THE GENERALIZED WIELANDT INEQUALITY IN INNER PRODUCT SPACES

M. Lin, G. Sinnamon

Communicated by E. Kissin

Abstract. A new inequality between angles in inner product spaces is formulated and proved. It leads directly to a concise statement and proof of the generalized Wielandt inequality, including a simple description of all cases of equality. As a consequence, several recent results in matrix analysis and inner product spaces are improved.

1 Introduction

The Wielandt and generalized Wielandt inequalities control how much angles can change under a given invertible matrix transformation of $\mathbb{C}^n$. The control is given in terms of the condition number of the matrix. Wielandt, in [13], gave a bound on the resulting angles when orthogonal complex lines are transformed. Subsequently, Bauer and Householder, in [1], extended the inequality to include arbitrary starting angles. These basic inequalities of matrix analysis were introduced to give bounds on convergence rates of iterative projection methods but have found a variety of applications in numerical methods, especially eigenvalue estimation. They are also applied in multivariate analysis, where angles between vectors correspond to statistical correlation. See, for example, [1], [5], [6], [7] and [5]. There are also matrix-valued versions of the inequality that are receiving attention, especially in the context of statistical analysis. See [2], [10], [12], and [16].

The condition number of an invertible matrix A is $\kappa(A) = \|A\| \|A^{-1}\|$, where $\|\cdot\|$ denotes the operator norm. If A is positive definite and Hermitian, $\kappa(A)$ is easily seen to be the ratio of the largest and smallest eigenvalues of A. The following statement of the generalized Wielandt inequality is taken from [5].

Theorem 1.1. *Let A be an invertible $n \times n$ matrix. If $x, y \in \mathbb{C}^n$ and $\Phi, \Psi \in [0, \pi/2]$ satisfy*

$$|y^*x| \le \|x\| \|y\| \cos \Phi \quad and \quad \cot(\Psi/2) = \kappa(A) \cot(\Phi/2),$$

then

$$|(Ay)^*(Ax)| \le \|Ax\| \|Ay\| \cos \Psi.$$

The generalized Wielandt inequality can be difficult to apply for several reasons. First, despite having various equivalent formulations, the inequality seems always to be

expressed in ways that hide the natural symmetry coming from the invertible transformation involved. Next, the conditions for equality are known, see [9], but are unwieldy and hard to apply. Finally, the angles involved are angles between complex lines rather than between individual vectors.

Although the last point seems minor, we found it to be the key to a symmetric formulation and a simple description of the cases of equality. In Theorem 2.1 and its matrix analytic counterpart, Theorem 3.1, we present a new inequality that gives sharp upper and lower bounds for the angle between a pair of transformed vectors. The conditions for equality are simple and easy to apply. This new inequality relates angles between vectors rather than between complex lines but it immediately implies a result for angles between complex lines that is equivalent to the generalized Wielandt inequality. Moreover, this version of the generalized Wielandt inequality retains the simple form of the new inequality and (most of) the simplicity of its conditions for equality.

In Section 2 we work in the context of an arbitrary real or complex vector space having two inner products. This approach preserves symmetry by avoiding the distinction between angles before and after a fixed transformation. Also, the main result is not restricted to $\mathbb{C}^n$ but holds for vectors in infinite-dimensional spaces. As an application of the unrestricted result, we improve a metric space inequality from [2]. The main results are then formulated in the language of matrix analysis in Section 3, and we apply them to improve inequalities from [15] and [11], and to settle a conjecture from [14].

To begin, a short discussion of angles in inner product spaces is in order. In a real inner product space $(V, \langle \cdot, \cdot \rangle)$ the angle $\theta = \theta(u, v)$ between two non-zero vectors is defined by, $0 \leq \theta \leq \pi$ and

$$\cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|}.$$

Here $\|u\| = \sqrt{\langle u, u \rangle}$ is the norm induced by the inner product. The angle between subsets S and T of V is the infimum of the angles between non-zero elements of S and T, so

$$\Theta(S, T) = \inf\{\theta(u, v) : 0 \neq u \in S, 0 \neq v \in T\}.$$

With this definition it is easy to check that the angle $\Theta = \Theta(\mathbb{R}u, \mathbb{R}v)$ between the lines $\mathbb{R}u$ and $\mathbb{R}v$ satisfies $0 \leq \Theta \leq \pi/2$ and

$$\cos \Theta = \frac{|\langle u, v \rangle|}{\|u\| \|v\|}.$$

A complex inner product space $(V, \langle \cdot, \cdot \rangle)$ may be viewed as the real inner product space $(V_\mathbb{R}, \mathrm{Re}\langle \cdot, \cdot \rangle)$ where $V_\mathbb{R} = V$ with the scalars restricted to $\mathbb{R}$. Since $\mathrm{Re}\langle v, v \rangle = \langle v, v \rangle$ for all $v \in V$, lengths in V are preserved and therefore so are angles. Thus, this real inner product is used to define the angle θ between the vectors u and v, and a computation gives the formula for the angle Θ between the complex lines $\mathbb{C}u$ and $\mathbb{C}v$. We have,

$$\cos \theta = \frac{\mathrm{Re}\langle u, v \rangle}{\|u\| \|v\|} \quad \text{and} \quad \cos \Theta = \frac{|\langle u, v \rangle|}{\|u\| \|v\|}.$$

The second formula is often used as a definition of the angle between vectors u and v in a complex inner product space. (Angles defined this way do not determine angles in triangles correctly but they have the advantage that complex orthogonality, namely $\langle u, v \rangle = 0$, is equivalent to the angle between u and v being $\pi/2$.)

We will make use of the simple observation that if $|\alpha| = 1$, then

$$\Theta(\mathbb{C}u, \mathbb{C}v) = \theta(\alpha u, v) \quad \text{if and only if} \quad |\langle u, v \rangle| = \alpha \langle u, v \rangle. \tag{1.1}$$

(Note that our inner products are taken to be linear in the first variable.) The above observation remains valid for $\Theta(\mathbb{R}u, \mathbb{R}v)$ in a real inner product space, where $\alpha = \pm 1$.

