STURM-LIOUVILLE THEORY, VARIATIONAL APPROACH

XIAO-BIAO LIN

1. Quadratic functional and the Euler-Jacobi Equation

The purpose of this note is to study the Sturm-Liouville problem. We use the variational problem as a tool – minimizing the functional is not the goal of this note.

Consider the quadratic functional

$$K(y) = \int_a^b [P(x)y^2 + R(x)y'^2]dx. \tag{1}$$

The Euler equation is the well-known Jacobi equation

$$Py - \frac{d}{dx}(Ry') = 0. \tag{2}$$

Consider the isoperimetric problem: find the stationary solution for K with the constrain

$$\int_a^b y^2 dx = 1. \tag{3}$$

The method of multiplier leads to the problem of finding the stationary solution for

$$\int_a^b (Ry'^2 + Py^2 - \lambda y^2)dx.$$

The corresponding Euler equation is

$$Py - \frac{d}{dx}(Ry') = \lambda y. \tag{4}$$

This is so called the Sturm-Liouville equation.

We will only consider the simplest boundary conditions

$$y(a) = y(b) = 0. \tag{5}$$

Assume that $R(x)$ and $P(x)$ are C^1 functions, and $R(x) > 0$ for $a \le x \le b$.

Introducing the notation

$$L(y) = Py - \frac{d}{dx}(Ry'). \tag{6}$$

Jacobi equation (2) and S-L equation (4) become

$$L(y) = 0, \quad L(y) = \lambda y.$$

Usually $L(y)$ is called the S-L operator. It is linear – for any C^2 functions y_1, y_2 and any constant α,

$$L(y_1 + y_2) = L(y_1) + L(y_2),$$
$$L(\alpha y) = \alpha L(y).$$

Quadratic and bilinear functionals

Define

$$K(y, z) = \int_a^b (Ry'z' + Pyz)dx. \tag{7}$$

$K(y, z)$ is bilinear with respect to $y = y(x)$ and $z = z(x)$. When $y = z$, we have

$$K(y, y) = K(y).$$

Observe that

$$K(a_1 y + a_2 z) = a_1^2 K(y) + 2a_1 a_2 K(y, z) + a_2^2 K(z),$$
$$K(\sum_{i=1}^{n} a_i y_i) = \sum_{i=1}^{n} a_i^2 K(y_i) + 2 \sum_{j>i} K(y_i, y_j).$$

Using the boundary condition (5),

$$\int_a^b Ry'z'dx = -\int_a^b (\frac{d}{dx}Ry')zdx = -\int_a^b (\frac{d}{dx}Rz')ydx.$$

Thus, (7) can be written as

$$K(y, z) = \int_a^b L(y)zdx, \tag{8}$$

where $L(y)$ is from (6). Obviously we also have

$$K(y, z) = \int_a^b L(z)ydx.$$

Therefore

$$\int_a^b L(y)z dx = \int_a^b L(z)y dx. \tag{9}$$

In particular, if $y = z$,

$$K(y) = \int_a^b L(y)y dx. \tag{10}$$

Self-adjoint operators Assume that a linear operator $Ay(x)$ is defined on the function space $\{y(x) : a \leq x \leq b\}$. If for any functions $y(x), z(x)$ in the space, we have

$$\int_z^b (Ay)z dx = \int_a^b (Az)y dx,$$

then Ay is called a self-adjoint operator. Equation (9) shows that the S-L operator is self-adjoint.

Similarly, we can show that in the space $y \in C^1$, $L(y)$ is self-adjoint under the conditions

$$y'(a) = 0, \quad y'(b) = 0.$$

Also, the operator $L(y)$ is self-adjoint under the periodic boundary condition with period $\omega = b - a$:

$$y(b) = y(a), \quad y'(b) = y'(a).$$

Orthogonality: Two functions $y_1(x)$ and $y_2(x)$ are orthogonal on $[a, b]$ with respect to the weight $\rho(x)$ if

$$\int_a^b \rho(x)y_1(x)y_2(x)dx = 0.$$

The function $y_1(x)$ is said to be normalized with respect to the weight $\rho(x)$ if

$$\int_a^b \rho(x)y_1^2(x)dx = 1.$$

We assume that $\rho(x) \geq 0$ and is not identically zero.

