# ON THE APPROXIMATION OF SYMMETRIC OPERATORS BY OPERATORS OF FINITE RANK

BY

SHMUEL KANIEL*

ABSTRACT

The following Theorem 1 and Theorem 2 are established. The proof utilizes the elementary properties of symmetric operators. It is shown that the symmetry condition is necessary for these theorems to hold.

In this work we shall prove the following theorem which stems from consideration of the minimal iteration method for eigenvalue problems ([1], [2]).

THEOREM 1. *Let $A$ be a bounded symmetric operator in Hilbert space $H$. Let $f \in H$ $f \neq 0$ be chosen and let $H_k$ denote the span of $f, Af, \cdots, A^k f$. Let $P$ be the orthogonal projection on $H_k$ and let $B$ denote the restriction of $PA$ to $H_k$.*

*Then there exists an absolute constant $c(k)$ tending to zero as $k \to \infty$ which does not depend on $A$ nor on $f$ so that for some $\lambda \in \sigma(A)$ and some $\mu \in \sigma(B)$:*

$$|\lambda - \mu| \leqq c(k) \|A\| \leqq \frac{2}{\sqrt{(k+1)}} \cdot \|A\|. \tag{1}$$

The proof of Theorem 1 depends on the following propositions:

PROPOSITION 1. *If $\|Af - \mu f\| \leqq \alpha \|f\|$, then there exists $\lambda \in \sigma(A)$ so that $|\lambda - \mu| \leqq \alpha$.*

This is a direct consequence of the spectral decomposition theorem.

PROPOSITION 2. *Let $\mu_i$ be the eigenvalues of $B$. Then $|\mu_i| \leqq \|A\|$.*

This is a direct consequence of the minimax principle. It is easy to see that $B$ is symmetric. Hence

$$\max_i |\mu_i| = \sup_{f \in H_k} \frac{(Bf,f)}{\|f\|} = \sup_{f \in H_k} \frac{|(PAf,f)|}{\|f\|} = \sup_{f \in H_k} \frac{|(Af,f)|}{\|f\|}$$
$$\leqq \sup_{f \in H} \frac{|(Af,f)|}{\|f\|} = \|A\|.$$

PROPOSITION 3. *If $p(x)$ is a polynomial having degree at most $k$, then $p(A)f = p(B)f$.*

Received November 11, 1964.

* This work was sponsored in part by NSF contract 2426.

**Proof.** For $i = 1, 2, \cdots, k$, $A^i f \in H_k$, so by a simple induction:

$$A^i f = PA^i f = PA \cdot A^{i-1} f = PA(PA)^{i-1} f = B^i f.$$

The passage to polynomials is trivial.

PROPOSITION 4. *If $B$ has a multiple eigenvalue, then all the eigenvalues of $B$ belong to $\sigma(A)$.*

**Proof.** We shall prove that in this case $H$ is an invariant subspace of $A$. So $B =$ restriction of $A$ to $H_k$ and hence $\sigma(B) \subset \sigma(A)$.

It is sufficient to prove that in this case the vectors $f, Af, \cdots, A^k f$, are dependent because then we have for some $l \leqq k-1$

$$A^{l+1} f = \sum_{i=0}^{l} \alpha_{il} A^i f \in H_k,$$

and we can prove inductively that:

$$A^j f = \sum_{i=0}^{l} \alpha_{ij} A^i f \in H_k, \; j = 1, 2, \cdots$$

$$\left( A^{j+1} f = AA^j f = A \sum_{i=0}^{l} \alpha_{ij} A^i f = \sum_{i=0}^{l-1} \alpha_{ij} A^{i+1} f + \alpha_{lj} \sum_{i=0}^{l} \alpha_{il} A^i f \right).$$

So let $g_1$ and $g_2$ be two different eigenvectors corresponding to the eigenvalue $\mu$, then some linear combination of $g_1$ and $g_2$ can be expressed by:

$$g = \sum_{i=0}^{k-1} \alpha_i A^i f, \qquad g \neq 0;$$

so:

$$0 = PAg - \mu g = PA \sum_{i=0}^{k-1} \alpha_i A^i f - \mu \sum_{i=0}^{k-1} \alpha_i A^i f = P \sum_{i=1}^{k} \alpha_{i-1} A^i f - \mu \sum_{i=0}^{k-1} \alpha_i A^i f$$

$$= \sum_{i=1}^{k-1} (\alpha_{i-1} - \mu\alpha_i) A^i f + \alpha_{k-1} A^k f - \mu\alpha_0 f.$$

Thus the linear dependence of $A^i f$, $i = 0, 1, \cdots, k$ is proved since not all coefficients in the last sum are zero. This leads to a contradiction and the proposition is established.

