Sturm-Liouville Theory

1 Self-adjoint Matrices

A real *self-adjoint* matrix is just another name for a symmetric matrix, i.e. satisfies $M^\top = M$, because $M^\top$ is also called the adjoint of M. This terminology is useful when we generalise matrices to linear operators.

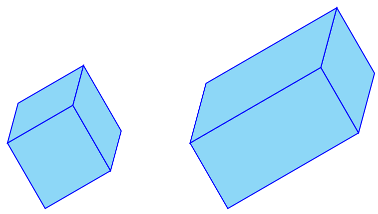
Solving a Linear Equation

Suppose we want to solve $M\mathbf{x} = \mathbf{b}$, where $\mathbf{x}, \mathbf{b} \in \mathbb{C}^n$ and $M = M^\top$ is an $n \times n$ real self-adjoint matrix. If we know how to diagonalise M, then this problem is quick!

First we will recall the geometric meaning of self-adjoint matrices, via the following theorem.

Theorem 1.1. A self-adjoint matrix has an orthonormal basis of eigenvectors.

The picture you should have in your head is a scaling of a cube into a cuboid in a rotated (or reflected) basis. The basis consists of eigenvectors, and the eigenvalues represent the stretch factor in each of the directions.



Since M can be diagonalised, we can write it in the form

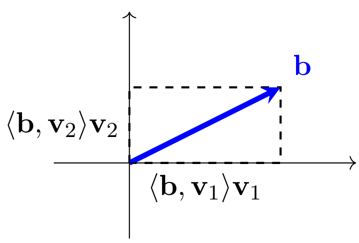
$$M = PDP^{-1} = PDP^\top,$$

with

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \quad P = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix},$$

where $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are the orthonormal eigenvectors with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. How does this help us solve $M\mathbf{x} = \mathbf{b}$? The key observation is that if we want to decompose $\mathbf{b}$ into a linear combination of $\mathbf{v}_i$, by orthogonality we can project $\mathbf{b}$ onto the span of $\mathbf{v}_i$, giving us

$$\mathbf{b} = \sum_{i=1}^n \langle \mathbf{b}, \mathbf{v}_i \rangle \mathbf{v}_i.$$



Now a solution to $M\mathbf{x} = \mathbf{b}$ can be found by inspection:

$$\mathbf{x} = \sum_{i=1}^{n} \frac{1}{\lambda_i} \langle \mathbf{b}, \mathbf{v}_i \rangle \mathbf{v}_i,$$

because M basically just scales $\mathbf{x}$ in each of the directions $\mathbf{v}_i$ by λ_i to get $\mathbf{b}$, so we *undo* this by *unscaling* by a factor of $1/\lambda_i$. We can verify this solution directly:

$$\begin{aligned}
M\mathbf{x} &= M \sum_{i=1}^{n} \frac{1}{\lambda_i} \langle \mathbf{b}, \mathbf{v}_i \rangle \mathbf{v}_i \\
&= \sum_{i=1}^{n} \frac{1}{\lambda_i} \langle \mathbf{b}, \mathbf{v}_i \rangle M \mathbf{v}_i && \text{linearity} \\
&= \sum_{i=1}^{n} \frac{1}{\lambda_i} \langle \mathbf{b}, \mathbf{v}_i \rangle \lambda_i \mathbf{v}_i && \text{scaling} \\
&= \sum_{i=1}^{n} \langle \mathbf{b}, \mathbf{v}_i \rangle \mathbf{v}_i \\
&= \mathbf{b}.
\end{aligned}$$

The result can also be put in a more suggestive form:

$$\begin{aligned}
\mathbf{x} &= \sum_{i=1}^{n} \frac{1}{\lambda_i} \left(\mathbf{v}_i^\top \mathbf{b} \right) \mathbf{v}_i \\
&= \left(\sum_{i=1}^{n} \frac{1}{\lambda_i} \mathbf{v}_i \mathbf{v}_i^\top \right) \mathbf{b}.
\end{aligned}$$

