TTIC 31150/CMSC 31150 Mathematical Toolkit (Fall 2024)

Avrim Blum

Lecture 4: Orthogonality and Adjoints

Recap

- Linear transformations, bases, connections of LTs to matrices, kernel (nullspace) and image, rank-nullity theorem.
- Eigenvectors and eigenvalues
- Eigenvectors of same eigenvalue form a subspace. Eigenvectors of different eigenvalues are linearly independent.
- Inner products, norm $||v|| = \sqrt{\langle v, v \rangle}$
- Cauchy-Schwartz: $|\langle u, v \rangle| \leq ||u||\ ||v||$.
- Triangle inequality of norm.

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.*

Proof:

- If u satisfies $\langle u, v \rangle = 0$ for all $v \in S \setminus \{u\}$, then it also has inner product 0 with any linear combination of $S \setminus \{u\}$, so it can't be in the span unless already the 0 vector.
- Since this holds for all $u \in S$, this means S must be linearly independent.

Proposition 1.3 (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*

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

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

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

define

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

- Unit-length is clear. Let's check orthogonality:

Using inductive assumption that $w_1, \ldots, w_k$ are orthonormal

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

Corollary 1.4 *Every finite dimensional inner product space has an orthonormal basis.*

Brief note on Hilbert spaces:

- Hilbert spaces also have a (countably infinite) orthonormal basis.
- Need to define basis a bit differently: span of a set of vectors is still the set of all finite linear combinations, but we only require that for any $v \in V$, we can get arbitrarily close to v using elements in the span.
- We will focus on finite-dimensional vector spaces.

Fourier Coefficients

Let V be a finite-dimensional inner-product space with orthonormal basis $\{w_1, \dots, w_n\}$.

- So, for any $v \in V$, there exist $c_1, \dots, c_n$ such that $v = \sum_{i=1}^{n} c_i w_i$.
- These c_i are called *Fourier Coefficients*.
- Note that $c_i = \langle w_i, v \rangle$. Why?
  - Let's compute $\langle w_i, v \rangle = \left\langle w_i, \sum c_j w_j \right\rangle = \sum_j \left\langle w_i, c_j w_j \right\rangle = \sum_j c_j \left\langle w_i, w_j \right\rangle = c_i$.
- So, $v = \sum_{i=1}^{n} \langle w_i, v \rangle w_i$.

This then gives us…

Parseval's identity

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

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

Proof:

- We know $v = \sum_i \langle w_i, v \rangle w_i$. Plug into LHS and distribute.

If working over $\mathbb{R}^n$, and w_i are the standard basis, with $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ then this says that $\langle u, v \rangle = \sum_i u_i v_i$.

Adjoint of a Linear Transform

Note that φ^* goes in reverse direction from φ.

Definition 2.1 *Let V, W be inner product spaces over the same field $\mathbb{F}$ and let $\varphi : V \to W$ be a linear transformation. A transformation $\varphi^* : W \to V$ 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 $\varphi : V \to W$ be represented by the matrix A. Then φ^* is represented by the matrix A^T. In particular, $\langle w, Av \rangle = w^T A v = (A^T w)^T v = \langle A^T w, v \rangle = \langle \varphi^*(w), v \rangle$. So, a symmetric matrix is "self-adjoint".*

Adjoint of a Linear Transform

Note that φ^* goes in reverse direction from φ.

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

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

Continuous functions from [0,1] to [-1,1]

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 $\varphi : V \to W$ be defined as $\varphi(f)(x) = f(2x)$. Then, $\varphi^* : W \to V$ can be defined as*

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

Let's calculate: $\langle g, \varphi(f) \rangle = \int_0^{1/2} g(x) f(2x) dx = \int_0^1 \frac{1}{2} g\left(\frac{y}{2}\right) f(y) dy$, using $y = 2x, dy = 2dx$.

Characterization of linear transformations from V to $\mathbb{F}$

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

In other words, the only linear transformations from V to $\mathbb{F}$ are those given by $\langle z, \cdot \rangle$ for some z.

Proof: Let $\{w_1, \ldots, w_n\}$ be an orthonormal basis for V. Given v, let $c_1, \ldots, c_n$ 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 $\alpha(v) = \sum_i c_i \alpha(w_i) = \sum_i \langle w_i, v \rangle \alpha(w_i) = \sum_i \left\langle \overline{\alpha(w_i)} w_i, v \right\rangle = \langle z, v \rangle$ for $z = \sum_i \overline{\alpha(w_i)} w_i$. ■

Scalar, so can move into 1st slot by taking conjugate

Using this, can show that any linear transformation has a unique adjoint.

Every linear transformation has a unique adjoint

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

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

Proof:

- For each w, the mapping $\psi_w(v) = \langle w, \varphi(v)\rangle$ is a linear transformation from V to $\mathbb{F}$.
- So, exists unique $z_w \in V$ s.t. $\psi_w(v) = \langle z_w, v\rangle$.
- Now, consider $\beta: W \to V$ defined as $\beta(w) = z_w$. So, we have $\langle w, \varphi(v)\rangle = \langle \beta(w), v\rangle$.
- Verify β is linear. In particular, for all w_1, w_2 have $\langle \beta(w_1 + w_2), v\rangle = \langle w_1 + w_2, \varphi(v)\rangle = \langle \beta(w_1) + \beta(w_2), v\rangle$ for all v, which implies $\beta(w_1 + w_2) = \beta(w_1) + \beta(w_2)$. Similar reasoning for $\beta(cw) = c\beta(w)$.

Self-adjoint Transformations

Definition 3.1 *A linear transformation $\varphi : V \to V$ 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.*

So, over the reals, square symmetric matrices are self-adjoint.

Self-adjoint Transformations

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

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

Proof (first part):

- Let v be an eigenvector and λ its associated eigenvalue.
- We know that $\langle\varphi(v), v\rangle = \langle v, \varphi(v)\rangle$, so $\langle\lambda v, v\rangle = \langle v, \lambda v\rangle$, so $\bar{\lambda} = \lambda$.

Self-adjoint Transformations

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

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

Proof (second part):

- Say w_1 has eigenvalue λ_1 and w_2 has eigenvalue λ_2, where $\lambda_1 \neq \lambda_2$.
- We know $\langle \varphi(w_1), w_2 \rangle = \langle w_1, \varphi(w_2) \rangle$, so $\langle \lambda_1 w_1, w_2 \rangle = \langle w_1, \lambda_2 w_2 \rangle$.
- This means $\lambda_1 \langle w_1, w_2 \rangle = \lambda_2 \langle w_1, w_2 \rangle$. So, must have $\langle w_1, w_2 \rangle = 0$.