Advanced Math. Models & Applications
Vol.2, No.1, 2017, pp.68-77

APPLICATION OF VARIATIONAL CALCULUS METHODS IN THE STURM-LIOUVILLE PROBLEM

S.M. Zeynalli[1*], S.V. Abbasova[1]

[1]Ganja State University, Ganja, Azerbaijan

Abstract. In the paper the application of the variational calculus methods to the Sturm-Liouville problem is considered. The Ritz method is used for the finding of the solution of the considered problem.

Keywords: *Sturm-Liouville problem, Ritz method, direct problem, minimizing sequence, variational calculus.*

AMS Subject Classification: *49R50.*

Corresponding Author: *Subhiya Zeynalli, Ganja State University, AZ2001, H. Aliyev str., 429, Ganja, Azerbaijan, e-mail: subhiya.zeynalli@rambler.ru*

Manuscript received: *16 December 2016*

1. Introduction

Throughout all the history, the thinking person is engaged in optimization, i.e. finding the minimum value of some quantity: the area of land, energy, profits, and cash costs. Nowadays variational methods are one of the most powerful tools for analyzing a wide variety of intensive problems. More intensively these methods have been used in the optimal design problems. The interest to these problems has increased due to the rapid development of aviation and space technology, shipbuilding, where it is essential to solve the structural weight reduction without compromising its strength and aerodynamic properties [1, 3, 5].

The main method of proving the existence of solutions and finding of the solution of a certain variational problem was the reduction of this problem to the question of the existence of solutions of the differential equation (or system differential equations). However, this method sometimes does not lead to the desired results. Its application is further complicated by the fact that for the solution of the variational calculus problems it is required to find solutions of the corresponding differential equations are not in a small neighborhood of a point, and in a fixed domain on the border of which the desired solution must satisfy certain boundary conditions. The arising difficulties forced to find new so-called direct methods [3, 6].

The development of direct methods of the calculus of variations has been useful not only for variational problems, but also for other areas of mathematics; in particular, they are widely used in the theory of differential equations [2,3].

If the differential equation can be considered as the Euler equation for some functional and if it is determined that the original equation has a solution satisfying the boundary conditions corresponding to the problem then one can make the following considerations. As is shown below, the direct methods of variational calculus give an opportunity not only to prove the existence of a solution, but actually find it with any degree of accuracy.

There are many different methods united under the general name "direct methods". One of the most used of them is the so-called Ritz method considered below. However most of these methods are based on the same general idea, which is as follows [2, 4].

Particularly, let's consider the problem of finding of the minimum of some functional $J[y]$, defined on some class $\mathfrak{M}$ of admissible curves. In order to this problem has a sense we assume that there exist the curves in the class $\mathfrak{M}$ such that $J[y] < +\infty$ and

$$Inf J[y] = \mu > -\infty\ . \tag{1}$$

In this case by the definition of the sharp lower bound, there exists a sequence $y_1, y_2, \dots, y_n$ called a minimizing sequence such that

$$\lim_{n\to\infty} J[y] = \mu \tag{2}$$

If there exists a limit curve $y^{(0)}$ for this sequence $\{y_n\}$ and if the passing to limit

$$J\left[y^{(0)}\right] = \lim_{n\to\infty}[y_n] \tag{3}$$

is valid, then

$$J\left[y^{(0)}\right] = \mu,$$

i.e. the limit curve $y^{(0)}$ would be a solution for the considered problem.

Thus, the solution of the variational problem by the direct method consists of

1) construction of the minimizing sequence $\{y_n\}$;
2) proof of the existence of the limit curve $y^{(0)}$ for this sequence;
3) proof of the legitimacy of the passing to limit (3).

The members of the minimizing sequence can be considered as the approximate solution of the corresponding variational problem.

2. Some preliminary facts

1. Construction of the minimizing sequence is obviously always possible unless $Inf J[y] > -\infty$. Each of the direct methods used in the calculus of variations is characterized, in fact, by the way of constructing of the minimizing sequences.
2. While minimizing sequence can be constructed in any variational problem, limit curve of such a sequence may not be exist. As an example consider the functional

$$J[y] = \int_{-1}^{1} x^2 y,'^2\ dx, y(-1) = -1, y(1) = 1;$$

that takes positive values and

$$Inf J[y] = 0.$$

As a minimizing sequence here the following sequence of functions may be taken

$$y_n(x) = \frac{arctgnx}{arctgn} \qquad (n = 1,2,\dots). \tag{4}$$

Really,

$$J[y_n] = \int_{-1}^{1} \frac{nx^2dx}{arctgn+(1+n^2x^2)} = \frac{1}{2n\,arctgn} \to 0 \quad \text{for } n \to \infty.$$

But the sequence (4) in the class of continuous functions, satisfying the boundary conditions $y(-1) = -1,\ y(1) = 1$ has no limit.

