JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 28, No. 1, January–February 2005

Technical Comments

Brief discussion of previous investigations in the aerospace sciences and technical comments on papers published in the Journal of Guidance, Control, and Dynamics *are presented in this special department. Entries must be restricted to a maximum of 1000 words, or the equivalent of one* Journal *page including formulas and figures. A discussion will be published as quickly as possible after receipt of the manuscript. Neither the AIAA nor its editors are responsible for the opinions expressed by the correspondents. Authors will be invited to reply promptly.*

Comment on "Determining If Two Solid Ellipsoids Intersect"

Thomas H. Kerr III*
TeK Associates, Lexington, Massachusetts 02420-5312

ALTHOUGH Alfano and Greer[1] have an elegant solution for assessing whether two three-dimensional ellipsoids overlap, which is both easy to understand and apparently straightforward to test for numerically, this Comment offers some clarification regarding the anticipated computational load and numerical sensitivity of the test, which is claimed in Ref. 1 to be light enough for it to constitute a real-time test. The implied eigenvalue–eigenvector calculation usually involves an iterative solution algorithm that, while almost instantaneous on general-purpose machines such as personal computers, may not be so readily available on embedded processors. Reference 1 advocates using explicit closed-form solutions for the quadratic surface (which arises from quadrics in four dimensions for three-dimensional ellipsoids) and for the conic curve (arising from quadrics in three dimensions for two-dimensional ellipses), but this path can be challenging when we seek to elucidate all possible situations for the polynomial coefficients to be encountered for the general case (as derived from the underlying matrices) and, we emphasize here, even for obtaining merely the defining characteristic equation that is to be solved for λ. Obtaining the characteristic equation involves expanding by minors and (Ref. 2, Sec. 2.4.3) identifies such operations as situations where we should "expect loss of correct significant digits when small numbers are additively computed from larger numbers" because "when calculations are performed on a computer, each arithmetic operation is generally affected by round-off error" (Ref. 2, Sec. 2.4.1). An exception is when only matrices with integer entries are present throughout all computations, but such examples are difficult to construct for the purpose of providing illustrative examples for eigenvalue–eigenvalue problems[3] (unless the matrices involved are merely diagonal and corresponding matrix inverses are obtained by merely taking the reciprocal of the diagonal terms, which yields proper fractions unless all original diagonal terms are 1).

Closed-form solutions of polynomial equations, such as are currently advocated in Ref. 1, where coefficients are derived from the determinants of more general (although positive definite) matrices, still involve the differences of large numbers and typically exhibit numerical sensitivities as a consequence. Only very simple special-case numerical examples with diagonal entries are treated in Ref. 1 to illustrate the behavioral trends and associated classifications, although the method is general [but messier for general three-dimensional covariance matrices exhibiting more arbitrary orientations and for machine-imposed floating point representations of the numbers (expected to be encountered within the application scenario as the more likely category of common formatting for matrix entries)]. According to Ref. 2, Sec. 7.2, "The act of computing eigenvalues is the act of computing the zeros of the characteristic polynomial. *Galois* theory tells us that such a process has to be iterative if $n > 4$ and so error will arise because of finite termination" (of such iterative algorithms and the computed answers).

Before we proceed, a distinction is made here between what is offered in Ref. 1 and what is offered in Ref. 4 as a test for ellipsoid containment before other historical connections and observations are made. Reference 2 provides a test for full containment of one ellipsoid within another only when they share a common center, $\bar{\boldsymbol{x}}$, as between

$$(\boldsymbol{x}-\bar{\boldsymbol{x}})^T\left(\tfrac{1}{2}\right)\boldsymbol{P}_1^{-1}(\boldsymbol{x}-\bar{\boldsymbol{x}}) \le 1 \qquad \text{and} \qquad (\boldsymbol{x}-\bar{\boldsymbol{x}})^T\left(\tfrac{1}{2}\right)\boldsymbol{P}_2^{-1}(\boldsymbol{x}-\bar{\boldsymbol{x}}) \le 1 \tag{1}$$

and the second is fully contained within the first if and only if

$$\boldsymbol{P}_1 < \boldsymbol{P}_2 \tag{2}$$

as a strict positive definiteness condition on matrices that themselves are each positive definite (as are all well-behaved, nondegenerate covariance matrices.[5,6]) A similar requirement on the two covariances participating in an earlier test for ellipsoid overlap (not containment) was encountered in Ref. 7 before test could be specified for ellipsoid overlap (in n dimensions) where the centers of the respective ellipsoids could differ, and where the particular covariance matrix, P_1 (in the case of Ref. 7, this was the solution of the Riccati equation) is so related to the other covariance matrix, P_2 (in the case of Ref. 7, this was the solution of the Lyapunov equation). The proof of Eq. (2) was easily obtained in Lemma 5.1 of Ref. 7 by just taking the synchronous difference of the two respective matrix difference equations that describe their evolution (in discrete time) by demonstrating that the difference is always positive definite (as it evolves for all time steps $k > 0$) as the positive definite matrix within the bracket below, as pre- and postmultiplied by a nonsingular matrix (Recall that the computed transition matrix is always nonsingular) and its transpose (yielding a positive semidefinite intermediary matrix as the first term) and added to a strictly positive definite matrix (the second term) to yield a strictly positive definite matrix result as