A sequence of functions y_i is said to be orthonormal with respect to $\rho(x)$ if

$$\int_a^b \rho(x)y_i(x)y_j(x)dx = \begin{cases} 0 & i \neq j, \\ 1 & i = j. \end{cases}$$

2. Eigenvalues and eigenfunctions

For any real or complex λ, equation (4) under condition (5) has a trivial solution

$$y(x) \equiv 0.$$

But for some λ, the system may have nontrivial solution $y \neq 0$. Those λ are called eigenvalues of the operator L, and the corresponding nontrivial functions are called eigenfunctions of $L(y)$. The eigenfunction is said to be normalized if

$$\int_a^b y^2(x)dx = 1.$$

Basic properties of eigenvalues and eigenfunctions

1. If $y(x)$ is an eigenfunction with $y(a) = y(b) = 0$, then

$$y'(a) \neq 0, \quad y'(b) \neq 0.$$

2. If y_1, y_2 are two eigenfunctions corresponding to the same eigenvalue, then the linear combination

$$y(x) = c_1 y_1(x) + c_2 y_2(x),$$

is either an eigenfunction or a trivial solution.

3. If y_1 and y_2 are eigenfunctions corresponding to the same eigenvalue λ, then y_1 and y_2 are linearly dependent. Moreover

$$y_2(x)y_1'(a) - y_1(x)y_2'(a) \equiv 0.$$

4. There are only two normalized eigenfunctions for the same eigenvalue λ. The two differ by a multiple of -1.

Theorem 1. *If y_1 and y_2 are eigenfunctions corresponding to two distinct eigenvalues $\lambda_1 \neq \lambda_2$, then y_1 and y_2 are orthogonal to each other*

$$\int_a^b y_1(x)y_2(x)dx = 0.$$

Theorem 2. *All the eigenvalues for L are real.*

Theorem 3. *If λ is an eigenvalue for L and y is a normalized eigenfunction, then*

$$K[y] = \lambda.$$

Theorem 4. *If $y(x)$ is an eigenfunction and if y_1 is orthogonal to y, then*

$$K(y, y_1) = 0.$$

We now consider the minimization of $K(y)$ under the condition

$$\int_1^b y^2(x)dx = 1, \quad y(a) = y(b) = 0.$$

It can be proved that there exists a C^1 function $y = y_1$ which solves the minimization problem. At some $\lambda = \lambda_1$, this function y_1 satisfies the Euler equation (4). Therefore, for the S-L equation there exist at least one eigenvalue λ_1 with corresponding y_1. From Theorem 3,

$$K(y_1) = \lambda_1.$$

By definition, under conditions (3), (5), λ_1 is the minimum of K with the function y_1. By the same theorem, any other eigenvalue is also a value of $K(y)$ with the corresponding normalized eigenfunction y. However, they are not the minimum of K. Thus, λ_1 is the smallest eigenvalue. The fact is summarized in the following

Theorem 5. *The smallest eigenvalue λ_1 is the conditional minimum of K under conditions (3) and (5).*

Theorem 6. *If λ_1 is the smallest eigenvalue with eigenfunction y_1, then for any C^1 function $y(x)$ satisfying (3) and (5),*

$$K(y) \geq \lambda_1 \int_a^b y^2(x)dx.$$

The equal sign holds iff $y(x) = \pm k y_1$.

3. Variational method on eigenvalues

Theorem 5 provides a method of getting the smallest eigenvalue λ_1. Other eigenvalues can also be obtained by a conditional minimization process.

