Probability Ellipse

1. Two-dimensional Normal Distribution and the Probability Ellipse

1.1 Fundamental statistics

$$x_m = \frac{\sum x_i}{n} \qquad \sigma_x = \sqrt{\frac{1}{n-1} \cdot \sum (x_i - x_m)^2} \tag{1}$$

$$y_m = \frac{\sum y_i}{n} \qquad \sigma_y = \sqrt{\frac{1}{n-1} \cdot \sum (y_i - y_m)^2} \tag{2}$$

$$\rho_{xy} = \frac{n \sum x_i y_i - \sum x_i \cdot \sum y_i}{\sqrt{[n \sum x_i{}^2 - (\sum x_i)^2] \cdot [n \sum y_i{}^2 - (\sum y_i)^2]}} \tag{3}$$

1.2 Probability density function

$$\left\{ \begin{array}{l} f(x,y) = \dfrac{1}{2\pi \cdot \sigma_x \cdot \sigma_y \cdot \sqrt{1-\rho^2}} \cdot \exp\left(-\dfrac{D^2}{2}\right) \\[4mm] D^2 = \dfrac{1}{1-\rho_{xy}^2} \left\{ \dfrac{(x-x_m)^2}{\sigma_x{}^2} + \dfrac{(y-y_m)^2}{\sigma_y{}^2} - \dfrac{2\rho_{xy}(x-x_m)(y-y_m)}{\sigma_x \sigma_y} \right\} \\[4mm] p = 1 - \exp\left(-\dfrac{D^2}{2}\right) \end{array} \right. \tag{4}$$

$$
\begin{array}{llll}
x_m,\, y_m & : \text{Average} & \rho_{xy} & : \text{Correlation coefficient} \\
\sigma_x,\, \sigma_y & : \text{Standard deviation} & p & : \text{Non-exceedance probability of plots} \\
D & : \text{Mahalabinos distance} & &
\end{array}
$$

■Relationship between D and p

○ Squared Mahalanobis distance D^2 is distributed according to the χ^2 (chi-squared) distribution.
○ In the two-dimensional probability distribution, it is distributed according to the χ^2 distribution withe two degrees of freedom.
○ The χ^2 cumulative distribution function F with degrees of freedom of k is expressed as follow:

$$F(x;k) = \frac{\gamma(k/2, x/2)}{\Gamma(k/2)}$$

where, $\gamma(s,t)$ is the lower incomplete Gamma function and $\Gamma(s)$ is the Gamma function.
○ In a special case of $k = 2$, this function has a simple form as follow:

$$F(x;2) = 1 - \exp(-x/2)$$

○ As a result, the following relationship between Mahalanobis distance D and non-exceedance probability p can be obtained.

$$p = 1 - \exp(-D^2/2) \quad (F \to p, x \to D^2)$$

1.3 Draw the probability ellipse

The procedure to draw the probability ellipse is shown below:

○ First, estimate the average (x_m, y_m), standard deviation (σ_x, σ_y) and the correlation coefficient (ρ_{xy}) using above relationship.
○ Next, define the probability p for the probability ellipse.
○ Finally, draw the probability ellipse using following equations.

$$
\begin{cases}
x = x_m + \sigma_x \cdot r \cdot \cos\phi \\
y = y_m + \sigma_y \cdot r \cdot \sin\phi \\
\\
r = \sqrt{\dfrac{-2(1 - \rho_{xy}^2) \cdot \ln(1 - p)}{1 - 2\rho_{xy} \cdot \sin\phi \cdot \cos\phi}} \\
\\
0 \leqq \phi \leqq 2\pi \qquad\qquad\qquad\quad : \text{Range of } \phi
\end{cases}
\tag{5}
$$

For only drawing the probability ellipse, we can make plots easily by obtaining the several coordinates on the probability ellipse.

But, when we want to use the Se option in GMT to draw the probability ellipse, we have to know the shortest radius, the longest radius and rotated angle from the horizontal axis.

After this, the techniques to obtain the required values are shown.

> (Notice) The arguments for Se option in GMT are actually x_m, y_m, θ, $2a$ and $2b$, where x_m and y_m are the coordinates of a center of an ellipse, θ is a rotated angle in degrees counter-clockwise from horizontal, $2a$ is the major axis (twice of the longest radius) and $2b$ is the minor axis (twice of the shortest radius).

2. Formula of Rotated Ellipse

2.1 General formula

General formula of a ellipse with the center of origin, the major axis of a and the minor axis of b is shown below.

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1 \tag{6}$$

The major axis of above ellipse is parallel to X-axis.

