Commentationes Mathematicae Universitatis Carolinae

Karel Najzar

On the method of least squares of finding eigenvalues of some symmetric operators

Commentationes Mathematicae Universitatis Carolinae, Vol. 9 (1968), No. 2, 311--323

Persistent URL: http://dml.cz/dmlcz/105183

Terms of use:

© Charles University in Prague, Faculty of Mathematics and Physics, 1968

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://project.dml.cz

Commentationes Mathematicae Universitatis Carolinae
9, 2 (1968)

ON THE METHOD OF LEAST SQUARES OF FINDING EIGENVALUES OF SOME SYMMETRIC OPERATORS

K. NAJZAR, Praha

Introduction

Among the numerical methods of finding the eigenvalues of a symmetric operator the variational methods are very important. The purpose of this paper is to prove the convergence of the method of least squares in the case of the symmetric operators with a discrete spectrum.

The principle of this method is simple and can be outlined as follows. Let A be a symmetric operator with a discrete spectrum and $\{\lambda_i\}_{i=0}^{\infty}$ be a set of eigenvalues of A. Let $\{\Psi_i\}_{i=1}^{\infty}$ be a system with properties which are described below. Let μ be a real number. Then $\mu + \eta_n$ or $\mu - \eta_n$ is the approximation to an eigenvalue of A, where

$$\eta_n = \min_{u \in \mathcal{L}\{\Psi_i\}_{i=1}^{n}} \frac{\|Au - \mu u\|}{\|u\|}$$

The approach to this problem is to be found in the book of Michlin [1] on page 390, where the problem is studied in the case in which the operator A is self-adjoint with a discrete spectrum. The proof of the basic theorem on the page 390 requires some remarks. For example:

The validity of the identity

$$\inf_{\mathfrak{D}_A} \frac{\|Au - \mu u\|}{\|u\|} = \inf_{\mathfrak{D}_{A^2}} \frac{\|Au - \mu u\|}{\|u\|}$$

is not obvious, because it is possible that for domains $\mathfrak{D}_A$, $\mathfrak{D}_{A^2}$ the following relations hold:

$$\mathfrak{D}_A \supset \mathfrak{D}_{A^2} \quad \text{and} \quad \mathfrak{D}_A \neq \mathfrak{D}_{A^2} \quad .$$

It is not obvious that the completeness and the A-completeness of the system $\{\Psi_i\}$ are sufficient conditions for the convergence in the case that A is an unbounded operator.

In the section 1 we shall introduce the notions and the terminology. Major results of this paper are summarized in Theorem 2 in the section 2 and Theorem 3 in the section 4. In conclusion we call the reader's attention to the Ritz's method and point out some of the advantages of the method of least squares in comparison with Ritz's method.

As to the mathematical formulation of the problem and to some assumptions we shall use the book of Achiezer-Glasmann [2] and Dunford-Schwartz [3].

1. In this section we collect several notations, notions and the terminology which will be used throughout the paper.

The symbol H will be used for the separable Hilbert space; I denotes the identity operator in H . Let H_i be a subspace of H , then the symbol $P_i u$ will be used to denote the projection u on H_i. The symbol $\sum_i \oplus H_i$ will be used for the direct sum of the Hilbert spaces H_i.

We shall be interested in operator A of the following types:

I) A is a symmetric operator in H, whose domain $\mathfrak{D}(A)$ is dense in H and range $\mathcal{R}(A)$ is a subset of H.

II) The spectrum $\sigma(A)$ of A is the closure of the set of eigenvalues of A.

Any operator having these properties we shall call PS-operator (operator with a point spectrum). If the operator A is a PS-operator and if the set of its eigenvalues is a set of the first category on the real axis, then we shall call it DS-operator (operator with a discrete spectrum).

By $\mathcal{L}\{\Psi_i\}_{i=1}^{n}$ we mean linear manifold generated by the vectors $\Psi_1, \Psi_2, \dots, \Psi_n$.

<u>Remark</u>. Let A be a PS-operator. Let λ_i $(i = 1, 2, \dots)$ be an enumeration of its distinct eigenvalues and let H_i $(i = 1, 2, \dots)$ be the closure of linear manifold generated by the eigenvectors of A associated with the eigenvalue λ_i. Then H may be broken into a direct sum of pairwise orthogonal subspaces H_i. If operator A is closed, then the linear manifold generated by the eigenvalues of A associated with the eigenvalue λ_i is closed and A is bounded and self-adjoint in H_i.

