Error Estimation in Fitting of Ellipses

Kirk T. McDonald
Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544
(July 17, 2014; updated August 1, 2014)

This note gives a prescription for fitting a set of m points, $\{x_j, y_j\}$, (perhaps from digitization of an image) to an ellipse, with the general form (with 5 parameters a_i),

$$a_1 x^2 + a_2 xy + a_3 y^2 + a_4 x + a_5 y - 1 = 0 \qquad (a_1,\ a_3 > 0,\ a_2^2 < 4a_1 a_3). \tag{1}$$

In addition, we give an estimate of the errors on the best-fit values of the parameters a_i.

1 Errors on the Parameters of the Quadratic Form (1)

We first define the auxiliary data set $\{z_{ij}\}$, $i = 1, 5, j = 1, m$,

$$z_{i,j} = (x_j^2, x_j y_j, y_j^2, x_j, y_j). \tag{2}$$

Among many possible measures of the goodness of fit of the data $\{z_{ij}\}$ to the ellipse (1), we adopt the simplest,[1] writing,

$$\chi^2 = \sum_{j=1}^{m} \frac{(\sum_{i=1}^{5} a_i z_{ij} - 1)^2}{\sigma_j^2}\,, \tag{3}$$

where σ_j is the measurement uncertainty associated with the data point (x_j, y_j). The best-fit parameters $\hat{a}_i$ are those that minimize the function χ^2 for a set of measurements $\{z_{ij}\}$.

We consider the case that the σ_j are not known, but can be assumed to have the common value σ. Then, by supposing that the function χ^2 is actually a chi-square [3, 4, 5, 6, 7, 8] with $m - 5$ degrees of freedom, the best-fit (minimum) χ^2 has most probable value $m - 5$. Assuming (naïvely) that the best-fit χ^2 has this value, the unknown σ is determined, and error estimates for the best-fit parameters $\hat{a}_i$ follow via standard procedures.[2]

A great insight is that $\exp(-\chi^2/2)$ can be thought of another way. It is also the (unnormalized) probability distribution that the polynomial coefficients have values a_i when their best-fit values are $\hat{a}_i$ with uncertainties due to the measurements $\{x_j, y_j\}$. Expressing this in symbols,

$$\exp(-\chi^2/2) = \text{const} \times \exp\left(-\sum_{k=1}^{5}\sum_{l=1}^{5} \frac{(a_k - \hat{a}_k)(a_l - \hat{a}_l)}{2\sigma_{kl}^2}\right), \tag{4}$$

or equivalently,

$$\chi^2 = \text{const} + \sum_{k=1}^{5}\sum_{l=1}^{5} \frac{(a_k - \hat{a}_k)(a_l - \hat{a}_l)}{\sigma_{kl}^2}. \tag{5}$$

[1] For a survey of 13 measures of goodness of fit, see [1]. For discussion of the measure (3), see [2].
[2] For a discussion of this approach for polynomial fitting, see Appendix D of [9].

The uncertainty on $\hat{a}_k$ is σ_{kk} in this notation. In eqs. (4) and (5) we have introduced the important concept that the uncertainties in the coefficients $\hat{a}_k$ are correlated. That is, the quantity σ_{kl}^2 is a measure of the probability that the values of $\hat{a}_k$ and $\hat{a}_l$ both have positive fluctuations at the same time. In fact, σ_{kl}^2 can be negative indicating that when $\hat{a}_k$ has a positive fluctuation then $\hat{a}_l$ has a correlated negative one.

One way to see the merit of minimizing the χ^2 is as follows. According to eq. (5) the derivative of χ^2 with respect to a_k is,

$$\frac{\partial \chi^2}{\partial a_k} = \sum_{l=1}^{5} \frac{\hat{a}_l - a_l}{\sigma_{kl}^2}, \tag{6}$$

so that all first derivatives of χ^2 vanish when all $a_l = \hat{a}_l$. That is, χ^2 is a minimum when the coefficients a_i take on their best-fit values $\hat{a}_i$. A further benefit is obtained from the second derivatives,

$$\frac{\partial^2 \chi^2}{\partial a_k \partial a_l} = \frac{1}{\sigma_{kl}^2}. \tag{7}$$

