

Lecture 4: October 6, 2016

Lecturer: Madhur Tulsiani

1 Inner Products

For the discussion below, we will take the field \mathbb{F} to be \mathbb{R} or \mathbb{C} since the definition of inner products needs the notion of a “magnitude” for a field element (these can be defined more generally for subfields of \mathbb{R} and \mathbb{C} known as Euclidean subfields, but we shall not do so here).

Definition 1.1 Let V be a vector space over a field \mathbb{F} (which is taken to be \mathbb{R} or \mathbb{C}). A function $\mu : V \times V \rightarrow \mathbb{F}$ is an inner product if

- The function $\mu(\cdot, w) : V \rightarrow \mathbb{F}$ is a linear transformation for every $w \in V$.
- The function satisfies $\mu(u, v) = \overline{\mu(v, u)}$.
- $\mu(v, v) \in \mathbb{R}_{\geq 0}$ for all $v \in V$ and is 0 only for $v = 0_V$.

We write the inner product corresponding to μ as $\langle u, v \rangle$.

Strictly speaking, the inner product should be written as $\langle u, v \rangle_\mu$ but we usually omit the μ when the function is clear from context (or we are referring to an arbitrary inner product).

Example 1.2 The following are all examples of inner products:

- The function $\int_{-1}^1 f(x)g(x)dx$ for $f, g \in C([-1, 1], \mathbb{R})$ (space of continuous functions from $[-1, 1]$ to \mathbb{R}).
- The function $\int_{-1}^1 \frac{f(x)g(x)}{\sqrt{1-x^2}} dx$ for $f, g \in C([-1, 1], \mathbb{R})$.
- For $x, y \in \mathbb{R}^2$, $\langle x, y \rangle = x_1y_1 + x_2y_2$ is the usual inner product. Check that $\langle x, y \rangle = 2x_1y_1 + x_2y_2 + x_1y_2/2 + x_2y_1/2$ also defines an inner product.

Exercise 1.3 Let $C > 4$. Check that

$$\mu(f, g) = \sum_{n=0}^{\infty} \frac{f(n) \cdot g(n)}{C^n}$$

defines an inner product on the space Fib .

An inner product also defines a norm $\|v\| = \sqrt{\langle v, v \rangle}$ and hence a notion of distance between two vectors in a vector space. This can be used to define convergence of sequences, and to define infinite sums and limits of sequences (which was not possible in an abstract vector space). However, it might still happen that the limit of a sequence of vectors in the vector space, which converges according to the norm defined by the inner product, may not converge to a vector in the space. Consider the following example.

Example 1.4 Consider the vector space $C([-1, 1], \mathbb{R})$ with the inner product defined by $\langle f, g \rangle = \int_{-1}^1 f(x)g(x)dx$. Consider the sequence of functions:

$$f_n(x) = \begin{cases} -1 & x \in [-1, \frac{-1}{n}) \\ nx & x \in [\frac{-1}{n}, \frac{1}{n}) \\ 1 & x \in [\frac{1}{n}, 1] \end{cases}$$

One can check that $\|f_n - f_m\|^2 = O(\frac{1}{n})$ for $m \geq n$. Thus, the sequence converges. However, the limit point is a discontinuous function not in the inner product space. To fix this problem, one can essentially include the limit points of all the sequences in the space (known as the completion of the space). An inner product space in which all (Cauchy) sequences converge to a point in the space is known as a Hilbert space. Many of the theorems we will prove will generalize to Hilbert spaces though we will only prove some of them for finite dimensional spaces. The following is an extremely useful inequality. Prove it for a general inner product space (not necessarily finite dimensional).

Exercise 1.5 (Cauchy-Schwarz-Bunyakowsky inequality) Let u, v be any two vectors in an inner product space V . Then

$$|\langle u, v \rangle|^2 \leq \langle u, u \rangle \cdot \langle v, v \rangle$$

Definition 1.6 Two vectors u, v in an inner product space are said to be orthogonal if $\langle u, v \rangle = 0$. A set of vectors $\{w_1, \dots, w_n\}$ is said to be orthonormal if $\langle w_i, w_j \rangle = 0$ for all $i \neq j$ and $\|w_i\| = 1$ for all $i \in [n]$.

