Intersection of Ellipsoids

David Eberly
Geometric Tools, LLC
http://www.geometrictools.com/

Created: October 11, 2000
Last Modified: March 2, 2008

Contents

1 Introduction

This article describes how to compute the points of intersection of two ellipsoids, a geometric query labeled *find intersections*. It also shows how to determine if two ellipsoids intersect without computing the points of intersection, a geometric query labeled *test intersection*. Specifically, the geometric queries for the ellipsoids E_0 and E_1 are:

- *Find Intersections*. If E_0 and E_1 intersect, find the points of intersection.
- *Test Intersection*. Determine if
  - E_0 and E_1 are separated (there exists a plane for which the ellipsoids are on opposite sides),
  - E_0 properly contains E_1 or E_1 properly contains E_0, or
  - E_0 and E_1 intersect.

An implementation of the *find* query, in the event of no intersections, might not necessarily determine if one ellipsoid is contained in the other or if the two ellipsoids are separated. Let the ellipsoids E_i be defined by the quadratic equations

$$\begin{aligned} Q_i(\mathbf{X}) &= \mathbf{X}^{\mathrm{T}} A_i \mathbf{X} + \mathbf{B}_i^{\mathrm{T}} \mathbf{X} + C_i \\ &= \left[\begin{array}{ccc} x & y & z \end{array}\right] \left[\begin{array}{ccc} a_{00}^{(i)} & a_{01}^{(i)} & a_{02}^{(i)} \\ a_{01}^{(i)} & a_{11}^{(i)} & a_{12}^{(i)} \\ a_{02}^{(i)} & a_{12}^{(i)} & a_{22}^{(i)} \end{array}\right] \left[\begin{array}{c} x \\ y \\ z \end{array}\right] + \left[\begin{array}{ccc} b_0^{(i)} & b_1^{(i)} & b_2^{(i)} \end{array}\right] \left[\begin{array}{c} x \\ y \\ z \end{array}\right] + c^{(i)} \\ &= 0 \end{aligned}$$

for $i = 0, 1$. It is assumed that the A_i are positive definite. In this case, $Q_i(\mathbf{X}) < 0$ defines the inside of the ellipsoid and $Q_i(\mathbf{X}) > 0$ defines the outside.

2 Find Intersection

2.1 Variation 1

The two polynomials $f(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2$ and $g(z) = \beta_0 + \beta_1 z + \beta_2 z^2$ have a common root if and only if the Bézout determinant is zero,

$$(\alpha_2\beta_1 - \alpha_1\beta_2)(\alpha_1\beta_0 - \alpha_0\beta_1) - (\alpha_2\beta_0 - \alpha_0\beta_2)^2 = 0.$$

This is constructed by the combinations

$$0 = \alpha_2 g(z) - \beta_2 f(z) = (\alpha_2\beta_1 - \alpha_1\beta_2)z + (\alpha_2\beta_0 - \alpha_0\beta_2)$$

and

$$0 = \beta_1 f(z) - \alpha_1 g(z) = (\alpha_2\beta_1 - \alpha_1\beta_2)z^2 + (\alpha_0\beta_1 - \alpha_1\beta_0),$$

solving the first equation for z and substituting it into the second equation. When the Bézout determinant is zero, the common root of $f(z)$ and $g(z)$ is

$$\bar{z} = \frac{\alpha_2\beta_0 - \alpha_0\beta_2}{\alpha_1\beta_2 - \alpha_2\beta_1}.$$

The common root to $f(z) = 0$ and $g(z) = 0$ is obtained from the linear equation $\alpha_2 g(z) - \beta_2 f(z) = 0$ by solving for z.

