A Simple Proof and Refinement of Wielandt's Eigenvalue Inequality

Lutz Dümbgen

published 1995 in *Statist. Prob. Letters* **25**, 113-115

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 $A \in \mathbf{R}^{p\times p}$ be a symmetric matrix of the form

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

with $B \in \mathbf{R}^{r\times r}$, $C \in \mathbf{R}^{r\times s}$ and $D \in \mathbf{R}^{s\times s}$ 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 $E \in \mathbf{R}^{q\times q}$. Wielandt (1967) showed that the eigenvalues of A can be approximated by the eigenvalues of B and D in the following sense:

$$\begin{aligned} 0 \leq \lambda_i(A) - \lambda_i(B) &\leq \frac{\lambda_1(CC')}{\lambda_i(B) - \lambda_1(D)} \quad \text{for } 1 \leq i \leq r \text{ and} \\ 0 \leq \lambda_j(D) - \lambda_{r+j}(A) &\leq \frac{\lambda_1(CC')}{\lambda_r(B) - \lambda_j(D)} \quad \text{for } 1 \leq j \leq s. \end{aligned} \tag{1}$$

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

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

where $\mathbf{V}$ stands for a linear subspace of $\mathbf{R}^q$; 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 $1 \leq j \leq s$,

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

Remark 1: Since

$$\sqrt{\alpha^2/4+\beta^2} \;-\; \alpha/2 \;\le\; \min\Bigl\{\beta, \beta^2/\alpha\Bigr\} \;\;\forall\, \alpha, \beta > 0,$$

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 $\mathbf{R}^{p\times p}$. Therefore one may assume without loss of generality that $\lambda_1(D) = 0$.

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

$$\lambda_i(A) \;\ge\; \max_{\mathbf{V}\subset\mathbf{R}^r\times\{0\}:\dim(\mathbf{V})=i} \;\; \min_{v\in\mathbf{V}:v'v=1} v'Av \;=\; \lambda_i(B). \tag{3}$$

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

$$\lambda_i(A) \;=\; \min_{v\in\mathbf{W}:v'v=1} v'Av.$$

If $v \in \mathbf{R}^p$ is written as $v = (v_{(1)}', v_{(2)}')'$ with $v_{(1)} \in \mathbf{R}^r$ and $v_{(2)} \in \mathbf{R}^s$, then

$$\mathbf{W}_{(1)} \;=\; \{v_{(1)} : v \in \mathbf{W}\}$$

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

$$\lambda_i(A) \;\le\; w'Aw \;=\; w_{(2)}'Dw_{(2)} \;\le\; 0,$$

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

$$v \;=\; \sqrt{(1+\rho)/2}\, u_{(1)} + \sqrt{(1-\rho)/2}\, u_{(2)}$$

for unit vectors $u_{(1)} \in \mathbf{W}_{(1)}, u_{(2)} \in \mathbf{R}^s$ and some $\rho \in [-1, 1]$. Then

$$\begin{aligned}
v'Av &= (1+\rho)\, u'_{(1)} B u_{(1)}/2 + \sqrt{1-\rho^2}\, u'_{(1)} C u_{(2)} + (1-\rho)\, u'_{(2)} D u_{(2)}/2 \\
&\leq (1+\rho)\, u'_{(1)} B u_{(1)}/2 + \sqrt{1-\rho^2}\, \sqrt{\lambda_1(CC')} \\
&= u'_{(1)} B u_{(1)}/2 + \left(\rho, \sqrt{1-\rho^2}\right) \begin{pmatrix} u'_{(1)} B u_{(1)}/2 \\ \sqrt{\lambda_1(CC')} \end{pmatrix} \\
&\leq u'_{(1)} B u_{(1)}/2 + \sqrt{(u'_{(1)} B u_{(1)})^2/4 + \lambda_1(CC')}.
\end{aligned}$$

Consequently, since $H(x) := x/2 + \sqrt{x^2/4 + \lambda_1(CC')}$ is nondecreasing in $x \geq 0$,

$$\begin{aligned}
\lambda_i(A) &\leq \min_{u_{(1)} \in \mathbf{W}_{(1)} : u'_{(1)} u_{(1)} = 1} H(u'_{(1)} B u_{(1)}) \\
&= H\left(\min_{u_{(1)} \in \mathbf{W}_{(1)} : u'_{(1)} u_{(1)} = 1} u'_{(1)} B u_{(1)} \right) \\
&\leq H(\lambda_i(B)) \\
&= \lambda_i(B) + \sqrt{(\lambda_i(B) - \lambda_1(D))^2/4 + \lambda_1(CC')} - (\lambda_i(B) - \lambda_1(D))/2.
\end{aligned}$$

Thus the first part of the theorem is true, and the second half follows by replacing A with $-A$ $\square$

References

Eaton, M.L. and D.E. Tyler (1991): On Wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix. *Ann. Statist.* **19**, 260-271

Eaton, M.L. and D.E. Tyler (1994): The asymptotic distribution of singular values with applications to canonical correlations and correspondence analysis. *J. Multivariate Anal.* **50**, 238-264

Rao, C.R. (1973): *Linear Statistical Inference and Its Applications (2nd edition).* Wiley, New York

Wielandt, H. (1967): *Topics in the Analytic Theory of Matrices.* (Lecture notes prepared by R. R. Meyer) Univ. Wisconsin Press, Madison