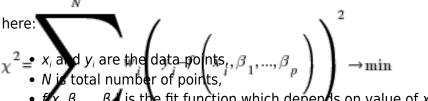# Table of Contents

Fitting Algorithm and Computational Formulas

MagicPlot uses iterative W Levenberg–Marquardt W nonlinear least squares curve fitting algorithm which is widely used in most software.

MagicPlot implementation of Levenberg–Marquardt algorithm is optimised for using with multi-core processors. MagicPlot successfully passed testing with NIST Nonlinear Regression datasets (see our report).

Nonlinear Least Squares Algorithm

Fit procedure iteratively varies the parameters β_k of fit function $f(x, \beta_1, ..., \beta_p)$ to minimize the residual sum of squares (RSS, χ^2):

here:

$$\chi^2 = \sum_i^N \left(w_i \left(y_i - f(x_i, \beta_1, ..., \beta_p) \right) \right)^2 \rightarrow \min$$

- x_i and y_i are the data points,
- N is total number of points,
- $f(x, \beta_1,...,\beta_p)$ is the fit function which depends on value of x and fit parameters β_k,
- p is the number of fit parameters β_k,
- w_i are normalized ($\Sigma w_i = 1$) data weighting coefficients for each point (x_i, y_i).

An initial guess for the parameters has to be provided to start minimization. Calculation of the new guess of parameters on each fit iteration is based on the fit function partial derivatives for current values of fit parameters for each x value:

$$\frac{\partial f}{\partial \beta_k}\left(x_i, \beta_1, ..., \beta_p \right)$$

Partial derivatives are computed using explicit formulas (for some predefined fit functions) or with finite difference (for custom equations).

Weighting of Data Points Using Y Errors

MagicPlot can use weighting of y values based on y errors s_i:

- If standard y errors are **not** specified: all $w_i=1$
- If standard y errors s_i are specified:

$$w_i = C \frac{1}{s_i^2}$$

here C is normalizing coefficient (to make the sum of w_i be equal to N):

$$C = N \sum_{i=1}^N \frac{1}{s_i^2}$$

In Fit Plot Properties dialog (Plot Data tab) you can set one of the following methods to evaluate standard y errors s_i:

- Get y errors from table column(s),
- Percent of data for every point,
- Fixed value or Standard deviation — do not use in weighting because in this case the error values are the same for all data points.

Iterations Stop Criteria

After each iteration except the first MagicPlot evaluates *deviation decrement D*:

$$D = \left| \frac{\chi^2_{curr.\,iter.}}{\chi^2_{prev.\,iter.}} - 1 \right|$$

Deviation decrement shows how the residual sum of squares (RSS) on current iteration relatively differs from that on the previous iteration.

The iterative fit procedure stops on one of two conditions:

- If the deviation decrements *D* for two last iterations is less than minimum allowable deviation decrement, which is 10^{-9} by default
- If the number of iterations exceeds maximum number of iterations, which is 100 by default

You can change the minimum allowable deviation decrement and maximum number of iterations in `Fitting` tab of `MagicPlot Preferences`.

Formulas Used in Fit Report

In the table below you can find the formulas which MagicPlot uses to calculate fit parameters and values in `Fit Report` tab.

Because of some confusion in the names of the parameters in different sources (books and software), we also give many different names of same parameter in *note* column.

Parameter Name	Symbol	Formula	Note
Original Data and Fit Model Properties			
Number of used data points	N	—	This is the number of data points inside specified fit Interval.
Fit parameters	$\beta_1,...,\beta_p$	—	For peak-like functions (Gauss, Lorentz) these parameters are amplitude, position and half width at half maximum. Only parameters with unset `Lock` checkbox are taken into account.

