Section 6: Self-adjoint mappings

Matthew Macauley

School of Mathematical & Statistical Sciences
Clemson University
http://www.math.clemson.edu/~macaule/

Math 8530, Advanced Linear Algebra

Self-adjoint and anti-self-adjoint maps

Throughout, let X be a finite-dimensional inner product space.

Definition

A linear map $M\colon X \to X$ is self-adjoint if $M^* = M$, and anti-self-adjoint if $M^* = -M$.

These are also called Hermitian and anti-Hermitian, respectively.

Remark

Every linear map $M\colon X \to X$ can be decomposed into a self-adjoint part and an anti-self-adjoint part:

$$M = H + A, \qquad H = \frac{M + M^*}{2}, \qquad A = \frac{M - M^*}{2}.$$

Compare/contrast this to:

- Every matrix can be written as a sum of a symmetric and skew-symmetric matrix.
- Every real-valued function can be written as a sum of an even and an odd function.

Why do we care about self-adjoint maps?

A real-valued matrix is self-adjoint if it is symmetric: $A^T = A$.

A complex-valued matrix is self-adjoint if it is Hermitian: $\overline{A}^T = A$.

Key idea (preview)

If $A\colon X \to X$ is self-adjoint, then

- all eigenvalues of A are real
- X has an orthonormal basis of eigenvectors of A.

In spaces of functions, self-adjoint differential operators are important because they guarantee an orthogonal basis of eigenfunctions, and a "generalized Fourier series."

Another source of self-adjoint maps are quadratic forms, which we will see in this lecture.

These arise in calculus, statistics, and many other branches of higher mathematics.

We'll begin by motivating them by revisiting the second derivative test from calculus.

Second order approximations

A common problem in Calculus 1 is:

use the tangent line to approximate a function $f(x)$ *near* $a \in \mathbb{R}$.

In Calculus 2, one learns about Taylor series, and higher-order approximations.

For example, consider the function

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \cdots$$

- the 0^{th} order term is $f(a)$
- the 1^{st} order term is $f'(a)$
- the 2^{nd} order term is $\frac{1}{2}f''(a)$.

If a is a critical point (i.e., $f'(a) = 0$), then the behavior of f is governed by $f''(x)$.

Multivariate Taylor series

Now, let $f(x_1,\dots,x_n)$ be a smooth function $\mathbb{R}^n \to \mathbb{R}$. Then near a point $a \in \mathbb{R}^n$,

$$f(x) = \sum_{k=1}^{\infty} \frac{D^k f(a)}{k!}(x-a)^k = f(a) + \ell(x) + \frac{1}{2}q(x) + \cdots .$$

- the 0^{th} order term is $f(a)$
- the 1^{st} order term is $\ell(y) = \langle g, y\rangle$, where $g = \nabla f(a) = \begin{bmatrix} \frac{\partial f(a)}{\partial x_1} \\ \vdots \\ \frac{\partial f(a)}{\partial x_n} \end{bmatrix}$.
- the 2^{nd} order term is

$$q(y) = \sum_{j=1}^{n}\sum_{i=1}^{n} h_{ij} y_i y_j, \qquad \text{where } H = (h_{ij}) = \left(\frac{\partial^2 f}{\partial x_i\,\partial x_j}\right)$$

is the Hessian of f. This can be written as

$$q(y) = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \begin{bmatrix} h_{11} & \cdots & h_{1n} \\ \vdots & \ddots & \vdots \\ h_{n1} & \cdots & h_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \langle y, Hy\rangle.$$

If $a \in \mathbb{R}^n$ is a critical point (i.e., $\nabla f = 0$), then the behavior of f is governed by $q(y)$.

Quadratic forms

Definition

A quadratic form is a function

$$q\colon X \to K, \qquad q(x) = \langle x, Hx \rangle$$

for some self-adjoint linear map $H\colon X \to X$.

Consider a quadratic form

$$q(x) = x^T Hx = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \langle x, Hx \rangle.$$

If we diagonalize H, i.e., write $H = PDP^{-1} = PDP^T$ (P is orthogonal), then

$$q(x) = \langle x, Hx \rangle = x^T Hx = x^T PDP^T x.$$

If we change variables by letting $z = P^T x$,

$$q(z) = z^T Dz = \begin{bmatrix} z_1 & \cdots & z_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \sum_{i=1}^n \lambda_i z_i^2 = \langle z, Dz \rangle.$$

Quadratic forms and conic sections

Consider the quadratic form

$$q(x) = \langle x, Ax \rangle = x^T A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 5x_1^2 - 6x_1x_2 + 5x_2^2.$$

It is easy to check that $A = PDP^T$ (or $D = P^T AP$), where

$$\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

Now, let $z = P^T x$, or $x = Pz$. In this new coordinate system,

$$q(x) = q(Pz) = \langle Pz, APz \rangle = (Pz)^T A(Pz) = z^T P^T APz = \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 8z_1^2 + 2z_2^2.$$

Let's sketch the graph of $f(x_1, x_2) = 5x_1^2 - 6x_1x_2 + 5x_2^2 = 1$, which is an ellipse.

