CIRCUMSCRIBING AN ELLIPSOID ABOUT THE INTERSECTION OF TWO ELLIPSOIDS

W. Kahan

(received December 27, 1967)

An ellipsoid $\mathcal{A}$ is associated uniquely with a positive definite matrix A via

$$x \in \mathcal{A} \text{ if and only if } x'Ax \leq 1 .$$

Note that all ellipsoids discussed here are centred at 0. Given $\mathcal{A}_1$ and $\mathcal{A}_2$ we seek another ellipsoid $\mathcal{H}$ circumscribed about $\mathcal{A}_1 \cap \mathcal{A}_2$. It is easy to see that

$$\mathcal{H} \supseteq \mathcal{A}_1 \cap \mathcal{A}_2 \text{ if and only if } x'Hx \leq \max_i x'A_ix \text{ for all vectors } x.$$

However, the last condition does not determine H uniquely; there are many circumscribing ellipsoids $\mathcal{H}$.

Let us say that $\mathcal{H}$ is <u>tight</u> whenever

$$\mathcal{H} \supseteq \mathcal{M} \supseteq \mathcal{A}_1 \cap \mathcal{A}_2 \text{ implies } \mathcal{M} = \mathcal{H};$$

in other words, $\mathcal{H}$ circumscribes $\mathcal{A}_1 \cap \mathcal{A}_2$ so tightly that no other ellipsoid $\mathcal{M}$ can be slipped between. Tightness does not determine $\mathcal{H}$ uniquely either, as we shall see, but yields a worthwhile simplification.

THEOREM. *$\mathcal{H}$ is tight if and only if $H = \alpha_1 A_1 + \alpha_2 A_2$ for some $\alpha_i \geq 0$ satisfying $\alpha_1 + \alpha_2 = 1$, except that if $\mathcal{A}_i \supseteq \mathcal{A}_j$ then only $\mathcal{H} = \mathcal{A}_j$ is tight.*

Proof. Suppose $\mathcal{H}$ is tight and let

$$\phi \equiv \min_{x \neq 0} \max_i x'A_ix/x'Hx .$$

Since ϕ is the minimum of a continuous function on a compact set ($x'Hx = 1$), the minimum is achieved at some vector $z \neq 0$; say

https://doi.org/10.4153/CMB-1968-050-2 Published online by Cambridge University Press

$$\frac{z'A_1 z}{z'H z} \le \phi = \frac{z'A_2 z}{z'H z} \le \max_i \frac{x'A_i x}{x'H x} \quad \text{for all} \quad x \ne 0 .$$

Clearly $\phi \ge 1$ since $\mathfrak{H} \supseteq \mathfrak{a}_1 \cap \mathfrak{a}_2$. Since $\mathfrak{H}$ is also tight, the inequality

$$\phi x'Hx \le x'Mx \le \max_i x'A_i x \quad \text{for all} \quad x$$

implies $H = M$. We shall prove the theorem by constructing an $M \equiv \alpha_1 A_1 + \alpha_2 A_2$ which satisfies the last inequality and has $\alpha_i \ge 0$ and $\alpha_1 + \alpha_2 = 1$.

There are three cases:

Case 1: Suppose $z'A_1 z < z'A_2 z = \phi z'Hz$. We shall show $\mathfrak{H} \supseteq \mathfrak{M} \equiv \mathfrak{a}_2$, whence tightness will imply $H = A_2$, by showing that $y'Hy \le y'A_2 y$ for all y. Indeed, we shall show that $\phi y'Hy \le y'A_2 y$, which is a stronger statement because $\phi \ge 1$.

Given any y, let $x \equiv z + \lambda y$; it is easily seen that

$$x'A_1 x < x'A_2 x \quad \text{and} \quad \phi x'Hx \le x' A_2 x$$

for all sufficiently small λ. For these λ we have $0 \le x'(A_2 - \phi H)x = 2\lambda y'(A_2 - \phi H)z + \lambda^2 y'(A_2 - \phi H)y$; and if we choose sign $(\lambda) \ne 0$ in such a way that $\lambda y'(A_2 - \phi H)z \le 0$, as can be done for any y, we obtain $y'(A_2 - \phi H)y \ge 0$. Therefore $\mathfrak{H} \supseteq \mathfrak{a}_2$ as claimed.