The matrix in bracket is therefore M^{-1}. You can see that it sends $\mathbf{v}_j \mapsto \frac{1}{\lambda_j} \mathbf{v}_j$:

$$\begin{aligned}
\left(\sum_{i=1}^{n} \frac{1}{\lambda_i} \mathbf{v}_i \mathbf{v}_i^\top \right) \mathbf{v}_j &= \sum_{i=1}^{n} \frac{1}{\lambda_i} \mathbf{v}_i (\mathbf{v}_i^\top \mathbf{v}_j) \\
&= \sum_{i=1}^{n} \frac{1}{\lambda_i} \mathbf{v}_i \delta_{ij} \\
&= \frac{1}{\lambda_j} \mathbf{v}_j,
\end{aligned}$$

which is exactly what you would expect M^{-1} to do; undo the scaling caused by M.

Dynamical Systems

Another use of diagonalisation is in solving

$$\frac{d}{dt} \mathbf{x} = M\mathbf{x}, \qquad \mathbf{x}(0) = \mathbf{x}_0,$$

where again M is self-adjoint. Here we want to find how $\mathbf{x}(t)$ evolves given the initial condition $\mathbf{x}_0$. Again, if we diagonalise M, the equation is easy to solve:

$$\frac{d}{dt} \mathbf{x} = PDP^{-1} \mathbf{x}.$$

Multiply both sides by P^{-1}:

$$\frac{d}{dt}P^{-1}\mathbf{x} = DP^{-1}\mathbf{x}.$$

Let $\mathbf{y} = P^{-1}\mathbf{x}$. Then we have a dynamical system for $\mathbf{y}$:

$$\frac{d}{dt}\mathbf{y} = D\mathbf{y}, \qquad \mathbf{y}(0) = P^{-1}\mathbf{x}_0.$$

Since D is diagonal, the system becomes decoupled, so we are just solving component-wise. For each i (no summation convention):

$$\frac{dy_i}{dt} = \lambda_i y_i \implies y_i(t) = c_i e^{\lambda_i t},$$

where c_i are constants of integration. Putting this altogether:

$$\mathbf{y}(t) = \sum_{i=1}^{n} e^{\lambda_i t} c_i \mathbf{e}_i,$$

and multiply through by P gives

$$\mathbf{x}(t) = \sum_{i=1}^{n} e^{\lambda_i t} c_i \mathbf{v}_i. \tag{$*$}$$

The constants of integration can be determined by the initial condition:

$$\mathbf{x}_0 = \sum_{i=1}^{n} c_i \mathbf{v}_i \implies c_i = \langle \mathbf{x}_0, \mathbf{v}_i \rangle.$$

Hence the solution is

$$\mathbf{x}(t) = \sum_{i=1}^{n} e^{\lambda_i t} \langle \mathbf{x}_0, \mathbf{v}_i \rangle \mathbf{v}_i.$$

If we draw the (hyper)cube with corner $\mathbf{x}(t)$, then its edge in the ith direction is expanding/contracting exponentially at rate λ_i.

Remark 1.2. The fact $M^\top = M$ was not used until the very last step. In fact, $(*)$ is still true for diagonalisable matrices, though the $\mathbf{v}_i$'s need not be orthonormal any more. In this case, $c_i = \langle \mathbf{x}_0, \mathbf{v}_i \rangle$ is no longer true. These c_i's will exist by the fact that $\mathbf{v}_i$'s form a basis, but determining these coefficients will require Gaussian elimination to solve

$$\begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{x}_0,$$

which takes far longer than computing a dot product.

2 Self-Adjoint Operator

Consider the Dirichlet boundary problem

$$\begin{cases} \mathcal{L}y = f & x \in [a,b] \\ y = 0 & x \in \{a,b\}, \end{cases} \tag{DBP}$$

where $\mathcal{L}$ is a differential operator. Examples to keep in mind include:

Heat operator. $\mathcal{L}y = -y''$;

Radial part of Polar Laplacian. $\mathcal{L}y(r) = -\frac{1}{r}\frac{d}{dr}\left(r\frac{dy}{dr}\right)$. This gives Bessel functions.