Eigenvalues and eigenvectors of self-adjoint maps

Theorem 6.1

A self-adjoint linear map $H\colon X \to X$ has only real eigenvalues, and a set of eigenvectors that forms an orthonormal basis of X.

Proof

We will show that:

1. H has only real eigenvalues
2. H has no (purely) generalized eigenvectors
3. eigenvectors corresponding to different eigenvalues are orthogonal.

Unitary diagonalization

Theorem 6.1

A self-adjoint linear map $H \colon X \to X$ has only real eigenvalues, and a set of eigenvectors that forms an orthonormal basis of X.

Corollary 6.2

If $H \colon X \to X$ is self-adjoint, then H is diagonalizable by a unitary matrix U. That is,

$$H = UDU^*, \qquad \text{where } U^*U = I.$$

Orthogonal projections onto eigenspaces

If $H\colon X \to X$ is self-adjoint with distinct eigenvectors $\lambda_1, \dots, \lambda_k$, then we can write

$$X = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}, \qquad \text{where } E_{\lambda_j} = N_{A-\lambda_j I},$$

i.e., E_{λ_j} is the eigenspace for λ_j.

This means we can write any $x \in X$ as

$$x = x^{(1)} + \cdots + x^{(k)}, \qquad \text{where } x^{(j)} \in E_{\lambda_j}.$$

Note that

$$Hx = \lambda_1 x^{(1)} + \cdots + \lambda_k x^{(k)}.$$

Denote the projection of $x \in X$ onto the eigenspace E_{λ_j} by

$$P_j\colon X \longrightarrow X, \qquad P_j\colon x \longmapsto x^{(j)}.$$

Remark

The orthogonal projection maps satisfy

(i) $P_i P_j = 0$ if $i \neq j$

(ii) $P_i^2 = P_i$

(iii) $P_i^* = P_i$.

Spectral resolutions

Definition

The decompositions

$$I = \sum_{j=1}^{k} P_j, \qquad H = \sum_{j=1}^{k} \lambda_j P_j$$

are called a resolution of the identity, and the spectral resolution of H, respectively.

Corollary 6.2 (self-adjoint maps are unitarily diagonalizable) can now be re-stated as:

Theorem 6.3

If $H: X \to X$ is self-adjoint, then there is a resolution of the identity, and a spectral resolution of H.

Functions of self-adjoint maps

Key idea

Spectral resolutions allow us to define functions on a self-adjoint map.

For example if $H\colon X \to X$ is self-adjoint with spectral resolution $H = \sum_{j=1}^{k} \lambda_j P_j$, then

- $H^2 = \sum_{j=1}^{k} \lambda_j^2 P_j$
- $H^m = \sum_{j=1}^{k} \lambda_j^m P_j$
- $p(H) = \sum_{j=1}^{k} p(\lambda_j) P_j$, for any polynomial $p(t)$
- $e^H = \sum_{j=1}^{k} e^{\lambda_j} P_j$
- $f(H) = \sum_{j=1}^{k} f(\lambda_j) P_j$, for any function $f(t)$ defined on $\lambda_1, \dots, \lambda_k$.

Commuting self-adjoint maps

When we studied Jordan canonical form, we proved the following:

Corollary 4.14

Let $A, B\colon X \to X$ be commuting diagonalizable linear maps. Then they are simultaneously diagonalizable. That is, for some invertible $P\colon X \to X$,

$$D_A = P^{-1}AP \qquad \text{and} \qquad D_B = P^{-1}BP.$$

This is *almost* enough to establish the following:

Theorem 6.4

Suppose H and K are self-adjoint commuting maps. Then they have a common spectral resolution. That is, there are orthogonal projections $P_j\colon X \to X$ such that

$$I = \sum_{j=1}^{k} P_j, \qquad H = \sum_{j=1}^{k} \lambda_j P_j, \qquad K = \sum_{j=1}^{k} \mu_j P_j$$

Proposition 6.5

Let $A\colon X \to X$ be an anti-self-adjoint map of an inner product space. Then

(i) the eigenvalues of A are purely imaginary,

(ii) X has an orthonormal basis of eigenvectors of A.