$$\begin{aligned}[P_2(k+1) - P_1(k+1\,|\,k)] &= \Phi(k+1,k)[P_2(k)\\ &\quad - P_1(k\,|\,k)]\Phi^T(k+1,k) + \Phi(k+1,k)P_1(k\,|\,k-1)H^T\\ &\quad \times\left[HP_1(k\,|\,k-1)H^T + R(k)\right]^{-1}HP_1(k\,|\,k-1)\Phi^T(k+1,k)\end{aligned} \tag{3}$$

The associated optimization problem in Ref. 7 had an intriguing similarity to that in Ref. 1, as explained later. In the case of posing the simpler problem of a one-dimensional test for the overlap of scalar Gaussian confidence intervals in Ref. 8 to show how the same test then generalizes to n dimensions, as a test for the overlap of Gaussian ellipsoidal confidence regions, the version of the test in Ref. 8 (simpler than that in Ref. 7) made possible a closed-form answer to the optimization.

The overlap test of Ref. 1 needs matrix positive definiteness/semidefiniteness tests along with an implied eigenvalue–eigenvector

Received 19 February 2004; revision received 25 June 2004; accepted for publication 30 June 2004.

*CEO/Principal Investigator, TK-MIP Software Algorithm Department, P.O. Box 459, 9 Meriam Street, Suite 7-R; thomas_h_kerr@msn.com. Senior Member AIAA.

calculation. The test is obtained by exploiting features of a three-dimensional ellipsoid translation represented as a rotation in four-dimensional space, a technique familiar in computer graphics applications (Ref. 9, pp. 479–481), included for the two ellipsoids of interest as

$$x^T \overbrace{M S_1 M^T}^{A} x = 0 \qquad \text{for} \qquad S_1 \triangleq \begin{bmatrix} \left(\frac{1}{2}\right) P_1^{-1} & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix}$$

with offset

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\bar{x}_1 & -\bar{x}_2 & -\bar{x}_3 & 1 \end{bmatrix}$$

where

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} \tag{4}$$

$$x^T \overbrace{S_2}^{B} x = 0, \qquad S_2 \triangleq \begin{bmatrix} \left(\frac{1}{2}\right) P_2^{-1} & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \quad \text{with no offset} \tag{5}$$

(without loss of generality, because coordinate origin can always be moved to perform this numerical test at the location of the second possible offset, thus causing it to be zeroed out). After Eqs. (4) and (5) are combined, the test consists of solving for λ in

$$x^T A \left[\lambda I_{4\times 4} - A^{-1} B\right] x = 0 \tag{6}$$

to determine whether the underlying two three-dimensional ellipsoids of primary interest either overlap. Corresponding compatible eigenvectors also need to be found and tested for consistency to complete the test of Ref. 1. Observe that a solution to the well-known generalized eigenproblem $\lambda A x = B x$ (Ref. 2, Sec. 7.7) is also a solution of the fundamental Eq. (12) of Ref. 1 because $\lambda A x = B x \Leftrightarrow [\lambda A - B] x = 0 \Rightarrow x^T [\lambda A - B] x = 0$. Use of Choleski factorization and the symmetric QR algorithm is offered in Ref. 2, Sec. 8.7.2, as a stable solution for the case of A, B being symmetric and A being positive definite, as is in fact the case for the matrices encountered in Ref. 1 and herein. Observe that Ref. 1 deduces overlap by focusing on how pairs of eigenvalues of nonsymmetric $A^{-1}B$ behave. Symmetric matrices have all real eigenvalues but nonsymmetric matrices sometimes have complex eigenvalues.

The clear result of Ref. 1 was obtained by embedding a test for the overlap of n-dimensional ellipsoids into a test that is performed in an associated $(n+1)$-dimensional space (which, coincidentally, the analysis of Refs. 7 and 8 also did). However, the resulting test in Ref. 1 appears to be simpler to implement as a lesser computational burden (than that of Ref. 7, obtained 30 years earlier) by Ref. 1 apparently avoiding any intermediate iterative techniques in solving for the implied eigenvalues and eigenvectors used in making the determination. However, additional logic still needs to be programmed for scaling the last component of the eigenvector x to be 1, consistent with the methodology's acknowledged constraint encountered after the n-dimensional problem has been embedded into $(n+1)$ dimensions, and for other aspects of unwinding or interpreting a final decision regarding the presence or absence of overlap). Reference 1, not needing any condition of Eq. (2) to be satisfied, is for a case more general than that treated in Refs. 7, 8, 10, and 11; however, the numerical calculations of Refs. 7 and 8 are tailored for a stand-alone real-time decision (which was used aboard U.S. submarines). If one were to attempt to generalize the results of Ref. 1 beyond two- and three- to n dimensions (as already done in Ref. 12 for just the theory and proofs[3]), a modified version of the computational approach of Refs. 7 and 8 may be useful in this endeavor (and perhaps even for two- and three dimensions as well) because the iterative algorithm used is a contraction mapping with a geometric rate of convergence (but needs to use double precision for all matrices and vectors involved).

