Information About Ellipses

David Eberly, Geometric Tools, Redmond WA 98052
https://www.geometrictools.com/

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Created: December 13, 2001
Last Modified: October 22, 2011

Contents

1 Discussion 2
1 Discussion


**Geometric Definition.** An ellipse is the set of points in a plane whose distances from two fixed points in that plane add to a constant. One of the fixed points is called a focal point of the ellipse. The two together are referred to as the foci of the ellipse.

**Standard Form.** Let the foci be $(\pm c, 0)$ where $c > 0$. Let $(x, y)$ be an ellipse point and let the sum of the distances from $(x, y)$ to the foci be denoted $2a$ for $a > 0$. The equation that $(x, y)$ must satisfy is

$$\sqrt{(x + c)^2 + y^2} + \sqrt{(x - c)^2 + y^2} = 2a.$$  

The points $(x, y)$, $(c, 0)$, and $(-c, 0)$ form a triangle. The sum of the lengths of two sides of a triangle must be larger than the length of the third side, so $2a > 2c$. Some algebraic manipulation of this equation leads to the standard form for an ellipse,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$  

(1)

where $b = \sqrt{a^2 - c^2}$. The argument of the square root is positive since earlier we argued that $a > c$. Moreover, $b < a$ is guaranteed since $b = \sqrt{a^2 - c^2} < \sqrt{a^2} = a$.

The center of the standard form ellipse is $(0, 0)$. The vertices are $(\pm a, 0)$. The major axis is the line segment that connects the vertices. The minor axis is the line segment with end points $(0, \pm b)$. The number $a$ is called the semimajor axis and the number $b$ is called the semiminor axis. [Note: I disagree with the use of the term “axis” to denote length.] The eccentricity is the ratio $c/a \in [0, 1]$ and is a measure of how stretched the ellipse is from a circle. A ratio of 0 occurs for a circle. A ratio nearly 1 indicates a long and narrow ellipse.

If the foci are chosen to be $(0, \pm c)$ and the sum of distances is $2b$, the standard form is also given by Equation (1), but now $b > c$ and $a = \sqrt{b^2 - c^2} < b$. The center is still $(0, 0)$, but the vertices are now $(0, \pm b)$, the major axis is the line segment connecting the vertices, the minor axis is the line segment with end points $(\pm b, 0)$, the semimajor axis is $b$, the semiminor axis is $a$, and the eccentricity is now defined as the ratio $c/b$.

If $a = b$, the foci are coincident with the origin $(0, 0)$ and the ellipse is really a circle. The concepts of major and minor axes do not apply here, but the eccentricity is 0.

**Area.** The area of an ellipse in standard form is

$$A = \pi ab.$$  

(2)

**Length.** The length of an ellipse is the total arc length of the curve. A closed form algebraic solution does not exist, but the length is given by an integral

$$L = 2 \int_{-a}^{a} \sqrt{1 + \frac{b^2 x^2}{a^2 (a^2 - x^2)}} \, dx = 2 \int_{-1}^{1} \sqrt{\frac{1 - (\lambda^2 - 1)t^2}{1 - t^2}} \, dt$$  

(3)

where $\lambda = b/a$. The integral can be approximated with a numerical integrator.
Center-Orient Form. An ellipse in the standard form given by Equation (1) can be oriented via a rotation so that the major and minor axes are not necessarily parallel to the coordinate axes. In vector/matrix form, the standard form is

\[
1 = \frac{x^2}{a^2} + \frac{y^2}{b^2} = \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} =: \mathbf{X}^T D \mathbf{X} \tag{4}
\]

where the last equality defines the $2 \times 1$ vector $\mathbf{X} = [x \ y]^T$, the $2 \times 2$ diagonal matrix $D = \text{Diag}(1/a^2, 1/b^2)$, and superscript $T$ denotes the transpose operation.

