Maximization of a Second-Degree Polynomial on the Unit Sphere

By James W. Burrows*

I. Introduction. Let A be a hermitian matrix of order n, and a be a known vector in C^n. The problem is to determine which vectors make $\Phi(x) = x^*Ax - 2 \operatorname{Re} \{x^*a\}$ (* denotes conjugate transpose) a maximum or minimum on the unit sphere $S = \{x: x^*x = 1\}$.

[1] considers finding the similarly constrained maximum or minimum of $(x - b)^*A(x - b)$ where b is a known vector. We have

$$\begin{aligned}(x - b)^*A(x - b) &= x^*Ax - b^*Ax - x^*Ab + b^*Ab \\ &= x^*Ax - 2 \operatorname{Re} \{x^*Ab\} + b^*Ab\end{aligned}$$

so with $a = Ab$, the problems are seen to be equivalent unless A is singular, in which case our formulation is more general. This formulation also seems to lead to simpler proofs.

II. Computation of Extremal Vectors. Let U be the unitary transformation which diagonalizes A, i.e., if $x = Uy$, then

$$x^*Ax - 2 \operatorname{Re} \{x^*a\} = y^*U^*AUy - 2 \operatorname{Re} \{y^*U^*a\} = y^*\Lambda y - 2 \operatorname{Re} \{y^*c\}, \tag{2.1}$$

where $c = U^*a$ and $\Lambda = \operatorname{diag} \{\lambda_1, \cdots, \lambda_n\}$, with real λ_i. It is thus equivalent to find the maximum or minimum of

$$\psi(y) = \sum_{i=1}^{n} \lambda_i |y_i|^2 - 2 \operatorname{Re} \left\{ \sum_{i=1}^{n} c_i \bar{y}_i \right\} \tag{2.2}$$

with the constraint

$$\sum_{i=1}^{n} | y_i |^2 = 1. \tag{2.3}$$

Construct

$$\chi(y) = \sum_{i=1}^{n} \lambda_i |y_i|^2 - 2 \operatorname{Re} \left\{ \sum_{i=1}^{n} c_i \bar{y}_i \right\} - \lambda \sum_{i=1}^{n} | y_i |^2 \tag{2.4}$$

where stationarity with respect to complex y requires that the Lagrange multiplier λ be real (cf. [1], p. 30). An extremal vector then satisfies the equation

$$0 = \tfrac{1}{2} \operatorname{grad} \chi(y) = \Lambda y - c - \lambda y = 0$$

or

$$(\lambda_i - \lambda) y_i = c_i, \qquad i = 1, \cdots, n. \tag{2.5}$$

If we solve this formally for y_i and substitute into (2.3) we are led to consider the

Received November 12, 1965.

* The author is with the Mathematical Analysis Staff, The Boeing Company, Seattle, Washington.

real roots of the equation

$$g(\lambda) = 1 \tag{2.6}$$

with

$$g(\lambda) = \sum_{i=1}^{n}{}' \frac{|c_i|^2}{(\lambda - \lambda_i)^2}. \tag{2.7}$$

A primed summation sign means terms with $c_i = 0$ are dropped, whatever the value of $\lambda - \lambda_i$. Two cases can occur:

Case I. λ is a real root of (2.6) and $\lambda \neq \lambda_i$ for all i. Then (2.5) gives the components of an extremal vector y_λ associated with λ.

Case II. For some k, $g(\lambda_k) \leqq 1$. This requires $c_i = 0$ for all i such that $\lambda_i = \lambda_k$. To obtain the components of an extremal vector y_{λ_k} associated with λ_k, solve (2.5) for y_i if $\lambda_i \neq \lambda_k$, then select any y_i for i such that $\lambda_i = \lambda_k$ so that

$$\sum_{i:\lambda_i=\lambda_k} |y_i| = 1 - g(\lambda_k). \tag{2.8}$$

Then both (2.5) and the constraint (2.3) are satisfied.

