17. Inner product spaces

Definition 17.1. *Let V be a real vector space. An* ***inner product*** *on V is a function*

$$\langle\ ,\ \rangle \colon V \times V \longrightarrow \mathbb{R},$$

which is

- ***symmetric****, that is*
$$\langle u, v\rangle = \langle v, u\rangle.$$
- ***bilinear****, that is linear (in both factors):*
$$\langle \lambda u, v\rangle = \lambda\langle u, v\rangle,$$
for all scalars λ and
$$\langle u_1 + u_2, v\rangle = \langle u_1, v\rangle + \langle u_2, v\rangle,$$
for all vectors u_1, u_2 and v.
- ***positive*** *that is*
$$\langle v, v\rangle \geq 0.$$
- ***non-degenerate*** *that is if*
$$\langle u, v\rangle = 0$$
for every $v \in V$ then $u = 0$.

We say that V is a ***real inner product space****. The* ***associated quadratic form*** *is the function*

$$Q \colon V \longrightarrow \mathbb{R},$$

defined by

$$Q(v) = \langle v, v\rangle.$$

Example 17.2. *Let $A \in M_{n,n}(\mathbb{R})$ be a real matrix. We can define a function*

$$\langle\ ,\ \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R},$$

by the rule

$$\langle u, v\rangle = u^t A v.$$

The basic rules of matrix multiplication imply that this function is bilinear. Note that the entries of A are given by

$$a_{ij} = e_i^t A e_j = \langle e_i, e_j\rangle.$$

In particular, it is symmetric if and only if A is ***symmetric*** *that is $A^t = A$. It is non-degenerate if and only if A is invertible, that is A has rank n. Positivity is a little harder to characterise.*

Perhaps the canonical example is to take $A = I_n$. In this case if $u = (r_1, r_2, \dots, r_n)$ and $v = (s_1, s_2, \dots, s_n)$ then $u^t I_n v = \sum r_i s_i$. Note *that if we take $u = v$ then we get $\sum r_i^2$. The square root of this is the Euclidean distance.*

Definition 17.3. *Let V be a real vector space. A* ***norm*** *on V is a function*

$$\|.\|: V \longrightarrow \mathbb{R},$$

what has the following properties

- $$\|kv\| = |k|\|v\|,$$
  for all vectors v and scalars k.
- ***positive*** *that is*
  $$\|v\| \geq 0.$$
- ***non-degenerate*** *that is if*
  $$\|v\| = 0$$
  then $v = 0$.
- *satisfies the* ***triangle inequality****, that is*
  $$\|u + v\| \leq \|u\| + \|v\|.$$

Lemma 17.4. *Let V be a real inner product space.*

Then

$$\|.\|: V \longrightarrow \mathbb{R},$$

defined by

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

is a norm on V.

Proof. It is clear that the norm satisfies the first property and that it is positive. Suppose that $u \in V$. By assumption there is a vector v such that

$$\langle u, v\rangle \neq 0.$$

Consider

$$\begin{aligned} 0 &\leq \langle u + tv, u + tv\rangle \\ &= \langle u, u\rangle + 2t\langle u, v\rangle + t^2\langle v, v\rangle. \end{aligned}$$

If $\langle v, v\rangle = 0$ then certainly $\langle u, u\rangle > 0$. Otherwise put

$$t = -\frac{\langle u, v\rangle}{\langle v, v\rangle} > 0.$$

Then

$$\langle u, u\rangle + 2t\langle u, v\rangle + t^2\langle v, v\rangle = \langle u, u\rangle + t\langle u, v\rangle.$$

Once again

$$\langle u, u\rangle > \frac{(\langle u, v\rangle)^2}{\langle v, v\rangle} > 0.$$

Thus the norm is non-degenerate.

Now suppose that u and $v \in V$. Then

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

Taking square roots gives the triangle inequality. □

Note that one can recover the inner product from the norm, using the formula

$$2\langle u, v\rangle = Q(u+v) - Q(u) - Q(v),$$

where Q is the associated quadratic form. Note the annoying appearence of the factor of 2.

Notice also that on the way we proved:

Lemma 17.5 (Cauchy-Schwarz-Bunjakowski)**.** *Let V be a real inner product space.*

If u and $v \in V$ then

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

Definition 17.6. *Let V be a real vector space with an inner product.*

We say that two vectors v and w are ***orthogonal*** *if*

$$\langle u, v\rangle = 0.$$

We say that a basis $v_1, v_2, \dots, v_n$ is an ***orthogonal basis*** *if the vectors $v_1, v_2, \dots, v_n$ are pairwise orthogonal. If in addition the vectors v_i have length one, we say that $v_1, v_2, \dots, v_n$ is an* ***orthonormal basis****.*

Lemma 17.7. *Let V be a real inner product space.*

(1) *If the vectors $v_1, v_2, \dots, v_m$ are pairwise orthogonal then they are independent. In particular if $m = \dim V$ then $v_1, v_2, \dots, v_m$ are an orthogonal basis of V.*
(2) *If $v_1, v_2, \dots, v_n$ are an orthonormal basis of V and $v \in V$ then*

$$v = \sum r_i v_i,$$

where

$$r_i = \langle v, v_i\rangle.$$

Proof. We first prove (1). Suppose that

$$r_1 v_1 + r_2 v_2 + \cdots + r_m v_m = 0.$$

Taking the inner product of both sides with v_j gives

$$\begin{aligned} 0 &= \langle r_1 v_1 + r_2 v_2 + \cdots + r_m v_m, v_j \rangle \\ &= \sum_{i=1}^{m} r_i \langle v_i, v_j \rangle \\ &= r_j \langle v_j, v_j \rangle. \end{aligned}$$

As

$$\langle v_j, v_j \rangle \neq 0,$$

it follows that $r_j = 0$. This is (1).