The symbol A_i will be used for the restriction of A to H_i.

2. One of our tools will be the following important Theorem 1.

Theorem 1: Let H be broken into a direct sum of pairwise orthogonal subspaces $\mathcal{H}_i$:

$$H = \sum_i \oplus \mathcal{H}_i \quad .$$

Let B be a linear closed operator on $\mathcal{D}_B \subset H$ satisfying the following conditions

I) $u \in \mathcal{D}_B \Rightarrow P_i u \in \mathcal{D}_B$.

II) The subspaces $\mathcal{H}_i$ and $H \ominus \mathcal{H}_i$ are invariant under B .

If we denote by B_i the restriction of B to $\mathcal{H}_i$, then $u \in \mathcal{D}_B$ if and only if

$$u_i = P_i u \in \mathcal{D}_{B_i} \quad \text{and} \quad \sum_{i=1}^{\infty} \| B_i u_i \|^2 < \infty \quad .$$

We have

$$B u = \sum_{i=1}^{\infty} B_i u_i \quad \text{for each} \quad u \in \mathcal{D}_B \quad .$$

Proof: Cf.Achiezer-Glasmann [2] on page 87.

The following lemma follows easily from Theorem 1.

Lemma 1: Let A be a closed PS-operator. Let λ_i ($i = 1, 2, \ldots$) be an enumeration of its distinct eigenvalues and let H_i ($i = 1, 2, \ldots$) be the closed linear manifold generated by the eigenvectors of A associated with the eigenvalue λ_i .

Then $u \in \mathcal{D}_A$ if and only if

$$\sum_{i=1}^{\infty} \| A_i u_i \|^2 < \infty \;, \quad \text{where } u_i = P_i u \,;\; A_i = A / H_i$$

and $A u = \sum_{i=1}^{\infty} A_i u_i = \sum_{i=1}^{\infty} \lambda_i u_i$.

Remark: Lemma 1 is a consequence of Riesz-Lerch's Lemma, cf. [5].

In the following the Theorem 2 will be found useful.

<u>Theorem 2</u>: Let A be a PS-operator. Let μ be an arbitrary real number. Then

$$\inf_{u \in \mathcal{D}_A} \frac{\| Au - \mu u \|}{\| u \|} = q ,$$

where $q = \inf_{j=1,2,\dots} |\lambda_j - \mu|$.

<u>Proof</u>: a) Firstly, we assume that A is closed PS-operator. Let A_i $(i = 1, 2, \dots)$ be a restriction of A to the subspace H_i . Then A_i is uniquely given by the formula

$$A_i u = Au = \lambda_i u , \quad u \in H_i$$

and thus A_i is the self-adjoint and bounded operator on H_i .

By lemma 1

$\mathcal{D}_A = \{ u \in H : \sum_{i=1}^{n} \| A_i u_i \|^2 = \sum_{i=1}^{\infty} \lambda_i^2 \| u_i \|^2 < \infty$, where $u_i = P_i u$ is the projection of u on H_i $\}$

and for each $u \in \mathcal{D}_A$

$$Au = \sum_{i=1}^{\infty} A_i u_i = \sum_{i=1}^{\infty} \lambda_i u_i \quad .$$

Thus

$$\frac{\| Au - \mu u \|^2}{\| u \|^2} = \frac{\sum_{i=1}^{\infty} (\lambda_i - \mu)^2 \| u_i \|^2}{\sum_{i=1}^{\infty} \| u_i \|^2} = q^2 + \frac{\sum_{i=1}^{\infty} [(\lambda_i - \mu)^2 - q^2] \cdot \| u_i \|^2}{\sum_{i=1}^{\infty} \| u_i \|^2} \geq q^2$$

for each $u \in \mathcal{D}_A$.

It follows that

$$\inf_{u \in \mathcal{D}_A} \frac{\| Au - \mu u \|}{\| u \|} \geq q .$$

To prove Theorem 1 it is sufficient to show that

$$\inf_{u \in \mathcal{D}_A} \frac{\| Au - \mu u \|}{\| u \|} \leqq q .$$

Now

$$\inf_{u \in \mathfrak{D}_A} \frac{\|Au - \mu u\|}{\|u\|} \leqq \frac{\|A\varphi_i - \mu\varphi_i\|}{\|\varphi_i\|} = |\lambda_i - \mu|, \; i = 1,2,3,\dots,$$

where φ_i is an eigenfunction corresponding to the eigenvalue λ_i,

and thus

(1) ..... $$\inf_{u \in \mathfrak{D}_A} \frac{\|Au - \mu u\|}{\|u\|} \leqq \varrho \; .$$

b) Secondly, let A be a PS-operator. The minimal closed symmetric extension of A we denote by $\bar{A}$ (cf.[3], p. 1226). Then $\bar{A}$ is closed PS-operator and the spectra of the operators A and $\bar{A}$ coincide.

