IJMMS 2003:63, 4015–4023
PII. S0161171203208334
http://ijmms.hindawi.com

OPTIMALLY ROTATED VECTORS

MORTEZA SEDDIGHIN

Received 5 August 2002

We study vectors which undergo maximum or minimum rotation by a matrix on the field of real numbers. The cosine of the angle between a maximally rotated vector and its image under the matrix is called the cosine or antieigenvalue of the matrix and has important applications in numerical methods. Using Lagrange multiplier technique, we obtain systems of nonlinear equations which represent these optimization problems. Furthermore, we solve these systems symbolically and numerically.

2000 Mathematics Subject Classification: 15A99, 47C99.

1. Introduction. The concept of cosine of an operator or a matrix was first introduced by Gustafson (see [2]). Given an operator on a Hilbert space, cosine of T is defined by

$$\cos T = \inf_{Tf \neq 0} \frac{\operatorname{Re}(Tf, f)}{\|Tf\| \|f\|}, \tag{1.1}$$

where $\cos T$ is also denoted by $\mu(T)$. This parameter has important applications in numerical analysis as well as pure matrix and operator theory. See [2] for more information on the applications of $\mu(T)$ to numerical techniques such as conjugate gradient and steepest descent methods. In recent years, many attempts have been made to compute or approximate $\cos T$ for operators on complex Hilbert spaces. In particular, computation and approximation of $\cos T$ for normal operators have been somewhat successful (see [2, 3, 4, 5]). A vector f for which the inf in (1.1) is attained is called a maximally rotated vector for T (a maximally rotated vector is called an antieigenvector in our earlier papers). On the other hand, vectors for which the sup in

$$\upsilon(T) = \sup_{Tf \neq 0} \frac{\operatorname{Re}(Tf, f)}{\|Tf\| \|f\|} \tag{1.2}$$

is attained are called minimally rotated vectors for T. A maximally or minimally rotated vector is called an optimally rotated vector. In the past the focus has been on the computation of maximally rotated vectors and $\mu(T)$. In the present paper we are also concerned with minimally rotated vectors and $\upsilon(T)$. Note that for a matrix on the real field, if the set of all negative eigenvalues of T is nonempty, then $\mu(T) = -1$. In this case the set of all maximally rotated vectors of T is simply the union of all eigenspaces corresponding to negative eigenvalues of T. However, if the set of negative eigenvalues of T is empty, then we have $\mu(T) \geq -1$. Likewise if the set of all positive eigenvalues of T is nonempty, then $\upsilon(T) = 1$. In this case the set of all minimally rotated vectors of T is simply the union of all eigenspaces corresponding to positive eigenvalues of T. However, if the set of positive eigenvalues of T is empty, then we have $\upsilon(T) \leq 1$. Unfortunately, a variational approach analogous to Rayleigh-Ritz variational theory of eigenvectors is not successful for computing optimally rotated vectors. Nevertheless, Gustafson has found an Euler equation that satisfies these vectors (see [2]). His approach is based on the direct computation of the left-hand side of

$$\lim_{h\to 0} \frac{\mathrm{Re}(T(f+hg), f+hg)/\|T(f+hg)\|\,\|f+hg\| - \mathrm{Re}(Tf,f)/\|Tf\|\,\|f\|}{h} = 0, \tag{1.3}$$

which yields the following equation:

$$2\|Tf\|^2\|f\|^2(\mathrm{Re}T)f - \|f\|^2\mathrm{Re}(Tf)T^*Tf - \|Tf\|^2\mathrm{Re}(Tf,f)f = 0. \tag{1.4}$$

In this paper we use Lagrange multipliers to compute the set of optimally rotated vectors for matrices on the real field. For a matrix T defined on the real field, we have

$$\mu(T) = \inf_{Tf\neq 0} \frac{(Tf,f)}{\|Tf\|\,\|f\|}, \qquad \nu(T) = \sup_{Tf\neq 0} \frac{(Tf,f)}{\|Tf\|\,\|f\|}. \tag{1.5}$$

Note that $\mu(T)$ and $\nu(T)$ can equivalently be defined by

$$\mu(T) = \inf_{\substack{Tf\neq 0\\ \|f\|=1}} \frac{(Tf,f)}{\|Tf\|}, \qquad \upsilon(T) = \sup_{\substack{Tf\neq 0\\ \|f\|=1}} \frac{(Tf,f)}{\|Tf\|}. \tag{1.6}$$

The following are three simple properties that result directly from definitions.

Property 1. For any real matrix T, we have $\mu(T) = \mu(T^t)$ and $\nu(T) = \nu(T^t)$, where T^t is the transpose of T.