Polar part of spherical Laplacian. Letting $x = \cos\theta$,

$$\mathcal{L}P(x) = -\frac{d}{dx}\left((1-x^2)\frac{d}{dx}\right)P(x).$$

This gives Legendre Polynomials.

In this set of notes, we will examine the heat operator $\mathcal{L} = -\frac{d^2}{dx^2}$ in detail. This means we are solving

$$\begin{cases} -y'' = f & x \in [a,b] \\ y = 0 & x \in \{a,b\}, \end{cases} \tag{DH}$$

which models the equilibrium temperature in $[a,b]$ subject to the Dirichlet boundary condition $y(a) = y(b) = 0$. We claim that solving $\mathcal{L}y = f$ is basically solving $M\mathbf{x} = \mathbf{b}$ in infinite dimension. First we need to say what we mean by infinite dimension.

A matrix is a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^n$. Here, the heat operator has domain

$$\text{domain} = \left\{y \in C^2([a,b]) : y(a) = y(b) = 0\right\}, \quad \text{codomain} = C([a,b]).$$

Both are vector spaces of infinite dimension. Note that the boundary condition is encoded in the domain of the linear operator.

Now we need to define a suitable notion of *self-adjointness*. A real symmetric matrix was defined to be $M^\top = M$, and here we generalise the idea of the transpose:

Definition 2.1. A matrix M is *self-adjoint* with respect to the inner product $\langle\cdot,\cdot\rangle$ if for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\langle \mathbf{x}, M\mathbf{y}\rangle = \langle M\mathbf{x}, \mathbf{y}\rangle.$$

In $\mathbb{R}^n$, we usually take the inner product to be the dot product, in which case we recover the definition $M^\top = M$, but in the infinite-dimensional case, it may sometimes be more convenient to choose a different inner product than the obvious one.

Remark 2.2. In $\mathbb{R}^n$, all inner products are of the form

$$\mathbf{x}^\top A\mathbf{y}, \tag{$*$}$$

where A is a positive-definite symmetric matrix. Positive-definite means its eigenvalues are all positive. This is to ensure $\mathbf{x}^\top A\mathbf{x} \geq 0$ with equality if and only if $\mathbf{x} = \mathbf{0}$. Matrices with this property are called *non-degenerate*. Here's a quick proof of eigenvalues all positive $\iff$ positive definite:

Proof. Suppose A is a symmetric matrix that is not positive-definite. Let $\mathbf{v}$ be a unit eigenvector of non-positive eigenvalue λ. Then

$$\mathbf{v}^\top A\mathbf{v} = \lambda\mathbf{v}^\top\mathbf{v} = \lambda \leq 0,$$

a contradiction to the non-degenerate condition.

Conversely, suppose A is positive-definite. Then all of its eigenvalues $\lambda_i > 0$. For any nonzero $\mathbf{x}$, we can write it as a linear combination of eigenvectors,

$$\mathbf{x} = \sum_{i=1}^n c_i\mathbf{v}_i,$$

where $c_i = \mathbf{x}^\top \mathbf{v}_i$ because $\mathbf{v}_i$'s are orthonormal. So

$$\begin{aligned}
\mathbf{x}^\top A\mathbf{x} &= \mathbf{x}^\top A \sum_{i=1}^n c_i \mathbf{v}_i \\
&= \mathbf{x}^\top \sum_{i=1}^n \lambda_i c_i \mathbf{v}_i \\
&= \sum_{i=1}^n \lambda_i c_i \mathbf{x}^\top \mathbf{v}_i \\
&= \sum_{i=1}^n \lambda_i c_i^2 > 0.
\end{aligned}$$