Therefore by a)

$$\inf_{u \in \mathfrak{D}_A} \frac{\|Au - \mu u\|}{\|u\|} \geq \inf_{u \in \mathfrak{D}_A} \frac{\|\bar{A}u - \mu u\|}{\|u\|} = \varrho$$

and by (1) we have

$$\inf_{u \in \mathfrak{D}_A} \frac{\|Au - \mu u\|}{\|u\|} = \varrho \; .$$

This completes the proof of the Theorem 2.

3. In this section we shall define A-complete system and total-complete system.

<u>Definition 1</u>: Let there be given linearly independent system $\{\Psi_i\}$, $\Psi_i \in \mathfrak{D}_A$, $i = 1, 2, \dots$. This system $\{\Psi_i\}$ will be said to be A-complete if for every $u \in \mathfrak{D}_A$ and $\varepsilon > 0$ there exists $n(u, \varepsilon)$ and $u_n \in \mathcal{L}\{\Psi_i\}_{i=1}^{n}$ such that

$$\| Au - Au_n \| < \varepsilon$$

(Cf.Michlin [1].)

Definition 2: Let there be given linearly independent system $\{\Psi_i\}$, $\Psi_i \in \mathcal{D}_A$, $i = 1, 2, \ldots$. This system $\{\Psi_i\}$ will be said to be total-complete if $\{\Psi_i\}$ is A_μ-complete for every real μ, where $A_\mu = A - \mu I$.

Remark: In case, in which $0 \bar{\in} \sigma(A)$ any complete and A-complete system is total-complete.

Lemma 2: Let A be a DS-operator. Then

I) There exists a real number μ , such that the operator $(A - \mu I)^{-1}$ is bounded. We denote by A_μ the operator $A - \mu I$.

II) Let there be given a linearly independent system $\{\Psi_i\}$, $\Psi_i \in \mathcal{D}_A$, $i = 1, 2, \ldots$. Then $\{\Psi_i\}$ is total-complete if and only if $\{\Psi_i\}$ is A_μ-complete, where A_μ is the operator defined in I.

Proof: The spectrum of A is real and not dense on the real axis. The statement I follows from it. To prove II, let u be in $\mathcal{D}_A$ and let μ_1 be a real number. Then

$$\begin{aligned} \|(A - \mu_1 I)u\| &\leq \|Au - \mu u + (\mu - \mu_1)u\| \leq \|A_\mu u\| + \\ &+ |\mu - \mu_1| \cdot \|A_\mu^{-1} A_\mu u\| \leq \|A_\mu u\| \cdot \\ &\cdot (1 + |\mu - \mu_1| \cdot \|A_\mu^{-1}\|) = k \cdot \|A_\mu u\| . \end{aligned} \tag{2}$$

Let us assume that the system $\{\Psi_i\}$ is A_μ-complete. Then for every $\varepsilon > 0$ there exists n and $u_n \in \mathcal{L}\{\Psi_i\}_{i=1}^n$ such that

$$\|A_\mu u - A_\mu u_n\| < \varepsilon .$$

By (2), we have

$$\|(A - \mu_1 I)(u - u_n)\| \leq k \cdot \|A_\mu (u - u_n)\| < k \cdot \varepsilon .$$

It follows the necessity of II.

The sufficiency follows easily from Definition 1 and 2.

Remark: If the operator A is PS-operator, then the statement of Lemma 2 is true. If the relation $\mu \overline{\in} \sigma(A)$ holds, then the proof is analogous to the proof of Lemma 2.

In the next section we shall construct the approximation q_n of the number q defined in Theorem 2 and prove the convergence q_n to q.

4. The next theorem gives useful information on the convergence of the method of least squares.

Theorem 3: Let A be a DS-operator. Let $\{\Psi_i\}$ be a total-complete system. Let the sequence $\{q_n\}$ be defined by formula

$$q_n = \min_{u \in \mathcal{L}\{\Psi_i\}_{i=1}^{n}} \frac{\| Au - \mu u \|}{\| u \|} .$$

Then the sequence $\{q_n\}$ is monotone decreasing and converging to the number q, where $q = \inf_{i=1,2,\dots} |\lambda_i - \mu|$.