The ellipse may be rotated to a different orientation by a $2 \times 2$ rotation matrix

\[
R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}
\]

The major axis direction (1, 0) is rotated to $(\cos \theta, \sin \theta)$ and the minor axis direction (0, 1) is rotated to $(-\sin \theta, \cos \theta)$. The general transformation is $\mathbf{Y} = R \mathbf{X}$ with inverse $\mathbf{X} = R^T \mathbf{Y}$. Substituting this into Equation (4) leads to

\[
\mathbf{Y}^T R D R^T \mathbf{Y} = 1. \tag{5}
\]

After orientation the ellipse can be additionally translated so that its old center, the origin 0, is mapped to a new center $\mathbf{K}$. The general transformation is $\mathbf{Y} = R \mathbf{X} + \mathbf{K}$; the rotation $R$ is applied first, followed by the translation $\mathbf{K}$. Equation (5) is modified to include the translation,

\[
(\mathbf{Y} - \mathbf{K})^T R D R^T (\mathbf{Y} - \mathbf{K}) = 1. \tag{6}
\]

General Quadratic Form. When the Equation (6) is expanded and all terms are grouped on the left-hand side of the equation, the resulting polynomial has $x$, $y$, $x^2$, $xy$, and $y^2$ terms. The general quadratic equation for an ellipse is

\[
a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + b_1x + b_2y + c = 0 \tag{7}
\]

or in vector/matrix form,

\[
\mathbf{Y}^T A \mathbf{Y} + \mathbf{B}^T \mathbf{Y} + c = 0 \tag{8}
\]

where

\[
\mathbf{Y} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}, \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.
\]

All conic sections are represented by these equations. The ellipses are those for which $a_{11}a_{22} - a_{12}^2 > 0$. Observe that this condition states the determinant of $A$ is positive, so $A$ is an invertible matrix with inverse denoted by $A^{-1}$. The matrix $A$ and its inverse $A^{-1}$ are both symmetric matrices since $A^T = A$ and $A^{-T} = (A^T)^{-1} = A^{-1}$.

A typical problem is to start with the general quadratic form and convert to the center-orient form. This can be done by first completing the square on the equation. Consider that

\[
(\mathbf{Y} - \mathbf{K})^T A (\mathbf{Y} - \mathbf{K}) = \mathbf{Y}^T A \mathbf{Y} - 2 \mathbf{K}^T A \mathbf{Y} + \mathbf{K}^T A \mathbf{K}
\]

\[
= (\mathbf{Y}^T A \mathbf{Y} + \mathbf{B}^T \mathbf{Y} + c) - (2\mathbf{K}^T + \mathbf{B})^T \mathbf{Y} + (\mathbf{K}^T A \mathbf{K} - c)
\]

\[
= -(2\mathbf{A} + \mathbf{B})^T \mathbf{Y} + (\mathbf{K}^T A \mathbf{K} - c).
\]
If you set $K = -A^{-1}B/2$, then $K^TAK = B^TA^{-1}B/4$ and

$$(Y - K)^T A (Y - K) = B^TA^{-1}B/4 - c.$$  

Dividing by the scalar on the right-hand side of the last equation and setting $M = A/(B^TA^{-1}B/4 - c)$ produces

$$(Y - K)^T M (Y - K) = 1.$$  

Finally, $M$ can be factored using an eigendecomposition into $M = RDR^T$ where $R$ is a rotation matrix and $D$ is a diagonal matrix whose diagonal entries are positive. The final equation obtained by substituting the factorization for $M$ is exactly Equation (6).

