Recap of Lecture 5

Classification of boundary conditions
- Dirichlet
- Neumann
- Mixed

Adjoint operator, self-adjoint operator

Sturm-Liouville problem in self-adjoint form:

\[
\begin{cases}
L(u) = -\lambda r(x)u, & x \in (a, b) \\
\alpha u(a) + \beta u'(a) = 0 \\
\gamma u(b) + \delta u'(b) = 0
\end{cases}
\]

with respect to the inner product

\[
\langle u, v \rangle \equiv \int_a^b \bar{u}(x)v(x)r(x)dx
\]

where the linear differential operator is

\[
L(u) \equiv \left( p(x)u_x \right)_x + q(x)u
\]

Results of Sturm-Liouville theory:

- \[
\left\langle u, \frac{1}{r(x)}L(v) \right\rangle = \left\langle \frac{1}{r(x)}L(u), v \right\rangle
\] (that is, \( \frac{1}{r(x)}L(i) \) is self-adjoint).
- All eigenvalues are real.
- For each eigenvalue, we can find a real eigenfunction.
- Eigenfunctions for different eigenvalues are orthogonal to each other.

Results of Sturm-Liouville theory (continued)

Since eigenvalues and eigenfunctions are both real, we only need to work with real functions.

All eigenvalues are simple.

(i.e., one corresponding eigenfunction for each eigenvalue.)
Suppose \( u(x) \) and \( v(x) \) are eigenfunctions corresponding to eigenvalue \( \lambda \).

Claim: \( u(x) \) and \( v(x) \) are proportional to each other.

Proof:

\[
L(u) = -\lambda r(x)u \\
L(v) = -\lambda r(x)v \\
\implies uL(v) - vL(u) = u(-\lambda r(x)v) - v(-\lambda r(x)u) = 0
\]

On the other hand, using the differential form of \( L(\bullet) \), we write

\[
uL(v) - vL(u) = u\left((pv_x)_x + qv\right) - v\left((pu_x)_x + qu\right)
\]

\[=
 u(pv_x)_x - v(pu_x)_x
\]

(Using Lemma 1 of Lecture 5)

\[=
 (p(uv_x - vu_x))_x
\]

Combining these two results, we have

\[
(p(uv_x - vu_x))_x = 0 \\
\implies p(uv_x - vu_x) = \text{const}
\]

Since both \( u(x) \) and \( v(x) \) satisfy the boundary conditions, we have

\[
\left.(uv' - vu')\right|_{x=b} = 0 \quad \text{(Lemma 2 of Lecture 5)}
\]

\[
\implies p(uv' - vu') = 0, \quad x \in [a, b]
\]

\[
\implies uv' - vu' = 0, \quad x \in [a, b]
\]

\[
\implies \left(\frac{v}{u}\right)' = 0, \quad x \in [a, b]
\]

\[
\implies \frac{v}{u} = \text{const}
\]

End of proof

Eigenvalue sequence

Eigenvalues form an unbounded strictly increasing infinite sequence

\[\lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots\]
$$\lim_{n \to +\infty} \lambda_n = +\infty$$

We will prove this after we describe the completeness of eigenfunctions.

### Completeness of eigenfunctions

The set of eigenfunctions \( \{v_0(x), v_1(x), v_2(x), \ldots, v_n(x), \ldots\} \) is a complete basis for all piecewise continuous functions on \([a, b]\).

That is, any piecewise continuous function \( f(x) \) can be expanded as

$$f(x) = \sum_{n=0}^{\infty} b_n v_n(x)$$

where coefficient \( b_n \) is given by

$$b_n = \frac{\int_{a}^{b} f(x)v_n(x)r(x)dx}{\int_{a}^{b} (v_n(x))^2 r(x)dx}$$

Below we do 3 things:

- prove the eigenvalue sequence.
- introduce the Rayleigh quotient, and then
- prove the completeness

#### Proof of eigenvalue sequence

Eigenvalue \( \lambda \) and eigenfunction \( u(x) \) satisfy the Sturm-Liouville equation

$$\left( p(x)u_{xx} \right)_x + q(x)u = -\lambda r(x)u, \quad x \in (a, b)$$

and boundary conditions

$$\begin{cases} \alpha u(a) + \beta u'(a) = 0 \\ \gamma u(b) + \delta u'(b) = 0 \end{cases}$$

The key step of the proof is \textbf{Prufer substitution}:

$$\begin{cases} p(x)u_x(x) = \rho(x) \cos(\theta(x)) \\ u(x) = \rho(x) \sin(\theta(x)) \end{cases}$$

Before the Prufer substitution, we have

a second order linear ODE for \( u(x) \).

