Lecture 3

Adjoint and self-adjoint operators and matrices

An inner product structure on a $\mathbb{C}$-vector spaces induces a "mirrored" twin for every linear transformation, called the adjoint. Linear operators equal their own adjoints have many important properties.

3.1 Adjoint of an operator or matrix

Definition 15 (Adjoint of a linear map)**.** *Let $f : U \to V$ be a linear map between two inner product spaces. The **adjoint** of f, denoted by $f^* : V \to U$, is the unique linear map such that $\langle f(u), v\rangle = \langle u, f^*(v)\rangle$ for all $u \in U$ and $v \in V$.*

Theorem 13 (Technical facts about adjoints)**.** *Let f and g be two linear operators on V.*

1. $(\alpha f + g)^* = \overline{\alpha} f^* + g^*$ *(conjugate linear)*
2. $(fg)^* = g^* f^*$ *(reverses composition)*
3. $(f^*)^*$ *(involutive)*
4. $I^* = I$ *(identity operator is its own adjoint)*

Proof.

1. Let $x, y \in V$. We need to show that $\langle (\alpha f + g)\, x, y\rangle = \langle x, (\overline{\alpha} f^* + g^*)\, y\rangle$.

$$\begin{aligned}
\langle (\alpha f + g)\, x, y\rangle &= \alpha \langle f x, y\rangle + \langle g x, y\rangle && (3.1)\\
&= \alpha \langle x, f^* y\rangle + \langle x, g^* y\rangle && (3.2)\\
&= \langle x, (\overline{\alpha} f^* + g^*)\, y\rangle && (3.3)\\
&= \langle x, \overline{\alpha} f^*\rangle + \langle x, g^* y\rangle && (3.4)\\
& && (3.5)
\end{aligned}$$

2. Same setup as before. $\langle fgx, y\rangle = \langle gx, f^* y\rangle = \langle x, g^* f^* y\rangle$.
3. Same setup as before. $\langle f^* x, y\rangle = \overline{\langle y, f^* x\rangle} = \overline{\langle f y, x\rangle} = \langle x, f y\rangle$.
4. Same setup as before. $\langle I x, y\rangle = \langle x, I y\rangle$.

□

Definition 16 (Adjoint of a matrix)**.** *Let $A \in \mathbb{C}^{m\times n}$. The **adjoint** or **conjugate transpose** of A is the matrix A^* such that $A^*_{ij} = \overline{A_{ji}}$.*

Theorem 14. *Over an inner product space of finite dimension, adjoints exist. In $\mathbb{C}^n$ with the standard inner product, an adjoint matrix is the matrix of the adjoint of the linear operator it represents.*

3.2 Self-adjoint operators

In this section we'll prove a condition for matrices to be orthogonally (-normally) diagonalizable.

Definition 17 (Self-adjoint)**.** *A linear operator or matrix is called **self-adjoint** or **Hermitian** if it is equal to its own adjoint.*

Lemma 1 (Real eigenvalues)**.** *Let $f : V \to V$ be a Hermitian operator and λ an eigenvalue of f. Then λ is real.*

Proof. Let v be a unit eigenvector of f satisfying $f(v) = \lambda v$. Then

$$\lambda = \langle fv, v \rangle \tag{3.6}$$
$$= \langle v, fv \rangle \tag{3.7}$$
$$= \overline{\langle fv, v \rangle} \tag{3.8}$$
$$= \overline{\lambda}. \tag{3.9}$$

□

A large class of complex matrices with complex eigenvectors, but real eigenvalues. Wow! Wow!

Definition 18 (Restriction)**.** *Let W be a subspace of V and f a linear operator on V. If $f(W) \subseteq W$, then W is called f**-invariant** or f**-stable**. The linear operator $f\big|_W : W \to W$ defined by $f\big|_W(w) = f(w)$ is called the **restriction** of f to W.*

Theorem 15 (Spectral[1] theorem)**.** *A linear operator $f : V \to V$ is self-adjoint if and only if it is diagonal and real in an orthonormal basis of V.*

Proof. First we show that (diagonal and real in an orthonormal basis) $\implies$ (self-adjoint). Suppose that the matrix of f is diagonal and real in an orthonormal basis. A diagonal real matrix is equal to its conjugate transpose. Therefore, as orthonormal bases faithfully represent inner product spaces and maps between them, f is self-adjoint.