2 Main results

Suppose V is a non-trivial real or complex vector space. Let $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ be inner products on V and define m, V_m, M, V_M, E_1 and E_2 by,

$$\begin{cases} m = \inf_{0 \neq v \in V} \|v\|_2 / \|v\|_1, & V_m = \{v \in V : \|v\|_2 = m\|v\|_1\}, \\[2mm] M = \sup_{0 \neq v \in V} \|v\|_2 / \|v\|_1, & V_M = \{v \in V : \|v\|_2 = M\|v\|_1\}, \\[2mm] E = E_j = \left\{ (u, v) : \dfrac{u}{\|u\|_j} + \dfrac{v}{\|v\|_j} \in V_m, \dfrac{u}{\|u\|_j} - \dfrac{v}{\|v\|_j} \in V_M \right\}, \end{cases} \tag{2.1}$$

for $j = 1, 2$. Here, as usual, $\|v\|_1 = \sqrt{\langle v, v \rangle_1}$ and $\|v\|_2 = \sqrt{\langle v, v \rangle_2}$. We anticipate the result of Corollary 2.2 in the definition of E above.

Evidently $0 \leq m \leq M \leq \infty$, $0 \in V_m$ and $0 \in V_M$. (The convention $0 \cdot \infty = 0$ ensures that $0 \in V_M$ when $M = \infty$.) A standard compactness argument shows that if V is finite dimensional then $0 < m \leq M < \infty$ and $V_m \neq \{0\} \neq V_M$. If $m = M$ then $V_m = V_M = V$ and, by polarization, $\langle u, v \rangle_2 = m^2 \langle u, v \rangle_1$ for all $u, v \in V$.

Lemma 2.1. *Let V be a real vector space equipped with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. Make definitions (2.1). If $m < M$, then V_m and V_M are subspaces and the two are mutually orthogonal with respect to both inner products.*

Proof. Suppose u is a non-zero vector in V_m and $v \in V$ is not a multiple of u. Then

$$f(t) = \frac{\|u + tv\|_2^2}{\|u + tv\|_1^2} = \frac{\langle u, u \rangle_2 + 2t\langle u, v \rangle_2 + t^2 \langle v, v \rangle_2}{\langle u, u \rangle_1 + 2t\langle u, v \rangle_1 + t^2 \langle v, v \rangle_1}$$

is defined and differentiable for $t \in \mathbb{R}$. Since f achieves its minimum value at $t = 0$, $f'(0) = 0$. That is, $\langle u, v \rangle_2 \langle u, u \rangle_1 = \langle u, u \rangle_2 \langle u, v \rangle_1$. Thus, for all $u \in V_m$ and all $v \in V$,

$$\langle u, v \rangle_2 = m^2 \langle u, v \rangle_1.$$

(The excluded case, $u = 0$ or v a multiple of u, is easily verified.) It follows that if $v \in V_m$ then f is the constant function with value m^2. In particular, $f(1) = m^2$, so $u + v \in V_m$. Since it is clearly closed under scalar multiplication, V_m is a subspace.

Repeating the argument for V_M shows that it, too, is a subspace and that for all $v \in V_M$ and $u \in V$,

$$\langle u, v \rangle_2 = M^2 \langle u, v \rangle_1.$$

If $u \in V_m$ and $v \in V_M$ then $m^2 \langle u, v \rangle_1 = \langle u, v \rangle_2 = M^2 \langle u, v \rangle_1$ and hence $\langle u, v \rangle_1 = \langle u, v \rangle_2 = 0$. Thus u and v are orthogonal with respect to both inner products. This completes the proof. □

Corollary 2.1. *Let V be a real vector space equipped with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. Make definitions (2.1). If V is two-dimensional, then there is a basis of V that is orthogonal with respect to both inner products.*

Proof. If $m = M$ then the two inner products are multiples of each other and any orthogonal basis will do. Otherwise, let $0 \neq b \in V_m$ and $0 \neq B \in V_M$. Then $\{b, B\}$ is the desired basis. □

The next result justifies the use of E to denote either E_1 or E_2.

Corollary 2.2. *Let V be a real vector space equipped with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. Make definitions (2.1). Then $E_1 = E_2$.*

Proof. By symmetry it is enough to show that $E_1 \subseteq E_2$. For $(u, v) \in E_1$, let

$$w = \frac{u}{\|u\|_1} + \frac{v}{\|v\|_1} \in V_m \quad \text{and} \quad W = \frac{u}{\|u\|_1} - \frac{v}{\|v\|_1} \in V_M.$$

By Lemma 2.1, w and W are orthogonal with respect to $\langle \cdot, \cdot \rangle_2$, so

$$\|u\|_2^2 / \|u\|_1^2 = \tfrac{1}{4} \|w + W\|_2^2 = \tfrac{1}{4}(\|w\|_2^2 + \|W\|_2^2) = \tfrac{1}{4} \|w - W\|_2^2 = \|v\|_2^2 / \|v\|_1^2.$$

Thus

$$\frac{u}{\|u\|_2} + \frac{v}{\|v\|_2} = \frac{\|u\|_1}{\|u\|_2} w \in V_m \quad \text{and} \quad \frac{u}{\|u\|_2} - \frac{v}{\|v\|_2} = \frac{\|u\|_1}{\|u\|_2} W \in V_m$$

and so $(u, v) \in E_2$. □

Having two inner products, the space V has two differing notions of the angle between vectors. Our main result provides a comparison between these angles in terms of the quantities m and M defined in (2.1).

Theorem 2.1. *Let V be a real or complex vector space equipped with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. Make definitions (2.1). For independent vectors u and v in V let φ and ψ be defined by, $0 \leq \varphi \leq \pi$, $0 \leq \psi \leq \pi$,*

$$\cos \varphi = \frac{\mathrm{Re}\langle u, v \rangle_1}{\|u\|_1 \|v\|_1} \quad \text{and} \quad \cos \psi = \frac{\mathrm{Re}\langle u, v \rangle_2}{\|u\|_2 \|v\|_2}.$$

Then

$$(m/M) \tan(\varphi/2) \leq \tan(\psi/2) \leq (M/m) \tan(\varphi/2). \tag{2.2}$$

Equality holds in the right-hand inequality if and only if $(u, v) \in E$. Equality holds in the left-hand inequality if and only if $(u, -v) \in E$.