The question of the validity of the limit passing (3) under the existing assumption of the limit of the minimizing sequence (if $y \to u$ means the convergence only the functions without their derivatives) for the functionals, is not trivial, as functionals considered in the calculus of variations, generally speaking, are continuous in the metric of C, and hence, the value of the functional J for the function $y^{(0)} = \lim_{n\to\infty} y_n$ generally speaking, is different from $y^{(0)} = \lim_{n\to\infty} J[y_n]$.

In some cases justification of the limit passing (3) may be done by the help of the following considerations.

For the validity of the equality (3) the continuity of the functional $J[y]$ is not necessary, and its lower boundedness is enough.

Actually, then

$$J[y^{(0)}] \geq \lim_{n\to\infty} [y_n] = infJ[y], \tag{5}$$

and from the other hand due to lower semi-continuity for large enough n,

$$J[y_n] - J[y^0] > -\varepsilon,$$

From this at $n\to\infty$ we get

$$J[y^{(0)}] \leq \lim_{n\to\infty} J[y^n] + \varepsilon,$$

i.e. due to arbitrariness of $\varepsilon > 0$,

$$J[y^{(0)}] \leq \lim_{n\to\infty} [y^n] \tag{6}$$

Thus from (5) and (6) we obtain that

$$J[y^{(0)}] \leq \lim_{n\to\infty} [y_n],$$

if the functional J is lower semi continuous.

3. Ritz method and the method of polygonal

As mentioned above, the main so-called direct methods of the calculus of variations are the construction of minimizing sequences of functions. One of the known direct methods is Ritz method that consists of the followings. Let the following minimization problem is considered

$$J[y]\to\min . \tag{7}$$

The functional is defined on a manifold from some linear normed space E.

Consider some sequence of functions

$$\varphi_1, \varphi_2, \dots, \varphi_n, \dots \tag{8}$$

from E such that the function and their linear combinations

$$y_n = c_1\varphi_1 + \dots + c_n\varphi_n \tag{9}$$

are addmissible for the functional (7). The problem is: given n to choose the coefficients c_k, k =1,2,..... n such that the value

$$J[c_1\varphi_1 + \cdots + c_n\varphi_n] \tag{10}$$

be as smaller as possible. This is a problem on finding of the minimum of the function of n variables $(c_1, c_2, \dots, c_n)$, is too simple than finding of the minimum of the functional (7). Thus for each n we get that the corresponding minimum cannot raise i.e.

$$\mu_1 \geq \mu_2 \dots.$$

since between the linear combinations of the first $n + 1$ functions contains all linear combinations of the first n functions.

Now let's consider the problem: within what conditions one can state the obtained by this way sequence of the functions $y_1, y_2, \dots, y_n, \dots$ is minimizing, *i.e*.

$$\lim_{n\to\infty} \mu_n = \mu$$

is a minimum of the functional (7).

Theorem. If the functional (7) is continuous and the system of the functions (8) is complete then

$$\lim_{n\to\infty} \mu_n = \mu$$

where μ is a minimum of the functional (7).

Proof. Let y^* be a curve on the functional (7) reaches its minimum and some $\varepsilon > 0$ be given. Since the functional (7) is assumed continuous then there exists δ >0 such that

$$|J[y] - J[y^*]| < \varepsilon \tag{11}$$

as $\|y\text{-}y^*\| < \delta$. Among the linear combinations of form (3) one can find such y_n that

$$\|y_n\text{-}y\| < \delta .$$

Then according to (11)

$$J[y_n] \leq \mu + \varepsilon.$$

If now $\bar{y}_n$ is the linear combination on which the function (10) reaches minimum by given n then

$$\left| J[\overline{y_n}] \leq J[y_n] \right| \leq \mu + \varepsilon,$$

From this by arbitrariness of ε we obtain that

$$\lim_{n\to\infty} J[\overline{y_n}] = \mu.$$

Theorem is proved.

This theorem is applicable for example, in the case when the functional type of

$$\int_a^b F(x,y,y')dx$$

is considered on some set belonging to D_1, since the functional of this type is continuous in this space.

