

C Self-adjoint operators and complete orthonormal bases

The concept of the *adjoint* of an operator plays a very important role in many aspects of linear algebra and functional analysis. Here we will primarily focus on self-adjoint operators and how they can be used to obtain and characterize complete orthonormal sets of vectors in a Hilbert space. The main fact we will present here is that self-adjoint operators typically have orthogonal eigenvectors, and under some conditions, these eigenvectors can form a complete orthonormal basis for the Hilbert space. In particular, we will use this technique to find a variety of orthonormal bases for $L^2[a, b]$ that satisfy different types of boundary conditions.

We begin with the algebraic aspects of self-adjoint operators and their eigenvectors. Those algebraic properties are identical in the finite and infinite dimensional cases. The completeness arguments in infinite dimensions are more delicate, we will only state those results without complete proofs. We begin first by defining the adjoint.

Definition C.1. (*Finite dimensional or bounded operator case*) Let $A : H_1 \rightarrow H_2$ be a bounded operator between the Hilbert spaces H_1 and H_2 . The adjoint $A^* : H_2 \rightarrow H_1$ of A is defined by the requirement

$$\langle v_2, Av_1 \rangle_2 = \langle A^*v_2, v_1 \rangle_1, \quad (\text{C.1})$$

for all $v_1 \in H_1$ and $v_2 \in H_2$.

Note that for emphasis, we have written $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ for the inner products in H_1 and H_2 respectively. It remains to be shown that the relation (C.1) defines an operator A^* from A uniquely. We will not do this here in general but will illustrate it with several examples.

Examples

Example C.2. In the finite dimensional case, i.e. when A is a matrix, A^* is simply the complex conjugate transpose of A as we now demonstrate. Let A be an $n \times m$ matrix, it can be viewed as a linear mapping $A : \mathbb{C}^m \rightarrow \mathbb{C}^n$. Let's temporarily denote the complex conjugate transpose of a matrix or a vector by † . Thus the inner product in \mathbb{C}^n is given by $\langle v, w \rangle = v^\dagger w$. Now, equation (C.1) becomes

$$v_1^\dagger (Av_2) = (A^*v_1)^\dagger v_2 = v_1^\dagger (A^*)^\dagger v_2,$$

for all vectors v_1 and v_2 . This in turn implies that $A = (A^*)^\dagger$ or equivalently that

$$A^* = A^\dagger,$$

i.e. the adjoint of a matrix is the complex conjugate transpose of that matrix.

Example C.3. Let $T : L^2[a, b] \rightarrow L^2[a, b]$ be the bounded operator defined by the continuous kernel function $T(\cdot, \cdot)$. Specifically, $T : f \mapsto g$ is given by

$$g(x) = \int_a^b T(x, \chi) f(\chi) d\chi.$$

We compute the kernel representation of T^* which we will denote by $T^*(\cdot, \cdot)$. Translating the requirement (C.1) using the inner product of $L^2[a, b]$ gives

$$\begin{aligned} \langle h, Tf \rangle &= \langle T^*h, f \rangle \\ &\Downarrow \\ \int_a^b h^\dagger(x) (Tf)(x) dx &= \int_a^b ((T^*h)(r))^\dagger f(r) dr, \end{aligned}$$

where again $(\cdot)^\dagger$ denotes complex conjugate transpose. Substituting the kernel representations for T and T^* , we obtain

$$\begin{aligned} \int_a^b h^\dagger(x) \left(\int_a^b T(x, \chi) f(\chi) d\chi \right) dx &= \int_a^b \left(\int_a^b T^*(r, \rho) h(\rho) d\rho \right)^\dagger f(r) dr \\ &\Downarrow \\ \int_a^b \int_a^b h^\dagger(x) T(x, \chi) f(\chi) d\chi dx &= \int_a^b \int_a^b h^\dagger(\rho) (T^*(r, \rho))^\dagger f(r) dr d\rho, \end{aligned}$$

which can then be rewritten as

$$\int_a^b \int_a^b h^\dagger(x) \left(T(x, \chi) - (T^*(\chi, x))^\dagger \right) f(\chi) d\chi dx.$$

This last equation has to hold for all functions h and f in $L^2[a, b]$. By using test functions (see problem C.15) for f and h , it is not difficult to verify that the kernels satisfy $T(x, \chi) = (T^*(\chi, x))^\dagger$. Taking transposes of both sides and relabeling the arguments x and χ we can then write

$$T^*(x, \chi) = T^\dagger(\chi, x).$$

Note the interesting pattern here; the kernel $T^*(\cdot, \cdot)$ is given by “transposing” the arguments of the kernel $T(\cdot, \cdot)$ and taking the complex conjugate transpose at each point. This result is consistent with the interpretation of kernel functions as continuum analogues of matrices (see Appendix B).

