Curve Fitting, Confidence Intervals and Envelopes, Correlations, and Monte Carlo Visualizations for Multilinear Problems in Chemistry: A General Spreadsheet Approach

ARTICLE *in* JOURNAL OF CHEMICAL EDUCATION · MAY 2001

Impact Factor: 1.11 · DOI: 10.1021/ed078p827

3 AUTHORS, INCLUDING:

Nick Guy
University of Wyoming
14 PUBLICATIONS **93** CITATIONS

SEE PROFILE

Computer Bulletin Board

edited by
Steven D. Gammon
University of Idaho
Moscow, ID 83844

Curve Fitting, Confidence Intervals and Envelopes, Correlations, and Monte Carlo Visualizations for Multilinear Problems in Chemistry: A General Spreadsheet Approach W

Paul Ogren,* Brian Davis, and Nick Guy
Department of Chemistry, Earlham College, Richmond, IN 47374-4095; *paulo@earlham.edu

Overview

A number of strategies for curve fitting and uncertainty assessment have been presented in standard laboratory textbooks (*1*, Sections III, XXII; *2*). Relevant articles in this *Journal* include the use of spreadsheet multilinear fitting tools (*3*), Solver (*4, 5*), and programming methods for treating multilinear (*6*) and more general nonlinear (*7*) problems. This paper addresses many of the same issues, but uses spreadsheets designed to acquaint students more directly with the logical principles and computational methods used to determine the best-fit parameters. We also use the geometrical properties of the χ^2 square distribution to develop visual depictions of parameter uncertainties and correlations. We use this approach in physical chemistry, hoping to develop a deeper understanding of some of the issues surrounding uncertainty assessment, the importance ("weight") of various parameters in a fit, and correlations between parameter choices. We use spreadsheets because students in our program will have had considerable previous work with these in analytical and general chemistry. In most programs, students in physical chemistry will have a calculus background sufficient to understand the derivations of the necessary equations.

The central features of our approach are:

1. A spreadsheet is used to directly calculate up to three parameters for multilinear examples. This case is simple enough for students to set up and compute the nine sums and the matrix inversion required to minimize the χ^2 merit function, the standard goal of least-square fits. No iterations are required for the multilinear case. Direct computations can provide insight into how this and more general least-squares problems work. They also demonstrate how parameter weights, uncertainties, and correlations are calculated for multilinear problems.
2. A Monte Carlo spreadsheet, which students can link to their initial spreadsheet, is used to demonstrate the dependence of best-fit parameters on the uncertainties of the original data set. This is done by creating 256 alternative "pseudo data" sets, computing best-fit parameters, and then producing scatter plots of the parameter sets. For the multilinear case these plots are elliptical, and often tilted owing to correlations. The boundaries and distributions help students to visually interpret computed uncertainties, and the tilt and eccentricity of the plots help in the understanding of parameter correlations.

Multilinear problems are intermediate between simple straight-line fits to weighted or unweighted data and nonlinear problems such as the fitting of kinetic data in complex models. An understanding of the multilinear solution procedure requires the introduction of concepts which can be easily generalized to nonlinear cases: the nature of the $\Delta\chi^2$ surfaces near the best-fit parameters, the calculation and geometric interpretation of the "curvature matrix", the relationship of best-fit solutions to the "error matrix", the concept of parameter "weights" in determining the quality of fit, and the algebraic and geometric properties of parameter correlations. Three-parameter problems are sufficient to suggest the geometric properties of more complex problems. In our approach, it is also quite simple to use a subset of the calculations to carry out two-parameter fits in order to explore, for example, the importance of a minor additional term. While our approach could be used for more than three parameters, the number of additional matrix terms that must be calculated quickly becomes cumbersome, with virtually no gain in insight into the methodology. For such cases, use of a "package" solution method is better.

Methodology

As illustrations we shall use three examples discussed in recent articles in this *Journal*. All three cases require a fit to a function of the general form

$$f = a_1X_1(x) + a_2X_2(x) + a_3X_3(x) \quad (1)$$

where the X_i may be simple or complex functions of an independent variable x, and where the total function f is *linear* in the parameters a_1, a_2, and a_3. This is the essential definition of "multilinear".

Examples

1. IR Spectrum of $D^{35}Cl$ (*3*)

The equation to be fitted is

$$\Delta\tilde{\nu}(m) = \tilde{\nu}(m+1) - \tilde{\nu}(m) = (2B_e - 3\alpha_e - 4D_e) - (2\alpha_e + 12D_e)m - 12D_em^2 \quad (2a)$$

or

$$\Delta\tilde{\nu}(m) = B_e(2) + \alpha_e(-3-2m) + D_e(-4-12m-12m^2) = B_eX_1(m) + \alpha_eX_2(m) + D_eX_3(m) \quad (2b)$$

where $\Delta\tilde{\nu}(m)$ is the separation in wavenumbers between successive lines in the rotation–vibration spectrum, m is an indexing integer assigned to each line, B_e is the rotational

constant, α_e is a rotation–vibration interaction constant, and D_e is a centrifugal distortion constant.

2. Vibrational Lines in the Electronic Spectrum of I_2 (8)

The equation to be fitted is

$$\tilde{\nu} = a + b(v' + \tfrac{1}{2}) + c(v' + \tfrac{1}{2})^2 \qquad (3)$$

where $\tilde{\nu}$ is the transition wavenumber, v' is the vibrational level in the upper state, and a, b, and c are combinations of spectroscopic constants: $c = -\omega'_e\chi'_e$, $b = \omega'_e$, and $a = E_{el} - \tfrac{1}{2}\omega''_e - \tfrac{1}{4}(\omega''_e\chi''_e)$.

3. Chromatographic Efficiency Described by van Deemter Equation (4, 5)

The equation to be fitted is

$$y = Ax + B/x + C \qquad (4)$$

where y is the plate height (HETP), x is the flow rate of mobile phase through a column, C is a multiple path parameter, A is an equilibration time parameter, and B is a longitudinal diffusion parameter.[1]

Analysis

Table 1 illustrates the meaning of the $\{X_1, X_2, X_3\}$ set for each of these cases. Note that we have chosen eq 2b for the first example, even though the simpler polynomial form (eq 2a) would seem easier. The reason is that fitting to eq 2b will lead to direct determination of uncertainties and correlations between individual parameters in the $\{B_e, \alpha_e, D_e\}$ set. Use of eq 2a will lead directly to uncertainties in *combinations* of the parameter set. Table 1 also provides individual σ_y values for the weighted van Deemter set of ref *4*. In the other examples, a constant-value σ_y is determined from the multilinear fits as described below.

The general strategy for the multilinear fitting problem is worked out in many sources. We repeat some of this analysis briefly, but in sufficient detail to permit spreadsheet development. Our notation parallels that found in the standard Bevington and Robinson text (*9*), with some modifications described in ref *10*. Following Bevington and Robinson, the goal in fitting eq 1 is to minimize the χ^2 merit function, which is defined by

$$\chi^2 = \sum_i \left(\frac{y_i - f_i(x_i)}{\sigma_i} \right)^2 \qquad (5)$$

χ^2 must be minimized with respect to each of the parameters $\{a_1, a_2, a_3\}$. This leads to three equations of the general form

$$\frac{\partial \chi^2}{\partial a_n} = -2\sum_i \frac{[y_i - f_i(x_i)]}{\sigma_i^2} \frac{\partial f_i(x_i)}{\partial a_n} = -2\sum_i \frac{y_i X_n(x_i)}{\sigma_i^2} + 2\sum_1^3 a_k \sum_i \frac{X_n(x_i)X_k(x_i)}{\sigma_i^2} \qquad (6)$$

or

$$-2b_n + 2(a_1\alpha_{n_1} + a_2\alpha_{n_2} + a_3\alpha_{n_3}) = 0 \qquad (7)$$

for $n = 1, 2, 3$, where[2]

$$b_n = \sum_i \frac{y_i X_n(x_i)}{\sigma_i^2} \quad \text{and} \quad \alpha_{nk} = \sum_i \frac{X_n(x_i)X_k(x_i)}{\sigma_i^2} \qquad (8)$$

