Lecture 9: Inner Product Spaces

Today we start Chapter 4.

Inner Product

Definition (Text, Definition 15.1)

An inner or dot product $(\ ,\ )$ on V is a function, which assigns to each pair of vectors u, v in V a real number. $(u,\, v)$ satisfies three axioms:

(i) **Bilinear**

$$\begin{aligned} (u+v,\, w) &= (u,\, w) + (v,\, w) \\ (u,\, v+w) &= (u,\, v) + (u,\, w) \\ (cu,\, v) &= (u,\, cv) = c\,(u,\, v)\,, \ \text{all } c \in \mathbb{R}. \end{aligned}$$

(ii) **Symmetric**

$$(u,\, v) = (v,\, u)\,, \ \text{all } u,\, v \in V.$$

(iii) **Positive Definite**
For all $u \in U$

$$(u,\, u) \geq 0$$

and

$$(u,\, u) = 0 \Longleftrightarrow u = 0.$$

Examples

(1) $\mathbb{R}^n$

$$((x_1,\, x_2,\, \ldots,\, x_n)\,,\, (y_1,\, y_2,\, \ldots,\, y_n)) = x_1y_1 + x_2y_2 + \ldots + x_ny_n.$$

(2) Let $C[0,\, 1]$ = continuous functions on $[0,\, 1]$.

$$(f,\, g) = \int_0^1 f(x)g(x)\,\mathrm{d}x$$

In any inner product space we can do Euclidean geometry, i.e., we can define lengths/distances and angles.

Definition

Let $v \in V$. We define the length of v, denoted $||v||$ by

$$||v|| = \sqrt{(v,\, v)}.$$

So in $\mathbb{R}^n$ with $v = (x_1,\, x_2,\, \ldots,\, x_n)$ we have

$$||v|| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

We define the distance between two vectors v and w by

$$\mathrm{d}\,(v,\, w) = ||v - w||.$$

This definition is motivated by the picture.

We define the unoriented angle $\measuredangle(u,\, v)$ (in $[0,\, \pi]$) between two vectors u and v by

$$\measuredangle(u,\, v) = \cos^{-1}\left(\frac{(u,\, v)}{||u||\,||v||}\right) \quad (*)$$

$\cos^{-1}$ has domain [-1, 1], so in order for $(*)$ to be a correct definition, we have to prove

$$-1 \leq \frac{(u,\, v)}{||u||\,||v||} \leq 1$$

The unoriented angle does not take into account the positive or negative rotation

$$\measuredangle(u,\, v) = \measuredangle(v,\, u)$$

The Cauchy-Schwartz Inequality for Inner Product Spaces

Theorem (Cauchy-Schwartz (CS))

$$|(u,\, v)| \leq ||u||\, ||v||$$

Proof. Let $u,\, v \in V$. Then for all $t \in \mathbb{R}$.

$$(u - tv,\, u - tv) \geq 0.$$

But

$$(u - tv,\, u - tv) = (u,\, u) - 2\,(u,\, v)\, t + (v,\, v)\, t^2.$$

Consider the quadratic function

$$f(t) = (v,\, v)\, t^2 - 2\,(u,\, v)\, t + (u,\, u)\,.$$

We have $f(t) \geq 0$.

But a quadratic function $f(t) = at^2 + bt + c$ satisfying $f(t) \geq 0$ has either two equal real roots or imaginary roots.
Hence

$$b^2 - 4ac = 4\,(u,\, v)^2 - 4\,(u,\, u)\,(v,\, v) \leq 0,$$

so

$$(u,\, v)^2 \leq (u,\, u)\,(v,\, v)\,,$$

and taking the square root of each side

$$|(u,\, v)| \leq ||u||\,||v||.$$

$\square$

Definition (1)

Two vectors u and v in V are said to be orthogonal if

$$(u,\, v) = 0.$$

Remark: Since $\cos^{-1} 0 = \frac{\pi}{2}$,

$$\begin{aligned} (u,\, v) \quad &\Longleftrightarrow \quad \measuredangle\,(u,\, v) = \frac{\pi}{2} \\ &\Longleftrightarrow \quad \text{they are perpendicular.} \end{aligned}$$

Definition

A basis $\mathscr{B} = \{u_1,\, \ldots,\, u_n\}$ for V is said to be orthonormal if

(1) $||u_j|| = 1,\ 1 \le j \le n.$

(2) $(u_i,\, u_j) = 0,\ i \neq j.$

Examples

(1) $\mathbb{R}^n$, $\mathscr{B} = \{e_1, \ldots, e_n\}$ so the standard basis for $\mathbb{R}^n$ is orthonormal.

(2) $L^2[0, 1]$

$$\mathscr{B} = \{1, \sin(nx), \cos(nx) : n \in \mathbb{Z}, n > 0\}$$

In the next lecture we will prove

Theorem

Every finite-dimensional vector space has an orthonormal basis (in fact, many).

