REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
Tome 23, N° 2, 1994, pp. 217–225

A NOTE ON THE STABILITY OF THE GENERALIZED RITZ METHOD

M. E. TITENSKY
(Beer-Sheva)

1. INTRODUCTION

The purpose of this note is to present effective conditions of the generalized Ritz method's stability in the $l_p(1 \leqslant p \leqslant \infty)$-norms (section 4). Another purpose is to underline that the assertions of this note can be in certain cases put into effect without any restrictions to the generalized Ritz matrix's elements (section 6).

As far as we know, a question of numerical stability was first put and considered by S. Mikhlin [5]. In the book he had established necessary and sufficient conditions of the Ritz method in the l_2-norm. Some sufficient conditions of the stability of this method in the uniform norm were obtained by the author [13–14].

The notions of the generalized Ritz method and its stability are introduced in sections 2 and 3 respectively. The applications of our results to the special class of coordinate systems are given in section 5.

2. PRELIMINARIES

Let H be a real separable Hilbert space. We consider the operator equation

(2.1) $$Au = f,$$

where A is a linear unbounded operator defined on a dense domain $D(A) \subset H$, $u \in D(A)$ is an unknown element, an element $f \in H$ is given.

DEFINITION 1 [3–4,9–11]. An operator A is said to be K-positively defined (K-p.d.) if there exists a closeable operator K with $D(K) \supseteq D(A)$ such that $KD(A)$ is dense in H and

(2.2) $$(Au,\ Ku) \geqslant \gamma_1^2 \|u\|^2,$$

(2.3) $$\|Ku\|^2 \leqslant \gamma_2^2(Au,\ Ku)$$

for some positive constants γ_1, γ_2 and all $u \in D(A)$. ((,) and $\| \|$ denote respectively the inner product and norm in H).

A is K-symmetric if

(2.4) $$(Au,\ Kv) = (Ku,\ Av)$$

for all $u,\ v \in D(A)$. If A is K-symmetric and K-p.d. and H_0 denotes the completion of $D(A)$ in the metric

(2.5) $$[u,\ v] = (Au,\ Kv),\quad |u|^2_{H_0} = [u,\ u],$$

then H_0 can be regarded as a subset of H and A has a closed K-p.d. and K-symmetric extension A_0 which is continuously invertible.

Let $(\varphi_i) \subset D(A)$ be a complete system of linearly independent elements in H_0 (a coordinate system). The approximate solution of the generalized Ritz method can be written in the form

(2.6) $$u_n = \sum_{i=1}^{n} a_i^{(n)} \varphi_i.$$

To compute the scalars $a_i^{(n)} (i = 1, \ldots, n)$, we must solve the following system of algebraic equations

(2.7) $$R_n a^{(n)} = f^{(n)}.$$

Here $a^{(n)} = (a_i^{(n)})$, $f^{(n)} = (f_i^{(n)})(f_i^{(n)} = (f,\ K\varphi_i))(i = 1, \ldots, n)$ are vectors and $(R_n)_{ij} = (A\varphi_i,\ K\varphi_j) = c_{ij} = c_{ji}(i,\ j = 1, \ldots, n)$ is the generalized Ritz matrix.

It is well known [9–11] that if (2.2)-(2.4) are hold, there exists one and only one solution of equation (2.1) in H_0 and

(2.8) $$|u_n - u|_{H_0} \to 0$$

as $n \to \infty$.

If $K = I$ we come to the usual Ritz method.

3. NUMERICAL STABILITY

When writing down the system (2.7) we make errors. Let $\gamma_{ij} = \gamma_{ji}$ denote the (small) errors arising in the evaluation of the inner products $(A\varphi_i,\ K\varphi_j)$ and Γ_n the matrix with elements $\gamma_{ij}(i,\ j = 1, \ldots, n)$. Let $\delta_i^{(n)}$ be the corresponding errors in $f_i^{(n)}$ and $\delta^{(n)}$ the vector with elements $\delta_i^{(n)}$. Instead of the „exact" system (2.7) we solve the „nonexact" system

(3.1) $$(R_n + \Gamma_n)b^{(n)} = f^{(n)} + \delta^{(n)},$$

where $b^{(n)} = (b_1^{(n)}, \ldots, b_n^{(n)})$ is the column-vector of the „nonexact" Ritz system. We assume that (2.7) and (3.1) are solved quite exactly, i.e., without round-off errors.