Property 2. For any invertible matrix T, we have $\mu(T) = \mu(T^{-1})$ and $\nu(T) = \nu(T^{-1})$.

Property 3. If H is a reducing subspace of T, then $\mu(T) \leq \mu(T|H)$ and $\upsilon(T) \geq \upsilon(T|H)$, where $T|H$ is the restriction of T on H.

2. Main results. If $T = [t_{ij}]$, $1 \leq i \leq n$, $1 \leq j \leq n$, is an $n \times n$ matrix on the real field and $f = (x_1, x_2, x_3, \ldots, x_n)$ is any vector in $\mathbb{R}^n$, direct computations show that the functional $(Tf,f)/\|Tf\|$ takes the form

$$J(f)=\frac{\sum_{i=1}^{n}\left(\sum_{j=1}^{n}t_{ij}x_j\right)x_i}{\sqrt{\sum_{i=1}^{n}\left(\sum_{j=1}^{n}t_{ij}x_j\right)^2}}. \tag{2.1}$$

Therefore, in order to compute $\mu(T)$ and $\upsilon(T)$ we must find the optimum values of the expression

$$\frac{\sum_{i=1}^{n}\left(\sum_{j=1}^{n}t_{ij}x_j\right)x_i}{\sqrt{\sum_{i=1}^{n}\left(\sum_{j=1}^{n}t_{ij}x_j\right)^2}} \tag{2.2}$$

on the unit sphere $\sum_{i=1}^{n}x_i^2=1$. Making use of Lagrange multiplier technique seems like a natural approach in computing optimally rotated unit vectors. However, as the next theorem illustrates, for a general matrix, the resulting equations are nonlinear and hard to solve even in the case of 2×2 matrices.

Theorem 2.1. *Let*

$$T=\begin{bmatrix} a & b\\ c & d\end{bmatrix} \tag{2.3}$$

be a 2×2 matrix on the field of real numbers; then any optimally rotated unit vector $f=(x,y)$ satisfies the following system of equations:

$$\begin{aligned}&\left[(ax+by)^2+(cx+dy)^2\right]\left[cy^2+by^2+2axy-cx^2-bx^2-2dxy\right]\\ &\quad=\left[ax^2+bxy+cxy+dy^2\right]\left[cdy^2+c^2xy+aby^2+a^2xy\right.\\ &\qquad\qquad\left.-d^2xy-cdx^2-b^2xy-abx^2\right],\\ &x^2+y^2=1.\end{aligned} \tag{2.4}$$

Proof. Finding

$$\mu(T)=\inf_{\substack{Tf\neq 0\\ \|f\|=1}}\frac{(Tf,f)}{\|Tf\|\|f\|} \tag{2.5}$$

or

$$\upsilon(T)=\sup_{\substack{Tf\neq 0\\ \|f\|=1}}\frac{(Tf,f)}{\|Tf\|} \tag{2.6}$$

is the same as finding the optimum values of the function

$$J(x,y)=\frac{ax^2+(b+c)xy+dy^2}{\sqrt{(ax+by)^2+(cx+dy)^2}} \tag{2.7}$$

on the sphere $x^2+y^2=1$. A necessary condition for $f=(x,y)$ to be an optimizing vector for $J(x,y)$ on the sphere is that the gradients of $J(x,y)$ and

$x^2+y^2=1$ be parallel. This means that we must have

$$\begin{aligned}\frac{\partial J}{\partial x}&=\frac{2ax+(b+c)y}{\sqrt{(ax+by)^2+(cx+dy)^2}}-\frac{ax^2+(b+c)xy+dy^2}{\left(\sqrt{(ax+by)^2+(cx+dy)^2}\right)^3}\\&\quad\times(a^2x+aby+c^2x+cdy)=2\lambda x,\\\frac{\partial J}{\partial y}&=\frac{(b+c)x+2dy}{\sqrt{(ax+by)^2+(cx+dy)^2}}-\frac{ax^2+(b+c)xy+dy^2}{\left(\sqrt{(ax+by)^2+(cx+dy)^2}\right)^3}\\&\quad\times(abx+b^2y+cxd+d^2y)=2\lambda y,\\&\qquad x^2+y^2=1,\end{aligned}\tag{2.8}$$

for some nonzero constant λ. Eliminating λ from (2.8) yields system (2.4). $\square$

As we will see later, system (2.4) can be solved algebraically for some special matrices. Nevertheless, we can solve that system numerically in all cases, as the following example shows.

Example 2.2. Find $\mu(T)$ and $\upsilon(T)$ for the matrix $T=\left[\begin{smallmatrix}2&-5\\1&3\end{smallmatrix}\right]$.