Proof. First consider the case that V is a real vector space. Note that the assumption of independence ensures $0 < \varphi < \pi$ and $0 < \psi < \pi$.

By Corollary 2.1, the span of u and v has a basis $\{b, B\}$ that is orthogonal with respect to both inner products. Without loss of generality we may assume that $\|b\|_1 = \|B\|_1 = 1$. For notational convenience, set $n = \|b\|_2$ and $N = \|B\|_2$ and suppose, by interchanging b and B if necessary, that $n \leq N$. Note that the definitions of m and M ensure that $m \leq n$ and $N \leq M$. Write $u = u_b b + u_B B$ and $v = v_b b + v_B B$ for some real numbers u_b, u_B, v_b, and v_B. In terms of these coordinates we have,

$$\|u\|_1^2 \|v\|_1^2 \sin^2 \varphi = \|u\|_1^2 \|v\|_1^2 - \langle u, v \rangle_1^2$$
$$= (u_b^2 + u_B^2)(v_b^2 + v_B^2) - (u_b v_b + u_B v_B)^2$$
$$= (u_b v_B - u_B v_b)^2$$

and

$$\|u\|_2^2 \|v\|_2^2 \sin^2 \psi = \|u\|_2^2 \|v\|_2^2 - \langle u, v \rangle_2^2$$
$$= (n^2 u_b^2 + N^2 u_B^2)(n^2 v_b^2 + N^2 v_B^2) - (n^2 u_b v_b + N^2 u_B v_B)^2$$
$$= n^2 N^2 (u_b v_B - u_B v_b)^2.$$

Thus,

$$\|u\|_2 \|v\|_2 \sin \psi = nN \|u\|_1 \|v\|_1 \sin \varphi. \tag{2.3}$$

The derivative of

$$g(x) = (u_b^2 + x u_B^2)^{1/2}(v_b^2 + x v_B^2)^{1/2} + (u_b v_b + x u_B v_B)$$

is

$$g'(x) = \frac{1}{2}\left(u_B \left(\frac{v_b^2 + x v_B^2}{u_b^2 + x u_B^2}\right)^{1/4} + v_B \left(\frac{u_b^2 + x u_B^2}{v_b^2 + x v_B^2}\right)^{1/4}\right)^2 \geq 0,$$

so $g(1) \leq g(N^2/n^2)$. Multiplying both sides of this by n^2 gives,

$$n^2 \|u\|_1 \|v\|_1 (1 + \cos \varphi) \leq \|u\|_2 \|v\|_2 (1 + \cos \psi). \tag{2.4}$$

Combining (2.3) and (2.4) gives,

$$\tan(\psi/2) = \frac{\sin \psi}{(1 + \cos \psi)} \leq \frac{nN \sin \varphi}{n^2(1 + \cos \varphi)} = (N/n)\tan(\varphi/2), \tag{2.5}$$

with equality if and only if $g'(x) = 0$ for $x \in (1, N^2/n^2)$. Since $m \leq n \leq N \leq M$, (2.5) proves the right-hand inequality of (2.2).

If equality holds in the right-hand inequality of (2.2), then equality holds in (2.5) and $n = m$, $N = M$, $b \in V_m$, and $B \in V_M$. If $m = M$ then $V_m = V_M = V$ and $\varphi = \psi$ so the last two statements of the theorem are trivial. Otherwise, equality in (2.5) implies that g' is zero on the non-trivial interval $(1, M^2/m^2)$. That is,

$$u_B \left(\frac{v_b^2 + x v_B^2}{u_b^2 + x u_B^2}\right)^{1/4} + v_B \left(\frac{u_b^2 + x u_B^2}{v_b^2 + x v_B^2}\right)^{1/4} = 0$$

and hence $u_B^2 v_b^2 = v_B^2 u_b^2$. Since u and v are independent, both u_B and v_B are non-zero, they have opposite signs, and $u_B v_b = -v_B u_b$. Therefore,

$$
\begin{aligned}
\frac{u}{\|u\|_1} + \frac{v}{\|v\|_1} &= \frac{u_b b + u_B B}{\sqrt{u_b^2 + u_B^2}} + \frac{v_b b + v_B B}{\sqrt{v_b^2 + v_B^2}} \\
&= \pm \left(\frac{(u_b/u_B)b + B}{\sqrt{(u_b/u_B)^2 + 1}} - \frac{(v_b/v_B)b + B}{\sqrt{(v_b/v_B)^2 + 1}} \right) = \pm \frac{2(u_b/u_B)b}{\sqrt{(u_b/u_B)^2 + 1}} \in V_m
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{u}{\|u\|_1} - \frac{v}{\|v\|_1} &= \frac{u_b b + u_B B}{\sqrt{u_b^2 + u_B^2}} - \frac{v_b b + v_B B}{\sqrt{v_b^2 + v_B^2}} \\
&= \pm \left(\frac{(u_b/u_B)b + B}{\sqrt{(u_b/u_B)^2 + 1}} + \frac{(v_b/v_B)b + B}{\sqrt{(v_b/v_B)^2 + 1}} \right) = \pm \frac{2B}{\sqrt{(u_b/u_B)^2 + 1}} \in V_M.
\end{aligned}
$$

That is, $(u, v) \in E_1 = E$.

