Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Generalized Wielandt and Cauchy-Schwarz Inequalities

Mohammed A. Hasan
Department of Electrical & Computer Engineering
University of Minnesota Duluth
Email:mhasan@d.umn.edu

Abstract

The Wielandt inequality is important in many applications. It involves functions of the extreme eigenvalues of a positive definite matrix. In this paper, we derive a few extensions of the Wielandt inequality and new inequalities involving the two largest and two smallest eigenvalues. The resulting inequalities are shown to be the best possible. A unified approach involving constrained optimization techniques are used to derive these results. The proposed inequalities are then utilized to obtain several bounds for the extremum eigenvalues and eigen spread of real symmetric matrices. A collection of bounds for functions of the eigenvalues of positive definite and general symmetric matrices are then derived in terms of the entries of the matrix. Additionally, lower bounds for the condition number of positive definite matrices as well as lower bounds for the minimum separation of eigenvalues are developed.

1 Introduction

The Wielandt inequality is an improvement on the general Cauchy-Schwarz inequality which asserts that if A is a positive definite matrix, then

$$|x^TAy| \leq \sqrt{(x^TAx)(y^TAy)}, \tag{1a}$$

for every pair of vectors x and y. There are several matrix versions of the Cauchy-Schwarz inequalities (1a) in the literature. In the last decade considerable progress has been made in deriving many equivalent forms. Mond and Pecaric [1] and Pecaric et al. [2] derived some general Cauchy-Schwarz type inequalities. Many other related versions can be found in [3]-[9].

In this paper, further improvement on the Cauchy-Schwarz and the Wielandt inequalities will be developed. The key idea here is to use optimization techniques to derive bounds for functions of the eigenvalues of a matrix. Some of these results apply to positive definite or semidefinite matrices while others apply to general symmetric matrices. The proposed approach is very effective in deriving many other matrix inequalities.

In the following sections we use the following notation. The vector e_i denotes the ith column of an identity matrix. The magnitude of a vector x will be denoted by $||x|| = \sqrt{x^Tx}$. The notation I denotes an identity matrix of appropriate size.

2. Wielandt-Like Inequalities

In this section we derive a few extensions of the well-known Wielandt inequality. Let $A \in \mathbb{R}^{n\times n}$ be a positive definite matrix and let x and $y \in \mathbb{R}^{n\times 1}$ be two vectors such that $|x| = |y| = 1$ and $x^Ty = 0$. Assume that the eigenvalues of A, in increasing order, are $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$, then the Wielandt inequality asserts that

$$|x^TAy| \leq \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\sqrt{(x^TAx)(y^TAy)}. \tag{1b}$$

Moreover, there exists an orthonormal pair of vectors x; y for which equality holds. This inequality is an improvement of that of Cauchy-Schwarz. A matrix version of (1a) and its applications to statistics is presented in [3]. Also, as in [4], it is an extension of the Kantorovich inequality. In the next development, we derive a few generalizations of (1b).

2.1 Generalization 1

Let x and y be unit vectors, then $\frac{x+y}{\sqrt{2(1+x^Ty)}}$ and $\frac{x-y}{\sqrt{2(1-x^Ty)}}$ are orthonormal vectors and therefore (1b) simplifies as follows

$$\frac{(x+y)^T}{\sqrt{2(1+x^Ty)}}A\frac{(x-y)}{\sqrt{2(1-x^Ty)}} \leq \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\times \sqrt{\frac{(x+y)^T}{\sqrt{2(1+x^Ty)}}A\frac{(x+y)}{\sqrt{2(1+x^Ty)}}} \times\sqrt{\frac{(x-y)^T}{\sqrt{2(1-x^Ty)}}A\frac{(x-y)}{\sqrt{2(1-x^Ty)}}}$$

or

$$|x^TAx - y^TAy| \leq \frac{\lambda_n - \lambda_1}{(\lambda_n + \lambda_1)}\sqrt{(x^TAx + y^TAy)^2 - 4(x^TAy)^2}. \tag{1c}$$

Consequently,

$$(x^TAx - y^TAy)^2 \leq (\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1})^2\times \{(x^TAx)^2 + 2(x^TAx)(y^TAy) + (y^TAy)^2 - 4(x^TAy)^2\}. \tag{1d}$$

By rearraning the terms of (1d), we obtain the following generalization of the Cauchy-Schwarz and Wielandt inequalities:

$$(x^TAy)^2 \leq \{\frac{x^TAx + y^TAy}{2} - \frac{\lambda_n + \lambda_1}{\lambda_n - \lambda_1}\frac{x^TAx - y^TAy}{2}\}\{\frac{x^TAx + y^TAy}{2} + \frac{\lambda_n + \lambda_1}{\lambda_n - \lambda_1}\frac{x^TAx - y^TAy}{2}\}. \tag{1e}$$

Note that this inequality holds for any two vectors having the same length. By setting $x = e_i$ and $y = e_j$, $i \neq j$, we obtain the following

$$a_{ii}a_{jj} - a_{ij}^2 \geq \frac{\lambda_1\lambda_n}{(\lambda_1 + \lambda_n)^2}(a_{ii} - a_{jj})^2.$$

0-7803-8335-4/04/$17.00 ©2004 AACC

The last inequality can be used to derive bounds for the condition number of A. Let $k = \frac{\lambda_n}{\lambda_1}$ be the condition number of A, then bounds on k can be obtained from solving the inequality

$$\frac{(a_{ii} - a_{jj})^2}{a_{ii}a_{jj} - a_{ij}^2} \leq \frac{(1+k)^2}{k}.$$

Since A is positive definite, the denominator of the left hand side is always positive.

