

Determination of Arrhenius Constants by Linear and Nonlinear Fitting

Ning Hsing Chen

Dept. of Chemical Engineering, University of Lowell, Lowell, MA 01852

Rutherford Aris

Dept. of Chemical Engineering and Material Science, University of Minnesota, Minneapolis, MN 55455

For nearly a century, the method of evaluating the constants, the frequency factor A and the activation energy E in the Arrhenius equation,

$$k = A \exp(-E/RT), \quad (1)$$

in which k is the rate constant, and T is the absolute temperature, has been to linearize Eq. 1 by taking its logarithm and writing:

$$\ln k = \ln A - E/RT. \quad (2)$$

The method of linear least squares is then applied to this equation to minimize the sum of the squares of the difference of the logarithmic experimental and calculated values of the rate constants:

$$S_L = \sum_{i=1}^n (\ln k_i - \ln A + E/RT_i)^2 = \sum_{i=1}^n [\ln(k_i/\hat{k}_i)]^2. \quad (3)$$

The set of data in (T_i, k_i) , $i = 1, 2, \dots, n$ and \hat{k}_i is the value of $k(T_i)$ calculated by Eq. 1.

In general, however, we are interested in k , rather than $\ln k$, so that the appropriate deviations, whose sum of squares is to be minimized, would seem to be $k_i - A \exp(-E/RT_i)$. We therefore define the sum

$$S_N = \sum_{i=1}^n (k_i - A \exp(-E/RT_i))^2 = \sum_{i=1}^n (k_i - \hat{k}_i)^2 \quad (4)$$

which can be minimized by any one of a number of two-dimensional search routines or by quasi-linearization.

We may first observe, however, that A is large—often of the order of 10^7 to 10^{14} , that the exponential is correspondingly small, and that it is easy to get into the situation of calculating small differences of large numbers, never one that makes for accuracy. We therefore reparametrize, taking

$$a = \ln A, \quad (5)$$

and rescale, taking,

$$b = E/1,000. \quad (6)$$

We will also take reduced temperatures:

$$t_i = RT_i/1,000. \quad (7)$$

Thus,

$$S_L = \sum_{i=1}^n (\ln k_i - a + b/t_i)^2 \quad (8)$$

and

$$S_N = \sum_{i=1}^n [k_i - \exp(a - b/t_i)]^2 \quad (9)$$

The minimization of the linear form S_L is well-known (for example, Neter et al., 1989) and the “best” values of a and b (in the least squares sense) are:

$$a = \frac{pl - qm}{nl - m^2}, \quad b = \frac{pm - qn}{nl - m^2} \quad (10)$$

where

$$m = \sum 1/t_i, \quad l = \sum 1/t_i^2, \quad p = \sum \ln k_i, \quad q = \sum (\ln k_i)/t_i. \quad (11)$$

Correspondence concerning this work should be addressed to N. H. Chen at Washington Tower, Suite 201, 9701 Fields Road, Gaithersburg, MD 20878.

The minimization of the nonlinear form, S_N , cannot be achieved analytically, but involves numerical iteration. However, we notice that it is quadratic in $e^a = A$ since

$$S_N = \sum k_i^2 - 2e^a \sum k_i e_i + e^{2a} \sum e_i^2, \quad e_i = \exp - b/t_i, \quad (12)$$

and hence, for any fixed b , S_N has a minimum

$$S(b) = \sum k_i^2 - (\sum k_i e_i)^2 / (\sum e_i^2) \quad (13)$$

when

$$e^a = (\sum k_i e_i) / (\sum e_i^2). \quad (14)$$

This reduces the search for the minimum to a one-dimensional search, a much easier task than the search in the a, b -plane. Indeed, the surface $S_N(a, b)$ is often a narrow canyon with steep sides and a long very shallow bottom whose contours are almost impossible to draw. Without reparametrization, that is, in the A, E -plane, the canyon is banana-shaped and would serve as a good test function in the manner of Rosenbrock's canyon (Rosenbrock and Storey, 1966, p. 70). However, one-dimensional search for the minimum can be accomplished in various ways (Chen, 1977; Rosenbrock and Storey, 1966; Chambers, 1973; Beveridge and Schechter, 1970; Avriel, 1976); otherwise, it can be found as the solution of $S'(b) = 0$: minimizing b is the root of

$$(\sum e_i^2) (\sum k_i e_i / t_i) = (\sum e_i^2 / t_i) (\sum k_i e_i) \quad (15)$$

which can be solved by Newton-Raphson iteration. The linear estimate from Eq. 10 makes a good starting point.

Illustration

A useful example is provided by the reaction of ethanol and acetic acid catalyzed by a cation exchange resin which was studied by Saleton and White (Ex. 3.6-1 of Chen, 1983). Their data are shown in the second and third columns of Table 1, and the scaled temperature is in the fourth. Applying the logarithmic linearization gives the fitted parameters shown in the first column of the lower part of the table, above which are the calculated best-fit values of k_i and the differences from the observed values. These differences are used to calculate S_N , the sum of their squares for comparison with the nonlinear fit. The latter, which is given in the second column of the lower

Table 1. Experimental Data and Comparison of Methods

<i>i</i>	Data			Log-linear		Nonlinear	
	$T_i^\circ\text{C}$	k_i	t_i	k_{iL}	$k_i - k_{iL}$	k_{iN}	$k_i - k_{iN}$
1	30	0.5	0.6023	0.5370	-0.0370	0.6812	-0.1812
2	40	1.1	0.6222	1.0690	+0.0310	1.2434	-0.1434
3	50	2.2	0.6421	2.0343	+0.1607	2.1866	0.0134
4	60	4.0	0.6619	3.7421	+0.2579	3.7173	0.2827
5	70	6.0	0.6818	6.6282	-0.6282	6.1271	-0.1271
Fixed Constants				S_N	0.4893		0.1496
				a	20.9397		18.46345
				b	12.9872		11.3525
				A	1.2416×10^9		1.0437×10^8
				E	12987		11352