For this matrix, the function $J(x,y)$ defined by (2.7) is

$$J(x,y)=\frac{2x^2-4xy+3y^2}{\sqrt{5x^2-14xy+34y^2}}.\tag{2.9}$$

Now if we solve the system

$$\frac{\partial J(x,y)}{\partial x}=2\lambda x,\quad \frac{\partial J(x,y)}{\partial y}=2\lambda y,\quad x^2+y^2=1,\tag{2.10}$$

numerically, we obtain the following sets of solutions: $(x=0.7437,\ y=0.66851)$, $(x=-0.7437,\ y=0.66851)$, $(x=-0.96038,\ y=0.27871)$, and $(-0.96038,0.27871)$. One can verify that

$$\mu(T)=J(0.7437,0.66851)=J(-0.7437,-0.66851)=0.13816,\tag{2.11}$$

and hence vectors $(0.7437,0.66851)$ and $(-0.7437,0.66851)$ are maximally rotated unit vectors. Similarly one can verify that

$$\upsilon(T)=J(-0.96038,0.27871)=J(96038,0.27871)=0.94926\tag{2.12}$$

with $(-0.96038,0.27871)$ and $(0.96038,0.27871)$ being minimally rotated vectors. We can obtain the same results graphically. If we substitute $y=\sqrt{1-x^2}$ in (2.9), we obtain

$$f(x)=J\left(x,\sqrt{1-x^2}\right)=\frac{-x^2-4x\sqrt{1-x^2}+3}{\sqrt{-29x^2-14x\sqrt{1-x^2}+34}}\tag{2.13}$$

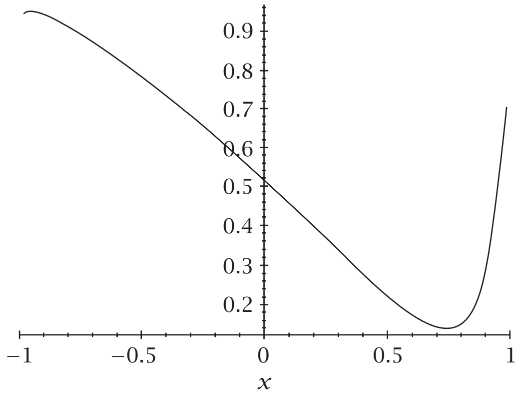

FIGURE 2.1

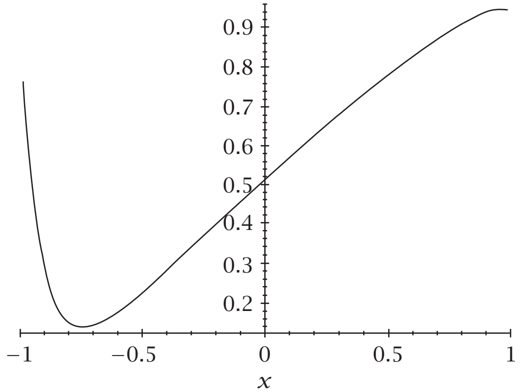

FIGURE 2.2

whose graph is shown in Figure 2.1. The y-axis represents the cosine of the angle between a vector $(x, f(x))$ and its image under T.

On the other hand, substituting $-\sqrt{1-x^2}$ in (2.9) gives us

$$g(x) = J\left(x, -\sqrt{1-x^2}\right) = \frac{-x^2+4x\sqrt{1-x^2}+3}{\sqrt{-29x^2+14x\sqrt{1-x^2}+34}} \tag{2.14}$$

whose graph is shown in Figure 2.2. The y-axis represents the cosine of the angle between a vector $(x, f(x))$ and its image under T.

Notice that the graphs of f and g are symmetric with respect to the vertical axis.

System (2.4) can be converted to a polynomial system. We omit the proof of the next corollary.

Corollary 2.3. *Let $T = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ be a 2×2 matrix on the field of real numbers; then an optimally rotated vector $f = (x,y)$ with $\|f\| = 1$ satisfies the following system of equations:*

$$\begin{gathered} px^4 + qx^3y + rx^2y^2 + sxy^3 + ty^4 = 0, \\ x^2 + y^2 = 1, \end{gathered} \tag{2.15}$$

where the coefficients p, q, r, s, and t are functions of a, b, c, and d as follows:

$$\begin{aligned} p &= a(cd+ab) - (b+c)(a^2+c^2), \\ q &= a^3 + ac^2 - 2a^2d - 3c^2d + ad^2 - bcd - abc, \\ r &= 3(a-d)(ab+cd), \\ s &= -d^3 + 3ab^2 + 2ad^2 - b^2d - a^2d + bcd + abc, \\ t &= (b+c)(b^2+d^2) - d(cd+ab). \end{aligned} \tag{2.16}$$