2.2 Generalization 2

In this section we derive a second extension of the Wielandt inequality.

Theorem 1. *Let $A \in \mathbb{R}^{n\times n}$ be a positive definite matrix and let x and $y \in \mathbb{R}^{n\times 1}$ be two vectors such that $|x| = |y| = 1$ and $x^T y = 0$. Assume that the eigenvalues of A, in increasing order, are $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$ with corresponding eigenvectors $q_1, \cdots, q_n$. Then*

$$|x^T A y|^2 \leq \max\{(\frac{\lambda_i - \lambda_j}{\lambda_i^r + \lambda_j^r})^2\}_{\substack{i,j=1\\ i\neq j}}^n (x^T A^r x)(y^T A^r y). \quad (2)$$

Proof: We prove this result for the case where all eigenvalues are distinct. Consider the optimization problem

$$\{\text{Maximize } \frac{(x^T Ay)(y^T Ax)}{(x^T A^r x)(y^T A^r y)} : x^T y = 0\}. \quad (3a)$$

which is equivalent to the optimization problem

$$\{\text{Maximize } (x^T Ayy^T Ax) : x^T A^r x = 1, y^T A^r y = 1 : x^T y = 0\}. \quad (3b)$$

The Lagrangian of this problem is given by

$$\mathcal{L} = \frac{1}{2}(x^T Ayy^T Ax) - \mu_1(x^T A^r x - 1) - \frac{\mu_2}{2}(y^T A^r y - 1) - \mu_3 x^T y. \quad (3c)$$

The first order necessary condition for optimality implies

$$\nabla_{x,y}\mathcal{L} = \begin{bmatrix} (y^T Ax)Ay - \mu_1 A^r x - \mu_3 y \\ (x^T Ay)Ax - \mu_2 A^r y - \mu_3 x \end{bmatrix} = 0. \quad (3d)$$

Thus the optimal solutions of (3b) satisfy the following relations:

$$\mu_1 = \mu_2 = \alpha^2 = (y^T Ax)^2 = \mu$$
$$\mu_3 = (y^T Ax)(y^T A^r y) = (y^T Ax)(x^T A^r x),$$

where $\alpha = y^T Ax = x^T Ay$. Note that $x^T A^r x = (y^T A^r y)$ provided that $y^T Ax \neq 0$. Clearly, if $y^T Ax = 0$, then x and y are not optimal solutions since one can always find two orthogonal vectors for which $y^T Ax \neq 0$. Therefore (3d) can be expressed as

$$(\alpha A - \mu_3 I)y = \mu A^r x,$$
$$(\alpha A - \mu_3 I)x = \mu A^r y. \quad (3e)$$

To solve these two equations, let $x = \sum_{i=1}^n a_i q_i$, $y = \sum_{i=1}^n b_i q_i$ and $D = \text{diag}(\lambda_1, \cdots, \lambda_n)$. This implies that

$$(\alpha D - \mu_3 I)b = \mu D^r a,$$
$$(\alpha D - \mu_3 I)a = \mu D^r b,$$

where $a = [\,a_1 \ \cdots \ a_n\,]$ and $b = [\,b_1 \ \cdots \ b_n\,]$. Equivalently,

$$(\alpha d_i - \mu_3)b_i = \mu d_i^r a_i,$$
$$(\alpha d_i - \mu_3)a_i = \mu d_i^r b_i,$$

for $i = 1, \cdots, n$. Since the eigenvalues of A are distinct, it follows that $a_k \neq 0$ for exactly two indices $k = i, j$ and $a_k = 0$ otherwise, and that $a_i^2 = b_i^2$, $a_j^2 = b_j^2$, and $a_i b_i + a_j b_j = 0$. Thus $a = \pm\frac{e_i + e_j}{\sqrt{2}}$ and $b = \pm\frac{e_i - e_j}{\sqrt{2}}$. Therefore the solution of this system of equation is of the form

$$\mu = \alpha\frac{\lambda_i - \lambda_j}{\lambda_i^r + \lambda_j^r},$$
$$\mu_3 = \alpha\lambda_i\lambda_j\frac{\lambda_i^{r-1} + \lambda_j^{r-1}}{\lambda_i^r + \lambda_j^r}, \quad (3f)$$

for some $i, j = 1, \cdots, n$. Since $\mu = \alpha^2$, it follows that $\mu = (\frac{\lambda_i - \lambda_j}{\lambda_i^r + \lambda_j^r})^2$. It can be shown that the Hessian matrix $\nabla_{x,y}\mathcal{L}$ is definite on the null space of the gradient of the constraints. This proves the generalized Wielandt inequality. It should be noted that the maximum of $\mu = (\frac{\lambda_i - \lambda_j}{\lambda_i^r + \lambda_j^r})^2$ occurs at $(i, j) = (1, n)$ when $r = 0, 1$. Howvere, this may not be the case if $r \geq 2$.