Cases 2 and 3: Suppose $z'A_1 z = z'A_2 z = \phi z'Hz$. Now define

$$y \equiv \phi Hz - \alpha_1 A_1 z - \alpha_2 A_2 z$$

by choosing α_1 and α_2 in such a way that

$$y'A_1 z = y'A_2 z = 0 .$$

Such a choice of α_1 and α_2 is easily seen to be possible, though not necessarily unique. Furthermore, we can deduce that $y = 0$ as follows.

Set $x \equiv \lambda z + y$; since $\phi x'Hx \le \max_i x'A_i x$ and $\phi z'Hz = z'A_i z$,

https://doi.org/10.4153/CMB-1968-050-2 Published online by Cambridge University Press

$$2\lambda\phi y'Hz + \phi y'Hy \leq \max_i (2\lambda y'A_i z + y'A_i y) \qquad (*)$$

for all y and λ. In particular, the y defined above must satisfy

$$2\lambda y'y + \phi y'Hy \leq \max_i y'A_i y$$

for all λ, and letting $\lambda \to +\infty$ implies $y = 0$ as claimed. Therefore there are some α_i such that

$$\phi Hz = \alpha_1 A_1 z + \alpha_2 A_2 z ;$$

pre-multiplying by z' shows that $\alpha_1 + \alpha_2 = 1$.

<u>Case</u> 2: Suppose $A_1 z \neq A_2 z$. We shall show next that both $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$. Note that the $A_i z$ are linearly independent; if $\beta_1 A_1 z = \beta_2 A_2 z$ then pre-multiplication by z' yields $\beta_1 = \beta_2$, which contradicts our supposition unless $\beta_1 = \beta_2 = 0$. Therefore we may so choose y that $y'A_1 z > 0 > y'A_2 z$; one such choice is $y \equiv A_1 z/\sqrt{(z'A_1^2 z)} - A_2 z/\sqrt{(z'A_2^2 z)}$. Substitute this y in (*) above and let λ approach first $-\infty$ and then $+\infty$; we find that

$$y'A_2 z \leq \alpha_1 y'A_1 z + \alpha_2 y'A_2 z \leq y'A_1 z .$$

Since $\alpha_1 + \alpha_2 = 1$, it soon follows that both $\alpha_i \geq 0$.

Now set $M \equiv \alpha_1 A_1 + \alpha_2 A_2$, and observe that $\mathfrak{M} \supseteq \mathfrak{A}_1 \cap \mathfrak{A}_2$. We shall show that $\mathfrak{H} \supseteq \mathfrak{M}$ by showing that $y'My \geq \phi y'Hy$ ($\geq y'Hy$) for all y. Given any y for which $y'(A_1 - A_2)z \neq 0$, set

$\lambda \equiv \frac{1}{2} y'(A_2 - A_1)y/y'(A_1 - A_2)z$ and $x \equiv \lambda z + y$. After verifying that $x'A_1 x = x'A_2 x = x'Mx$, we infer from the definition of ϕ that $x'Mx \geq \phi x'Hx$, and from the suppositions about z and $Mz = \phi Hz$ that $y'My \geq \phi y'Hy$. This inequality has been proved for all y not in the plane $y'(A_1 - A_2)z = 0$; the inequality is true for all y by virtue of continuity.

