A Simple Proof and Refinement of Wielandt’s Eigenvalue Inequality

Lutz Dümbgen


Abstract: Wielandt (1967) proved an eigenvalue inequality for partitioned symmetric matrices, which turned out to be very useful in statistical applications. A simple proof yielding sharp bounds is given.

Keywords and phrases: eigenvalue inequality, partitioned matrix

Correspondence to: Lutz Duembgen, Institute of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland; e-mail: duembgen@stat.unibe.ch
Let $A \in \mathbb{R}^{p \times p}$ be a symmetric matrix of the form

$$A = \begin{pmatrix} B & C \\ C' & D \end{pmatrix}$$

with $B \in \mathbb{R}^{r \times r}$, $C \in \mathbb{R}^{r \times s}$ and $D \in \mathbb{R}^{s \times s}$ such that

$$\lambda_r(B) > \lambda_1(D);$$

generally $\lambda_1(E) \geq \lambda_2(E) \geq \cdots \geq \lambda_q(E)$ denote the ordered eigenvalues of a symmetric matrix $E \in \mathbb{R}^{q \times q}$. Wielandt (1967) showed that the eigenvalues of $A$ can be approximated by the eigenvalues of $B$ and $D$ in the following sense:

$$0 \leq \lambda_i(A) - \lambda_i(B) \leq \frac{\lambda_1(CC')}{\lambda_i(B) - \lambda_1(D)} \text{ for } 1 \leq i \leq r \text{ and } \quad (1)$$

$$0 \leq \lambda_j(D) - \lambda_{r+j}(A) \leq \frac{\lambda_1(CC')}{\lambda_r(B) - \lambda_j(D)} \text{ for } 1 \leq j \leq s.$$  

These inequalities can be used to compute derivatives and pseudo-derivatives of eigenvalues. They are also very useful in statistical problems involving eigenvalues of random symmetric matrices; see Eaton and Tyler (1991, 1994). In my opinion the original proof, described in Eaton and Tyler (1991), is somewhat complicated. The main ingredient seems to be the Courant-Fischer minimax representation

$$\lambda_k(E) = \max_{\text{dim}(V)=k} \min_{v \in V: v^Tv=1} v'Ev \text{ for } 1 \leq k \leq q, \quad (2)$$

where $V$ stands for a linear subspace of $\mathbb{R}^q$; see section 1f.2 of Rao (1973). In this note (2) is used directly to derive the following refinement of (1):

**Theorem.** For $1 \leq i \leq r$,

$$0 \leq \lambda_i(A) - \lambda_i(B) \leq \sqrt{\frac{(\lambda_i(B) - \lambda_1(D))^2}{4} + \lambda_1(CC')} - \frac{\lambda_i(B) - \lambda_1(D)}{2},$$

and for $1 \leq j \leq s$,

$$0 \leq \lambda_j(D) - \lambda_{r+j}(A) \leq \sqrt{\frac{(\lambda_r(B) - \lambda_j(D))^2}{4} + \lambda_1(CC')} - \frac{\lambda_r(B) - \lambda_j(D)}{2}.$$  

2
Remark 1: Since 
\[
\sqrt{\frac{\alpha^2}{4} + \beta^2} - \frac{\alpha}{2} \leq \min\{\beta, \frac{\beta^2}{\alpha}\} \quad \forall \alpha, \beta > 0,
\]
this result implies Wielandt’s bounds (1).

Remark 2: The upper bounds are sharp. For if \(p = 2\) one can compute the eigenvalues of \(A\) explicitly and obtains

\[
\lambda_1(A) - \lambda_1(B) = \lambda_1(D) - \lambda_2(A) = \sqrt{\frac{(B - D)^2}{4} + C^2} - \frac{B - D}{2}.
\]

For general \(p\) one has to consider diagonal matrices \(B, D\) and suitable matrices \(C\) with only one nonzero coefficient.

Proof of the Theorem: One easily verifies that the asserted inequalities are invariant under the transformation \(A \mapsto A - \lambda_1(D)I\), where \(I\) is the identity matrix in \(\mathbb{R}^{p \times p}\). Therefore one may assume without loss of generality that \(\lambda_1(D) = 0\).

For \(1 \leq i \leq r\) it follows from (2) that

\[
\lambda_i(A) \geq \max_{V \subset \mathbb{R}^r \times \{0\}: \dim(V) = i} \min_{v \in V : v'v = 1} v'Av = \lambda_i(B). \tag{3}
\]

On the other hand, let \(W\) be an \(i\)-dimensional subspace of \(\mathbb{R}^p\) such that

\[
\lambda_i(A) = \min_{v \in W : v'v = 1} v'Av.
\]

If \(v \in \mathbb{R}^p\) is written as \(v = (v_1', v_2')'\) with \(v_1 \in \mathbb{R}^r\) and \(v_2 \in \mathbb{R}^s\), then

\[
W_{(1)} = \{v_1 : v \in W\}
\]

is an \(i\)-dimensional subspace of \(\mathbb{R}^r\). For if \(\dim(W_{(1)}) < i\), then \(w_{(1)} = 0\) for some unit vector \(w \in W\), and

\[
\lambda_i(A) \leq w'Aw = w_2'Dw_2 \leq 0,
\]

which would contradict (3). Any unit vector \(v \in W\) can be written as

\[
v = \sqrt{(1 + \rho)/2} u_{(1)} + \sqrt{(1 - \rho)/2} u_{(2)}
\]
for unit vectors $u^{(1)} \in W(1), u^{(2)} \in R^s$ and some $\rho \in [-1, 1]$. Then
\[
v'Av = (1 + \rho) u'^{(1)}Bu^{(1)}/2 + \sqrt{1 - \rho^2} u'^{(1)}Cu^{(2)} + (1 - \rho) u'^{(2)}Du^{(2)}/2
\leq (1 + \rho) u'^{(1)}Bu^{(1)}/2 + \sqrt{1 - \rho^2} \sqrt{\lambda_1(CC')}
\leq u'^{(1)}Bu^{(1)}/2 + \sqrt{(u'^{(1)}Bu^{(1)})^2/4 + \lambda_1(CC')}.
\]
Consequently, since $H(x) := x/2 + \sqrt{x^2/4 + \lambda_1(CC')}$ is nondecreasing in $x \geq 0$,
\[
\lambda_i(A) \leq \min_{u^{(1)} \in W(1); u'^{(1)}u^{(1)} = 1} H(u'^{(1)}Bu^{(1)})
= H \left( \min_{u^{(1)} \in W(1); u'^{(1)}u^{(1)} = 1} u'^{(1)}Bu^{(1)} \right)
\leq H(\lambda_i(B))
= \lambda_i(B) + \sqrt{(\lambda_i(B) - \lambda_1(D))^2/4 + \lambda_1(CC') - (\lambda_i(B) - \lambda_1(D))/2}.
\]
Thus the first part of the theorem is true, and the second half follows by replacing $A$ with $-A$.

References