For a $2 \times 2$ matrix, the eigendecomposition can be done symbolically. An eigenvector $V$ of $M$ corresponding to an eigenvalue $\lambda$ is a nonzero vector such that $M V = \lambda V$. The eigenvalues are solutions to the quadratic equation $\det(M - \lambda I) = 0$ where $I$ is the identity matrix. Since $M$ is a symmetric matrix, the eigenvalues must be real numbers. For each eigenvalue, a corresponding eigenvector $V$ is a nonzero solution to $(M - \lambda I) V = 0$. Let $M = [m_{ij}]$. The quadratic equation is

$$0 = \det(M - \lambda I) = \det \begin{bmatrix} m_{11} - \lambda & m_{12} \\ m_{12} & m_{22} - \lambda \end{bmatrix} = (m_{11} - \lambda)(m_{22} - \lambda) - m_{12}^2 = \lambda^2 - (m_{11} + m_{22})\lambda + (m_{11}m_{22} - m_{12}^2).$$

The roots are

$$\lambda = \frac{(m_{11} + m_{22}) \pm \sqrt{(m_{11} + m_{22})^2 - 4(m_{11}m_{22} - m_{12}^2)}}{2} = \frac{(m_{11} + m_{22}) \pm \sqrt{(m_{11} - m_{22})^2 + 4m_{12}^2}}{2}. \quad (9)$$

The argument of the square root is nonnegative, so the roots must be real-valued. The only way for the roots to be equal is if $m_{11} = m_{22}$ and $m_{12} = 0$, in which case $M$ must have been a scalar multiple of the identity matrix (the ellipse is really a circle). I assume for the remainder of the construction that the two eigenvalues are different.

Define $\lambda_1$ to be the eigenvalue in Equation (9) that uses the plus sign and define $\lambda_2$ to be the one that uses the minus sign. It is the case that $\lambda_1 > \lambda_2$. An eigenvector corresponding to $\lambda_1$ is perpendicular to one of the rows of the matrix

$$\begin{bmatrix} m_{11} - \lambda_1 & m_{12} \\ m_{12} & m_{22} - \lambda_1 \end{bmatrix} = \begin{bmatrix} \frac{(m_{11} - m_{22}) - P}{2} & m_{12} \\ m_{12} & \frac{(m_{11} - m_{22}) + P}{2} \end{bmatrix}$$

where $P = \sqrt{(m_{11} - m_{22})^2 + 4m_{12}^2} > 0$. We need to be certain that the selected row is not the zero vector. If $m_{12} \neq 0$, then either row will suffice. In a floating-point system, though, $m_{12}$ might be nearly zero. It is better to devise a selection scheme that does not suffer from numerical round-off errors. Specifically, if $m_{11} \geq m_{22}$, then

$$| - (m_{11} - m_{22}) - P | \geq |(m_{11} - m_{22}) - P|$$

The best choice is to use the second row to generate the eigenvector. If $m_{11} \leq m_{22}$, then

$$| - (m_{11} - m_{22}) - P | \leq |(m_{11} - m_{22}) - P|$$
and the best choice is to use the first row to generate the eigenvector. Let \( \mathbf{U}_1 = (\alpha, \beta) \) be a normalized vector that is perpendicular to the selected row. The eigenvector corresponding to \( \lambda_2 \) is chosen to be \( \mathbf{U}_2 = (-\beta, \alpha) \).

By definition of eigenvectors, \( M\mathbf{U}_1 = \lambda_1 \mathbf{U}_1 \) and \( M\mathbf{U}_2 = \lambda_2 \mathbf{U}_2 \). We can write the two equations jointly by using a matrix \( R = [\mathbf{U}_1 \; \mathbf{U}_2] \) whose columns are the unit-length eigenvectors. The columns are unit length and perpendicular to each other, so \( R \) is an orthogonal matrix. In fact, by the choice of \( \mathbf{U}_2 \), \( R \) happens to be a rotation matrix (no reflection component so to speak). The joint equation is \( MR = RD \) where \( D = \text{Diag}(\lambda_1, \lambda_2) \). Multiplying on the right by \( R^T \) leads to the decomposition \( M = RDR^T \).

