle = 0$$

again by **(L2)** for all $s \in S$, and this implies $\alpha \cdot w_1 \in S^\perp$. □

Notation. *If $S = \{v\}$ consists of a unique vector we will write instead of $S^\perp$ just simply $v^\perp$.*

Lemma. *Let $v_1, \ldots, v_n$ nonzero vectors in the inner product space $(V, \langle -, - \rangle)$ which are pairwise orthogonal, i.e. $v_i \perp v_j$ for all $1 \leq i \neq j \leq n$. Then $v_1, \ldots, v_n$ are linear independent.*

Proof. We prove this by induction on $n \geq 1$. The case $n = 1$ is clear.

Let $n \geq 2$. Assume we have a linear combination

$$a_1 \cdot v_1 + a_2 \cdot v_2 + \ \ldots \ + a_n \cdot v_n \ = \ 0 \, .$$

Let $w = \sum_{i=1}^{n-1} a_i \cdot v_i$. Then

$$\langle v_n, w \rangle \ = \ \sum_{i=1}^{n-1} \bar{a}_i \cdot \langle v_n, v_i \rangle$$

by **(L2)**, and therefore $\langle v_n, w \rangle = 0$ since $\langle v_n, v_i \rangle = 0$ for all $1 \leq i \leq n-1$.

By construction, we have $a_n \cdot v_n + w = 0$ and so

$$\begin{aligned} 0 &= \langle a_n \cdot v_n + w, w \rangle \\ &= a_n \cdot \langle v_n, w \rangle + \langle w, w \rangle \\ &= 0 + \langle w, w \rangle \\ &= \langle w, w \rangle \, . \end{aligned}$$

This implies $w = 0$ by (**P**). Hence by induction

$$a_1 = a_2 = \ldots = a_{n-1} = 0$$

and we are done. □

The norm

Definition. *Let V be an inner product space with an inner product $\langle -, - \rangle$. For $v \in V$ the norm or length of v is then defined as*

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

Lemma. *Let $v \in V$ and $c \in F$. Then we have*
(a) $\|v\| \geq 0$ *and* $\|v\| = 0$ *if and only if* $v = 0$*;*
(b) $\|c \cdot v\| = |c| \cdot \|v\|$ *where* $|c|$ *is the usual absolute value of the real respectively complex number* c*.*

Proof. This follows from the axioms of the inner product space. $\square$

Example. Let $V = \mathbb{R}^3$ be the 3-dimensional real vector space with the usual scalar product

$$\langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \rangle := x_1 \cdot y_1 + x_2 \cdot y_2 + x_3 \cdot y_3\,.$$

Then we have

$$\left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right\| = \sqrt{x_1^2 + x_2^2 + x_3^2}\,,$$

and so the norm of a vector is the usual euclidean length of it.

We now prove the following important properties of an inner product and its associated norm.

Theorem. *Let $(V, \langle -, - \rangle)$ be an inner product space and $\| - \|$ the associated norm. Then:*

(i) **(Cauchy-Schwarz inequality)** *For all $v, w \in V$ one has*

$$| < v, w > | \leq \|v\| \cdot \|w\|.$$

(ii) **(Triangle inequality)** $\|v + w\| \leq \|v\| + \|w\|$ *for all $v, w \in V$.*

Proof. (i) If $v = 0$ or $w = 0$ this is obvious. So assume $v \neq 0$. For any scalar α we have then by **(P)**

$$0 \leq \|\alpha \cdot v - w\|^2 = \langle \alpha \cdot v - w, \alpha \cdot v - w \rangle\,,$$

and so by **(L1)**, **(L2)**, and **(H)** we get

$$0 \leq (\alpha \cdot \bar{\alpha}) \cdot \langle v, v \rangle - \alpha \cdot \langle v, w \rangle - \bar{\alpha} \cdot \langle w, v \rangle + \langle w, w \rangle\,.$$

This inequality is also true for $\alpha := \frac{\langle w, v \rangle}{\langle v, v \rangle}$, which gives using that

$$|\langle v, w \rangle|^2 = \langle v, w \rangle \cdot \overline{\langle v, w \rangle} = \langle v, w \rangle \cdot \langle w, v \rangle\,,$$

where the latter equality is a consequence of axiom **(H)**, the inequality

$$0 \leq -\frac{|\langle v, w\rangle|^2}{\langle v, v\rangle} + \langle w, w\rangle\,.$$

Therefore

$$|\langle v, w\rangle|^2 \;\leq\; \|v\|^2 \cdot \|w\|^2\,,$$

from which the claimed inequality follows by taking square roots on both sides.