**Proof of Theorem** 1. In view of Proposition 4 it is sufficient to consider the case where $\mu_i$, the eigenvalues of $B$, are distinct. So let $g_i$, $i = 1, \cdots, k+1$ be an orthonormal base of eigenvectors corresponding to the eigenvalues $\mu_i$. Let $f = \sum_{i=1}^{k+1} \beta_i g_i$. We shall prove that $\beta_i \neq 0$. Indeed, $H_k$ is spanned by

$$A^j f = B^j f = \sum_{i=1}^{k+1} \beta_i \mu_i{}^j g_i, \quad j = 0, 1, \cdots, k.$$

If for some $m$: $\beta_m = 0$, it follows that $A^j f$ are orthogonal to $g_m$ or that $H_k$ is orthogonal to $g_m$ which is impossible.

Consider now the following polynomials:

$$q_j(x) = \sum_{\substack{n=1 \\ n \neq j}}^{k+1} c_{nj} \frac{\prod_{i \neq j,n} (x - \mu_i)}{\prod_{i \neq j,n} (\mu_n - \mu_i)}, \quad j = 1, 2, \cdots, k+1,$$

where $|c_{nj}| = |1/\beta_n|$ and the sign of $c_{nj}$ is chosen so that for any $n$ and $j$:

$$\beta_j c_{nj} \frac{\prod_{i \neq j,n} (\mu_j - \mu_i)}{\prod_{i \neq j,n} (\mu_n - \mu_i)} > 0.$$

The polynomials $q_j(x)$ satisfy:

$$(2) \qquad |\beta_n q_j(\mu_n)| = 1, \qquad n \neq j.$$

We shall prove that:

$$(3) \qquad \max_j |\beta_j q_j(\mu_j)| \geqq k.$$

In fact:

$$\sum_{j=1}^{k+1} \beta_j q_j(\mu_j) = \sum_{\substack{n,j=1 \\ n \neq j}}^{k+1} \beta_j c_{nj} \frac{\prod_{i \neq j,n} (\mu_j - \mu_i)}{\prod_{i \neq j,n} (\mu_n - \mu_i)}$$

$$\geqq (k+1)(k)^{(k+1)(k)} \sqrt{\sum_{\substack{n,j \neq 1 \\ n \neq j}}^{k+1} \frac{\beta_i}{\beta_n} \frac{\prod_{i \neq j,n} |\mu_j - \mu_i|}{\prod_{i \neq j,n} |\mu_n - \mu_i|}} = (k+1)(k),$$

because the expression under the root sign is symmetric in $j$ and $n$.

Suppose that $|\beta_m q_m(\mu_m)| \geqq k$ and consider the vector $g = q_m(A)f$. Since $q_m(x)$ has the degree $k-1$, it follows that $q_m(A)f = q_m(B)f$ and:

$$g = \sum_{i=1}^{k+1} q_m(\mu_i)\beta_i g_i.$$

By (2) and (3)

$$\|g\|^2 = \sum_{i=1}^{k+1} |q_m(\mu_i)\beta_i|^2 = \sum_{i \neq m} |q_m(\mu_i)\beta_i|^2 + |\beta_m q_m(\mu_m)|^2 \geqq k + k^2.$$

Consider now $Ag - \mu_m g$. Since the degree of $p(x) = (x - \mu_m)q_m(x)$ is $k$, we have:

$$\|Ag - \mu_m g\|^2 = \|(A - \mu_m I)q_m(A)f\|^2 = \|(B - \mu_m I)q_m(B)f\|^2$$

$$= \sum_{i=1}^{k+1} |(\mu_i - \mu_m) \cdot q_m(\mu_i) \cdot \beta_i|^2 \leqq k \max_i |\mu_i - \mu_m|^2 \leqq 4k\|A\|^2.$$

So:

$$\| Ag - \mu_m g \| \leqq \frac{2}{\sqrt{(k+1)}} \| A \| \| g \|$$

which by Proposition 1 means that there exists $\lambda \in \sigma(A)$ so that (1) holds. Thus the proof is complete.

