Andrzej Pokrzywa – Seminar note

A posteriori eigenvalue error estimations for nonselfadjoint operator approximation

Introduction. In the papers [1, 2, 3] the eigenvalue problem is studied:

Differential problem. Find $u \neq 0$, $\lambda \in C$ such that

$$-\Delta u + \beta \cdot \nabla u = \lambda u \quad \text{in } \Omega$$
$$u(x) = 0 \quad \text{for } x \in \partial\Omega$$

where $\operatorname{div} \beta = 0$.

This problem has a variational formulatione find $u \in V = H_0^1(\Omega) \subset L^2(\Omega) = H$ such that

$$a(u, \phi) = \lambda(u, \phi) \quad \forall \phi \in V,$$

where

$$a(u, v) = \int_\Omega (\nabla u \cdot \nabla \bar{v}) + (\beta \cdot \nabla u)\bar{v},$$

and $(\cdot, \cdot)$ is the scalar product in H.

Because

$$(\beta \cdot \nabla u, v) = -(v, \beta \cdot v)$$

hence

$$\|u\|_V^2 = \int_\Omega (\nabla u \cdot \nabla \bar{b}) = Re\, a(v, v).$$

The dual problem consists in finding λ^*, u^* such that

$$a(\phi, u^*) = \bar{\lambda}^*(\phi, u^*) \quad \forall \phi \in V$$

The authors do not define explicitly operators in V, this will be done in this paper. The solution u of the variational problem: for $f \in H$ find $u \in V$ such that

$$a(u, \phi) = (f, \phi) \quad \forall \phi \in V$$

$u = Af$ depends linearly on f. The equality

$$a(Af, Af) = (f, Af)$$

implies that

$$\|Au\|_V^2 = Re\, a(Af, Af) \leq |(f, Af)| \leq \|f\|\|Af\| \leq \|f\|\|Af\|_V.$$

Therefore

$$\|Au\|_V \leq \|f\| \leq \|f\|_V.$$

This shows that $A : V \to V$ is a bounded operator, and because the inclusion V in H is compact the operator A is compact.

If $Au = \lambda u$ then the equality $f = \lambda^{-1}u$ implies that $Af = u$ and we see that

$$a(u, \phi) = (\lambda^{-1}u, \phi) = \lambda^{-1}(u, \phi).$$

Thus the eigenvalues of the form a are reciprocals of the eigenvalues of the operator A while eigenvectors remain the same.

What is the dual problem ?

Similary as A we define the operator $C : V \to V$ by the equality

$$a(\phi, Cg) = (\phi, g) \quad \forall \phi \in V$$

Is it true that $A^* = C$?

As for $u \in V$ $a(u, \cdot)$ is a continous antylinear functional $a(u, \cdot)$ equals to the scalar product in V whith some vector in V, which we denote by Bu

$$a(u, v) = \langle Bu, v\rangle \quad u, v \in V \tag{1}$$

($\langle \cdot, \cdot \rangle$ denotes scalar product in V)
Because the form a is bounded we have

$$\|B\| \le \|a\| = \sup\{|a(u, v); u, v \in V, \|u\|_V \le 1, \|v\|_V \le 1\}$$

Similarily for $u \in V$ $(u, \cdot)$ is a continous antylinear functional, and may be interpreted as a scalar product in V. Thus there exists linear operator T, such that

$$(u, v) = \langle Tu, v\rangle \quad u, v \in V.$$

Because $\langle Tu, u\rangle = \|u\|^2 \ge 0$,we have

$$T = T^* \ge 0$$

Having defined the above operators we may find new definitions of A i C. The identities

$$a(Af, \phi) = (f, \phi), \quad a(\phi, Cg) = (\phi, g) \quad \forall \phi \in V$$

are equivalent to

$$\langle BAf, \phi\rangle = \langle Tf, \phi\rangle, \quad \langle B\phi, Cg\rangle = \langle T\phi, g\rangle \quad \forall \phi \in V$$

This holds for all $f, g \in V$ and therefore we have the equalities:

$$BA = T, \quad C^*B = T.$$

Hence

$$T = BA = C^*B \quad \text{and } A = B^{-1}C^*B$$

The operator A is similar to C^*, or A^* is similar to C. Simple statements concerning the eigenvalues and eigenvectors of these operators may be formulated.

Approximation.

Let $V_h \subset V$ be a subspace of V.

Approximaton problem in terms of the form a is the following:
find $v_h \in V_h$ such that

$$a(v_h, \phi) = (f, \phi) \quad \forall \phi \in V_h$$

If v is the solution of the original problem this is equivalent to

$$a(v_h, \phi) = a(v, \phi) \quad \forall \phi \in V_h$$

The transformation $v \to v_h$ is linear and is a projection, let us denote it by P. We have $P^2 = P$, $\operatorname{ran} P = V_h$.

Thus

$$a(Pv, \phi) = a(v, \phi) \quad \forall \phi \in V_h, \forall v \in V$$

or

$$a(Pv, P\phi) = a(v, P\phi) \quad \forall \phi, v \in V$$

Using (1) we have

$$\langle P^* BPv, \phi\rangle = \langle BPv, P\phi\rangle = \langle Bv, P\phi\rangle = \langle P^* Bv, \phi\rangle$$

This means $P^*BP = P^*B$, taking adjoint operators we have

$$P^* B^* P = B^* P$$

Thus $\operatorname{ran} P^* = B^* V_h$.

We may pose the questions:

a) Does $\operatorname{ran} P$, $\operatorname{ran} P^*$ are sufficient for defining P, in particular, is the role of B essential here?

b) If yes, are then the projection defined uniequely?

A little bit about projections

The following fact may be usefull: if P is a projection (finite-dimensional) the there exists orthonormal sequences $e_1, \ldots, e_n$, $f_1, \ldots, f_n$ such that

$$P = \sum_1^n \lambda_j \langle \cdot, f_j\rangle e_j \tag{2}$$

where $\lambda_h \langle e_j, f_i\rangle = \delta_{i,j}$.