4. Eigenfunctions and eigenvalues of the boundary problem for the Sturm-Liouville operator

Consider the application of the direct methods of the variational calculus to the differential equations on the example of the following problem. Let the Sturm-Liouville equation

$$(Py')' + \mathrm{Q}\mathrm{y} = \lambda \mathrm{y} \tag{12}$$

and boundary conditions

$$y(a) = y(b) = 0. \tag{13}$$

be given, where $\mathrm{P}(x)>0$ has continuous derivatives. It needs to find the solution of the equation (12), satisfying conditions (13), and additionally define the values of the parameter λ, at the the problem has non-zero solution.

The equation (12) together with conditions (13) is called the Sturm-Liouville boundary problem. The values of the parameter at which the equation (12) has non-zero solution satisfying (13) is called eigenvalues and corresponding solution- eigenfunctions of the considered boundary problem.

For the boundary problem (12), (13) there exist infinite number of the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n, \dots.$ and to each λ_n corresponds the only with the accuracy of constant factor eigenfunction.

Simultaneously with the proof of this fact we obtain a method for approximate determination of the eigenfunctions.

Note that the equation (12) is an Euler equation corresponding to the following conditional extremum problem: find the minimum of the functional

$$J[y] = \int_a^b \left(Py' + Qy^2\right)dx \tag{14}$$

within the condition

$$\int_a^b y^2 dx = 1 \tag{15}$$

and (13)

To apply to this variation problem the direct methods we show first the integral (14) is lower bounded. Since $P(x)>0$ then

$$\int_a^b \left(Py'^2 + Qy^2\right)dx \ge \int_a^b Qy^2 dx;$$

But

$$\int_a^b Qy^2 dx \ge M\int_a^b y^2 dx = M,$$

where

$$M = \min_{a \le x \le b} Q(x).$$

Thus the integral (14) really is lower bounded.

Now we use Ritz method. For the sake of simplicity instead of the interval (a, b) we consider (0, π). Let's take in this interval some complete system of functions $\{\varphi_n(x)\}$ satisfying conditions (13), for example, $\{\sin nx\}$.

Consider the possible linear combination of the first m functions of this system

$$y_m(x)=\sum_{n=1}^{m} a_n \sin nx \ . \tag{16}$$

The functional (14) may be written in the following quadratic form

$$J_m \equiv J[a_1, a_2, ..., a_m] = \int_0^{\pi} \left(P\left(\sum_{n=1}^{m} a_n \sin nx\right)'^2 + Q\left(\sum_{n=1}^{m} a_n \sin nx\right)^2 \right) dx, \tag{17}$$

and the condition (15) as follows

$$\int_0^{\pi} \left(\sum_{n=1}^{m} a_n \sin nx\right)^2 dx = 1 \ . \tag{18}$$

The boundary condition are satisfied automatically due to the choice of the functions $\varphi_n(x) = \sin nx$.

Integrating by terms the left hand side of (18) we get

$$\frac{\pi}{2}\sum_{n=1}^{m} a_n^2 = 1. \tag{19}$$

This means that the quadratic form (17) to which the functional (14) is reduced on the set of the functions form of (16) is considered on the surface of the sphere in the m-dimensional space. By the Weierstrass (17) reaches on the sphere (19) its minimum in some point. Let as this point be $(a_1, a_2, \dots, a_m)$ and

$$y_m = \sum_{n=1}^{m} a_n \sin nx.$$

Putting $m = 1,2, \dots$, we get the sequence of minimums

$$\lambda_1^{(1)}, \lambda_2^{(1)}, \dots \tag{20}$$

of the corresponding quadratic forms. It is easy to see that

$$\lambda_{m-1}^{(1)} \le \lambda_m^{(1)}. \tag{21}$$

Actually,

$$J_m = J[a_1, \dots a_m] = J[a_1, \dots, a_m, 0]$$

and adding one more argument can only decrease the minimum. From this and boundedness of the functional J follows that there exists the limit

$$\lim_{m\to\infty} \lambda_m^{(1)} = \lambda^{(1)}. \tag{22}$$

Thus we proved the convergence of the numerical sequence $\{\lambda_m^{(1)}\}$ consisting of the minimums of the functional

$$\int_0^{\pi} \left(Py'^2 + Qy^2\right) dx,$$

on the set of the function

$$\sum_{n=1}^{m} a_n \sin nx$$

at $m = 1,2, \ldots$

Now it is natural to try to show the convergence of the sequence of the functions

$$y_m(x) = \sum_{n=1}^{m} a_n \sin nx,$$

on which the corresponding minimal values are reached. First we show that the set $\{y_m(x)\}$ contains some uniformly converging subsequence.

For this purpose we how that the set of functions $\{y_m(x)\}$ is uniformly bounded equicontinuously continuous.