Which maps have orthonormal eigenvectors?

Notice that the following linear maps all have orthonormal bases of eigenvectors:

1. self-adjoint: $H^* = H$
2. anti-self-adjoint: $A^* = -A$
3. orthogonal: $Q^* = Q^T = Q^{-1}$
4. unitary: $U^* = \overline{U}^T = U^{-1}$

The following generalizes all of these:

Definition

A linear map $N \colon X \to X$ is normal if $N^*N = NN^*$.

Note that NN^* and N^*N are self-adjoint, and hence normal.

Theorem 6.6

If $N \colon X \to X$ is normal, then X has an orthonormal basis of eigenvectors of N.

The reason *why* this holds is because $N = \dfrac{N+N^*}{2} + \dfrac{N-N^*}{2} = H + A$.

Properties of normal linear maps

Proposition 6.7

For a linear map $M\colon X \to X$ on an inner product space,

(i) if $\langle Mx, x\rangle = 0$ for all $x \in X$, then $M = 0$.

(ii) M is normal if and only if

$$||Mx|| = ||M^*x||, \qquad \text{for all } x \in X.$$

Corollary 6.8

If $N\colon X \to X$ is normal, then N and N^* have the same nullspace.

Unitary linear maps

Proposition 6.9

Let $U \colon X \to X$ be unitary. Then

1. X has an orthonormal basis of eigenvectors
2. each eigenvalue has norm 1.

The Rayleigh quotient

We derived the spectral resolution of self-adjoint maps using the spectral theory of linear maps.

In this lecture, we'll give an alternate proof that has several advantages:

1. It doesn't assume the fundamental theorem of algebra.
2. Over $\mathbb{R}$, it avoids complex numbers.
3. It leads to a "min-max principle" which characterizes eigenvalues and eigenvectors as critical points of a particular function.

Throughout, let $H\colon X \to X$ be self-adjoint, with

- eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$
- orthonormal eigenvectors $v_1, \ldots, v_n$.

Recall that

$$\langle x, x\rangle = \sum_{j=1}^{n} a_j^2 \qquad \text{and} \qquad \langle x, Hx\rangle = \sum_{j=1}^{n} \lambda_j a_j^2.$$

The Rayleigh quotient

Definition

For a self-adjoint map $H\colon X \to X$, define the Rayleigh quotient of H as

$$R\colon X \setminus \{0\} \longrightarrow \mathbb{R}, \qquad R(x) = R_H(x) = \frac{\langle x, Hx\rangle}{\langle x, x\rangle} = \Big\langle \tfrac{x}{||x||},\ H\tfrac{x}{||x||}\Big\rangle.$$

Note that if $Hv_i = \lambda_i v_i$, then $R(v_i) = \lambda_i$.

Goal

Show that the critical points occur at the eigenvectors of H, and deduce that H has a full set of eigenvectors.

The Rayleigh quotient's minimum value

Since $R(x) = \dfrac{\langle x, Hx\rangle}{\langle x, x\rangle} = R(kx)$, we can think of R as being a map from the unit sphere.

This is compact (closed and bounded), so $R(x)$ achieves a minimum and maximum value.

Let $v \in X$ satisfy $R(v) = \min\limits_{||u||=1} R(u) := \lambda$.

Goal

Show that $Hv = \lambda v$, and that λ is the smallest eigenvalue of H.

Pick any other vector $w \in X$, a parameter $t \in \mathbb{R}$, and consider $R(v + tw)$.

The second-smallest eigenvalue of H

Let $v_1 \in X$ satisfy $R(v_1) = \min\limits_{||u||=1} R(u) := \lambda_1$.

We just showed that $Hv_1 = \lambda_1 v_1$, and λ_1 is the smallest eigenvalue.

Now, let

$$X_1 := \mathrm{Span}(v_1)^{\perp}, \qquad \text{and so} \qquad X = X_1 \oplus \mathrm{Span}(v_1), \qquad \dim X_1 = n-1.$$

Goal

(i) Show that X_1 is H-invariant

(ii) Repeat the previous step (minimize the Rayleigh quotient) on X_1

(iii) Define $X_2 = \mathrm{Span}(\{v_1, v_2\})^{\perp}$, and iterate this process.

