

[Events \(http://www.truman.edu/event/\)](http://www.truman.edu/event/)

[Directory \(http://www.truman.edu/directory/\)](http://www.truman.edu/directory/)

[TruView Login \(http://www.truman.edu/truview-login/\)](http://www.truman.edu/truview-login/)

[Search](#)

<http://www.truman.edu>

[About \(http://www.truman.edu/about/\)](http://www.truman.edu/about/)

[Majors & Programs \(http://www.truman.edu/majors-programs/\)](http://www.truman.edu/majors-programs/)

[Admission & Cost \(http://www.truman.edu/admission-cost/\)](http://www.truman.edu/admission-cost/)

[Student Life \(http://www.truman.edu/student-life/\)](http://www.truman.edu/student-life/)

[Alumni & Donors \(http://www.truman.edu/alumni-donors/\)](http://www.truman.edu/alumni-donors/)

[Athletics \(http://www.trumanbulldogs.com\)](http://www.trumanbulldogs.com)

Advanced Regression with Microsoft Excel

[Data Analysis \(https://chemlab.truman.edu/data-analysis/\)](https://chemlab.truman.edu/data-analysis/) / Advanced Regression with Microsoft Excel

Truman ChemLab

[\(https://chemlab.truman.edu/\)](https://chemlab.truman.edu/)

General Information

The Laboratory

Notebook

[\(https://chemlab.truman.edu/the-laboratory-notebook/\)](https://chemlab.truman.edu/the-laboratory-notebook/)

Before Laboratory

[\(https://chemlab.truman.edu/before-laboratory/\)](https://chemlab.truman.edu/before-laboratory/)

Laboratory Safety

[\(https://chemlab.truman.edu/laboratory-safety/\)](https://chemlab.truman.edu/laboratory-safety/)

SDS and MSDS

[\(https://chemlab.truman.edu/sds-and-msds/\)](https://chemlab.truman.edu/sds-and-msds/)

Glassware and

Equipment

[\(https://chemlab.truman.edu/glassware-and-equipment/\)](https://chemlab.truman.edu/glassware-and-equipment/)

Instrumentation

[\(https://chemlab.truman.edu/instrumentation/\)](https://chemlab.truman.edu/instrumentation/)

Data Analysis

[\(https://chemlab.truman.edu/data-analysis/\)](https://chemlab.truman.edu/data-analysis/)

[The Laboratory Report](https://chemlab.truman.edu/the-laboratory-report/)
(<https://chemlab.truman.edu/the-laboratory-report/>).

[Assessment](https://chemlab.truman.edu/assessment/)
(<https://chemlab.truman.edu/assessment/>).

Course Specific Information

[Chemistry for Contemporary Living](https://chemlab.truman.edu/chem100/)
(<https://chemlab.truman.edu/chem100/>).

[Chemical Principles](https://chemlab.truman.edu/chemical-principles/)
(<https://chemlab.truman.edu/chemical-principles/>).

[Inorganic Chemistry](https://chemlab.truman.edu/inorganic/)
(<https://chemlab.truman.edu/inorganic/>).

[Instrumental Analysis](https://chemlab.truman.edu/chem322manual/)
(<https://chemlab.truman.edu/chem322manual/>).

[Organic Chemistry](https://chemlab.truman.edu/organic-chemistry/)
(<https://chemlab.truman.edu/organic-chemistry/>).

[Physical Chemistry](https://chemlab.truman.edu/physical-chemistry/)
(<https://chemlab.truman.edu/physical-chemistry/>).

[Quantitative Analysis](https://chemlab.truman.edu/quantitative-analysis/)
(<https://chemlab.truman.edu/quantitative-analysis/>).

[Undergraduate Research](https://www.truman.edu/majors-programs/majors-minors/chemistry-major/chemistry-major-research-opportunities/)
(<https://www.truman.edu/majors-programs/majors-minors/chemistry-major/chemistry-major-research-opportunities/>).

[Chemistry of Art \(JINS\)](https://chemlab.truman.edu/chemistry-of-art-jins/)
(<https://chemlab.truman.edu/chemistry-of-art-jins/>).