Really from the convergity of the sequence

$$\lambda_m^{(1)} = \int_0^{\pi} \left(Py'^2 + Qy_m^2 \right) dx$$

we show that the limit function satisfies the Sturm-Liouville equation (12).

The peoble is that in the integral

$$\int_0^{\pi} \left(Py_{m_k}'^2 + Qy_{m_k}^2 \right) dx \tag{23}$$

We cannot pass to limit immediately at $k \to \infty$, since we have no information on the convergence of of the derivatives $y'_{m_k}(x)$. Therefore from the fact that for each k the function $y_{m_k}(x)$ provide minimum for the integral (23) in the proper finite dimensional space we do not get that the limit function *y(x)* gives minimum to the functional (12). To get around this difficulty, we prove the following lemma.

Lemma. If for any function $\zeta\,(x)$ having continuous first and second derivatives and satisfying the boundary conditions (13), the equality

$$\int_0^{\pi} (L[\zeta] y) dx = 0,$$

is valid, where

$$L[\zeta] = -(P\zeta')' + Q_1\zeta,$$

then y is twice differentiable and $L[y] = 0$.

Now we turn back to our main problem and show that the function *y(x)*, that is a limit of the cobnstucted subsequence $\{y_{m_k}(x)\}$ satisfies the equation (12) for λ being equal to

$$\lambda^1 = \lim_{m \to \infty} \lambda_m^{(1)}.$$

The point $(a_1, a_2, \ldots, a_m)$ in which the quadratic form J_m reaches its minimum is defined according to the theory of the conditional extremum by the equation

$$\frac{\partial}{\partial a_k}\left(J[y_m]-\lambda_m\int_0^\pi\left(\sum_{n=1}^m a_n\sin nx\right)^2dx\right)=$$

$$=\int_0^\pi\left\{P(x)\left(\sum_{n=1}^m a_n\sin' nx\right)\sin' kx\mp Q(x)-\lambda_m\left(\sum_{n=1}^m a_n\sin nx\right)\sin kx\right\}dx=0, \qquad (24)$$

$(k=1,2,...,m)$

Multiplying all these equalities by the arbitrary constants $A_k^{(m)}$ and summing over k from 1 to m, we get

$$\int_0^\pi\left\{Py'_m\zeta'_m+\left(Q-\lambda_m^{(1)}\right)y_m\zeta_m\right\}dx=0\,, \qquad (25)$$

where

$$\zeta_m(x)=\sum_{k=1}^m A_k^{(m)}\sin kx. \qquad (26)$$

Let ζ be any twice differentiable function satisfying the conditions (2). Then the coefficients $A_k^{(m)}$ for any m=1,2,.... may be chosen such that

$$\zeta_m\to\zeta,\ \zeta'_{\mathrm{m}}\to\zeta',\ \zeta''_m\to\zeta''.$$

It follows from the last that

$$L[\zeta_m]\to L[\zeta] \quad \text{at} \quad m\to\infty\,.$$

Now suppose that in the inequality (25) that may be written in the form

$$\int_0^\pi L[\zeta_m]y_m dx=\lambda_m\int_0^\pi y_m\zeta_m dx, \qquad (27)$$

m varies as the sequence m_k, corresponding to the converging to the function $y(x)$ subsequence $\{y_{mk}(x)\}$. We can pass to limit in (27). This gives

$$\int_0^\pi L[\zeta]ydx=\lambda^{(1)}\int_0^\pi y\zeta dx \qquad (28)$$

for the any continuously differentiable function $\zeta(x)$. Due to the proved above lemma and the last we obtain that $y(x)$ has twice derivative and

$$L[y]=\lambda^{(1)}y.$$

Really, it is enough here to take

$$Q_1(x)=Q(x)-\lambda^{(1)}.$$

Thus we showed that $y(x)$ satisfies the equation (12).

We defined above $y(x)$ as a limit of some subsequence $\{y_{mk}(x)\}$ of the sequence $\{y_m(x)\}$. Let's show that the sequence $\{y_m(x)\}$ also converges to $y(x)$. For this purpose we use the fact that if λ is given the solution of the equation

$$-(Py')+Qy=\lambda y$$

satisfying the boundary conditions

$$y(0)=y(\pi)=0$$

and norming condition

$$\int_0^{\pi} y^2(x)dx = 1,$$

is defined up to sign. Consider such a solution and let in the point x_0 this solution is differ from zero $y(x_0) \neq 0$. Choose the sign of $y(x)$ such that $y(x_0) > 0$. The sign of $y_m(x)$ is chosen such that $y_m(x_0) \geq 0$ for all m. If $\{y_m(x)\}$ does not converge to $y(x)$, then from $\{y_m(x)\}$ may be chosen the second subsequence converging to the solution of $\bar{y}(x) \neq y(x)$. Due to the above mentionned uniquiness of the solution satisfying conditions (13), $\bar{y}(x)=y(x)$, but then $y_m(x_0) < 0$, that is impossible, since $y_m(x_0) \geq 0$. Thus $y_m(x) \to y(x)$, if onle to choose the proper sign of $y_m(x)$.