The min-max principle

Theorem 6.10

Let $H\colon X \to X$ be self-adjoint with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$. Then

$$\lambda_k = \min_{\dim S = k} \left\{ \max_{x \in S \setminus 0} R_H(x) \right\}.$$

Summary and applications of the Rayleigh quotient

For a self-adjoint map $H\colon X \to X$, the Rayleigh quotient of H is

$$R\colon X \setminus \{0\} \longrightarrow \mathbb{R}, \qquad R(x) = R_H(x) = \frac{\langle x, Hx\rangle}{\langle x, x\rangle} = \Big\langle \tfrac{x}{||x||},\ H\tfrac{x}{||x||}\Big\rangle.$$

Summary of the Rayleigh quotient

(i) The eigenvectors of H are the critical points of $R_H(x)$, i.e., the first derivatives of $R_H(x)$ are zero iff x is an eigenvector.

(ii) $R_H(v_i) = \lambda_i$ for any $Hv_i = \lambda_i v_i$.

(iii) In particular,

$$\lambda_1 = \min_{x\neq 0} R_H(x), \qquad \lambda_n = \max_{x\neq 0} R_H(x).$$

Application to numerical linear algebra

Let H be real-symmetric with $Hv = \lambda v$. If $||v - w|| \leq \epsilon$, then $|\lambda - R_H(w)| \leq \mathcal{O}(\epsilon^2)$.

That is, $R_H(w)$ is a 2nd order Taylor approximation of the eigenvalue.

Self-adjoint differential operators

In an earlier lecture, we gave examples of orthogonal functions arising from differential equations (ODEs).

The reason *why* they exist is because they are eigenfunctions of a self-adjoint differential operator.

This is the idea of Sturm-Liouville theory, which we will summarize here.

We will not assume any knowledge about differential equations, other than what they are.

For more detailed information, see my series of lectures on *Advanced Engineering Mathematics*.

Self-adjointness of the SL operator

Definition

A Sturm-Liouville equation is a 2nd order ODE of the following form:

$$-\frac{d}{dx}\Big(p(x)y'\Big)+q(x)y=\lambda w(x)y, \qquad \text{where } p(x),\ q(x),\ w(x)>0.$$

We are usually interested in solutions $y(x)$ on $[a,b]$, under homogeneous BCs:

$$\begin{aligned} \alpha_1 y(a)+\alpha_2 y'(a)&=0 \qquad \alpha_1^2+\alpha_2^2>0\\ \beta_1 y(b)+\beta_2 y'(b)&=0 \qquad \beta_1^2+\beta_2^2>0. \end{aligned}$$

Together, this BVP is called a Sturm-Liouville (SL) problem.

Remark

Consider the linear differential operator $L=\dfrac{1}{w(x)}\Big(-\dfrac{d}{dx}\Big[p(x)\dfrac{d}{dx}\Big]+q(x)\Big)$.

$$\mathbb{C}^\infty[a,b] \xrightarrow{L_1=p(x)\frac{d}{dx}} \mathbb{C}^\infty[a,b] \xrightarrow{L_2=-\frac{1}{w(x)}\frac{d}{dx}+\frac{q(x)}{w(x)}} \mathbb{C}^\infty[a,b]$$

$$y \longmapsto p(x)y'(x) \longmapsto \frac{-1}{w(x)}\frac{d}{dx}\Big[p(x)y'(x)\Big]+\frac{q(x)}{w(x)}y(x)$$

An SL equation is just an eigenvalue equation: $Ly=\lambda y$, and $L=L_2\circ L_1$ is self-adjoint!

Main theorem

The SL operator $L = \frac{1}{w(x)}\Big(-\frac{d}{dx}\Big[p(x)\frac{d}{dx}\Big] + q(x)\Big)$ is self-adjoint on $\mathcal{C}^{\infty}_{\alpha,\beta}[a,b]$ with respect to the inner product

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}w(x)\,dx.$$

This means that:

(a) The eigenvalues are real and can be ordered so $\lambda_1 < \lambda_2 < \lambda_3 < \cdots \to \infty$.

(b) Each eigenvalue λ_i has a unique (up to scalars) eigenfunction $y_i(x)$.

(c) W.r.t. the inner product $\langle f, g\rangle := \int_a^b f(x)\overline{g(x)}w(x)\,dx$, the eigenfunctions form an orthogonal basis on the subspace of functions $\mathcal{C}^{\infty}_{\alpha,\beta}[a,b]$ that satisfy the BCs.

