Mathematical Toolkit	**Autumn 2023**

Lecture 4: October 5, 2023

Lecturer: Madhur Tulsiani

1 Inequalities for inner products and distances

We start with the following extremely useful inequality.

Proposition 1.1 (Cauchy-Schwarz-Bunyakovsky inequality) *Let u, v be any two vectors in an inner product space V. Then*

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

Proof: To prove for general inner product spaces (not necessarily finite dimensional) we will use the only inequality available in the definition i.e., $\langle w, w\rangle \geq 0$ for all $w \in V$. Taking $w = a \cdot u + b \cdot v$ and using the properties from the definition gives

$$\langle w, w\rangle \;=\; \langle (a \cdot u + b \cdot v), (a \cdot u + b \cdot v)\rangle \;=\; a\bar{a} \cdot \langle u, u\rangle + b\bar{b} \cdot \langle v, v\rangle + \bar{a}b \cdot \langle u, v\rangle + a\bar{b}\,\langle v, u\rangle$$

Taking $a = \langle v, v\rangle$ and $b = -\langle v, u\rangle = -\overline{\langle u, v\rangle}$ gives

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

If $v = 0_V$, then the inequality is trivial. Otherwise, we must have $\langle v, v\rangle > 0$. Thus,

$$\langle w, w\rangle \geq 0 \;\;\Rightarrow\;\; \langle u, u\rangle \cdot \langle v, v\rangle - |\langle u, v\rangle|^2 \;\geq\; 0\,,$$

which proves the desired inequality. ■

An inner product also defines a norm $\|v\| = \sqrt{\langle v, v\rangle}$ and a hence a notion of distance between two vectors in a vector space. This is a "distance" in the following sense.

Exercise 1.2 (Triangle inequality) *Prove that for any inner product space V, and any vectors $u, v, w \in V$*

$$\|u - w\| \;\leq\; \|u - v\| + \|v - w\| \;.$$

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 1.3 *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:*

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

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.

2 Orthogonality and orthonormality

Definition 2.1 *Two vectors u,v in an inner product space are said to be* orthogonal *if $\langle u,v\rangle = 0$. A set of vectors $S \subseteq V$ is said to consist of* mutually orthogonal vectors *if $\langle u,v\rangle = 0$ for all $u \neq v$, $u,v \in S$. A set of $S \subseteq V$ is said to be* orthonormal *if $\langle u,v\rangle = 0$ for all $u \neq v$, $u,v \in S$ and $\|u\| = 1$ for all $u \in S$.*

Proposition 2.2 *A set $S \subseteq V \setminus \{0_V\}$ consisting of mutually orthogonal vectors is linearly independent.*

Proof: Let $v_1,\ldots,v_n \in S$ and $c_1,\ldots,c_n \in \mathbb{F}$ be such that $\sum_{i\in[n]} c_i \cdot v_i = 0_V$. Taking inner product with a vector v_j for $j \in [n]$, we get that $\sum_i c_i \cdot \langle v_j, v_i\rangle = 0$. Since vectors in S are mutually orthogonal, we get that $\langle v_j, v_i\rangle = 0$ when $i \neq j$, which implies using the previous equality that that $c_j \langle v_j, v_j\rangle = 0$. Since $v_j \neq 0_V$, we must have $\langle v_j, v_j\rangle > 0$, and thus $c_j = 0$. Also, since our choice of j was arbitrary, this is true for all $j \in [n]$, implying $c_1 = \cdots = c_n = 0$. Thus, the only way a finite linear combination of vectors from S equals 0_V, if all coefficients are 0 , which implies that S is linearly independent. ■

Proposition 2.3 (Gram-Schmidt orthogonalization) *Given a finite set $\{v_1, \ldots, v_n\}$ of linearly independent vectors, there exists a set of orthonormal vectors $\{w_1, \ldots, w_n\}$ such that*

$$\mathrm{Span}\left(\{w_1, \ldots, w_n\}\right) \;=\; \mathrm{Span}\left(\{v_1, \ldots, v_n\}\right) .$$

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

$$\mathrm{Span}\left(\{w_1, \ldots, w_k\}\right) \;=\; \mathrm{Span}\left(\{v_1, \ldots, v_k\}\right) ,$$

define

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

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

Corollary 2.4 *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 is 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.

2.1 Fourier coefficients

Let V be a finite dimensional inner product space and let $B = \{w_1, \ldots, w_n\}$ be an orthonormal basis for V. Then for any $v \in V$, there exist $c_1, \ldots, c_n \in \mathbb{F}$ such that $v = \sum_i c_i \cdot w_i$. The coefficients c_i are often called Fourier coefficients (of v, with respect to the basis B). Using the orthonormality and the properties of the inner product, we get

Proposition 2.5 *Let $B = \{w_1, \ldots, w_n\}$ be an orthonormal basis for V, and let $v \in V$ be expressible as $v = \sum_{i=1}^{n} c_i \cdot w_i$. Then, for all $i \in [n]$, we must have $c_i = \langle w_i, v \rangle$.*

This can be used to prove the following

Proposition 2.6 (Parseval's identity) *Let V be a finite dimensional inner product space and let $\{w_1, \ldots, w_n\}$ be an orthonormal basis for V. Then, for any $u, v \in V$*

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

Exercise 2.7 *Prove that the set of functions*

$$S \;=\; \{1/2\} \cup \{\sin(k\pi x) \mid k \in \mathbb{N}, k \geq 1\} \cup \{\cos(k\pi x) \mid k \in \mathbb{N}, k \geq 1\} \;,$$

is an orthonormal set in the Hilbert space of continuous real-valued functions mapping $[-1,1]$ to $\mathbb{R}$ (denoted $C([-1,1],\mathbb{R})$) under the inner product $\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx$.

In fact, the above functions form an orthonormal (Hilbert) basis for the space $C([-1,1],\mathbb{R})$, and are often referred to as the Fourier basis functions in signal analysis.

Parseval's identity can also be used to show that the size of the set large Fourier coefficients is small, and this is often very useful in "truncating" a signal to its large Fourier coefficients.

Exercise 2.8 *Let $B = \{w_1, \ldots, w_n\}$ be an orthonormal basis for a Hilbert space V. Let $v \in V$ with $\|v\| = 1$ be expressed as $v = \sum_i c_i \cdot w_i$. Then show that for any $\delta > 0$*

$$|\{i \;\mid\; |c_i| \geq \delta\}| \;\leq\; \frac{1}{\delta^2} \,.$$