Proof. Let $P = UA$ be a polar decoompoosition of P, i.e. U is a unitary operator, $A = A^* > 0$. Let $A = \sum \lambda_j \langle \cdot, f_j\rangle f_j$ be the spectral representaton of A, (f_j are orthonormal eigenvectors, λ_j – eigenvalues of A). As U is unitary the vectors $e_j = Uf_j$ $j = 1, \ldots, n$

form also an orthonormal sequence. Now the equality $e_k = Pe_k = \sum_1^n \lambda_j \langle e_k, f_j\rangle e_j$ and the linear independence of e_j implies that $\lambda_k\langle e_k, f_k\rangle = 1$, and $\langle e_k, f_j\rangle = 0$ for $j \neq k$.

Does $\operatorname{ran} P$, $\operatorname{ran} P^*$ define P uniquely ?

If there exists a projection Q such that

$$\operatorname{ran} Q = \operatorname{ran} P, \tag{3}$$

then

$$PQ = Q, \quad QP = P$$

If

$$\operatorname{ran} Q^* = \operatorname{ran} P^* \tag{4}$$

then

$$P^*Q^* = Q^*, \quad Q^*P^* = P^*$$

if (3) i (d) hold then

$$P = (P^*)^* = (Q^*P^*)^* = PQ = Q.$$

Next question: If we have to subspaces M, N such that $\dim M = \dim N$ does there exist a projection P such that $\operatorname{ran} P = M$, $\operatorname{ran} P^* = N$?

Simple counterexample: let $e \perp f$ span respectively the subspaces M, N. Then the projection P may be represented by $P = \alpha\langle\cdot, f\rangle e$ with some α. But then $Pe = 0$.

What condition on M, N assures the existence of a desired projection? Any projection P has a representation (2). Then $(\langle e_i, f_j\rangle)_{i,j=1}^n$ is a nonsingular matrix. If in this matirx the vectors e_i are substitued by any other basis $\tilde{e}_i$, then the columns of the matrix $(\langle \tilde{e}_i, f_j\rangle)_{i,j=1}^n$ are linear combination of the previous matrix. If we repeat the same with f_j then the new matrix rows are linear combinatons of the previously constucted matrix. Any way the last matrix remains nonsingular. Thus the condition is that for some basis $\{e_j\}_1^n$ of M, $\{f_j\}_1^n$ of N the matrix $(\langle e_i, f_j\rangle)_{i,j=1}^n$ is nonsingular.

The next question (more closer to the main problem) is: How to charecterize the operators B such that for any subspace M there exists a projection P such that

$$\operatorname{ran} P = M, \quad BP = P^*BP, \tag{5}$$

i.e. $\operatorname{ran} P^* = BM$.

If the subspace M is one dimensional, spanned by the vector e the condition is $\langle Be, e\rangle \neq 0$, or $0 \notin W(B)$, where $W(B)$ denotes the numerical range of B. The numerical range is a convex set, therefore there exists a straight line in the complex plane such that the point 0 and $W(B)$ lie on other sides of this line. Thus there exists a complex number α such that

$$Re\,\alpha\langle Be, e\rangle \geq 1 \quad \forall e \in V, \|e\| = 1. \tag{6}$$

We can say, that up to some multiplicative complex constant α these are the operators B for which $Re\,B = \frac{1}{2}(B + B^*) > I$. The condition (6) is a necessary one. We shall show it is also sufficient. Let $\{e_j\}_1^n$ be an orthonormal basis in M. It suffices to show that the matrix $(\langle Be_i, e_j\rangle)_{i,j=1}^n$ is nonsingular. This matrix represents the operator $QB|_M$, where Q is an orthogonal projection on M. The desired conclusion is implied by

$$\sigma(QB|_M) \subset W(QB|_M) \subset W(B)$$

Note also, that if $Re\, B > 1$ and P satisfies (5) then

$$\begin{aligned}\|Px\|^2 &\leq \langle Re\, BPx, Px\rangle = Re\langle BPx, Px\rangle \leq |\langle BPx, Px\rangle| \\ &= |\langle P^*BPx, x\rangle| = |\langle BPx, x\rangle| \leq \|B\|\|Px\|\|x\|\end{aligned}$$

therefore $\|P\| \leq \|B\|$, thus all such projections a uniformly bounded.

We have thus shown.

Lemma 1 *If $Re\, B \geq 1$ then for any subspace $M \subset V$ there exists exactly one projection P such that* $\operatorname{ran} P = V$, $\operatorname{ran} P^* = BV$. *Moreover* $\|P\| \leq \|B\|$.

Left multiplying the equality $BP = P^*BP$ by B^{-1} and setting $\bar{P} = B^{-1}P^*B$ we get $P = \bar{P}P$. Hence $\operatorname{ran} P \subset \operatorname{ran} \bar{P}$, and at least in the finite dimensional case $\dim P = \dim \bar{P}$, what implies $\operatorname{ran} P = \operatorname{ran} \bar{P} = M$. Is it true that $P = \bar{P}$? We check it with one dimensional subspace M spanned by e. We search for f such that $P = \langle \cdot, f\rangle e$ satisfies (5). We count $Px = \langle x, f\rangle e$, $BPx = \langle x, f\rangle Be$, $P^*BPx = \langle x, f\rangle\langle Be, e\rangle f$. Thus

$$f = \frac{Be}{\langle Be, e\rangle}, \quad P^*Bx = \langle Bx, e\rangle f = \frac{\langle Bx, e\rangle}{\langle Be, e\rangle} Be,$$

and

$$\bar{P}x = B^{-1}P^*Bx = \frac{\langle Bx, e\rangle}{\langle Be, e\rangle} e, \quad Px = \langle x, f\rangle e = \left\langle x, \frac{Be}{\langle Be, e\rangle}\right\rangle e = \frac{\langle x, Be\rangle}{\langle e, Be\rangle} e$$

If $P = \bar{P}$ then

$$\frac{\langle Bx, e\rangle}{\langle Be, e\rangle} = \frac{\langle x, Be\rangle}{\langle e, Be\rangle} \quad \forall x \in V$$

or equivalently

$$\frac{B^*e}{\langle Be, e\rangle} = \frac{Be}{\langle e, Be\rangle}$$

Thus not always $P = \bar{P}$, but when $B = B^*$ then $P = \bar{P}$, not only in the case of onedimensional projection. Moreover for the projection P satisfying (5) we have $P^*B = (P^*BP)^* = P^*BP = BP$, and $\bar{P} = B^{-1}P^*B = P$.