For our particular χ^2 (3), with $\sigma_j = \sigma$, the first derivatives are,

$$\frac{\partial \chi^2}{\partial a_k} = \sum_{j=1}^{m} \frac{z_{kj}(\sum_{i=1}^{5} a_i z_{ij} - 1)}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^{5} \sum_{j=1}^{m} a_i z_{ij} z_{kj} - \frac{1}{\sigma^2} \sum_{j=1}^{m} z_{kj}, \tag{8}$$

and the second derivatives are,

$$\frac{\partial^2 \chi^2}{\partial a_k \partial a_l} = \frac{1}{\sigma^2} \sum_{j=1}^{m} z_{kj} z_{lj} \equiv \frac{M_{kl}}{\sigma^2}. \tag{9}$$

Using the matrix M_{kl} introduced in eq. (9), the condition that the first derivatives (8) vanish at the best-fit coefficients $\hat{a}_k$ can be written as,

$$\sum_{i=5}^{5} M_{ik} \hat{a}_i = \sum_{j=1}^{m} z_{kj} \equiv V_k. \tag{10}$$

We then calculate the inverse matrix M^{-1} and apply it to find the best-fit coefficients $\hat{a}_k$ (which do not depend on the as-yet-unknown value of σ),

$$\hat{a}_k = \sum_{l=1}^{5} M_{kl}^{-1} V_l. \tag{11}$$

Comparing eqs. (7) and (9) we have,

$$\frac{1}{\sigma_{kl}^2} = \frac{M_{kl}}{\sigma^2}, \qquad i.e., \qquad \sigma_{kl}^2 = \frac{\sigma^2}{M_{kl}}. \tag{12}$$

The uncertainty in best-fit coefficient $\hat{a}_i$ is then reported as,

$$\sigma_{\hat{a}_i} = \sigma_{ii} = \frac{\sigma}{\sqrt{M_{ii}}}. \tag{13}$$

All that remains is to find the value of the unknown uncertainty $\sigma = \sigma_j$ on the measurements. For this, we set the χ^2 for the best-fit parameters $\hat{a}_i$ equal to the number of degrees of freedom, $m - 5$,

$$\chi^2(\hat{a}_i) = m - 5 = \sum_{j=1}^{m} \frac{(\sum_{i=1}^{5} \hat{a}_i z_{ij} - 1)^2}{\sigma^2}\,, \tag{14}$$

such that σ is determined to be,

$$\sigma = \sqrt{\frac{\sum_{j=1}^{m}(\sum_{i=1}^{5} \hat{a}_i z_{ij} - 1)^2}{m-5}}\,, \qquad \text{and} \qquad \sigma_{\hat{a}_i} = \sigma_{ii} = \sqrt{\frac{\sum_{j=1}^{m}(\sum_{k=1}^{5} \hat{a}_k z_{kj} - 1)^2}{(m-5)\sum_{j=1}^{m} z_{ij}^2}}\,. \tag{15}$$

2 Errors on the Conventional Ellipse Parameters

An alternative description of the ellipse of eq. (1) is that it has semimajor axis of length a which makes angle θ to the x-axis, semiminor axis of length b, and center at (x_0, y_0). The shape parameters a, b and θ depend only on a_1, a_2 and a_3, while the center of the ellipse depends on all five of the a_i. We now deduce the alternative parameters, and their fit errors, in terms of the a_i and the errors on the latter as found in sec. 1.

We first translate the coordinates according to $x' = x - x_0$ and $y' = y - y_0$ such that the resulting parameters a'_i of the quadratic form have,

$$a'_1 = a_1, \quad a'_2 = a_2, \quad a'_3 = a_3, \quad a'_4 = 2a_1x_0 + a_2y_0 + a_4, \quad a'_5 = a_2x_0 + 2a_3y_0 + a_5. \tag{16}$$

For the ellipse to be centered at $x' = 0 = y'$ we need $a'_4 = 0 = a'_5$, which leads to,

$$x_0 = \frac{a_2a_5 - 2a_3a_4}{4a_1a_3 - a_2^2}\,, \qquad y_0 = \frac{a_2a_4 - 2a_1a_5}{4a_1a_3 - a_2^2}\,. \tag{17}$$

As a check, we note that if $a_4 = 0 = a_5$ then the original ellipse was centered on the origin, and indeed eq. (17) implies that $x_0 = 0 = y_0$.

To deduce the error on, say, x_0 we first consider the differential,

$$dx_0 = \frac{a_5da_2 + a_2da_5 - 2a_4da_3 - 2a_3da_4 - x_0(4a_3da_1 + 4a_ada_3 - 2a_2da_2)}{4a_1a_3 - a_2^2}\,. \tag{18}$$

Then, on squaring this we can identify $(dx_0)^2$ with the squared error $\sigma_{x_0}^2$ when we identify the products $da_i\,da_j$ with the σ_{ij}^2 found in eq. (12).

