Iranian Journal of Numerical Analysis and Optimization
Vol. 15, No. 2, 2025, pp 645–654
https://doi.org/10.22067/ijnao.2024.89575.1500
https://ijnao.um.ac.ir/

Research Article

Numerical study of the Sturm–Liouville problem

S.D. Algazin, A.A. Sinitsyn*,

Abstract

The article discusses the general Sturm–Liouville problem. To solve it numerically, a new algorithm is proposed, which is based on the variational principle and does not use saturation. The problem of constructing numerical methods for solving eigenvalue problems can be divided into two stages. First, we need to reduce the infinite-dimensional problem into a finite-dimensional one, and then find a method for solving this finite-dimensional algebraic eigenvalue problem. In this paper, we only consider the first stage, and solve the resulting algebraic problem using the QR algorithm. A comparison with the results of other authors is also carried out. Methodical calculations confirm the correctness of the new approach.

*Corresponding author

Received 26 August 2024; revised 22 September 2024; accepted 11 December 2024

Sergey Algazin
Doctor of Physical and Mathematical Sciences, Leading Researcher, Laboratory of Mechanics and Optimization of Structures, Ishlinsky Institute for Problems of Mechanics of the Russian Academy of Sciences, e-mail: algazinsd@mail.ru

Artemiy Sinitsyn
Student of the MSU Faculty of Mechanics and Mathematics, Department of Elasticity Theory, e-mail: art@sinitsyn.info

How to cite this article
Algazin, S.D. and Sinitsyn, A.A., Numerical study of the Sturm–Liouville problem. *Iran. J. Numer. Anal. Optim.*, 2025; 15(2): 645-654. https://doi.org/10.22067/ijnao.2024.89575.1500

AMS subject classifications (2020): Primary 34B24; Secondary 65Y20, 74S99.

Keywords: Sturm–Liouville problem, Numerical algorithm without saturation, Computational experiment.

1 Introduction

There are a number of competing methods for the numerical solution of the eigenvalue problem. These are, first of all, projection methods: the Ritz method, the Bubnov–Galerkin method, and so on. We know quite a lot about the accuracy provided by these methods. For example, the approximations for the eigenvalues of self-adjoint problems given by the Ritz method lie on top of the exact values. A number of convergence results are known, and in some special cases, error estimates of projection methods are obtained [10]. Along with projection methods, difference methods have also become widespread [9]. However, when designing these numerical methods, a number of important circumstances are not taken into account, which significantly reduces their effectiveness. Usually, when solving an eigenvalue problem, we have colossal a priori information. Most often, the solutions sought are infinitely differentiable or even analytical. Therefore, they are elements of functional compacts, quite simply arranged. As a rule, the asymptotics of their diameters are well known for such compacts. On the other hand, any projection method is based on the choice of a certain set of finite-dimensional subspaces and thereby some way of approximating the desired solution (and this method, as a rule, is not consistent with the optimal methods mentioned above). This naturally leads to the fact that the numerical algorithm based on such a projection method is far from optimal in its properties. At the same time, by basing the numerical algorithm on a rational way of approximating the desired element, we obtain an algorithm close to the optimal one. This approach will be developed below, and it is based on the ideas of the work [4]. The different methods have significant disadvantages [4] and, in particular, the fact that they are methods with saturation (quite a lot of works have been devoted to the accuracy of these methods, and of them we will point only to [9, 5]). Therefore, the difference method of solving the eigenvalue problem again ignores a priori information about the smoothness of the solution, and taking into account the loss of smoothness inherent in difference methods, we obtain algorithms that are far from optimal. The problem of constructing numerical methods for solving the eigenvalue problem is divided into two. First of all, you need to reduce an infinite-dimensional problem to a finite-dimensional problem, and then specify a method for solving the resulting algebraic eigenvalue problem. In this paper, only the first stage is considered, and the resulting algebraic problem is solved by the QR method.

Abstract theorems on error estimation in eigenvalue problems are published in [2, 7]. Note that in [7] only compact operators are considered, and in [2] arbitrary closed operators are considered.

A special case of the Sturm–Liouville problem was considered earlier in [3].