Conversely, suppose that $(u, v) \in E$, set

$$
w = \frac{u}{\|u\|_1} + \frac{v}{\|v\|_1} \in V_m \quad \text{and} \quad W = \frac{u}{\|u\|_1} - \frac{v}{\|v\|_1} \in V_M,
$$

and observe that $w + W$ is in the direction of u and $w - W$ is in the direction of v. By Lemma 2.1, w and W are orthogonal with respect to both inner products. Thus,

$$
\cos \varphi = \frac{\langle w + W, w - W \rangle_1}{\|w + W\|_1 \|w - W\|_1} = \frac{\|w\|_1^2 - \|W\|_1^2}{\|w\|_1^2 + \|W\|_1^2}
$$

and

$$
\tan^2(\varphi/2) = \frac{1 - \cos \varphi}{1 + \cos \varphi} = \frac{\|W\|_1^2}{\|w\|_1^2}.
$$

A similar calculation yields the corresponding formula for ψ and leads to the conclusion,

$$
\tan^2(\psi/2) = \frac{\|W\|_2^2}{\|w\|_2^2} = \frac{M^2 \|W\|_1^2}{m^2 \|w\|_1^2} = (M/m)^2 \tan^2(\varphi/2).
$$

Taking square roots establishes equality in the right-hand inequality of (2.2).

Applying the right-hand inequality of (2.2) to the vectors u and $-v$ replaces φ by $\pi - \varphi$ and ψ by $\pi - \psi$ to give the conclusion,

$$
\cot(\psi/2) = \tan(\pi/2 - \psi/2) \le (M/m) \tan(\pi/2 - \varphi/2) = (M/m) \cot(\varphi/2).
$$

This proves the left-hand inequality of (2.2), with equality if and only if $(u, -v) \in E$. This completes the proof in the case that V is a real vector space.

If V is a complex space and $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are complex inner products, the conclusion of the theorem follows by applying the result just proved to the real vector space $V_{\mathbb{R}}$ equipped with the real inner products $\mathrm{Re}\langle \cdot, \cdot \rangle_1$ and $\mathrm{Re}\langle \cdot, \cdot \rangle_2$. This completes the proof. $\qquad\square$

The angle between two subsets of V is defined as an infimum of angles between pairs of vectors. The inequality (2.2) remains valid when we take an infimum of all three terms so we have the following result. Note that since the cosine function is decreasing, the cosine of an infimum of angles is achieved by taking the supremum of their cosines.

Corollary 2.3. *Let V be a real or complex vector space equipped with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. Make definitions (2.1). For $S, T \subseteq V$, each containing at least one non-zero vector, let Φ and Ψ be the angles between the subsets S and T with respect to $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$, respectively. That is, $0 \leq \Phi \leq \pi$, $0 \leq \Psi \leq \pi$,*

$$\cos \Phi = \sup_{\substack{0 \neq u \in S \\ 0 \neq v \in T}} \frac{\mathrm{Re}\langle u, v \rangle_1}{\|u\|_1 \|v\|_1}, \quad and \quad \cos \Psi = \sup_{\substack{0 \neq u \in S \\ 0 \neq v \in T}} \frac{\mathrm{Re}\langle u, v \rangle_2}{\|u\|_2 \|v\|_2}. \qquad (2.6)$$

Then

$$(m/M) \tan(\Phi/2) \leq \tan(\Psi/2) \leq (M/m) \tan(\Phi/2).$$

The following theorem is our version of the generalized Wielandt inequality in inner product spaces. As pointed out earlier, the angles between the (real or complex) lines determined by u and v are often taken as alternative definitions of the angle between vectors themselves. We show that with this definition the results of Theorem 2.1 still hold, but the conditions for equality become slightly more complicated.

Theorem 2.2. *Let V be a real or complex vector space equipped with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. Make definitions (2.1). For independent vectors u and v in V let Φ and Ψ be defined by, $0 \leq \Phi \leq \pi/2$, $0 \leq \Psi \leq \pi/2$,*

$$\cos \Phi = \frac{|\langle u, v \rangle_1|}{\|u\|_1 \|v\|_1} \quad and \quad \cos \Psi = \frac{|\langle u, v \rangle_2|}{\|u\|_2 \|v\|_2}.$$

Then

$$(m/M) \tan(\Phi/2) \leq \tan(\Psi/2) \leq (M/m) \tan(\Phi/2). \qquad (2.7)$$

Let α_1 and α_2 be solutions to $|\langle u, v \rangle_1| = \alpha_1 \langle u, v \rangle_1$ and $|\langle u, v \rangle_2| = \alpha_2 \langle u, v \rangle_2$. Equality holds in the right-hand inequality of (2.7) if and only if $(\alpha_1 u, v) \in E$ and either $\alpha_1 = \alpha_2$ or $\langle u, v \rangle_2 = 0$. Equality holds in the left-hand inequality of (2.7) if and only if $(\alpha_2 u, -v) \in E$ and either $\alpha_1 = \alpha_2$ or $\langle u, v \rangle_1 = 0$.

Proof. Apply Corollary 2.3 to the lines $S = \mathbb{C}u$ and $T = \mathbb{C}v$ ($S = \mathbb{R}u$ and $T = \mathbb{R}v$ in the real case) to obtain (2.7). By (1.1), Φ is the angle between $\alpha_1 u$ and v with respect to $\langle \cdot, \cdot \rangle_1$ and Ψ is the angle between $\alpha_2 u$ and v with respect to $\langle \cdot, \cdot \rangle_2$. To analyse the right-hand inequality of (2.7), let θ be the angle between $\alpha_1 u$ and v with respect to $\langle \cdot, \cdot \rangle_2$. The infimum definition of Ψ and Theorem 2.1 show that

$$\tan(\Psi/2) \leq \tan(\theta/2) \leq (M/m) \tan(\Phi/2). \qquad (2.8)$$

By (1.1), the first of these is equality if and only if either $\alpha_1 = \alpha_2$ or $\langle u, v \rangle_2 = 0$. By Theorem 2.1, the second is equality if and only if $(\alpha_1 u, v) \in E$. Thus equality holds

in the right-hand inequality of (2.7) if and only if $(\alpha_1 u, v) \in E$ and either $\alpha_1 = \alpha_2$ or $\langle u, v \rangle_2 = 0$.