The ellipsoid equations can be written as quadratics in z whose coefficients are polynomials in x and y,

$$Q_i(x, y, z) = \left(a_{00}^{(i)} x^2 + 2a_{01}^{(i)} xy + a_{11}^{(i)} y^2 + b_0^{(i)} x + b_1^{(i)} y + c^{(i)}\right) + \left(2a_{02}^{(i)} x + 2a_{12}^{(i)} y + b_2^{(i)}\right) z + \left(a_{22}^{(i)}\right) z^2.$$

Using the notation of the previous paragraph with f corresponding to Q_0 and g corresponding to Q_1,

$$\begin{aligned}
\alpha_0 &= a_{00}^{(0)} x^2 + 2a_{01}^{(0)} xy + a_{11}^{(0)} y^2 + b_0^{(0)} x + b_1^{(0)} y + c^{(0)}, \;\; \alpha_1 = 2a_{02}^{(0)} x + 2a_{12}^{(0)} y + b_2^{(0)}, \;\; \alpha_2 = a_{22}^{(0)}, \\
\beta_0 &= a_{00}^{(1)} x^2 + 2a_{01}^{(1)} xy + a_{11}^{(1)} y^2 + b_0^{(1)} x + b_1^{(1)} y + c^{(1)}, \;\; \beta_1 = 2a_{02}^{(1)} x + 2a_{12}^{(1)} y + b_2^{(1)}, \;\; \beta_2 = a_{22}^{(1)}.
\end{aligned}$$

Let $\alpha_2\beta_1 - \alpha_1\beta_2 = \sum_{i+j\le 1} v_{0ij} x^i y^j$, $\alpha_1\beta_0 - \alpha_0\beta_1 = \sum_{i+j\le 3} v_{1ij} x^i y^j$, and $\alpha_2\beta_0 - \alpha_0\beta_2 = \sum_{i+j\le 2} v_{2ij} x^i y^j$. The Bézout determinant is a polynomial in x and y of degree at most 4, $R(x, y) = \sum_{i+j\le 4} r_{ij} x^i y^j$ where

$$\begin{aligned}
r_{40} &= v_{010}v_{130} - v_{220}^2 \\
r_{31} &= v_{010}v_{121} + v_{001}v_{130} - 2v_{211}v_{220} \\
r_{22} &= v_{010}v_{112} + v_{001}v_{121} - v_{211}^2 - 2v_{202}v_{220} \\
r_{13} &= v_{010}v_{103} + v_{001}v_{112} - 2v_{202}v_{211} \\
r_{04} &= v_{001}v_{103} - v_{202}^2 \\
r_{30} &= v_{010}v_{120} + v_{000}v_{130} - 2v_{210}v_{220} \\
r_{21} &= v_{010}v_{111} + v_{001}v_{120} + v_{000}v_{121} - 2v_{210}v_{211} - 2v_{201}v_{200} \\
r_{12} &= v_{010}v_{102} + v_{001}v_{111} + v_{000}v_{112} - 2v_{202}v_{210} - 2v_{201}v_{211} \\
r_{03} &= v_{001}v_{102} + v_{000}v_{103} - 2v_{201}v_{202} \\
r_{20} &= v_{010}v_{110} + v_{000}v_{120} - v_{210}^2 - 2v_{200}v_{220} \\
r_{11} &= v_{010}v_{101} + v_{001}v_{110} + v_{000}v_{111} - 2v_{201}v_{210} - 2v_{200}v_{211} \\
r_{02} &= v_{001}v_{101} + v_{000}v_{102} - v_{201}^2 - 2v_{200}v_{202} \\
r_{10} &= v_{010}v_{100} + v_{000}v_{110} - 2v_{200}v_{210} \\
r_{01} &= v_{001}v_{100} + v_{000}v_{101} - 2v_{200}v_{201} \\
r_{00} &= v_{000}v_{100} - v_{200}^2
\end{aligned}$$