The determine the shape parameters a, b and θ we perform a coordinate rotation by angle θ with respect to the x'-axis[3] (which is parallel to the x-axis),

$$x'' = x'\cos\theta + y'\sin\theta, \qquad x' = x''\cos\theta - y''\sin\theta, \tag{19}$$
$$y'' = -x'\sin\theta + y'\cos\theta, \qquad y' = x''\sin\theta + y''\cos\theta, \tag{20}$$

[3]We could also make the rotation directly from the (x, y) coordinates, with no affect on the shape parameters as these don't depend on a_4 and a_5. However, if parameters x_0 and y_0 are deduced only after this rotation, they appear to depend on θ, which complicates the expressions for their errors.

and require that $a''_2 = 0$, in which case $a''_1 = 1/a^2$ and $a''_3 = 1/b^2$. This leads to,[4]

$$\tan 2\theta = \frac{a_2}{a_1 - a_3}\,, \quad \cos 2\theta = \frac{a_1 - a_3}{\sqrt{a_2^2 + (a_1 - a_3)^2}}\,, \quad \sin 2\theta = \frac{a_2}{\sqrt{a_2^2 + (a_1 - a_3)^2}}\,, \tag{21}$$

$$\sigma_\theta = \frac{\cos^2 2\theta}{2\,|a_1 - a_3|}\sqrt{\tan^2 2\theta(\sigma_{11}^2 + \sigma_{33}^2 - 2\sigma_{13}^2) + \sigma_{22}^2 - 2\tan 2\theta(\sigma_{12}^2 - \sigma_{23}^2)}, \tag{22}$$

and,

$$\frac{1}{a^2} = a_1 \cos^2\theta + a_2 \sin\theta\cos\theta + a_3 \sin^2\theta = \frac{a_1 + a_3 + (a_1 - a_3)\cos 2\theta + a_2 \sin 2\theta}{2}$$
$$= \frac{a_1 + a_3 + \sqrt{a_2^2 + (a_1 - a_3)^2}}{2}\,, \tag{23}$$
$$\frac{1}{b^2} = a_1 \sin^2\theta - a_2 \sin\theta\cos\theta + a_3 \cos^2\theta = \frac{a_1 + a_3 - (a_1 - a_3)\cos 2\theta - a_2 \sin 2\theta}{2}$$
$$= \frac{a_1 + a_3 - \sqrt{a_2^2 + (a_1 - a_3)^2}}{2}\,. \tag{24}$$

Note that $1/b^2 \le 1/a^2$, which means that b is the semimajor axis, and a is the semiminor axis. Note also that $\tan 2\theta = \tan 2(\theta - \pi/2)$, so there is an ambiguity in eq. (21) as to whether θ is the angle to the semimajor or the semiminor axis.

As a measure of the departure of the ellipse from a circle we introduce the ellipticity (flattening) ϵ,[5]

$$\epsilon \equiv \frac{b}{a} \ge 1\,, \qquad \epsilon^2 = \frac{a_1 + a_3 + \sqrt{a_2^2 + (a_1 - a_3)^2}}{a_1 + a_3 - \sqrt{a_2^2 + (a_1 - a_3)^2}} \equiv \frac{C_+}{C_-}\,. \tag{25}$$

Taking the differential, we have,

$$2\epsilon\, d\epsilon = \frac{dC_+ - \epsilon^2\, dC_-}{C_-} \equiv \frac{C_1\, da_1 + C_2\, da_2 + C_3\, da_3}{C_-}\,, \tag{26}$$

where,

$$C_\pm = a_1 + a_3 \pm \sqrt{a_2^2 + (a_1 - a_3)^2} = a_1 + a_3 \pm S, \qquad S \equiv \sqrt{a_2^2 + (a_1 - a_3)^2}, \tag{27}$$
$$C_1 = 1 - \epsilon^2 + \frac{a_1 - a_3}{S}(1 + \epsilon^2) = -\frac{2[a_2^2 - 2a_3(a_1 - a_3)]}{SC_-} \equiv -2\frac{D_1}{SC_-}\,, \tag{28}$$
$$C_2 = \frac{a_2}{S}(1 + \epsilon^2) = \frac{2a_2(a_1 + a_3)}{SC_-} \equiv 2\frac{D_2}{SC_-}\,, \tag{29}$$
$$C_3 = 1 - \epsilon^2 - \frac{a_1 - a_3}{S}(1 + \epsilon^2) = -\frac{2[a_2^2 + 2a_1(a_1 - a_3)]}{SC_-} \equiv -2\frac{D_3}{SC_-}\,. \tag{30}$$