Help

[Chemistry Contact Center](https://chemlab.truman.edu/cccl/)
(<https://chemlab.truman.edu/cccl/>).

[AXE Tutoring](http://axe.truman.edu/tutoring.htm)
(<http://axe.truman.edu/tutoring.htm>).

[Installing LoggerPro](https://wp-internal.truman.edu/chemistry/installing-loggerpro/)
(<https://wp-internal.truman.edu/chemistry/installing-loggerpro/>).

[Review Guides](https://chemlab.truman.edu/main-review-page/)
(<https://chemlab.truman.edu/main-review-page/>).

[Proofreader's Marks](https://chemlab.truman.edu/proofreaders-marks/)
(<https://chemlab.truman.edu/proofreaders-marks/>).

Forms

Crystal Model Request

Form

(<https://chemlab.truman.edu/files/2015/07/Crystal-Model-Request-Form.pdf>).

NMR Class Submission

Form

(<https://chemlab.truman.edu/files/2015/07/NMRClassSubmissionForm.pdf>).

Databases and References

AIST Spectral Database

(http://riodb01.ibase.aist.go.jp/sdbs/cgi-bin/cre_index.cgi?lang=eng).

NIST WebBook

(<http://webbook.nist.gov/chemistry/>).

NMR Solvents

(<https://chemlab.truman.edu/nmr-solvent-properties/>).

Advanced Regression with Microsoft Excel

Author: J. M. McCormick

Last Update: December 29, 2010

It is possible to have Excel perform a non-linear least square regression. One simple trick is to create columns each containing the variable of interest to the requisite power. For example, if we wanted to fit a set of data to a third order polynomial (i. e., $y = ax + bx^2 + cx^3 + d$, where a , b , c and d are constants that we need to find), then we would create columns containing the independent variable to the desired powers, as shown in Fig. 1. Note that in addition to x^1 , x^2 and x^3 there must be a column containing x^0 , which contains only ones. Once the spreadsheet is set up as shown below, select Tools, Data Analysis from the menu bar and scroll down to Regression, select it and click OK.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		x^0	x^1	x^2	x^3	y									
2		1	-13	169	-2197	2571.7									
3		1	-12	144	-1728	2598.8									
4		1	-11	121	-1331	2625.5									
5		1	-10	100	-1000	2651.7									
6		1	-9	81	-729	2677.5									
7		1	-8	64	-512	2702.8									
8		1	-7	49	-343	2727.5									
9		1	-6	36	-216	2751.6									
10		1	-5	25	-125	2775.3									
11		1	-4	16	-64	2798.4									
12		1	-3	9	-27	2821									
13		1	-2	4	-8	2843.1									
14		1	-1	1	-1	2864.7									
15		1	1	1	1	2905.9									
16		1	2	4	8	2925.4									
17		1	3	9	27	2944.4									
18		1	4	16	64	2962.7									
19		1	5	25	125	2980.5									
20		1	6	36	216	2997.6									
21		1	7	49	343	3014.1									
22		1	8	64	512	3029.8									
23		1	9	81	729	3044.8									
24		1	10	100	1000	3059.1									
25		1	11	121	1331	3072.6									
26		1	12	144	1728	3085.4									
27		1	13	169	2197	3097.5									
28		1	14	196	2744	3108.8									
29															
30															
31															

Figure 1. Sample spreadsheet that is ready to be fit to the cubic expression $y = ax + bx^2 + cx^3 + d$ using Excel's regression package.

The regression pop up will appear (Fig. 2). In the Input Y Range box type in the cell addresses that contains your y values (or click and drag to select them from the worksheet). In Fig. 1 this would be the cells F2 to F28 (enter F2:F28 in the Input Y Rangebox). While entering the y values is no different than for a linear regression, inputting the x values is very different in that the Input X Range box must contain all of the columns containing a power of x. In Fig. 1 this would be columns B through E (cells B2 to E28). Once the x and y ranges are set, you can set any other parameters as desired, and click OK.