For the proper definition of the adjoint in the unbounded operator case, domain considerations are important.

Definition C.4. (*Unbounded operator case*) Let $A : H_1 \rightarrow H_2$ be an operator between the Hilbert spaces H_1 and H_2 with domain $\mathcal{D}(A) \subset H_1$. The adjoint $A^* : H_2 \rightarrow H_1$ of A is defined by the requirement

$$\langle v_2, Av_1 \rangle_2 = \langle A^*v_2, v_1 \rangle_1, \tag{C.2}$$

for all $v_1 \in \mathcal{D}(A)$ and v_2 for which it is finite. The set of such vectors v_2 is the domain $\mathcal{D}(A^*) \subset H_2$ of A^*

Examples

Example C.5. Consider the following differential operator on $L^2[a, b]$

$$D := \frac{d}{dx}, \quad \mathcal{D}(D) := \left\{ f \in L^2; \frac{df}{dx} \in L^2, f(a) = 0 \right\}.$$

For $f \in \mathcal{D}(D)$ and $g \in L^2$, we evaluate $\langle g, Df \rangle$ and use integration by parts to obtain

$$\langle g, Df \rangle = \int_a^b g^\dagger(x) \frac{df}{dx}(x) dx \tag{C.3}$$

$$= g(x) f(x) \Big|_a^b - \int_a^b \frac{dg^\dagger}{dx}(x) f(x) dx \tag{C.4}$$

$$= g(b) f(b) - \left\langle \frac{d}{dx} g, f \right\rangle. \tag{C.5}$$

Thus it is clear that for (C.2) to hold, D^* must be given by

$$D^* = -\frac{d}{dx}, \quad \mathcal{D}(D^*) := \left\{ f \in L^2; \frac{df}{dx} \in L^2, f(b) = 0 \right\}. \quad \checkmark$$

Note the difference between the domains of D and D^* .

Example C.6. Consider Δ_D , the second order derivative operator with homogeneous Dirichlet boundary conditions defined on $L^2[a, b]$ by

$$\Delta_D := \frac{d^2}{dx^2}, \quad \mathcal{D}(\Delta_D) := \left\{ f \in L^2; \frac{d^2 f}{dx^2} \in L^2, f(a) = f(b) = 0 \right\}.$$

For $f \in \mathcal{D}(\Delta_D)$ and $g \in L^2$, we evaluate¹

¹ Here we use the notation f' and f'' for the first and second derivatives of a function f respectively

$$\langle g, \Delta_D f \rangle = \int_a^b g^\dagger(x) f''(x) dx \quad (\text{C.6})$$

$$= g(x)f'(x)|_a^b - \int_a^b g^\dagger(x)f'(x) dx \quad (\text{C.7})$$

$$= g(x)f'(x)|_a^b - g'(x)f(x)|_a^b + \int_a^b g''^\dagger(x)f(x) dx \quad (\text{C.8})$$

$$= g(b)f'(b)|_a^b + \left\langle \frac{d^2}{dx^2} g, f \right\rangle. \quad (\text{C.9})$$

Thus for (C.2) to hold, Δ_D must be defined as $\frac{d^2}{dx^2}$ on the domain of twice differentiable functions $g \in L^2$ with $g(a) = g(b) = 0$. This is precisely the same definition and domain as Δ_D , and we therefore conclude

$$\Delta_D^* = \Delta_D.$$

We have concluded that Δ_D is equal to its adjoint. Such operators are called self-adjoint.

Self-adjoint operators

Definition C.7. Let A be an operator from a dense domain $\mathcal{D}(A) \subseteq H$ to the same Hilbert space H . A is called self-adjoint if $\mathcal{D}(A^*) = \mathcal{D}(A)$ and $A^* = A$.

In the finite dimensional case, a self adjoint matrix is a matrix that is equal to its complex conjugate transpose, i.e. a Hermitian matrix. Sometimes self-adjoint operators are also called Hermitian operators. There are two very important properties of self-adjoint operators that are easy to demonstrate.

Lemma C.8. Let A be a self-adjoint operator on a Hilbert space H , then

- All the eigenvalues of A are real.
- If v_1 and v_2 are eigenvectors of A corresponding to distinct eigenvalues $\lambda_1 \neq \lambda_2$, then v_1 and v_2 are orthogonal.

