

Lecture 4: April 8, 2021

Lecturer: Avrim Blum (notes based on notes from Madhur Tulsiani)

1 Orthogonality and orthonormality

Definition 1.1 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 1.2 A set $S \subseteq V \setminus \{0_V\}$ consisting of mutually orthogonal vectors is linearly independent.

Proposition 1.3 (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

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

We can now check that the set $\{w_1, \dots, w_{k+1}\}$ satisfies the required conditions. Unit length is clear, so let's check orthogonality:

$$\langle u_{k+1}, w_j \rangle = \langle v_{k+1}, w_j \rangle - \sum_{i=1}^k \overline{\langle w_i, v_{k+1} \rangle} \cdot \langle w_i, w_j \rangle = \langle v_{k+1}, w_j \rangle - \overline{\langle w_j, v_{k+1} \rangle} = 0.$$

■

Corollary 1.4 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. 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 $c_i = \langle w_i, v \rangle$. This can be used to prove the following

Proposition 1.5 (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 $u, v \in V$*

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

Proof: Just plug in $v = \sum_i \langle w_i, v \rangle w_i$ in the left-hand side and distribute out the inner product. ■

Let's consider \mathbb{R}^n . If the w_i are the "standard basis", then this is just writing the inner product $\langle u, v \rangle$ in the usual way as the sum of the products of the coordinate values $\sum_j u_j v_j$ where $u_j = \langle u, w_j \rangle$ and $v_j = \langle v, w_j \rangle$. Parseval's identity says you can do this using any orthonormal basis you want. Plugging in the case of $v = u$, we get $\|u\|^2 = \sum_i u_i^2$.

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*

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

Example 2.2 *Let $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ with the usual inner product, and let $\varphi : V \rightarrow W$ be represented by the matrix A . Then φ^* is represented by the matrix A^T . In particular, $\langle w, Av \rangle = w^T Av = (A^T w)^T v = \langle A^T w, v \rangle = \langle \varphi^*(w), v \rangle$. So, a symmetric matrix is "self-adjoint".*

Example 2.3 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 .

Example 2.4 Let $V = C([0, 1], [-1, 1])$ with the inner product $\langle f_1, f_2 \rangle = \int_0^1 f_1(x)f_2(x)dx$, and let $W = C([0, 1/2], [-1, 1])$ with the inner product $\langle g_1, g_2 \rangle = \int_0^{1/2} g_1(x)g_2(x)dx$. Let $\varphi : V \rightarrow W$ be defined as $\varphi(f)(x) = f(2x)$. Then, $\varphi^* : W \rightarrow V$ can be defined as

$$\varphi^*(g)(y) = (1/2) \cdot g(y/2).$$

We will prove that every linear transformation has a unique adjoint. However, we first need the following characterization of linear transformations from an inner product space V to the field \mathbb{F} it is over.

Proposition 2.5 (Riesz Representation Theorem) Let V be a finite-dimensional inner product space over \mathbb{F} and let $\alpha : V \rightarrow \mathbb{F}$ be a linear transformation. Then there exists a unique $z \in V$ such that $\alpha(v) = \langle z, v \rangle \forall v \in V$.

We only prove the theorem here for finite-dimensional spaces. However, the theorem holds for any Hilbert space.

Proof: Let $\{w_1, \dots, w_n\}$ be an orthonormal basis for V . Given v , let c_1, \dots, c_n be its Fourier coefficients, so $v = \sum_i c_i w_i$, and $c_i = \langle w_i, v \rangle$. Since α is a linear transformation, we must have $\alpha(v) = \sum_i c_i \alpha(w_i) = \sum_i \langle w_i, v \rangle \alpha(w_i) = \sum_i \langle \overline{\alpha(w_i)} w_i, v \rangle = \langle z, v \rangle$ for $z = \sum_i \overline{\alpha(w_i)} w_i$. ■

This can be used to prove the following:

Proposition 2.6 Let V, W be finite dimensional inner product spaces and let $\varphi : V \rightarrow W$ be a linear transformation. Then there exists a unique $\varphi^* : W \rightarrow V$, such that

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

Proof: For each $w \in W$, the map $\langle w, \varphi(\cdot) \rangle : V \rightarrow \mathbb{F}$ is a linear transformation (check!) and hence there exists a unique $z_w \in V$ satisfying $\langle w, \varphi(v) \rangle = \langle z_w, v \rangle \forall v \in V$. Consider the map $\beta : W \rightarrow V$ defined as $\beta(w) = z_w$. By definition of β ,

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

To check that β is linear, we note that $\forall v \in V, \forall w_1, w_2 \in W$,

$$\langle \beta(w_1 + w_2), v \rangle = \langle w_1 + w_2, \varphi(v) \rangle = \langle w_1, \varphi(v) \rangle + \langle w_2, \varphi(v) \rangle = \langle \beta(w_1), v \rangle + \langle \beta(w_2), v \rangle,$$

which implies $\beta(w_1 + w_2) = \beta(w_1) + \beta(w_2)$. $\beta(c \cdot w) = c \cdot \beta(w)$ follows similarly. ■

Note that the above proof only requires the Riesz representation theorem (to define z_w) and hence also works for Hilbert spaces.

3 Self-adjoint transformations

Definition 3.1 A linear transformation $\varphi : V \rightarrow V$ is called self-adjoint if $\varphi = \varphi^*$. Linear transformations from a vector space to itself are called linear operators.

Example 3.2 The transformation represented by matrix $A \in \mathbb{C}^{n \times n}$ is self-adjoint if $A = \overline{A^T}$. Such matrices are called Hermitian matrices.

Proposition 3.3 Let V be an inner product space and let $\varphi : V \rightarrow V$ be a self-adjoint linear operator. Then

- All eigenvalues of φ are real.
- If $\{w_1, \dots, w_n\}$ are eigenvectors corresponding to distinct eigenvalues then they are mutually orthogonal.