2 Problem statement and variational principle

Consider the eigenvalue problem (Sturm–Liouville problems):

$$-\frac{d}{dx}\left(p(x)\frac{du}{dx}\right)+r(x)u(x)-\lambda\rho(x)u(x)=0, \qquad a\le x\le b, \tag{1}$$

$$\alpha u'(a)-\beta u(a)=0, \quad \gamma u'(b)+\delta u(b)=0, \tag{2}$$

with $\alpha>0$, $\gamma>0$, $\beta\ge 0$, $\delta\ge 0$, and at least one of the coefficients β and δ is different from zero, $p(x)\ge 0, r(x)\ge 0, A=\int_a^b \frac{dx}{p(x)}<\infty,\ \rho_0\le\rho(x)\le\rho_1,\ \rho_0,\rho_1>0$. Then the problem (1)–(2) has a discrete spectrum [8].

Variational principle. Denote $Au=-\frac{d}{dx}\left(p(x)\frac{du}{dx}\right)+r(x)u(x)$.

Differential operator is defined on functions satisfying boundary conditions (2). Then

$$|u|^2=(Au,u)=\int_a^b(pu'^2+ru^2)dx+\frac{\beta}{\alpha}p(a)u^2(a)+\frac{\delta}{\gamma}p(b)u^2(b)\equiv J(u).$$

Let $\bar{u}=u+\varepsilon\eta(x)$. Then

$$J(\bar{u})=\int_a^b(p(u'+\varepsilon\eta')^2+r(u+\varepsilon\eta)^2)dx$$

$$
\begin{aligned}
&+\frac{\beta}{\alpha}p(a)(u(a)+\varepsilon\eta(a))^2+\frac{\delta}{\gamma}p(b)(u(b)+\varepsilon\eta(b))^2\\
\equiv&\Phi(\varepsilon).
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
\Phi'(0)=&\int_a^b(p(x)u'(x)(u'+\eta')+ru(u+\eta))dx+\frac{\beta}{\alpha}p(a)u(a)(u(a)+\eta(a))\\
&+\frac{\delta}{\gamma}p(b)u(b)(u(b)+\eta(b))\\
=&\int_a^b p(x)(u'(x))^2dx+\int_a^b pu'\eta'dx+\int_a^b ru^2dx\\
&+\int_a^b ru\eta dx+\frac{\beta}{\alpha}p(a)u(a)(u(a)+\eta(a))+\frac{\delta}{\gamma}p(b)u(b)(u(b)+\eta(b)).
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\int_a^b pu'\eta'dx=\int_a^b pu'd\eta=pu'\eta|_a^b-\int_a^b(pu')'\eta dx,\\
&\int_a^b pu'^2dx=\int_a^b pu'du=pu'u|_a^b-\int_a^b(pu')'udx,\\
&\int_a^b[-(pu')'+ru]udx=0.
\end{aligned}
$$

By virtue of the equation, it remains without η:

$$
\begin{aligned}
&pu'u|_a^b+\frac{\beta}{\alpha}p(a)u^2(a)+\frac{\delta}{\gamma}p(b)u^2(b)\\
&=p(u'(b)+\frac{\delta}{\gamma}u(b))u(b)+p(-u'(a)+\frac{\beta}{\alpha}u(a))u(a)=0.
\end{aligned}
$$

Due to the boundary conditions, it remains with η:

$$
\int_a^b\eta[-(pu')'+ru]dx+p(b)[u'(b)+\frac{\delta}{\gamma}u(b)]\eta(b)+p(a)[-u'(a)+\frac{\beta}{\alpha}u(a)]\eta(a)=0.
$$

Because the function η is arbitrary, then we get from the condition $J(u)\to\min$, equations and boundary conditions (1)–(2).

3 Interpolation formula

Let $u=u(y), y\in[a,b]$, and replacement $y=\frac{b-2}{2}x+\frac{a+b}{2},\ x\in[-1,+1]$.