Lemma 2 *If $\{P_n\}$ is a sequence of projections such that*
a) $P_n \to I$ strongly,
b) $\operatorname{ran} P_n^* = \operatorname{ran} BP_n$,
where B is a bounded operator such that $Re\, B \geq I$ then $P_n^* \to I$ strongly.

Proof Fix x, and let $x_n = BP_nB^{-1}x$, then

$$\|x_n - x\| = \|B(P_nB^{-1}x - B^{-1}x)\| \le \|B\|\|(P_n - I)B^{-1}x\| \to 0$$

Note that $x_n \in \operatorname{ran} BP_n = \operatorname{ran} P_n^*$, therefore $x_n = P_n^* x_n$. This and Lemma 1 imply

$$\|P_n^*x - x\| = \|P_n^*(x - x_n) - (x - x_n)\| \le (\|P_n^*\| + 1)\|x - x_n\| \le (\|B\| + 1)\|x - x_n\| \to 0.$$

Known Fact 1. *If $E = E^2$ and $E \ne 0$, $E \ne I$ then $\|E\| = \|I - E\|$.*

We can represent E as a matrix

$$E = \begin{pmatrix} I & A \\ 0 & 0 \end{pmatrix}, \quad I - E = \begin{pmatrix} 0 & -A \\ 0 & I \end{pmatrix}$$

then

$$EE^* = \begin{pmatrix} I & A \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I & 0 \\ A^* & 0 \end{pmatrix} == \begin{pmatrix} I + AA^* & 0 \\ 0 & 0 \end{pmatrix},$$
$$(I - E)^*(I - E) = \begin{pmatrix} 0 & 0 \\ -A^* & I \end{pmatrix}\begin{pmatrix} 0 & -A \\ 0 & I \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & I + A^*A \end{pmatrix}.$$

Hence

$$\|E\|^2 = \|I + AA^*\| = \|I + A^*A\| = \|I - E\|^2.$$

Moreover we see that the singular values exeeding 1 of the projections E i $I - E$ are the same. The above beutiful proof of Corach may be found in [4].

Approximation error of eigenvalules

Let λ_0 be an isolated eigenvalue of the operator A with finite geometrical multiplicity n_0, and let Γ be a closed curve (for example circle) inside which lies λ_0 and outside the rest of the spectrum of A. Let $M = \sup_{\lambda\in\Gamma} \|(A - \lambda)^{-1}\|$.

If $\|A - B\| < M^{-1}/2$ then the operator $B - \lambda$ is invertible for $\lambda \in \Gamma$, because

$$B - \lambda = A - \lambda + (B - A) = (A - \lambda)(I - (A - \lambda)^{-1}(B - A))$$

hence

$$(B - \lambda)^{-1} = \sum_{k=0}^{\infty} \left((A - \lambda)^{-1}(B - A)\right)^k (A - \lambda)^{-1}$$

This implies

$$\|(B - \lambda)^{-1}\| \le \|(A - \lambda)^{-1}\| \frac{1}{1 - \|(A - \lambda)^{-1}\|\|B - A\|} \le 2M$$

The identity

$$(B - \lambda)^{-1} - (A - \lambda)^{-1} = \sum_{k=1}^{\infty} \left((A - \lambda)^{-1}(B - A)\right)^k (A - \lambda)^{-1}$$

implies that

$$\begin{aligned}\|(B-\lambda)^{-1}-(A-\lambda)^{-1}\| &\le \sum_{k=1}^{\infty} M^{k+1}\|B-A\|^k = M^2\|B-A\|\sum_{k=0}^{\infty} M^k\|B-A\|^k \\ &\le 2M^2\|B-A\|\end{aligned}$$

Denoting by E, F spectral projections defined by

$$E = \frac{-1}{2\pi i}\int_\Gamma (A-\lambda)^{-1}, \quad F = \frac{-1}{2\pi i}\int_\Gamma (B-\lambda)^{-1}$$

we can estimate the norm of their difference

$$\|E-F\| \le \frac{1}{\pi}|\Gamma| M^2\|B-A\|. \tag{7}$$

We shall use the trace properties: $\operatorname{tr} EAE = n_0\lambda_0$ and that $\operatorname{tr} FBF$ is the sum of the eigenvalues of B which lie inside Γ and are counted according to their geometrical multiplicities. Thus the quantity $\operatorname{tr}(EAE - FBF)$ measures the perturbation λ_0.

We have the identity

$$\begin{aligned}(E-F)^2E &= (E+F-EF-FE)E = (E+FE-EF-FE)E \\ &= E(E-F)E = E(E-F)^2\end{aligned} \tag{8}$$

We shall answer the question: *What condition should satisfy the projections* $\{Q_n\}_1^\infty$, $\{P_n\}_1^\infty$ *that for any compact operator* A $\|A - Q_nAP_n\| \to 0$ *?*

It is rather obvious, that the norms of these projections should be uniformly bounded by some constant m. If A is compact, then for every $\varepsilon > 0$ there exist operators C, D such that $A = C + D$, $\|C\| \le \frac{\varepsilon}{2m}$, and D is finite dimensional. Obviousely $\|Q_nCP_n\| \le \varepsilon/2$. Because any finite dimensional operator is a sum of one dimensional operators, and each such operator G may be written as $G = \langle\cdot, f\rangle e$ our question restricts to: *What conditions the projection* $\{Q_n\}_1^\infty$, $\{P_n\}_1^\infty$ *should satisfy that for any operator* $G = \langle\cdot, f\rangle e$ $\|G - Q_nGP_n\| \to 0$ *?*

As

$$Gx - Q_nGP_nx = \langle x, f\rangle - \langle P_nx, f\rangle Q_ne = \langle x, f\rangle - \langle x, P_n^*f\rangle Q_ne$$

thus for any $f, e \in V$ we have to have $\|f - P_n^*f\| \to 0$, $\|e - Q_ne\| \to 0$, this means that the operators P_n^*, Q_n should strongly converge to the identity operator.