REMARK. This estimate has no significance from the numerical analysis point of view. In general the estimate will be much better provided that some additional information about $\sigma(A)$ is given. For so-called "practical" estimates the reader is referred to [2].

Let us now apply Theorem 1 to get:

THEOREM 2. *Under the conditions of Theorem* 1 *the eigenvalues of PA can be ordered so that for any* $\mu_j$, $j = 1,2,\cdots,k$ *there exists* $\lambda(j) \in \sigma(A)$:

$$|\mu_j - \lambda(j)| \leqq \frac{2}{\sqrt{(k+2-j)}} \| A \|. \tag{4}$$

REMARK. This means that for large $k$ most of the eigenvalues of $B$ will be close to points in the spectrum of $A$. As in Theorem 1 the estimates depend on $k$ and $\| A \|$ only.

**Proof.** As in Theorem 1 it is sufficient to consider the case where the eigenvalues of $B$ are distinct. We shall prove the theorem by induction. Suppose that the theorem is true for $j = 1,2,\cdots,m$ and consider the vector

$$f_m = \prod_{j=1}^{m} (A - \mu_j I) f.$$

Denote by $H_k^m$ the span of $f_m, Af_m, \cdots, A^{k-m} f_m$; by $P_m$ the projection on $H_k^m$, and by $B_m$ the restriction of $P_m A$ to $H_k^m$.

By Theorem 1 there exists $\mu \in \sigma(B_m)$ and $\lambda \in \sigma(A)$ so that:

$$|\mu - \lambda| \leqq \frac{2}{\sqrt{k - m + 1}} \| A \|. \tag{5}$$

The induction will be complete if we show that $\mu \in \sigma(B)$ and $\mu \neq \mu_i, i = 1,2,\cdots,m$.*
Indeed, by Proposition 3 for $i = 0,\cdots,k-m$:

$$A^i f_m = A^i \prod_{j=1}^{m} (A - \mu_j I) f = B^i \prod_{j=1}^{m} (B - \mu_j I) f = \sum_{l=m+1}^{k+1} \beta_l \mu_l^i \prod_{j=1}^{m} (\mu_l - \mu_j) \cdot g_l.$$

This means that $A^i f_m$, $i = 0,1,\cdots,k-m$ are spanned by $g_j$, $j = m+1,\cdots,k+1$. $A^i f_m$ are independent because $A^i f$, $i = 0,1,\cdots,k+1$ are independent; therefore

* In this case we define $\mu = \mu_{m+1}$, establishing (4) for $m+1$.

by dimension argument $H_k^m$ is *the span* of $g_j$, $j = m+1,\cdots,k+1$. Hence $B$ is invariant in $H_k^m$ and the eigenvalues of $B_m$ coincide with $\mu_{m+1},\cdots,\mu_{k+1}$ which are different from $\mu_1,\cdots,\mu_m$. Thus the proof is complete.

The condition that $A$ is symmetric is necessary as the following example will show.

Define $H$ to be the space of double sequences $(\cdots,a_{-i},\cdots,a_0,\cdots,a_i,\cdots)$ so that $\sum_{i=-\infty}^{\infty}|a_i|^2 < \infty$. Define $A$ to be the left shift operator, i.e. if $\phi_i$ is a unit vector $(\cdots,0,\cdots,0,1_i,\cdots,0,\cdots)$ $-\infty < i < \infty$, then $A\phi_i = \phi_{i-1}$. $A$ is unitary so for any $\lambda \in \sigma(A)$: $|\lambda| = 1$.

Let $f = \phi_k$; then $H_k$ will be spanned by $\phi_0,\phi_1,\cdots,\phi_k$. The point 0 is the only point in $\sigma(B)$ ($\phi_0$ only is an eigenvector so $|0-\lambda| = 1$ contrary to the theorem's conclusion.

## References

1. D. K. Faddeev and V. N. Faddeeva, *Computational Methods of Linear Algebra*, Translated from Russian by R. C. Williams, W. H. Freeman and Co., San Francisco and London (1963).

2. S. Kaniel, *Estimates for some computational techniques used in linear algebra*, to appear.

STANFORD UNIVERSITY
STANFORD, CALIFORNIA