Scientific
Research
Publishing

Two Nonzero Component Lemma and Matrix Trigonometry

Morteza Seddighin

Indiana University East, Richmond, USA
Email: mseddigh@iue.edu

Abstract

In this paper we show that the author's Two Nonzero Lemma (TNCL) can be applied to present a simple proof for a very useful equality which was first proved by Karl Gustafson in 1968. Gustafson used Hilbert space methods, including convexity of the Hilbert space norm, to prove this identity which was the basis of his matrix trigonometry. By applying TNCL, we will reduce the problem to a simple problem of ordinary calculus.

Keywords

Matrix Trigonometry

1. Introduction

Given a positive matrix T, in 1968 Gustafson proved

$$\inf_{\epsilon > 0} \left\| \epsilon T - I \right\| = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \tag{1}$$

where

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \geq \lambda_n \tag{2}$$

are eigenvalues of T such that λ_1 and λ_n are the largest and the smallest eigenvalues of T respectively. Please see [1] [2] [3].

The equality (1) played an important role in establishing what Gustafson calls "operator trigonometry". In fact, for a positive matrix T he defined $\sin T$ to be

$$\sin T = \inf_{\epsilon > 0} \left\| \epsilon T - I \right\| = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \tag{3}$$

He proved (1) by using the convexity of the Hilbert space norm and other Hilbert space properties.

Later, in his investigation on problems of antieigenvalue theory, this author discovered a useful lemma which he calls the Two Nonzero Component Lemma or TNCL, for short (see [4] [5] [6]). The antieigenvalue of an accretive operator T acting on a complex Hilbert space is defined to be

$$\mu = \mu_1(T) = \inf_{\substack{Tf \neq 0 \\ \|f\|=1}} \frac{\mathrm{Re}(Tf, f)}{\|Tf\|} \tag{4}$$

For positive matrices, there is a relationship between the antieigenvalue of T and $\mathrm{Sin}T$. In a series of papers this author applied his TNCL to compute antieigenvalues of different types of operators, including normal operators. He also applied TNCL to compute other types of antieigenvalue quantities such as total antieigenvalues, higher order antieigenvalues, and joint antieigenvalues. Furthermore, he applied TNCL to solve some optimization problems in statistics, econometrics, and resource allocations. Please see [5]-[11]. Although this Lemma is implicitly used in all of the author's earlier papers up to 2008, it was not until 2008 that he stated a formal description of the Lemma in his paper titled, "Antieigenvalue Techniques in Statistics." Below is the statement of the lemma. For an early proof of the lemma please see the author's work in [5].

Lemma 1 (*The Two Nonzero Component Lemma*) *Let l_1^+ be the set of all sequences with nonnegative terms in the Banach Space l_1. That is, let*

$$l_1^+ = \left\{ \boldsymbol{t} = (t_i) \epsilon l_1, t_i \geq 0 \right\} \tag{5}$$

Let

$$F(x_1, x_2, \cdots, x_m) \tag{6}$$

be a function from R^m to R. Assume $g_k(\boldsymbol{t}) = \sum c_i^k t_i$ for $(c_i^k) \epsilon l_1^+$, $\boldsymbol{t} \epsilon l_1^+$, and $1 \leq k \leq m$. Then the minimizing vectors for the function

$$F(g_1(\boldsymbol{t}), g_2(\boldsymbol{t}), \cdots, g_m(\boldsymbol{t})) \tag{7}$$

on the convex set $C = \left\{ (t_i) \epsilon l_1 : \sum t_i = 1 \right\}$ have at most two nonzero components.

What make the proof of the Lemma possible are the following two facts: First, the convexity of the set

$$C = \left\{ (t_i) \epsilon l_1^+ : \sum t_i = 1 \right\} \tag{8}$$

Second, a special property that the functions

$$F(g_1(\boldsymbol{t}), g_2(\boldsymbol{t}), \cdots, g_m(\boldsymbol{t}).) \tag{9}$$

involved possess. If we set

$$D(t_1, t_2, t_3, \cdots) = F(g_1(\boldsymbol{t}), g_2(\boldsymbol{t}), \cdots, g_m(\boldsymbol{t})), \tag{10}$$

then all restrictions of the form

$$D(t_1, t_2, \cdots t_{i-1}, 0, t_{i+1}, \cdots) \tag{11}$$

of

$$D(t_1, t_2, t_3, \cdots) \tag{12}$$

have the same algebraic form as $D(t_1, t_2, t_3, \cdots)$ itself. For example if

$$D(t_1, t_2, \cdots, t_n) = \frac{\beta_1 t_1 + \beta_2 t_2 + \cdots + \beta_n t_n}{\left|\lambda_1\right|^2 t_1 + \left|\lambda_2\right|^2 t_2 + \cdots + \left|\lambda_3\right|^2 t_3} \tag{13}$$

then we have

$$D(0, t_2, \cdots, t_n) = \frac{\beta_2 t_2 + \cdots + \beta_n t_n}{\left|\lambda_2\right|^2 t_2 + \cdots + \left|\lambda_3\right|^2 t_3} \tag{14}$$

which has the same algebraic form as

$$D(t_1, t_2, \cdots, t_n) = \frac{\beta_1 t_1 + \beta_2 t_2 + \cdots + \beta_n t_n}{\left|\lambda_1\right|^2 t_1 + \left|\lambda_2\right|^2 t_2 + \cdots + \left|\lambda_3\right|^2 t_3} \tag{15}$$