Proof: 1) Firstly, assume that $q > 0$. It follows from Theorem 2 that

(3) .... $\| Au - \mu u \| \geq q \cdot \| u \|$ for $u \in \mathcal{D}_A$.

Now let φ_j be an eigenfunction corresponding to the eigenvalue λ_j such that $\| \varphi_j \| = 1$.

Select ε, $0 < \varepsilon < q$ and select j such that $|\lambda_j - \mu| < q + \varepsilon$. Since $\{\Psi_i\}$ is total-complete, there exists for ε a number n and $u_n \in \mathcal{L}\{\Psi_i\}_{i=1}^{n}$ such that

(4) ..... $\| (A - \mu I)(u_n - \varphi_j) \| < \varepsilon$.

By (3) we have

(5) ..... $\|(A-\mu I)(u_n-\varphi_j)\| \geq q\cdot\|u_n-\varphi_j\|$

and so

(6) ..... $\|(A-\mu I)u_n\| \leq \|(A-\mu I)(u_n-\varphi_j)\|+\|(A-\mu I)\varphi_j\| \leq \varepsilon+|\lambda_j-\mu|.$

From (4) and (5) it follows

$$\|u_n-\varphi_j\| \leq \frac{\varepsilon}{q} < 1$$

(7) so that

$$\|u_n\| = \|u_n-\varphi_j+\varphi_j\| \geq \big|\,\|\varphi_j\|-\|u_n-\varphi_j\|\,\big| \geq 1-\frac{\varepsilon}{q} .$$

Therefore, by (6) and (7) we find

$$q \leq q_n \leq \frac{\|(A-\mu E)u_n\|}{\|u_n\|} \leq \frac{\varepsilon+|\lambda_j-\mu|}{1-\frac{\varepsilon}{q}} \leq \frac{2\varepsilon+q}{1-\frac{\varepsilon}{q}} .$$

It follows immediately the statement.

B) Secondly, assume that $q = 0$. Select $\varepsilon > 0$. The set of eigenvalues of A is of the first category on the real axis. It follows that there exists real number μ_1 such that

$$|\mu-\mu_1| < \varepsilon \quad \text{and} \quad q_1 = \inf_{i=1,2,\dots} |\lambda_i-\mu_1| > 0 .$$

Then

$$q_1 = \inf_{i=1,2,\dots} |\lambda_i-\mu+\mu-\mu_1| \leq \inf_{i=1,2,\dots} |\lambda_i-\mu|+|\mu-\mu_1| = |\mu-\mu_1| < \varepsilon .$$

Consequently, from a) there follows:

for $\varepsilon > 0$ there exists n such that

$$0 < q_1 \leq \min_{u\in\mathcal{L}\{\Psi_i\}_{i=1}^{n}} \frac{\|Au-\mu_1 u\|}{\|u\|} \leq q_1+\varepsilon .$$

Thus

$$q_n = \min_{u\in\mathcal{L}\{\Psi_i\}_{i=1}^{n}} \frac{\|Au-\mu u\|}{\|u\|} \leq \min_{u\in\mathcal{L}\{\Psi_i\}_{i=1}^{n}} \frac{\|(A-\mu_1 I)u\|+|\mu-\mu_1|\cdot\|u\|}{\|u\|} \leq$$

$$\leq q_1+\varepsilon+|\mu-\mu_1| \leq 3\varepsilon .$$

Therefore $\lim_{n \to \infty} q_n = 0$ and the proof is completed.

For the method of the determining of q_n cf. Michlin [1].

5. Proceeding in an analogous manner, we may also establish the convergence of the Ritz's method for the finding eigenvalues of A. The valuable results are summarized in detail in the book of Michlin [1]. The rest of the paper is devoted to reformulations and extensions of the results of Michlin in case, in which DS-operator A is bounded below.

<u>Theorem 4</u>: Let A be a DS-operator which is bounded below. Let $\lambda_1 < \lambda_2 < \lambda_3 \dots$ be an enumeration of its distinct eigenvalues increasing order of values and let μ be such a number that $\mu < \lambda_1$. Let there be given a total-complete system $\{\Psi_i\}$. Then

a) $$\inf_{u \in \mathcal{D}_A} \frac{(Au - \mu u, u)}{\|u\|^2} = \lambda_1 - \mu ,$$

b) Denote

$$Q_n = \min_{u \in \mathcal{L}\{\Psi_i\}_{i=1}^n} \frac{(Au - \mu u, u)}{\|u\|^2} \quad \text{(Ritz's method)} .$$

The sequence $\{\lambda_n\}$ is monotone decreasing and it converges to $(\lambda_1 - \mu)$.

c) Q_n satisfies the inequality

$Q_n \leq q_n$, where q_n is defined in Theorem 3.