The inequality is strict because c_i's are not all zero. □

The natural generalisation of self-adjointness is therefore:

Definition 2.3. A linear operator $\mathcal{L}$ is *self-adjoint* with respect to the inner product $\langle \cdot, \cdot \rangle$ if for all functions f, g,

$$\langle f, \mathcal{L}g \rangle = \langle \mathcal{L}f, g \rangle.$$

The inner products on the vector space of functions we consider will be of the form

$$\langle f, g \rangle_w \coloneqq \int_a^b w(x) f(x) g(x)\, dx,$$

where $w\colon [a, b] \to \mathbb{R}_{\geq 0}$ is a continuous function which is only zero at finitely many points. It is called the *weight function* and takes the role of A in equation $(*)$ above. The condition on w is to ensure non-degeneracy of the inner product.

The self-adjoint operators, like symmetric matrices, have an orthonormal eigenbasis.

Theorem 2.4 (Spectral Theorem)**.** If $\mathcal{L}$ is self-adjoint, then there exists $y_1, y_2, \ldots$ of eigenfunctions such that any suitably nice function can be written uniquely as an *infinite* linear combination of these functions.

This theorem can be thought of as saying $\mathcal{L}$ can be decomposed as

$$\mathcal{L} = PDP^{-1},$$

where D is the infinite-dimensional matrix $\operatorname{diag}(\lambda_1, \lambda_2, \ldots)$ and P is the infinite-dimensional orthogonal matrix with infinitely many columns $y_1, y_2, \ldots$.

3 Sturm-Liouville

We can now solve $\mathcal{L}y = f$ for self-adjoint $\mathcal{L}$. There are basically 3 steps, with the last one being "done by inspection".

Step 1. Show that $\mathcal{L}$ is self-adjoint with respect to some inner product $\langle \cdot, \cdot \rangle$ which you should specify.

Step 2. Find the eigenfunctions $y_i(x)$. If it's complicated you will probably need to use the series method.

Step 3. Put it all together: we are attempting to solve $\mathcal{L}y = f$. Like the finite-dimensional analogue, we know that f can be written as an "infinite" linear combination of the eigenfunctions:

$$\mathcal{L}y(x) = f(x) = \sum_{n=1}^{\infty} \langle f, y_n \rangle_w y_n(x),$$

so by inspection, the solution is

$$\boxed{y(x) = \sum_{n=1}^{\infty} \frac{\langle f, y_n \rangle_w}{\lambda_n} y_n(x),}$$

which we think as undoing the scaling caused by $\mathcal{L}$. That's all. In practice you just substitute w found in Step 1 and y_i found in Step 2 into this formula.

Remark 3.1. We can also put this equation is a more suggestive form:

$$\begin{aligned} y(x) &= \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \left(\int_a^b w(\xi) f(\xi) y_n(\xi) \, d\xi \right) y_n(x) \\ &= \int_a^b \left(w(\xi) \sum_{n=1}^{\infty} \frac{y_n(x) y_n(\xi)}{\lambda_n} \right) f(\xi) \, d\xi \end{aligned}$$

$$\implies \boxed{y(x) = \int_a^b \left(w(\xi) \sum_{i=1}^{\infty} \frac{y_i(x) y_i(\xi)}{\lambda_i} \right) f(\xi) \, d\xi.}$$

There is a name for the term in parantheses:

$$G(x; \xi) \coloneqq w(\xi) \sum_{n=1}^{\infty} \frac{y_n(x) y_n(\xi)}{\lambda_n}$$

is called the *Green's function* of $\mathcal{L}$, and can be regarded as $\mathcal{L}^{-1}$.