After the Prufer substitution, (as we will see) we have
a first order non-linear ODE system for $\theta(x)$ and $\rho(x)$.

Let us look at the advantages of Prufer substitution.

**Advantage #1:** $|\rho(x)| \neq 0, \quad x \in [a, b]

Suppose $\rho(x_0) = 0$ at some point $x_0$ in $[a, b]$.

\[
\begin{align*}
\Rightarrow & \quad \begin{cases} p(x_0)u_x(x_0) = 0 \\ u(x_0) = 0 \end{cases} \\
\Rightarrow & \quad \begin{cases} u_x(x_0) = 0 \\ u(x_0) = 0 \end{cases}
\]

That is, $u(x)$ satisfies the zero initial conditions at $x_0$.

The Sturm-Liouville equation is a second order ODE of $u(x)$

\[
\Rightarrow \quad u(x) \equiv 0 \quad \text{(due to the zero initial conditions at } x_0)
\]

In the Sturm-Liouville problem, we are only interested in non-trivial solutions.

**Advantage #2:** Boundary conditions affect only $\theta$

Boundary condition at $x = a$:

\[
\alpha u(a) + \beta u'(a) = 0
\]

\[
\Rightarrow \quad \alpha p(a)\sin(\theta(a)) + \beta \frac{p(a)}{p(a)} \cos(\theta(a)) = 0
\]

\[
\Rightarrow \quad \alpha p(a)\sin(\theta(a)) + \beta \cos(\theta(a)) = 0
\]

\[
\Rightarrow \quad \sin(\theta(a) - \eta) = 0
\]

where $\eta$ satisfies

\[
\cos(\eta) = \frac{\alpha p(a)}{\sqrt{(\alpha p(a))^2 + \beta^2}}, \quad \sin(\eta) = \frac{-\beta}{\sqrt{(\alpha p(a))^2 + \beta^2}}
\]

The boundary condition becomes

\[
\theta(a) - \eta = n\pi
\]

\[
\Rightarrow \quad \theta(a) = \eta + n\pi
\]

We can shift the whole function $\theta(x)$ to write the boundary condition at $x = a$ as
After the shifting, we write the boundary condition at \( x = a \) as

\[
BC \text{ at } x = a : \quad \theta(a) = \eta, \quad 0 \leq \eta < \pi
\]

Note: we use shifting and pick suitable value of \( n \) to make \( 0 \leq \eta < \pi \) and \( 0 < \mu \leq \pi \), which is important in the proof of Lemma 1 and proof of eigenvalue sequence.

**Advantage #3:** The evolution equation of \( \theta(x) \) is independent of \( \rho(x) \).

\( \theta(x) \) satisfies the first order non-linear ODE

\[
\theta_x = (\lambda r(x) + q(x))\sin^2 \theta + \frac{1}{p(x)}\cos^2 \theta
\]

**Key steps in derivation:** 2 equations for \( \rho(x) \) and \( \theta(x) \):

**Equation 1:**

\[
\left(p(x)u_x\right)_x + \left(\lambda r(x) + q(x)\right)u = 0 \quad \text{(the Sturm-Liouville equation)}
\]

\[\Rightarrow \quad (\rho \cos \theta)_x + (\lambda r(x) + q(x))\rho \sin \theta = 0\]

**Equation 2:**

\[
\rho \cos \theta = p(x)u_x = p(x)(\rho \sin \theta)_x
\]

\[
\ldots
\]

(See Exercise #6 for derivation).

**Lemma 1:** Let \( \theta(x, \lambda) \) be the solution of

\[
\theta_x = (\lambda r(x) + q(x))\sin^2 \theta + \frac{1}{p(x)}\cos^2 \theta
\]

\[
\theta(a) = \eta, \quad 0 \leq \eta < \pi
\]

where \( p(x) > 0 \) and \( r(x) > 0 \) for \( x \) in \([a, b]\) (from the Sturm-Liouville problem).