Proof. Let \dagger denote complex conjugation. If λ is an eigenvalue of A , then \exists a nonzero eigenvector v with $Av = \lambda v$. Starting from the middle equality in the following equations (which follows from $A = A^*$), we compute

$$\lambda^\dagger \|v\|^2 = \langle \lambda v, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \langle v, \lambda v \rangle = \lambda \|v\|^2.$$

Since $\|v\| \neq 0$, we conclude that $\lambda^\dagger = \lambda$, i.e. λ is real.

To prove the second statement, we again start from the middle equality (which again follows from $A = A^*$) in

$$\lambda_1 \langle v_1, v_2 \rangle = \langle Av_1, v_2 \rangle = \langle v_1, Av_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle,$$

which implies that $\langle v_1, v_2 \rangle = 0$ since $\lambda_1 \neq \lambda_2$.

In the case of repeated eigenvalues, one can still make statements similar to the ones above. If A has two linearly independent eigenvectors w, v with the same eigenvalue λ , then any vector in the two dimensional subspace spanned by $\{w, v\}$ is an eigenvector with eigenvalue λ . In this case, we can pick two orthonormal vectors in that space that span it. This procedure can be applied to eigenvalues of any multiplicity. The result is that we can choose a complete orthonormal basis from amongst the eigenvectors of a self-adjoint adjoint operator. The precise statement follows, though its complete proof is not given here.

Theorem C.9. *Let A be a self adjoint operator defined on a dense domain in a Hilbert space H . If A has a discrete spectrum, then there exists an orthonormal set of eigenvectors of A that span all of H .*

This result is normally used to construct bases for function spaces like $L^2(\mathbb{X})$ that satisfy certain boundary conditions. To do this, one must be able to construct a self-adjoint operator whose domain is specified by those boundary conditions. This is only possible for certain kinds of boundary conditions. We illustrate this technique with two examples.

Example C.10. Consider the operator Δ_D described in example C.6. This operator is self-adjoint as demonstrated in that example. We now determine its spectrum. If it turns out to be discrete, then we can use the eigenfunctions as a complete orthogonal basis of $L^2[a, b]$ which have the property of being zero at the boundary points a and b . For computational simplicity, let's consider the problem over $L^2[-\pi, \pi]$ though the procedure is applicable to any finite interval.

A non-zero function $\phi \in L^2[-\pi, \pi]$ is an eigenfunction of Δ_D if $\phi(\pm\pi) = 0$ (i.e. it is in $\mathcal{D}(\Delta_D)$) and $\Delta_D\phi = \lambda\phi$ or equivalently

$$\frac{d^2}{dx^2}\phi(x) = \lambda\phi(x), \quad \phi(\pm\pi) = 0.$$

This is a second order ODE with constant coefficients. It may not have solutions that satisfy the boundary conditions for all λ . We first characterize the numbers λ for which there exists solutions, and then construct those solutions. The characteristic roots are given as the roots of the polynomial (in α) $(\alpha^2 - \lambda) = 0$. The form of the solutions depends on whether λ is positive or negative, they are

$$\begin{aligned} \phi(x) &= a e^{\sqrt{\lambda} x} + b e^{-\sqrt{\lambda} x} && \text{if } \lambda > 0, \\ \phi(x) &= a \cos(\sqrt{|\lambda|} x) + b \sin(\sqrt{|\lambda|} x) && \text{if } \lambda < 0. \end{aligned}$$

We first note that the case $\lambda > 0$ can not satisfy the given boundary conditions since in that case

$$\begin{aligned} 0 &= \phi(\pi) = a e^{\pi\sqrt{\lambda}} + b e^{-\pi\sqrt{\lambda}} \\ 0 &= \phi(-\pi) = a e^{-\pi\sqrt{\lambda}} + b e^{\pi\sqrt{\lambda}} \end{aligned} \Rightarrow \begin{bmatrix} e^{\pi\sqrt{\lambda}} & e^{-\pi\sqrt{\lambda}} \\ e^{-\pi\sqrt{\lambda}} & e^{\pi\sqrt{\lambda}} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

for which the only solution is $a = b = 0$. Now, considering the case $\lambda < 0$, we need to satisfy

$$\begin{aligned} 0 &= a \cos(\pi\sqrt{|\lambda|}) + b \sin(\pi\sqrt{|\lambda|}) \\ 0 &= a \cos(-\pi\sqrt{|\lambda|}) + b \sin(-\pi\sqrt{|\lambda|}). \end{aligned}$$