We proved the existence of the function $y(x)$, that we denote by $y^{(1)}(x)$, corresponding to one eigenvalue of the Sturm-Liouville equation. The next eigenfunction $y^2(x)$ and corresponding eigenvalue $y^{(2)}$ may be found as follows. We seek the minimum of the integral (14) subject to (13) and additional orthogonality condition

$$\int_0^{\pi} \mathrm{y}^{(1)}(x)y^{(2)}(x)dx = 0.$$

Taking

$$y_m^{(2)}(x) = \sum_{k=1}^{m} b_k \sin kx$$

we put this expression into the integral (14) together with $y(x)$. We obtain new quadratic form We consider this form on the set of the functions of form

$$\sum_{k=1}^{m} b_k \sin kx,$$

satisfying orthogonality condition to the above constructed functions $y_m(x)$, and get

$$\sum_{k=1}^{m} b_k \int_0^{\pi} \sin kx, \left(\sum_{k=1}^{m} a_n \sin nx \right) dx = 0. \tag{29}$$

The equality (29) presents the equation of the $(m-1)$–dimensional plane in the m – dimensional space, passing through origin. Its intersection with the sphere defined by the condition (15) is a sphere of dimension $m-1$. On this sphere our functional (12) is reduced to the quadratic form. Applying Weierstrass theorem we see that the quadratic form reaches on this sphere its minimum that we denote as $\lambda_m^{(2)}$. It is clear that

$$\lambda_{m+1}^{(2)} \leq \lambda_m^{(2)},$$

and since the functional (14) is lower bounded there exists a limit

$$\lambda^{(2)} = \lim_{m\to\infty} \lambda_m^{(2)}.$$

So,

$$\lambda^{(1)} \leq \lambda^{(2)}.$$

Constructing the sequence of the functions

$$y_m^{(2)}(x) = \sum_{k=1}^{m} b_k \sin kx \qquad (m = 1, 2, \ldots),$$

Each of which gives minimum to $\lambda_m^{(2)}$ и and satisfies the orthogonality condition

$$\sum_{k=1}^{m} b_k \int_0^{\pi} \sin kx \left(\sum_{n=1}^{m} a_n \sin nx \right) dx = 0,$$

we can show that this sequence uniformly converges to some limit function $y^{(2)}(x)$, satisfying the condition

$$-(Py')' + Qy = \lambda^{(2)} y,$$

boundary conditions

$$y(0) = y(\pi) = 0,$$

normalizing condition

$$\int_0^{\pi} y^2(x)dx = 1$$

and orthogonality condition and to $y^{(1)}(x)$:

$$\int_0^{\pi} y^{(1)}(x) y(x) dx = 0. \qquad (30)$$

Thus $y^{(2)}(x)$ presents the eigenfunction for the equation (12), corresponding to the eigenvalue $\lambda^{(2)}$. Since orthogonal functions cannot be linearly dependent and to each eigenvalue λ corresponds the only eigenfunction, it is valid

$$\lambda_2 > \lambda_1.$$

Repeating the similar considerations we can obtain the eigenvalues $\lambda_3, \lambda_4, \ldots$and corresponding eigenfunctions $y^{(3)}(x)$, $y^{(4)}(x), \ldots$.

References

1. Banichuk N.V., (1980) Optimization of the Forms of the Elastic Bodies, Moscow, Nauka, 255 p.
2. Gabasov R.F., (1997) Extremal problems in modern science and applications, *Soros Educational Journal*, 7, 115-120.
3. Gelfand I.M., Fomin S.V., (1961) Variational Calculus, Moscow, Fizmatiz, 228p.
4. Kolmanovskii V.B. (1997) Optimal Control Problems, *Soros Educational Journal*, 6, 121-127.
5. Pontryagin L.S., Boltiansky V.G., Gamkrelidze R.V., Mishchenko E.F., (1969) The Mathematical Theory of the Optimal Processes, Moscow, Nauka, 384 p.
6. Young L., (1974) Lectures on the Calculus of Variations and the Theory of Optimal Control, Moscow, Mir, 488 p.