Convergence of approximate operator methods for eigenvectors

A. L. Andrew

This paper examines a large class of common numerical methods
for computing the eigenvectors of a compact linear operator.
Special cases include all projection methods and the Weinstein
intermediate problems method. Simple sufficient conditions are
established for the sequence of approximate eigenvectors
obtained by any of these methods to converge to an exact
eigenvector. In the most general case only the convergence of
a subsequence was previously known.

1.   Introduction

Approximate operator methods for numerical solution of the
eigenvalue problem

(1) $$T\psi = \mu\psi$$

in infinite dimensional space, are those methods in which successive
approximations $\{\mu_n, \psi_n\}$ to the solution $\{\mu, \psi\}$ are obtained by solving
simpler problems

(2) $$T_n\psi_n = \mu_n\psi_n \; ,$$

where the sequence of operators $\{T_n\}$ converges uniformly to T .
Important special cases are the various projection methods [1], [2] (see
Section 3) and the Weinstein intermediate problems method (and its

---

subsequent modifications) [3]. Convergence of the approximate
eigenvalues, μ_n , to solutions of (1) has been proved under very general
conditions [1; Lemmas 3.3 and 3.4], [2; Theorem 2], [3; Theorem 1. IX].
This paper strengthens a known result for convergence of the corresponding
eigenvectors, ψ_n .

The following notation is used here:

H is a (real or complex) infinite dimensional separable Hilbert
space;

E is the set of all unit vectors in H ;

Γ is the set of all scalars of unit modulus;

Z^+ is the set of all positive integers.

The following theorem is proved in Section 2, and some of its
implications noted in Section 3.

THEOREM 1. *Let* T_n *(n = 1, 2, ...) and T be compact linear*
operators in H . Let μ *be a simple nonzero eigenvalue of T*
satisfying (1) with $\psi \in E$. *For all n $\in Z^+$, let* μ_n, ψ_n *satisfy (2)*
with $\psi_n \in E$ *, and let*

(3) $\|T_n - T\| \to 0$ *and* $\mu_n \to \mu$ *as n $\to \infty$.*

Then there is a sequence of scalars $\alpha_n \in \Gamma$ *such that* $\|\alpha_n\psi_n - \psi\| \to 0$ *as*
n $\to \infty$.

In the case of projection methods, a weaker version of Theorem 1 is
well known. This weaker version requires the additional hypothesis that
all but finitely many members of the sequence $\{\mu_n\}$ are simple
eigenvalues of the corresponding T_n . As noted by Petryshyn [1; p. 431]
this weaker result follows immediately from a theorem of Polsky [2;
Theorem 4] concerning the invariant spaces of T and T_n . The proof
given here is an extension of that used for another of Polsky's theorems
[2; Theorem 3]. (Whereas Theorem 1 here proves the convergence of the
whole sequence, Polsky's Theorem 3 proved only the convergence of a

subsequence.)

The requirement that μ be simple is known to be satisfied for many important operators in mathematical physics. It may be less easy to check the extra requirement (that the μ_n be simple also) which is dispensed with by Theorem 1. Note that Theorem 1 does not require that *all* the eigenvalues of T be simple.

In one sense Theorem 1 is best possible, as Polsky [2; p. 61] showed that when the space, V_μ, of eigenvectors of T corresponding to μ, has dimension more than one, there may be vectors in V_μ which cannot be obtained as the limit of any sequence of eigenvectors of the T_n. However when the dimension of V_μ is greater than one, analogues of Theorem 1 may be proved provided the T_n satisfy certain additional hypotheses. These results are stronger than that proved by Polsky [2; Theorem 3] and that noted as a remark in Section 2. It is hoped to return to this matter in a later paper.

In the special case of the Ritz method, an important quantitative version of Theorem 1 has been proved by Weinberger [4]. His proof uses the fact that in this special case T must be hermitian.