To analyse the left-hand inequality of (2.7), let θ be the angle between $\alpha_2 u$ and v with respect to $\langle \cdot, \cdot \rangle_1$. The infimum definition of Φ and Theorem 2.1 show that

$$(m/M) \tan(\Phi/2) \leq (m/M) \tan(\theta/2) \leq \tan(\Psi/2). \tag{2.9}$$

By (1.1), the first of these is equality if and only if either $\alpha_1 = \alpha_2$ or $\langle u, v \rangle_1 = 0$. By Theorem 2.1, the second is equality if and only if $(\alpha_2 u, -v) \in E$. Thus equality holds in the left-hand inequality of (2.7) if and only if $(\alpha_2 u, -v) \in E$ and either $\alpha_1 = \alpha_2$ or $\langle u, v \rangle_1 = 0$. $\qquad\square$

Inequalities (2.2) and (2.7) can be expressed in various equivalent forms. In terms of cosines (2.2) becomes, with $\chi = (M^2 - m^2)/(M^2 + m^2)$,

$$\frac{-\chi + \cos\varphi}{1 - \chi\cos\varphi} \leq \cos\psi \leq \frac{\chi + \cos\varphi}{1 + \chi\cos\varphi}. \tag{2.10}$$

Replace φ and ψ by Φ and Ψ to get the expression for (2.7). In terms of inner products instead of angles, the inequalities (2.2) of Theorem 2.1 and (2.7) of Theorem 2.2 become, in the case $\|u\|_1 = \|v\|_1 = 1$,

$$\frac{-\chi + \operatorname{Re}\langle u, v \rangle_1}{1 - \chi\operatorname{Re}\langle u, v \rangle_1} \leq \frac{\operatorname{Re}\langle u, v \rangle_2}{\|u\|_2\|v\|_2} \leq \frac{\chi + \operatorname{Re}\langle u, v \rangle_1}{1 + \chi\operatorname{Re}\langle u, v \rangle_1}. \tag{2.11}$$

and

$$\frac{-\chi + |\langle u, v \rangle_1|}{1 - \chi|\langle u, v \rangle_1|} \leq \frac{|\langle u, v \rangle_2|}{\|u\|_2\|v\|_2} \leq \frac{\chi + |\langle u, v \rangle_1|}{1 + \chi|\langle u, v \rangle_1|}, \tag{2.12}$$

respectively.

The special case $\Phi = \pi/2$ in Theorem 2.2 gives an inner product formulation of Wielandt's inequality that includes all cases of equality. Note that the right-hand inequality of (2.12) is equivalent to the left-hand inequality of (2.7).

Corollary 2.4. *Let V be a real or complex vector space equipped with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. Make definitions (2.1). Suppose the non-zero vectors $u, v \in V$ are orthogonal with respect to $\langle \cdot, \cdot \rangle_1$ and α satisfies $|\langle u, v \rangle_2| = \alpha\langle u, v \rangle_2$. Then,*

$$\frac{|\langle u, v \rangle_2|}{\|u\|_2\|v\|_2} \leq \frac{M^2 - m^2}{M^2 + m^2} \tag{2.13}$$

with equality if and only if $(\alpha u, -v) \in E$.

The following theorem gives upper and lower bounds on the difference between the cosines of φ and ψ. It improves the estimates given in Theorems 1 and 2 of [2].

Theorem 2.3. *Let V be a real or complex vector space equipped with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. Make definitions (2.1). For independent vectors u and v in V,*

$$-2\frac{M - m}{M + m} \leq \frac{\operatorname{Re}\langle u, v \rangle_2}{\|u\|_2\|v\|_2} - \frac{\operatorname{Re}\langle u, v \rangle_1}{\|u\|_1\|v\|_1} \leq 2\frac{M - m}{M + m} \tag{2.14}$$

and, if $\mathrm{Re}\langle u, v\rangle_1 \geq 0$, *then*

$$\frac{\mathrm{Re}\langle u, v\rangle_2}{\|u\|_2\|v\|_2} - \frac{\mathrm{Re}\langle u, v\rangle_1}{\|u\|_1\|v\|_1} \leq \frac{M^2 - m^2}{M^2 + m^2}. \tag{2.15}$$

Also,

$$-\frac{M^2 - m^2}{M^2 + m^2} \leq \frac{|\langle u, v\rangle_2|}{\|u\|_2\|v\|_2} - \frac{|\langle u, v\rangle_1|}{\|u\|_1\|v\|_1} \leq \frac{M^2 - m^2}{M^2 + m^2}. \tag{2.16}$$

Proof. Suppose φ and ψ are the angles between u and v with respect to $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$. Since,

$$\cos\psi - \cos\varphi = 2/(1 + \tan^2(\psi/2)) - 2/(1 + \tan^2(\varphi/2)),$$

Theorem 2.1 gives

$$\frac{2}{1 + (M/m)^2 x} - \frac{2}{1 + x} \leq \cos\psi - \cos\varphi \leq \frac{2}{1 + (m/M)^2 x} - \frac{2}{1 + x},$$

where $x = \tan^2(\varphi/2)$. A little calculus shows that the minimum value, over all $x \in [0, \infty]$, of the expression on the left occurs at $x = m/M$ and the maximum value, over all $x \in [0, \infty]$, of the expression on the right occurs at $x = M/m$. This gives (2.14). If $\mathrm{Re}\langle u, v\rangle_1 \geq 0$ then $\varphi \leq \pi/2$ and so $x = \tan^2(\varphi/2) \leq 1$. The maximum value on the right now occurs at $x = 1$, giving (2.15).

The same analysis, applied to the angles Φ and Ψ between the lines $\mathbb{C}u$ and $\mathbb{C}v$ (or $\mathbb{R}u$ and $\mathbb{R}v$ in the real case) includes the restriction $\tan^2(\Phi/2) \leq 1$ and gives the right-hand inequality in (2.16). The left-hand inequality follows from the right-hand one by interchanging the inner products $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$. Besides interchanging the angles φ and ψ, this has the effect of replacing m by $1/M$ and M by $1/m$ to give

$$\frac{|\langle u, v\rangle_1|}{\|u\|_1\|v\|_1} - \frac{|\langle u, v\rangle_2|}{\|u\|_2\|v\|_2} \leq \frac{(1/m)^2 - (1/M)^2}{(1/m)^2 + (1/M)^2} = \frac{M^2 - m^2}{M^2 + m^2}.$$