Regression

Input

Input Y Range:

Input X Range:

☐ Labels ☐ Constant is Zero

☐ Confidence Level: %

Output options

☐ Output Range:

☒ New Worksheet Ply:

☐ New Workbook

Residuals

☐ Residuals ☐ Residual Plots

☐ Standardized Residuals ☐ Line Fit Plots

Normal Probability

☐ Normal Probability Plots

OK Cancel Help

Figure 2. Regression set-up window

If you have selected to have the results of the fit put in a new worksheet ply, then you will see something like what is shown in Fig. 3 (note that the columns in this spreadsheet have been modified to clearly show their contents). Each of the rows labeled X Variable, except X Variable 1, correspond to one of the constants in the polynomial. The row labeled Intercept will always be the constant in the polynomial that is not multiplied by x. In this case where the polynomial is $y = ax + bx^2 + cx^3 + d$, the Intercept is d, X Variable 2 is a, X Variable 3 is b and X Variable 4 is c. X Variable 1 will always be meaningless. The values under the Standard Error and Upper and Lower 95% for each variable have the same meaning as they do in a simple regression.

Note that the values of R and R^2 are not going to be particularly useful in defining how well this expression fit the data, or in comparing this fit to a fit using another expression. For this one must use the standard error about the regression (listed as the Standard Error under the Regression Statistics heading). Also note that we could set up a column of calculated y values based on this fit (for example, in column G of the spreadsheet shown in Fig. 1) by entering the function in the appropriate column and using absolute references to the values in the output spreadsheet (Fig. 3).

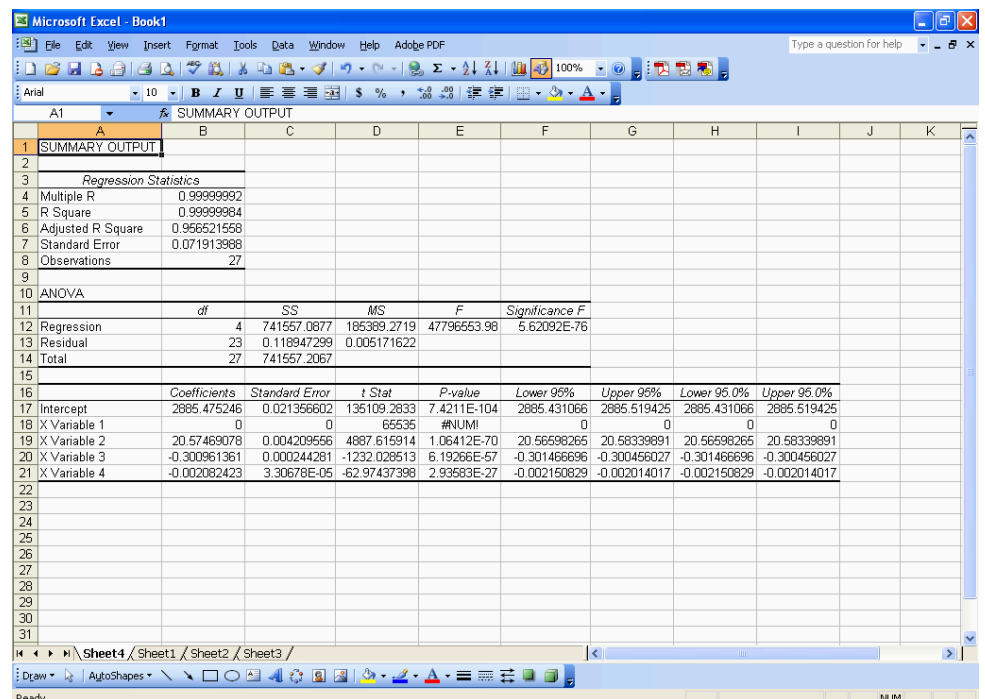


Figure 3. Excel worksheet showing the results for the fit of the data in Fig. 1 to a cubic function.