<u>Case</u> 3: Suppose $A_1 z = A_2 z$. We may restrict our attention to the subspace L of vectors y satisfying $y'Hz = y'A_i z = 0$, because every vector x can be decomposed uniquely into a sum $x = \lambda z + y$ with $y \in L$, and $x'(A_i - \phi H)x = y'(A_i - \phi H)y$. Therefore the theorem's proof is reduced to the problem of finding $M \equiv \alpha_1 A_1 + \alpha_2 A_2$ with both $\alpha_i \geq 0$

https://doi.org/10.4153/CMB-1968-050-2 Published online by Cambridge University Press

and $\alpha_1 + \alpha_2 = 1$ such that $\phi y'Hy \le y'My$ for all $y \in L$, given only that $\phi y'Hy \le \max_i y'A_i y$ for all $y \in L$. This is just the problem with which the proof began. The problem is obviously solved if L is 1-dimensional, and otherwise the problem is solved by repeating the foregoing calculations in L.

Therefore, if $\mathcal{H}$ is tight then $H = \alpha_1 A_1 + \alpha_2 A_2$. Conversely, suppose $H = \alpha_1 A_1 + \alpha_2 A_2$ with both $\alpha_i \ge 0$ and $\alpha_1 + \alpha_2 = 1$; is $\mathcal{H}$ tight ? The foregoing calculations provide $M \equiv \beta_1 A_1 + \beta_2 A_2$ with both $\beta_i \ge 0$, $\beta_1 + \beta_2 = 1$, and $x'Hx \le x'Mx$ for all x; in other words, $(\alpha_1 - \beta_1)x'(A_2 - A_1)x \ge 0$. If $\mathcal{Q}_2 \not\supseteq \mathcal{Q}_1 \not\supseteq \mathcal{Q}_2$ then $x'(A_2 - A_1)x$ takes on both positive and negative values, so $\alpha_i = \beta_i$; otherwise if $\mathcal{Q}_1 \supseteq \mathcal{Q}_2$ then obviously only $\mathcal{Q}_2$ can be tight (cf. Case 1). So ends the proof.

Must the proof be so long?

Remarks. The theorem has applications to certain schemes for circumscribing complicated regions by simple ones in an electronic computer. Ellipsoids are regarded as simple because they are representable by matrices. However, complicated regions are often far smaller than the simplest region circumscribed around them, and therefore we are sometimes forced to manipulate unions, intersections, sums and other combinations of simple regions. Storage capacity limits the complexity achievable in practice, and forces simplifications of which one kind has been discussed above. For further work along these lines see Schweppe (1967) and Kahan (1967-8).

There are several questions requiring further study. If several ellipsoids $\mathcal{Q}_i$ are given, then $H \equiv \Sigma_i \alpha_i A_i$ with all $\alpha_i \ge 0$ and $\Sigma_i \alpha_i = 1$ produces $\mathcal{H} \supseteq \bigcap_i \mathcal{Q}_i$; but what characterizes the tight ellipsoids $\mathcal{H}$? If ellipsoids $\mathcal{Q}_i$ are not centred at 0, but instead consist of points x for which $(x - c_i)' A_i (x - c_i) \le 1$, then the characterization of tight ellipsoids is further complicated. Indeed, even to tell whether two such ellipsoids intersect seems to demand the solution of an eigenproblem; see Forsythe and Golub (1965) and Burrows (1966).

Acknowledgement: The author's work has been assisted by grants from the National Research Council of Canada.

REFERENCES

J. W. Burrows (1966), Maximization of a second-degree polynomial on the unit sphere. Math. of Comp. 20, 441-444.

G. E. Forsythe and G. H. Golub (1965), On the stationary values of a second-degree polynomial on the unit sphere. J. Soc. Indust. Appl. Math. 13, 1050-1068.

W. Kahan (1967), Circumscribing an ellipsoid about the Minkowski sum of given ellipsoids. (Submitted to J. Linear Algebra). *

W. Kahan (1968), An ellipsoidal error bound for linear systems of ordinary differential equations. (Manuscript to appear).

F. C. Schweppe (1967), Recursive state estimation when observation errors and system inputs are bounded. (Sperry Rand Research Centre Report RR-67-25, Sudbury, Mass.)

University of Toronto

* Added in proof: Related problems are considered by D.K. Faddeèv and V.N. Faddeèva (1968) "Stability in Linear Algebra Problems". Proc. IFIP Congress 68 in Edinburgh.