Multiplying through by -1 completes the proof. $\square$

In our notation, Dragomir's results from [2] are

$$1 - \frac{M^2}{m^2} \leq \frac{|\langle u, v\rangle_2|}{\|u\|_2\|v\|_2} - \frac{|\langle u, v\rangle_1|}{\|u\|_1\|v\|_1} \leq 1 - \frac{m^2}{M^2},$$

and, if $\mathrm{Re}\langle u, v\rangle_1 \geq 0$, then

$$1 - \frac{M^2}{m^2} \leq \frac{\mathrm{Re}\langle u, v\rangle_2}{\|u\|_2\|v\|_2} - \frac{\mathrm{Re}\langle u, v\rangle_1}{\|u\|_1\|v\|_1} \leq 1 - \frac{m^2}{M^2}.$$

Since

$$1 - \frac{M^2}{m^2} \leq -2\frac{M - m}{M + m} \leq -\frac{M^2 - m^2}{M^2 + m^2} \quad \text{and} \quad \frac{M^2 - m^2}{M^2 + m^2} \leq 1 - \frac{m^2}{M^2},$$

Theorem 2.3 improves on both of these statements.

Estimate (2.14), on the difference between the cosines of φ and ψ readily gives a lower bound on the product of those cosines.

Corollary 2.5. *Let V be a real or complex vector space equipped with inner products $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$. Make definitions (2.1). For independent vectors u and v in V,*

$$\frac{\mathrm{Re}\langle u,v\rangle_1}{\|u\|_1\|v\|_1}\frac{\mathrm{Re}\langle u,v\rangle_2}{\|u\|_2\|v\|_2} \geq -\left(\frac{M-m}{M+m}\right)^2. \tag{2.17}$$

Proof. Let $\mu = (M-m)/(M+m)$,

$$x = \frac{\mathrm{Re}\langle u,v\rangle_1}{\|u\|_1\|v\|_1}, \quad \text{and} \quad y = \frac{\mathrm{Re}\langle u,v\rangle_2}{\|u\|_2\|v\|_2}.$$

Note that $0 \leq \mu < 1$. By the Cauchy-Schwarz inequality and (2.14), the point (x, y) lies in the region defined by $-1 \leq x \leq 1$, $-1 \leq y \leq 1$, and $-2\mu \leq x - y \leq 2\mu$. Minimizing xy over this hexagonal region easily yields $(x, y) = (-\mu, \mu)$ or $(x, y) = (\mu, -\mu)$. Thus, $xy \geq -\mu^2$ as required. $\square$

3 Formulation in terms of matrices

The angle θ between vectors $x, y \in \mathbb{C}^n$ is defined by $0 \leq \theta \leq \pi$ and

$$\cos\theta = \frac{\mathrm{Re}\,y^*x}{\|x\|\|y\|}$$

and the angle Θ between the complex lines $\mathbb{C}x$ and $\mathbb{C}y$ satisfies $0 \leq \Theta \leq \pi/2$ and

$$\cos\Theta = \frac{|y^*x|}{\|x\|\|y\|}.$$

Let A be an invertible $n \times n$ matrix and consider the two inner products

$$\langle x,y\rangle_1 = y^*x \quad \text{and} \quad \langle x,y\rangle_2 = (Ay)^*(Ax) \tag{3.1}$$

on $\mathbb{C}^n$. Then the definitions in (2.1) show that $M = \|A\|$ and $1/m = \|A^{-1}\|$ so the condition number of A is $\kappa(A) = M/m$. Theorem 2.1 becomes the following.

Theorem 3.1. *Let A be an invertible $n \times n$ matrix. For independent $x, y \in \mathbb{C}^n$ let φ be the angle between x and y and let ψ be the angle between Ax and Ay. Then,*

$$\kappa(A)^{-1}\tan(\varphi/2) \leq \tan(\psi/2) \leq \kappa(A)\tan(\varphi/2).$$

Let λ_n and λ_1 denote the smallest and largest eigenvalues of A^*A. Then equality holds in the right-hand inequality above if and only if $x/\|x\| + y/\|y\|$ is in the λ_n-eigenspace of A^*A and $x/\|x\| - y/\|y\|$ is in the λ_1-eigenspace of A^*A. Also, equality holds in the left-hand inequality above if and only if $x/\|x\| - y/\|y\|$ is in the λ_n-eigenspace of A^*A and $x/\|x\| + y/\|y\|$ is in the λ_1-eigenspace of A^*A.

Theorem 2.2 gives a concise reformulation of the generalized Wielandt inequality. Since $\kappa(A) = \kappa(A^{-1})$, the symmetry between the angles Φ and Ψ is clear.

Theorem 3.2. *Let A be an invertible $n \times n$ matrix. For independent $x, y \in \mathbb{C}^n$ let Φ be the angle between the complex lines $\mathbb{C}x$ and $\mathbb{C}y$ and let Ψ be the angle between the complex lines $\mathbb{C}(Ax)$ and $\mathbb{C}(Ay)$. Then*

$$\kappa(A)^{-1}\tan(\Phi/2) \leq \tan(\Psi/2) \leq \kappa(A)\tan(\Phi/2).$$

It takes a bit of care to show the equivalence of this theorem with Theorem 1.1 because the angles Φ and Ψ represent subtly different concepts in the two statements. In Theorem 3.2, Φ and Ψ represent angles between given complex lines, while in Theorem 1.1 they represent bounds on those angles rather than the angles themselves. Also, one must apply Theorem 1.1 to A and to A^{-1} (or else to x, y and to $x, -y$) to obtain both sides of the inequality above.

The conclusion of Theorems 3.1 and 3.2 may be rewritten as

$$\frac{-\chi + \cos\varphi}{1 - \chi\cos\varphi} \leq \cos\psi \leq \frac{\chi + \cos\varphi}{1 + \chi\cos\varphi}, \tag{3.2}$$

where $\chi = (\kappa(A)^2 - 1)/(\kappa(A)^2 + 1)$. (Of course, φ and ψ should be replaced by Φ and Ψ when rewriting Theorem 3.2.)