We prove now (ii). We have

$$\begin{aligned}\|v+w\|^2 &= \langle v+w, v+w\rangle \\ &= \|v\|^2 + \langle v, w\rangle + \langle w, v\rangle + \|w\|^2 \quad \text{by } \textbf{(L1)} \text{ and } \textbf{(L2)} \\ &= \|v\|^2 + \|w\|^2 + \langle v, w\rangle + \overline{\langle v, w\rangle} \quad \text{by } \textbf{(H)}.\end{aligned}$$

Side Remark. For a complex number z one has $z+\bar{z} \leq 2\cdot|z|$. Indeed, setting $z = x + i\cdot y$, we have

$$z + \bar{z} = (x + i\cdot y) + (x - i\cdot y) = 2x$$

and $|z| = \sqrt{x^2+y^2} \geq x$.

Using this remark we obtain

$$\langle v, w\rangle + \overline{\langle v, w\rangle} \;\leq\; 2\cdot|\langle v, w\rangle| \;\leq\; 2\cdot\|v\|\cdot\|w\|$$

(the latter inequality by Cauchy-Schwarz, *i.e.* the already proven part (i) of the theorem). Inserting this in above equation we get

$$\|v+w\|^2 \;\leq\; \|v\|^2 + 2\cdot(\|v\|\cdot\|w\|) + \|w\|^2 \;=\; (\|v\| + \|w\|)^2\,,$$

from which the triangle inequality follows by taking square roots on both sides. □

Orthogonal bases

Definition. *Let V be a F-vector space with an inner product $\langle -, -\rangle$. A basis $\{v_1, \ldots, v_n\}$ of V is called an orthogonal basis if*

- **(O)** *The vectors $v_1, \ldots, v_n$ are pairwise orthogonal,* i.e. *$v_i \perp v_j$ for all $1 \leq i \neq j \leq n$.*

If moreover

- **(N)** *all v_i have norm 1,* i.e. *$\|v_i\| = \sqrt{\langle v_i, v_i\rangle} = 1$ for all $1 \leq i \leq n$*

then the basis $v_1, \ldots, v_n$ is called an orthonormal basis.

Remark. *We observe that if $v_1, \dots, v_n$ is an orthogonal basis of $(V, \langle -, - \rangle)$ then*

$$\frac{1}{\|v_1\|} \cdot v_1 \,,\ \dots\ ,\ \frac{1}{\|v_n\|} \cdot v_n$$

is an orthonormal basis. Indeed, we have $\|\alpha \cdot v\| = |\alpha| \cdot \|v\|$ for all $\alpha \in F$ and $v \in V$.

Given an orthogonal basis $v_1, \dots, v_n$ of an inner product space $(V, \langle -, - \rangle)$ and $v \in V$ we can write $v = \sum_{i=1}^{n} a_i \cdot v_i$ for some uniquely determined scalars $a_1, \dots, a_n \in F$.

Lemma. *The above coefficients a_i are given by $a_i = \frac{\langle v_i, v \rangle}{\langle v_i, v_i \rangle}$.*

Proof. Using **(L1)** we get

$$\langle v, v_j \rangle \ = \ \langle \sum_{i=1}^{n} a_i \cdot v_i, v_j \rangle \ = \ \sum_{i=1}^{n} a_i \cdot \langle v_i, v_j \rangle \ = \ a_j \langle v_j, v_j \rangle$$

(the last equation since $v_i \perp v_j$ for $i \neq j$), and so

$$a_j \ = \ \frac{\langle v, v_j \rangle}{\langle v_j, v_j \rangle}$$

for all $1 \leq j \leq n$. □

Theorem. *Let $(V, \langle -, - \rangle)$ be an inner product space, and $w_1, \dots, w_n$ be a set of linear independent vectors in V. Then there exists vectors $v_1, \dots, v_n$ which are orthogonal to each other and whose span is equal the span of $w_1, \dots, w_n$:* $\mathrm{Span}\{w_1, \dots, w_n\} = \mathrm{Span}\{v_1, \dots, v_n\}$. *In particular, V has an orthogonal basis.*

Proof. We prove this by induction on $n \geq 1$ using the so-called *Gram-Schmidt orthogonalization process.*

If $n = 1$ we set $v_1 = w_1$. So let $n \geq 2$. By induction we can assume that there is a system of pairwise orthogonal vectors $v_1, \dots, v_{n-1}$ which span the subspace generated by the first $(n-1)$ vectors $w_1, \dots, w_{n-1}$. We set then

$$v_n \ := \ w_n - \sum_{i=1}^{n-1} \frac{\langle w_n, v_i \rangle}{\langle v_i, v_i \rangle} \cdot v_i \,.$$

Since $w_n = v_n + \sum_{i=1}^{n-1} \frac{\langle w_n, v_i \rangle}{\langle v_i, v_i \rangle} \cdot v_i$ the vector w_n is in the span of $v_1, \dots, v_n$ and so the vectors $v_1, \dots, v_{n-1}, v_n$ generate the same subspace as the vectors $w_1, \dots, w_{n-1}, w_n$.