Proposition 1.7 A set $S \subseteq V$ consisting of mutually orthogonal vectors is linearly independent.

Proposition 1.8 (Gram-Schmidt orthogonalization) *Given a finite set $\{v_1, \dots, v_n\}$ of linearly independent vectors, there exists a set of orthonormal vectors $\{w_1, \dots, w_n\}$ such that*

$$\text{Span}(\{w_1, \dots, w_n\}) = \text{Span}(\{v_1, \dots, v_n\}) .$$

Proof: By induction. The case with one vector is trivial. Given the statement for k vectors and orthonormal $\{w_1, \dots, w_k\}$ such that

$$\text{Span}(\{w_1, \dots, w_k\}) = \text{Span}(\{v_1, \dots, v_k\}) ,$$

define

$$u_{k+1} = v_{k+1} - \sum_{i=1}^k \langle v_{k+1}, w_i \rangle \cdot w_i \quad \text{and} \quad w_{k+1} = \frac{u_{k+1}}{\|u_{k+1}\|} .$$

It is easy to check that the set $\{w_1, \dots, w_{k+1}\}$ satisfies the required conditions. ■

Corollary 1.9 *Every finite dimensional inner product space has an orthonormal basis.*

In fact, Hilbert spaces also have orthonormal bases (which are countable). The existence of a maximal orthonormal set of vectors can be proved by using Zorn's lemma, similar to the proof of existence of a Hamel basis for a vector space. However, we still need to prove that a maximal orthonormal set is a basis. This follows because we define the basis slightly differently for a Hilbert space: instead of allowing only finite linear combinations, we allow infinite ones. The correct way of saying this is that we still think of the span as the set of all *finite* linear combinations, then we only need that for any $v \in V$, we can get arbitrarily close to v using elements in the span (a converging sequence of finite sums can get arbitrarily close to its limit). Thus, we only need that the span is *dense* in the Hilbert space V . However, if the maximal orthonormal set is not dense, then it is possible to show that it cannot be maximal. Such a basis is known as a Hilbert basis.

Let V be a finite dimensional inner product space and let $\{w_1, \dots, w_n\}$ be an orthonormal basis for V . Then for any $v \in V$, there exist $c_1, \dots, c_n \in \mathbb{F}$ such that $v = \sum_i c_i \cdot w_i$. The coefficients c_i are often called Fourier coefficients. Using the orthonormality and the properties of the inner product, we get $c_i = \langle v, w_i \rangle$. This can be used to prove the following

Proposition 1.10 (Parseval's identity) *Let V be a finite dimensional inner product space and let $\{w_1, \dots, w_n\}$ be an orthonormal basis for V . Then, for any $u, v \in V$*

$$\langle u, v \rangle = \sum_{i=1}^n \langle u, w_i \rangle \cdot \langle v, w_i \rangle .$$

2 Adjoint of a linear transformation

Definition 2.1 Let V, W be inner product spaces over the same field \mathbb{F} and let $\varphi : V \rightarrow W$ be a linear transformation. A transformation $\varphi^* : W \rightarrow V$ is called an adjoint of φ if

$$\langle \varphi(v), w \rangle = \langle v, \varphi^*(w) \rangle \quad \forall v \in V, w \in W.$$

Example 2.2 Let $V = W = \mathbb{C}^n$ with the inner product $\langle u, v \rangle = \sum_{i=1}^n u_i \cdot \overline{v_i}$. Let $\varphi : V \rightarrow V$ be represented by the matrix A . Then φ^* is represented by the matrix A^T .

Exercise 2.3 Let $\varphi_{\text{left}} : \text{Fib} \rightarrow \text{Fib}$ be the left shift operator as before, and let $\langle f, g \rangle$ for $f, g \in \text{Fib}$ be defined as $\langle f, g \rangle = \sum_{n=0}^{\infty} \frac{f^{(n)}g^{(n)}}{C^n}$ for $C > 4$. Find φ_{left}^* .