If $B = PAP$ and the projections E, F are sufficiently close we may derive much simpler estimate of $\operatorname{tr}(EAE - FBF)$. Because P commutes with B then also $PF = FP$, we have also $\operatorname{ran} F \subset \operatorname{ran} P$ hence $PF = FP = F$. Therefore

$$\operatorname{tr} FBF = \operatorname{tr} FPAPF = \operatorname{tr} FAF. \tag{11}$$

using the splitting $A = EAE + (I-E)A(I-A)$ we can write

$$\operatorname{tr} FBF = \operatorname{tr} F(EAE + (I-E)A(I-A))F. \tag{12}$$

We calculate the traces $\operatorname{tr}(FEAEF - EAE)$ i $\operatorname{tr} F(I-E)A(I-A)F$.

$$\operatorname{tr}(FEAEF - EAE) = \operatorname{tr}(FEA - EA)E = \operatorname{tr}(F-E)EAE = \operatorname{tr} E(F-E)EA \qquad (13)$$

and

$$\operatorname{tr}(F(I-E)A(I-E)F = \operatorname{tr}(I-E)F(F-E)A = \operatorname{tr}(F-E)F(F-E)A. \qquad (14)$$

(8) – (14) imply now the estimation

$$|\operatorname{tr}(FBF - EAE)| \leq \|E-F\|^2\|A\|(\|F\| + \|E\|). \qquad (15)$$

Let us resume our consideration in the way usefull for further ones.

Lemma3. *If A is a compact operator, $\lambda \neq 0$ its eigenvalue, with spectral projection E, and $\{P_h\}$ is a sequence of projections such that,*
$P_h \to I$ and $P_h^* \to I$ strongly,
$\|(I-P_h)|_{\operatorname{ran} E}\| \leq h$,
$\|(I-P_h^*)|_{\operatorname{ran} E}\| \leq h$,
$\|(I-P_h)|_{\operatorname{ran} E^*}\| \leq h$,
$\|(I-P_h^*)|_{\operatorname{ran} E^*}\| \leq h$
then for any $r > 0$ such that $\sigma(A) \cap K(\lambda, r) = \{\lambda\}$ and sufficiently small h the disc $K(\lambda, r)$ contains exactly $\dim E$ eigenvalues of the operator $P_h A|_{\operatorname{ran} P_h}$ counted according to their multiplicities, and their arithmetic mean λ_h satisfies $|\lambda - \lambda_h| \leq O(h^2)$.

Approximating problems and projections

Let us come back to our problems - primal and dual. We use the notation B^{-*} for $(B^*)^{-1}$.

Primal problem	**Dual problem**
find $v \in V$ such that	find $w \in V$ such that
$a(v, \psi) = (f, \psi), \quad \forall \psi \in V$	$a(\psi, w) = (\psi, g) \quad \forall \psi \in V$

we define operators A, B by

$\langle BAf, \psi \rangle = (f, \psi), \quad \forall \psi \in V$	$\langle B\psi, Cg \rangle = (\psi, g) \quad \forall \psi \in V$

approximating problems

find $v_h \in V_h$ such that	find $w_h \in V_h$ such that
$a(v_h, \psi) = (f, \psi), \quad \forall \psi \in V_h$	$a(\psi, w_h) = (\psi, g) \quad \forall \psi \in V_h$

or equivalently

$a(v_h, \psi) = a(v, \psi), \quad \forall \psi \in V_h$	$a(\psi, w_h) = a(\psi, w) \quad \forall \psi \in V_h$

we define operators acting in V_h

$$A_h : V_h \to V_h \qquad \Big| \qquad C_h : V_h \to V_h$$

by

$$\langle BA_h v, \psi\rangle = \langle BAv, \psi\rangle, \quad \forall \psi \in V_h \qquad \Big| \qquad \langle B\psi, C_h w\rangle = \langle B\psi, Cw\rangle, \quad \forall \psi \in V_h$$

if P_h, Q_h are projections on V_h we have

$$\langle BA_h v, P_h\psi\rangle = \langle BAv, P_h\psi\rangle, \quad \forall \psi \in V \qquad \Big| \qquad \langle BQ_h\psi, C_h w\rangle = \langle BQ_h\psi, Cw\rangle, \quad \forall \psi \in V$$

we may assume that

$$A_h : V_h \to V \qquad \Big| \qquad C_h : V_h \to V$$

$$A_h = P_h A_h \qquad \Big| \qquad C_h = Q_h C_h$$

then

$$P_h^* BA_h = P_h^* BA : V_h \to V \qquad \Big| \qquad Q_h^* B^* C_h = Q_h^* B^* C : V_h \to V$$

we may require the projections to satisfy

$$\operatorname{ran} P_h^* = BV_h \qquad \Big| \qquad \operatorname{ran} Q_h^* = B^* V_h$$

then according to (5)

$$P_h^* BP_h = BP_h \qquad \Big| \qquad Q_h^* B^* Q_h = B^* Q_h$$

$$P_h^* BA = P_h^* BP_h A_h = BP_h A_h \qquad \Big| \qquad Q_h^* B^* C = Q_h^* B^* Q_h C_h = B^* Q_h C_h$$

left multiplying by operator B^{-1} or B^{-*} we get

$$B^{-1} P_h^* BA = P_h A_h \qquad \Big| \qquad B^{-*} Q_h^* B^* C = Q_h C_h$$

the operator

$$\tilde{P}_h = B^{-1} P_h^* B \qquad \Big| \qquad \tilde{Q}_h = B^{-*} Q_h^* B^*$$

is a projection and its easy to check that

$$\operatorname{ran} \tilde{P}_h = V_h \qquad \Big| \qquad \operatorname{ran} \tilde{Q}_h = V_h$$