Remark 1: The classical Wielandt inequality of (1b) can be obtained by setting $r = 1$, in which case (3e) can be rewritten as

$$(A^{-1} - \frac{\alpha}{\mu_3 I})y = \frac{-\mu}{\mu_3}x$$
$$(A^{-1} - \frac{\alpha}{\mu_3 I})x = \frac{-\mu}{\mu_3}y. \quad (3g)$$

As shown in Lemma 3 below, it follows that

$$\frac{\alpha}{\mu_3} = \frac{\lambda_i^{-1} + \lambda_j^{-1}}{2},$$
$$\frac{-\mu}{\mu_3} = \frac{\lambda_i^{-1} - \lambda_j^{-1}}{2}.$$

Solving these two equations for μ and μ_3, we obtain

$$\mu = (\frac{\lambda_i^{-1} - \lambda_j^{-1}}{\lambda_i^{-1} + \lambda_j^{-1}})^2 = (\frac{\lambda_j - \lambda_i}{\lambda_i + \lambda_j})^2,$$

for some $i, j = 1, \cdots, n$. The maximum occurs at $(i, j) = (1, n)$ or $(i, j) = (n, 1)$ since for each (i, j)

$$\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} - \frac{\lambda_i - \lambda_j}{\lambda_i + \lambda_j} = \frac{2(\lambda_n\lambda_j - \lambda_i\lambda_1)}{(\lambda_i + \lambda_j)(\lambda_n + \lambda_1)} > 0,$$

i.e., $\max\{(\frac{\lambda_i - \lambda_j}{\lambda_i + \lambda_j})^2\}_{i,j=1}^n$ occurs at $(i, j) = (1, n)$ or $(i, j) = (n, 1)$.

For $r = 0$ another useful inequality is obtained as shown next.

Proposition 2. *Let A, $x, y, \lambda_1, \lambda_n$ be as in Theorem 1, then*

$$\min_{\substack{i,j=1\\ i\neq j}}\{|\frac{\lambda_i - \lambda_j}{2}|\} \leq |x^T Ay| \leq \frac{\lambda_n - \lambda_1}{2}. \quad (4)$$

To prove Proposition 2, the following Lemma is needed.

Lemma 3. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix of size n. Assume that the eigenvalues of A are $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ with corresponding eigenvectors $q_1, \cdots, q_n$. Let $x, y \in \mathbb{R}^n$ such that $||x|| = ||y|| = 1$ and $x^T y = 0$. Then the solution of the system*

$$(A - \mu_1 I)x = \mu_2 y,$$
$$(A - \mu_1 I)y = \mu_2 x.$$

has the form

$$\mu_1 = \frac{\lambda_i+\lambda_j}{2},\ x = \frac{q_i+q_j}{\sqrt{2}},$$
$$\mu_2 = \frac{\lambda_i-\lambda_j}{2},\ y = \frac{q_i-q_j}{\sqrt{2}},$$

where $i,j = 1,\cdots,n$.

Proof of Proposition 2: We solve the optimization problem

$$\{\text{Optimize } x^TAy : x^Tx = 1, y^Ty = 1, x^Ty = 0\}. \qquad (5a)$$

Let $x = \sum_{k=1}^n \alpha_k q_k$ and $y = \sum_{k=1}^n \beta_k q_k$, where $\{q_k\}_{k=1}^n$ is the set of eigenvectors of A so that $Aq_k = \lambda_k q_k$, for $k = 1,\cdots,n$. Thus the above optimization problem can be expressed as

$$\{\text{Optimize } \alpha^TD\beta : \alpha^T\alpha = 1, \beta^T\beta = 1, \alpha^T\beta = 0\}. \qquad (5b)$$

Let $\mathcal{L}$ be the Lagrangian of this problem given by

$$\mathcal{L} = \alpha^TD\beta - \frac{\mu_1}{2}(\alpha^T\alpha-1) - \frac{\mu_2}{2}(\beta^T\beta-1) - \frac{\mu_3}{2}\alpha^T\beta, \quad (5c)$$

where $D = diag\{\lambda_1,\cdots,\lambda_n\}$. We will assume that eigenvalues are simple. The first order necessary condition for optimality implies

$$\nabla_{\alpha,\beta}\mathcal{L} = \begin{bmatrix} D\beta-\mu_1\alpha-\mu_3\beta \\ D\alpha-\mu_2\beta-\mu_3\alpha \end{bmatrix} = 0. \qquad (5d)$$

To solve these equations for α and β, we note that at optimal solutions the following holds:

$$\mu_1 = \alpha^TD\beta,\ \mu_2 = \beta^TD\alpha,$$
$$\mu_3 = \alpha^TD\alpha,\ \mu_3 = \beta^TD\beta.$$

Consequently, $\mu_1 = \mu_2 = \mu$, and $\alpha^TD\alpha = \beta^TD\beta$. Since eigenvalues of D are distinct, the last equality implies that α and β can not be eigenvectors of D. Now,

$$\begin{aligned}(D-\mu_3I)\beta &= \mu\alpha,\\ (D-\mu_3I)\alpha &= \mu\beta.\end{aligned} \qquad (5e)$$

Equations (5d) yield

$$\begin{aligned}(D-\mu I)^2\alpha &= \mu^2\alpha\\ (D-\mu I)^2\beta &= \mu^2\beta.\end{aligned} \qquad (5f)$$

This means that

$$\begin{aligned}(\lambda_i-\mu_3)^2\alpha_i &= \mu^2\alpha_i,\\ (d_i-\mu_3)^2\beta_i &= \mu^2\beta_i.\end{aligned} \qquad (5g)$$

for $i = 1,\cdots,n$. If $\alpha_k \neq 0$ and $\beta_k \neq 0$ for some $k = i,j$, then $(\lambda_i-\mu_3)^2 = \mu^2$, $(\lambda_j-\mu_3)^2 = \mu^2$, $\alpha_i^2 = \beta_i^2$, and $\alpha_j^2 = \beta_j^2$. The assumption that the eigenvalues of A are distinct implies that $a_k = \beta_k = 0$ for $k \neq i,j$. Hence, $\mu = \frac{\lambda_i+\lambda_j}{2}$ and $\mu_3 = \frac{\lambda_i-\lambda_j}{2}$. Now, since α and β are unit vectors and orthogonal,

$$\alpha = \alpha_ie_i + \alpha_ie_j = \frac{e_i+e_j}{\sqrt{2}},$$

and

$$\beta = \beta_ie_i + \beta_je_j = \alpha_ie_i - \alpha_ie_j = \frac{e_i-e_j}{\sqrt{2}}. \qquad (5h)$$