Let $z^{(n)}$ denote a vector with elements $z_1^{(n)}, \ldots, z_n^{(n)}$. Let

$$\|z^{(n)}\|_{l_p^{(n)}} = \left(\sum_{i=1}^{n} |z_i^{(n)}|^p\right)^{1/p} \quad (1 \leqslant p < \infty),$$

$$\|z^{(n)}\|_{l_\infty^{(n)}} = \max_{1 \leqslant i \leqslant n} |z_i^{(n)}|.$$

DEFINITION 2. The solution of (2.7) is stable, of there exist positive constants M, Q and P, independent of n, such that for $\|\Gamma_n\|_{l_p^{(n)}} \leqslant M$ and arbitrary $\delta^{(n)}$ the system (3.1) is solvable and the following inequality holds

$$\|b^{(n)} - a^{(n)}\|_{l_p^{(n)}} \leqslant Q\,\|\Gamma_n\|_{l_p^{(n)}} + P\,\|\delta^{(n)}\|_{l_p^{(n)}},$$

where $\|\Gamma_n\|_{l_p^{(n)}}$ are the norms of the matrix $(\gamma_{ij})_{ij=1}^{n}$ in $l_p^{(n)}$. It should be noted that

(3.2) $$\|\Gamma_n\|_{l_\infty^{(n)}} = \|\Gamma_n\|_{l_1^{(n)}} = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |\gamma_{ij}|.$$

DEFINITION 3. The aproximative solution of the generalized Ritz method (2.6) is said to be stable, if there exist positive constants M_1, Q_1 and P_1, independent of n, such that for $\|\Gamma_n\|_{l_p^{(n)}} \leqslant M_1$ and arbitrary $\delta^{(n)}$ the system (3.1) is solvable and the following inequality holds

$$|v_n - u_n|_{H_0} \leqslant Q_1\,\|\Gamma_n\|_{l_p^{(n)}} + P_1\,\|\delta^{(n)}\|_{l_p^{(n)}},$$

where $v_n = \sum_{i=1}^{n} b_i^{(n)} \varphi_i$.

DEFINITION 4. If the solution of (2.7) and the approximate solution of the generalized Ritz method (2.6) are stable, the generalized Ritz method is called stable.

4. RESULTS

In this section some results on the generalized method's Ritz stability will be formulated and proved.

THEOREM 1. *Let a coordinate system* (φ_i) *be such that*

(4.1) $$c_{ii} = 1 \quad (i = 1, \ldots,)$$

and

(4.2) $$\max_{1 \leqslant i \leqslant n} \sum_{j=1,\ j \neq i}^{n} |c_{ij}| \leqslant \lambda_0 < 1,$$

where λ_0 is a positive constant, independent of n. Then the solution of (2.7) is stable in the l_∞-norm.

Proof. Following [13–14] rewrite the system (2.7) as

$$a_i^{(n)} = f_i^{(n)} - \sum_{j=1,\ j \neq i}^{n} c_{ij}\, a_j^{(n)} \quad (i = 1, \ldots, n).$$

Then

(4.3) $$\|a^{(n)}\|_{l_\infty^{(n)}} \leqslant \|f^{(n)}\|_{l_\infty^{(n)}} + \max_{1 \leqslant i \leqslant n} \sum_{j=1,\ j \neq i}^{n} |c_{ij}|\, \|a^{(n)}\|_{l_\infty^{(n)}}.$$

From (4.3), taking in account (4.2), we obtain

$$(4.4) \qquad \|a^{(n)}\|_{l_\infty^{(n)}} \leqslant (1-\lambda_0)^{-1}\|f^{(n)}\|_{l_\infty^{(n)}}.$$

Now

$$(4.5) \qquad \|f^{(n)}\|_{l_\infty^{(n)}} = \max_{1\leqslant i\leqslant n} |(f, K\varphi_i)| \leqslant \max_{1\leqslant i\leqslant n} \|K\varphi_i\|\|f\| \leqslant \gamma_2\|f\|$$

because of (2.3) and (4.1). (4.4) and (4.5) imply that $\|R_n^{-1}\|_{l_\infty^{(n)}}$ and $\|a^{(n)}\|_{l_\infty^{(n)}}$ are bounded above independently of n. Hence by Theorem 13.1 [5] we immediately come to the desired result.

THEOREM 2. *Assume that the assumptions of Theorem 1 are fulfilled and there exists a positive constant N_1, independent of n, such that*

$$(4.6) \qquad \|f^{(n)}\|^p_{l_p^{(n)}} \leqslant N_1 \quad (1 \leqslant p < \infty).$$

Then the solution of (2.7) is stable in the $l_p^{(n)}(1 \leqslant p < \infty)$-norms.