Another way to have Excel perform a non-linear least squares fit is to use the Solver tool. Before getting started, you will need to set up the spreadsheet, but in a different way than for the Regression

package. Again let us assume that we want to fit a set of y values to a cubic expression in x (i. e., $y = ax + bx^2 + cx^3 + d$). We only need a single column for the x values, but we will need to have two columns for the y values (see Fig. 4). The first column of y values (column C in Fig. 4) will contain the y values that we actually measured. It is labeled y_{obs} (for observed values) to avoid confusion with the calculated y values (column D in Fig. 4), which we label y_{calc} (for calculated values). They_{calc} are determined using the values of a , b , c and d . Thus, to find the values of y_{calc} we will have to have cells containing trial values of a , b , c and d (these trial values are sometimes referred to as seed values). These can be placed in any convenient place in the spreadsheet. In Fig. 4 they are in cells J2 through J5 (note that to minimize confusion the cells I2 through I5 contain labels for the adjacent cells). We could type in any number that we want as trial values, but to maximize the chance of quickly obtaining a successful fit, the trial values should be reasonable guesses. The last column that we will need is one that contains the square of each y_{obs} minus the corresponding y_{calc} (hence the name least squares). In the current example, this is column E in Fig. 4. The final step is to create a cell containing the sum of the values in column E, which is cell H6 in Fig. 4.

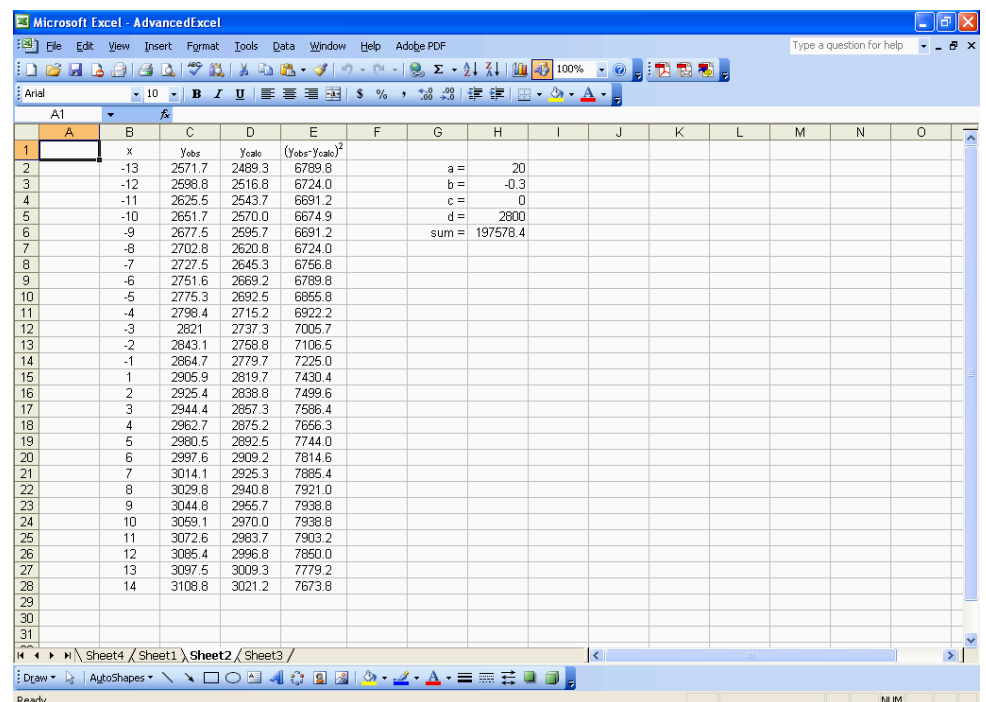


Figure 4. Excel spreadsheet set up to fit a data set to any expression in the variable x using the Solver tool.