---

[4] As $z_{2j}^2 = x_j^2 y_j^2 = z_{1j} z_{3j}$, the definition (9) implies that $M_{13} = M_{22} = M_{31}$. Furthermore, $x^2 + y^2 \ge 2xy$, so that $x^4 + y^4 \ge 2x^2y^2$, and hence $M_{11} + M_{33} \ge 2M_{13}$. Then, it can be that $\sigma_{11}^2 + \sigma_{33}^2 = \sigma^2/M_{11} + \sigma^2/M_{33} < 2\sigma^2/M_{13} = 2\sigma_{13}^2$, and the argument of the square root in eq. (22) can be negative (as verified in numerical examples). Similarly, the argument of the square root in eq. (31) can be negative. This suggests that the matrix σ_{kl}^2 of eq. (16) may not be a good representation of the correlation in the uncertainties on the five parameters $\hat{a}_i$.

In numerical calculations, one could use the absolute value of the arguments of the square roots.

[5] If we define $\epsilon' = (b - a)/a = \epsilon - 1$, so that $\epsilon' = 0$ for a circle, then $\sigma_{\epsilon'} = \sigma_\epsilon$.

Then, the error σ_ϵ on the ellipticity ϵ is given by,

$$\sigma_\epsilon = \frac{1}{\epsilon S C_-^2} \sqrt{D_1^2 \sigma_{11}^2 + D_2^2 \sigma_{22}^2 + D_3^2 \sigma_{33}^2 - 2D_1 D_2 \sigma_{12}^2 - 2D_2 D_3 \sigma_{23}^2 + 2D_1 D_3 \sigma_{13}^2} \,. \tag{31}$$

Of course, the "error" computed this way assumes that the fit is "good", which might not be the case. The user should make a separate judgment as to whether the fit is indeed "good" before taking seriously the error estimates presented here (which in any case are subject to the doubt raised in footnote 4).

References

[1] P.L. Rosin, *Analysing error of fit functions for ellipses*, Patt. Rec. Lett. **17** 1461 (1996),
http://physics.princeton.edu/~mcdonald/mumu/target/Rosin/rosin_ellipse_fit_functions.pdf
Assessing Error of Fit Functions for Ellipses, Comp. Graph. Info. Proc. **58**, 494 (1996),
http://physics.princeton.edu/~mcdonald/examples/statistics/rosin_gmip_58_494_96.pdf

[2] A. Albano, *Representation of Digitized Contours in Terms of Conic Arcs and Straight-Line Segments*, Comp. Graph. Info. Proc. **3**, 23 (1974),
http://physics.princeton.edu/~mcdonald/examples/statistics/albano_cgip_3_23_74.pdf

[3] K. Pearson, *On the Criterion that a given System of Deviations from the Probable in the case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from random sampling*, Phil. Mag. **1**, 157 (1900),
http://physics.princeton.edu/~mcdonald/examples/statistics/pearson_pm_1_157_00.pdf

[4] R.A. Fisher, *On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P*, J. Roy. Stat. Soc. **85**, 87 (1922),
http://physics.princeton.edu/~mcdonald/examples/statistics/fisher_jrss_85_87_22.pdf

[5] R.A. Fisher, *The Conditions under which χ^2 Measures the Discrepancy between Observation and Hypothesis*, J. Roy. Stat. Soc. **87**, 422 (1924),
http://physics.princeton.edu/~mcdonald/examples/statistics/fisher_jrss_87_442_24.pdf

[6] W.G. Cochran, *The χ^2 Test of Goodness of Fit*, Ann. Math. Soc. **23**, 315 (1952),
http://physics.princeton.edu/~mcdonald/examples/statistics/cochran_ams_23_315_52.pdf

[7] R.L. Plackett, *Karl Pearson and the Chi-squared Test*, Intl. Stat. Rev. **51**, 59 (1983),
http://physics.princeton.edu/~mcdonald/examples/statistics/plackett_isr_51_59_83.pdf

[8] S.D. Bolboacă *et al.*, *Pearson-Fisher Chi-Square Statistic Revisited*, Information **2**, 528 (2011), http://physics.princeton.edu/~mcdonald/examples/statistics/bolboaca_info_2_528_11.pdf

[9] K.T. McDonald, *Princeton Physics 101 Lab Manual* (2013),
http://physics.princeton.edu/~mcdonald/examples/ph101_2013/Labs/PHY101_2013_Lab_Manual.pdf