Technical Comments

Comment on “Determining If Two Solid Ellipsoids Intersect”

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

Although Alfano and Greer¹ 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 additive 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-eigenvector problems⁵ (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, \( \mathbf{e} \), as between

\[
(x - \mathbf{x})^T \left( \frac{1}{2} P_1 \right)^{-1} (x - \mathbf{x}) \leq 1 \quad \text{and} \quad (x - \mathbf{x})^T \left( \frac{1}{2} P_2 \right)^{-1} (x - \mathbf{x}) \leq 1
\]

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

\[ P_1 < P_2 \]

as a strict positive definiteness condition on matrices that themselves are each positive definite (as are all well-behaved, nondegenerate covariance matrices).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:

\[
P_2 (k + 1) - P_1 (k + 1) [k] = \Phi (k + 1, k) [P_2 (k) - P_1 (k + 1, k)] + \Phi (k + 1, k) P_1 (k + 1, k) (k) H^T \\
\times [H P_2 (k) (k) (k) H^T + R(k)]^{-1} H P_2 (k) (k) (k) \Phi^T (k + 1, k)
\]

(3)

The associated optimization problem in Ref. 7 had an intriguing similarity to that in Ref. 1, as explained below. In the case of using 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 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

---


² Refer to Ref. 2, Sec. 7.2 for the Galois theory.


of interest as
\[ A \] for \( S_1 \triangleq \begin{bmatrix} \frac{1}{2} P_1^{-1} & 0_{3 \times 1} \\ 0_{1 \times 3} & 1 \end{bmatrix} \]

with offset \( M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\tilde{x}_1 & -\tilde{x}_2 & -\tilde{x}_3 & 1 \end{bmatrix} \)

where \( x = \begin{bmatrix} x_2 \\ x_3 \\ 1 \end{bmatrix} \) (4)

(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 \( \lambda \) in

\[ x^T \lambda [I_{4 \times 4} - \lambda^{-1} B] x = 0 \] (6)

to determine whether or not the underlying two- or three-dimensional ellipsoids of primary interest above either overlap or not. Corresponding compatible eigenvectors also need to be found and tested for consistency to complete the test of Ref. 1. [Observe that the 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. 10 and 12 also did). However, the resulting test in Ref. 1 appears to be simpler to implement as a lesser computational burden (that is to say, obtained 30 years earlier) by Ref. 1 apparently not 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. 1, 2, 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 proofs3), a modified version of the computation 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).

References
Queries

Q1. References must be listed in the order they are called out in the text. Please check renumbering.
Q2. Give job/position title./Give department if applicable.
Q3. OK as reworded?
Q4. The use of footnotes to text is strongly discouraged per AIAA style. Ok as set?
Q5. The use of footnotes to text is strongly discouraged per AIAA style. Ok as set?
Q6. Spell out acronym.
Q7. References must be listed in the order they are called out in the text. Please check renumbering.
Q8. Give volume number./Give issue number (or month if there is no issue number).
Q10. Verify acronym spell-out.