We are left to show that the vectors $v_1, \dots, v_n$ are pairwise orthogonal. This is true by induction for the subset $v_1, \dots, v_{n-1}$, and so we have only to show $v_n \perp v_i$ for all $i = 1, \dots, n-1$. But

$$\begin{aligned}\langle v_n, v_i\rangle &= \langle w_n - \textstyle\sum_{j=1}^{n-1} \frac{\langle w_n, v_j\rangle}{\langle v_j, v_j\rangle} \cdot v_j, v_i\rangle \\ &= \langle w_n, v_i\rangle - \textstyle\sum_{j=1}^{n-1} \frac{\langle w_n, v_j\rangle}{\langle v_j, v_j\rangle} \cdot \langle v_j, v_i\rangle \\ &= \langle w_n, v_i\rangle - \frac{\langle w_n, v_i\rangle}{\langle v_i, v_i\rangle} \cdot \langle v_i, v_i\rangle \\ &= 0\end{aligned}$$

for all $1 \leq i \leq n-1$. We are done. □

Example. We consider the 3-dimensional real vector space $\mathbb{R}^3$ with the usual scalar product $\langle -, -\rangle$. The vectors

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \ w_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \text{ and } w_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

form a basis of $\mathbb{R}^3$ which is however not an orthogonal basis as *e.g.*

$$\langle w_1, w_2\rangle = 2.$$

We use the Gram-Schmidt process to construct out of it an orthogonal basis v_1, v_2, v_3. We set $v_1 := w_1$. Then

$$v_2 \ := \ w_2 - \frac{\langle w_2, v_1\rangle}{\langle v_1, v_1\rangle} \cdot v_1 \ = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{pmatrix},$$

and

$$v_3 \ := \ w_3 - \sum_{i=1}^{2} \frac{\langle w_3, v_i\rangle}{\langle v_i, v_i\rangle} \cdot v_i \ = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix}.$$

Definition. *Let $(V, \langle -, -\rangle)$ be an inner product space, and $W \subseteq V$ be a subspace. The subspace*

$$W^{\perp} \ := \ \left\{ v \in V \,|\, v \perp w \text{ for all } w \in W \right\}$$

is called the orthogonal complement of W in V.

Remark. We proved before that $W^{\perp}$ is in fact a subspace of V.

Remark. *Let $v \in V$. Then it is easy to see that $v^{\perp} = (F \cdot v)^{\perp}$.*

The name "orthogonal complement" is justified by the following fact.

Theorem. *Let W be a subspace of the inner product space $(V, \langle -, - \rangle)$. Then any $v \in V$ can be uniquely written*

$$v = x + y$$

with $x \in W$ and $y \in W^{\perp}$. In particular, we have $\dim W + \dim W^{\perp} = \dim V$.

Proof. **Uniqueness**: We have seen before that if v, w are nonzero vectors which are orthogonal to each other then v and w are linear independent and so in particular then $v \neq w$. Hence if

$$x + y = v = x' + y'$$

with $x, x' \in W$ and $y, y' \in W^{\perp}$ then $x - x' = y' - y$. It is easy to see that $\langle x - x', y' - y \rangle = 0$, i.e. the vector $z = x - x' = y' - y$ is orthogonal to itself. This can happen if and only if $z = 0$ implying $x = x'$ and $y = y'$. This prove the uniqueness.

We now show the **existence** of this decomposition. Let $v_1, \ldots, v_d$ be an orthogonal basis of W. We can extend this basis to a basis $v_1, \ldots, v_d, w_{d+1}, \ldots, w_n$ of the whole space V. Applying the Gram-Schmidt process to this basis we get an orthogonal basis

$$v_1, \ldots, v_d, v_{d+1}, \ldots, v_n$$

of V. Hence the vectors $v_{d+1}, \ldots, v_n$ are in the orthogonal complement $W^{\perp}$ of W.

Let now $v \in V$. Since $v_1, \ldots, v_n$ is a basis of V we can write

$$v = \sum_{i=1}^{d} a_i \cdot v_i + \sum_{i=d+1}^{n} a_i \cdot v_i$$

for some $a_1, \ldots, a_n$ in F. By construction, the vector $x = \sum_{i=1}^{d} a_i \cdot v_i$ is in W, and $v_{d+1}, \ldots, v_n \in W^{\perp}$ and so $y = \sum_{i=d+1}^{n} a_i \cdot v_i$ is in $W^{\perp}$. Therefore $v = x + y$ is the claimed decomposition.

