# Sturm-Liouville operators

February 19, 2021

In a previous lecture, we discussed complete orthogonal systems. In infinite dimensional spaces they play a role similar to the role of bases in finite-dimensional spaces.

Now we address the question how to construct them. Two main methods exist.

**1. First method. Orthogonalization.** Suppose that we have a sequence of linearly independent vectors $(v_n)$. The process of (Gram–Schmidt) orthogonalization produces a new sequence $(u_n)$, such that each $u_n$ is a linear combination of the first $n$ vectors $v_k$, and $u_n$ are orthogonal. The algorithm works as follows.

Step 1. Set $u_1 = v_1$.

Step 2. Set $u_2 = v_2 - cu_1$. To make it orthogonal to $u_1$, we write

$$0 = (u_2, u_1) = (v_2, u_1) - c(u_1, u_1),$$

so we choose $c = (v_2, v_1)/(u_1, u_1)$.

Then continue in the similar way:

Step $n$. Set

$$u_n = v_n - \sum_{k=1}^{n-1} c_{n,k} u_k,$$

where $u_k$ for $1 \leq k \leq n-1$ are already constructed in the previous steps. Since $u_1, \ldots, u_{n-1}$ are orthogonal, we choose

$$c_{n,k} = \frac{(v_n, u_k)}{(u_k, u_k)}, \quad 1 \leq k \leq n-1.$$

With this choice, $u_n$ will be orthogonal to all $u_k$ with $k \leq n-1$.

Example. Apply this procedure to the system $v_n(x) = x^n$ on the interval $(-1, 1)$. Construct $u_n$ for $n \leq 3$.

**2. Second method. Spectral Theorem for Hermitian operators.**

I recall that a linear operator on a vector space $V$ is a map $L : V \to V$ which satisfies

$$L(c_1x + c_2y) = c_1L(x) + c_2L(y), \quad \text{for all vectors} \quad x, y \in V,$$

and all numbers $c_1, c_2$. On the space $\mathbf{C}^n$ all linear operators have the form $L(x) = Ax$, where $A$ is an $n \times n$ matrix.

If for some vector $v$ and a number $\lambda$ we have

$$L(v) = \lambda c, \quad \text{and} \quad v \neq 0,$$

then we say that $\lambda$ is an *eigenvalue* of $L$, and $v$ is an *eigenvector* corresponding to this eigenvalue.

It is useful to have a basis of the space consisting of eigenvectors of a given operator $L$. Once this is done, the action of $L$ on any vector becomes completely clear:

Suppose that $v_1, \ldots, v_n$ are eigenvectors of $L$ with eigenvalues $\lambda_1, \lambda_n$, and $v_1, \ldots, v_n$ form a basis. Then every vector $x$ can be expanded

$$x = c_1v_1 + c_2v_2 + \ldots + c_nv_n$$

and we have

$$L(x) = c_1\lambda_1v_1 + \ldots + c_n\lambda_nv_n.$$

Unfortunately such a basis cannot be found for all linear operators.

The important class of operators for which it exists consists of *Hermitian operators*. To define them, I recall the notion of adjoint operator. We say that $L^*$ is adjoint to $L$ if

$$(L(x), y) = (x, L^*(y)) \quad \text{for all} \quad x, y.$$

If $L$ is represented by a matrix $A$, and the dot product is the standard one, then $L^*$ is represented by *Hermitian transpose* which we denote $A^*$. It is obtained from $A$ by transposition and complex conjugation.

Indeed, if $\mathbf{C}^n$ is the space of column vectors then the standard dot product is

$$(x, y) = x^T\overline{y} = x_1\overline{y_1} + \ldots + x_n\overline{y_n},$$

and if $L(x) = Ax$, then we have

$$(L(x), y) = (Ax)^T \overline{y} = x^T A^T \overline{y} = x^T \overline{(\overline{A^T} y)} == x^T (A^* y) = (x, L^*(y)).$$

An operator $L$, or a matrix $A$ is called *self-adjoint* or *Hermitian* if

$$L = L^* \quad \text{or} \quad A^* = A.$$

For example, a real matrix is Hermitian if and only if it is symmetric. Then we have the following fundamental theorem:

**Spectral theorem for Hermitian operators in a finite-dimensional space.** *Let $L$ be an Hermitian operator. Then:*
*a) All eigenvalues are real.*
*b) Eigenvectors with distinct eigenvalues are orthogonal.*
*c) There is an orthogonal basis of the space consisting of eigenvectors.*

Let me prove a) and b) since these statements do not depend on the assumption that the space is finite-dimensional.

To prove a), suppose that $L(v) = \lambda v$. Then

$$(L(v), v) = (\lambda v, v) = \lambda(v, v).$$

On the other hand

$$(L(v), v) = (v, L^*(v)) = (v, \lambda v) = \overline{\lambda}(v, v).$$

Since $v \neq 0$, we have $(v, v) \neq 0$ and conclude that $\lambda = \overline{\lambda}$, so $\lambda$ must be real.