When the ellipse is rotated by the angle of θ in the direction of counter-clockwise from horizontal, a relationship between the coordinate (X, Y) before rotation and the coordinate (x, y) after rotation becomes following.

$$\begin{cases} X = x\cos\theta + y\sin\theta \\ Y = y\cos\theta - x\sin\theta \end{cases} \tag{7}$$

The following formula is obtained by substituting above relation to the general formula of ellipse.

$$\frac{a^2\sin^2\theta + b^2\cos^2\theta}{a^2b^2} \cdot x^2 - 2\sin\theta\cos\theta\frac{a^2 - b^2}{a^2b^2} \cdot xy + \frac{a^2\cos^2\theta + b^2\sin^2\theta}{a^2b^2} \cdot y^2 = 1 \tag{8}$$

$$A \cdot x^2 + B \cdot xy + C \cdot y^2 = 1 \quad \begin{cases} A = \dfrac{a^2\sin^2\theta + b^2\cos^2\theta}{a^2b^2} \\[2mm] B = -2\sin\theta\cos\theta\dfrac{a^2 - b^2}{a^2b^2} \\[2mm] C = \dfrac{a^2\cos^2\theta + b^2\sin^2\theta}{a^2b^2} \end{cases} \tag{9}$$

The upper expression is a formula of a ellipse after rotation.

2.2 Formula of Rotated Ellipse passing through given three coordinates

Set given coordinates which are located on the rotated ellipse as (x_1, y_1), (x_2, y_2) and (x_3, y_3). Using these coordinates, coefficients of A, B and C can be obtained as follow.

$$\begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 \\ x_2^2 & x_2y_2 & y_2^2 \\ x_3^2 & x_3y_3 & y_3^2 \end{bmatrix} \begin{Bmatrix} A \\ B \\ C \end{Bmatrix} = \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} \quad \Rightarrow \quad \begin{Bmatrix} A \\ B \\ C \end{Bmatrix} = \begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 \\ x_2^2 & x_2y_2 & y_2^2 \\ x_3^2 & x_3y_3 & y_3^2 \end{bmatrix}^{-1} \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} \tag{10}$$

2.3 Major axis, minor axis and rotated angle of a ellipse

Find the major axis, minor axis and rotated angle, where the major axis is twice of the longest radius and the minor axis is also twice of the shortest radius.

Those values can be obtained by solving following non-linear simultaneous equations with three variables.

$$
\begin{cases}
f(a,b,\theta) = a^2 \sin^2\theta + b^2 \cos^2\theta - a^2 b^2 A & = 0 \\[2mm]
g(a,b,\theta) = -2\sin\theta\cos\theta(a^2 - b^2) - a^2 b^2 B & = 0 \\[2mm]
h(a,b,\theta) = a^2 \cos^2\theta + b^2 \sin^2\theta - a^2 b^2 C & = 0
\end{cases}
\tag{11}
$$

In order to obtain the values of a, b and θ which satisfy the conditions of $f = 0$, $g = 0$ and $h = 0$, following iterative calculation is carried out. When the values of Δa, Δb and $\Delta \theta$ become smaller than specified error, the iterative calculation can be finished.

$$
\begin{cases}
a_{i+1} = a_i + \Delta a \\
b_{i+1} = b_i + \Delta b \\
\theta_{i+1} = \theta_i + \Delta\theta
\end{cases}
\tag{12}
$$

$$
\begin{bmatrix} \frac{\partial f}{\partial a} & \frac{\partial f}{\partial b} & \frac{\partial f}{\partial \theta} \\ \frac{\partial g}{\partial a} & \frac{\partial g}{\partial b} & \frac{\partial g}{\partial \theta} \\ \frac{\partial h}{\partial a} & \frac{\partial h}{\partial b} & \frac{\partial h}{\partial \theta} \end{bmatrix} \begin{Bmatrix} \Delta a \\ \Delta b \\ \Delta\theta \end{Bmatrix} = \begin{Bmatrix} -f \\ -g \\ -h \end{Bmatrix} \quad \Rightarrow \quad \begin{Bmatrix} \Delta a \\ \Delta b \\ \Delta\theta \end{Bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial a} & \frac{\partial f}{\partial b} & \frac{\partial f}{\partial \theta} \\ \frac{\partial g}{\partial a} & \frac{\partial g}{\partial b} & \frac{\partial g}{\partial \theta} \\ \frac{\partial h}{\partial a} & \frac{\partial h}{\partial b} & \frac{\partial h}{\partial \theta} \end{bmatrix}^{-1} \begin{Bmatrix} -f \\ -g \\ -h \end{Bmatrix}
\tag{13}
$$

$$\frac{\partial f}{\partial a} = 2a(\sin^2\theta - b^2 A) \qquad \frac{\partial f}{\partial b} = 2b(\cos^2\theta - a^2 A) \qquad \frac{\partial f}{\partial \theta} = 2\sin\theta\cos\theta(a^2 - b^2)$$

$$\frac{\partial g}{\partial a} = -2a(2\sin\theta\cos\theta + b^2 B) \qquad \frac{\partial g}{\partial b} = 2b(2\sin\theta\cos\theta - a^2 B) \qquad \frac{\partial g}{\partial \theta} = -2(\cos^2\theta - \sin^2\theta)(a^2 - b^2)$$

$$\frac{\partial h}{\partial a} = 2a(\cos^2\theta - b^2 C) \qquad \frac{\partial h}{\partial b} = 2b(\sin^2\theta - a^2 C) \qquad \frac{\partial h}{\partial \theta} = -2\sin\theta\cos\theta(a^2 - b^2)$$

3. Programing

3.1 f90_SREG.f90

This is a program to obtain the fundamental statistics and basic characteristics of the probability ellipse from original plot data. Non-exceedance probability for finding the characteristics of probability ellipse is inputted as a command line argument.