Clearly, $\alpha^TD\beta = \frac{\lambda_i-\lambda_j}{2}$, which is maximum if $i = n$ and $j = 1$. Thus this maximum is attained at $\alpha = \frac{e_n+e_1}{\sqrt{2}}$, and $\beta = \frac{e_n-e_1}{\sqrt{2}}$. It can be shown that the Hessian $\nabla^2\mathcal{L}$ is negative definite on the null space of the gradient of the constraints. Thus this solution is a maximizer of (5a). Similar analysis may apply to show that the minimum of x^TAy is of the form $\frac{\lambda_i-\lambda_j}{2}$.

Remark 2: Assume that $x^Ty = 0$, then $x^T(A+cI)y = x^TAy$ for each positive number c. Thus Proposition 2 is also valid for general symmetric matrices.

Corollary 4. *Let A, λ_1, λ_n be as in Theorem 1 and let x, y be unit vectors, then*

(a) $\min_{\substack{i,j=1\\ i\neq j}}\{|\frac{\lambda_i-\lambda_j}{2}|\} \leq x^TA^2x - (x^TAx)^2 \leq (\frac{\lambda_n-\lambda_1}{2})^2$

(b) $|x^TAx - y^TAy| \leq (\lambda_n-\lambda_1)\sqrt{1-(x^Ty)^2}$

(c) *Let B be a symmetric matrix, then for each unit vectors x and z the following hold*

$$|x^TABz - (x^TAx)(x^TBz| \leq (\frac{\lambda_n-\lambda_1}{2})(\frac{\mu_n-\mu_1}{2}), \qquad (6)$$

where μ_n and μ_1 is the largest and smallest eigenvalues of B.

(d) $\frac{x^TA^3x-(x^TAx)(x^TA^2x)}{\sqrt{x^TA^4x-(x^TA^2x)^2}} \leq \frac{\lambda_n-\lambda_1}{2}$.

Proof: To prove (a), assume that $||x|| = 1$ and set $y = Ax - x^TAxx$. Then $x^Ty = 0$ and $||y|| = \sqrt{x^TA^2x-(x^TAx)^2}$. Thus,

$$x^TA\frac{y}{|y|} = x^TA\frac{(Ax-x^TAxx)}{\sqrt{x^TA^2x-(x^TAx)^2}} = \frac{x^TA^2x-(x^TAx)^2}{\sqrt{x^TA^2x-(x^TAx)^2}}$$
$$= \sqrt{x^TA^2x-(x^TAx)^2} \leq \frac{\lambda_n-\lambda_1}{2}.$$

(b) Let x and y be unit vectors, then $\frac{x+y}{\sqrt{2(1+x^Ty)}}$ and $\frac{x-y}{\sqrt{2(1-x^Ty)}}$ are orthonormal vectors and therefore (4) simplifies to

$$\frac{(x+y)^T}{\sqrt{2(1+x^Ty)}}A\frac{(x-y)}{\sqrt{2(1-x^Ty)}} \leq \frac{\lambda_n-\lambda_1}{2}.$$

(c) Let $y = Bz - (x^TBz)x$, then $x^Ty = 0$ and $||y|| = \sqrt{z^TB^2z-(x^TBz)^2}$. It follows from Proposition 2 that

$$x^TA\frac{y}{||y||} = |\frac{x^TABz-(x^TAx)(x^TBz)}{\sqrt{z^TB^2z-(x^TBz)^2}}| \leq \frac{\lambda_n-\lambda_1}{2}.$$

(d) follows from (c) by setting $B = A^2$ ans $z = x$.

3. Inequalities Involving the Largest Two and Smallest Two Eigenvalues

In this section we use optimization techniques to derive matrix inequalities which are then used to provide bounds for the extremum eigenvalues of hermitian matrices.

Theorem 5. *Let $A \in R^{n\times n}$ $(n \geq 2)$ be a positive definite matrix and let x and $y \in R^{n\times 1}$ be two vectors such that $|x| = |y| = 1$ and $x^Ty = 0$. Assume that the eigenvalues of A, in increasing order, are $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$, then*

$$\left(\frac{\lambda_1+\lambda_2}{2}\right)^2 \leq (x^TAx)(y^TAy) \leq \left(\frac{\lambda_{n-1}+\lambda_n}{2}\right)^2. \quad (7a)$$

Proof: Consider the optimization problem

$$\{\text{Maximize } (x^TAx)(y^TAy) : x^Tx = 1, y^Ty = 1, x^Ty = 0\}.$$

Then the Lagrangian of this problem is given by

$$\mathcal{L} = \frac{1}{2}(x^TAx)(y^TAy) - \mu_1 x^Ty - \frac{\mu_2}{2}(x^Tx - 1) - \frac{\mu_3}{2}(y^Ty - 1), \tag{7b}$$

where μ_1, μ_2 and μ_3 are Lagrange multipliers. The first order necessary condition for optimality is

$$\nabla_{x,y}\mathcal{L} = \begin{bmatrix} (y^TAy)Ax - \mu_1 y - \mu_2 x \\ (x^TAx)Ay - \mu_1 x - \mu_3 y \end{bmatrix} = 0. \tag{7c}$$