2. Proof of Theorem 1

DEFINITION 1. For all r in Z^+, define S_r to be the sequence obtained by deleting from $\{\psi_n\}$ those elements for which

(4) $(\exists \alpha \in \Gamma)(\forall \omega \in E)|(\alpha\psi_n - \psi, \omega)| \leq 2^{-r}$.

LEMMA 1. *For all r in Z^+, S_r contains at most finitely many elements.*

Proof. Suppose the contrary. Then, since the ψ_n are bounded, S_r contains an infinite subsequence $S_r^* = \{\psi_{\nu(n)}\}$ converging weakly to an element $\psi_0 \in H$, and by (3), $\|T_{\nu(n)} - T\| \to 0$ and $\mu_{\nu(n)} \to \mu$ as

$n \to \infty$. Hence, since $\mu_{\nu(n)}\psi_{\nu(n)} = T_{\nu(n)}\psi_{\nu(n)}$,

$$\left\| T\psi_0 - \mu_{\nu(n)}\psi_{\nu(n)} \right\| \leq \left\| T\psi_0 - T\psi_{\nu(n)} \right\| + \left\| T - T_{\nu(n)} \right\| \left\| \psi_{\nu(n)} \right\| \to 0 ,$$

since T is compact and $\left\| \psi_{\nu(n)} \right\| = 1$. Hence $\mu_{\nu(n)}\psi_{\nu(n)} \to T\psi_0$.
Moreover, since $\mu_{\nu(n)} \to \mu$ and $\psi_{\nu(n)} \rightharpoonup \psi_0$, $\mu_{\nu(n)}\psi_{\nu(n)} \rightharpoonup \mu\psi_0$. Since
the weak limit is unique it follows that

(5) $$\mu_{\nu(n)}\psi_{\nu(n)} \to T\psi_0 = \mu\psi_0 .$$

Hence $\psi_0 \in V_\mu$.

Since $\mu_{\nu(n)} \to \mu \neq 0$, it follows from (5) that $\psi_{\nu(n)} \to \psi_0$ and that
$\psi_0 \in E$. Hence for all α in Γ , and all w in E ,
$\left| \left(\alpha\psi_{\nu(n)} - \alpha\psi_0, w \right) \right| \leq \left\| \psi_{\nu(n)} - \psi_0 \right\| \to 0$ as $n \to \infty$. It follows that

(6) $(\forall \alpha \in \Gamma)\left(\exists \psi_{\nu(n)} \in S_r^*\right)(\forall w \in E)\left| \left(\alpha\psi_{\nu(n)} - \alpha\psi_0, w \right) \right| < 2^{-r}$.

Since $S_r^* \subset S_r$, it follows from Definition 1 that

(7) $(\forall \alpha \in \Gamma)\left(\forall \psi_{\nu(n)} \in S_r^*\right)(\exists w \in E)\left| \left(\alpha\psi_{\nu(n)} - \psi, w \right) \right| > 2^{-r}$.

Combining (6) and (7) shows that

(8) $(\forall \alpha \in \Gamma)\alpha\psi_0 \neq \psi$.

Since the eigenvectors ψ_0 and ψ are both in E , it follows from (8)
that they are linearly independent. This contradicts the fact that μ is
a simple eigenvalue of T . Lemma 1 follows.

Since clearly $S_r \subset S_{r+1}$, it follows from Lemma 1 that there is an
infinite sequence of integers $n_1 < n_2 < \ldots$ such that for all $r \in Z^+$
and all integers $n > n_r$, $\psi_n \notin S_r$. It follows from the definitions of
n_r and S_r that there is a sequence $\{\alpha_n\}$ of scalars in Γ such that
if $n_r < n \leq n_{r+1}$ then $\left| \left(\alpha_n\psi_n - \psi, w \right) \right| \leq 2^{-r}$ for all w in E . Since
$r \to \infty$ as $n \to \infty$, it follows that $\alpha_n\psi_n \rightharpoonup \psi$. Hence by (1) and (2)

$$|\mu| \, \left\| \psi - \alpha_n \psi_n \right\| = \left\| T\psi - \mu\alpha_n \psi_n \right\| \leq \left\| T\psi - T\alpha_n \psi_n \right\| + \left\| T - T_n \right\| \left\| \alpha_n \psi_n \right\| + |\mu_n - \mu| \left\| \alpha_n \psi_n \right\| \to 0$$

by (3), since T is compact and $\left\| \alpha_n \psi_n \right\| = 1$. Since $\mu \neq 0$, this completes the proof of Theorem 1.