with

$$
\begin{aligned}
v_{000} &= a_{22}^{(0)} b_2^{(1)} - a_{22}^{(1)} b_2^{(0)} \\
v_{010} &= 2a_{22}^{(0)} a_{02}^{(1)} - 2a_{22}^{(1)} a_{02}^{(0)} \\
v_{001} &= 2a_{22}^{(0)} a_{12}^{(1)} - 2a_{22}^{(1)} a_{12}^{(0)} \\
v_{100} &= b_2^{(0)} c^{(1)} - b_2^{(1)} c^{(0)} \\
v_{110} &= \left(2a_{02}^{(0)} c^{(1)} + b_2^{(0)} b_0^{(1)}\right) - \left(2a_{02}^{(1)} c^{(0)} + b_2^{(1)} b_0^{(0)}\right) \\
v_{101} &= \left(2a_{12}^{(0)} c^{(1)} + b_2^{(0)} b_1^{(1)}\right) - \left(2a_{12}^{(1)} c^{(0)} + b_2^{(1)} b_1^{(0)}\right) \\
v_{120} &= \left(2a_{02}^{(0)} b_0^{(1)} + b_2^{(0)} a_{00}^{(1)}\right) - \left(2a_{02}^{(1)} b_0^{(0)} + b_2^{(1)} a_{00}^{(0)}\right) \\
v_{111} &= \left(2a_{02}^{(0)} b_1^{(1)} + 2a_{12}^{(0)} b_0^{(1)} + 2b_2^{(0)} a_{01}^{(1)}\right) - \left(2a_{02}^{(1)} b_1^{(0)} + 2a_{12}^{(1)} b_0^{(0)} + 2b_2^{(1)} a_{01}^{(0)}\right) \\
v_{102} &= \left(2a_{12}^{(0)} b_1^{(1)} + b_2^{(0)} a_{11}^{(1)}\right) - \left(2a_{12}^{(1)} b_1^{(0)} + b_2^{(1)} a_{11}^{(0)}\right) \\
v_{130} &= 2a_{02}^{(0)} a_{00}^{(1)} - 2a_{02}^{(1)} a_{00}^{(0)} \\
v_{121} &= 2a_{12}^{(0)} a_{00}^{(1)} - 2a_{12}^{(1)} a_{00}^{(0)} \\
v_{112} &= 2a_{02}^{(0)} a_{11}^{(1)} - 2a_{02}^{(1)} a_{11}^{(0)} \\
v_{103} &= 2a_{12}^{(0)} a_{11}^{(1)} - 2a_{12}^{(1)} a_{11}^{(0)} \\
v_{200} &= a_{22}^{(0)} c^{(1)} - a_{22}^{(1)} c^{(0)} \\
v_{210} &= a_{22}^{(0)} b_0^{(1)} - a_{22}^{(1)} b_0^{(0)} \\
v_{201} &= a_{22}^{(0)} b_1^{(1)} - a_{22}^{(1)} b_1^{(0)} \\
v_{220} &= a_{22}^{(0)} a_{00}^{(1)} - a_{22}^{(1)} a_{00}^{(0)} \\
v_{211} &= 2a_{22}^{(0)} a_{01}^{(1)} - 2a_{22}^{(1)} a_{01}^{(0)} \\
v_{202} &= a_{22}^{(0)} a_{11}^{(1)} - a_{22}^{(1)} a_{11}^{(0)}
\end{aligned}
$$

Points of intersection are solutions to $R(x, y) = 0$. If there are solutions, then there must be at least one solution (x, y) that is closest to the origin. This problem can be set up as a constrained minimization: Minimize $|\mathbf{X}|^2$ subject to the constraint $R(\mathbf{X}) = 0$. Applying the method of Lagrange multipliers, define $F(\mathbf{X}, t) = |\mathbf{X}|^2 + tR(\mathbf{X})$. Setting $\partial F/\partial t = 0$ reproduces the constraint $R = 0$. Setting the spatial derivatives $\nabla F = \mathbf{0}$ yields $2(x, y) + t(R_x, R_y) = (0, 0)$ where $R_x = \partial R/\partial x$ and $R_y = \partial R/\partial y$. Therefore, $0 = S(x, y) =$