At optimal solutions, the following hold:

$$\mu_2 = \mu_3 = (x^TAx)(y^TAy)$$

$$\mu_1 = (y^TAy)(x^TAy) = (x^TAx)(y^TAx),$$

i.e. $x^TAx = y^TAy$ provided that $x^TAy = y^TAx \neq 0$. It follows from Lemma 3 that $x = \alpha_1 q + \alpha_2 p$ and $y = \beta_1 q + \beta_2 p$, where p and q are two distinct unit eigenvectors of A corresponding to the eigenvalues λ and μ. Note that $||p|| = ||q|| = 1$ and $p^Tq = 0$. The optimality conditions yield:

$$\lambda\alpha_1^2 + \mu\alpha_2^2 = \lambda\beta_1^2 + \mu\beta_2^2,$$

$$\alpha_1^2 + \alpha_2^2 = 1, \quad \beta_1^2 + \beta_2^2 = 1, \quad \alpha_1\beta_1 + \alpha_2\beta_2 = 0.$$

Hence,

$$\lambda\alpha_1^2 + \mu(1 - \alpha_1^2) = \lambda\beta_1^2 + \mu(1 - \beta_1^2).$$

$\alpha_1^2(\lambda - \mu) = \beta_1^2(\lambda - \mu)$ or $\alpha_1^2 = \beta_1^2$ which implies that $\alpha_1 = \mp\beta_1$. Similarly, $\alpha_2^2 = \beta_2^2$ or $\alpha_2 = \mp\beta_2$,

$$0 = \alpha_1\beta_1 + \alpha_2\beta_2 = \alpha_1^2 - \alpha_2^2$$

and therefore, $\alpha_1 = \mp\alpha_2$. This yields $x = \pm\frac{q+p}{\sqrt{2}}, \quad y = \pm\frac{q-p}{\sqrt{2}}$. Hence $(x^TAx)(y^TAy) = (\frac{\mu+\lambda}{2})^2$. This quantity is maximum if $x = \pm\frac{q_n + q_{n-1}}{\sqrt{2}}$, and $y = \pm\frac{q_n - q_{n-1}}{\sqrt{2}}$, in which case $(x^TAx)(y^TAy) = (\frac{\lambda_n + \lambda_{n-1}}{2})^2$. Similarly, the minimum of $(x^TAx)(y^TAy) = (\frac{\lambda_1+\lambda_2}{2})^2$, which is attained at $x = \pm\frac{q_1+q_2}{\sqrt{2}}, \quad y = \pm\frac{q_1-q_2}{\sqrt{2}}$.

It can be verified that the second order optimality condition implies $\nabla^2\mathcal{L}$ is negative definite on the null space of the gradient of the constraints.

Corollary 6. *Let $A \in R^{n\times n}$ be a positive definite matrix and let x and $y \in R^{n\times 1}$ be two vectors such that $||x|| = ||y|| = 1$. Assume that the eigenvalues of A, in increasing order, are $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$, then*

$$(\frac{\lambda_1 + \lambda_2}{2})^2\sqrt{1 - (x^Ty)^2} \leq (x^TAx + y^TAy)^2 - 4(x^TAy)^2 \leq (\frac{\lambda_n + \lambda_{n-1}}{2})^2\sqrt{1 - (x^Ty)^2}. \tag{8a}$$

Proof: The vectors $\frac{x+y}{\sqrt{2(1+x^Ty)}}$ and $\frac{x-y}{\sqrt{2(1-x^Ty)}}$ are orthonormal and hence Theorem 5 guarantees that

$$(\frac{(x+y)^T}{\sqrt{2(1+x^Ty)}}A\frac{x+y}{\sqrt{2(1+x^Ty)}})(\frac{(x-y)^T}{\sqrt{2(1-x^Ty)}}A\frac{x-y}{\sqrt{2(1-x^Ty)}}) \leq (\frac{\lambda_n + \lambda_{n-1}}{2})^2.$$

Equivalently,

$$\{x^TAx + y^TAy + 2(x^TAy)\}\{(x^TAx + y^TAy) - 2(x^TAy)\} \leq (\lambda_n + \lambda_{n-1})^2\sqrt{1 - (x^Ty)^2}. \tag{8b}$$

Theorem 5 and Corollary 6 can be applied to deduce lower bounds for the maximum eigenvalues of a positive definite matrix. Let $A = [a_{ij}]$ and let $x = \frac{e_i + e_j}{\sqrt{2}}$ and $y = \frac{e_i - e_j}{\sqrt{2}}$, then it implies from Corollary 6 that

$$(\lambda_n + \lambda_{n-1})^2 \geq (a_{ii} + a_{jj})^2 - 4a_{ij}^2.$$

This also implies that

$$\lambda_n \geq \sqrt{(a_{ii} + a_{jj})^2 - 4a_{ij}^2}.$$

Thus we have the following result:

Proposition 7. *Let A be a positive definite matrix of size n. Assume that the eigenvalues of A are $\lambda_1 < \lambda_2 \leq \cdots < \lambda_n$, then*

$$\lambda_n + \lambda_{n-1} \geq \max\{\sqrt{(a_{ii} + a_{jj})^2 - 4a_{ij}^2}\}_{i,j=1}^n,$$

$$\lambda_1 + \lambda_2 \leq \min\{\sqrt{(a_{ii} + a_{jj})^2 - 4a_{ij}^2}\}_{i,j=1}^n.$$

As a result, the following hold:

$$\lambda_n \geq \max\{\sqrt{(a_{ii} + a_{jj})^2 - 4a_{ij}^2}\}_{i,j=1}^n,$$

$$\lambda_1 \leq \min\{\sqrt{(a_{ii} + a_{jj})^2 - 4a_{ij}^2}\}_{i,j=1}^n.$$

The next result provides another version of Wielandt-like inequality.