Once the spreadsheet is ready, select **Tools, Solver** from the menu bar. The Solver window will pop up (Fig. 5). Under **Set Target Cell**, enter the cell containing the value of $\sum(y_{\text{obs}} - y_{\text{calc}})^2$, which in our example is cell H6. Select the **Min (minimum)** radio button under the **Equal To** heading (note that if you select **Value Of** and enter zero, you may not get a reasonable answer). In the **By Changing Cells**

box, enter the cell(s) containing the trial values that you want to change. Since we want to fit the data by adjusting all of the values in cells H2 to H5, we would enter H2:H5 in this box (or by highlighting all of the cells in the spreadsheet). Note that you can choose to vary all or only some of the trial values by which cells you select at this point. Sometimes it is useful to vary only a subset of all the parameters, while holding others constant, but this is highly dependent on the situation. You may also apply additional constraints to the fit in the Subject to Constraints box by clicking the Add button. For example, to constrain the value in cell H2 to be positive, one would click Add and then in the new pop-up enter H2 in the Cell Reference box, select the "> =" in the middle pull-down box, and then enter "0" in the Constraint box.

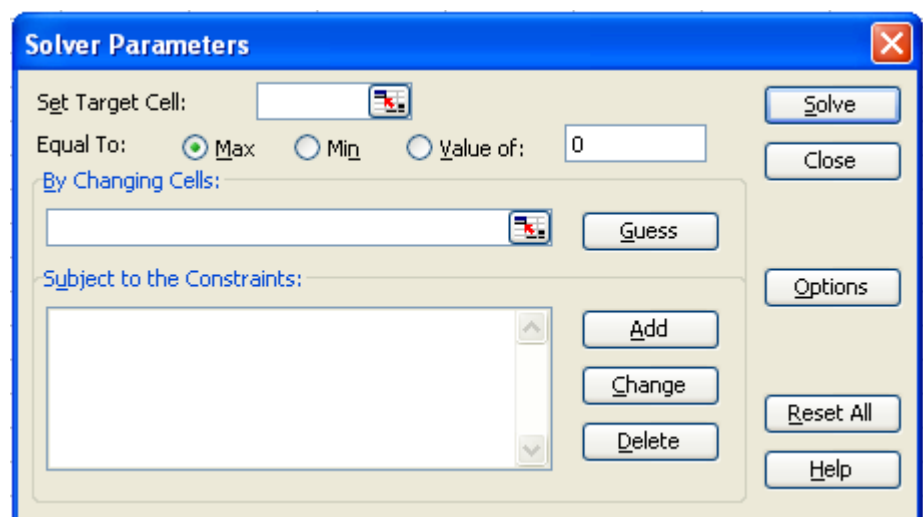


Figure 5. Solver set-up window

Once everything is set, click Solve. Solver will then attempt change the cells that you told it to change, subject to any constraints that you specified, to minimize the difference between the observed and calculated y values. If it is successful it will display a window stating that it found a solution and asking whether you want to keep it (note that the best fit values of your variables and the function that you were minimizing will be displayed in the spreadsheet behind the pop-up). If you want to keep the fit, click OK, if not click CANCEL. If a satisfactory fit was not found, a similar message will be displayed and you can either keep or discard the result.

Sometimes you will need to restart the Solver tool (that is reinitiate the procedure with the values from a previous fit) before a satisfactory fit is obtained. In any case, it is good practice to perform the fit with different values of the initial parameters to be sure that the procedure converges to a unique solution. It may also be necessary to adjust the fitting procedure itself (done by clicking the

Options button in the main Solver window) to get a good result. However, if this occurs it may be advisable to find a more robust fitting algorithm or to find a simpler model.

Unlike the Regression package, Solver does not give uncertainties in the fit parameters. However, there is a way to estimate the uncertainties,¹ but it is somewhat time consuming.

References

1. Harris, D. C. *J. Chem. Educ.* **1998**, 75, 119-121. Click **here** (<http://pubs.acs.org/doi/pdf/10.1021/ed075p119>) to view this article as a PDF file (for Truman addresses and *J. Chem. Educ.* subscribers only).

(<https://chemlab.truman.edu>)

Contact Us

(<http://www.truman.edu/contact-us/>)

Maps & Directions

(<http://www.truman.edu/admission-cost/visit-truman/campus-map-directions/>)

Campus Offices

(<http://www.truman.edu/about/our-people/administrative-offices/>)

Bookstore

Copyright ©2024

(<http://www.truman.edu/copyright-notice/>)

Truman State University
100 East Normal Avenue
Kirksville, Missouri 63501
660-785-4000
(tel:16607854000)