THEOREM. *Let λ_j be the largest eigenvalue of A for which $g(\lambda_j) \leqq 1$. Let $\underline{\lambda}$ be the largest root of* (2.6) *with $\underline{\lambda} \neq \lambda_i$, $i = 1, \cdots, n$. The quadratic polynomial $\psi(y)$ is maximized by a vector associated with the larger of $\underline{\lambda}$ and λ_j.*

PROOF. For real $\lambda \neq \lambda_i$, $i = 1, \cdots, n$, let the components of y_λ be given by (2.5), then

$$\begin{aligned} \psi(y_\lambda) &= \sum_{i=1}^{n} \lambda_i \frac{|c_i|^2}{(\lambda_i - \lambda)^2} - 2\,\mathrm{Re}\left\{\sum_{i=1}^{n} \frac{|c_i|^2}{\lambda_i - \lambda}\right\} \\ &= \sum_{i=1}^{n} |c_i|^2 \left[\frac{\lambda_i}{(\lambda_i - \lambda)^2} - \frac{2}{\lambda_i - \lambda}\right] \\ &= \lambda \sum_{i=1}^{n} \frac{|c_i|^2}{(\lambda - \lambda_i)^2} + \sum_{i=1}^{n} \frac{|c_i|^2}{\lambda - \lambda_i} \\ &= \lambda g(\lambda) + \sum_{i=1}^{n} \frac{|c_i|^2}{\lambda - \lambda_i}. \end{aligned} \tag{2.9}$$

If λ is a root of (2.6), then

$$\psi(y_\lambda) = \lambda + \sum_{i=1}^{n} \frac{|c_i|^2}{\lambda - \lambda_i}. \tag{2.10}$$

If $\lambda = \lambda_k$ and the other conditions of Case II are fulfilled, then the value of $\psi(y_\lambda)$ for $\lambda = \lambda_k$ is calculated by priming the summation sign in (2.9) and adding

$$\lambda_k \sum_{i:\lambda_i=\lambda_k} |y_i|^2.$$

We then have

$$\psi(y_\lambda) = \lambda g(\lambda) + \sum_{i=1}^{n}{}' \frac{|c_i|^2}{\lambda - \lambda_i} + \lambda_k \sum_{i:\lambda_i=\lambda_k} |y_i|^2 = \lambda + \sum_{i=1}^{n}{}' \frac{|c_i|^2}{\lambda - \lambda_i}. \tag{2.11}$$

When $\lambda \neq \lambda_i$ for all i, (2.11) is the same as (2.10). Therefore, (2.11) is true for all extremal vectors. To complete the proof, let μ, ν be two values of λ which satisfy the conditions of either Case I or Case II, and suppose $\mu > \nu$. Then

$$\begin{aligned}
\psi(y_\mu) - \psi(y_\nu) &= \mu + \sum_{i=1}^{n}{}' \frac{|c_i|^2}{\mu - \lambda_i} - \nu - \sum_{i=1}^{n}{}' \frac{|c_i|^2}{\nu - \lambda_i} \\
&= \mu - \nu + \sum_{i=1}^{n}{}' |c_i|^2 \left(\frac{1}{\mu - \lambda_i} - \frac{1}{\nu - \lambda_i} \right) \\
&= (\mu - \nu) \left[1 - \sum_{i=1}^{n}{}' \frac{|c_i|^2}{(\mu - \lambda_i)(\nu - \lambda_i)} \right] \\
&\geqq (\mu - \nu) \left[\frac{1}{2} g(\mu) + \frac{1}{2} g(\nu) - \sum_{i=1}^{n}{}' \frac{|c_i|^2}{(\mu - \lambda_i)(\nu - \lambda_i)} \right] \\
&\geqq \frac{1}{2} (\mu - \nu) \sum_{i=1}^{n}{}' |c_i|^2 \left[\frac{1}{(\mu - \lambda_i)^2} + \frac{1}{(\nu - \lambda_i)^2} - \frac{2}{(\mu - \lambda_i)(\nu - \lambda_i)} \right] \\
&\geqq 0.
\end{aligned}$$

Therefore, $\psi(y_\lambda)$ increases for increasing λ which satisfy either Case I or Case II. This proves the theorem; a similar statement about the minimum of the polynomial is easily proven.

III. An Application. Let (x, y, z) be the position vector of a target in a coordinate system attached to a rolling ship and $(\dot{x}, \dot{y}, \dot{z})$ the target's inertial velocity vector in the same coordinates. Consider the angular accelerations of a gun tracking this target. The gun has the usual two degrees of freedom: a train axis perpendicular to the deck and an elevation axis perpendicular to the train axis. Let θ be the train angle. The parts of the train angular acceleration $\ddot{\theta}$ which contain the target velocity are