Proof. Let R_n^{-1} be the matrix with elements $g_{ji} = g_{ij}(i, j = 1, \ldots, n)$ and $1 < p < \infty$. We have by Holder's inequality

$$\left|\sum_{i=1}^n g_{ij}x_i^{(n)}\right| \leqslant \left(\sum_{i=1}^n |g_{ij}\|x_i^{(n)}|^p\right)^{1/p}\left(\sum_{i=1}^n |g_{ij}|\right)^{1/q} \quad (p^{-1}+q^{-1}=1)$$

for all $x^{(n)} = (x_i^{(n)})(i = 1, \ldots, n)$. Since by Theorem 1 $\|R_n^{-1}\|_{l_\infty^{(n)}} \leqslant (1-\lambda_0)^{-1}$, we obtain because of (3.2)

$$\sum_{j=1}^n\left|\sum_{i=1}^n g_{ij}x_i^{(n)}\right|^p \leqslant (1-\lambda_0)^{-(p/q+1)}\sum_{i=1}^n |x_i^{(n)}|^p.$$

Hence

$$(4.7) \qquad \|R_n^{-1}x^{(n)}\|_{l_p^{(n)}} \leqslant (1-\lambda_0)^{-1}\|x^{(n)}\|_{l_p^{(n)}}$$

and $\|R_n^{-1}\|_{l_p^{(n)}}(1 \leqslant p < \infty)$ are bounded above independently of n.

Now we estimate

$$a^{(n)} = R_n^{-1}f^{(n)}$$

in the l_p-norm. Then using (4.6) and (4.7) we establish that $\|a^{(n)}\|_{l_p^{(n)}}$ are bounded above independently of n too. Q.E.D.

THEOREM 3. *Let under the assumptions of Theorem 2 the inequality (4.6) be true if $p = 1$. Then the approximate solution of the generalized Ritz method (2.6) is stable.*

Proof. As

$$u_n - v_n = \sum_{i=1}^n (a_i - b_i)\varphi_i$$

and $|\varphi_i|_{H_0} = 1$ $(i = 1, \ldots)$, the desired result follows from Theorem 2.

5. SPEARED SYSTEMS AND STABILITY

In this section we consider a class of coordinate systems for which (4.1) and (4.2) are easily verifiable.

DEFINITION 5 [2]. Let (φ_i) be a coordinate system and

(a) $$|\varphi_i|_{H_0} = 1 \quad (i = 1, \ldots);$$

(b) $$\sup_j \sum_{i=1,\, i\neq j}^{\infty} |[\varphi_i, \varphi_j]| \leqslant \lambda_0$$

(λ_0 is a positive constant);

(c) $$\lim_{j\to\infty} \sum_{i=1,\, i\neq j}^{\to\infty} |[\varphi_i, \varphi_j]| = 0.$$

Then this system is called a speared system with respect to operators A and K.

DEFINITION 6 [12]. If (φ_i) satisfies the assumptions (a) and (b) only (see Definition 5), then this system is called a quasi speared system with respect to operators A and K.

THEOREM 4. *Suppose that (φ_i) is a speared system with respect to A and K. Then the solution of (2.7) is stable in l_∞. If, moreover, (4.6) is fulfilled, then the stability takes place in the $l_p(1 \leqslant p \leqslant \infty)$-norms.*

Proof. As (φ_i) is a speared system with respect to A and K we conclude that $\|R_n^{-1}\|_{l_\infty^{(n)}} \leqslant C$, where C is a constant, independent of n [2]. Besides $\|f^{(n)}\|_{l_\infty}$ are bounded above independently of n as well (see the proof of Theorem 1). Thus the assertions follow now from Theorem 13.1 [5] and Theorem 2.

THEOREM 5. *Let (φ_i) be a quasi speared system with respect to A and K and $\lambda_0 < 1$ (see Definition 5). Then the solution of (2.7) is stable in the l_∞-norm. If, moreover, condition (4.6) is true, then the stability takes place in the l_p-norms $(1 \leqslant p < \infty)$.*

This result follows immediatedly from Theorems 1 and 2.

6. EXAMPLE

It is possible in certain cases to choose an operator K so as to ensure the stability.

One of these cases is embodied in the following example.