Table 1. Illustration of the Three Cases

$D^{35}Cl$ Infrared		I_2 Electronic/Vibrational		van Deemter Data		
$X_1 = 2$		$X_1 = 1$		$X_1 = x$		
$X_2 = -3 - 2m$		$X_2 = v' + 1/2$		$X_2 = 1/x$		
$X_3 = -4 - 12m - 12m^2$		$X_3 = (v' + 1/2)^2$		$X_3 = 1$		
$x = m$	$y = \Delta\tilde{\nu}/cm^{-1}$	$x = m$	$y = \tilde{\nu}/cm^{-1}$	x/(mL min^{-1})	y/mm	σ_y
15	6.81	18	17,702	3.4	9.59	0.48
14	7.08	19	17,797	7.1	5.29	0.26
13	7.33	20	17,889	16.1	3.63	0.18
12	7.63	21	17,979	20.0	3.42	0.17
11	7.85	22	18,064	23.1	3.46	0.17
10	8.16	23	18,149	34.4	3.06	0.15
9	8.36	24	18,235	40.0	3.25	0.16
8	8.69	25	18,318	44.7	3.31	0.17
7	8.86	26	18,396	65.9	3.50	0.18
6	9.17	27	18,471	78.9	3.86	0.19
5	9.37	28	18,546	96.8	4.24	0.21
4	9.64	29	18,618	115.4	4.62	0.23
3	9.84	30	18,688	120.0	4.67	0.23
2	10.12	31	18,755			
1	10.28	32	18,825			
-2	11.04	33	18,889			
-3	11.21	34	18,954			
-4	11.42	35	19,019			
-5	11.69	36	19,077			
-6	11.81	37	19,131			
-7	12.07	38	19,186			
-8	12.26	39	19,238			
-9	12.47	40	19,286			
-10	12.63	41	19,339			
-11	12.86	42	19,384			
-12	13.03	43	19,429			
-13	13.20	44	19,467			
-14	13.39	45	19,512			
-15	13.56	46	19,546			
-16	13.60	47	19,585			

For a three-parameter problem, three b_n sums and six α_{nk} sums must be calculated (note that $\alpha_{nk} = \alpha_{kn}$). After this is done, the three resulting expressions within eq 7 can be combined in the matrix form:

$$(\boldsymbol{b}) = [\boldsymbol{\alpha}](\boldsymbol{a}) \qquad (9)$$

The $(\boldsymbol{a})$ vector elements are the best-fit solution set $\{a_1, a_2, a_3\}$ for eq 1. The square matrix $[\boldsymbol{\alpha}]$ is called the *curvature matrix* because it describes the shape and orientation of ellipsoidal surfaces of constant $\Delta\chi^2$ near the minimum value of χ^2. Some of these geometric ideas are reviewed more extensively in the Appendix. The elements of the $(\boldsymbol{b})$ vector and the curvature matrix, given by eq 8, are easily calculated on a spreadsheet. The $(\boldsymbol{a})$ vector is then obtained by inverting the curvature matrix, using the Excel spreadsheet MINVERSE function.[3]

Formally the solution to eq 9 is written

$$(\boldsymbol{a}) = [\boldsymbol{\alpha}]^{-1}(\boldsymbol{b}) = [\boldsymbol{\varepsilon}](\boldsymbol{b}) \qquad (10)$$

where $[\boldsymbol{\varepsilon}]$, known as the "error matrix", is the inverse of $[\boldsymbol{\alpha}]$. Individual elements of $(\boldsymbol{a})$ are obtained from the $[\boldsymbol{\varepsilon}]$ matrix and $(\boldsymbol{b})$ vector elements by equations such as

$$a_1 = \varepsilon_{11}b_1 + \varepsilon_{12}b_2 + \varepsilon_{13}b_3 \qquad (11)$$

Table 2. $\Delta\chi^2$ for Different Confidence Levels and Numbers of Parameters

Confidence Level	No. of Parameters	$\Delta\chi^2$
68.3% (1σ)	1	1
	2	2.30
	3	3.53
95.4% (2σ)	1	4
	2	6.17
	3	8.02

Once the set of constants $\{a_1,a_2,a_3\}$ and the error matrix elements have been determined, we also have enough information to obtain uncertainties and correlations. Before discussing these, however, we find it useful to introduce an additional step, which uses the best-fit values to define a "dimensionless" or "reduced-variable" curvature matrix [$\boldsymbol{\alpha}'$]. In more general cases (*10*), a dimensionless approach allows one to treat parameters that have very different units and meanings on a comparable statistical basis. The range of values in the dimensionless matrix elements also tends to be smaller than for the original curvature matrix, and this reduces the problem of round-off error in the computations. To convert to a dimensionless format, we designate the best-fit parameter set as $\{a_1^*,a_2^*,a_3^*\}$.[4] Noting that y_i, σ_i, and a_iX_i all must have the same units, we then define elements of a dimensionless ($\boldsymbol{b}'$) and [$\boldsymbol{\alpha}'$] matrix by

$$b'_n = \sum_i \frac{a_n^* y_i X_n(x_i)}{\sigma_i^2} \quad \text{and} \quad \alpha'_{nk} = \sum_i \frac{a_n^* a_k^* X_n(x_i) X_k(x_i)}{\sigma_i^2} \tag{12}$$

The reasons for this definition are discussed in the Appendix. The dimensionless curvature matrix and its associated dimensionless error matrix [$\boldsymbol{\varepsilon}'$] now lead to the following additional results.

Parameter weights. It has been shown elsewhere (*10*) that the diagonal elements of the [$\boldsymbol{\alpha}'$] matrix are proportional to the "weights" of the parameters in determining the fit. Thus, simple examination of the [$\boldsymbol{\alpha}'$] matrix provides insight to interpreting which parameters are most important in determining the "goodness of fit".

Parameter uncertainty estimates. For multilinear problems, parameter uncertainties are tied to the multivariate normal distribution of possible parameter sets (*11*, p 290). Ellipsoidal volume regions that include a certain percentage of these sets are defined by the magnitude of $\Delta\chi^2_{np}$ in the multidimensional space of the p independent parameters. Values of $\Delta\chi^2_{np}$ have been tabulated for several cases of n and p (*12*, p 536), and a few of these results are reproduced in Table 2. Combining the proper $\Delta\chi^2_{np}$ value with the appropriate diagonal element of the dimensionless error matrix gives

$$n\sigma_{a_j} = a_j^* \sqrt{\Delta\chi^2_{np}} \sqrt{\boldsymbol{\varepsilon}'_{jj}} \tag{13}$$

where p is the number of fitted parameters and $\pm n\sigma_{a_j}$ are the confidence limits for the parameter a_j at the confidence level shown in Table 2. For n = 1 (1σ) and p = 3 parameters, $\Delta\chi^2_{np}$ is 3.53. If different n and p values are involved, the required $\Delta\chi^2_{np}$ values could be entered in the spreadsheet as a lookup table.