System (2.15) can be easily solved for some special cases. For example, if a matrix is such that $q = 0$ and $s = 0$, then the optimally rotated unit vectors $f = (x,y)$ satisfy

$$x^2 = \frac{(2p-r) \pm \sqrt{r^2 - 4pt}}{2(p-r+t)}, \qquad y^2 = \frac{(2t-r) \pm \sqrt{r^2 - 4pt}}{2(p-r+t)}. \tag{2.17}$$

Corollary 2.4. *For the matrix $T = \left[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right]$,*

$$\mu(T) = \nu(T) = \frac{a}{2\sqrt{a^2+b^2}}. \tag{2.18}$$

Proof. In this case all of the coefficients in the first equation of system (2.15) are zero. Also direct substitutions show that for any vector $f = (x,y)$, the value of $J(x,y)$ defined by (2.7) is $a/2\sqrt{a^2+b^2}$. $\square$

Corollary 2.5. *For the matrix $T = \left[\begin{smallmatrix} a & b \\ b & -a \end{smallmatrix}\right]$, the unit vectors $f = (x,y)$ whose components satisfy $x^2 = 1/2 + a/\sqrt{a^2+b^2}$ and $y^2 = 1/2 - a/\sqrt{a^2+b^2}$ are minimally rotated vectors giving $\nu(T) = 1$. Also the unit vectors whose components satisfy $x^2 = 1/2 - a/\sqrt{a^2+b^2}$ and $y^2 = 1/2 + a/\sqrt{a^2+b^2}$ are maximally rotated vectors giving $\mu(T) = (b^2-a^2)/(b^2+a^2)$.*

Proof. If we substitute entries of this matrix in system (2.15) and eliminate y in it, we obtain the equation

$$4(a^2+b^2)x^4 - 4(a^2+b^2)x^2 + b^2 = 0. \tag{2.19}$$

This will give us two sets of the solutions. One set satisfies $x^2 = 1/2 + a/\sqrt{a^2+b^2}$ and $y^2 = 1/2 - a/\sqrt{a^2+b^2}$. The other set satisfies $x^2 = 1/2 - a/\sqrt{a^2+b^2}$ and $y^2 = 1/2 + a/\sqrt{a^2+b^2}$. If we substitute any vector from the first set in (2.7), we obtain 1. If we substitute any vector from the second set in (2.7), we obtain $(b^2-a^2)/(b^2+a^2)$. $\square$

In many applications one needs only to find upper bounds for $\mu(T)$ and lower bounds for $\upsilon(T)$. If K is a reducing subspace for T, then $\mu(T|K)$ is an upper bound for $\mu(T)$ and $\nu(T|K)$ is a lower bound for $\nu(T)$. In particular, if J is any elementary Jordan matrix in the elementary Jordan form of T, then $\mu(T) \le \mu(J)$. It is in general easier to compute μ and υ for an elementary Jordan matrix than for a general matrix.

Theorem 2.6. *If $J = \left[\begin{smallmatrix} k & 0 \\ 1 & k \end{smallmatrix}\right]$ is a 2×2 elementary Jordan matrix, then*

(1) *if $k > 0$, the unit vector $(0,1)$ is a minimally rotated vector and $\upsilon(J) = 1$. The two vectors*

$$\left(\frac{2k}{\sqrt{4k^2+1}}, \frac{-1}{\sqrt{4k^2+1}}\right), \qquad \left(\frac{-2k}{\sqrt{4k^2+1}}, \frac{1}{\sqrt{4k^2+1}}\right) \tag{2.20}$$

are maximally rotated unit vectors and $\mu(J) = (4k^2-1)/(4k^2+1)$;

(2) *if $k < 0$, the unit vector $(0,1)$ is a maximally rotated vector and $\mu(J) = -1$. The two vectors*

$$\left(\frac{2k}{\sqrt{4k^2+1}}, \frac{-1}{\sqrt{4k^2+1}}\right), \qquad \left(\frac{-2k}{\sqrt{4k^2+1}}, \frac{1}{\sqrt{4k^2+1}}\right) \tag{2.21}$$

are minimally rotated unit vectors and $\upsilon(J) = (4k^2-1)/(4k^2+1)$.