Let $H = L_2\ [0, \pi]$ and

$$(6.1) \qquad Au = -\,\mathrm{d}/\mathrm{d}x\,(b(x)\,\mathrm{d}u/\mathrm{d}x)$$

is defined on the domain $D(A)$ of twice continuously differentiable functions with the boundary conditions

$$(6.2) \qquad u(0) = u'(\pi) = 0.$$

Let $b(x)$ be continuously differentiable on $[0, \pi]$ and $b(x) \geq b_0 > 0$. Let us assume that K is given by the formula

$$Ku = \int_0^x a(t)\, b^{-1}(t) u'(t)\, dt \tag{6.3}$$

(a function $a(x) \geq a_0 > 0$ will be chosen below). Now we write down the scalar product (Au, Ku) in $L_2\,[0, \pi]$ (A and K are as defined in (6.1) and (6.3) respectively)

$$(Au, Kv) = -\int_0^\pi \left[d/dx(b(x)\, du/dx) \int_0^x a(t)\, b^{-1}(t) v'(t)\, dt \right] dx. \tag{6.4}$$

Integrating (6.4) by parts we have for all $u, v \in D(A)$

$$(Au, Kv) = -\left[b(x)\, du/dx \int_0^x a(t)\, b^{-1}(t) v'(t)\, dt \right]\Big|_0^\pi +$$

$$+\int_0^\pi a(x) \frac{du}{dx}\frac{dv}{dx}\, dx = (Av, Ku)$$

because of (6.2), i.e., A is K-symmetric operator. Next

$$u(x) = \int_0^x u'(t)\, dt = \int_0^x \sqrt{a(t)}\, u'(t) \sqrt{a^{-1}(t)}\, dt.$$

Hence, by Cauchy's inequality,

$$\|u\|_{L_2}^2 \leq \pi \int_0^\pi a^{-1}(x)\, dx \int_0^\pi a(x)(u'(x))^2\, dx = \gamma_1^{-2}(Au, Ku),$$

where $\gamma_1 = \left(\pi \int_0^\pi a^{-1}(x)\, dx \right)^{-1/2}$.

At last we use Cauchy's inequality once more to estimate

$$\|Ku\|_{L_2}^2 = \int_0^\pi \left| \int_0^x a(t) b^{-1}(t)\, u'(t)\, dt \right|^2 dx.$$

Thus we come to inequality (2.3), where $\gamma_2 = \left(\pi \int_0^\pi a(x)\, b^{-2}(x)\, dx \right)^{1/2}$, i.e., A is K-p.d. operator.

Let H_0 denote the linear space of all real functions $u(x)$ absolutely continuous on $[0, \pi]$, satisfying the boundary conditions (6.2) and such that $u'(x) \in L_2\,[0, \pi]$ [6]. Evidently that

$$|u|_{H_0} = \sqrt{\int_0^\pi a(x)(u'(x))^2\, dx}.$$

We choose the next system as a coordinate one

$$\varphi_i = r_i^{-1} \sin r_i x \left(\int_0^\pi a(x) \cos^2 r_i x)\, dx \right)^{-1/2} \quad (i = 1, \ldots,), \tag{6.5}$$

where $r_i = (2i-1)/2$. It is clear that $(\varphi_i) \subset D(A)$ and this system is complete in H_0.

We make sure that

$$c_{ii} = (A\varphi_i, K\varphi_i) = 1 \quad (i = 1, \ldots,).$$

Now we have

$$c_{ij} = l_i\, l_j \int_0^\pi a(x) \cos r_i x \cos r_j x dx, \tag{6.6}$$

where $l_i = \left(\int_0^\pi a(x) \cos^2 r_i\, x\, dx \right)^{-1/2}$.

Integrating (6.6) twice by parts we obtain

$$(g_{ij} = i + j - 1, \quad \alpha_{ij} = i - j)$$

$$c_{ij} = 2^{-1} l_i\, l_j [g_{ij}^{-2} [a'(x) \cos g_{ij}\, x]|_0^\pi - \alpha_{ij}^{-2} \int_0^\pi a''(x) \cos \alpha_{ij}\, x\, dx -$$

$$- g_{ij}^{-2} \int_0^\pi a''(x) \cos g_{ij}\, x\, dx + \alpha_{ij}^{-2} [a'(x) \cos \alpha_{ij} x]|_0^\pi \Big]. \tag{6.7}$$