Proposition 8. *Let $A \in R^{n\times n}$ be a positive definite matrix and let x and $y \in R^{n\times 1}$ be two vectors such that $||x|| = ||y|| = 1$ and $x^Ty = 0$. Assume that the eigenvalues of A, in increasing order, are $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$, then*

$$|x^TA^2y| \leq (\lambda_n - \lambda_1)\sqrt{(x^TAx)(y^TAy)}.$$

It is known that $|x^TAy| \leq \frac{x^TAx + y^TAy}{2}$ for every two vectors x and y. In the next proposition we generalize this inequality for the case where x and y are orthogonal.

Proposition 9. *Let $A \in R^{n\times n}$ be a positive definite matrix and let x and $y \in R^{n\times 1}$ be two vectors such that $||x|| = ||y||$ and $x^Ty = 0$. Assume that the eigenvalues of A, in increasing order, are $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$, then*

$$|x^TAy| \leq \frac{1}{2}\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}(x^TAx + y^TAy).$$

It is know that if A is a positive definite matrix of size n, then all diagonal elements of A are positive, and therefore $|a_{ii} - a_{jj}| < a_{ii} + a_{jj}$ for all $i, j = 1, \cdots n$. The next result is a generalization of this simple observation.

Corollary 10. *Assume that A is a positive definite matrix of size n, and let the eigenvalues of A, in increasing order, be $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$, then*

$$|x^TAx - y^TAy| \leq \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}(x^TAx + y^TAy).$$

Corollary 11. *Assume that $A = [a_{ij}]$ is positive definite matrix of size n, and let the eigenvalues of A, in increasing order, be $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$, then for every $i \neq j$*

1. $|a_{ij}| \leq \frac{1}{2}\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}(a_{ii}+a_{jj})$.
2. *If $k = \frac{\lambda_n}{\lambda_1}$ is the condition number of A, then*
$$k \geq \frac{a_{ii}+2a_{ij}+a_{jj}}{a_{ii}-2a_{ij}+a_{jj}}.$$
3. $|a_{ii}-a_{jj}| \leq \frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}(a_{ii}+a_{jj})$ *and hence*
$$k \geq \max_{\substack{i=1\\ i\neq j}}^{n} \frac{a_{ii}}{a_{jj}}.$$

4. Miscellaneous Results and Conjectures

In this section, we state some results that can be developed using the adopted approach of constrained optimization. The following results is a generalization of Proposition 2.

Proposition 12. *Let $A \in R^{n\times n}$ be a positive definite matrix and let x and $y, b \in R^{n\times 1}$ be three vectors such that $||x|| = ||y|| = ||b|| = 1$, $x^Ty = 0$, $x^Tb = 0$, and $y^Tb = 0$. Then*
$$|x^TAy| \leq \frac{\lambda_n(A_b)-\lambda_1(A_b)}{2},$$
where $A_b = (I-bb^T)A(I-bb^T)$.

Remark 3: It can be shown that $\lambda_n(A_b) \leq \lambda_n(A)$ and $\lambda_1(A_b) \geq \lambda_1(A)$, hence
$$|x^TAy| \leq \frac{\lambda_n(A_b)-\lambda_1(A_b)}{2} \leq \frac{\lambda_n(A)-\lambda_1(A)}{2}.$$

Proposition 13. *Assume that $A =$ is a positive definite matrix of size n, and let the eigenvalues of A, in increasing order, be $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$, then*
$$(x^TAx)(y^TAy) - (x^TAy)^2 \leq \lambda_n\lambda_{n-1}.$$

Proof: Let $x = \sum_{i=1}^n \alpha_i q_i$ and $y = \sum_{i=1}^n \beta_i q_i$, where $\{q_1, \cdots, q_n\}$ is the set of eigenvectors of A. The proof follows directly from the identity:
$$(x^TAx)(y^TAy) - (x^TAy)^2 = (\sum_{i=1}^n \lambda_i\alpha_i^2)(\sum_{j=1}^n \lambda_j\beta_j^2)$$
$$-(\sum_{k=1}^n \lambda_k\alpha_k\beta_k)^2 = \sum_{i=1}^n \lambda_i\lambda_j(\alpha_i\beta_j-\alpha_j\beta_i)^2.$$

Another generalization of Waielandt inequality is given next.

Proposition 14. *Let $A \in R^{n\times n}$ be a positive definite matrix and let x and $y \in R^{n\times 1}$ be two vectors such that $||x|| = ||y||$ and $x^Ty = 0$. Assume that the eigenvalues of A, in increasing order, are $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$, then for every two positive integers r and s*
$$|x^TA^sy| \leq \max_{\substack{i,j=1\\ i\neq j}} \frac{1}{2}|\frac{\lambda_i^s-\lambda_j^s}{\lambda_i^r+\lambda_j^r}|(x^TA^rx)(y^TA^ry).$$

Corollary 15. *Let $A \in \mathbb{R}^{n\times n}$ be a positive definite matrix of size n. Assume that the eigenvalues of A are $\lambda_1 < \lambda_2 \leq \cdots < \lambda_n$, then for any unit vector $x \in \mathbb{R}^n$:*
$$\frac{\{x^TA^2x-(x^TAx)^2\}^{\frac{3}{2}}}{x^TA^3x-(x^TAx)(x^TA^2Ax)} \leq \frac{1}{2}\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}.$$

The following two conjectures are generalizations of Theorem 5 and Proposition 9.