We have omitted the characterization of the cases of equality in Theorem 3.2 but they can be readily obtained from Theorem 2.2. Conditions for equality in Theorem 2.1 are simpler than those in Theorem 2.2 because the former deals with angles between a single pair of vectors and the latter with an infimum of angles between vectors in two one-dimensional subspaces. To recognize when equality occurs in Theorem 2.1 one only has to consider the placement of the vectors u and v relative to the eigenspaces V_m and V_M. But equality in Theorem 2.2 requires that this infimum of angles be achieved for u and v in addition to requiring their correct placement with respect to these eigenspaces. In [9], Kolotilina gave the following characterization of the cases of equality in the generalized Wielandt inequality, without explicit recognition of this two-stage requirement. We give an alternative proof using Theorem 2.2. (Notice that the complex numbers ξ and η appearing in the Theorem of [9] are unnecessary as they may be absorbed into the eigenvectors x_1 and x_n.)

Proposition 3.1. *Let B be an $n \times n$ invertible Hermitian matrix, suppose $\lambda_1 > \lambda_n > 0$ are its largest and smallest eigenvalues, respectively, and set $\chi = (\lambda_1 - \lambda_n)/(\lambda_1 + \lambda_n)$. Fix independent $x, y \in \mathbb{C}^n$ and let $\cos\varphi = |y^*x|/(\|x\|\|y\|)$. Then*

$$|y^*Bx| = \frac{\chi + \cos\varphi}{1 + \chi\cos\varphi}\sqrt{x^*Bx}\sqrt{y^*By} \tag{3.3}$$

if and only if

$$\begin{aligned}
\frac{x}{\|x\|} &= \frac{1}{\sqrt{2}}\left(\sqrt{1 + \cos\varphi}\,x_1 + \sqrt{1 - \cos\varphi}\,x_n\right), \quad and \\
\frac{y}{\|y\|} &= \frac{\varepsilon}{\sqrt{2}}\left(\sqrt{1 + \cos\varphi}\,x_1 - \sqrt{1 - \cos\varphi}\,x_n\right)
\end{aligned} \tag{3.4}$$

for some complex number ε of unit modulus and some unit eigenvectors x_1 and x_n satisfying $Bx_1 = \lambda_1 x_1$ and $Bx_n = \lambda_n x_n$.

Proof. With $A = B^{1/2}$ we have $B = A^*A$. Apply Theorem 2.2 to inner products (3.1) and note that $M = \lambda_1$ and $m = \lambda_n$ so V_M and V_m are the λ_1- and λ_n-eigenspaces of B, respectively. Using (2.10), we see that (3.3) is equivalent to equality in the left hand inequality of (2.7). Thus, Theorem 2.2 shows that (3.3) holds if and only if $(\alpha_2 x, -y) \in E$ and either $\alpha_1 = \alpha_2$ or $y^*x = 0$. As in Theorem 2.2, $|y^*x| = \alpha_1 y^*x$ and $|(Ay)^*(Ax)| = \alpha_2 (Ay)^*(Ax)$.

First suppose that x and y satisfy (3.4). A calculation, using the fact that x_1 and x_n are orthogonal, shows that $\varepsilon y^*x \geq 0$ and $\varepsilon (Ay)^*(Ax) \geq 0$. It follows that either $\alpha_1 = \alpha_2 = \varepsilon$ or $y^*x = 0$. Also,

$$\frac{\varepsilon x}{\|\varepsilon x\|} + \frac{-y}{\| - y\|} = \sqrt{2}\varepsilon\sqrt{1 - \cos\varphi}\, x_n \in V_m$$

and

$$\frac{\varepsilon x}{\|\varepsilon x\|} - \frac{-y}{\| - y\|} = \sqrt{2}\varepsilon\sqrt{1 + \cos\varphi}\, x_1 \in V_M$$

so $(\alpha_2 x, -y) \in E$.

Conversely, suppose that $(\alpha_2 x, -y) \in E$ and either $\alpha_1 = \alpha_2$ or $y^*x = 0$. Set $\varepsilon = \alpha_2$. Then there exist $w \in V_m$ and $W \in V_M$ such that

$$\frac{\varepsilon x}{\|x\|} - \frac{y}{\|y\|} = w \quad \text{and} \quad \frac{\varepsilon x}{\|x\|} + \frac{y}{\|y\|} = W.$$

Since w and W are orthogonal, the parallelogram law gives $\|W\|^2 + \|w\|^2 = 4$ and the definition of φ gives $\|W\|^2 - \|w\|^2 = 4\cos\varphi$. Solving these two equations yields, $\|W\| = \sqrt{2}\sqrt{1 + \cos\varphi}$ and $\|w\| = \sqrt{2}\sqrt{1 - \cos\varphi}$. With $x_1 = \bar{\varepsilon}W/\|W\|$ and $x_n = \bar{\varepsilon}w/\|w\|$ we have (3.4). This completes the proof. □

In Theorem 3 of [15], Yeh gave a different generalization of the Wielandt inequality for angles between complex lines. Here we show that Theorem 3.2 gives the stronger inequality.

Theorem 3.3. *[15] Let A be an invertible $n \times n$ matrix. For independent $x, y \in \mathbb{C}^n$ let Φ be the angle between the complex lines $\mathbb{C}x$ and $\mathbb{C}y$ and let Ψ be the angle between the complex lines $\mathbb{C}(Ax)$ and $\mathbb{C}(Ay)$. Define θ by $0 \leq \theta \leq \pi/2$ and $\cot(\theta/2) = \kappa(A)$. If $\cos\Phi \leq 1/\kappa(A)^2$, then*

$$\cos\Psi \leq \cos\theta + 2\cos^2(\theta/2)\cos\Phi. \tag{3.5}$$

Proof. By Theorem 3.2 and (3.2), it is enough to show that

$$\frac{\chi + \cos\Phi}{1 + \chi\cos\Phi} \leq \cos\theta + (1 + \cos\theta)\cos\Phi,$$

where

$$\chi = \frac{\kappa(A)^2 - 1}{\kappa(A)^2 + 1} = \frac{\cot^2(\theta/2) - 1}{\cot^2(\theta/2) + 1} = \cos\theta.$$

But both χ and $\cos\Phi$ are positive, so

$$\frac{\chi + \cos\Phi}{1 + \chi\cos\Phi} \leq \chi + \cos\Phi \leq \chi + (1 + \chi)\cos\Phi$$

as required. □

In Theorem 3.1 of [14], Yan generalized the Wielandt inequality for real symmetric matrices as follows.