$yR_x - xR_y = \sum_{i+j \le 4} s_{ij} x^i y^j$, another polynomial of degree at most four. The coefficients are

$$
\begin{aligned}
s_{40} &= -r_{31} \\
s_{31} &= 4r_{40} - 2r_{22} \\
s_{22} &= 3r_{31} - 3r_{13} \\
s_{13} &= 2r_{22} - 4r_{04} \\
s_{04} &= r_{13} \\
s_{30} &= -r_{21} \\
s_{21} &= 3r_{30} - 2r_{12} \\
s_{12} &= 2r_{21} - 3r_{03} \\
s_{03} &= r_{12} \\
s_{20} &= -r_{11} \\
s_{11} &= 2r_{20} - 2r_{02} \\
s_{02} &= r_{11} \\
s_{10} &= -r_{01} \\
s_{01} &= r_{10} \\
s_{00} &= 0
\end{aligned}
$$

We now have two polynomial equations in two unknowns, $R(x, y) = 0$ and $S(x, y) = 0$.

Consider $f(y) = \sum_{i=0}^{4} \alpha_i y^i$ and $g(y) = \sum_{i=0}^{4} \beta_i y^i$. The Bézout matrix for f and g is the 4×4 matrix $M = [M_{ij}]$ with

$$
M_{ij} = \sum_{k=\max(4-j,4-i)}^{\min(4,7-i-j)} w_{k,7-i-j-k}
$$

for $0 \le i \le 3$ and $0 \le j \le 3$, with $w_{i,j} = \alpha_i \beta_j - \alpha_j \beta_i$ for $0 \le i \le 4$ and $0 \le j \le 4$. In expanded form,

$$
M = \begin{bmatrix}
w_{4,3} & w_{4,2} & w_{4,1} & w_{4,0} \\
w_{4,2} & w_{3,2} + w_{4,1} & w_{3,1} + w_{4,0} & w_{3,0} \\
w_{4,1} & w_{3,1} + w_{4,0} & w_{2,1} + w_{3,0} & w_{2,0} \\
w_{4,0} & w_{3,0} & w_{2,0} & w_{1,0}
\end{bmatrix}.
$$

Both $R(x, y) = 0$ and $S(x, y) = 0$ can be written as polynomials in y whose coefficients are polynomials in x. That is, the α_i and β_i are polynomials in x of degree $4 - i$. The degree of $w_{i,j}$ is $8 - i - j$. The Bézout determinant is $D(x) = \det(M(x))$, a polynomial of degree 16 in x.

The roots of $D(x) = 0$ are computed. For each root x, the coefficients of $f(y)$ are computed and the roots for the fourth degree polynomial equation $f(y) = 0$ are computed. The pairs (x, y) are tested to make sure $R(x, y) = 0$ and $S(x, y) = 0$. Once such a pair $(\bar{x}, \bar{y})$ is found, then the problem now is to traverse the curve of intersection, or verify that the initial point is isolated, in which case the two ellipsoids are tangent at that point. The point is isolated if $\nabla Q_0(\bar{x}, \bar{y})$ and $\nabla Q_1(\bar{x}, \bar{y})$ are parallel. A simple verification that the cross product of the gradient vectors is the zero vector will suffice. If the point is not isolated, then the curve can be traversed by solving a system of differential equations

$$\frac{dx}{dt} = R_y(x, y), \quad \frac{dy}{dt} = -R_x(x, y), \quad (x(0), y(0)) = (\bar{x}, \bar{y}).$$

The vector (R_x, R_y) is normal to the level curve defined by $R = 0$, so the vector $(R_y, -R_x)$ is tangent to the level curve. The differential equations just specify to traverse the curve by following the tangent vector.

2.2 Variation 2

The main problem with Variation 1 is that numerically finding the roots of a degree 16 polynomial is usually an ill-conditioned problem. An alternative is to set up a system of differential equations that allows you to walk along one ellipsoid in search of a point of intersection with the other ellipsoid. The search will either find a point or determine that there is none.