4 Heat Operator

We will apply Sturm-Liouville Theory to solve

$$\begin{cases} -y'' = f & x \in [0, L] \\ y = 0 & x \in \{0, L\}. \end{cases}$$

Step 1. This operator is self-adjoint with respect to the weight $w \equiv 1$. We can verify this by taking any function y_1, y_2 in the domain of $\mathcal{L}$, i.e.

$$y_1, y_2 \in \left\{ y \in C^2([0, L]) : y(0) = y(L) = 0 \right\},$$

and compute:

$$\begin{aligned} \int_0^L y_1(x) \mathcal{L} y_2(x) \, dx &= \int_0^L y_1(x)(-y_2''(x)) \, dx \\ &= \cancelto{0}{[-y_1(x) y_2'(x)]_0^L} + \int_0^L y_1'(x) y_2'(x) \, dx \end{aligned}$$

$$= \cancelto{0}{[y_1'(x)y_2(x)]_0^L} - \int_0^L y_1''(x)y_2(x)\,dx$$
$$= \int_0^L \mathcal{L}y_1(x)y_2(x)\,dx.$$

Note that it is at this step where we use the boundary data. The trick was to encode the boundary data in the domain of $\mathcal{L}$. This doesn't have a finite-dimensional analogue.

Step 2. The eigenfunctions can be found by solving

$$-y'' = \lambda y, \qquad y(0) = y(L) = 0.$$

We can solve this using ordinary second-order methods. This gives countably many solutions:

$$y_n(x) \propto \sin\frac{n\pi x}{L},$$

with eigenvalue $\lambda_n = \frac{n^2\pi^2}{L^2}$, where $n = 1, 2, 3, \ldots$. We normalise this:

$$\left\|\sin\frac{n\pi x}{L}\right\|^2 = \int_0^L \sin^2\frac{n\pi x}{L}\,dx = \frac{L}{2}$$
$$\implies y_n(x) = \sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}.$$

Step 3. Thus the solution is

$$y(x) = \int_0^L \left(w(\xi)\sum_{n=1}^{\infty}\frac{y_n(x)y_n(\xi)}{\lambda_n}\right) f(\xi)\,d\xi$$
$$= \int_0^L \left(\sum_{n=1}^{\infty}\frac{\sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}\sqrt{\frac{2}{L}}\sin\frac{n\pi\xi}{L}}{n^2\pi^2/L^2}\right) f(\xi)\,d\xi$$
$$= \int_0^L \left(\sum_{n=1}^{\infty}\frac{2L}{n^2\pi^2}\sin\frac{n\pi x}{L}\sin\frac{n\pi\xi}{L}\right) f(\xi)\,d\xi,$$

so we just need to substitute a concrete forcing term f to get the solution.

5 Physical Meaning of Eigenfunctions

The eigenfunctions of the heat operator satisfies $-y'' = \lambda y$. We saw that the eigenfunctions are

$$\sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}.$$

What is the physical significance of these functions?

These are the heat distributions which decay without changing the nature of their shape in the time-dependent heat equation

$$\frac{\partial y}{\partial t} - \frac{\partial^2 y}{\partial x^2} = 0 \qquad \text{or} \qquad \frac{\partial y}{\partial t} + \mathcal{L}y = 0. \tag{5.1}$$

Let me explain what I mean by "nature of their shape". If we initialise the heat as

$$y(x, 0) = \sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L},$$

then

$$y(x,t) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L} e^{-n\pi t/L},$$

so the shape remains the same (still a sine) but the amplitude decays exponentially. The higher modes (large n) decay faster than the lower modes (small n).

If we initialise the heat as something that is not an eigenfunction, then the evolution of the heat will not be as simple as an exponential decay of the amplitude. To calculate the evolution, we should decompose it into eigenfunctions, and make each of those coefficients decay exponentially, then add them back up:

$$y(x,0) = \sum_{n=1}^{\infty} c_n y_n(x) \tag{5.2}$$

$$\implies y(x,t) = \sum_{n=1}^{\infty} c_n y_n(x) e^{-\lambda_n t}. \tag{5.3}$$

The Heat Equation $\frac{\partial y}{\partial t} + \mathcal{L}y = 0$ (Equation 5.1) with initial condition $y(x,0) = \sum_{n=1}^{\infty} c_n y_n(x)$ (Equation 5.2) is the infinite dimensional analogue of solving