because of

$$\tilde{P}_h^* = B^* P_h B^{-*} \qquad \Big| \qquad \tilde{Q}_h^* = BQ_h B^{-1}$$

we have

$$\operatorname{ran} \tilde{P}_h^* = B^* \operatorname{ran} P_h = B^* V_h = \operatorname{ran} Q_h^* \qquad \Big| \qquad \operatorname{ran} \tilde{Q}_h^* = B \operatorname{ran} Q_h = BV_h = \operatorname{ran} P_h^*$$

this and Lemma 1 imply

$\tilde{P}_h = Q_n = B^{-1}P_h^*B$	$\tilde{Q}_h = P_h = B^{-*}Q_h^*B^*$
$Q_hA = P_hA_h = A_h$	$P_hC = Q_hA_h = C_h$

we extend the operators for all the space V identyfying

A_h z A_hQ_h	C_h z C_hP_h

then

$A_h = Q_hAQ_h$	$C_h = P_hCP_h$

Approximating operators are

$$A_h = Q_hAQ_h, \quad C_h = P_hCP_h$$

where P_h, Q_h are projections on V_h such that

$$\operatorname{ran} P_h^* = BV_h, \quad \operatorname{ran} Q_h^* = B^*V_h.$$

$A = B^{-1}C^*B$ implies $A^* = B^*CB^{-*}$ and

$$\begin{aligned} A_h^* &= Q_h^*A^*Q_h^* = Q_h^*B^*CB^{-*}Q_h^* = B^*(B^{-*}Q_h^*B^*)C(B^{-*}Q_h^*B^*)B^{-*} \\ &= B^*P_hCP_hB^{-*} = B^*C_hB^{-*} \end{aligned}$$

Now we see, that with additional approximation properties of the subspaces V_h we may use Lemma 3 to estimate the approximation error of eigenvalues.

A posteriori estimates

Example 1. If $A = A^*$ is a compact operator it may be represented by a series $A = \sum_j \lambda_j \langle \cdot, e_j \rangle e_j$, where λ_j are eigenvalues corresponding to eigenvectors e_j. Then for any $\lambda \in C$, $e \in V$ z $\|e = \| = 1$

$$\begin{aligned} \|(A-\lambda)e\|^2 = \|\sum_j(\lambda_j - \lambda)\langle e, e_j\rangle e_j\|^2 &= \sum_j |\lambda_j - \lambda|^2 |\langle e, e_j\rangle|^2 \\ &\geq \min_j |\lambda_j - \lambda|^2 |\sum_j |\langle e, e_j\rangle|^2 = \min_j |\lambda_j - \lambda|^2. \end{aligned}$$

If λ, e are counted by some approximation algorithm and we have a possibility to calculate $\|Ae\|$ we get in this way a posteriori estimation of eiganvalue error. In practise it suffices to count Ae with greater precision.

Example 2. Let $A = E = \langle \cdot, f \rangle g$, where $\lambda_0 = \langle g, f \rangle = 1$ is the eigenvalue of A. Let P_h be an orthogonal projection. $A_h = P_h A P_h = \langle \cdot, P_h f \rangle P_h g$ is an approximating operator. The only nonzero eigenvalue of it is $\lambda_h = \langle Pg, Pf \rangle = \langle g, Pf \rangle$, what follows from $A_h Pg = \langle Pg, Pf \rangle Pg$. The normed eigenvector is $u_h = \frac{Pg}{\|Pg\|}$ and we can write

$$(A - \lambda_h)u_h = \langle u_h, f \rangle g - \lambda_h u_h = \|Pg\|^{-1}(\langle Pg, f \rangle g - \lambda_h Pg) = \|Pg\|^{-1}\lambda_h(g - Pg)$$

$$\|(A - \lambda_h)u_h\| = \|Pg\|^{-1}|\lambda_h|\|(g - Pg)\|$$

We may consider adjoint operators A^* i $A_h^* = PA^*P$. The nonzero eigenvalue of A_h^* is $\bar{\lambda}_h$ and corresponds to the normed eigenvector $v_h = \frac{Pf}{\|Pf\|}$, and we have

$$(A^* - \bar{\lambda}_h)v_h = \langle v_h, g \rangle f - \bar{\lambda}_h v_h = \|Pf\|^{-1}(\langle Pg, f \rangle f - \bar{\lambda}_h Pf) = \|Pf\|^{-1}\bar{\lambda}_h(f - Pf)$$

$$\|(A^* - \bar{\lambda}_h)v_h\| = \|Pf\|^{-1}|\lambda_h|\|(f - Pf)\|$$

How these quantities $\|(A - \lambda_h)u_h\|$, $\|(A^* - \bar{\lambda}_h)v_h\|$ obtained a posteriori may be used to estimate the error $\lambda - \lambda_h$? In the Example 1 we have the additional information about the operator A – namely, that is selfadjoint, this has given us information about the resolvent norm of A, namely $\|(A - \lambda)^{-1}\| = (\operatorname{dist}(\lambda, \sigma(A)))^{-1}$. Let us find a similar information for the operator in the example. Without any additional information we cannot get a posteriori further conclusions.

We may estimate the norm of $(A - \lambda)^{-1}$. With the equality $A = E = E^2$ it is easy to verify that

$$\begin{aligned}(A - \lambda)\left(((1 - \lambda)^{-1}E - \lambda^{-1}(I - E)\right) &= (1 - \lambda)^{-1}E - \lambda\left(((1 - \lambda)^{-1}E - \lambda^{-1}(I - E)\right) \\ &= \left((1 - \lambda)^{-1} - \lambda(1 - \lambda)^{-1} + 1\right)E + I = I.\end{aligned}$$

With Fact 1 we get

$$\|(A - \lambda)^{-1}\| \leq \|(1 - \lambda)^{-1}E\| + \|\lambda^{-1}(I - E)\| \leq \left(\frac{1}{|1 - \lambda|} + \frac{1}{|\lambda|}\right)\|E\|$$