Then, \( \theta(x, \lambda) \) has the 3 properties below. For \( x > a \),

a) \( \theta(x, \lambda) \) is a strictly increasing function of \( \lambda \).

b) \( \lim_{\lambda \to +\infty} \theta(x, \lambda) = +\infty \)

c) \( \lim_{\lambda \to -\infty} \theta(x, \lambda) = 0 \)

(For proof of Lemma 1, see Appendices of Lecture 6 in a separate PDF file).
Now we use the results of Lemma 1 to prove the eigenvalue sequence.

\[
\lim_{\lambda \to -\infty} \theta(b, \lambda) = 0 < \mu \quad \text{(since } 0 < \mu \leq \pi) \]

\[
\implies \mu - \pi \leq 0 < \theta(b, \lambda) < \mu \quad \text{for } \lambda \text{ negative and large.}
\]

As \( \lambda \) increases to infinity, \( \theta(b, \lambda) \) also increases to infinity.

At some value, \( \lambda_0 \), we will have \( \theta(b, \lambda_0) = \mu \).

\[
\implies \lambda_0 \text{ is the smallest value of } \lambda \text{ for which } \theta(b, \lambda_0) = \mu + n\pi \text{ is satisfies.}
\]

\[
\implies \lambda_0 \text{ is the smallest eigenvalue.}
\]

As \( \lambda \) increases further, at some value, \( \lambda_1 \), we will have \( \theta(b, \lambda_1) = \mu + \pi; \)

\[
\implies \lambda_1 \text{ is the smallest eigenvalue after } \lambda_0.
\]

At some value, \( \lambda_2 \), we will have \( \theta(b, \lambda_2) = \mu + 2\pi; \)

\[
\implies \lambda_2 \text{ is the smallest eigenvalue after } \lambda_1.
\]

At some value, \( \lambda_3 \), we will have \( \theta(b, \lambda_3) = \mu + 3\pi; \)

\[
\implies \lambda_3 \text{ is the smallest eigenvalue after } \lambda_2.
\]

\[
\vdots
\]

Thus, we have a strictly increasing sequence of eigenvalues

\[
\lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots
\]

Next we prove \( \lim_{n \to +\infty} \lambda_n = +\infty \).

Since \( \{ \lambda_n \} \) increases monotonically, we have either \( \lim_{n \to +\infty} \lambda_n = +\infty \) or \( \lim_{n \to +\infty} \lambda_n = \text{finite} \).

Suppose \( \lim_{n \to +\infty} \lambda_n = \lambda^{(c)} = \text{finite} \). We have

\[
\lambda^{(c)} > \lambda_n \quad \text{for any } n
\]

\[
\implies \theta(b, \lambda^{(c)}) > \theta(b, \lambda_n) = \mu + n\pi \quad \text{for any } n
\]

\[
\implies \theta(b, \lambda^{(c)}) = +\infty, \quad \text{which contradicts with } \lambda^{(c)} = \text{finite}.
\]

End of proof of eigenvalue sequence

Rayleigh quotient

Definition:
Rayleigh quotient is defined as
\[
R(u) = \frac{-\left\langle u, \frac{1}{r(x)} L(u) \right\rangle}{\left\langle u, u \right\rangle} = \frac{-\int_a^b u L(u) dx}{\int_a^b u^2 r(x) dx}
\]

Rayleigh quotient on eigenfunctions:

Suppose
\[
\lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots
\]
is the eigenvalue sequence and
\[
\{v_0(x), v_1(x), v_2(x), \ldots, v_n(x), \ldots\}
\]
is the sequence of corresponding eigenfunction.

It is straightforward to verify that Rayleigh quotient satisfies

- \[R(v_n) = \lambda_n, \quad n = 0, 1, 2, \ldots\]
- \[R \left( \sum_{n=0}^{N} b_n v_n \right) = \frac{\sum_{n=0}^{N} \lambda_n b_n^2 \left\langle v_n, v_n \right\rangle}{\sum_{n=0}^{N} b_n^2 \left\langle v_n, v_n \right\rangle}\]
- \[\lambda_0 = \min_{\{b_n\}} R \left( \sum_{n=0}^{N} b_n v_n \right)\]

(See Exercise #6 for derivation).