Correlations between parameters. The elements of the dimensionless curvature matrix [$\boldsymbol{\alpha}'$] contain information about pairwise correlations in two commonly used forms (*10*):

$$\rho_{jk} = \frac{-\alpha'_{jk}}{\sqrt{\alpha'_{jj}\alpha'_{kk}}} \tag{14}$$

where ρ_{jk} is the "correlation coefficient" and

$$\frac{a_j - a_j^*}{a_j^*} = \frac{\delta a_j}{a_j^*} = \delta a'_j = \rho_{jk}\sqrt{\frac{\alpha_{kk}}{\alpha_{jj}}} = -\frac{\alpha'_{jk}}{\alpha'_{jj}}\delta a'_k \tag{15}$$

where $\delta a'_i = \delta a_i/a_i^*$ is the *relative error* in parameter a'_i. The correlation coefficient in eq 14, whose value ranges from 0 (no correlation) to ±1 (strong correlation), is closely associated with the shape and orientation of the distribution of parameter uncertainties. For multilinear problems, $\rho \approx \pm 1$ corresponds to a strongly tilted and highly eccentric $\Delta\chi^2$ ellipsoid. Equation 15 provides a simple recipe for estimating the influence of changes in one parameter on another. Two particularly important points should be made about the use of eq 15, however. First, consider the possibility that parameter a'_k has little influence (low weight α'_{kk}) on a fit relative to the influence of parameter a'_j. Then changes in a'_k will have little impact on a new best-fit value of a'_j even if the correlation between these parameters is quite large. This reasonable expectation is in complete accord with eq 15, which states that mutual influences of parameter values are a function of both the correlation coefficient *and* relative weights. Second, when more than two parameters have comparable weights and strong mutual correlations, the use of eq 15 for pairwise correlations may be quite misleading, as the third example will illustrate. A more general version of eq 15, which includes mutual correlations between several parameters, is given by

$$\delta a'_j = \delta a'_k \frac{\text{cof}(\alpha'_{jk})}{\text{cof}(\alpha'_{kk})} \tag{16}$$

where "cof" refers to the cofactor of a designated curvature matrix element. This equation is geometrically related to the ellipsoidal description of confidence limits and "confidence limit vectors" and is discussed more fully in the Appendix.

Example 1: $D^{35}Cl$ Infrared Spectrum Analysis

Table 3 illustrates spreadsheet results for typical data obtained on a moderate-resolution FTIR instrument. The original spectrum, which includes $D^{37}Cl$ lines as well, is shown in Figure 1, with the m index labels used in the analysis.[5] Data of this sort can be treated as a multilinear 4-parameter (cubic) fit problem (*1*, p 400), or even a 5-parameter problem when overtone transitions are included (*3*). For our purposes however, we used a frequency difference method to reduce the data to a three-parameter problem described by eq 2b. Table 3 includes partial spreadsheet results for terms in the summations of eq 8 and the nine summation values needed to construct ($\boldsymbol{b}$) and [$\boldsymbol{\alpha}$]. (It does not show the columns for the $\{x_i,y_i\}$ data and the $X_n(x_i)$ values obtained from the formulas in Table 1.) It then presents results for the original and dimensionless curvature and error matrices, best-fit parameters, and their 1σ uncertainties and normalized weights. The last, totaling 1.000,

Table 3. Spreadsheet Results for Typical Moderate-Resolution FTIR Data

Polyfit — D35Cl IR data — f = Be(2) + αe(-3-2m) + De(-4-12m-12m²)

[α] Curvature Matrix

120	-120	-65520
-120	11030	65520
-65520	65520	60266976

[ε] Error Matrix

0.020596	9.17E-05	2.23E-05
9.17E-05	9.20E-05	1.99E-23
2.23E-05	1.99E-23	4.08E-08

[α'] Dimensionless Curvature Matrix

2870319	-58767.5	-42571.3
-58767.5	110595.9	871.6138
-42571.3	871.6138	1063.693

[ε'] Dimensionless Error Matrix

8.61E-07	1.87E-07	3.43E-05
1.87E-07	9.10E-06	-2.40E-21
3.43E-05	-2.40E-21	0.002313

Be (cm⁻¹)	αe (cm⁻¹)	De (cm⁻¹)	I (kg m²)	r (pm)	
5.44957	0.111576	0.000148	5.14E-47	127.45	values
0.0095	0.000634	1.34E-05		0.11	uncertainties
0.96256	0.037088	0.000357			weights
				127.46	literature

Linear Error Matrix

0.008425	9.17E-05
9.17E-05	9.17E-05

Linear fit Be (cm⁻¹)	αe (cm⁻¹)
5.368742	0.111576

number of data points	30	2.014	mass D
sigma from quadratic fit	0.0352361	34.969	mass Cl-35
chi-square, 3 deg. freedom	27.0000		

SUMS

α11	α22	α33	α12	α13	α23	b1	b2	b3
120	11030	60266976	-120	-65520	65520	630.86	586.43	-340824

X1X1/σ²	X2X2/σ²	X3X3/σ²	X1X2/σ²	X1X3/σ²	X2X3/σ²	yX1/σ²	yX2/σ²	yX3/σ²
4	1089	8317456	-66	-5768	95172	13.62	-224.73	-19640
4	961	6370576	-62	-5048	78244	14.16	-219.48	-17869.9
4	841	4787344	-58	-4376	63452	14.66	-212.57	-16038
.....	30	data	points	.....	.....	.....	.....	.....

are obtained by dividing each diagonal curvature matrix element by the trace of this matrix.

In the initial treatment of the data, we assumed a constant uncertainty for each data point and used $\sigma_i = 1$. This is fine for determining the best-fit parameter values, but not for determining parameter uncertainties. Table 3 also reports a value of 0.035 determined from the standard deviation of the data points from the fit curve.[6] This value is used for σ in eq 12 to calculate the *dimensionless* curvature matrix elements. In doing this, we are choosing deviations that are "externally consistent" with the fitted function, as opposed to deviations that are "internally consistent" with and determined by sources of error in the experimental measurements (*13*). When the externally consistent choice is made, the value of χ^2 should equal the number of data points minus the degrees of freedom from the fit, 30 – 3 = 27 in this example. This can be used as a check on the validity of the rather complex computations.[7]

A common objective of this particular experiment is to determine the internuclear bond distance r from B_e using the definitions:

$$B_e = \frac{h}{8\pi^2 I_e c} \quad \text{and} \quad I_e = \mu r^2 \tag{17}$$

The r value and the uncertainty, calculated from the uncertainty in B_e, are reported in Table 3. In this example it is also straightforward to determine the best linear fit by dropping the D_e and associated X_3 terms. This results in a smaller curvature matrix, the upper left 2×2 submatrix of the quadratic [α] matrix in Table 3, and a corresponding 2×2 linear-fit error matrix.

All of these results may be used to construct the plots in Figure 2, which compare the two-parameter (linear) and three-parameter fits. It is obvious that the D_e term improves the fit, even though its relative weight is quite small (0.04%). The low weight of the D_e term is associated with the fact that the relative uncertainty in D_e is rather high, almost 10%. In addition to improving the fit, the inclusion of the D_e term clearly raises the intercept and the B_e value, as noted in ref *3*. Ignoring D_e introduces systematic error in B_e and consequently in r. The results in Table 3 indicate that including the D_e term will raise B_e by 1.5%. This lowers the r value by about 0.7%, leading to much better agreement with literature data (*14*).

The systematic error introduced by ignoring D_e is closely tied to the issue of parameter correlations and weights, and eq 15 can sometimes be used to estimate the effect of a parameter change without a detailed fitting process. For the previous example, eq 15 becomes

$$\frac{\delta B_e}{B_e^*} = \frac{\delta B_e}{5.450} = -\frac{\alpha'_{BD}}{\alpha'_{BB}} \frac{\delta D_e}{D_e^*} = \frac{42{,}571}{2{,}870{,}319} \frac{\delta D_e}{0.00148} \tag{18}$$

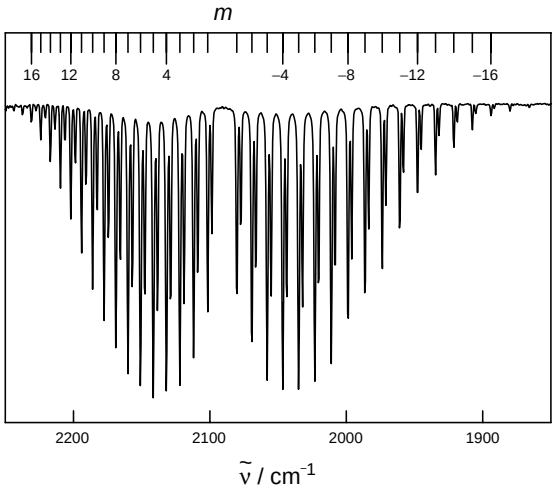

Figure 1. IR transmission spectrum of gaseous DCl, taken with a Perkin-Elmer Paragon 1000 PC FTIR instrument at 1 cm^{-1} resolution (points at 0.5-cm^{-1} intervals). The spectrum shows partially resolved ^{35}Cl and ^{37}Cl isotopic pairs. The m index labels, corresponding to eq 2b, are indicated for the more abundant ^{35}Cl isotope.

