Frequently Asked Questions

1. What do you mean by “Principle of least squares?

Answer:
The “Principle of Least Squares” states that the most probable values of a system of unknown quantities upon which observations have been made are obtained by making the sum of the squares of the errors a minimum. This statement however, is indefinite in that the term “error” is not rigorously defined; and further, the ordinary proof of the principle is defective from the fact that the same indefinite nomenclature, and loose reasoning founded thereon, are used throughout.

2. Explain different categories of least squares.

Answer:
Least squares problems fall into two categories: linear or ordinary least squares and non-linear least squares, depending on whether or not the residuals are linear in all unknowns.
The linear least-squares problem occurs in statistical regression analysis; it has a closed-form solution.
A closed-form solution (or closed-form expression) is any formula that can be evaluated in a finite number of standard operations.
The non-linear problem has no closed-form solution and is usually solved by iterative refinement; at each iteration the system is approximated by a linear one.

3. Who described the basis for least square method?

Answer:
The fundamental basis for least-squares method was first described by Carl Friedrich Gauss around 1794 at the age of eighteen.

4. Who developed the criterion of minimum error while finding the solution in the method of least squares?

Answer:
The development of a criterion that can be evaluated to determine when the solution with the minimum error has been achieved, developed by Laplace in his *Method of Least Squares*.

5. Give a brief history of least squares.

Answer:
The fundamental basis for least-squares method was first described by Carl Friedrich Gauss around 1794 at the age of eighteen.
The method of least squares grew out of the fields of astronomy and geodesy as scientists and mathematicians sought to provide solutions to the challenges of navigating the Earth's oceans during the Age of Exploration.
The method was the culmination of several advances that took place during the course of the eighteenth century :
The combination of different observations taken under the *same* conditions contrary to simply trying one's best to observe and record a single observation accurately. This approach was notably used by Tobias Mayer while studying the librations of the moon.
The combination of different observations as being the best estimate of the true value; errors decrease with aggregation rather than increase, perhaps first expressed by Roger Cotes.

The combination of different observations taken under *different* conditions as notably performed by Roger Joseph Boscovich in his work on the shape of the earth and Pierre-Simon Laplace in his work in explaining the differences in motion of Jupiter and Saturn.
The development of a criterion that can be evaluated to determine when the solution with the minimum error has been achieved, developed by Laplace in his *Method of Least Squares*.
Gauss did not publish the method until 1809, when it appeared in volume two of his work on celestial mechanics, *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium*. In 1822, Gauss was able to state that the least-squares approach to regression analysis is optimal in the sense that in a linear model where the errors have a mean of zero, are uncorrelated, and have equal variances, the best linear unbiased estimator of the coefficients is the least-squares estimator. This result is known as the Gauss–Markov theorem.
The idea of least-squares analysis was also independently formulated by the Frenchman Adrien-Marie Legendre in 1805 and the American Robert Adrain in 1808.

6. Explain fundamental principle of least square.

Answer:
In order to obtain most probable values (MPVs), the sum of squares of the residuals must be minimized. (See book for derivation.) In the weighted case, the weighted squares of the residuals must be minimized.
That is $\sum v^2 = v_1^2 + v_2^2 + v_3^2 + \cdots + v_n^2 = \text{minimum}$

7. When do we have stochastic model?

Answer:
- The covariances (including variances) and hence the weights as well, form the *stochastic model*
- Even an “unweighted” adjustment assumes that all observations have equal weight which is also a stochastic model

8. Explain the mathematical model.

Answer:
- The mathematical model is a set of one or more equations that define an adjustment condition
- Models also include collinearity equations in photogrammetry and the equation of a line in linear regression
- It is important that the model properly represents reality – for example the angles of a plane triangle should total 180^0, but if the triangle is large, spherical excess cause a systematic error so a more elaborate model is needed.

9. Explain different types of models.

Answer: There are two types of models namely conditional and parametric.
- A conditional model enforces geometric conditions on the measurements and their residuals
- A parametric model expresses equations in terms of unknowns that were not directly measured, but relate to the measurements (e.g. a distance expressed by coordinate inverse) Parametric models are more commonly used because it can be difficult to express all of the conditions in a complicated measurement network

10. What do you mean by observation equations?

Answer:

- Observation equations are written for the parametric model
- One equation is written for each observation
- The equation is generally expressed as a function of unknown variables (such as coordinates) equals a measurement plus a residual
- We want more measurements than unknowns which gives a redundant adjustment

11. How do you find the least square solution for a given problem having an independent variable?

Answer:

The objective of adjusting the parameter o a model function to best fit a data set. Let us consider a simple problem, where a data set consist of n points (x_i, y_i), i = 1, 2, . . . n, where x_i is an independent variable and y_i is a dependent variable whose value is found by observation. The model function has the form f(x,β), where the m adjustable parameters are held in the vector β. The goal is to find the parameter values for the model which best fits the data. The least squares method finds its optimum when the sum, S of squared residuals

$$S = \sum_{i=1}^{n} r_i^2$$

is minimum where, $r_i=y_i-f(x_i,\beta)$, the difference between the actual value of the dependent variable and the value predicted by the model.