Indeed, for any j, $1 \le j < n$; all restrictions of the function

$$D(t_1, t_2, \cdots, t_n) = \frac{\beta_1 t_1 + \beta_2 t_2 + \cdots + \beta_n t_n}{\left|\lambda_1\right|^2 t_1 + \left|\lambda_2\right|^2 t_2 + \cdots + \left|\lambda_3\right|^2 t_3} \tag{16}$$

obtained by setting an arbitrary set of j components of $D(t_1, t_2, \cdots, t_n)$ equal to zeros have the same algebraic form as $D(t_1, t_2, \cdots, t_n)$. Obviously, not all functions have this property. For instance, for the function $G(t_1, t_2) = 2t_1 + t_1 t_2$, $G(t_1, 0) = t_1$, which does not have the same algebraic form as $G(t_1, t_2)$.

In the next section we prove that Gustafson's identity (1) can be obtained using this author's the Two Nonzero Component Lemma or TNCL. Our proof is elementary (comparing to Gustafson's proof) in the sense that we use only TNCL and techniques of calculus.

2. A Proof of (1) Based on TNCL

Theorem 2 *Let* T *be a positive matrix where*

$$\lambda_1 \ge \lambda_2 \ge \lambda_3 \cdots \ge \lambda_n$$

are eigenvalues of T such that λ_1 and λ_n are the largest and the smallest eigenvalues of T, then

$$\inf_{\in > 0} \left\| \in T - I \right\| = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \tag{17}$$

Proof. Note that if we square the left hand side of (17) we get

$$\inf_{\in > 0} \sup_{\|x\| = 1} \left\| (\in T - I) x \right\|^2 \tag{18}$$

Thus, we need to show

$$\inf_{\in > 0} \sup_{\|x\| = 1} \left\| (\in T - I) x \right\|^2 = \frac{(\lambda_1 - \lambda_n)^2}{(\lambda_1 + \lambda_n)^2} \tag{19}$$

Now to follow notations usually used in differential calculus, let's substitute $\in$ with y and consider

$$\inf_{\in > 0} \sup_{\|x\| = 1} \left\| (yT - I) x \right\|^2 \tag{20}$$

instead. With this change of notation. now we apply spectral theorem to the positive matrix T and assume

$$x = (x_1, x_2, \cdots, x_n)$$

are components of x with respect to an orthogonal basis corresponding to

$$\lambda_1, \lambda_2, \lambda_3 \cdots \lambda_n$$

Therefore, we can rewrite (20) as

$$\inf_{y>0} \sup_{\|x\|=1} \sum_{i=1}^{n} (y\lambda_i - 1)^2 \|x_i\|^2 \tag{21}$$

Applying TNCL we can assume any optimizing vector

$$x = (x_1, x_2, \cdots, x_n)$$

is so that only two of its components, say x_1 and x_j are nonzero and the rest of them are zero. Keeping that in mind, for such optimizing vectors x (21) will be reduced to

$$\inf_{y>0} \sup_{\|xi\|^2 + \|x_j\|^2 = 1} (y\lambda_i - 1)^2 \|x_i\|^2 + (y\lambda_j - 1)^2 \|xj\|^2 \tag{22}$$

To compute (22), let's do some change of variables first. Substitute a for λ_i, b for λ_j, and x for $\|xi\|^2$. (22) then becomes

$$\inf_{y>0} \sup_{0 \le x \le 1} \left[(ya - 1)^2 x + (yb - 1)^2 (1 - x) \right] \tag{23}$$

For a fixed x we compute the inf in (23) with respect to y first. Consider the expression

$$(ya - 1)^2 x + (yb - 1)^2 (1 - x) \tag{24}$$

We next find the derivative of (24) with respect to y and set it equal to zero

$$2b^2 y - 2b - 2ax + 2bx + 2a^2 xy - 2b^2 xy = 0 \tag{25}$$

and then solve it for y. The solution is

$$y = \frac{a(b + ax - bx)}{a^2 x - b^2 x + b^2} \tag{26}$$

Assume $a > b$ and note that the second derivative of (24) is

$$b^2 + a^2 x - b^2 x \tag{27}$$

which is positive. This shows

$$y = \frac{a(b + ax - bx)}{a^2 x - b^2 x + b^2} \tag{28}$$

is indeed a minimizing value. If we substitute y from (26) in 24) and simplify we get

$$-x(a - b)^2 \frac{x - 1}{a^2 x - b^2 x + b^2} \tag{29}$$

The derivative of (29) with respect to x is

$$-\frac{(a - b)^2}{(a^2 x - b^2 x + b^2)^2} (a^2 x^2 - b^2 x^2 + 2b^2 x - b^2) \tag{30}$$