Theorem 7. *Eigenvalues of L can be arranged as a increasing infinite sequence*

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n < \ldots,$$

with corresponding eigenfunctions $y_1, y_2, \dots$. For each n, the eigenvalue λ_n is the conditional minimum of $K(y)$ under the conditions

$$\int_a^b y^2 dx = 1, \int_a^b y y_i dx = 0,\ i = 1, 2, \dots, n-1, \qquad (11)$$
$$y(a) = y(b) = 0.$$

Theorem 8. *For any C^1 function y that is orthogonal to the first $n-1$ eigenfunctions and satisfies (5), we have*

$$K(y) \geq \lambda_n \int_a^b y^2 dx. \qquad (12)$$

The equal sign happens iff $y = ky_n$.

Theorem 9. *If the coefficients $P(x)$ and $R(x)$ increase by positive $\delta P(x)$ and $\delta R(x)$, then the nth eigenvalue λ_n, $n = 1, 2, 3, \dots$ increases.*

Theorem 10. *(Courant) The nth eigenvalue, denoted $\lambda_n(b)$ is a monotone decreasing function of the right boundary b. more over*

$$\lambda_n(b) \to \infty, \text{ as } b \to a.$$

Theorem 11. *(Oscillation theorem) The eigenfunction $y_n(x)$ corresponding to the nth eigenvalue $\lambda_n(b)$ has $n-1$ zeros in (a, b).*

4. Completeness of the eigenfunctions

Let $\{y_n\}$ be the orthonormal set of eigenfunctions corresponding to eigenvalues

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n < \dots.$$

Then

$$\int_a^b (\sum_{i=1}^n a_i y_i)^2 dx = \sum_{i=1}^n a_i^2.$$

For any $y \in C[a,b]$, $c_i = \int_a^b y y_i dx$ is the Fourier coefficient of y with respect to y_i, $\sum_{i=1}^\infty a_i y_i(x)$ is the Fourier series for $y(x)$ and $\sum_{i=1}^n a_i y_i(x)$ is the partial sum. The remainder is defined as

$$R_n(x) := y(x) - \sum_{i=1}^n c_i y_i(x).$$

$$\int_a^b y^2 dx = \sum_{i=1}^n c_i^2 + \int_a^b R_n^2 dx.$$

(Parseval's inequality) $$\sum_{i=1}^\infty c_i^2 \le \int_a^b y^2 dx.$$

If the equal sign holds for any continuous function $y(x)$, then we say $\{y_n(x)\}$ is complete.

Theorem 12. *Eigenfunctions of the S-L equation form a complete orthonormal basis.*

Proof. Consider $K(y) = \int_a^b (Py^2 + Ry'^2) dx$ and its first n eigenfunctions: $y_1, y_2, \dots, y_n$. Let $y(x) = \sum_1^n c_i y_i + R_n$.

$$K(y) = K(\sum_1^n c_i y_i + R_n) = K(R_n) + \sum_1^n c_i^2 K(y_i)$$
$$+ 2\sum_1^n c_i K(y_i, R_n) + 2\sum_{i \ne j} c_i c_j K(y_i, y_j).$$

Based on $K(y_i) = \lambda_i$ and the orthogonality,

$$K(y_i, R_n) = 0, \quad K(y_i, y_j) = 0, \text{ if } i \ne j.$$
$$K(y) = \sum_1^n \lambda_i c_i^2 + K(R_n).$$

Since R_n is orthogonal to $y_1, \dots, y_n$, from Theorem 8,

$$K(R_n) \ge \lambda_{n+1} \int_a^b R_n^2 dx.$$

$\lambda_i > 0$ starting from some index $i > i_0$, thus,

$$K(R_n) = K(y) - \sum_1^n \lambda_i c_i^2,$$

decreases if $n \ge i_0$, hence bounded with respect to n.

$$\int_a^b R_n^2 dx \le \frac{1}{\lambda_{n+1}} (K(y) - \sum_1^n \lambda_i c_i^2).$$

Since $\lambda_{n+1} \to \infty$, we have

$$\lim_{n\to\infty} \int_a^b R_n^2 dx = 0.$$

□