To prove b), suppose that $L(u) = \lambda u$ and $L(v) = \mu v$, and $\mu \neq \lambda$. Then we have

$$\lambda(u, v) = (L(u), v) = (u, L^*(v)) = (u, \mu v) = \mu(u, v),$$

where we used that $\mu$ is real, which was proved in part a). Since $\lambda \neq \mu$, this implies that $(u, v) = 0$. This proves b). So Hermitian operators can serve as a source of orthogonal systems.

We will also need a slight generalization of this theorem. Let $L$ be a Hermitian operators in $\mathbf{C}^n$ and $M$ a diagonal matrix with *positive entries*[1]. Together with the usual Hermitian product we consider another product

$$(x, y)_M = x^T M \overline{y} = (x, My) = m_{1,1} x_1 \overline{y_1} + \ldots + m_{n,n} x_n \overline{y_n}.$$

[1]Actually, one can take any *positive definite* Hermitian matrix for $M$.

Then we may consider the *generalized eigenvalue problem*

$$L(v) = \lambda M v, \quad v \neq 0. \tag{1}$$

In this generalized setting, we have a complete analog of the Spectral theorem:

a) All generalized eigenvalues are real.

b) Generalized eigenvectors with distinct eigenvalues are $M$-orthogonal, that is $(x, y)_M = 0$ for any such eigenvectors.

c) There is an $M$-orthogonal basis consisting of generalized eigenvectors.

em Exercise. Prove a) and b) by modifying the argument given above for the special case.

**3. Differential operators**. Now we discuss how this theory generalizes to infinite-dimensional spaces. We are interested in a special kind of operators, namely second order differential operators. Let us write such an operator in the form

$$L(y) = r(x)y'' + p(x)y' + q(x)y,$$

where $r$, $p$ and $q$ are some *real* continuous functions on an interval $(a, b)$, and the vector space consists of functions on an interval $[a, b]$ with the dot product

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

We want to know when such an operator is Hermitian, that is under what conditions on $p, q$, and on the space of functions, we have $(L(f), g) = (f, L(g))$. So we have to compute the adjoint.

First of all, since $q$ is real, we evidently have

$$(qf, g) = (f, qg).$$

Then, integrating by parts,

$$\begin{aligned} (pf', g) &= \int_a^b pf'\overline{g} = \int_a^b p\overline{g}df = pf\overline{g}|_a^b - \int_a^b f(p\overline{g})' \\ &= pf\overline{g}|_a^b - \int_a^b fp'\overline{g} - \int_a^b fp\overline{g}' = -(f, pg' + p'g) + pf\overline{g}|_a^b, \end{aligned}$$

where we used that $p$ is real.

And integrating by parts twice, we obtain

$$\begin{aligned}(rf'', g) &= \int_a^b rf''\overline{g} = \int_a^b r\overline{g}\, df' = rf'\overline{g}|_a^b - \int_a^b f'(r\overline{g})' \\ &= rf'\overline{g}|_a^b - \int_a^b (r\overline{g})'\, df = rf'\overline{g}|_a^b - (r\overline{g})'f|_a^b + \int_a^b f(r\overline{g})''.\end{aligned}$$

Putting this together, we obtain

$$(L(f), g)) = (f, L^*(g)) + \text{terms coming from endpoints},$$

where

$$L^*(y) = (ry)'' - py' - p'y + qy = ry'' + (2r' - p)y' + (r'' - p' + q)y.$$

This is called the *formal adjoint* operator to $L$ ("formal" because we ignored the non-integrated terms). Operator $L$ is called *formally self-adjoint* if $L = L^*$, that is if

$$p = r'.$$

So the general form of a formally self-adjoint operator can we written as

$$L(y) = (ry')' + qy. \tag{2}$$

*Exercise.* According to our calculation, the operator $L(y) = y'$ is not formally self-adjoint. Check that the operator $L(y) = iy'$, where $i = \sqrt{1}$ is formally self-adjoint. The finite-dimensional analog of this statement is that if $A$ is a skew-symmetric matrix, that is $A^T = -A$, then $iA$ uis Hermitian.

Now consider the non-integrated terms. If an operator is formally self-adjoint, then some of them vanish, and what remains gives the **Lagrange formula**:

*For a formally self-adjoint operator L, we have*

$$(L(f), g) = (f, L(g)) + r(f'\overline{g} - f\overline{g'})\Big|_a^b$$

To obtain a true self-adjoint operator, we have to restrict our space of functions, so that the non-integrated terms in the Lagrange formula vanish for all $f, g$ in our space. This is where the boundary conditions start playing a role.

We state two kinds of boundary conditions which ensure that the non-integrated term in Lagrange's formula vanishes.