$$\frac{d}{dt}\mathbf{x} + M\mathbf{x} = \mathbf{0}, \qquad \mathbf{x}(0) = \sum_{n=1}^{\infty} c_n \mathbf{v}_n,$$

which has solution

$$\mathbf{x}(t) = \sum_{n=1}^{\infty} c_n \mathbf{v}_n e^{-\lambda_n t}.$$

6 Bessel Example

For the 1D heat equation, we did not need to make use of the weight function. We simply set it to 1. In this example, we will see how the weight function becomes important.

We are solving 2D Poisson's Equation for the case that the forcing function f is radially symmetric.

$$\begin{cases} \mathcal{L}u \coloneqq -\dfrac{1}{r}\dfrac{d}{dr}\left(r\dfrac{du}{dr}\right) = f(r) & r < 1 \\ \qquad\qquad\qquad\quad u = 0 & r = 1. \end{cases}$$

This solves the equilibrium temperature in the disk with radially symmetric heat source f and 0 temperature at the wall $r = 1$.

Step 1. We set $w(r) = r$ this time (the next section will say how we found this). Now $\mathcal{L}$ is self-adjoint with respect to $\langle \cdot, \cdot \rangle_r$, as we can verify: take any

$$u_1, u_2 \in \{u \in C^2([0,1]) : u(1) = 0\} = \operatorname{dom} \mathcal{L}.$$

Then

$$\begin{aligned} \int_0^1 u_1(r)\mathcal{L}u_2(r) r \, dr &= \int_0^1 u_1(r)(-ru_2'(r))' \, dr \\ &= -[\cancelto{0}{u_1(1)} 1 u_2'(1) - u_1(0) 0 u_2'(0)] + \int_0^1 r u_1'(r) u_2'(r) \, dr, \end{aligned}$$

$$\begin{aligned}\int_0^1 u_2(r)\mathcal{L}u_1(r)r\,dr &= \int_0^1 u_2(r)(-ru_1'(r))'\,dr \\ &= -[\cancelto{0}{u_2(1)}1u_1'(1) - u_2(0)0u_1'(0)] + \int_0^1 ru_2'u_1'\,dr, \\ \implies \langle u_1, \mathcal{L}u_2\rangle_r &= \int_0^1 ru_2'(r)u_1'(r)\,dr = \langle \mathcal{L}u_1, u_2\rangle_r.\end{aligned}$$

Hence $\mathcal{L}$ is self-adjoint with respect to $\langle\cdot,\cdot\rangle_r$. Note we did not need $u_1(0) = u_2(0) = 0$ as a boundary condition, since the weight function took care of that in the boundary term in $u_1(0)0u_2'(0)$. It also physically does not make sense to impose $u_1(0) = u_2(0) = 0$ as the centre of the disk is not part of the boundary.

Step 2. We have to find the eigenfunctions by solving

$$\begin{aligned}-\frac{1}{r}\frac{d}{dr}\left(r\frac{du}{dr}\right) &= \lambda u \\ \implies ru'' + u' + \lambda ru &= 0.\end{aligned}$$

We use the method of power series to solve this. Since $r = 0$ is singular, we try

$$u(r) = \sum_{n=0}^{\infty} a_n r^{n+\sigma}.$$

We will not provide the details here, but you should get $\sigma = 0, 0$ repeated root, so putting in $\sigma = 0$ and recursively finding a_n, we get the Bessel function of the zeroth order. Normalising gives:

$$u_n(r) = \frac{\sqrt{2}}{J_0'(j_n)} J_0(j_n r),$$

for $n = 1, 2, \ldots$, with eigenvalues $\lambda_n = j_n^2$, where

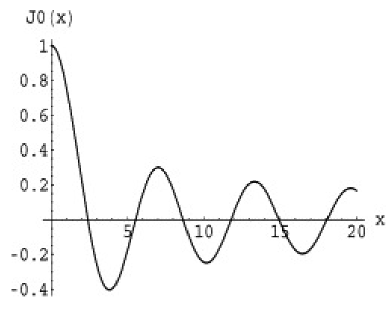
- J_0 is a function that cannot be written as elementary functions. It is called the Bessel function of order 0:

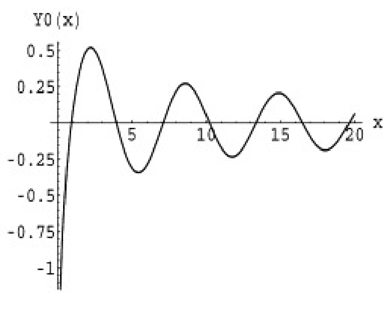
$$J_0(r) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^{2n}n!^2} r^{2n}.$$

  This is the function for which if you initialise as heat, the nature of its shape will stay the same (and the amplitude decays exponentially).
- j_n is the nth zero of J_0.
- The factor in front comes from normalisation (remember to use the weighted norm):

$$\int_0^1 rJ_0(j_n r)^2\,dr = \frac{1}{2}J_0'(j_n)^2.$$

- As this was a second-order ODE, we expect two linearly independent solutions. The other one, called Y_0, was rejected because it is not defined at $r = 0$.





Step 3. Putting it all together, we have

$$\begin{aligned} u(r) &= \int_0^a \left(w(\rho) \sum_{n=1}^{\infty} \frac{u_n(r)u_n(\rho)}{\lambda_n} \right) f(\rho)\, d\rho \\ &= \int_0^a \left(\rho \sum_{n=1}^{\infty} \frac{2}{j_n^2 J_0'(j_n)^2} J_0(j_n r) J_0(j_n \rho) \right) f(\rho)\, d\rho. \end{aligned}$$

This is the Sturm-Liouville solution for radially symmetric solutions to the heat equation: given a radially symmetric heat source f, we have solved for the temperature u.

7 Sturm-Liouville Form

Student: Okay, for the heat example and the Bessel example you found the weight function by inspection, but how do you do it in general?
Professor: A general second-order differential operator looks like

$$\mathcal{L} = \alpha(x) \frac{d^2}{dx^2} + \beta(x) \frac{d}{dx} + \gamma(x).$$

We will now show that

$$\boxed{w(x) = \exp \int^x \frac{\beta - \alpha'}{\alpha}\, dx}$$

will work.

Proof. We have

$$\begin{aligned} w\mathcal{L}f &= w\alpha f'' + w\beta f' + w\gamma f \\ &= (w\alpha f')' - w'\alpha f' - w\alpha' f' + w\beta f' + w\gamma f \\ &= (w\alpha f')' + (w\beta - w'\alpha - w\alpha')f' + w\gamma f. \end{aligned}$$

It would be nice if the f' term vanishes, so we require

$$w'\alpha = w(\beta - \alpha').$$

This is a separable equation:

$$\frac{dw}{w} = \frac{\beta - \alpha'}{\alpha} \implies w = \exp \int^x \frac{\beta - \alpha'}{\alpha}\, dx.$$

Now

$$w\mathcal{L}f = (w\alpha f')' + w\gamma f.$$

This is a very nice form, called the *Sturm-Liouville form*, which is (clearly, in the perspective of the professor) self-adjoint. For the student, we can check for self-adjointness in more detail: for any f, g with Dirichlet boundary conditions ($f(a) = f(b) = g(a) = g(b) = 0$):

$$\begin{aligned} (\mathcal{L}f, g)_w &= \int_a^b w\mathcal{L}f g\, dx \\ &= \int_a^b (w\alpha f')' g\, dx + \int_a^b w\gamma f g\, dx \end{aligned}$$

$$= \cancelto{0}{[w\alpha f'g]_a^b} - \int_a^b w\alpha f'g'\,dx + \int_a^b w\gamma fg\,dx.$$