And this program outputs the coordinates of center and the coordinates of selected three points which are on the curve of probability ellipse with the center of the origin (0,0). As these three points, the vertex of the major axis, the vertex of the minor axis and minimum x-value and corresponding y-value are selected.

The characteristics of ellipse which is obtained from this program can not use normally in GMT drawing, because they are special values in case of the same scale of x-axis and y-axis.

3.2 f90_ELLIP.f90

This is a program to create the arguments for 'Se' option use in GMT. 'Se' is an option of the command of 'psxy' in GMT to draw a ellipse without many plot values.

In GMT, the size of graph is specified in unit of 'cm'. When we want to draw a ellipse using 'Se' option in GMT, it is required to input the coordinates of center, rotated angle of the ellipse, major axis and minor axis. These input values for 'Se' option in GMT shall be corresponded to the graph size and scale of graph.

Therefore, I made this program for GMT drawing using 'Se' option.

The function of this program is as follow.

○ Input filename and the values for scale arrangement shall be inputted as command line arguments. The values for scale arrangement include the size of x and y axis on the graph in unit of 'cm', length of x and y-axis (maximum value minus minimum value).

○ The coordinates of center of ellipse and the coordinates of three points on the ellipse with the center of the origin (0,0) are inputted from the specified input file.

○ The scale arrangement is carried out for inputted three points taking into consideration of the graph size and range of coordinates.
○ The coefficients for ellipse formula is calculated from the arranged coordinates of three points.
○ The calculation for the rotated angle, major axis and minor axis is carried out taking into consideration of actual size of graph.
○ The center coordinates, rotated angle, major axis and minor axis are outputted for GMT drawing.

## 4. Example of GMT drawing of a probability ellipse

Two drawings with the same size and same plot values are shown below. The difference between these graphs is only the scale or range of y-axis.

In the graph of left side, the rotated angle and the length of major and minor axis are about 46 degrees, 3.6cm and 1.8cm respectively.

Whereas, in the graph of right side, the rotated angle and the length of major and minor axis are about 25 degrees, 3.0cm and 1.3cm respectively.

In actual drawing job, the ellipse line by 'Se' option is superimposed on the plots obtained by one degree interval calculation.

The degree of coincidence between two drawing methods is fine, and it means the expected result can be obtained.
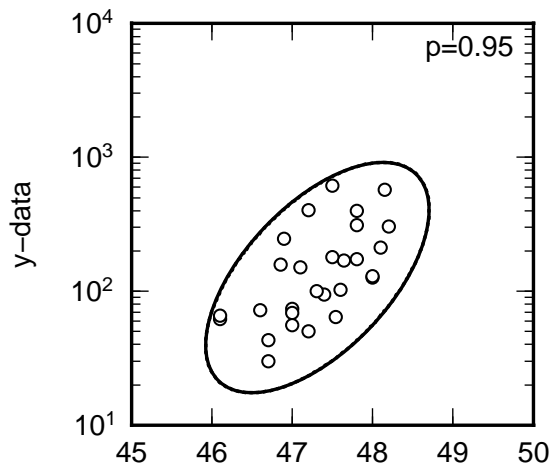
Length of x-axis (cm)	5.0
Length of y-axis (cm)	5.0
x-coordinate of center	0.4731310E+02
y-coordinate of center	0.2102538E+01
Direction (degree)	0.4648623E+02
Major axis (cm)	0.3563558E+01
Minor axis (cm)	0.1803970E+01
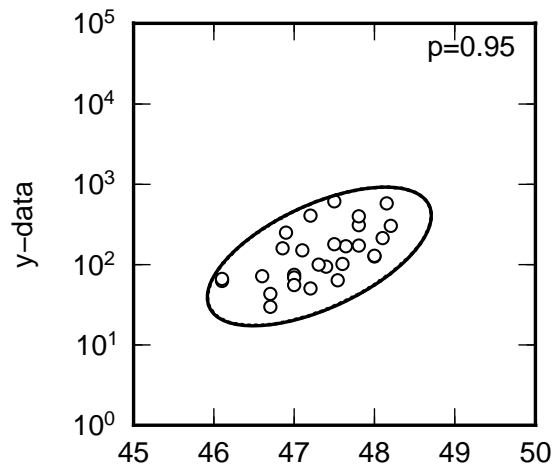
Length of x-axis (cm)	5.0
Length of y-axis (cm)	5.0
x-coordinate of center	0.4731310E+02
y-coordinate of center	0.2102538E+01
Direction (degree)	0.2493061E+02
Major axis (cm)	0.3007828E+01
Minor axis (cm)	0.1282364E+01