Correlation Coefficient of Ellipse (Section 2)

Nino

June 4, 2020

My name is Nino (pen name). This article is an English version of the Japanese article I wrote at the following URL: http://hooktail.sub.jp/contributions/ellipse_cor_02v2.pdf

Introduction

In the previous article [1], I calculated the correlation coefficient of an ellipse using covariance and standard deviation. In this article, we use vector representations and linear combinations of trigonometric functions to derive the correlation coefficient of an ellipse.

Derivation of Correlation Coefficient of Ellipse using Dot Product of Two Vectors

In the previous article [1], when the ellipse with the major radius a and the minor radius b was rotated 45 degrees counterclockwise around the origin, the correlation coefficient r of the rotated ellipse was given by

$$r = (a^2 - b^2) / (a^2 + b^2) \tag{1}$$

We can rewrite equation (1) as

$$r = \frac{aa + b(-b)}{\sqrt{a^2 + b^2}\sqrt{a^2 + (-b)^2}} \tag{2}$$

Equation (2) shows that the numerator is the dot product of vector $\mathbf{P}(a, \mathrm{b})$ and vector $\mathbf{Q}(a, -b)$, and that the denominator is the product of the lengths of each vector.

An example of the ellipse and the two vectors is shown in Figure 1.

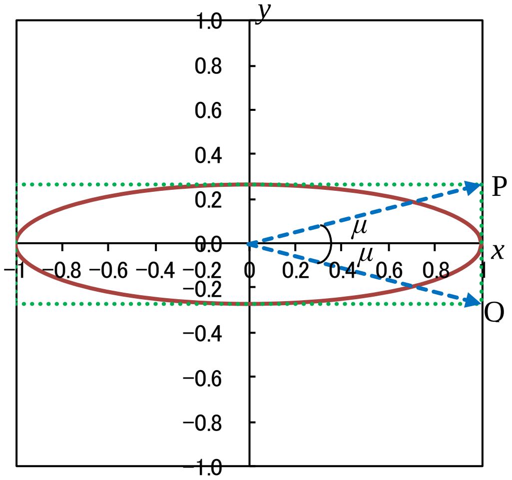


Figure.1 Scatterplot of the ellipse (a=1, b=0.268), vector **P** (1, 0.268) and vector **Q** (1, -0.268).
(2μ is the angle between **P** and **Q**, $\mu = \angle \mathrm{P}0x = \angle \mathrm{Q}0x$)

The dot product of two vectors **P** and **Q** is defined by

$$\cos 2\mu = \mathbf{P} \cdot \mathbf{Q} / \|\mathbf{P}\| \|\mathbf{Q}\| = (a^2 - b^2)/(a^2 + b^2) \quad (3)$$

Where:

P · **Q** is the dot product of two vectors **P** and **Q**

‖**P**‖ is the length of vector **P.**

‖**Q**‖ is the length of vector **Q.**

According to equation (3), $\cos 2\mu$ is equal to the correlation coefficient of an ellipse rotated 45 degrees counterclockwise around the origin [1]. Therefore, the correlation coefficient of an ellipse is determined by the major and minor radii of the ellipse or the angle between two vectors, regardless of whether the ellipse is rotated or not. For example, we put a=1 and b=0.268 in equation (3), we get $\cos 2\mu$ = 0.866 or 2μ = 30 degrees.

Derivation of Correlation Coefficient using Linear Combination of Trigonometric Functions

The standard parametric equation for an ellipse $x^2/a^2 + y^2/b^2 = 1$ is

$$x = a\cos\theta, \qquad y = b\sin\theta \quad (4)$$

The new coordinates(X,Y) of the point (x, y) after rotating counterclockwise around the origin by 45 degrees are given by, [2]

$$\begin{aligned} X &= 1/\sqrt{2}(a\cos\theta - b\sin\theta) \\ Y &= 1/\sqrt{2}(a\cos\theta + b\sin\theta) \end{aligned} \quad (5)$$

And a linear combination of sine and cosine with the same argument can be expressed as a single trigonometric function with an amplitude and phase [3].

$$\begin{aligned} a\cos\theta - b\sin\theta &= \sqrt{a^2 + b^2}\cos(\theta + \phi) \\ a\cos\theta + b\sin\theta &= \sqrt{a^2 + b^2}\cos(\theta - \phi) \end{aligned} \quad (6)$$

Where, ϕ is given by

$$\cos\phi = \frac{a}{\sqrt{a^2 + b^2}}, \qquad \sin\phi = \frac{b}{\sqrt{a^2 + b^2}} \quad (7)$$

Once we put equations (6) in equations (5), we get

$$\begin{aligned} X &= \sqrt{\frac{a^2 + b^2}{2}}\cos(\theta + \phi) \\ Y &= \sqrt{\frac{a^2 + b^2}{2}}\cos(\theta - \phi) \end{aligned} \quad (8)$$

The amplitudes of X and Y are the same value $\sqrt{(a^2+b^2)/2}$, and the phase difference between them is 2ϕ. We calculate the cosine of the phase difference 2ϕ using the cosine double angle formula.

$$\begin{aligned}\cos 2\phi &= 2\cos^2\phi - 1 \\ &= 2\left(\frac{a}{\sqrt{a^2+b^2}}\right)^2 - 1 \\ &= \frac{a^2-b^2}{a^2+b^2}\end{aligned} \qquad (9)$$

$\cos 2\phi$ is equal to the correlation coefficient r of the rotated ellipse. Therefore, it was found that the angle 2μ between the two vectors in Figure 1 is equal to the phase difference 2ϕ between X and Y having the same amplitude.
We examine the time series of x, y, X and Y (Figure.2).

