



Math for the people, by the people.

least squares

Canonical name	LeastSquares
Date of creation	2013-03-22 12:05:47
Last modified on	2013-03-22 12:05:47
Owner	akrowne (2)
Last modified by	akrowne (2)
Numerical id	8
Author	akrowne (2)
Entry type	Definition
Classification	msc 15-00
Synonym	least squares problem
Synonym	least-squares
Synonym	least-squares problem

The general problem to be solved by the least squares method is this: given some direct measurements y of random variables, and knowing a set of equations f which have to be satisfied by these measurements (possibly involving unknown parameters x), find the set of x which comes closest to satisfying

$$f(x, y) = 0$$

where “closest” is defined by a Δy such that

$$f(x, y + \Delta y) = 0 \text{ and } \Delta y^2 \text{ is minimized}$$

The sum of squares of elements of a vector can be written in different ways

$$\Delta y^2 = \Delta y^T \Delta y = \|\Delta y\|_2^2 = \sum_i \Delta y_i^2$$

The assumption has been made here that the elements of y are statistically uncorrelated and have equal variance. For this case, the above solution results in the most efficient estimators for x , Δy . If the y are correlated, correlations and variances are defined by a covariance matrix C , and the above minimum condition becomes

$$\Delta y^T C^{-1} \Delta y \text{ is minimized}$$

Least squares solutions can be more or less simple, depending on the constraint equations f . If there is exactly one equation for each measurement, and the functions f are linear in the elements of y and x , the solution is discussed under linear regression. For other linear models, see Linear Least Squares. Least squares methods applied to few parameters can lend themselves to very efficient algorithms (e.g. in real-time processing), as they reduce to simple matrix operations.

If the constraint equations are non-linear, one typically solves by linearization and in iterations, using approximate values of x , Δy in every step, and linearizing by forming the matrix of derivatives, df/dx (the Jacobian matrix) and possibly also df/dy at the last point of approximation.

Note that as the iterative improvements $\delta x, \delta y$ tend towards zero (if the process converges), Δy converges towards a final value which enters the minimum equation above.

Algorithms avoiding the explicit calculation of df/dx and df/dy have also been investigated, e.g. [?]; for a discussion, see [?]. Where convergence (or control over convergence) is problematic, use of a general package for minimization may be indicated.

References

- [1] M.L. Ralston and R.I. Jennrich, *Dud, a Derivative-free Algorithm for Non-linear Least Squares*, *Technometrics* 20-1 (1978) 7.
- [2] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C*, Second edition, Cambridge University Press, 1995.

*Note: This entry is based on content from the <http://rkb.home.cern.ch/rkb/titleA.html> *Data Analysis Briefbook**