Next we show that (self-adjoint) $\implies$ (diagonal and real in an orthonormal basis). We will use induction on n, the dimension of V. If $n = 1$ then f is already diagonal in any basis.

Next we need to show that if the Spectral Theorem holds on vector spaces of dimension $n - 1$, then it holds on vector spaces of dimension n.

By the Fundamental Theorem of Algebra, f has an eigenvalue λ. Because f is self-adjoint, λ is real. Let v be an eigenvector such that $f(v) = \lambda v$. Both $\operatorname{Span} v$ and its orthogonal complement $(\operatorname{Span} v)^\perp$ are stable under f, the former because it is an eigenspace, the latter in this way: let $\langle v', v \rangle = 0$. Then $\langle f(v'), v \rangle = \langle v', f(v) \rangle = \overline{\lambda} \langle v', v \rangle = 0$.

By the induction hypothesis, $f\big|_{(\operatorname{Span} v)^\perp}$ has $n - 1$ orthogonal eigenvectors in W. They are still eigenvectors, and still orthogonal, when treated as members of V. Furthermore, they are orthogonal to v by construction.

As such f has a basis of orthogonal eigenvectors, so it has a basis of orthonormal eigenvectors as well. □

Lemma 2 (Spectral theorem, factorization version)**.** *Let $A \in \mathbb{C}^{n \times n}$ satisfy $A = A^*$. Then there exist a unitary matrix U and a real diagonal matrix Λ such that $A = U \Lambda U^*$.*

Notice that $U^* = U^{-1}$, so no manual inversion necessary.

[1]The word *spectral* imputes an aura of magic and mystery. That is fitting because this theorem is very, very powerful.

Lemma 3 (Spectral theorem, projection version)**.** *Let $f : V \to V$ satisfy $f = f^*$. Let $\{v_1, v_2, \dots, v_n\}$ be orthonormal eigenvectors of f with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$.*

Then $f = \sum_{i=1}^{n} \lambda_i \operatorname{proj}_{v_i}$. That is, every self-adjoint operator is a real *linear combination of orthogonal projections.*

Lemma 4 (Spectral theorem, dyad version)**.** *The last and final form of the Spectral Theorem can be seen either as an expansion of the factorization version into outer products or as a translation of the projection version into orthonormal coordinates. Let $A \in \mathbb{C}^n$ satisfy $A = A^*$. Let $\{v_1, v_2, \dots, v_n\}$ be orthonormal eigenvectors of f with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$.*

Then $f = \sum_{i=1}^{n} \lambda_i v_i v_i^*$.

Application: direction of maximum amplification

Given a nonzero matrix $A \in \mathbb{C}^{n \times n}$, we might wonder what spatial direction gives you the best bang for your buck under left multiplication by A. That is, we are interested in the maximum amplification that A can exert on any vector:

$$\max_{v} \frac{\|Av\|}{\|v\|} \tag{3.10}$$

We can narrow our search to unit vectors.

$$\max_{\|v\|=1} \sqrt{\langle Av, Av\rangle} = \max_{\|v\|=1} \sqrt{\langle A^*Av, v\rangle} \tag{3.11}$$

A^*A is self-adjoint: $(A^*A)^* = (A^*)(A^*)^* = A^*A$. Diagonalize it as $A^*A = U\Lambda U^*$.

$$= \max_{\|v\|=1} \sqrt{\langle U\Lambda U^* v, v\rangle} \tag{3.12}$$

$$= \max_{\|v\|=1} \sqrt{\langle \Lambda U^* v, U^* v\rangle} \tag{3.13}$$

Change variables to $w = U^*v$.

$$= \max_{\|w\|=1} \sqrt{\langle \Lambda w, w\rangle} \tag{3.14}$$

This maximum is achieved when $w = e_i$, where λ_i is a maximal entry of Λ.

$$= \sqrt{\lambda_{\max}(A^*A)} \tag{3.15}$$

The furthest that A can magnify any vector is the square root of $\lambda_{\max}$, a maximal eigenvalue of A^*A. It quantifies how "big" A is, and is sometimes called are σ_1 or $\|A\|_2$, the operator norm of A.