Start with a point $\mathbf{X}_0$ such that $Q_0(\mathbf{X}_0) = 0$. If $Q_1(\mathbf{X}_0) = 0$, you got lucky and have a point of intersection. If $Q_1(\mathbf{X}_0) < 0$, then $\mathbf{X}_0$ is inside the other ellipsoid. The idea is to walk tangent to the first ellipsoid while increasing the value of Q_1 to zero. In space, the direction of largest increase of Q_1 is ∇Q_1. However, this vector is usually not tangent to the first ellipsoid, so you need to project it onto the first ellipsoid's tangent space by projecting out the contribution by ∇Q_0. The path on the first ellipsoid with largest increase in Q_1 locally is determined by

$$\frac{d\mathbf{X}}{dt} = \nabla Q_1 - \frac{\nabla Q_1 \cdot \nabla Q_0}{|\nabla Q_0|^2} \nabla Q_0, \quad \mathbf{X}(0) = \mathbf{X}_0.$$

In the event that $Q_1(\mathbf{X}_0) > 0$, the tangent direction must be reversed so that Q_1 is decreased as rapidly as possible to 0. The differential equations for this case are

$$\frac{d\mathbf{X}}{dt} = -\nabla Q_1 + \frac{\nabla Q_1 \cdot \nabla Q_0}{|\nabla Q_0|^2} \nabla Q_0, \quad \mathbf{X}(0) = \mathbf{X}_0.$$

Whether or not the ellipsoids not intersect, eventually the traversal will lead to a point for which the gradients are parallel. In this case the right-hand side of the differential equation reduces to the zero vector. The length of the right-hand side vector can be used as a termination criterion in the numerical solver. Another concern is that the numerical solver will produce a new position from an old one. Because of numerical error, the new position might not be on the first ellipsoid. A correction can be made to adjust the new position so that it is on the first ellipsoid. The corrected value can be used to generate the next iterate in the solver.

Once a point $\mathbf{X}_1 = \mathbf{X}(T)$ is found for which $Q_1(\mathbf{X}_1) = 0$, the 2D level curve traverse mentioned in Variation 1 can be applied. However, it is possible to traverse the curve of intersection in 3D. A tangent vector for the curve is perpendicular to both ∇Q_0 and ∇Q_1. The system of equations to solve is

$$\frac{d\mathbf{X}}{dt} = \nabla Q_0 \times \nabla Q_1, \quad \mathbf{X}(0) = \mathbf{X}_1.$$

3 Test Intersection

3.1 Variation 1

All level curves defined by $Q_0(\mathbf{X}) = \lambda$ are ellipsoids, except for the minimum (negative) value λ for which the equation defines a single point, the center of every level curve ellipsoid. The ellipsoid defined by $Q_1(\mathbf{X}) = 0$ is a surface that generally intersects many level surfaces of Q_0. The problem is to find the minimum level value λ_0 and maximum level value λ_1 attained by any $\mathbf{X}$ on the ellipsoid E_1. If $\lambda_1 < 0$, then E_1 is properly contained in E_0. If $\lambda_0 > 0$, then E_0 and E_1 are separated. Otherwise, $0 \in [\lambda_0, \lambda_1]$ and the two ellipsoids intersect.

This can be formulated as a constrained minimization that can be solved by the method of Lagrange multipliers: Minimize $Q_0(\mathbf{X})$ subject to the constraint $Q_1(\mathbf{X}) = 0$. Define $F(\mathbf{X}, t) = Q_0(\mathbf{X}) + tQ_1(\mathbf{X})$. Differentiating yields $\nabla F = \nabla Q_0 + t\nabla Q_1$ where the gradient indicates the derivatives in $\mathbf{X}$. Also, $\partial F/\partial t = Q_1$. Setting the t-derivative equal to zero reproduces the constraint $Q_1 =$. Setting the $\mathbf{X}$-derivative equal to zero yields $\nabla Q_0 + t\nabla Q_1 = \mathbf{0}$ for some t. Geometrically this means that the gradients are parallel.