Thus for the computed eigenvalue λ_h and eigenvector u_h

$$1 = \|u_h\| \leq \left(\frac{1}{|1 - \lambda_h|} + \frac{1}{|\lambda_h|}\right)\|E\|\|(A - \lambda_h)u_h\|$$

or

$$|\lambda_0 - \lambda_h| \leq \frac{(|\lambda_0 - \lambda_h| + |\lambda_h|)\|E\|}{|\lambda_h|}\|(A - \lambda_h)u_h\|.$$

$$|\lambda_0 - \lambda_h| \leq \|E\|\|(A - \lambda_h)u_h\|\left(1 - \frac{\|E\|}{|\lambda_h|}\|(A - \lambda_h)u_h\|\right)^{-1}$$

This is an aposteriori extimation, however we do not know the constants on the right side We have chances to estimate them, for example . $\lambda_0 - \lambda_h$ is close to zero, may be estimated by $\varepsilon > 0$ – for small h it becomes true. $E = \langle \cdot, f \rangle g$ is eigenprojection, its norm equals $\|f\|\|g\|$, we may hope that for h small enough it close to $|\langle u_h, v_h \rangle|^{-1}$, where u_h, v_h are unit eigenvectors of operators A i A^* ocorresponding to eigenvalue λ_h ($\bar{\lambda}_h$), $E_h = \langle \cdot, v_h \rangle u_h / \langle u_h, v_h \rangle$ aproximates E and $\|E_h\| = |\langle u_h, v_h \rangle|^{-1}$.

May we a posteriori extimate $\|u - u_h\|$ and $\|v - v_h\|$?

We have the equality

$$(A - \lambda_h)u_h = (A - \lambda_h)\left(Eu_h + (I - E)u_h\right) = (\lambda_0 - \lambda_h)Eu_h + (A - \lambda_h)(I - E)u_h$$

(15) implies that λ_h is aproximated with rate $O(h^2)$, while u_h i v_h are aproximated with rate $O(10)$, the second term should dominate in the above sum. In our example $A(I - E) = 0$ and consequently

$$|\lambda_h|\|(I - E)u_h\| \leq \|(A - \lambda_h)u_h\| + O(h^2).$$

We may use (15) now, to do it we have to estimate the norm of the error of eigenprojection

$$E - E_h = \frac{\langle \cdot, v \rangle u}{\langle u, v \rangle} - \frac{\langle \cdot, v_h \rangle u_h}{\langle u_h, v_h \rangle},$$

where u, v normalized vectors g, f.

We may expect, that with some constant c

$$\|E - E_h\| \leq c(\|(A - \lambda_h)u_h\| + \|(A^* - \bar{\lambda}_h)v_h\|),$$

then using (15) we get

$$\|\lambda - \lambda_h\| \leq c_1(\|(A - \lambda_h)u_h\| + \|(A^* - \bar{\lambda}_h)v_h\|)^2.$$

Example 2 is in fact not so far from the general case.

General case of simple eigenvalue.

Let λ be an isolated eigenvalue of A and $E = E(\lambda)$ the spectral projection. Then the operator $(A - \mu)|_{\ker E}$ is invertible for μ sufficiently close to λ.

$$\|\left((A - \mu I)|_{\ker E}\right)^{-1}\| = \|\left((A^* - \bar{\mu} I)|_{\ker E^*}\right)^{-1}\| \leq C_\lambda \quad \text{for } |\lambda - \mu| \leq r_\lambda, \tag{16}$$

C_λ is called in [1-3] a week stability constant. In the sequel we assume that $|\lambda - \lambda_h| \leq r_\lambda$ and that each $u \in \operatorname{ran} E$ is an eigenvector, i.e. there are not nontrivial Jordan cells. (16) implies that

$$\|(I - E)u_h\| = \|(A - \lambda_h)^{-1}(A - \lambda_h)(I - E)u_h\| \leq C_\lambda \|(A - \lambda_h)(I - E)u_h\|.$$

By the triangle inequality we have also

$$\|(A - \lambda_h)(I - E)u_h\| \leq \|(A - \lambda_h)u_h\| + \|(A - \lambda_h)Eu_h\| = \|(A - \lambda_h)u_h\| + |\lambda - \lambda_h|\|Eu_h\|$$

The above inequalities show that

$$\|(I-E)u_h\| \leq C_\lambda(\|(A-\lambda_h)u_h\| + |\lambda-\lambda_h|\|Eu_h\|) \tag{17}$$

Similar inequality may be written for adjoint operators.

$$\|(I-E^*)v_h\| \leq C_\lambda(\|(A^*-\bar{\lambda}_h)v_h\| + |\lambda-\lambda_h|\|E^*v_h\|) \tag{18}$$

To apply (15) and get aposteriori eigenvalue error estimation we need to estimate

$$\|E-E_h\| = \left\| \frac{\langle\cdot,v\rangle u}{\langle u,v\rangle} - \frac{\langle\cdot,v_h\rangle u_h}{\langle u_h,v_h\rangle} \right\|$$

The quantity $|\langle u,v\rangle|^{-1}$ is called a condtion number of eigenvalue λ, it is the norm $\|E\|$.

Let c stands for the right hand side of (17). We may take $u = \frac{Eu_h}{\|Eu_h\|}$. $u_h = (u_h - Eu_h) + Eu_h$ with the triangle inequality implies that $\|u_h\| = 1 \leq \|u_h - Eu_h\| + \|Eu_h\|$, $-Eu_h = u_h - Eu_h + u_h$ implies that $\|Eu_h\| \leq \|u_h - Eu_h\| + 1$. Therefore

$$1 - c \leq \|Eu_h\| \leq 1 + c.$$

Thus

$$\left\| Eu_h - \frac{Eu_h}{\|Eu_h\|} \right\| = \|Eu_h\| \left| 1 - \frac{1}{\|Eu_h\|} \right| = |\|Eu_h\| - 1| \leq c$$

this with (17) shows that

$$\|u_h - u\| \leq 2C_\lambda(\|(A-\lambda_h)u_h\| + |\lambda-\lambda_h|\|E\|)$$

of course we have to assume that $c \leq 1$. Similarily we show that

$$\|v_h - v\| \leq 2C_\lambda(\|(A^*-\bar{\lambda}_h)v_h\| + |\lambda-\lambda_h|\|E\|)$$