The interpolation formula by x reads as follows:

$$
\begin{aligned}
P_{n+2}(x;u) = &\sum_{j=0}^{n+1} \frac{(x^2-1)T_n(x)}{[\dots]'_{x=x_j}(x-x_j)} u_j,\ u_j = u(y_j),\\
u_a =& u(a),\ u_b = u(b).\\
y_j =& \frac{b-2}{2}x_j + \frac{a+b}{2},\ x_j = \cos\theta_j,\ \theta_j = \frac{2j-1}{2n}\pi,\\
x_0 =& -1, \quad x_{n+1} = +1, \quad T_n(x) = \cos(n \arccos x).
\end{aligned}
$$

From here we get that because

$[\dots]'_{\ x=x_j} = 2xT_n(x) + (x^2-1)T_n'(x)\big|_{x=x_j} = (x_j^2-1)T_n'(x_j)$, then

$$
\frac{T_n(x)}{T_n'(x_j)(x-x_j)} = \sum_{k=0}^{n-1}{}' a_k^{(n)} T_k(x),\ a_k^{(n)} = \frac{2}{n}T_k(x_j) = \frac{2}{n}\cos k\theta_j.
$$

where symbol "′" of the sign of the sum indicates that the summand for $k=0$ is taken with the coefficient $\frac{1}{2}$.

Quadrature formula. Let us define the coefficients of the quadrature formula of the integral. They are the sum of the weights on the function values in the nodes as follows:

$$
\begin{aligned}
&\int_a^b f(y)dy \qquad (y = \frac{b-a}{2}x + \frac{a+b}{2}, \quad dy = \frac{b-a}{2}dx)\\
&\qquad = \frac{b-a}{3}\int_{-1}^{+1} f(x)dx\\
&\qquad = \frac{b-a}{2} - \frac{u_a}{2T_n(-1)}I_n^a + - \frac{u_b}{2T_n(1)}I_n^b + \int_{-1}^{+1}(\dots)dx,\\
&T_n(1) = 1, \qquad T_n(-1) = (-1)^n.
\end{aligned}
$$

Here $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$ which implies

$$
\int_{-1}^{+1} T_k(x)dx = \frac{(-1)+(-1)^{k+1}}{k^2-1} = \begin{cases} 0, & k\text{-odd},\\ \frac{-2}{k^2-1}, & k\text{ -even}, \end{cases}
$$

$$
I_n^a = \int_{-1}^{+1}(x-1)T_n(x)dx = \begin{cases} \frac{2}{n^2-1}, & k\text{-even},\\ \frac{-2}{n^2-4}, & k\text{ -odd}, \end{cases}
$$

$$I_n^b = \int_{-1}^{+1} (x+1) T_n(x) dx = \begin{cases} \frac{-2}{n^2-1}, & k\text{-even}, \\ \frac{1}{n^2-4}, & k \text{ -odd}, \end{cases}$$

$$I_k = \int_{-1}^{+1} (x^2-1) T_k(x) dx = \begin{cases} \frac{1}{k^2-1} - \frac{1}{2}\{\frac{1}{(k+1)(k+3)} + \frac{1}{(k-1)(k-3)}\}, & k\text{-even}, \\ 0, & k \text{ -odd}, \end{cases}$$

$$c_a = \frac{b-a}{2}\{\frac{(-1)^{n+1}}{2} I_n^a\},$$

$$c_b = \frac{b-a}{2}\{\frac{1}{2} I_n^b\},$$

$$c_j = \frac{b-a}{2}\{\frac{-2}{n} \sum_{k=0(2)}^{n-1} \frac{\cos k\theta_j}{\sin^2 \theta_j} I_k\},$$

$$I_0 = -\frac{4}{3},$$

$$\begin{aligned} \int_a^b f(y) dy =& \frac{b-a}{2} \int_{-1}^{+1} f\left(\frac{b-a}{2} x + \frac{a+b}{2}\right) dx \\ =& c_a f_a + c_b f_b + \sum_{j=1}^{n} c_j f_j, \qquad f_j = \cos\theta_j,\ j = 1, 2, \ldots, n. \end{aligned}$$