Here the boundary term disappears because $g(a) = g(b) = 0$. Similarly,

$$\begin{aligned}(f, \mathcal{L}g)_w &= \int_a^b fw\mathcal{L}g\,dx \\ &= \int_a^b f(w\alpha g')'\,dx + \int_a^b w\gamma fg\,dx \\ &= \cancelto{0}{[fw\alpha g']_a^b} - \int_a^b w\alpha f'g'\,dx + \int_a^b w\gamma fg\,dx.\end{aligned}$$

This time we used $f(a) = f(b) = 0$. Comparing the two results, we have

$$(\mathcal{L}f, g)_w = (f, \mathcal{L}g)_w,$$

thus demonstrating that $\mathcal{L}$ is self-adjoint with respect to $(\cdot, \cdot)_w$. □

Remark 7.1. The Bessel function is associated to the linear operator

$$\mathcal{L} = -\frac{1}{r}\frac{d}{dr}\left(r\frac{du}{dr}\right) = -\frac{d^2u}{dr^2} - \frac{1}{r}\frac{du}{dr},$$

so $\alpha = -1, \beta = -\frac{1}{r}$, so the formula gives

$$w = \exp\int^r \frac{-\frac{1}{r} - 0}{-1}\,dr = \exp\ln r = r.$$

Remark 7.2. Let us define Sturm-Liouville form more officially. A second-order differential operator is in *Sturm-Liouville form* if

$$\mathcal{L}f = -(pf')' + qf,$$

for some functions $p(x), q(x)$. Note the minus sign which was absent in the proof above. It is introduced by convention/convenience, but does not affect whether the operator is in Sturm-Liouville form or not, since the minus sign can be readily absorbed by p.

Remark 7.3. Many differential equations already come in Sturm-Liouville form, e.g. Legendre's differential equation

$$\mathcal{L} = -((1 - x^2)P'(x))'.$$

This corresponds to

$$\begin{aligned}p(x) &= 1 - x^2 \\ q(x) &= 0.\end{aligned}$$

In this case, the weight function will automatically be 1, because $\mathcal{L}$ is already self-adjoint.

Remark 7.4. In the case of the Hermite differential equation,

$$\mathcal{L}f = -f'' + 2xf',$$

the weight function is not 1. The weight is

$$w(x) = \exp\int^x \frac{\beta - \alpha'}{\alpha}\,dx = \exp\int^x \frac{2x - 0}{-1}\,dx = e^{x^2}.$$

However, we don't like memorising random formulae, so we usually derive it again. It is traditional to derive the weight function by requiring f to be an eigenfunction:

$$-f'' + 2xf' = \lambda f.$$

Now LHS resembles an integrating factor problem, with integrating factor

$$\mu = e^{\int^x 2x} = e^{x^2}.$$

Hence we multiply both sides by the integrating factor to get

$$-e^{x^2} f'' + 2xe^{x^2} f' = -(e^{x^2} f')' = e^{x^2} \lambda f.$$

8 Other Boundary Conditions

So far we have only looked at Dirichlet boundary conditions. These were sufficient for $\mathcal{L}$ to be self-adjoint, as the boundary term $[\cdots]_a^b$ vanishes when we do integration by parts. Recall how this worked:

$$\begin{aligned} (\mathcal{L}f, g)_w &= \int_a^b w\mathcal{L}fg\,dx \\ &= \int_a^b (w\alpha f')'g\,dx + \int_a^b w\gamma fg\,dx \\ &= \cancelto{0}{[w\alpha f'g]_a^b} - \int_a^b w\alpha f'g'\,dx + \int_a^b w\gamma fg\,dx. \end{aligned}$$

But the boundary term can vanish for other reasons. For example, if we had Neumann boundary conditions instead, i.e. $f'(a) = f'(b) = g'(a) = g'(b) = 0$, then the term would vanish too.