As the norm $\|E\|$ may be very large the scalar product $\langle u,v\rangle$ may be very small. However $\langle u_h,v_h\rangle \to \langle u,v\rangle$. To have $|\langle u_h,v_h\rangle| \geq \frac{1}{2}|\langle u,v\rangle|$ we have to make additional assumptions. From the identity

$$\langle u_h,v_h\rangle = \langle u+(u_h-u), v+(v_h-v)\rangle = \langle u,v\rangle + \langle u_h-u,v\rangle + \langle u,v_h-v\rangle + \langle u_h-u,v_h-v\rangle$$

we have

$$|\langle u_h,v_h\rangle - \langle u,v\rangle| \leq \|u_h-u\| + \|v_h-v\| + \|u_h-u\|\|v_h-v\|$$

To proceed further we assume that

$$|\langle u_h,v_h\rangle - \langle u,v\rangle| \leq \frac{1}{2}|\langle u,v\rangle| = \frac{1}{2}\|E\|^{-1}$$

to have

$$|\langle u_h,v_h\rangle| \geq \frac{1}{2}|\langle u,v\rangle| = \frac{1}{2}\|E\|^{-1}.$$

Then

$$\left|\frac{1}{\langle u_h, v_h\rangle} - \frac{1}{\langle u, v\rangle}\right| \leq \left|\frac{\langle u_h, v_h\rangle - \langle u, v\rangle}{\langle u_h, v_h\rangle\langle u, v\rangle}\right| \leq 2\|E\|2(\|u_h - u\| + \|v_h - v\|)$$

(we add the assumption $\|u_h - u\|\|v_h - v\| < 1$). Now

$$\begin{aligned}\|E - E_h\| &= \left\|\left(\frac{1}{\langle u, v\rangle} - \frac{1}{\langle u_h, v_h\rangle}\right)\langle\cdot, v\rangle u + \frac{\langle\cdot, v_h\rangle u_h - \langle\cdot, v\rangle u}{\langle u_h, v_h\rangle}\right\| \\ &\leq 4\|E\|(\|u_h - u\| + \|v_h - v\|) + 2\|E\|\|\langle\cdot, v_h\rangle(u_h - u) + \langle\cdot, v_h - v\rangle u\| \\ &\leq 6\|E\|(\|u_h - u\| + \|v_h - v\|)\end{aligned}$$

Now we use trace estimation (15) and we get (using $\|E_h\| \leq 2\|E\|$)

$$|\lambda - \lambda_h| \leq 3\|E - E_h\|^2\|A\|\|E\|$$

Thus with new constants $c_1 = 2C_\lambda$, $c_2 = 2C_\lambda\|E\|$, $c_3 = 18\|A\|\|E\|^2$ we have the system of inequalities

$$\begin{aligned}\|u_h - u\| &\leq c_1\|(A - \lambda_h)u_h\| + c_2|\lambda - \lambda_h| \\ \|v_h - v\| &\leq c_1\|(A^* - \bar{\lambda}_h)v_h\| + c_2|\lambda - \lambda_h| \\ |\lambda - \lambda_h| &\leq c_3(\|u_h - u\| + \|v_h - v\|)^2\end{aligned}$$

with unknowns $\|u_h - u\|, \|v_h - v\|, |\lambda - \lambda_h|$. Solving it we get a posteriori estimations

$$\begin{aligned}\|u_h - u\| &\leq c_4(\|(A - \lambda_h)u_h\| + |(A^* - \bar{\lambda}_h)v_h\|) \\ \|v_h - v\| &\leq c_4(\|(A - \lambda_h)u_h\| + |(A^* - \bar{\lambda}_h)v_h\|) \\ |\lambda - \lambda_h| &\leq c_6(|(A - \lambda_h)u_h\| + |(A^* - \bar{\lambda}_h)v_h\|)^2\end{aligned}$$

which are valid for h sufficiently small.

The "general case" studied above is in fact not general – we have considered the case of an eigenvalue with geometric multiplicity one. We have used the fact that the eigenprojection $E_h = \langle\cdot, v_h\rangle u_h$, where u_h, v_h are eigenvectors of A_h and A_h^*. Even if A is a projection A_h may have nontrivial Jordan blocks. This shown in the next example.

Example 3. Let A be m-dimensional orthogonal projection,

$$A = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}$$

and G some $m \times m$ matrix. Let

$$P_n = \begin{pmatrix} I_m - \varepsilon G & 0 & B_\varepsilon & 0 \\ 0 & I_n & 0 & 0 \\ C_\varepsilon & 0 & \varepsilon G & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where the bottom right block is infinite dimensional. $B_\varepsilon = \sqrt{\varepsilon} I_m$, $C_\varepsilon = \sqrt{\varepsilon} G(I_m - \varepsilon G)$. If $\varepsilon = \varepsilon_n \to 0$ then $P_n \to I$ and $P_n^* \to I$ strongly. P_n are projections, this follows from

$$\begin{aligned}
&\begin{pmatrix} I_m - \varepsilon G & \sqrt{\varepsilon} I_m \\ \sqrt{\varepsilon} G(I_m - \varepsilon G) & \varepsilon G \end{pmatrix}^2 \\
&= \begin{pmatrix} (I_m - \varepsilon G)^2 + \varepsilon G(I_m - \varepsilon G) & \sqrt{\varepsilon}(I_m - \varepsilon G) + \sqrt{\varepsilon}\varepsilon G \\ \sqrt{\varepsilon} G(I_m - \varepsilon G)^2 + \varepsilon\sqrt{\varepsilon} G^2 (I_m - \varepsilon G) & \varepsilon G(I_m - \varepsilon G) + \varepsilon^2 G^2 \end{pmatrix} \\
&= \begin{pmatrix} I_m - \varepsilon G & \sqrt{\varepsilon} I_m \\ \sqrt{\varepsilon} G(I_m - \varepsilon G) & \varepsilon G \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
P_n A P_n &= \begin{pmatrix} I_m - \varepsilon G & \sqrt{\varepsilon} I_m \\ \sqrt{\varepsilon} G(I_m - \varepsilon G) & \varepsilon G \end{pmatrix} \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_m - \varepsilon G & \sqrt{\varepsilon} I_m \\ \sqrt{\varepsilon} G(I_m - \varepsilon G) & \varepsilon G \end{pmatrix} \\
&= \begin{pmatrix} I_m - \varepsilon G & 0 \\ \sqrt{\varepsilon} G(I_m - \varepsilon G) & 0 \end{pmatrix} \begin{pmatrix} I_m - \varepsilon G & \sqrt{\varepsilon} I_m \\ 0 & 0 \end{pmatrix} \\
&= (I_m - \varepsilon G) \begin{pmatrix} I_m & 0 \\ \sqrt{\varepsilon} G & 0 \end{pmatrix} \begin{pmatrix} I_m - \varepsilon G & \sqrt{\varepsilon} I_m \\ 0 & 0 \end{pmatrix} = (I_m - \varepsilon G) P_n
\end{aligned}$$