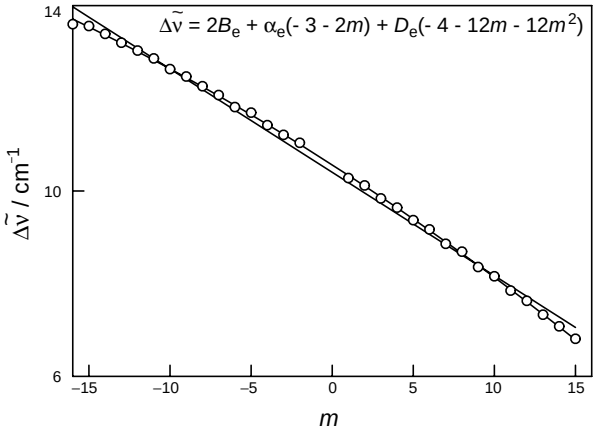

Figure 2. A comparison of 2-parameter (linear) and 3-parameter fits of the $D^{35}Cl$ IR data to eq 2. The fit parameter results are summarized in Table 3.

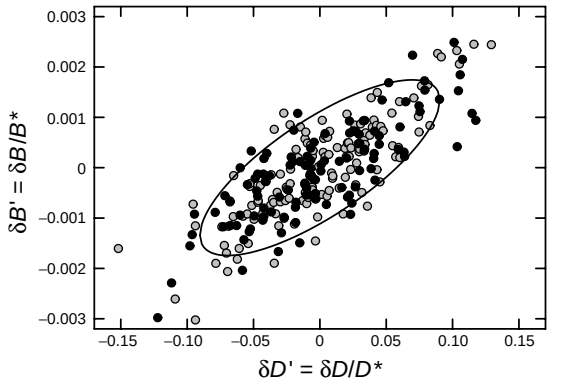

Figure 3. Monte Carlo scatter plot for variations in the D_e and B_e parameters of eq 2; 256 Monte Carlo data sets were used in calculating the individual points. The ellipse is the projection of the $\Delta\chi^2$ = 3.53 ellipsoidal surface onto the $\delta D_e' - \delta B_e'$ plane. The rather even distribution of the solid (+z) and gray (-z) points implies negligible tilt of the full 3-D distribution with respect to the $\delta D_e' - \delta B_e'$ plane.

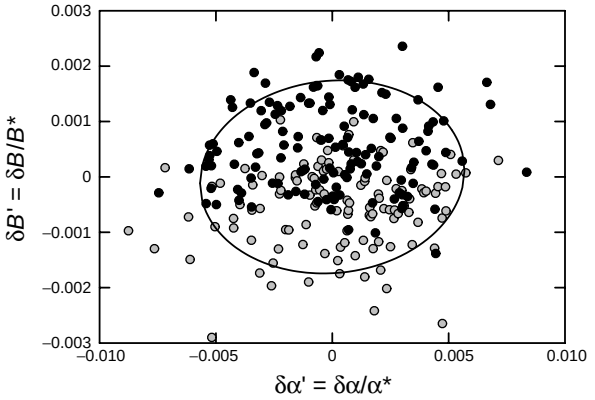

Figure 4. Monte Carlo scatter plot for variations in the α_e and B_e parameters of eq 2. The ellipse is again the projection of the $\Delta\chi^2$ = 3.53 surface onto the plane. The low eccentricity and nearly vertical orientation are associated with low correlation between these two parameters. The solid ($z = +\delta D_e'$) and gray ($z = -\delta D_e'$) point distribution indicates that positive $\delta B_e'$ values tend to be associated with positive $\delta D_e'$ values, consistent with Figure 3.

using the best-fit values of B_e^* and D_e^*, and the curvature matrix elements from Table 3. If one chooses $\delta D_e = -0.00148$, the effect is to shift D_e to a value of zero. The equation then predicts a 1.48% decrease in $\delta B_e/B_e^*$, in excellent agreement with the fitting calculation. The more complex expression (eq 16) also gives 1.48%. We can "get away with" using the simpler pairwise correlation formula (eq 15), which ignores the influence of α_e, because changes in D_e have virtually no effect on α_e, so there will be no additional influence of a shifted α_e on the value of B_e. It is probably obvious that these special conditions may not apply in many other cases.

Monte Carlo Analysis

Equations 13–16 answer important questions about uncertainties and correlations for multilinear fitting problems. The numbers alone, however, do not provide as much insight into the problem as pictures constructed from alternative data sets produced by Monte Carlo methods (*11*, Chapter 21). For the multilinear problem, each alternative data set produces a set of fit parameters that can be associated with a point in parameter space. For three fitted variables, a multitude of such points begins to resemble a three-dimensional ellipsoidal probability distribution, and projections or cross-sections of this surface in 2-parameter planes are easily visualized.

Our students use the Monte Carlo analysis by pasting the data of Table 3 into a provided spreadsheet that will accommodate up to 50 data points. They also paste in the X_1, X_2, and X_3 values, the σ_y values, and the dimensionless curvature and error matrices. The output consists of scatter plots that indicate confidence intervals and correlations between fitted parameters. The general strategy of the spreadsheet is as follows.[8]

First, use the data uncertainties to create new "pseudo data sets" $\{y_i\}$. The x_i values (and hence the X_i values) are left unchanged. We make the usual assumption that the y_i uncertainty distribution is Gaussian, and then generate random deviates with a normal distribution. The deviates are generated in (y_1,y_2) pairs using the Box–Muller algorithm (*12*, pp 200–203):

$$y_i = \cos 2\pi z_2\sqrt{-2\ln z_1} \quad \text{and} \quad y_2 = \sin 2\pi z_2\sqrt{-2\ln z_1} \qquad (19)$$

where z_1 and z_2 are random points in the (0,1) interval. These deviates are multiplied by the standard deviation estimates of the original experimental points and then added to the experimental point values. Our spreadsheet creates 256 new data sets in this fashion.

Second, compute a "best-fit" parameter set for each of the 256 data sets. This requires three new summations involving X_i and y_i, followed by matrix calculations for each set. However, with current PC computational speeds, the entire calculation, including plot updates, takes on the order of one second. The parameter sets are saved and presented in the dimensionless relative error format described in eq 15: $\delta a_i' = \delta a_i/a_i^*$. Plot axes depict shifts in the $\delta a_i'$ away from the original best-fit values of 0.

Third, construct scatter plots of the fit parameters. These are most readily comprehended as two-dimensional plots involving a pair of parameters. When the spreadsheet is run repeatedly, using the F9 key on PC systems, slight fluctuations in the scatter patterns give an even better feeling for the range and stability of the results.