REMARK. When the dimension of V_μ is greater than one, a slight modification of the above proof shows that the orthogonal projection of ψ_n on the orthogonal complement of V_μ still converges strongly to zero as $n \to \infty$. The analogue of (4) in the modified proof is

$$\left(\exists \varphi \in V_\mu \cap E \right) \left(\forall w \in E \right) \left| \left(\varphi - \psi_n, \, w \right) \right| \leq 2^{-r} .$$

An argument exactly analogous to that used in the proof of Theorem 1 shows that there is a sequence $\{\varphi_n\}$ of elements in $V_\mu \cap E$ such that $\psi_n - \varphi_n \to 0$ as $n \to \infty$. The result follows.

3. Applications

Let $\{v_r\}$ be a sequence of linearly independent vectors complete in H. Let U_n be the space spanned by the first n vectors in the sequence and P_n the orthogonal projector of H onto U_n. Projection methods for solving (1) are those methods in which approximations for μ and ψ are obtained by solving

$$(9) \qquad P_n T\psi_n = \mu_n \psi_n , \quad P_n \psi_n = \psi_n$$

in U_n. Important special cases are the Ritz method, the generalised Ritz method, the Galerkin method and the method of moments [1]. Since $\{v_r\}$ is a complete sequence and T is compact, it can be shown [1] that $\{P_n T\}$ converges uniformly to T, so that all projection methods for solving (1) with T compact are special cases of the approximate operator methods considered here. Polsky [2; Theorem 2] showed that the set of limit points of all sequences of eigenvalues of the $P_n T$ is precisely the set of all eigenvalues of T. It follows from Theorem 1 that, if the eigenvalues of T are all simple, the set of all limit points of the

corresponding sequences of appropriately normalized eigenvectors of the $P_n T$ is precisely the set of all normalized eigenvectors of T . In the special case where T is hermitian and strictly positive, it is known [3] that the sequence whose n-th member is the k-th largest eigenvalue of $P_{n+k} T$ converges monotonically to the k-th largest eigenvalue of T , so that it is easy to identify the corresponding convergent sequence of approximate eigenvectors, even when convergence is slow.

The commonest applications of projection methods involve a differential operator T^{-1} with compact inverse T . Since the eigenvalue problem $T^{-1}\psi = \mu\psi$ may be considered as $T(T^{-1}\psi) = \mu^{-1}(T^{-1}\psi)$, the theory developed for compact operators is readily extended to such problems.

In addition to the projection methods, there are other important methods to which Theorem 1 applies, including some in which the T_n operate on infinite dimensional spaces (see, for example, [3; Theorem 3. III]).

References

[1]   W.V. Petryshyn, "On the eigenvalue problem $Tu - \lambda Su = 0$ with unbounded and nonsymmetric operators T and S", *Philos. Trans. Roy. Soc. London Ser. A* **262** (1967/68), 413-458.

[2]   N.I. Polskiĭ, "On the convergence of certain approximate methods of analysis", (Russian), *Ukrain. Math. Ž.* **7** (1955), 56-70.

[3]   Gaetano Fichera, "Approximation and estimates for eigenvalues", *Proc. Sympos. Univ. Maryland,* (1965), 317-352. *Numerical solution of partial differential equations*, edited by James H. Bramble (Academic Press, New York, London, 1966).

[4] H.F. Weinberger, "Error bounds in the Rayleigh-Ritz approximation of
          eigenvectors", *J. Res. Nat. Bur. Standards Sect. B* 64B (1960),
          217-225.

La Trobe University,
Bundoora, Victoria.