$$\begin{aligned}
\ddot{\theta}(\dot{x}, \dot{y}, \dot{z}) &= 2(x^2 + y^2)^{-2}\{xy(\dot{x}^2 - \dot{y}^2) - \dot{x}\dot{y}(x^2 - y^2) \\
&\qquad + \dot{R}[z(y^2 - x^2)\dot{x} - 2xyz\dot{y}]\} + 2\dot{R}x(x^2 + y^2)^{-1}\dot{z},
\end{aligned} \tag{3.1}$$

where $\dot{R}$ is the roll rate (assumed to be about the x-axis). The last term can be recognized as a component of the Coriolis acceleration. The remaining terms can be computed by considering the relative motion in a nonrotating system (i.e., take two derivatives of $y = x \tan \theta$). The problem of maximizing the entire expression as a function of $\dot{x}, \dot{y}, \dot{z}$ with $\dot{x}^2 + \dot{y}^2 + \dot{z}^2 = 1$ and fixed $x, y, z, \dot{R}$ is of the type considered, with A singular. In fact,

$$A = 2(x^2 + y^2)^{-2} \begin{pmatrix} xy & -\frac{1}{2}(x^2 - y^2) & 0 \\ -\frac{1}{2}(x^2 + y^2) & -xy & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.2}$$

and

$$a^* = -(x^2 + y^2)^{-2}\dot{R}(z(y^2 - x^2), -2xyz, x(x^2 + y^2)). \tag{3.3}$$

Further computation yields

$$\lambda_1 = -(x^2 + y^2)^{-1}, \qquad \lambda_2 = (x^2 + y^2)^{-1}, \qquad \lambda_3 = 0; \tag{3.4}$$

$$U = [2(x^2 + y^2)]^{-1/2} \begin{pmatrix} x - y & x + y & 0 \\ x + y & y - x & 0 \\ 0 & 0 & [2(x^2 + y^2)]^{1/2} \end{pmatrix}; \tag{3.5}$$

$$c^* = a^*U = -(2)^{-1/2}(x^2 - y^2)^{-3/2}\dot{R}(-z(x + y), z(y - x), x[2(x^2 + y^2)]^{1/2}). \tag{3.6}$$

Therefore,

$$(x^2 + y^2)\psi = y_2^2 - y_1^2 + \dot{R}(2)^{1/2}(x^2 + y^2)^{-1/2} \cdot(-z(x + y)y_1 + z(y - x)y_2 + x[2(x^2 + y^2)]^{1/2}y_3). \tag{3.7}$$

After neglecting the fixed factor $x^2 + y^2$,

$$\frac{2g(\lambda)}{\dot{R}^2} = \frac{z^2(x + y)^2}{(x^2 + y^2)(\lambda + 1)^2} + \frac{z^2(y - x)^2}{(x^2 + y^2)(\lambda - 1)^2} + \frac{2x^2}{\lambda^2}. \tag{3.8}$$

In the general case, when none of the numerators are zero, the problem is solved by finding the largest real root of (3.8) with $g(\lambda) = 1$. Classical root calculation procedures, such as Newton's method, should encounter no difficulty. If one or more of the numerators are zero, the computation is simpler. For example, if $z = 0$, $g(\lambda) = \dot{R}^2x^2/\lambda^2$ and Case I applies if $\lambda = |\dot{R}x| \geqq 1$. Then $y_1 = y_2 = 0, y_3 = \pm 1$; if $|\dot{R}x| < 1$, then Case II applies and $y_1 = 0$, $y_2 = (1 - \dot{R}^2x^2)^{1/2}$, $y_3 = \dot{R}x$. The geometric interpretation of this is that the Coriolis term predominates for large x values.

1. GEORGE E. FORSYTHE & GENE H. GOLUB, *Maximizing a second-degree polynomial on the unit sphere*, Tech. Rep. CS16, Stanford University Computer Science Department, Stanford, Calif., 1965.

Questions Concerning Khintchine's Constant and the Efficient Computation of Regular Continued Fractions

By John W. Wrench, Jr. and Daniel Shanks

Let x be a real number whose regular continued fraction is given by

$$x = a_0 + \frac{1}{a_1 +} \frac{1}{a_2 +} \frac{1}{a_3 +} \cdots, \tag{1}$$

with a_0 an integer, and $a_1, a_2, a_3, \cdots$ positive integers. Let

$$G_n(x) = (a_1 \cdot a_2 \cdot a_3 \cdot \cdots \cdot a_n)^{1/n}. \tag{2}$$

Then Khintchine's famous theorem states that, for almost all x,

$$\lim_{n\to\infty} G_n(x) = K, \tag{3}$$

Received January 24, 1966.