**Separated boundary conditions:**

$$\alpha y(a) + \beta y'(a) = 0, \quad \gamma y(b) + \delta y'(b) = 0,$$

where $\alpha, \beta, \gamma, \delta$ are some given real numbers, $\alpha^2 + \beta^2 \neq 0$ and $\gamma^2 + \delta^2 \neq 0$. Suppose that $f$ and $g$ both satisfy these conditions. Then, on the left end we have

$$\begin{aligned} \alpha f(a) + \beta f'(a) &= 0 \\ \alpha \overline{g(a)} + \beta \overline{g'(a)} &= 0, \end{aligned}$$

where we used the assumption that $\alpha$ and $\beta$ are real. Considering this as a system of equations with respect to $\alpha$ and $\beta$, we conclude that its determinant must be 0, since $(\alpha, \beta) \neq (0, 0)$. So

$$f'(a)\overline{g(a)} - f(a)\overline{g'(a)} = 0$$

and this exactly means that the contribution from the left end $a$ in the non-integrated term in Lagrange's formula vanishes. Similarly, the contribution of the right end vanishes. Such boundary conditions are called *self-adjoint*.

**Periodic boundary conditions:**

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

If both $f$ and $g$ satisfy these conditions, then the non-integrated terms in Lagrange's formula cancel (contribution from $a$ cancels with the contribution from $b$).

**Anti-periodic boundary conditions:**

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

Again, if both $f$ and $g$ satisfy these conditions, then the non-integrated terms in Lagrange's formula cancel.

Now we give a precise formulation of the eigenvalue problem.

**Regular Sturm–Liouville boundary value problem.** *Let $L$ be a formally self-adjoint operator of the form (2) on a* **finite interval** $[a, b]$*, with $r(x) > 0$ for $x \in [a, b]$. Suppose that $r, r'$ and $q$ are continuous on $[a, b]$.*

*Let $w(x)$ be another strictly positive continuous function on $[a, b]$.*

*Find numbers $\lambda$ (eigenvalues) for which there exists a function $y$ on $[a, b]$ satisfying given self-adjoint boundary conditions, and $L(y) + \lambda w y = 0$.*

(Following the book, we switched the sign convention for the eigenvalue in comparison with finite dimensional case).

Notice the important assumption $r(x) > 0$ on the whole interval, including the ends. The function $w$ is called the *weight*. It is the analog of the positive diagonal matrix $M$ in (1).

The spectral theorem for this problem is the following:

**Spectral theorem for regular Sturm-Liouville problems.**

*a) All eigenvalues $\lambda$ are real.*

*b) Eigenfunctions $\phi_j$ with different eigenvalues are $w$-orthogonal, that is*

$$(\phi_j, \phi_k) := \int_a^b \phi_j(x)\overline{\phi_k(x)}w(x)dx = 0, \quad k \neq j.$$

*c) Eigenvalues form an infinite sequence $\lambda_n \to +\infty$, and the corresponding eigenfunctions form a complete orthogonal system in $L^2_w(a, b)$.*

*d) If $f$ is in $C^2[a, b]$ (twice continuously differentiable) and satisfies the boundary conditions, then the Fourier series*[2]

$$\sum_n (f, \phi_n)_w \phi_n$$

*converges to $f$ uniformly.*

*e) For separated boundary conditions, to each eigenvalue corresponds one eigenfunction (up to a constant multiple), and for periodic boundary conditions there are at most two linearly independent eigenfunctions.*

The proofs of statements a) and b) is exactly the same as in the finite-dimensional case. Statement e) follows from the basic facts about ODE. The set of *all* solutions of a linear homogeneous ODE is two-dimensional. So under any boundary conditions, the eigenspace can have dimension at most two.

If we have separated boundary conditions, one of them is of the form $\alpha f(a) + \beta f'(a) = 0$, but from the existence and uniqueness theorem for ODE we know that $f(a)$ and $f'(a)$ can be chosen arbitrarily. Thus not all solutions satisfy this boundary condition, so the subspace of those which do satisfy has dimension 1.

[2] Notice the misprint in the book in this place (p. 90, Theorem 3.10).

The most difficult part of the theorem is statement c). The whole Chapter 10 of the book is dedicated to the proof of it. Once c) is proved, the proof of d) is similar to the proof of convergence theorem of Fourier series.

Read the example on p. 91 where the operator $L(y) = y''$ with general separated boundary conditions is discussed.

If one or both ends of the interval $[a, b]$ is infinite, or if it is finite and function $r$, equals 0 at this end, the Sturm–Liouville boundary value problem is called *singular*. Unfortunately the theory for this case is much more complicated (it is not included in our textbook), but on the other hand, this case is most frequently encountered in applications.

The problem is how to state the correct boundary condition on an irregular end. The left end is called irregular if either $a = -\infty$ or $r(a) = 0$. Similarly for the right end.

The good news is that physical considerations and common sense usually help to state a correct condition. Read section 3.6 of the book.