Conjecture 1. *Let $A \in R^{n\times n}$ $(n \geq 3)$ be a positive definite matrix and assume that the eigenvalues of A, in increasing order, are $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$. Then for any three orthonormal vectors $x, y, z \in R^{n\times 1}$*
$$\left(\frac{\lambda_1+\lambda_2+\lambda_3}{3}\right)^3 < (x^TAx)(y^TAy)(z^TAz)$$
$$< \left(\frac{\lambda_{n-2}+\lambda_{n-1}+\lambda_n}{3}\right)^3,$$
and that the strick inequalities always hold.

Conjecture 2. *Let $A \in R^{n\times n}$ be a positive definite matrix and assume that the eigenvalues of A, in increasing order, are $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$. Then for any three orthonormal vectors $x, y, z \in R^{n\times 1}$*
$$|x^TAy+x^TAz+y^TAz| \leq \frac{1}{2}\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}(x^TAx+y^TAy+z^TAz).$$

5. Miscellaneous Inequalities

In this section, we list a few inequalities that can be derived from the framework of the previous sections. Some of these results are known in the literature.

Corollary 15. *Let $A \in \mathbb{R}^{n\times n}$ be a positive definite matrix of size n. Assume that the eigenvalues of A are $\lambda_1 < \lambda_2 \leq \cdots < \lambda_n$, then for any two unit vectors $x, y \in \mathbb{R}^n$*

(a) $x^TAy \leq \cos(\theta)(\frac{x^TAx+y^TAy}{2}) + \frac{\lambda_n-\lambda_1}{2}\sin(\theta)$, *where θ is the angle between x and y.*

(b) $|\lambda_k - a_{ii}| \leq \lambda_n - \lambda_1,\ k, i = 1, \cdots, n$

(c) $|a_{ii}a_{jj}| \leq (\frac{\lambda_n+\lambda_{n-1}}{2})^2,\ i, j = 1, \cdots, n$

(d) $|\frac{Trace(A)}{n} - \lambda_k| \leq \lambda_n - \lambda_1,\ k = 1, \cdots, n$

(e) $\sqrt{\sum_{i\neq j} a_{ij}^2} \leq \frac{\lambda_n-\lambda_1}{2}$

(f) $|a_{ii} - a_{jj}| \leq \lambda_n - \lambda_1,\ i, j = 1, \cdots, n$

(g) $\frac{\sqrt{\sum_{i<j}(C_i-C_j)^2}}{n} \leq \frac{\lambda_n-\lambda_1}{2}$, *where C_i denotes the sum of the elements of the ith column of A.*

(h) $|a_{ij}| = e_j^TAe_i \leq \frac{\lambda_n-\lambda_1}{2},\ i \neq j$, *and hence*
$$\lambda_n - \lambda_1 \geq 2\max_{i\neq j}\{|a_{ij}|\}.$$

(i) *If the vectors $\frac{\alpha e_i+e_j}{\sqrt{1+\alpha^2}}$ and $\frac{e_i-\alpha e_j}{\sqrt{1+\alpha^2}}$ are considered, then*
$$\max_{\alpha,i\neq j}\{\frac{\alpha(a_{ii}-a_{jj})+(1-\alpha^2)a_{ij}}{1+\alpha^2}\} \leq \frac{\lambda_n-\lambda_1}{2}.$$

(j) *Assume that n is even and let $C_k = \sum_{i=1}^n a_{ik}$, then*
$$\frac{\sum_{i=1}^n(-1)^{k+1}C_k}{n} \leq \frac{\lambda_n-\lambda_1}{2}$$

(k) $\frac{1}{n^2}[n\ trace(A^2) - (trace(A))^2] \leq (\frac{\lambda_1-\lambda_n}{2})^2$

(l) $x^TA^3x - 2x^TAx(x^TA^2x) + (x^TAx)^3 \leq (\frac{\lambda_n+\lambda_{n-1}}{2})^2\frac{\sqrt{x^TA^2x-(x^TAx)^2}}{x^TAx}$

(m) $\frac{trace(A^3)-\frac{1}{n}trace(A^2)trace(A)}{\sqrt{n\ trace(A^4)-(trace(A^2))^2}} \leq \frac{\lambda_n-\lambda_1}{2}$, *and thus*

$$trace(A^3) - \frac{1}{n}trace(A^2)trace(A) \leq$$
$$n(\frac{\lambda_n^2-\lambda_1^2}{2})\frac{(\lambda_n-\lambda_1)}{2} \leq \frac{n}{4}(\lambda_n^2-\lambda_1^2)(\lambda_n-\lambda_1)$$
$$= \frac{n}{4}(\lambda_1+\lambda_n)(\lambda_n-\lambda_1)^2.$$

Proof: The proof of (a) from Proposition 2 and the observation that the vectors x and $z = \frac{y-x^Ty)x}{\sqrt{1-(x^Ty)^2}}$ are orthogonal for any unit vector x and that $||z|| = 1$. Specifically,

$$x^TAz = x^TA\frac{y}{||y||} = x^TA\frac{y-x^Ty)x}{\sqrt{1-(x^Ty)^2}} \leq \frac{\lambda_n-\lambda_1}{2}.$$

Parts (b), (d), and (f) follows from Proposition 2 by setting $(x,y) = (q_k, e_i), (e_i, e_j), (\frac{\sum_{i=1}^n q_i}{\sqrt{n}}, q_k)$, respectively. Part (c) is a direct result of Theorem 5 where $x = e_i, y = e_j$.
Proof of (j): Assume that n is even and set $x = \frac{1}{\sqrt{n}}\sum_{k=1}^n e_k$ and $y = \frac{1}{\sqrt{n}}\sum_{k=1}^n(-1)^k e_k$, then $x^Ty = 0$ and thus

$$\frac{\sum_{i=1}^n(-1)^{k+1}C_k}{n} \leq \frac{\lambda_n-\lambda_1}{2}$$

Proof of (k): Let $x = \frac{\sum_{i=1}^n q_i}{\sqrt{n}}$, then $x^TA^kx = \frac{trace(A^k)}{n}$ for each integer k. The conclusion is a direct result of Corollay 4, Part (a).