Definition

If $f \in \mathcal{C}^{\infty}_{\alpha,\beta}[a,b]$, then f can be written uniquely as a linear combination of the eigenfunctions. That is,

$$f(x) = \sum_{n=1}^{\infty} c_n y_n(x), \qquad \text{where } c_n = \frac{\langle f, y_n\rangle}{\langle y_n, y_n\rangle} = \frac{\int_a^b f(x)\overline{y_n(x)}w(x)\,dx}{\int_a^b ||y_n(x)||^2 w(x)\,dx}.$$

This is called a generalized Fourier series with respect to the orthogonal basis $\{y_n(x)\}$ and weighting function $w(x)$.

Fourier series

Dirichlet BCs

$-y'' = \lambda y, \quad y(0) = 0, \quad y(\pi) = 0$ is an SL problem with:

- Eigenvalues: $\lambda_n = n^2, \quad n = 1, 2, 3, \ldots$.
- Eigenfunctions: $y_n(x) = \sin(nx)$.

The orthogonality of the eigenvectors means that

$$\langle y_m, y_n \rangle := \int_0^\pi y_m(x) y_n(x) w(x)\, dx = \int_0^\pi \sin(mx)\, \sin(nx)\, dx = \begin{cases} 0 & \text{if } m \neq n \\ \pi/2 & \text{if } m = n. \end{cases}$$

Note that this means that $||y_n|| := \langle y_n, y_n \rangle^{1/2} = \sqrt{\pi/2}$.

Fourier series: *any* function $f(x)$, continuous on $[0, \pi]$ satisfying $f(0) = 0$, $f(\pi) = 0$ can be written *uniquely* as

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx$$

where

$$b_n = \frac{\langle f, \sin nx \rangle}{\langle \sin nx, \sin nx \rangle} = \frac{\int_0^\pi f(x) \sin nx\, dx}{|| \sin nx||^2} = \frac{2}{\pi} \int_0^\pi f(x) \sin nx\, dx.$$

Fourier series

Neumann BCs

$-y'' = \lambda y, \quad y'(0) = 0, \quad y'(\pi) = 0$ is an SL problem with:

- Eigenvalues: $\lambda_n = n^2, \quad n = 0, 1, 2, 3, \ldots.$
- Eigenfunctions: $y_n(x) = \cos(nx)$.

The orthogonality of the eigenvectors means that

$$\langle y_m, \, y_n \rangle := \int_0^\pi y_m(x) y_n(x) w(x)\, dx = \int_0^\pi \cos(mx)\, \cos(nx)\, dx = \begin{cases} 0 & \text{if } m \neq n \\ \pi/2 & \text{if } m = n > 0. \end{cases}$$

Note that this means that $||y_n|| := \langle y_n, \, y_n \rangle^{1/2} = \begin{cases} \sqrt{\pi/2} & n > 0 \\ \sqrt{\pi} & n = 0. \end{cases}$

Fourier series: *any* function $f(x)$, continuous on $[0, \pi]$ satisfying $f'(0) = 0, \ f'(\pi) = 0$ can be written *uniquely* as

$$f(x) = \sum_{n=0}^{\infty} a_n \cos nx$$

where

$$a_n = \frac{\langle f, \, \cos nx \rangle}{\langle \cos nx, \, \cos nx \rangle} = \frac{\int_0^\pi f(x) \cos nx \, dx}{|| \cos nx||^2} = \frac{2}{\pi} \int_0^\pi f(x) \cos nx \, dx.$$

More complicated Sturm-Liouville problems

Every 2nd order linear homogeneous ODE, $y'' + P(x)y' + Q(x)y = 0$ can be written in self-adjoint or "Sturm-Liouville form":

$$-\frac{d}{dx}\Big(p(x)y'\Big) + q(x)y = \lambda w(x)y, \qquad \text{where } p(x),\ q(x),\ w(x) > 0.$$

Examples from physics and engineering

- Legendre's equation: $(1-x^2)y'' - 2xy' + n(n+1)y = 0$. *Used for modeling spherically symmetric potentials in the theory of Newtonian gravitation and in electricity & magnetism (e.g., the wave equation for an electron in a hydrogen atom).*
- Parametric Bessel's equation: $x^2y'' + xy' + (\lambda x^2 - \nu^2)y = 0$. *Used for analyzing vibrations of a circular drum.*
- Chebyshev's equation: $(1-x^2)y'' - xy' + n^2y = 0$. *Arises in numerical analysis techniques.*
- Hermite's equation: $y'' - 2xy' + 2ny = 0$. *Used for modeling simple harmonic oscillators in quantum mechanics.*
- Laguerre's equation: $xy'' + (1-x)y' + ny = 0$. *Arises in a number of equations from quantum mechanics.*
- Airy's equation: $y'' - k^2xy = 0$. *Models the refraction of light.*