Figures 3 and 4 show Monte Carlo scatter plots for the $D^{35}Cl$ parameters. In both cases, the distributions are approxi-

Table 4. Best-Fit Parameter Values Using I_2 Visible Spectral Data

Polyfit — Iodine Data — $f=a+b(v'+0.5)+c(v'+0.5)^2$

[α] Curvature Matrix

30	990	34917.5
990	34917.5	1300613
34917.5	130061	5056542

[ε] Error Matrix

8.177908	-0.51322	0.007554
-0.51322	0.03289	-0.000492
0.007554	-0.000492	7.45E-06

[α′] Dimensionless Curvature Matrix

1.67E+09	4.68E+08	-1.30E+08
4.68E+08	140092698	-4.03E+07
-1.30E+08	-4.03E+07	12114135

[ε′] Dimensionless Error Matrix

1.47E-07	-1.10E-06	-2.07E-06
-1.10E-06	8.20E-06	1.59E-05
-2.07E-06	1.59E-05	3.11E-05

a (cm^{-1})	b (cm^{-1})	c (cm^{-1})	
15603.65	132.41782	-1.02325	values
11.2323	0.7123457	0.01072	uncertainties
0.91653	0.0768262	0.006643	weights

30	# of data points
27	chi-square
2.09055	sigma, quadratic fit

SUMS

α_{11}	α_{22}	α_{33}	α_{12}	α_{13}	α_{23}	b_1	b_2	b_3
30	34917.5	50565421	990	34917.5	1300613	563474	18740468	6.70E+08

X_1X_1/σ^2	X_2X_2/σ^2	X_3X_3/σ^2	X_1X_2/σ^2	X_1X_3/σ^2	X_2X_3/σ^2	yX_1/σ^2	yX_2/σ^2	yX_3/σ^2
1	342.25	117135.06	18.5	342.25	6331.625	17702	327487	6058510
1	380.25	144590.06	19.5	380.25	7414.875	17797	347041.5	6767309
1	420.25	176610.06	20.5	420.25	8615.125	17889	366724.5	7517852
.....	30	data	points	.....		.....	.....	.....

mately elliptical, as expected (cf. Appendix). To help visualize the third parameter dimension, solid circles are used for points lying above the plane of the page, and gray-scale circles for points below the page plane. In Figure 3, the tilt of the plot implies positive (+ρ) correlation between the B_e and D_e parameters; that is, $\delta D_e/D_e^*$ values that are slightly higher than the original value at the origin are associated with slightly higher $\delta B_e/B_e^*$ values. From eq 14, we find a correlation coefficient for this pair of parameters of ρ = .77. By contrast, in Figure 4, there is very little tilt or eccentricity in the elliptical distribution. This is completely consistent with the low correlation between the B_e and α_e (ρ = .10). Here, if $\delta\alpha_e/\alpha_e^*$ is slightly higher, the possible B_e values are still rather evenly distributed about $\delta\alpha_e/\alpha_e^* = 0$. Both figures also show an ellipse derived from the three-dimensional $\Delta\chi^2 = 3.53$ surface containing 68% of the points. These ellipses are projections of the 3-D surface onto the $\delta D_e' - \delta B_e'$ or $\delta\alpha_e' - \delta B_e'$ plane, and their derivation is briefly described in the Appendix. Their maximum extensions in the $\delta B_e'$, $\delta\alpha_e'$, and $\delta D_e'$ directions are equal to the *relative* uncertainties for each parameter listed in Table 3.

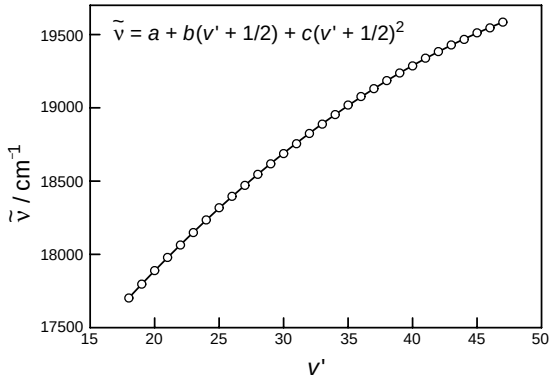

Figure 5. Fit of the iodine vapor spectral data from ref *8* to eq 3. This plot is identical to that in the earlier publication.

Example 2: An Illustration Using I_2 Visible Spectral Data

A data set for an analysis published in this *Journal* was kindly provided by the authors (*8*). Our results for best-fit parameter values (Table 4) and the curve fitting shown in Figure 5 are identical to those of ref *8*. Externally consistent uncertainties of 2.091 cm^{-1} were determined for the *y* values from the fit. These were used for the dimensionless matrices and for obtaining the scatter plots. As expected, $\chi^2 = 30 - 3 = 27$ for 30 data points. The *a* parameter corresponds to the *y* intercept in Figure 5 and the *b* parameter corresponds roughly to the average slope of the plot.

A Monte Carlo scatter plot for the $\delta a' - \delta b'$ plane is shown in Figure 6. At first glance, one might expect that this would be similar to the plot in Figure 4. In fact, however, there is strong negative correlation as shown ($\rho_{ab} = -.97$). Since all of the points in Figure 5 are to the right of the *y* axis, shifting the slope while maintaining a reasonable data fit will neces-

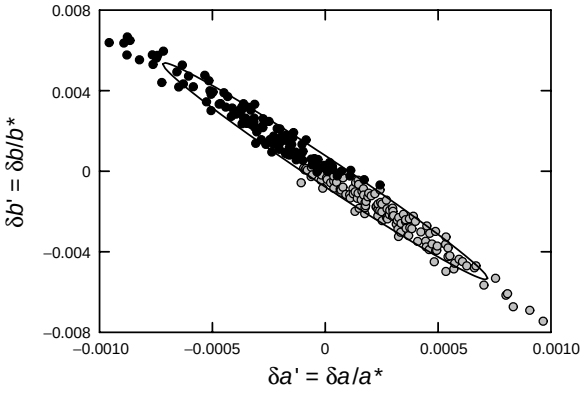

Figure 6. Monte Carlo scatter plot for the iodine spectrum example. The projection ellipse from the $\Delta\chi^2 = 3.53$ surface is also shown. About 81 points (32%) should lie outside this ellipsoidal surface. In addition to ca. 40 points that clearly fall outside the projection ellipse, a number of points that are well above or below the $\Delta\chi^2$ surface appear to fall inside when they are projected onto the graph's plane. The distribution of solid (+*z*) and gray (-*z*) points implies a strong tilt of the $\Delta\chi^2$ surface with respect to the plane of the figure.

Table 5. Computations Using Data from Table 1 (*4*)

y = HETP / mm *x* = flow rate / (mL/min)

Polyfit — van Deemter data — $y = Ax+B/x+C$

[α] Curvature Matrix

1268245	356.3585	17610.61
356.3585	1.035327	12.9228
17610.61	12.9228	356.3585

[ε] Error Matrix

3.89E-06	0.001935	-0.000262
0.001935	2.72852	-0.194584
-0.000262	-0.194584	0.0228196

[α′] Dimensionless Curvature Matrix

724.3386	223.2579	678.5363
223.2579	711.5059	546.1805
678.5363	546.1805	926.2873

[ε′] Dimensionless Error Matrix

0.006803	0.00309	-0.006805
0.00309	0.00397	-0.004604
-0.006805	-0.004604	0.0087791

A	B	C	
0.023898	26.21503	1.612239	values
0.003703	3.103496	0.283819	uncertainties
0.306646	0.301214	0.39214	weights

13 # of data points
2.7949 chi-square

SUMS

α_{11}	α_{22}	α_{33}	α_{12}	α_{13}	α_{23}	b_1	b_2	b_3
1268245	1.035327	356.3585	356.3585	17610.61	12.9228	68043.5	56.492179	1334.17

X_1X_1/σ^2	X_2X_2/σ^2	X_3X_3/σ^2	X_1X_2/σ^2	X_1X_3/σ^2	X_2X_3/σ^2	yX_1/σ^2	yX_2/σ^2	yX_3/σ^2
50.17361	0.375457	4.340278	4.340278	14.75694	1.276552	141.519	12.242136	41.6233
745.7101	0.293452	14.7929	14.7929	105.0296	2.083507	555.607	11.021752	78.2544
8000.309	0.11907	30.8642	30.8642	496.9136	1.917031	1803.8	6.9588222	112.037
.....	13	data	points	.....	.....	.....	.....	.....

sarily shift the intercept. In Figure 2, by contrast, the experimental points are almost evenly distributed on both sides of the *y* axis, so shifting the slope will not greatly alter the intercept, which is dominated by the B_e term. A more thorough analysis of the effect of point distribution may be found in ref *11*, Chapter 32. In addition to the strong *a*–*b* correlation demonstrated by Figure 6, the distribution of closed and gray-scale points shows that the full ellipsoidal surface is strongly tilted in the δ*c*′ parameter direction.