To find the optimizing value x, we solve the following equation with respect

to x.

$$a^2x^2 - b^2x^2 + 2b^2x - b^2 = 0 \qquad (31)$$

The solution of (31) is

$$x = \frac{b}{a+b} \qquad (32)$$

If we substitute the value of x from (32) in (29) and simplify we get

$$\frac{(a-b)^2}{(a+b)^2} \qquad (33)$$

The second derivative of (29) is

$$-2a^2b^2 \frac{(a-b)^2}{(a^2x - b^2x + b^2)^3} \qquad (34)$$

which is negative, under our assumption that $a > b$. This indicates that x given by (32) is indeed a maximizing vector. Thus we have proved

$$\inf_{\in > 0} \|\in T - I\| = \frac{\lambda_i - \lambda_j}{\lambda_i + \lambda_j} \qquad (35)$$

Finally, we show that $i = 1$ and $j = n$. To show this note that

$$\frac{\lambda_i - \lambda_j}{\lambda_i + \lambda_j} = \frac{1 - \dfrac{\lambda_j}{\lambda_i}}{1 - \dfrac{\lambda_j}{\lambda_i}} \qquad (36)$$

Now define

$$f(t) = \left(\frac{1-t}{1+t}\right) \qquad (37)$$

and notice that

$$f'(t) = \frac{-2}{(t+1)^2} \qquad (38)$$

Hence $f(t)$ is decreasing and $\dfrac{\lambda_i - \lambda_j}{\lambda_i + \lambda_j} = \dfrac{1 - \dfrac{\lambda_j}{\lambda_i}}{1 - \dfrac{\lambda_j}{\lambda_i}}$ has the largest value when

$\dfrac{\lambda_j}{\lambda_i}$ takes the smallest value. That is when $j = n$ and $i = 1$.

Remark 3 The equality (35) is valid even if T is an infinite dimensional positive operator acting on a separable Hilbert space. The reason is that TNCL is valid both when $t = (t_i)$ has a finite or infinite number of components. However, in the case of an infinite dimensional positive operator, we do not know for what pair of i and j (35) holds.

3. Conclusion

We showed that TNCL can be used to prove an identity which was proved by

Karl Gustafson in 1968. This identity was part of his min-max theorem. The identity was the basis of operator trigonometry. The original proof was based on Hilbert space techniques and convexity of operator norm. Using TNCM we reduced the problem to a very simple problem in elementary calculus. This indeed shows the power of this dimension reducing optimization lemma which is used by this author in many of his previous work. The lemma not only proved equality (1) but, as we noted in the remark above, it extended it to the case of positive operators on an infinite dimensional Hilbert space.

Acknowledgement

References

[1] Gustafson, K. (2010) On My Min-Max Theorem (1968) and Its Consequences. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, **14**, 45-51.

[2] Gustafson, K. (1968) A Min-Max Theorem. *Notices of the American Mathematical Society*, **15**, 799.

[3] Gustafson, K. and Rao, D. (1997) Numerical Range. Springer.
https://doi.org/10.1007/978-1-4613-8498-4

[4] Gustafson, K. and Seddighin, M. (1989) Antieigenvalue Bounds. *Journal of Mathematical Analysis and Applications*, **143**, 327-340.
https://doi.org/10.1016/0022-247X(89)90044-9

[5] Seddighin, M. (2002) Antieigenvalues and Total Antieigenvalues of Normal Operators. *Journal of Mathematical Analysis and Applications*, **274**, 239-254.
https://doi.org/10.1016/S0022-247X(02)00295-0

[6] Seddighin, M. (2009) Antieigenvalue Techniques in Statistics. *Linear Algebra and Its Applications*, **430**, 2566-2580. https://doi.org/10.1016/j.laa.2008.05.007

[7] Seddighin, M. and Gustafsib, K. (2005) On the Eigenvalues which Express Antieigenvalues. *International Journal of Mathematics and Mathematical Sciences*, **2005**, 1543-1554.

[8] Seddighin, M. (2010) Gustafson K. Slant Antieigenvalues and Slant Antieigenvectors of Operators. *Journal of Linear Algebra and Applications*, **432**, 1348-1362.
https://doi.org/10.1016/j.laa.2009.11.001

[9] Seddighin, M. (2011) Slant Joint Antieigenvalues and Antieigenvectors of Operators in Normal Subalgebras. *Journal of Linear Algebra and its Applications*, **434**, 1395-1408. https://doi.org/10.1016/j.laa.2010.11.020

[10] Seddighin, M. (2014) Proving and Extending Greub-Reinboldt Inequality Using the Two Nonzero Component Lemma. *Advances in Linear Algebra & Matrix Theory*, **4**, 120-127. https://doi.org/10.4236/alamt.2014.42010

[11] Seddighin, M. (2014) Application of the Two Nonzero Component Lemma in Resource Allocation. *Journal of Applied Mathematics and Physics*, **2**, 653-661.