Legendre's differential equation

Consider the following Sturm-Liouville problem, defined on $(-1, 1)$:

$$-\frac{d}{dx}\left[(1-x^2)\frac{d}{dx}y\right] = \lambda y, \qquad \left[p(x) = 1-x^2, \quad q(x) = 0, \quad w(x) = 1\right].$$

The eigenvalues are $\lambda_n = n(n+1)$, $n \in \mathbb{N}$, and the eigenfunctions solve Legendre's equation:

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0.$$

For each n, one solution is a degree-n "Legendre polynomial"

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\left[(x^2-1)^n\right].$$

They are orthogonal with respect to the inner product $\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)\,dx$.

It can be checked that

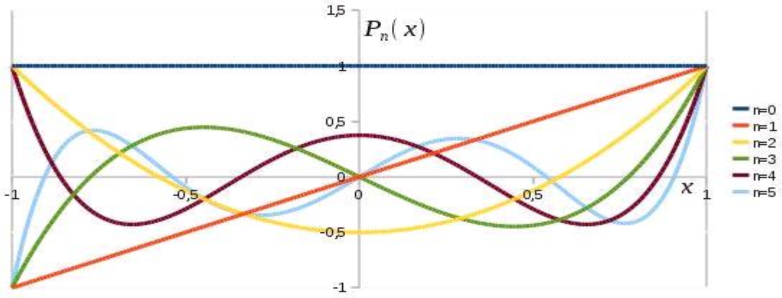
$$\langle P_m, P_n \rangle = \int_{-1}^{1} P_m(x)P_n(x)\,dx = \frac{2}{2n+1}\delta_{mn}.$$

By orthogonality, every function f, continuous on $-1 < x < 1$, can be expressed using Legendre polynomials:

$$f(x) = \sum_{n=0}^{\infty} c_n P_n(x), \qquad \text{where} \quad c_n = \frac{\langle f,\, P_n \rangle}{\langle P_n,\, P_n \rangle} = (n + \tfrac{1}{2})\,\langle f,\, P_n \rangle$$

Legendre polynomials

$$
\begin{aligned}
P_0(x) &= 1 \\
P_1(x) &= x \\
P_2(x) &= \tfrac{1}{2}(3x^2 - 1) \\
P_3(x) &= \tfrac{1}{2}(5x^3 - 3x) \\
P_4(x) &= \tfrac{1}{8}(35x^4 - 30x^2 + 3) \\
P_5(x) &= \tfrac{1}{8}(63x^5 - 70x^3 + 15x) \\
P_6(x) &= \tfrac{1}{8}(231x^6 - 315x^4 + 105x^2 - 5) \\
P_7(x) &= \tfrac{1}{16}(429x^7 - 693x^5 + 315x^3 - 35x)
\end{aligned}
$$



Parametric Bessel's differential equation

Consider the following Sturm-Liouville problem on $[0, a]$:

$$-\frac{d}{dx}(xy') - \frac{\nu^2}{x}y = \lambda xy, \qquad \left[p(x) = x, \quad q(x) = -\frac{\nu^2}{x}, \quad w(x) = x\right].$$

For a fixed ν, the eigenvalues are $\lambda_n = \omega_n^2 := \alpha_n^2/a^2$, for $n = 1, 2, \dots$.

Here, α_n is the n^{th} positive root of $J_\nu(x)$, the Bessel functions of the first kind of order ν.

The eigenfunctions solve the parametric Bessel's equation:

$$x^2y'' + xy' + (\lambda x^2 - \nu^2)y = 0.$$

Fixing ν, for each n there is a solution $J_{\nu n}(x) := J_\nu(\omega_n x)$.

They are orthogonal with repect to the inner product $\langle f, g\rangle = \int_0^a f(x)g(x)\, x\, dx$.