This last result motivates the Rayleigh principle below.

**Rayleigh principle:**

Let \[C_{bc}^2 = \left\{ u(x) \middle| u \in C^2 \text{ and satisfies } \alpha u(a) + \beta u'(a) = 0, \gamma u(b) + \delta u'(b) = 0 \right\}\]

Consider the minimization of \(R(u)\) over \(C_{bc}^2\).

Let \[u_0 = \arg \min_{u \in C_{bc}^2, u \neq 0} R(u)\]

**Claim 1:**

\(u_0\) is an eigenfunction corresponding to eigenvalue \(\lambda_0\) and
\[\lambda_0 = \min_{u \in C_{bc}^2, u \neq 0} R(u)\]

**Proof of Claim 1:**

Consider a function of a scalar variable \(s\)
\[ g(s) = R(u_0 + s u) \quad \text{where } u \in C^2_{bc} \]

By definition of \( u_0 \), we have \( g'(0) = 0 \) for any \( u \in C^2_{bc} \).

The derivatives of numerator and denominator of \( R(u_0 + s u) \) are

\[
\frac{d}{ds} \left( u_0 + s u, u_0 + s u \right) \bigg|_{s=0} = 2 \langle u, u_0 \rangle
\]

\[
\frac{d}{ds} \left( u_0 + s u, \frac{1}{r(x)} L(u_0 + s u) \right) \bigg|_{s=0} = 2 \left( u, \frac{1}{r(x)} L(u_0) \right)
\]

(See Exercise #6 for derivation).

The derivative of \( R(u_0 + s u) \) has the expression

\[
\frac{d}{ds} R(u_0 + s u) \bigg|_{s=0} = -2 \left( u, \frac{1}{r(x)} L(u_0) + R(u_0) u_0 \right)
\]

\[
= -2 \langle u_0, u_0 \rangle \left( u, \frac{1}{r(x)} L(u_0) + R(u_0) u_0 \right)
\]

(See Exercise #6 for derivation).

The derivative of \( R(u_0 + s u) \) is zero for any \( u \in C^2_{bc} \).

\[
\frac{d}{ds} R(u_0 + s u) \bigg|_{s=0} = 0 \quad \text{for any } u \in C^2_{bc}
\]

\[
\Rightarrow \quad \left( u, \frac{1}{r(x)} L(u_0) + R(u_0) u_0 \right) = 0 \quad \text{for any } u \in C^2_{bc}
\]

\[
\Rightarrow \quad \frac{1}{r(x)} L(u_0) + R(u_0) u_0 = 0
\]

\[
\Rightarrow \quad L(u_0) = -R(u_0) r(x) u_0
\]

\[
\Rightarrow \quad R(u_0) \text{ is an eigenvalue and } u_0 \text{ is a corresponding eigenfunction.}
\]

Next we show \( R(u_0) = \lambda_0 \).

By definition of \( u_0 \), we have

\[
R(u_0) = \min_{u \in C^2_{ac}} R(u)
\]

\[
\Rightarrow \quad R(u_0) \leq R(v_0) = \lambda_0 \quad (v_0 \text{ is an eigenfunction corresponding to } \lambda_0).
\]

On the other hand, \( R(u_0) \) is an eigenvalue and \( \lambda_0 \) is the smallest eigenvalue. Thus, we have
Combining these two results, we conclude

\[ \lambda_0 = R(u_0) = \min_{u \in C^2_{sc}} R(u) \]

End of proof of Claim 1

Next we look at the minimum of Rayleigh quotient over subset of \( C^2_{bc} \).

Let \( W_N = \text{span}\{v_0, \ldots, v_N\} \)

\[ W_N^\perp = \{u | u \in C^2_{bc} \text{ and } u \perp W_N\} \]

Consider the minimization of \( R(u) \) over \( W_N^\perp \).

Let \( u_{N+1} = \arg \min_{u \in W_N^\perp, u \neq 0} R(u) \)

Claim 2:

\( u_{N+1} \) is an eigenfunction corresponding to eigenvalue \( \lambda_{N+1} \) and

\[ \lambda_{N+1} = \min_{u \in W_N^\perp, u \neq 0} R(u) \]

(For proof of Claim 2, see Appendices of Lecture 6 in a separate PDF file)

Rayleigh principle provides a way of approximating the lowest few eigenvalues, especially the smallest eigenvalue.