Theorem 3.4. *[14] Let B be a real $n \times n$ symmetric positive definite matrix with eigenvalues $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n > 0$. For independent $x, y \in \mathbb{R}^n$ define Φ by $0 \leq \Phi \leq \pi/2$ and $\|x\|\|y\| \cos \Phi = |y^T x|$. Then,*

$$|x^T By| \leq \left(\max_{i,j} \frac{\lambda_i \cos^2(\Phi/2) - \lambda_j \sin^2(\Phi/2)}{\lambda_i \cos^2(\Phi/2) + \lambda_j \sin^2(\Phi/2)} \right) \sqrt{x^T Bx} \sqrt{y^T By}. \qquad (3.6)$$

It was left as a conjecture in [14] that the theorem remains true for complex vectors x and y and a positive definite Hermitian matrix B.

It is routine to verify that the expression

$$\frac{s \cos^2(\Phi/2) - t \sin^2(\Phi/2)}{s \cos^2(\Phi/2) + t \sin^2(\Phi/2)}$$

is increasing in s and decreasing in t. Thus, the maximum in (3.6) is achieved when $i = 1$ and $j = n$, where it takes the value,

$$\frac{\lambda_1 \cos^2(\Phi/2) - \lambda_n \sin^2(\Phi/2)}{\lambda_1 \cos^2(\Phi/2) + \lambda_n \sin^2(\Phi/2)} = \frac{\chi + \cos \Phi}{1 + \chi \cos \Phi}.$$

Here $\chi = (\lambda_1/\lambda_n - 1)/(\lambda_1/\lambda_n + 1)$. If $A = B^{1/2}$, then $\kappa(A)^2 = \kappa(B) = \lambda_1/\lambda_n$ so Theorem 3.2 and (3.2) implies that Theorem 3.4 holds in both the real and complex cases, confirming Yan's conjecture.

We end with an improvement of Lemma 2.2 from [11]. It follows directly from Corollary 2.5 with $\langle x, y \rangle_1 = y^T Ax$ and $\langle x, y \rangle_2 = y^T Bx$.

Lemma 3.1. *Suppose A and B are real symmetric positive definite $n \times n$ matrices and let $\kappa = \kappa(A^{-1/2}BA^{-1/2})$. Then for $x, y \in \mathbb{R}^n$ with $y \neq 0$,*

$$\frac{y^T Ax}{\sqrt{x^T Ax}\sqrt{y^T Ay}} \frac{y^T Bx}{\sqrt{x^T Bx}\sqrt{y^T By}} \geq -\left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^2.$$

The above inequality followed by the AM-GM inequality give the conclusion of Lemma 2.2 from [11]:

$$2\frac{y^T Ax}{y^T Ay} \frac{y^T Bx}{y^T By} \geq -2\left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^2 \left(\frac{x^T Ax}{y^T Ay} \frac{x^T Bx}{y^T By} \right)^{1/2}$$

$$\geq -\left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^2 \left(\frac{x^T Ax}{y^T Ay} + \frac{x^T Bx}{y^T By} \right).$$

References

[1] F.L. Bauer and A.S. Householder, Some inequalities involving the euclidean condition of a matrix. Math. 2 (1960), 308–311.

[2] R. Bhatia and C. Davis, More operator versions of the Schwarz inequality. Commun. Math. Phys. 215 (2000), 239–244.

[3] D. Boyd, Best constants in a class of integral inequalities. Pac. J. Math. 30 (1969), 367–383.

[4] S.S. Dragomir, Inner product inequalities for two equivalent norms and applications. Acta. Math. Vietnam. 34 (2009), 361–369.

[5] M.L. Eaton, A maximization problem and its application to canonical correlation. J. Multivariate Anal. 6 (1976), 422–425.

[6] M.L. Eaton and D. Tyler, The asymptotic distribution of singular values with applications to canonical correlations and correspondence analysis. J. Multivariate Anal. 50 (1994), 238–264.

[7] R.A. Horn and C.R. Johnson, *Matrix Analysis.* Cambridge University Press, London, 1985.

[8] A.S. Househoulder, *The Theory of Matrices in Numerical Analysis.* Blaisdell, New York, 1964.

[9] L.Yu. Kolotilina, The case of equality in the generalized Wielandt inequality. J. Math. Sci. (N. Y.) 114 (2003), 1803–1807.

[10] S. Liu, C. Lu, and S. Puntanen, Matrix trace Wielandt inequalities with statistical applications. J. Statist. Plann. Inference 139 (2009), 2254–2260.

[11] M. Lin and G. Sinnamon, A condition for convexity of a product of positive definite quadratic forms. SIAM J. Matrix Anal. Appl. 32 (2011), 457–462.

[12] S.-G. Wang and W.-C. Ip, A matrix version of the Wielandt inequality and its applications to statistics. Linear Algebra Appl. 296 (1999), 171–181.

[13] H. Wielandt, Inclusion theorems for eigenvalues., National Bureau of Standards Appl. Math. Series 29 (1953), 75–78 .

[14] Z. Yan, A unified version of Cauchy-Schwarz and Wielandt inequalities. Linear Algebra Appl. 428 (2008), 2079–2084.

[15] L. Yeh, A note on Wielandt's inequality. Appl. Math. Lett. 8 (1995), 29–31.

[16] B. X. Zhang and X. H. Zhu, Generalized matrix versions of the constrained Kantorovich and Wielandt inequalities. Acta Math. Sinica (Chin. Ser.) 45 (2002), 151–156.

Minghua Lin
Department of Combinatorics and Optimization
University of Waterloo
Waterloo, Canada
E-mail: mlin87@ymail.com

Gord Sinnamon
Department of Mathematics
University of Western Ontario
London, Canada
E-mail: sinnamon@uwo.ca