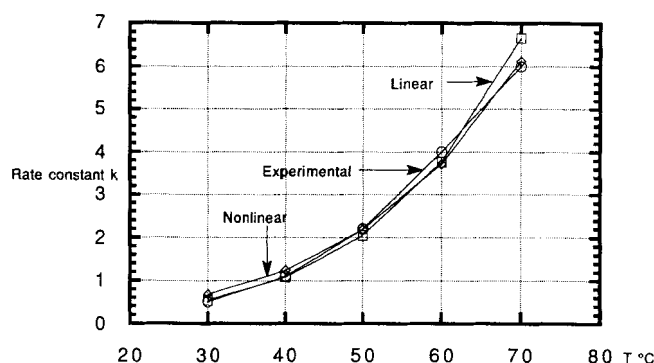


Figure 1. Comparison of methods.

part of the table, clearly shows that the two sets of parameters differ considerably— E by more than 1,600 calories (some 12.5%) and A by an order of magnitude. The sum of the squares is diminished by a factor of 3.27. The data and the two fittings are shown in Figure 1.

How can the estimates be compared? As was mentioned, the surface $S_N(A, E)$ has a steep-sided canyon with a very shallow bottom. We obtained the position of the minimum, $A = 1.0437 \times 10^8$, $E = 11,352$, by solving for $S'(b) = 0$ to give the lowest point of the floor of the canyon. The contour of $S_N = 0.15$, a value slightly greater than the minimum 0.1496, stretches from $E = 11,289$ to $E = 11,416$; that of $S_N = 0.175$ from $E = 10,827$ to $11,898$; and that of $S_N = 0.2$ from $E = 10,617$ to $12,127$. This shows how difficult it is to set a stopping criterion for a search method, since the 0.15 contour, differing by only 0.25% from the minimum covers a range of 1.1% of E . The Newton-Raphson iteration, starting from $E = 10,000$ ($b = 10$), converged in five steps (11,183, 11,349, 11,352, 11,352.505, and 11,352.505). The formula for this iteration is:

$$b_{k+1} = b_k + \frac{(\sum e_i^2) (\sum k_i e_i / t_i) - (\sum e_i^2 / t_i) (\sum k_i e_i)}{(\sum e_i^2 / t_i) (\sum k_i e_i / t_i) + (\sum e_i^2) (\sum k_i e_i / t_i^2) - 2(\sum e_i^2 / t_i^2) (\sum k_i e_i)} \quad (16)$$

where the k th iterate b_k is used in evaluating the $e_i = \exp - b/t_i$.

In linear regression $\sum (k_i - \bar{k})^2 / (n - 2)$ is an unbiased estimator of the variance of observed deviations. If the same estimator is used in nonlinear fitting, the two cases give variances of 0.16 (L) and 0.05 (N), the nonlinear being the better fit.

Conclusion

Accurate values of rate constants are essential to applied chemical kinetics. The proposed method of evaluating the Arrhenius constants from experimental data is more reasonable and more accurate than the method of logarithmic linearization and is recommended for use.

Notation

$a = \ln A$
 $A =$ Arrhenius pre-exponential factor
 $b = E/1,000$
 $b_k =$ k th iterate for minimizing value of b
 $E =$ activation energy
 $e_i = \exp -E/RT_i = \exp -b/t_i = \hat{k}_i/A$
 $k_i =$ i th observed value of k
 $\hat{k}_i = Ae^{-E/RT_i} = \exp(a - b/t_i)$, calculated $k(T_i)$
 $l = \Sigma 1/t_i^2$
 $m = \Sigma 1/t_i$
 $n =$ number of data points
 $p = \Sigma \ln k_i$
 $q = \Sigma (\ln k_i)/t_i$
 $S(b) = \Sigma k_i^2 - (\Sigma k_i e_i)^2 / \Sigma e_i^2$
 $S_L = \Sigma (\ln k_i - a + b/t_i)^2$
 $S_N = \Sigma (k_i - \hat{k}_i)^2$
 $T_i =$ temperature of i th observation
 $t_i = RT_i/1,000$, scaled temperature

Subscripts

$i =$ i th observation, $i = 1, \dots, n$

$L =$ logarithmic-linearization case
 $N =$ nonlinear case

Literature Cited

- Avriel, M., *Nonlinear Programming Analysis and Methods*, Prentice-Hall, Englewood Cliffs, NJ (1976).
Beveridge, G. S. G., and R. S. Schechter, *Optimization: Theory and Practice*, McGraw Hill, New York (1970).
Chambers, J., "Fitting Nonlinear Models: Numerical Techniques," *Biometrics*, **60**, 1 (1973).
Chen, N. H., *New Mathematics for Chemical Engineers*, Hoover, Taiwan (1977).
Chen, N. H., *Process Reactor Design*, Allyn and Bacon, Newton, MA (1983).
Neter, J., W. Wasserman, and M. H. Kutner, *Applied Linear Regression Models*, Irwin, Boston (1989).
Rattowsky, D. A., *Nonlinear Regression Modeling*, Dekker, New York (1983).
Rosenbrock, H. H., and C. Storey, *Computational Techniques for Chemical Engineers*, Pergamon, Oxford (1966).

Manuscript received Oct. 31, 1991, and revision received Feb. 10, 1992