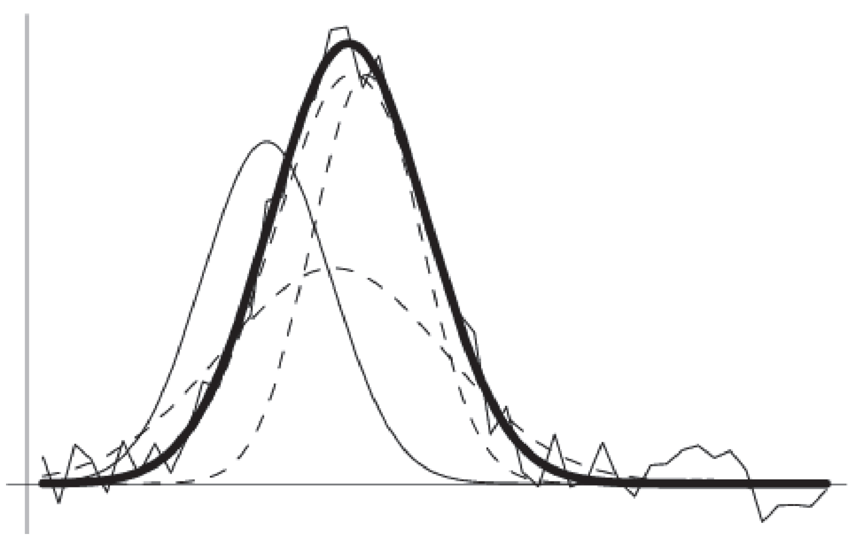
The example of a model is that of a straight line. Denoting the intercepts as β_0 and the slope as β_1, the model function is given by $f(x,\beta)=\beta_0+\beta_1 x$.

The data point may consist of more than one independent variable.

12. Give an example for non linear least square fit.

Answer:

An example of a nonlinear least square is fit to a noisy Gaussian function $f(x; A, x_0, \sigma) = Ae^{-(x-x_0)^2/(2\sigma^2)}$ is shown below.



Here the thin solid curve is the initial guess, the dotted curves are intermediate iterations, and the heavy solid curve is the fit to which the solution converges.
The actual parameters are (A, x_0, σ) = (1, 20, 5) , the initial guess was (0.8, 15, 4), and the converged values are (1.03105, 20.1369, 4.86022), with R^2 = 0.1485 . The partial derivatives used to construct the matrix A are

$$\frac{df}{dA} = e^{-(x-x_0)^2/(2\sigma^2)}$$

$$\frac{df}{dx_0} = \frac{A(x-x_0)}{\sigma^2} e^{-(x-x_0)^2/(2\sigma^2)}$$

$$\frac{df}{d\sigma} = \frac{A(x-x_0)^2}{\sigma^3} e^{-(x-x_0)^2/(2\sigma^2)}$$

13. Determine the least squares solution for the following:

$$F(x,y) = x + y - 2y^2 = -4$$
$$G(x,y) = x^2 + y^2 = 8$$
$$H(x,y) = 3x^2 - y^2 = 7.7$$

Use x_0 = 2, and y_0 = 2 for initial approximations.

Answer:
Take partial derivatives and form the Jacobian matrix.

$$\frac{\partial F}{\partial x} = 1 \qquad \frac{\partial G}{\partial x} = 2x \qquad \frac{\partial H}{\partial x} = 6x$$
$$\frac{\partial F}{\partial y} = 1-4y \qquad \frac{\partial G}{\partial y} = 2y \qquad \frac{\partial H}{\partial y} = -2y$$

$$J = \begin{bmatrix} 1 & 1-4y_0 \\ 2x_0 & 2y_0 \\ 6x_0 & -2y_0 \end{bmatrix} = \begin{bmatrix} 1 & -7 \\ 4 & 4 \\ 12 & -4 \end{bmatrix}$$

Form K matrix and set up least squares solution.