Finally we show the formula

$$\dim W + \dim W^{\perp} = \dim V.$$

By construction we have $W = \mathrm{Span}\{v_1, \ldots, v_d\}$ so that $\dim(W) = d$. Let also $W' = \mathrm{Span}\{v_{d+1}, \ldots, v_n\}$. Then $\dim(W') = n - d$ (because $v_{d+1}, \ldots, v_n$ are linearly independent). This implies

$$\dim(V) = \dim(W) + \dim(W').$$

We saw above that $W' \subset W^\perp$. Hence it remains to prove that $W^\perp = W'$. Let $v \in W^\perp$. Write $v = x + y$ where $x = a_1 v_1 + \cdots + a_d v_d \in W$ and $y \in W'$. Since v is orthogonal to all vectors in W it is orthogonal to $v_1, \ldots, v_d$. It follows that for every $1 \le i \le d$ one has

$$0 = \langle x+y, v_i\rangle = \langle x, v_i\rangle + \langle y, v_i\rangle = \langle x, v_i\rangle + 0 = \langle \sum a_j v_j, v_i\rangle = a_i\langle v_i, v_i\rangle$$

and this implies a_i. Thus $x = 0$, i.e $v = y \in W'$ as required. □

Corollary. *Let $(V, \langle -, -\rangle)$ be an inner product space over F and let $f : V \to F$ an F-linear map. Then there exists a unique vector $y \in V$, such that*

$$f(v) = \langle v, y\rangle$$

for all $v \in V$.

Proof. **Uniqueness:** Let y_1 and y_2 be two vectors such that

$$\langle v, y_1\rangle = \langle v, y_2\rangle.$$

Then we have

$$0 = \langle v, y_1\rangle - \langle v, y_2\rangle = \langle v, y_1 - y_2\rangle$$

for all $v \in V$. In particular this is also true for $v = y_1 - y_2$, *i.e.*

$$0 = \langle y_1 - y_2 . y_1 - y_2\rangle\,,$$

and so by **(P)** we have $y_1 - y_2 = 0$, *i.e.* $y_1 = y_2$.

Existence: let $W = \operatorname{Ker} f$. If $W = V$ then $f(v) = 0$ for all vectors $v \in V$ and we can chose $y = 0$.

Assume now that $W \neq V$. Recall that

$$\dim V = \dim \operatorname{Ker}(f) + \dim \operatorname{Image}(f)$$

and that the image of f has dimension at most 1 (because it lives in 1-dimensional vector space F). Then automatically $\operatorname{Im}(f) = F$ and hence $\dim(W) = \dim(V) - 1$. Applying the theorem above we get $\dim W^\perp = 1$.

Let v be a basis of $W^\perp$, *i.e.* $F \cdot v = W^\perp$. We claim that

$$y = \frac{\overline{f(v)}}{\langle v, v\rangle} \cdot v$$

does the job. (Note that $v \neq 0$ since $W^\perp \neq \{0\}$ and so by axiom **(P)** we have $\langle v, v\rangle > 0$.)

Take any $x \in V$. Write $x = w + w'$ with $w \in W$ and $w' \in W^\perp$, say $w' = \lambda \cdot v$ for some $\lambda \in F$. Then

$$
\begin{aligned}
\langle x, y\rangle &= \langle w + \lambda \cdot v\,,\ \tfrac{\overline{f(v)}}{\langle v,v\rangle} \cdot v\rangle \\
&= \langle w, \tfrac{\overline{f(v)}}{\langle v\,,\, v\rangle} \cdot v\rangle \ +\ \lambda \cdot \langle v, \tfrac{\overline{f(v)}}{\langle v,v\rangle} \cdot v\rangle && \text{by } \textbf{(L1)} \\
&= \lambda \cdot \langle v\,,\ \tfrac{\overline{f(v)}}{\langle v,v\rangle} \cdot v\rangle && \text{since } w \in W \text{ and } \tfrac{\overline{f(v)}}{\langle v,v\rangle} \cdot v \in W^\perp \\
&= \lambda \cdot \tfrac{f(v)}{\langle v,v\rangle} \cdot \langle v, v\rangle && \text{by } \textbf{(L2)}\text{, note that } \langle v, v\rangle \in \mathbb{R} \\
&= \lambda \cdot f(v) \ =\ f(\lambda \cdot v) && \text{since } f \text{ is linear} \\
&= f(\lambda \cdot v) + f(w) && \text{since } f(w) = 0 \text{ as } w \in W = \mathrm{Ker}(f) \\
&= f(\lambda \cdot v + w) \ =\ f(x) && \text{since } f \text{ is linear.}
\end{aligned}
$$

$\square$