LETTERS TO THE EDITOR

To the Editor:

In the article titled the "Determination of Arrhenius Constants by Linear and Nonlinear Fitting," Chen and Aris (April 1992, p. 626) compared fitting the Arrhenius temperature dependence of the rate constant k by minimizing the sum of squared deviations (SSD) of ln$k$, which they called log-linear fitting, to minimizing the SSD of k, which they called nonlinear fitting, and concluded that the latter is *in general* better. Their conclusion is incorrect.

They compared the SSD of k with log-linear and nonlinear fitting and found that nonlinear fitting gave a SSD value a factor of 3.3 smaller. However, if they had compared the SSD of ln$k$ in the same fashion, they would have found that log-linear fitting gave a value of SSD of ln$k$ a factor of 4.4 smaller. They also argued that "we are interested in k, rather than in ln$k$." These are not criteria for comparing the two methods of regression. We are interested in A and E, and the choice of the method of regression *depends on the error structure of the data.*

Least-squares regression arises logically from maximum likelihood estimation of the parameters of variables with normal errors. If the variance of the errors in k is constant (constant arithmetic error variance), the maximum likelihood method yields the nonlinear fitting preferred by Chen and Aris, while if the variance in ln$k$ is constant (constant percent error variance), the maximum likelihood method yields the log-linear fitting. The choice of method, therefore, depends on the nature of the errors.

This problem can be approached somewhat more generally by assuming that the variance of measured k is proportional to the γ power of k, so that $\gamma = 0$ corresponds to constant variance in k, while $\gamma = 2$ corresponds to constant variance in ln$k$. The maximum likelihood method can then be applied to the (assumed) normal variate:

$$\frac{[k - \langle k \rangle]}{\langle k \rangle^{\gamma/2}}$$

where $\langle k \rangle$ is the expected value of k, given by $\langle k \rangle = A\exp(-E/RT)$, to estimate A, E *and* γ. This procedure has been implemented in the modeling and simulation software called SimuSolv (Steiner et al., 1986), in which γ is constrained $0<\gamma<2$. I have applied this to the data given by Chen and Aris with the result that the maximum likelihood estimates of the parameters are $\gamma = 2.00$, $A = 1.27$ E09, and $E = 13{,}007$. The data are found to be fitted best by parameters closer to those of the log-linear model.

SimuSolv also calculates all the statistics of the regression. The standard deviation of the estimate of A is 1.25 E09 (!), and that of E is 628. The correlation coefficient between the estimates of A and E is 0.9989. These are rather poor data for determining A and, for that matter, for comparing methods of regression.

Literature cited

Chen, N. H., and R. Aris, "Determination of Arrhenius Constants by Linear and Nonlinear Fitting," *AIChE J.*, **38**, 626 (1992).

Steiner, E. C., G. E. Blau, and G. L. Agin, *Introductory Guide to SimuSolv*, The Dow Chemical Co., Midland (1986).

Rane L. Curl
Dept. of Chemical Engineering
The University of Michigan
Ann Arbor, MI 48109

Reply:

Our article was based on the following:

1. The fundamental principle that nonlinear equations should be solved by the nonlinear method.
2. The simple reasoning, $2 + 3 = 5$, without considering the error in 2 and/or 3, and the reliability of 5.
3. Interest in k rather than ln(k).
4. The objective function of Eq. 4, rather than Eq. 3.

Consequently, the statistical aspect of the topic was not considered in detail and therefore was beyond the scope of our method.

Curl, however, proposed the statistical error structure approach in his letter, and concluded that the maximum likelihood method applied to his assumed "normal variate" fitted our data with the log-linear model better without showing numerical justification. He, however, neglected to compare it with the nonlinear case by the same method. Unfortunately, he forgot to point out that his result was obtained by minimizing the SSD of ln(k). As mentioned before, we should minimize the SSD of k, rather than the SSD of ln(k). To account for the experimental error, the present weighting factor method can be employed.

In addition, his comments are not very clear and need explanation and/or mathematical proofs such as:

1. "If the variance of the errors in k is constant (constant arithmetic error variance), the maximum likelihood method yields the nonlinear fitting preferred by Chen and Aris, while if the variance in ln(k) is constant (constant percent variance), the maximum likelihood method yields the log-linear fitting." These new terms need clarification with numerical illustration.
2. His assumption, "the variance of measured k is proportional to the power of k," needs numerical illustration.
3. Which procedure does he refer to by "this procedure"?
4. How can the maximum likelihood method be applied to his normal variate?
5. How can the figures 1.25 E09 and 628 be obtained?

In conclusion, I believe our proposed method is still valid.

Ning Hsing Chen
9701 Fields Road (Apt. 201)
Gaithersburg, MD 20878