The above shows that $A_n = P_n A P_n$ is unitarily equivalent to the direct sum of $I_m - \varepsilon G$ and null operator in the infinite dimensional space. and A_n restricted to the range of the spectral projection E_h is just $I_m - \varepsilon G$. In particular it may be one Jordan block, and we cannot argue is in "General case".

Note that $A_n = \{0\} \cup (1 - \varepsilon\sigma(G))$.

Note that $= \operatorname{ran} E_n = \operatorname{ran} P_n$ and

$$P_n \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (I_m - \varepsilon G)x + \sqrt{\varepsilon} G y \\ \sqrt{\varepsilon}(I_m - \varepsilon G) G x + \varepsilon G y \end{pmatrix} = \begin{pmatrix} z \\ \sqrt{\varepsilon} G z \end{pmatrix}$$

with $z = (I_m - \varepsilon G)x + \sqrt{\varepsilon} G y$.

Thus $\operatorname{ran} E_n = \left\{ \begin{pmatrix} z \\ \sqrt{\varepsilon} G z \end{pmatrix} ; z \in \operatorname{ran} A \right\}$. and

$$\left\| \begin{pmatrix} z \\ \sqrt{\varepsilon} G z \end{pmatrix} - A \begin{pmatrix} z \\ \sqrt{\varepsilon} G z \end{pmatrix} \right\| = \left\| \begin{pmatrix} 0 \\ \sqrt{\varepsilon} G z \end{pmatrix} \right\| = \sqrt{\varepsilon} \|Gz\|$$

Thus $\operatorname{ran} E_n$ the generalized space of eigenvectors is approximated with order $\sqrt{\varepsilon}$ while eigenvalues are approximated with order ε as expected.

General case – multiple eigenvalue.

Instead of eigenvectors we may try to use eigenspaces. Let E be the eigenprojection corresponding to $\lambda \in \sigma(A)$, and E_h be the approximating eigenspace. We may compute

$$\operatorname{dist}(\operatorname{ran} E_h, \operatorname{ran} A E_h), \quad \operatorname{dist}(\operatorname{ran} E_h^*, \operatorname{ran} A^* E_h^*)$$

With E instead of E_h these distances equal to zero. $\operatorname{ran} AE_h$ should be close to $\operatorname{ran} EE_h = \operatorname{ran} E$ – as $\|(I-E)E_h\|$ tends to zero.

We may hope that the inequality

$$\|E-F\| \leq c \max\{\operatorname{dist}(\operatorname{ran} E, \operatorname{ran} F), \operatorname{dist}(\operatorname{ran} E^*, \operatorname{ran} F^*)$$

holds for sufficiently close projections E, F. Then with (15) we should get a posteriori estimation of the eigenvalue error.

For two subspaces M, N we define

$$\delta(M,N) = \sup\{x \in N, \|x\| = 1; \inf_{y \in M} \|x-y\|\}, \quad \operatorname{dist}((M,N) = \max\{\delta(M,N), \delta(N,M)\}.$$

Lemma. If E, F are projections then

$$\|E-F\| \leq \max\{\|E\|, \|F\|\} \max\{\operatorname{dist}(\operatorname{ran} E, \operatorname{ran} F), \operatorname{dist}(\operatorname{ran} E^*, \operatorname{ran} F^*)\}.$$

Proof. Let $\|x\| = 1$ and $\|(E-F)x\| = \|E-F\|$. Then

$$\|Ex - Fx\| = \|Ex - FEx + FEx - Fx\| \leq \|Ex - FEx\| + \|Fx - FEx\|$$

$$\|Ex - FEx\| = \|Ex\| \left\| \frac{Ex}{\|Ex\|} - F\frac{Ex}{\|Ex\|} \right\| \leq \|E\| \delta(\operatorname{ran} E, \operatorname{ran} F).$$

Let $\|y\| = 1$ and

$$\|Fx - FEx\| = \langle F(I-E)x, y\rangle = \langle x, (I-E^*)F^*y\rangle$$

then

$$\|Fx - FEx\| \leq \|x\| \|F^*y\| \left\| \frac{F^*y}{\|F^*y\|} - E^* \frac{F^*y}{\|F^*y\|} \right| \leq \|F^*\| \delta(\operatorname{ran} F^*, \operatorname{ran} E^*)$$

The above inequalities imply the thesis.

References

[1] Joscha Gedicke and Carsten Carstensen, *A posteriori error estimators for non-symetric eigenvalue problems*, August 2009. DFG Research Center MATHEON, preprint 659.

[2] Vincent Heuvelline and Folf Rannacher *A posteriori error control for finite element approximation of elliptic eigenvalue problems*, Adv. Comp. Math. 15 (2001) 107–138

[3] Carstensen, C.; Gedicke, J.; Mehrmann, V.; Miedlar, A. *An adaptive homotopy approach for non-selfadjoint eigenvalue problems* . Numer. Math. 119 (2011), no. 3, 557–583

[4] Daniel B. Szyld, *The many proofs of an identity on the norm of oblique projections.* Numer.Algor (2006) 42, 309–323