Example 3: Fitting the van Deemter Equation

In contrast to the earlier illustrations, the experimental points used to fit eq 4 have different weighting factors.[9] In addition, the standard deviations have been set by some means other than comparison with the van Deemter fit—perhaps by observed experimental variation in the HETP value for repeated trials at a set flow rate. This is an example of *internally consistent* data. In such cases, one can expect only approximate agreement between the best-fit χ^2 value (2.8) and the value expected from the number of data points and fit parameters (13 – 3 = 10). That the agreement is poor in this example may suggest that the *y*-value uncertainties have been overestimated. For example, if we reduce each uncertainty by a scaling factor of 1.9, the χ^2 value is quite close to 10.

Table 5 summarizes the computations of best-fit parameters, uncertainties, and weights based upon the data of Table 1 taken from ref *4*. The parameter values agree completely with ref *4* for the variable-weight case. In contrast to the first two examples, reasonable σ_y values are available at the outset for calculating the curvature matrix elements, so new σ_y values are not required for the dimensionless elements calculated using eq 12. In this example the weights of all three fit parameters are comparable, a result that is not surprising to anyone who has tried to fit this type of data by guesswork.

Figure 7 shows the data fit, which is quite similar to the results published in ref *4* for the unweighted data case. Figure 7

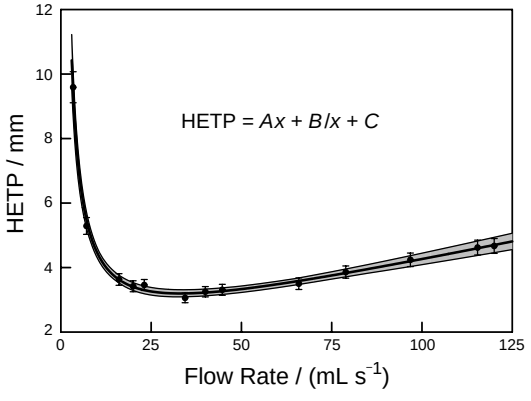

Figure 7. Fit of the van Deemter data to eq 4. The boundary lines above and below the central best-fit line are upper and lower 1σ confidence limits. All of the curves described by points inside the $\Delta\chi^2$ = 3.53 ellipsoidal surface will fall within the confidence interval over the range shown.

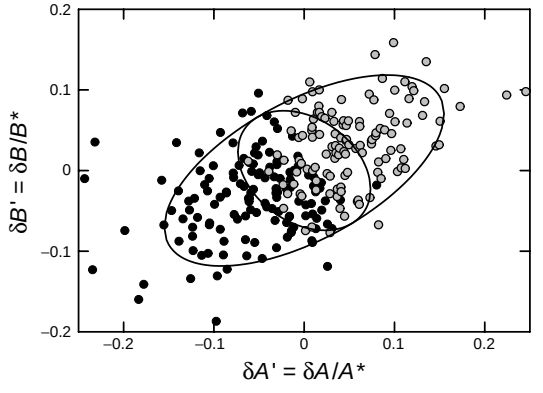

Figure 8. A Monte Carlo scatter plot for the *A* and *B* parameters of the van Deemter equation. The distribution of solid (+*z*) and gray (-*z*) implies that the right side of the ellipsoid tilts below the plane of the figure. The apparently opposite correlation trends of the larger projection ellipse and the smaller intersection ellipse are discussed in the text.

also illustrates the use of Monte Carlo methodology to construct curves that provide "confidence envelopes". We first calculate $\Delta\chi^2$ for the 256 pseudo data set results;[10] 175 (68.4%) of these fall within the 1σ ellipsoid surface, and there are also 175 predicted curves associated with this smaller set. For any value on the *x* axis, we now calculate a set of 175 Monte Carlo curve predictions. The lowest such prediction defines a point on the lower confidence limit boundary and the highest a point on the upper boundary. All 175 curves would fall within the boundaries shown. This concept has minor relevance to this example and is of no particular relevance in the first two examples. However, it is very useful in problems such as kinetic modeling (*15*).

From Table 5, the three pairwise correlation coefficients calculated using eq 15 are $\rho_{AB} = -.31$, $\rho_{AC} = -.83$, and $\rho_{BC} = -.67$. The Monte Carlo data of Figure 8 are particularly interesting in view of the negative correlations between all parameter *pairs*. The larger elliptical curve is again the projection of the ellipsoidal $\Delta\chi^2$ surface onto the $\delta A' - \delta B'$ plane and it correctly suggests a *positive* correlation between *B* and *A* when all three parameters are adjusted; that is, higher values of *A* will be associated with higher values of *B*. The smaller ellipse represents the intersection of the 3-D ellipsoidal surface with the $\delta A' - \delta B'$ plane, and it implies just the opposite, a *negative* correlation between *B* and *A*. Within the intersection plane, eq 15 predicts $\delta B' = -(223/712)\delta B' = -0.31\delta A'$. However, the latter conclusion is true *only* if $\delta C'$ is fixed at a constant value.[11] Equation 16 is the correct algebraic prescription for predicting the influence of a change in *A* on *B* with three adjustable parameters, and for this example, if $\delta A' = 0.030$ (3.0% increase in *A*), the new best-fit value of $\delta B'$ will be given by

$$\delta B' = \delta A' \frac{\mathrm{cof}(\alpha'_{AB})}{\mathrm{cof}(\alpha'_{BB})} = 0.030 \frac{163{,}802}{210{,}534} = +0.023 \qquad (20)$$

(a 2.3% *increase* in *B*). In effect, if $\delta A'$ is increased, a new best-fit analysis will try to stay near the middle of the 3-D ellipsoid by *decreasing* $\delta C'$ while slightly *increasing* $\delta B'$. This example demonstrates that it can be quite misleading to consider only pairwise correlations!

The parameter uncertainties in this example are determined both by the goodness of fit with the van Deemter function and the *y*-value uncertainties themselves, since these were set prior to the fitting. The uncertainties based upon the weighted data points are significantly higher than those reported by Harris for an "unweighted" treatment (*4*). This is partially due to the unusually large internally consistent uncertainties noted above.[12]

Conclusions and Extensions

We hope that these examples illustrate some useful educational advantages of a more detailed introduction to error analysis in upper-level chemistry. The multilinear case introduces several powerful ideas and strategies, but a few comments should be made about more general applications.

First, when fit parameters have different units, a "dimensionless parameter" approach has major advantages that may not be apparent in these simple examples (*10*). A disadvantage is that some prior estimate or knowledge of a reasonable set of parameter values is required, although this is usually a minor issue.

Second, all of these examples make the usual assumption that a Gaussian distribution describes the uncertainties in experimental points. If this is not the case, then the analysis given here will be incorrect. Often data are analyzed in a linearized form, for example, $\ln c = \ln c_0 - kt$ for a first-order kinetic decay process. If the uncertainty in *c* follows a Gaussian distribution, then the uncertainty in $\ln c$ will not even be symmetrical (*16*). However, this may be a minor problem if the relative standard deviation in *c* is small. The effect of predictable deviations from a Gaussian distribution can be tested by a modification of the Monte Carlo methods presented here.

Third, in addition to the "internal" and "external" choice for setting parameter uncertainty, errors can be treated as "fixed" (constant weight), "instrumental" (proportional to measured magnitude), or variable in some other fashion (*9*). It is important to provide clear statements about how uncertainties are being assigned in a given experiment.