The proof of (2) is similar. □

So how does one find an orthonormal basis?

Algorithm 17.8 (Gram-Schmidt)**.** *Let $v_1, v_2, \dots, v_n$ be independent vectors in a real inner product space V.*

(1) *Let $1 \leq k \leq n$ be the largest index such that $v_1, v_2, \dots, v_k$ are orthonormal.*
(2) *If $k = m$ then* ***stop****.*
(3) *Otherwise let*

$$u_{k+1} = v_{k+1} - r_1 v_1 - r_2 v_2 - \cdots - r_m v_m,$$

where $r_i = \langle v_{k+1}, v_i \rangle$. Replace v_{k+1} by

$$\frac{u_{k+1}}{\|u_{k+1}\|},$$

and return to (1).

In practice the algorithm works as follows. First we replace v_1 by

$$\frac{v_1}{\|v_1\|},$$

so that v_1 has unit length. Then we consider v_2. We have to subtract some of v_1 to ensure that it is orthogonal to v_1. So consider a vector of the form

$$u = v_2 + \lambda v_1,$$

where λ is chosen to make u orthogonal to v_1. We have

$$0 = \langle u, v_1 \rangle = \langle v_2, v_1 \rangle + \lambda \langle v_1, v_1 \rangle,$$

so that

$$\lambda = -\langle v_2, v_1 \rangle.$$

Then we rescale to get a vector of unit length. At the next stage, we can choose λ and μ so that

$$v_3 + \lambda v_1 + \mu v_2,$$

is orthogonal v_1 and v_2. The key thing is that since v_1 and v_2 are orthogonal, our choice of λ and μ are independent of each other.

For example, consider the vectors

$$v_1 = (1, -1, 1), \qquad v_2 = (1, 0, 1) \qquad \text{and} \qquad v_3 = (1, 1, 2),$$

in $\mathbb{R}^3$ with the usual inner product. The first step is to replace v_1 by

$$v_1 = \frac{1}{\sqrt{3}}(1, -1, 1).$$

Now let

$$\begin{aligned} u &= (1, 0, 1) - \frac{2}{3}(1, -1, 1) \\ &= \frac{1}{3}(1, 2, 1). \end{aligned}$$

Then we replace u by a vector parallel to u of unit length

$$v_2 = \frac{1}{\sqrt{6}}(1, 2, 1).$$

Finally we put

$$\begin{aligned} u &= (1, 1, 2) - \frac{2}{3}(1, -1, 1) - \frac{5}{6}(1, 2, 1) \\ &= \frac{1}{2}(-1, 0, 1). \end{aligned}$$

Finally we replace u by a vector of unit length,

$$v_3 = \frac{1}{\sqrt{2}}(-1, 0, 1).$$

Thus

$$v_1 = \frac{1}{\sqrt{3}}(1, -1, 1) \qquad v_2 = \frac{1}{\sqrt{6}}(1, 2, 1) \qquad \text{and} \qquad v_3 = \frac{1}{\sqrt{2}}(-1, 0, 1),$$

is the orthonormal basis produced by Gram-Schmidt.

One very useful property of inner products is that we get canonically defined complimentary linear subspaces:

Lemma 17.9. *Let V be a finite dimensional real inner product space.*
If $U \subset V$ is a linear subspace, then let

$$U^{\perp} = \{\, w \in V \mid \langle w, u \rangle = 0, \forall u \in U \,\},$$

the set of all vectors orthogonal to every element of U. Then

- *$U^{\perp}$ is a linear subspace of V.*
- *$U \cap U^{\perp} = \{0\}$.*
- *U and $U^{\perp}$ span V.*

In particular V is isomorphic to $U \oplus U^{\perp}$.

Proof. We first prove (1). Suppose that w_1 and $w_2 \in U^{\perp}$. Pick $u \in U$. Then

$$\begin{aligned} \langle w_1 + w_2, u \rangle &= \langle w_1, u \rangle + \langle w_2, u \rangle \\ &= 0 + 0 = 0. \end{aligned}$$

It follows that $w_1 + w_2 \in U^{\perp}$ and so $U^{\perp}$ is closed under addition. Now suppose that $w \in U^{\perp}$ and λ is a scalar. Then

$$\begin{aligned} \langle \lambda w, u \rangle &= \lambda \langle w, u \rangle \\ &= \lambda 0 = 0. \end{aligned}$$

Thus $\lambda w \in U^{\perp}$ and so $U^{\perp}$ is closed under scalar multiplication. Thus $U^{\perp}$ is a linear subspace of V. This is (1).

Suppose that $w \in U \cap U^{\perp}$. Then

$$\langle w, w \rangle = 0.$$

But then $w = 0$. This is (2).

Suppose that $v \in V$. If $v \in U$ there is nothing to prove. Otherwise pick an orthonormal basis $u_1, u_2, \dots, u_k$ of U. Then the vectors v, $u_1, u_2, \dots, u_k$ are independent. By Gram-Schmidt we may find scalars $r_1, r_2, \dots, r_k$ such that

$$w = v - \sum r_i u_i,$$

is orthogonal to $u_1, u_2, \dots, u_k$ (in fact $r_i = \langle v, u_i \rangle$). But then w is orthogonal to U, that is $w \in U^{\perp}$. Let $u = \sum r_i u_i \in U$. Then $v = u + w$. This is (3). $\square$