Example:

\[
\begin{cases}
  u'' = -\lambda u \\
  u(0) = 0, \quad u(1) = 0
\end{cases}
\]

This is a Sturm–Liouville problem with

\[ p(x) = 1, \quad r(x) = 1, \quad q(x) = 0, \quad L(u) = u'' \]

We solved it previously. The eigenvalues are

\[ \lambda_n = (n + 1)^2 \pi^2, \quad n = 0, 1, 2, \ldots \]

Now we use Rayleigh principle to approximate the smallest eigenvalue.

(Skip the details in Lecture)

We try

\[ w(x) = x(1 - x) \]

(it has to satisfy the boundary conditions)
The numerator and denominator of Rayleigh quotient are

\[
- \left< w, \frac{1}{r(x)} L(w) \right> = - \int_0^1 w(x) w''(x) \, dx = 2 \int_0^1 x(1-x) \, dx = \frac{1}{3}
\]

\[
\left< w, \ w \right> = \int_0^1 (w(x))^2 \, dx = \int_0^1 x^2(1-x)^2 \, dx = \frac{1}{30}
\]

\[
\Rightarrow R(w) = \frac{- \left< w, \frac{1}{r(x)} L(w) \right>}{\left< w, \ w \right>} = \frac{1}{3} = 10
\]

\[
\Rightarrow \lambda_0 = \min_{u \in \mathcal{C}_2^2} R(u) \leq R(w) = 10
\]

The true value is \( \lambda_0 = \pi^2 \approx 9.87 \).

**Proof of the completeness of eigenfunctions**

For a piecewise continuous function \( f(x) \), we show that

\[
\lim_{N \to \infty} \left\| f - \sum_{n=0}^N b_n v_n \right\| = 0
\]

where

\[
b_n = \frac{\left< f, v_n \right>}{\left< v_n, v_n \right>}
\]

\[
\left< u, v \right> = \int_a^b u(x) v(x) r(x) \, dx
\]

\[
\| u \| = \sqrt{\left< u, u \right>}
\]

The proof consists of results 1-4 below.

**Result 1:**

\[
\left( f - \sum_{n=0}^N b_n v_n \right) \perp v_j \quad \text{for } 0 \leq j \leq N
\]

This follows directly from the definition of \( b_n \).

**Result 2:**

\[
\left\| f - \sum_{n=0}^N b_n v_n \right\| = \min_{\{c_n\}} \left\| f - \sum_{n=0}^N c_n v_n \right\|
\]
This follows directly from Result 1 above

\[
\left\langle \left( f - \sum_{n=0}^{N} b_n v_n \right) + \sum_{n=0}^{N} \beta_n v_n, \left( f - \sum_{n=0}^{N} b_n v_n \right) + \sum_{n=0}^{N} \beta_n v_n \right\rangle \\
= \left\langle \left( f - \sum_{n=0}^{N} b_n v_n \right), \left( f - \sum_{n=0}^{N} b_n v_n \right) \right\rangle + \left\langle \sum_{n=0}^{N} \beta_n v_n, \sum_{n=0}^{N} \beta_n v_n \right\rangle
\]

\[\Rightarrow \| f - \sum_{n=0}^{N} c_n v_n \| \geq \| f - \sum_{n=0}^{N} b_n v_n \|\]

Result 3:

We only need to show that for a piecewise continuous function \( f(x) \), we have

for any \( \varepsilon > 0 \) there exists \( N \) and \( \{c_n\} \) such that

\[\| f - \sum_{n=0}^{N} c_n v_n \| < \varepsilon \quad (***)
\]

Result 4:

Since \( C^2_{\text{bc}} \) is dense in the set of piecewise continuous functions, we only need to show that for \( f(x) \in C^2_{\text{bc}} \), we have

for any \( \varepsilon > 0 \) there exists \( N \) and \( \{c_n\} \) such that

\[\| f - \sum_{n=0}^{N} c_n v_n \| < \varepsilon \quad (***)
\]

For proof of (***), see Appendices of Lecture 6 in a separate PDF file.