Fourth, for multilinear problems, the surfaces of constant $\Delta\chi^2$ will always be ellipsoidal, with a center at $\Delta\chi^2 = 0$, corresponding to a single minimum χ_0^2 value. For nonlinear cases, these surfaces will be approximately ellipsoidal near the minimum χ^2, but they can have very different shapes for larger values. Multiple local minima may also be possible. Simple standard formulas for confidence limits and correlations are valid only for ellipsoidal behavior of the surfaces. The ellipsoidal condition is often known as the *quadratic approximation* because it is associated with a general Taylor expansion of χ^2 that is truncated after quadratic terms. Deviations from ellipsoidal behavior can also be assessed by Monte Carlo methods (*10*), but such cases are not explored here. Generally, nonlinear problems require iterative searches for solutions (*9, 15*), and these are much easier to carry out with standard programming languages such as Fortran.

Finally, all cases discussed here assume that for $y = f(x)$, there is uncertainty in the measured *y* values but no uncertainty in the *x* values. Even the straight-line-fit case for variable weights in both parameters requires iteration (*17*), and more complex cases require a much more versatile computational method than the one presented here.

Acknowledgments

We thank Jan. P. Hessler for helpful comments and suggestions, the authors of ref *8* for their data set, and the Earlham College Ford–Knight Fund for supporting this work. We also thank the reviewers of this paper for detailed checking and a number of useful suggestions for clarifications.

[W]Supplemental Material

The spreadsheets used for Tables 3–5 and a Monte Carlo spreadsheet used for Figure 6 are available in this issue of *JCE Online*. They were developed in a PC/Windows environment.

Notes

1. The standard version of this equation, presented in several analytical chemistry textbooks, is $y = A + B/x + Cx$; that is, the

A and *C* definitions are reversed with respect to eq 4. The reversal is unfortunate, but we make this choice to remain consistent in comparing our results with an identical equation in ref *4*.

2. Bevington and Robinson use β_n rather than b_n in their discussion of multilinear analysis. We use b_n to avoid confusion with another common usage of β_n:

$$\beta_n = -\frac{1}{2}\frac{\partial \chi^2}{\partial a_n}$$

3. Suppose the curvature matrix elements are in cells J4:L6. To create the error matrix array, highlight a 3×3 target cell array on the spreadsheet; then type = MINVERSE(J4:L6) in the formula bar and use <Shift-Control-Enter> to create the inverse array on a PC system.

4. These a_i^* values are the same as the a_i values determined by eq 11. However, we will continue to use a_i in two more general ways: as a variable to be located in a best-fit analysis, and as one of several alternative parameter choices in a Monte Carlo procedure.

5. As described in ref *1*, $m = J + 1$ for the higher energy R branch transitions ($\Delta J = +1$), and $m = -J$ for the P ($\Delta J = -1$) branch, where J is the lower state rotational quantum number for a linear rigid rotor.

6. The Excel function STDEV assumes $N - 1$ degrees of freedom, whereas there are $N - 3$ degrees of freedom in these examples. Therefore we multiply the STDEV result by $[(N - 1)/(N - 3)]^{1/2}$ to obtain the reported table values.

7. We use the σ label in two standard but somewhat different ways. The $n\sigma_{aj}$ of eq 13 is a fit parameter confidence limit. σ_y in Table 1 ($=\sigma_i$) describes the y uncertainty of a data point and is presumed to be the standard deviation of a Gaussian distribution of possible point values in later Monte Carlo calculations.

8. The computation of 256 pseudo-data sets for up to 50 data points generates a very large Monte Carlo spreadsheet, approximately 1300 rows by 150 columns and 3 MB in size. The data entry requires only a small fraction of this space. We use the Excel functions INDEX, INDIRECT, NAME, and IF to simplify selection of variables for *xy* plots and to automatically sort points into the solid (+*z*) and gray-scale (–*z*) subsets.

9. Our analysis only concerns the variable-weight case presented in ref *4*. Both ref *4* and ref *5* have also treated the case of equal weights for these data.

10. Use each of the 256 $\{a_1, a_2, a_3\}$ sets and the *original* data and σ values to calculate a χ^2. These will all be higher than χ_0^2 for the original data and original best-fit parameters. The differences provide a set of 256 $\Delta\chi^2$ values.

11. A slice plane through the $\Delta\chi^2$ surface at any fixed $\delta C'$ value will define an intersection ellipse parallel to the $\delta A' - \delta B'$ plane. This intersection ellipse will have the same tilt and eccentricity as the ellipse shown in the figure for $\delta C' = 0$.

12. Harris (*4*) only reports parameter uncertainty values for the *unweighted* data. When we treat that case, we find an externally consistent standard deviation of 0.1065 for the y values and $\{A,B,C\}$ uncertainties of $\{0.0018, 0.92, 0.14\}$. The comparable Harris values are $\{0.0015, 2.18, 0.16\}$. de Levie (*5*) also used the unweighted data with spreadsheet macros or with a standard nonlinear least-squares routine to determine parameter uncertainties of $\{0.0010, 0.49, 0.076\}$. We get the latter values for the unweighted data *if* we use $\Delta\chi^2_{np} = 1.00$ in eq 13. However, $\Delta\chi^2_{np} = 3.53$ is correct for a three-parameter fit. For this example, both our methods and the de Levie procedures provide simpler and more precise methods of uncertainty computation than the jackknife method of ref *4*.

Literature Cited

1. Shoemaker, D. P.; Garland, C. W.; Nibler, J. W. *Experiments in Physical Chemistry*, 6th ed.; McGraw-Hill: New York, 1996.
2. Halpern, A. M. *Experimental Physical Chemistry*, 2nd ed.; Prentice-Hall: Englewood Cliffs, NJ, 1997; Part 1.
3. Schwenz, R. W.; Polik, W. E. *J. Chem. Educ*. **1999,** *76*, 1302–1307.
4. Harris, D. C. *J. Chem. Educ*. **1998,** *75*, 119–121.
5. de Levie, R. *J. Chem. Educ*. **1999,** *76*, 1594–1598.
6. O'Neill, R. T.; Flaspohler, D. C. *J. Chem. Educ*. **1990,** *67*, 40–42.
7. Zielinski, T. J.; Allendoerfer, R. D. *J. Chem. Educ*. **1979,** *74*, 1001–1007.
8. Pursell, C. J.; Doezema, L. *J. Chem. Educ*. **1999,** *76*, 839–841.
9. Bevington, P. R.; Robinson, D. K. *Data Reduction and Error Analysis for the Physical Sciences*, 2nd ed.; McGraw-Hill: New York, 1993.
10. Hessler, J. P.; Current, D. H.; Ogren, P. J. *Comp. Phys*. **1996,** *10*, 186–199.
11. Meyer, S. L. *Data Analysis for Scientists and Engineers*; Wiley: New York, 1975.
12. Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes*; Cambridge University Press: New York, 1989.
13. Cvetanović, R. J.; Singleton, D. L.; Paraskevopoulos, G. *J. Phys. Chem*. **1979,** *83*, 50–60.
14. Huber, K. P.; Herzberg, G. *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules*; Van Nostrand Reinhold: New York, 1978; p 286.
15. Hessler, J. P. *Int. J. Chem. Kinet*. **1997,** *29*, 803–817.
16. Chong, D. P. *J. Chem. Educ*. **1994,** *71*, 489–490.
17. Williamson, J. A. *Can. J. Phys*. **1968,** *46*, 1845–1847.