It can be checked that

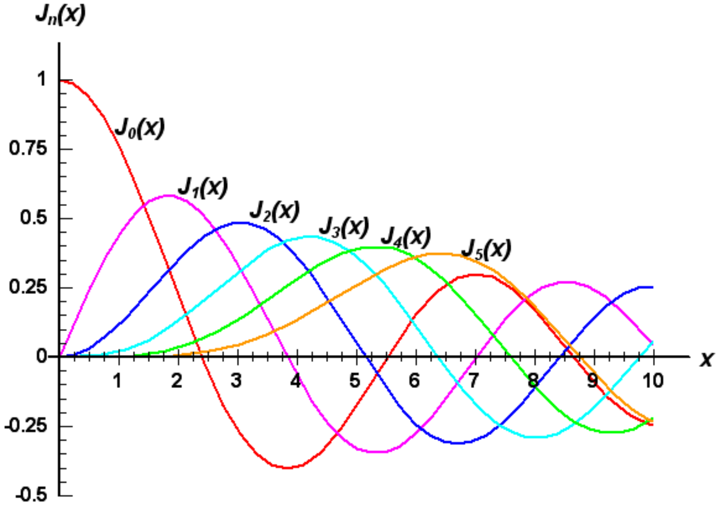
$$\langle J_{\nu n}, J_{\nu m}\rangle = \int_0^a J_\nu(\omega_n x)J_\nu(\omega_m x)\, x\, dx = 0, \quad \text{if } n \neq m.$$

By orthogonality, every continuous function $f(x)$ on $[0, a]$ can be expressed in a "Fourier-Bessel" series:

$$f(x) \sim \sum_{n=0}^{\infty} c_n J_\nu(\omega_n x), \qquad \text{where} \quad c_n = \frac{\langle f,\, J_{\nu n}\rangle}{\langle J_{\nu n},\, J_{\nu n}\rangle}.$$

Bessel functions (of the first kind)

$$J_\nu(x) = \sum_{m=0}^{\infty} (-1)^m \frac{1}{m!(\nu+m)!} \left(\frac{x}{2}\right)^{2m+\nu} .$$



Chebyshev's differential equation

Consider the following Sturm-Liouville problem on $[-1,1]$:

$$-\frac{d}{dx}\left[\sqrt{1-x^2}\,\frac{d}{dx}y\right]=\lambda\frac{1}{\sqrt{1-x^2}}y, \qquad \left[p(x)=\sqrt{1-x^2},\quad q(x)=0,\quad w(x)=\tfrac{1}{\sqrt{1-x^2}}\right].$$

The eigenvalues are $\lambda_n=n^2$ for $n\in\mathbb{N}$, and the eigenfunctions solve Chebyshev's equation:

$$(1-x^2)y''-xy'+n^2y=0.$$

For each n, one solution is a degree-n "Chebyshev polynomial," defined recursively by

$$T_0(x)=1,\qquad T_1(x)=x,\qquad T_{n+1}(x)=2xT_n(x)-T_{n-1}(x).$$

They are orthogonal with repect to the inner product $\langle f,g\rangle=\int_{-1}^{1}\frac{f(x)g(x)}{\sqrt{1-x^2}}\,dx$.

It can be checked that

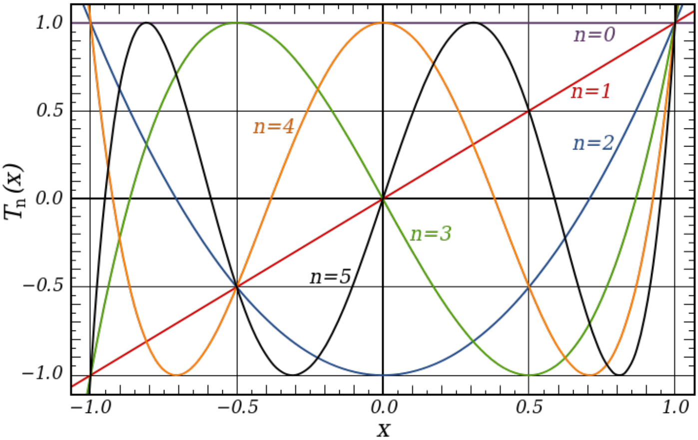
$$\langle T_m,T_n\rangle=\int_{-1}^{1}\frac{T_m(x)T_n(x)}{\sqrt{1-x^2}}\,dx=\begin{cases}\frac{1}{2}\pi\delta_{mn} & m\neq 0, n\neq 0\\ \pi & m=n=0\end{cases}$$

By orthogonality, every function $f(x)$, continuous for $-1<x<1$, can be expressed using Chebyshev polynomials:

$$f(x)\sim\sum_{n=0}^{\infty}c_nT_n(x),\qquad \text{where}\quad c_n=\frac{\langle f,\,T_n\rangle}{\langle T_n,\,T_n\rangle}=\frac{2}{\pi}\langle f,\,T_n\rangle,\ \text{ if } n>0.$$