In summary, for an ellipse specified as \( a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + b_1x + b_2y + c = 0 \), first verify that \( a_{11}a_{22} - a_{12}^2 > 0 \) so that you really do have an ellipse. Then

1. The center is
   \[
   \mathbf{K} = (k_1, k_2) = \frac{(a_{22}b_1 - a_{12}b_2, a_{11}b_2 - a_{12}b_1)}{2(a_{11}^2 - a_{12}a_{22})}.
   \]

2. Set \( \mu = 1/(\mathbf{K}^T \mathbf{A} \mathbf{K} - c) = 1/(a_{11}k_1^2 + 2a_{12}k_1k_2 + a_{22}k_2^2 - c) \) and define \( m_{11} = \mu a_{11} \), \( m_{12} = \mu a_{12} \), and \( m_{22} = \mu a_{22} \).

3. Set \( \lambda_1 = ((m_{11} + m_{22}) + \sqrt{(m_{11} - m_{22})^2 + 4m_{12}^2})/2 \). The semiminor axis of the ellipse is
   \[
   b = \frac{1}{\sqrt{\lambda_1}}.
   \]

Set \( \lambda_2 = ((m_{11} + m_{22}) - \sqrt{(m_{11} - m_{22})^2 + 4m_{12}^2})/2 \). The semimajor axis of the ellipse is
   \[
   a = \frac{1}{\sqrt{\lambda_2}}.
   \]

4. If \( m_{11} \geq m_{22} \), choose the minor axis direction of the ellipse to be
   \[
   \mathbf{U}_1 = \frac{(\lambda_1 - m_{22}, m_{12})}{|\lambda_1 - m_{22}, m_{12}|}.
   \]

If \( m_{11} < m_{22} \), choose the minor axis direction to be
   \[
   \mathbf{U}_1 = \frac{(m_{12}, \lambda_1 - m_{11})}{|m_{12}, \lambda_1 - m_{11}|}.
   \]

If \( \mathbf{U}_1 = (\alpha, \beta) \), choose the major axis direction to be \( \mathbf{U}_2 = (-\beta, \alpha) \).

5. If all you need is the angle formed by the major axis with the positive \( x \)-axis, that angle satisfies the equation
   \[
   \tan(2\theta) = -\frac{2a_{12}}{a_{22} - a_{11}}
   \]

This is obtained by making the change of variables \( x = \bar{x}\cos\theta - \bar{y}\sin\theta \) and \( y = \bar{x}\sin\theta + \bar{y}\cos\theta \) and substituting into the original quadratic equation. After expanding all terms, the coefficient of \( \bar{x}\bar{y} \) is
   \[
   -2a_{11}\sin\theta\cos\theta + 2a_{12}(\cos^2\theta - \sin^2\theta) + 2a_{22}\sin\theta\cos\theta = 2a_{12}\cos(2\theta) + (a_{22} - a_{11})\sin(2\theta)
   \]

Setting this coefficient to zero gives you an axis-aligned ellipse in the \((\bar{x}, \bar{y})\) coordinate system, so the angle \( \theta \) represents how much you must rotate the original ellipse to the axis-aligned one.
6. For $R = [U_1 \ U_2]$ where $U_1$ and $U_2$ are written as columns and $D = \text{Diag}(1/a^2, 1/b^2)$, the ellipse is represented by the factored form

$$(Y - K)^T R D R^T (Y - K) = (Y - K)^T \left( \frac{1}{a^2} U_1 U_1^T + \frac{1}{b^2} U_2 U_2^T \right) (Y - K) = 1. \quad (15)$$

7. Observe that $Y = K + RX = K + xU_1 + yU_2$. Replacing this in the factored form leads to $(x/a)^2 + (y/b)^2 = 1$, as expected since originally $Y$ was selected to be the coordinates representing the rotation and translation of the standard form ellipse with coordinates $X$.

8. The bounding rectangle for the ellipse that has the same directions as the major and minor axes of the ellipse has center $K$. The four corners are $K \pm aU_1 \pm bU_2$. 