Appendix

The two common ways for visualizing the meaning of the χ^2 statistic in curve fitting involve χ^2 paraboloid surfaces or $\Delta\chi^2$ ellipsoidal surfaces. Both approaches are based upon a Taylor expansion of χ^2 near its minimum:

$$\chi^2(\boldsymbol{a}) = \chi_0^2(\boldsymbol{a}^*) + \sum_n \frac{\partial \chi^2}{\partial a_n}\delta a_n + \frac{1}{2}\sum_n\sum_k \frac{\partial^2 \chi^2}{\partial a_n \partial a_k}\delta a_n \delta a_k + \ldots \quad \text{(A1)}$$

where $(\boldsymbol{a}^*)$ and $(\boldsymbol{a})$ are, respectively, vectors defined by the best-fit parameter set $\{a_1^*, a_2^* \ldots\}$ and a parameter set in the vicinity of the best-fit set, and where $\delta a_n = a_n - a_n^*$. With rearrangement and substitution, this can be written as

$$\chi^2 = \chi_0^2(\boldsymbol{a}^*) = -2(\boldsymbol{\beta})(\boldsymbol{a}) + (\boldsymbol{\delta a})[\boldsymbol{\alpha}](\boldsymbol{\delta a}) + \ldots \quad \text{(A2a)}$$

or

$$\Delta\chi^2 = \chi^2(\boldsymbol{a}) - \chi_0^2(\boldsymbol{a}^*) = -2(\boldsymbol{\beta})(\boldsymbol{\delta a}) + (\boldsymbol{\delta a})[\boldsymbol{\alpha}](\boldsymbol{\delta a}) + \ldots \quad \text{(A2b)}$$

where the elements of the $(\boldsymbol{\beta})$ vector are given by

$$\beta_n = -\frac{1}{2}\frac{\partial \chi^2}{\partial a_n} \quad \text{(A3)}$$

and the elements of the $[\boldsymbol{\alpha}]$ vector are given by

$$\alpha_{nk} = -\frac{1}{2}\frac{\partial^2 \chi^2}{\partial a_n \partial a_k} \quad \text{(A4)}$$

The $[\boldsymbol{\alpha}]$ vector is once again the curvature matrix, and its inverse is the error matrix $[\boldsymbol{\varepsilon}]$. If the first derivatives in eq A1 are taken at the minimum, then all terms in the first sum are 0. If the next term in the

expansion is kept, but higher order terms are dropped, the χ^2 surface is a *paraboloid* in the parameter space. This can be readily visualized in 3-D space for two independent parameters and the dependent variable χ^2 (*2*; *11*, Chapter 32). For the multilinear case, it is rather easy to show that all terms beyond the quadratic term are 0, and that eq A4 reduces to expression 8 for α_{nk}. For nonlinear problems, truncation after the quadratic terms is known as the "*quadratic approximation*". In nonlinear problems, eq A2b is used as a basis for an iterative approach to determining the best-fit parameter set $\{a_1^*, a_2^* \ldots\}$: starting with an initial set of parameter estimates near the best-fit values, the derivatives of eq A2b with respect to each δa_i are set equal to zero, and the resulting matrix equation is solved to determine the set of changes $\{\delta a_1, \delta a_2, \ldots\}$ needed to improve the fit (*7, 9, 10*).

At the χ^2 minimum, the ($\boldsymbol{\beta}$) vector terms of eq A2 are all zero. For a multilinear problem, eq A2b then becomes

$$\Delta\chi^2 = (\boldsymbol{\delta a})[\boldsymbol{\alpha}](\boldsymbol{\delta a}) \tag{A5}$$

and for a three-parameter problem, this translates into

$$\begin{aligned}\Delta\chi^2 = {} & \delta a_1^2\alpha_{11} + \delta a_1\delta a_2\alpha_{12} + \delta a_1\delta a_3\alpha_{13} \\ & + \delta a_2\delta a_1\alpha_{21} + \delta a_2^2\alpha_{22} + \delta a_2\delta a_3\alpha_{23} \\ & + \delta a_3\delta a_1\alpha_{31} + \delta a_3\delta a_2\alpha_{32} + \delta a_3^2\alpha_{33}\end{aligned} \tag{A6}$$

This corresponds to a family of ellipsoidal surfaces of constant $\Delta\chi^2$ in the 3-D parameter space defined by δa_1, δa_2, and δa_3. This ellipsoidal interpretation can obviously be extended to additional parameters and dimensions. Even for nonlinear problems, where terms of higher order in eq A1 are often important, the quadratic approximation serves as a useful point of comparison.

Equations A1 and A2a,b can be readily converted to an equivalent dimensionless format by defining

$$\beta'_n = -\frac{a_n^*}{2}\frac{\partial\chi^2}{\partial a_n}$$

and, as in eq 12,

$$\alpha'_{nk} = \sum_i \frac{a_n^* a_k^* X_n(x_i) X_k(x_i)}{\sigma_i^2}$$

and by replacing δa_i with $\delta a'_i = (a_i - a_i^*)/a_i^*$. $\delta a'_i$ is the familiar *relative uncertainty* in parameter a_i. With these changes, for a three-parameter multilinear problem, eq A6 becomes

$$\begin{aligned}\Delta\chi^2 = {} & (\delta a'_1)^2\alpha_{11} + \delta a'_1\delta a'_2\alpha'_{12} + \delta a'_1\delta a'_3\alpha'_{13} \\ & + \delta a'_2\delta a'_1\alpha'_{21} + (\delta a'_2)^2\alpha'_{22} + \delta a'_2\delta a'_3\alpha'_{23} \\ & + \delta a'_3\delta a'_1\alpha'_{31} + \delta a'_3\delta a'_2\alpha'_{32} + (\delta a'_3)^2\alpha'_{33}\end{aligned} \tag{A7}$$

Both two-dimensional projections of the full ellipsoidal surface and planar slices through this surface provide useful visual aids in understanding the behavior of complex fitting problems. The general ellipse equation in the *xy* plane is

$$ax^2 + by^2 + cxy = f \tag{A8}$$

and eq A7 obviously has this form when one of the parameters, say $\delta a'_3$, is set to zero. This reduced version of eq A7 was used with the appropriate curvature matrix elements to construct the intersection ellipse of Figure 8. Finding the *a*, *b*, *c*, and *f* values of eq A8 for the projection ellipse in the *xy* plane is more challenging. We first note that the full 1σ $\Delta\chi^2$ surface has a maximum extent in the *x* direction given by the coordinates

$$x_{\max} = \sqrt{3.53\varepsilon'_{xx}};\quad y = x_{\max}\frac{\mathrm{cof}(\alpha'_{xy})}{\mathrm{cof}(\alpha'_{xx})};\quad \text{and}\quad z = x_{\max}\frac{\mathrm{cof}(\alpha'_{xz})}{\mathrm{cof}(\alpha'_{xx})} \tag{A9}$$

The $x_{\max}$ and $y(x_{\max})$ values of eq A9 must also define the maximum extent of the projection ellipse in the *xy* plane. A similar pair of equations defines y_{max} and $x(y_{\max})$ for the projection ellipse. One of the constants in eq A8 may be chosen arbitrarily, so we set $a = 1$ and then look for two equations in addition to eq A8 that can be used to calculate *b*, *c* and *f*. The additional equations are obtained from the differential of eq A8: $2ax\,dx + 2by\,dy + cy\,dx + cx\,dy = 0$, and the relationships $dy_{\max}/dx = 0$ and $dx_{\max}/dy = 0$ in the *xy* projection plane.

Finally, we note that eq A9 also defines the upper confidence limit for *x*. The line from the ellipsoid center to this $\{x_{\max}, y, z\}$ point is called the confidence limit vector (*10*). This line can be extended in the opposite direction to the symmetrically located lower confidence limit point. It is interesting to notice what happens when the best-fit parameter associated with "*x*" is shifted from the origin to a new fixed value x_{new}. Geometrically, the finding new optimum *y* and *z* values involves locating the center of a thin elliptical disk slicing through the full $\Delta\chi^2$ surface at a distance x_{new} from the *yz* plane. This disk center will be on the confidence limit vector, and so the new *y* and *z* values will be given by the last two parts of eq A9 where x_{new} replaces $x_{\max}$. This result is identical to eq 16.