Parameter Name	Symbol	Formula	Note
Original Data and Fit Model Properties			
Number of fit function parameters β	p	—	This is the total number of unlocked parameters of fit curves which are summarized to fit.
W Degrees of freedom	ν	$N-p$	
Estimated mean of data	$\overline{y}$	$\dfrac{1}{N}\sum\limits_{i=1}^{N} y_i$	
Estimated variance of data	s^2	$\dfrac{1}{N-1}\sum\limits_{i=1}^{N}\left(y_i-\overline{y}\right)^2$	Not used by fit algorithm, only for comparison.
Data total sum of squares (TSS, SST)	TSS	$\sum\limits_{i=1}^{N} w_i\left(y_i-\overline{y}\right)^2$	**Also known as:** • Sum of squares about the mean.
Fit Result			
Residual sum of squares (RSS)	χ^2	$\sum\limits_{i=1}^{N} w_i\left(y_i-f\left(x_i,\beta_1,...,\beta_p\right)\right)^2$	This value is minimized during the fit to find the optimal fit function parameters. **Also known as:** • 'Chi-squared' • Sum of squared residuals (SSR), • Error sum of squares (ESS), • Sum of squares due to error (SSE).
Reduced χ^2	$\chi^2_{red.}$	$\dfrac{\chi^2}{\nu}=\dfrac{\chi^2}{N-p}$	The advantage of the reduced chi-squared is that it already normalizes for the number of data points and model (fit function) complexity. **Also known as:** • Mean square error (MSE), • Residual mean square.
Residual standard deviation	s	$\sqrt{\left(\chi^2_{red.}\right)}$	**Also known as:** • Root mean square of the error (Root MSE)

Parameter Name	Symbol	Formula	Note
Original Data and Fit Model Properties			
w Coefficient of determination	R^2	$1 - \dfrac{\chi^2}{TSS}$	R^2 will be equal to one if fit is perfect, and to zero otherwise. This is a biased estimate of the population R^2, and will never decrease if additional fit parameters (fit curves) are added, even if they are irrelevant.
Adjusted R^2	$\overline{R}^2$	$1 - \dfrac{N-1}{N-p-1}\left(1 - R^2\right)$	Adjusted R^2 (or *degrees of freedom adjusted R-square*) is a slightly modified version of R^2, designed to penalize for the excess number of fit parameters which do not add to the explanatory power of the regression. This statistic is always smaller than R^2, can decrease as you add new fit curves or introduce parameters, and even be negative for poorly fitting models.
Covariance matrix of parameters β_k	Σ	$\displaystyle\sum_{i=1}^{N} w_i \frac{\partial f}{\partial \beta_m}\left(x_i, \beta_1, ..., \beta_p\right)\frac{\partial f_m}{\partial \beta_n}\left(x_i, \beta_1, ..., \beta_p\right)$ $\dfrac{\chi^2}{N-p}\left[\alpha\right]^{-1}$, $where \left[\alpha\right] =$	Here α is the matrix of partial derivatives of fit function with respect to parameters β_m and β_n which is also used by fitting algorithm to compute parameters for next iteration.

Parameter Name	Symbol	Formula	Note
Original Data and Fit Model Properties			
Standard deviation of parameters β_k (std. dev.)	s_k	$\sqrt{\left[\Sigma\right]_{k,k}}$	These values are displayed in `Std. Dev.` column in parameters table.
Correlation matrix of parameters β_k	C	$\left[C\right]_{i,j} = \dfrac{\left[\Sigma\right]_{i,j}}{s_i s_j}$	This matrix shows if the parameters are linked. The values lie in range -1...1, diagonal elements are always 1. If two parameters are linked the corresponding matrix value will be close to 1. It means that changing the first parameter compensates changing of the second one so that the fitting algorithm cannot select between them.

See Also

- [Nonlinear Curve Fitting: Fit Plot](#)
- [Specifying Custom Fit Equation](#)
- [Using Spline for Baseline Subtraction](#)
- [Joining the Parameters of Fit Curves](#)
- [Guessing Peaks](#)
- [Predefined Fit Curves Equations](#)
- [Calculating Integrals and Statistics on Intervals using Fit Plot](#)