The Triangle Inequalities

Before doing this we will prove the triangle inequalities. Given three vectors $u,\, v,\, w \in V$

$$\left.\begin{array}{rcl} \mathrm{d}\,(u,\, v) & \leq & \mathrm{d}\,(u,\, w) + \mathrm{d}\,(w,\, v) \\ \mathrm{d}\,(u,\, w) & \leq & \mathrm{d}\,(u,\, v) + \mathrm{d}\,(v,\, w) \\ \mathrm{d}\,(v,\, w) & \leq & \mathrm{d}\,(v,\, u) + \mathrm{d}\,(u,\, w) \end{array}\right\} (T1)$$

The point is that the length of any side of a triangle is less than the sums of the lengths of the other two sides.

From the definition of distance the triangle inequalities are equivalent to

$$\left.\begin{array}{rcl} ||u-v|| & \leq & ||u-w|| + ||w-v|| \\ ||u-w|| & \leq & ||u-v|| + ||v-w|| \\ ||v-w|| & \leq & ||v-u|| + ||u-w|| \end{array}\right\} (T2)$$

Put $a = v - w$, $b = v - u$, $c = u - w$.
Then $a = b + c$ and the triangle inequality is equivalent to proving

Theorem (Text, Theorem 15.6)

Suppose b, $c \in V$. *Then*

$$||b + c|| \leq ||b|| + ||c|| \quad (T3)$$

(Put $b = v - u$ and $c = u - w$ to get $(T2)$ and hence $(T1)$.)

Proof. Square both sides of $(T3)$ to get

$$||b+c||^2 \leq ||b||^2 + 2||b||||c|| + ||c||^2 \quad (\flat)$$

But

$$\begin{aligned} ||b+c||^2 &= (b+c,\, b+c) = (b,\, b) + 2\,(b,\, c) + (c,\, c) \\ &= ||b||^2 + 2\,(b,\, c) + ||c||^2. \end{aligned}$$

So $(\flat)$ is equivalent to

$$||b||^2 + 2\,(b,\, c) + ||c||^2 \leq ||b||^2 + 2||b||\,||c|| + ||c||^2.$$

But this inequality holds because

$$(b,\, c) \leq ||b||||c||$$

$\square$

Orthonormal Basis

Definition

A subset $\{u_1, \ldots, u_n\}$ of V is an orthonormal set if

$$(u_i, u_i) = 1 \text{ and } (u_i, u_j) = 0, \, i \neq j.$$

Lemma

Every orthonormal set in an independent set.

Proof. Suppose $\{u_1, \ldots, u_n\}$ is an orthonormal set and

$$\sum_{i=1}^{n} c_i u_i = 0 \qquad (*)$$

Orthonormal Basis

Take the dot product of each side of $(*)$ with u_j

$$\text{LHS} = \left(\sum_{i=1}^{n} c_i u_i,\, u_j \right) == \sum_{i=1}^{n} c_i \left(u_i,\, u_j \right).$$

But $(u_i,\, u_j) = 0$ unless $i = j$, so

$$\text{LHS} = c_j \left(u_j,\, u_j \right) = c_j.$$

(because $(u_j,\, u_j) = 1$).

$$\text{RHS} = (0,\, u_j) = 0.$$

Hence $c_j = 0$, all j and $\{u_1,\, \ldots,\, u_n\}$ is an independent set. $\square$

Next we prove the very useful formula for the coordinates of a vector v relative to an orthonormal basis $\mathscr{U} = \{u_1, \ldots, u_n\}$.

Proposition

Suppose $\mathscr{U} = \{u_1, \ldots, u_n\}$ *is an orthonormal basis. Let* $v \in V$*. Then the coordinates of* v *relative to [?] are*

$$((v, u_1), \ldots, (v, u_n)).$$

Proof. Let $(c_1, \ldots, c_n)$ be coordinates of v relative to $\mathscr{U}$. Hence

$$v = \sum_{i=1}^{n} c_i u_i \qquad (**)$$

Take the inner product of each side of $(**)$ with u_j. Then LHS$= (v, u_j)$ and as for the case of $(*)$ we get

$$\mathrm{RHS} = c_j$$

Hence

$$c_j = (v, u_j)\,.$$

$\square$

Finally, we will need a formula for the matrix $M(T)$ (or ${}_{\mathscr{U}}[T]_{\mathscr{U}}$) for the matrix of a linear transformation $T \in L(V, V)$ relative to a orthonormal basis $\mathscr{U} = \{u_1, \ldots, u_n\}$.

Proposition

$$M(T) = (a_{ij})$$

where $a_{ij} = (Te_j, e_i)$.

Proof. The entries a_{ij} of $M(T)$ are defined by the equation

$$T(u_j) = \sum_{k=1}^{n} a_{kj} u_i, \quad 1 \leq j \leq n.$$

Take the inner product of each side of this equaiton with u_i. We get

$$\begin{aligned} (T(u_j),\, u_j) &= \left(\sum_{i=1}^{n} a_{kj} u_k,\, u_i \right) = \sum_{i=1}^{n} a_{kj} \left(u_k,\, u_i \right) \\ &= a_{ij} \end{aligned}$$

since $(u_k,\, u_i) = 0$ unless $k = i$. $\square$