$$K = \begin{bmatrix} -4-F(x_0,y_0) \\ 8-G(x_0,y_0) \\ 7.7-H(x_0,y_0) \end{bmatrix} = \begin{bmatrix} -4-(-4) \\ 8-8 \\ 7.7-8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -0.3 \end{bmatrix}$$

$$J^T J = \begin{bmatrix} 1 & 4 & 12 \\ -7 & 4 & -4 \end{bmatrix} \begin{bmatrix} 1 & -7 \\ 4 & 4 \\ 12 & -4 \end{bmatrix} = \begin{bmatrix} 161 & -39 \\ -39 & 81 \end{bmatrix}$$

$$J^T K = \begin{bmatrix} 1 & 4 & 12 \\ -7 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -0.3 \end{bmatrix} = \begin{bmatrix} -3.6 \\ 1.2 \end{bmatrix}$$

$$X = \begin{bmatrix} 161 & -39 \\ -39 & 81 \end{bmatrix}^{-1} \begin{bmatrix} -3.6 \\ 1.2 \end{bmatrix} = \begin{bmatrix} -0.2125 \\ 0.00458 \end{bmatrix}$$

Add the corrections to get new approximations and repeat.

x_0 = 2.00 – 0.02125 = 1.97875 y_0 = 2.00 + 0.00458 = 2.00458

$$X = \begin{bmatrix} 157.61806 & -38.75082 \\ -38.75082 & 81.40354 \end{bmatrix}^{-1} \begin{bmatrix} -0.12393 \\ 0.75219 \end{bmatrix} = \begin{bmatrix} 0.00168 \\ 0.01004 \end{bmatrix}$$

Add the new corrections to get better approximations.

x_0 = 1.97875 + 0.00168 = 1.98043 y_0 = 2.00458 + 0.01004 = 2.01462

Further iterations give negligible corrections so the final solution is:

x = 1.98 y = 2.01

14. Fit a least square line to the following data. Also find trend values and show that $\sum\left(Y-\hat{Y}\right)=0$

X	1	2	3	4	5
Y	2	5	3	8	7

Answer:

X	Y	XY	X^2	$\hat{Y}=1.1+1.3X$	$Y-\hat{Y}$
1	2	2	1	2.4	-0.4
2	5	10	4	3.7	+1.3
3	3	9	9	5.0	-2
4	8	32	16	6.3	1.7
5	7	35	25	7.6	-0.6
$\sum X=15$	$\sum Y=25$	$\sum XY=88$	$\sum X^2=55$	**Trend Values**	$\sum\left(Y-\hat{Y}\right)=0$

The equation of least square line $Y=a+bX$

Normal equation for a, $\sum Y = na + b\sum X$ $25=5a+15b$

Normal equation for b, $\sum XY = a\sum X + b\sum X^2$ $88=15a+55b$

Eliminate a from both equation and solve for b and substitute value of b to get a.

Here a=1.1 and b=1.3. hence the equation of least square line becomes

Y=1.1+1.3x

The 5th column in the table gives the trend values.

15. Differentiate between linear and nonlinear least squares.

Answer:

- The model function, *f*, in LLSQ (linear least squares) is a linear combination of parameters of the form $f = X_{i1}\beta_1 + X_{i2}\beta_2 + \cdots$ The model may represent a straight line, a parabola or any other linear combination of functions. In NLLSQ (non-linear least squares) the parameters appear as functions, such as $\beta^2, e^{\beta x}$ and so forth. If the derivatives $\partial f/\partial \beta_j$ are either constant or depend only on the values of the independent variable, the model is linear in the parameters. Otherwise the model is non-linear.

- Algorithms for finding the solution to a NLLSQ problem require initial values for the parameters, LLSQ does not.
- Like LLSQ, solution algorithms for NLLSQ often require that the Jacobian be calculated. Analytical expressions for the partial derivatives can be complicated. If analytical expressions are impossible to obtain either the partial derivatives must be calculated by numerical approximation or an estimate must be made of the Jacobian.
- In NLLSQ non-convergence (failure of the algorithm to find a minimum) is a common phenomenon whereas the LLSQ is globally concave so non-convergence is not an issue.
- NLLSQ is usually an iterative process. The iterative process has to be terminated when a convergence criterion is satisfied. LLSQ solutions can be computed using direct methods, although problems with large numbers of parameters are typically solved with iterative methods, such as the Gauss–Seidel method.
- In LLSQ the solution is unique, but in NLLSQ there may be multiple minima in the sum of squares.
- Under the condition that the errors are uncorrelated with the predictor variables, LLSQ yields unbiased estimates, but even under that condition NLLSQ estimates are generally biased.