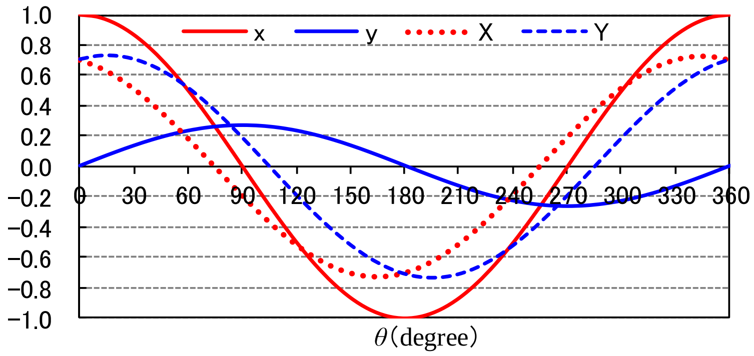


Figure.2 Time series of x, y, X and Y for the case of a=1 and b=0.268

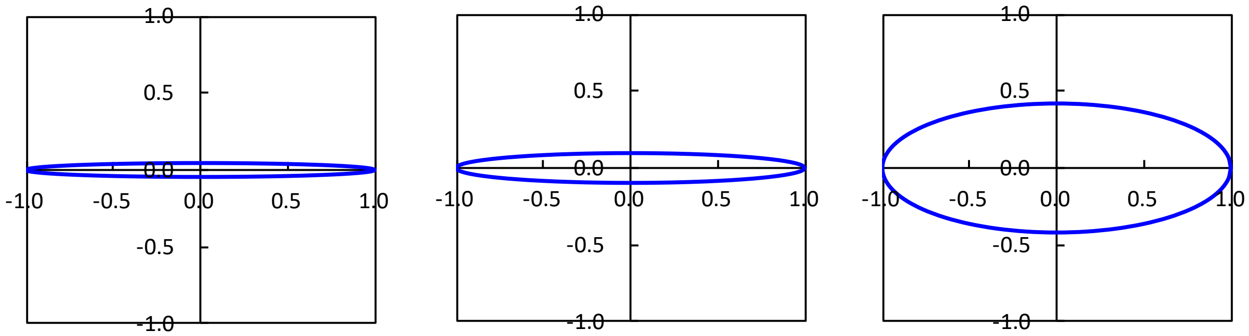
The amplitude of x is a=1 and the amplitude of y is b=0.268, and the phase difference between them is 90 degrees. The amplitudes of x and y are equal to the major and minor radii of the ellipse in Figure 1, respectively. On the other hand, the amplitudes of X and Y are both equal at 0.732 which is $\sqrt{(a^2+b^2)/2}$ in equations (8). The phase difference between them is 30 degrees. Using these phase differences, we got the correlation coefficients which are cos90°=0 and cos30°= 0.866 respectively [1].

Relationship between Correlation Coefficient and b/a Ratio

We can also rewrite equation (1) as

$$r = \frac{1-(b/a)^2}{1+(b/a)^2} \qquad (10)$$

The b/a ratio can be regarded as a fluctuation range in the y-axis direction.
Specific examples are shown in Figure 3.



(A) $b/a = 0.05$ $r = 0.995$, $2\mu = 2.9°$

(B) $b/a = 0.10$ $r = 0.980$, $2\mu = 5.7°$

(C) $b/a = 0.42$ $r = 0.700$, $2\mu = 22.8°$

Figure 3. Scatterplots of ellipses with the b/a ratios of 0.05, 0.10, and 0.42, and their correlation coefficients r and angles 2μ

In general, if the correlation coefficient is 0.7 or higher, it is evaluated as having a strong positive relationship. However, Figure 3 shows that the b/a ratio with a correlation coefficient of 0.700 is 4.5 times (＝0.42／0.10） larger than the b/a ratio with a correlation coefficient of 0.980. Thus, it was shown that there is a large difference even in the range where the correlation coefficient is 0.7 or more.

There are a few things to keep in mind when applying the results obtained so far to the measured data. They are, for example, whether the shape of the distribution is close to an ellipse and whether it is affected by outliers and errors.

REFERENCES

1. Correlation Coefficient of Ellipse (Section 1)
http://hooktail.sub.jp/contributions/ellipse_cor_01_en.pdf
2. Wikipedia, Rotation matrix
https://en.wikipedia.org/wiki/Rotation_matrix
3. Transformations of Trig Functions
http://www2.me.rochester.edu/~clark/ME163Web/webexamp/trigtran.pdf