Remark 4: In Part (i), let

$$g(\alpha) = \frac{\alpha}{1+\alpha^2}(a_{ii}-a_{jj}) + \frac{(1-\alpha^2)}{1+\alpha^2}a_{ij},$$

then

$$g'(\alpha) = \frac{-(a_{ii}-a_{jj})\alpha^2 - 4a_{ij}\alpha + (a_{ii}-a_{jj})}{(1+\alpha^2)^2}.$$

If $a_{ii} \neq a_{jj}$, then $g'(\alpha) = 0$ if and only if

$$\alpha = \frac{-4a_{ij} \pm \sqrt{(a_{ii}-a_{jj})^2 + 16a_{ij}^2}}{2(a_{ii}-a_{jj})}.$$

One of these values can be shown to be a local and global maxima of the function $g(\alpha)$. Thus an improvement on the estimates of (f) and (h) can be obtained.

Remark 5: Assume that $x = \sum_{i=1}^n \alpha_i q_i$, then $\sum_{i=1}^n \alpha_i^2 = 1$, where the q_is are the set of eigenvectors of A. Then

$$(x^TA^2x)(x^Tx) = \sum \lambda_i^2\alpha_i^2 = (\sum_{i=1}^n \lambda_i^2\alpha_i^2)(\sum_{i=1}^n \alpha_i^2),$$

and

$$(x^TAx)^2 = (\sum_{i=1}^n \lambda_i\alpha_i^2)(\sum_{j=1}^n \lambda_j\alpha_j^2).$$

Therefore,

$$(x^TA^2x)(x^Tx) - (x^TAx)^2 = \sum_{i=1}^n\sum_{j=1}^n \lambda_i^2\alpha_i^2\alpha_j^2$$
$$-\sum_{i=1}^n\sum_{j=1}^n(\lambda_i\alpha_i^2)(\lambda_j\alpha_j^2) = \sum_{i=1}^n\sum_{j=1}^n(\lambda_i^2-\lambda_i\lambda_j)\alpha_i^2\alpha_j^2 \quad (9)$$
$$= \sum_{i=1}^n\sum_{j<i}^n(\lambda_j-\lambda_i)^2\alpha_i^2\alpha_j^2.$$

This identity can be utilized to prove that the maximum of $x^TA^2x - (x^TAx)^2$ is $\frac{(\lambda_n-\lambda_1)^2}{4}$ and $\alpha = \begin{bmatrix} \pm\frac{1}{\sqrt{2}} & 0\cdots & \pm\frac{1}{\sqrt{2}} \end{bmatrix}$. Note that the expression in (9) can be rewritten as

$$(x^TA^2x)(x^Tx) - (x^TAx)^2 = c^TBc$$

where $B = [b_{ij}]$, $b_{ij} = (\lambda_j-\lambda_i)^2$, and $c = \begin{bmatrix} \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \end{bmatrix}$. An interesting property of the matrix B is that it is of rank 3 at most regardless of its size.

6. Conclusion

Wielandt-type inequalities are derived using equality constraints optimization techniques. These inequalities are then utilized to develop bounds for functions of eigenvalues of positive semidefinite matrices. Some of these bounds are related to functions of extreme eigenvalues and others to the largest two or smallest two eigenvalues. In this work, although all matrices involved are real most of the results can be extended to positive definite hermitian matrices with minor modifications. The proposed methods are also applicable for deriving bounds for the singular values of matrices, however, these bounds are not reported here due to space limitation.

References

[1] J.K. Baksalary, S. Puntanen, Generalized matrix versions of the Cauchy-Schwarz and Kantorovich inequalities, Aequationes Math. 41 (1991) Pages:103-110.

[2] B. Mond, J.E. Pecaric, Matrix versions of some means inequalities, Aust. Math. Soc. Gaz. 20 (1993) Pages:117-120.

[3] Song-Gui Wang and Wai-Cheung Ip, A matrix version of the Wielandt inequality and its applications to statistics, Linear Algebra and its Applications, Volume 296, Issues 1-3, 15 July 1999, Pages 171-181.

[4] M. Fujii, Y. Katayama, R. Nakamoto, Generalizations of the Wielandt theorem, Math. Japan. 49 (1999) 217-222.

[5] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, MA, 1985.

[6] Jun Ichi Fujii, Ritsuo Nakamoto and Yuki Seo, Operator inequalities of Malamud and Wielandt, Linear Algebra and its Applications, Volume 326, Issues 1-3, 15 March 2001, Pages 101-109.

[7] Karl Gustafson, The geometrical meaning of the KantorovichWielandt inequalities, Linear Algebra and its Applications, Volume 296, Issues 1-3, 15 July 1999, Pages 143-151.

[8] Lina Yeh, A note on Wielandt's inequality, Applied Mathematics Letters, Volume 8, Issue 3, May 1995, Pages 29-31.

[9] A.S. Householder, The Theory of Matrices in Numerical Analysis, Reprint Edition, Dover, New York, pp. 81-85. Original Version: Blaisdell, New York, 1964.