Hence we get the estimate

$$|c_{ij}| \leq 2\pi^{-1} a_0^{-1} [\,|a'(0)| + |a'(\pi)| + \pi \max_{0 \leq x \leq \pi} |a''(x)|\,](i-j)^{-2},$$

i.e.,

$$\sup_i \sum_{j=1,\, j \neq i}^{\infty} |c_{ij}| \leq \lambda_0,$$

where

$$\lambda_0 = 4\pi^{-1} a_0^{-1} [\,|a'(0)| + |a'(\pi)| + \pi \max_{0 \leq x \leq \pi} |a''(x)|\,] \sum_{m=1}^{\infty} m^{-2}.$$

So the system (6.5) is a quasi speared one with respect to operators A and K. Let us choose, fir example, $a(x) = e^{\varepsilon x}$, where ε is a positive number, $a_0 = 1$. If ε is too small, condition (4.2) is always valid, i.e., the generalized Ritz method of problem (2.1) (A is as defined in (6.1)) is stable.

7. CONCLUDING REMARKS

Remark 1. An important class of equations are those of the form

$$\frac{du}{dt} + Au = f(t), \quad u(0) = u_0, \quad 0 \leqslant t \leqslant T. \tag{7.1}$$

Here an operator A is K-p.d. and K-symmetric acting in a real separable Hilbert space H_0; an element $u_0 \in H_0$, $u(t)$ is an unknown function, $f(t)$ is a continuous one on $[0, T]$ ($f(t) \in H_0$ for all $t \in [0, T]$).

We are going to investigate the method's Galerkin stability of problem (7.1) in one of the following papers.

Remark 2. All the results of our note are true for the coordinate systems

$$(\varphi_i^{(m_k)})(i = 1, \ldots, m_k; \ 0 < m_1 <, \ldots, < m_k < m_{k+1} <, \ldots,)$$

as well. Elements $(\varphi_i^{(m_k)})$ do not, in general, appear amongst elements $(\varphi_i^{(m_{k+1})})$ for $k = 1, 2, \ldots,$ [1,7 – 8]. These systems are widespread in applications (see the finite elements method).

REFERENCES

1. R. S. Anderssen and A. R. Mitchell, *Analysis of generalized Galerkin methods in the numerical solution of elliptic equations*, Math. Meth. Appl. Sci., **1**, (1979), 3–15.
2. D. S. Jones, *Galerkin's method and stability*, Math. Meth. Appl. Sci., **2** (1980), 347–377.
3. A. E. Martyniuk, *Variational methods in boundary value problems for weakly elliptic equations*, Dokl. Akad. Nauk SSSR, **126** (1959), 1222–1225 (Russian).
4. A.E. Martyniuk, *On the method's Galerkin-Krilov inverse variants*, Vichislitelnaya i prikladnaya mat., **14** (1971), 18–35 (Russian).
5. S. G. Mikhlin, *The numerical performance of variational methods*, Wolters-Noordhoff Publishing, Groningen, 1971.
6. S. G. Mikhlin, *Variational methods in mathematical physics*, The Macmillan Co, New York, 1964.
7. S. G. Mikhlin, *Errors of computational processes*, Tbilis Gos. Univ., Tbilisi, 1983 (Russian).
8. S. G. Mikhlin, *Some problems in error theory*, Leningrad Univ., Leningrad, 1988 (Russian).
9. W. V. Petryshyn, *Direct and iterative methods for the solution of linear operator equations in Hilbert space*, Trans. Amer. Math. Soc., **105** (1962), 136–175.
10. W. V. Petryshyn, *On a class of K-p.d. and non K-p.d. operators and operator equations*, J. Math. Anal. Appl., **10** (1965), 1–24.
11. W. V. Petryshyn, *Projection methods in nonlinear numerical functional analysis*, Journal of Math. and Mech., **17** (1967), 353–372.
12. M. E. Titensky, *On the stability of the Galerkin method for certain coordinate systems*, Izv. Vyssh. Uchebn. Zaved. Mat., **6** (1987), 57–65 (Russian). Engish translation : Soviet Math (Iz VUZ) **31** (1987) no. 6, 73–82.

13. M. E. Titensky, *On the stability of the Ritz method*, The Institute of Scientific and Technical Information, Moscow, on May 18, no. 1782–79, (1979), 12 p. (Russian)
14. M. E. Titensky), *On some sufficient conditions of the stability of the Ritz and Bubnov-Galerkin methods*, Ibid., on May 19, no. 7924–80, (1980), 16 p. (Russian).

Received 20 VIII 1993

Department of Mathematics and Computer Science
Ben-Gurion University of the Negev, P.O.B. 653
84105, Beer-Sheva, Israel