Numerical differentiation. Let us differentiate the interpolation formula by x as follows:

$$\begin{aligned} P_{n+2}(x,u) =& \frac{1}{2}(x+1) T_n(x) u_b - \frac{(-1)^n}{2}(x-1) T_n(x) u_a \\ &+ \frac{2}{n} \sum_{j=1}^{n} \frac{1-x^2}{\sin^2\theta_j} \{\sum_{k=0}^{n-1} \cos k\theta_j T_k(x)\} u_j, \end{aligned}$$

$$f(y) = f\{\frac{b-a}{2} x + \frac{a+b}{2}\} f'_x = f'_y \frac{b-a}{2} \Rightarrow f'_y = \frac{2}{b-a} f'_x,$$

$$\begin{aligned} P'_{n+2}(x,u) =& \frac{u_b}{2}\{(x+1) T_n{}'(x) + T_n(x)\} - \frac{(-1)^n}{2}\{(x-1) T_n(x) + T_n(x)\} u_a \\ &+ \frac{2}{n} \sum_{j=1}^{n} \{\frac{-1}{\sin^2\theta_j}\{\sum_{k=0}^{n-1} \{2x T_k(x) + (x^2-1) T'_k(x)\} \cos k\theta_j\} u_j, \end{aligned}$$

Let $x = \pm 1$, or let $x = \cos\theta_i,\ i = 1, 2, \ldots, n$. Then

$$\begin{aligned}P'_{n+2}(-1,u) =& \frac{u_b}{2}(-1)^n - \frac{1+2n^2}{2}u_a \\ &+ \frac{2}{n}\sum_{j=1}^{n}\{\frac{1}{\sin^2\theta_j}\{\sum_{k=0}^{n-1}\{2(-1)^k\}\cos k\theta_j\}\}hu_j,\end{aligned}$$

$$\begin{aligned}P'_{n+2}(+1,u) =& \frac{1+2n^2}{2}u_b - \frac{(-1)^n}{2}u_a \\ &+ \frac{2}{n}\sum_{j=1}^{n}\{\frac{-1}{\sin^2\theta_j}\{\sum_{k=0}^{n-1}2\cos k\theta_j\}\}u_j,\end{aligned}$$

$$\begin{aligned}P'_{n+2}(x_i,u) =& \frac{u_b}{2}\{n(x_i+1)\frac{\sin n\theta_i}{\sin\theta_i}\} - \frac{(-1)^n}{2}\{n(x_i-1)\frac{\sin n\theta_i}{\sin\theta_i}\}u_a \\ &+ \frac{2}{n}\sum_{j=1}^{n}\{\frac{-1}{\sin^2\theta_j}\{\sum_{k=0}^{n-1}\{2x_i\cos k\theta_i - k\sin\theta_i\sin k\theta_i\}\cos k\theta_j\})\}u_j.\end{aligned}$$

4 Discretization

The problem of discretizing a quadratic functional can be reduced to the problem of finding the minimum of a quadratic function as follows:

$$\int_a^b (pu'^2 + ru^2)dx + \frac{\beta}{\alpha}p(a)u^2(a) + \frac{\delta}{\gamma}p(b)u^2(b) \equiv J(u) \to \min.$$

Calculating the integral by the quadrature formula, we reduce the problem of the minimum of the quadratic functional to the problem of the minimum of the quadratic form:

$$\begin{aligned}&J(u) \to J(u_b, u_1, u_2, \ldots, u_n, u_a),\\ &\frac{\partial J(u_b, u_1, u_2, \ldots, u_n, u_a)}{\partial u_l} = 0, \quad l = b, 1, 2, \ldots, n, a.\end{aligned}$$

Calculation results. The calculations use $Q = 0, N_2 = 1026$. Then the calculations give $\lambda_0 = -0.878$; $k = 1$, $\lambda_1 = 1.000002$; $k = 72$, $\lambda_{72} = 5184.001$.

Table 1: $p = 1,\ q = 20\cos 2y + 100\sin^2 2y\,, u'(0) = 0,\ \ u'(\pi) = 0$

i	λ_i	λ_i	λ_i	λ_i
N2	34	64	134	[6]
1	$2.0 * 10^-5$	$5.2 * 10^-8$	$5.1 * 10^-8$	0.0000000
2	37.7733909214894	37.8059004212607	37.8059002738378	37.7596285
3	41.3283618161036	45.6727315410478	46.6101985290947	37.8059002
4	69.8880664248824	69.8212185270856	69.8010881122681	37.8525995
5	70.7581077855011	70.6037307664520	70.5599120331492	70.5475097
6	71.5270508988527	71.4361363622447	71.4119790617174	92.6538177
7	96.2500910443853	96.2105986913062	96.2068706709564	96.2058159
8	110.785329891999	110.714479891137	110.695004912174	102.254347