Note that $\nabla Q_i = 2A_i\mathbf{X} + \mathbf{B}_i$, so

$$\mathbf{0} = \nabla Q_0 + t\nabla Q_1 = 2(A_0 + tA_1)\mathbf{X} + (\mathbf{B}_0 + t\mathbf{B}_1).$$

Formally solving for $\mathbf{X}$ yields

$$\mathbf{X} = -(A_0 + tA_1)^{-1}(\mathbf{B}_0 + t\mathbf{B}_1)/2 = \frac{1}{\delta(t)}\mathbf{Y}(t)$$

where $\delta(t)$ is the determinant of $(A_0 + tA_1)$, a cubic polynomial in t, and $\mathbf{Y}(t)$ has components cubic in t. Replacing this in $Q_1(\mathbf{X}) = 0$ yields

$$\mathbf{Y}(t)^{\mathrm{T}}A_1\mathbf{Y}(t) + \delta(t)\mathbf{B}_1^{\mathrm{T}}\mathbf{Y}(t) + \delta(t)^2C_1 = 0,$$

a degree 6 polynomial in t. The roots can be computed, the corresponding values of $\mathbf{X}$ computed, and $Q_0(\mathbf{X})$ evaluated. The minimum and maximum values are stored as λ_0 and λ_1, and the earlier comparisons with zero are applied.

3.2 Variation 2

An iterative method can be set up that attempts to find a separating plane between the two ellipsoids. This does not directly handle proper containment of one ellipsoid by the other, but a similar algorithm can be derived for the containment case. Let the ellipsoids be in factored form, $(\mathbf{X} - \mathbf{C}_i)^{\mathrm{T}}M_i(\mathbf{X} - \mathbf{C}_i) = 1$ where M_i is positive definite and $\mathbf{C}_i$ is the center of the ellipsoid, $i = 0, 1$. A potential separating axis is $\mathbf{C}_0 + t\mathbf{N}$ where $\mathbf{N}$ is a unit length vector. The t-interval of projection of E_0 onto the axis is $I_0(\mathbf{N}) = [-r_0, r_0]$ where $r_0 = \sqrt{\mathbf{N}^{\mathrm{T}}M_0^{-1}\mathbf{N}}$. The t-interval of projection of E_1 onto the axis is $I_1(\mathbf{N}) = [\mathbf{N}\cdot\mathbf{\Delta} - r_1, \mathbf{N}\cdot\mathbf{\Delta} + r_1]$ where $\mathbf{\Delta} = \mathbf{C}_1 - \mathbf{C}_0$ and $r_1 = \sqrt{\mathbf{N}^{\mathrm{T}}M_1^{-1}\mathbf{N}}$.

Select an initial $\mathbf{N}$. If the intersection $F(\mathbf{N}) := I_0(\mathbf{N}) \cap I_1(\mathbf{N}) = \emptyset$, then the ellipsoids are separated. If $F(\mathbf{N}) \neq \emptyset$, then the given axis does not separate the ellipsoids. When the intervals overlap, $F(\mathbf{N}) = [f_0, f_1]$ where $f_0 = \max\{\mathbf{N}\cdot\mathbf{\Delta} - r_1, -r_0\}$ and $f_1 = \min\{\mathbf{N}\cdot\mathbf{\Delta} + r_1, r_0\}$. The function $D(\mathbf{N}) = f_1 - f_0 > 0$ when there is overlap. If the two intervals have a single point of intersection, then $f_0 = f_1$. If the intervals are disjoint, then $f_1 < f_0$ and $D(\mathbf{N}) < 0$. The problem now is to search the space of unit length vectors, starting at the initial $\mathbf{N}$, to determine if there is such a vector that makes $D < 0$. It is enough to determine if $D = 0$ and the graph of D has a transverse crossing at that location.