The solution offered in Ref. 13 is a precedent for what is speculated later in Ref. 12, Conclusions, as likely being possible in the future: to be able to solve successfully for the simultaneous intersection of several quadratic surfaces. However, intersection of four (or more) quadratic surfaces in 4-space (consisting of three coordinates of position and receiver's time clock offset) was obtained in closed form in Ref. 13. However, despite a closed-form solution path now being available, global positioning system receivers continue to use an earlier iterative solution approach that also yields, as a by-product, an evaluation of associated geometric dilution of precision, which assesses the goodness of the satellite geometry.

The use of an iterative solution technique is not necessarily at odds with providing real-time answers and may be the simplest path to follow. A navigation application using an even easier criterion of ellipsoid containment in dimensions higher than three is discussed in Ref. 14. Some missile-chasing-interceptor tracking-sensor processing and also some multiplatform rendezvous control strategies may benefit by focusing on use of six-dimensional ellipsoids that conjoin three-dimensional position ellipsoids together with three-dimensional velocity ellipsoids within a combined six-dimensional hypothesis test, because both position- and velocity-tracking errors originate from a common source, being computed outputs of the same on-line estimation algorithm from which the target tracks were generated (i.e., spawned from the same common sensor measurement data as fundamental stimulus) and the associated cross covariances, as computed, can be seen and verified to be nonzero, as further evidence that both computed velocity errors and computed position errors are cross-correlated and should be considered together jointly in a test (which does not preclude also using them separately afterward in individual three-dimensional tests).

References

[1]Alfano, S., and Greer, M. L., "Determining If Two Solid Ellipsoids Intersect," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 1, 2003, pp. 106–110.

[2]Golub, G. H., and Van Loan, C. F., *Matrix Computations*, 3rd ed., Johns Hopkins Univ. Press, Baltimore, MD, 1996.

[3]Galvin, W. P., "Matrices with 'Custom-Built' Eigenspaces," *American Mathematical Monthly*, Vol. 91, No. 5, 1984, pp. 308, 309.

[4]Kalandros, M., Pao, L Y., "Covariance Control for Multisensor Systems," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, No. 4, 2002, pp. 1138–1157.

[5]Kerr, T. H., "On Misstatements of the Test for Positive Semidefinite Matrices," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 3, 1990, pp. 571, 572.

[6]Kerr, T. H., "Fallacies in Computational Testing of Matrix Positive Definiteness/Semi-definiteness," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 26, No. 2, 1990, pp. 415–421.

[7]Kerr, T. H., "Real-Time Failure Detection: A Static Nonlinear Optimization Problem That Yields a Two Ellipsoid Overlap Test," *Journal of Optimization Theory and Applications*, Vol. 22, No. 4, 1977, pp. 509–536.

[8]Kerr, T. H., "Statistical Analysis of a Two Ellipsoid Overlap Test for Real-Time Failure Detection," *IEEE Transactions on Automatic Control*, Vol. 25, No. 4, 1980, pp. 762–773.

[9]Kalatchin, P., Chebotko, I., Podvoiski, E., Velikevich, S., Kiselev, Y., and Ryabov, T., *The Revolutionary Guide to Bitmapped Graphics*, Wrox Press, Birmingham, England, U.K., 1994.

[10]Kerr, T. H., "False Alarm and Correct Detection Probabilities over a Time Interval for Restricted Classes of Failure Detection Algorithms," *IEEE Transactions on Information Theory*, Vol. 28, No. 4, 1982, pp. 619–631.

[11]Kerr, T. H., "Comments on 'An Algorithm for Real-Time Failure Detection in Kalman Filters,'" *IEEE Transactions on Automatic Control*, Vol. 43, No. 5, 1998, pp. 682–683.

[12]Chan, K., "A Simple Mathematical Approach for Determining Intersection of Quadratic Surfaces," *Proceedings of the American Astronautical Society*, Vol. 109, Pt. 1, 2001, pp. 785–801.

[13]Leva, J. L., "An Alternative Closed-Form Solution to the GPS Pseudo-range Equations," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 32, No. 4, 1996, pp. 1430–1439.

[14]Kerr, T. H., "Decentralized Filtering and Redundancy Management for Multisensor Navigation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 23, No. 1, 1987, pp. 83–119.