Table 2: $p = 1, q = \exp(y), u'(0) = 0,\ \ u'(\pi) = 0$

i	λ_i	λ_i	λ_i	λ_i
N2	34	66	1026	[6]
1	4.9013	4.8978	4.8966739510209	4.89571
2	5.1951	5.2204	5.22886153287197	-
3	10.053	10.047	10.0451980276767	9.99955
4	16.033	16.022	16.0192812195823	15.4685
5	23.292	23.292	23.2662960864049	21.0371
6	32.306	32.274	32.2637489027225	28.1893
7	43.284	43.236	43.2200825906882	37.7907
8	56.271	56.204	56.1816819519642	49.6137
9	71.272	71.182	71.1531142297211	63.5205
10	88.285	88.170	88.1322684247181	79.4646
11	107.30	107.16	107.116861697407	97.4279
12	128.33	128.16	128.105246952299	117.402
13	151.37	151.16	151.096313344083	139.384
14	176.41	176.17	176.089314130927	163.370
15	203.46	203.17	203.083739890564	189.359
16	232.51	232.18	232.079236175438	217.351
17	263.65	263.19	263.075551299445	247.344
18	296.87	296.21	296.072503036626	279.338
19	334.82	331.22	331.069957054295	313.334
20	373.85	368.24	368.067812678323	349.330
21	437.56	407.26	407.065993320228	387.326

Table 3: $p = 1, q = \frac{e\cos(2\pi y)}{1+e\cos(2\pi y)}, u'(0) = 0, u' = 0$

N2	34	64	[1]
e = 0.1	9.83337323117731	9.81926933696682	$\lambda_1^s = 9.81382$
e = 0.2	9.76485503817489	9.75081106707683	$\lambda_1^s = 9.74579$

5 Discussion of the results

Methodical calculations at $q = 0$ demonstrate the correctness of the methodology. Comparison with the results of [1] (see Table 3) confirms its correctness. However a comparison with the results of [6] demonstrates a discrepancy. In Table 2, in the cited work, the second eigenvalue is omitted. The remaining eigenvalues match satisfactorily only for small numbers of eigenvalues. In Table 1, only individual eigenvalues match satisfactorily.

Acknowledgements

Authors are grateful to there anonymous referees and editor for their constructive comments. The work was carried out on the topic of state assignment no. 124012500437-9.

References

[1] Akulenko, L.D. and Nesterov. S.V. *High-precision methods in eigenvalue problems and their applications*, Chapman and Hall/CRC, 2004.

[2] Algazin, S.D. *Localization of eigenvalues of closed linear operators*, Sib. math. J. 24 (2) (1983) 155–159.

[3] Algazin, S.D., Babenko, K.I. and Kosorukoye, A.L. *The numerical solution of the eigenvalue problem*, IPMatem Akad. Nauk S.S.S.R. Moscow, 1975.

[4] Babenko, K.I. *Fundamentals of numerical analysis*, Moscow, Nauka, 744, 1986 [in Russian].

[5] Hubbard, B.E. *Bounds for eigenvalues of the Sturm-Liouville problem by finite difference methods*. Arch. Ration. Mech. Anal. 10 (2) (1962) 171–179.

[6] Ghelardoni, P. *Approximations of Sturm-Liouville eigenvalues using Boundary Value Methods*, Journal of Computational and Applied Mathematics, Volume 23, Issue 3, May 1997, Pages 311-325.

[7] Mercier, B., Osborn, J., Rappaz, J. and Raviart, P.A. *Eigenvalues approximation by mixed and hybrid methods*, Math. Comp. 36 (154) (1981) 427–453.

[8] Mikhlin, S.G. *Variational methods in mathematical physics*, Pergamon Press, Oxford, England, 1964.

[9] Prikazchikov, V.G. *High-accuracy homogeneous difference schemes for the Sturm-Liouville problem*, JUSSR Comput. Math. Math. Phys. 9 (2) (1964) 76–106.

[10] Vainiko, G.M. *Asymptotic evaluations of the error of projection methods for the eigenvalue problem*, USSR Comput. Math. Math. Phys. 4 (3) (1964) 9–36.