The existence of a nonzero pair (a, b) that satisfy the above is equivalent to

$$-\cos(\pi\sqrt{|\lambda|}) \sin(\pi\sqrt{|\lambda|}) - \cos(\pi\sqrt{|\lambda|}) \sin(\pi\sqrt{|\lambda|}) = 0,$$

which is equivalent to

$$\sin(2\pi\sqrt{|\lambda|}) = 0.$$

This means that $2\pi\sqrt{|\lambda|} = i\pi$, where $i = 1, 2, \dots$. From this we conclude that the set of $\lambda < 0$ for which we can satisfy the boundary conditions is given by

$$\lambda_i = -\frac{i^2}{4}, \quad i = 1, 2, \dots$$

Thus we have found the eigenvalues of Δ_D and they indeed form a discrete set. The corresponding eigenfunctions are given by

$$\phi_i(x) = a_i \cos\left(\frac{i}{2}x\right) + b_i \sin\left(\frac{i}{2}x\right).$$

Enforcing the boundary conditions at $\pm\pi$

$$a_i \cos(\pm i\frac{\pi}{2}) + b_i \sin(\pm i\frac{\pi}{2}) = 0 \Rightarrow \begin{cases} a_i = 0 & \text{for } i \text{ even} \\ b_i = 0 & \text{for } i \text{ odd} \end{cases}.$$

These functions can be listed in a more compact manner by noticing that for i even $\sin(\frac{i}{2}x) = \pm \sin(\frac{i}{2}(x + \pi))$, and for i odd $\cos(\frac{i}{2}x) = \pm \sin(\frac{i}{2}(x + \pi))$. Thus all the eigenfunctions of Δ_D can be listed as

$$\phi_i(x) = \sin\left(\frac{i}{2}(x + \pi)\right).$$

A graph of the first four such functions is shown in Figure 1.4. The fact that these are mutually orthogonal and span all of $L^2[-\pi, \pi]$ follows the self-adjointness of Δ_D and that we have shown the discreteness of its spectrum. The functions are not normalized. A simple computation shows that $\|\phi_i\| = \sqrt{\pi}$, and so we redefine those functions to make them orthonormal by

$$\phi_i(x) = \frac{1}{\sqrt{\pi}} \sin\left(\frac{i}{2}(x + \pi)\right).$$

Example C.11. In this example, we would like to construct a self-adjoint operator whose domain is specified by some given boundary conditions. Once such an operator is constructed, its eigenfunctions will then give the desired orthonormal basis.

Suppose we would like to find an orthonormal basis $\{\phi_i\}$ for $L^2[-1, 1]$ such that

$$\phi_i(\pm 1) = 0, \quad \frac{d\phi}{dx}(\pm 1) = 0. \quad (\text{C.10})$$

One way to achieve this is to construct a self-adjoint operator on $L^2[-1, 1]$ whose domain is given by those boundary conditions. Since we have four boundary conditions, said operator will need to be a fourth order differential operator. The general form of a fourth order differential operator with constant coefficients is

$$A := a_4 D^4 + a_3 D^3 + a_2 D^2 + a_1 D + a_0 I,$$

where for simplicity of notation we use $D := \frac{d}{dx}$. As shown in Problem C.16, a necessary condition for A to be self-adjoint is that it contains only even order terms, hence $a_3 = a_1 = 0$. We investigate whether a pure fourth order term by itself is sufficient. From Problem C.16 we can state

$$\langle g, D^4 f \rangle = (g f''' - g' f'' + g'' f' - g''' f)|_{-1}^1 + \langle D^4 g, f \rangle.$$

By inspecting the boundary terms, we see that if f is such that $f(\pm 1) = f'(\pm 1) = 0$, then for all functions g with $g(\pm 1) = g'(\pm 1) = 0$ we get that $\langle g, D^4 f \rangle = \langle D^4 g, f \rangle$. This implies that if we define

$$A := \frac{d^4}{dx^4}, \quad \mathcal{D}(A) := \left\{ f \in L^2; f^{(4)} \in L^2, f(\pm 1) = f'(\pm 1) = 0 \right\},$$

then it follows that $A^* = \frac{d^4}{dx^4}$ and that $\mathcal{D}(A^*) = \mathcal{D}(A)$. The reader is asked to verify this carefully as an exercise. We have thus found a self-adjoint operator whose domain is given by the desired boundary conditions (C.10). To find the eigenvalues of A , one needs to find the real values λ for which the equation

$$\frac{d^4}{dx^4} \phi(x) - \lambda \phi(x) = 0, \quad \phi_i(\pm 1) = \frac{d\phi}{dx}(\pm 1) = 0,$$

has a solution. Those values turn out to form a discrete set. The eigenfunctions are then found by solving the differential equation at those values of λ . A complete orthogonal basis for $L^2[-1, 1]$ satisfying the boundary conditions (C.10) is thus obtained.