Proof. First note that k is an eigenvalue for J. One can also verify that in this case system (2.15) is reduced to

$$x + 2ky = 0, \qquad x^2 + y^2 = 1, \tag{2.22}$$

which has two sets of solutions

$$\left(x = \frac{2k}{\sqrt{4k^2+1}}, y = \frac{-1}{\sqrt{4k^2+1}}\right), \qquad \left(x = \frac{-2k}{\sqrt{4k^2+1}}, y = \frac{1}{\sqrt{4k^2+1}}\right). \tag{2.23}$$

Substituting the values of x and y in either set of solutions in (2.7) yields $(4k^2-1)/(4k^2+1)$. □

Although it becomes impossible to algebraically solve the system of equations resulting from the Lagrange multipliers for general matrices of dimension greater than two, numerical solutions of these systems are always accessible.

Example 2.7. Find $\mu(T)$ and $\upsilon(T)$ for the matrix

$$T = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 1 \\ 2 & 2 & 3 \end{bmatrix}. \tag{2.24}$$

Substituting the entries of this matrix in functional (2.1), we obtain

$$g(x,y,z) = \frac{x^2+5xy+xz+y^2+3yz+3z^2}{\sqrt{14x^2+18xy+16xz+9y^2+10yz+11z^2}}, \tag{2.25}$$

where $f=(x,y,z)$ is a vector. To optimize (2.25) on the unit sphere $x^2+y^2+z^2=1$, we need to solve the system

$$\begin{aligned} \frac{\partial g(x,y,z)}{\partial x} &= 2\lambda x, & \frac{\partial g(x,y,z)}{\partial y} &= 2\lambda y, \\ \frac{\partial g(x,y,z)}{\partial z} &= 2\lambda z, & x^2+y^2+z^2 &= 1. \end{aligned} \tag{2.26}$$

Solving this system numerically, we find two minimally rotated vectors

$$(-0.12416, 0.28283, 0.9511), \qquad (0.12416, -0.28283, -0.9511), \tag{2.27}$$

yielding

$$\begin{aligned} \upsilon(T) &= g(-0.12416, 0.28283, 0.9511) \\ &= g(0.12416, -0.28283, -0.9511) \\ &= 0.99929. \end{aligned} \tag{2.28}$$

We also find two maximally rotated vectors

$$\begin{aligned} &(0.63262, -0.77202, 6.1442\times 10^{-2}), \\ &(-0.63262, 0.77202, -6.1442\times 10^{-2}), \end{aligned} \tag{2.29}$$

yielding

$$\begin{aligned} \mu(T) &= g(0.63262, -0.77202, 6.1442\times 10^{-2}) \\ &= g(-0.63262, 0.77202, -6.1442\times 10^{-2}) \\ &= -1. \end{aligned} \tag{2.30}$$

Notice that this matrix has only one real eigenvalue which is negative, and the last two vectors are actually eigenvectors of T corresponding to this negative eigenvalue.

This example underlines the very important fact that our techniques can be used to find eigenvectors that correspond to real eigenvalues and hence the real eigenvalues themselves. To find the negative eigenvalue of the matrix T above, note that

$$\begin{aligned} Tf &= (-0.97286, 1.1873, -9.4474\times 10^{-2}) \\ &= -1.5379(-0.63262, 0.77202, -6.1442\times 10^{-2}). \end{aligned} \tag{2.31}$$

Therefore, the negative eigenvalue of T is -1.5379.

Remark 2.8. In [1] Gustafson has developed an extended operator trigonometry by redefining $\mu(T)$ for invertible operators based on their polar decomposition $T = U|T|$. He has replaced the definition of $\mu(T)$ given by expression (1.1) with

$$\mu(T) = \inf_{Tf \neq 0} \frac{(|T|f, f)}{\||T|f\|\|f\|}. \tag{2.32}$$

Acknowledgment. The author wishes to thank Professor Karl Gustafson for introducing this topic to him.

References

[1] K. Gustafson, *An extended operator trigonometry*, Linear Algebra Appl. **319** (2000), no. 1–3, 117–135.
[2] K. Gustafson and D. K. M. Rao, *Numerical Range*, Universitext, Springer-Verlag, New York, 1997.
[3] K. Gustafson and M. Seddighin, *Antieigenvalue bounds*, J. Math. Anal. Appl. **143** (1989), no. 2, 327–340.
[4] ______, *A note on total antieigenvectors*, J. Math. Anal. Appl. **178** (1993), no. 2, 603–611.
[5] M. Seddighin, *Antieigenvalues and total antieigenvalues of normal operators*, J. Math. Anal. Appl. **274** (2002), no. 1, 239–254.

Morteza Seddighin: Division of Natural Science and Mathematics, Indiana University East, Richmond, IN 47374, USA
E-mail address: mseddigh@indiana.edu