Chebyshev polynomials (of the first kind)

$$T_0(x) = 1 \qquad T_4(x) = 8x^4 - 8x^2 + 1$$
$$T_1(x) = x \qquad T_5(x) = 16x^5 - 20x^3 + 5x$$
$$T_2(x) = 2x^2 - 1 \qquad T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1$$
$$T_3(x) = 4x^3 - 3x \qquad T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x$$



Hermite's differential equation

Consider the following Sturm-Liouville problem on $(-\infty, \infty)$:

$$-\frac{d}{dx}\left[e^{-x^2}\frac{d}{dx}y\right] = \lambda e^{-x^2}y, \qquad \left[p(x) = e^{-x^2}, \quad q(x) = 0, \quad w(x) = e^{-x^2}\right].$$

The eigenvalues are $\lambda_n = 2n$ for $n = 1, 2, \dots,$ and the eigenfunctions solve Hermite's equation:

$$y'' - 2xy' + 2ny = 0.$$

For each n, one solution is a degree-n "Hermite polynomial," defined by

$$H_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2} = \left(2x - \frac{d}{dx}\right)^n \cdot 1$$

They are orthogonal with repect to the inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)e^{-x^2}\,dx$.

It can be checked that

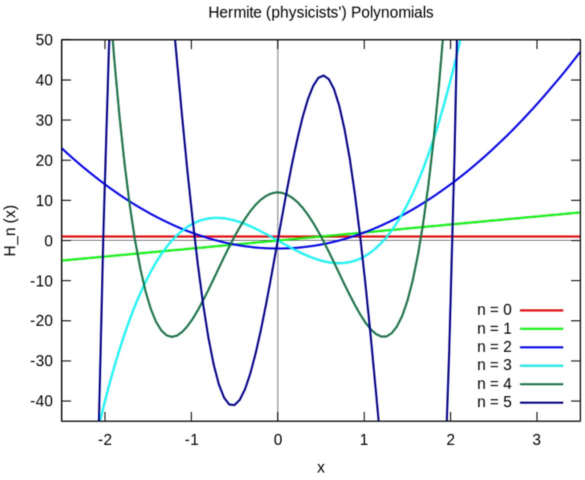
$$\langle H_m, H_n\rangle = \int_{-\infty}^{\infty} H_m(x)H_n(x)e^{-x^2}\,dx = \sqrt{\pi}2^n n!\delta_{mn}.$$

By orthogonality, every function $f(x)$ satisfying $\int_{-\infty}^{\infty} f^2e^{-x^2}dx < \infty$ can be expressed using Hermite polynomials:

$$f(x) \sim \sum_{n=0}^{\infty} c_nH_n(x), \qquad \text{where} \quad c_n = \frac{\langle f,\ H_n\rangle}{\langle H_n,\ H_n\rangle} = \frac{\langle f, H_n\rangle}{\sqrt{\pi}2^n n!}.$$

Hermite polynomials

$$H_0(x) = 1 \qquad H_4(x) = 16x^4 - 48x^2 + 12$$
$$H_1(x) = 2x \qquad H_5(x) = 32x^5 - 160x^3 + 120x$$
$$H_2(x) = 4x^2 - 2 \qquad H_6(x) = 64x^6 - 480x^4 + 720x^2 - 120$$
$$H_3(x) = 8x^3 - 12x \qquad H_7(x) = 128x^7 - 1344x^5 + 3360x^3 - 1680x$$



Hermite functions

The Hermite functions can be defined from the Hermite polynomials as

$$\psi_n(x) = \left(2^n n! \sqrt{\pi}\right)^{-\frac{1}{2}} e^{-\frac{x^2}{2}} H_n(x) = (-1)^n \left(2^n n! \sqrt{\pi}\right)^{-\frac{1}{2}} e^{-\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-x^2}.$$

They are orthonormal with respect to the inner product

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) g(x)\, dx.$$

Every real-valued function f such that $\int_{-\infty}^{\infty} f^2\, dx < \infty$ "can be expressed uniquely" as

$$f(x) \sim \sum_{n=0}^{\infty} c_n \psi_n(x)\, dx, \qquad \text{where } c_n = \langle f, \psi_n \rangle = \int_{-\infty}^{\infty} f(x) \psi_n(x)\, dx.$$

These are solutions to the time-independent Schrödinger ODE: $-y'' + x^2 y = (2n+1)y$.