<u>Proof</u>: a) By lemma 1 we have

$$\frac{(Au - \mu u, u)}{\|u\|^2} = \frac{\sum_{i=1}^{\infty} (\lambda_i - \mu) u_i^2}{\sum_{i=1}^{\infty} u_i^2} = \lambda_1 - \mu + \frac{\sum_{i=1}^{\infty} (\lambda_i - \lambda_1) u_i^2}{\sum_{i=1}^{\infty} u_i^2} \geq \lambda_1 - \mu$$

for each $u \in \mathcal{D}_A$.

Applying Theorem 3 to the operator A we obtain

$$\inf_{u \in \mathfrak{D}_A} \frac{\|Au - \mu u\|}{\|u\|} = \lambda_1 - \mu .$$

Since

$$(8) \ldots\ldots \quad \frac{\|Au - \mu u\|}{\|u\|} \geq \frac{|(Au - \mu u, u)|}{\|u\|^2} ,$$

it follows that

$$\inf \frac{(Au - \mu u, u)}{\|u\|^2} = \lambda_1 - \mu .$$

The statement c) follows immediately from (8).

The statement b) follows immediately from c) and from Theorem 3.

For the following eigenvalues we have

Theorem 5: Let A be a DS-operator which is bounded below. Let $\lambda_1 < \lambda_2 < \ldots$ be an enumeration of its distinct eigenvalues increasing order of values. Let μ be a real number such that $\lambda_m \leq \mu \leq \lambda_{m+1}$.

Let there be given a total-complete system $\{\Psi_i\}$. We denote by H^1 the direct sum of the Hilbert spaces $H_1, \ldots$ $\ldots, H_m$:

$$H^1 = \sum_{i=1}^{m} \oplus H_i$$

and by H^2 the orthocomplement of H^1 in H. Then

a) $$\inf_{u \in \mathfrak{D}_A \cap H^2} \frac{(Au - \mu u, u)}{\|u\|^2} = \lambda_{m+1} - \mu .$$

b) Denote

$$Q_n = \min_{u \in \mathcal{L}\{\Psi_i\}_{i=1}^{n} \cap H^2} \frac{(Au - \mu u, u)}{\|u\|^2} .$$

Then the sequence $\{Q_n\}$ is monotone decreasing and converging to $(\lambda_{m+1} - \mu)$.

c) If $\lambda_{m+1} - \mu < \mu - \lambda_n$ and $\Psi_i \in H^2$ then $Q_n \leq q_n$.

Proof: By lemma 3, we have

$$\frac{(Au - \mu u, u)}{\|u\|^2} = \frac{\sum_{i=m+1}^{\infty} (\lambda_i - \mu) u_i^2}{\sum_{i=m+1}^{\infty} u_i^2} \quad \text{for } u \in H^2$$

and the proof of this Theorem is similar to that of Theorem 4.

In conclusion we shall point out some of the advantages of using the method of least squares in comparison with the Ritz's method. The principal disadvantage of Ritz's method lies in the necessity of evaluating eigenfunctions associated with the eigenvalue λ_i, $\lambda_i < \mu$. To obtain the approximation of λ_{m+1} we must know the subspace $H^2 = \sum_{i=1}^{m} \oplus H_i$ (cf.Theorem 4). The Ritz's method gives only upper bound of the eigenvalue. In case of the method of least squares we can obtain upper or lower bound of the eigenvalue for some particular choosing of μ (cf.Theorem 3).

The numerical aspects of the method of least squares including the stability of appertaining processes were studied. The results will be published elsewhere.

R e f e r e n c e s

[1] C.P. MICHLIN: Prjamyje metody v matematičeskoj fizike,1950.

[2] N.I. ACHIEZER - I.M. GLASMANN: Theorie der linearen Operatoren im Hilbert-Raum,1960.

[3] N. DUNFORD and J.T. SCHWARTZ: Linear Operators,part II,1963.

[4] A.E. TAYLOR: Introduction to functional analysis, 1958.
[5] K. MAURIN: Methods of Hilbert spaces, 1967.

(Received May 16, 1968)