Journal of Thermal Analysis, Vol. 38 (1992) 1283-1291

CALCULATION TECHNIQUES FOR SOLVING NON-ISOTHERMAL KINETIC PROBLEMS

V. A. Sipachev and I. V. Arkhangel'skii

DEPARTMENT OF INORGANIC CHEMISTRY, MOSCOW STATE UNIVERSITY, MOSCOW, RUSSIA

A calculation technique based on the SVD algorithm is suggested for solving non-isothermal kinetics problems. The uncertainties in the sought parameter values are obtained by superimposing random (Gauss) noise on experimental dependences.

Keywords: non-isothermal kinetics techniques

The so-called non-isothermal kinetics techniques are used extensively to determine the kinetic parameters of heterogeneous reactions, which is explained by the apparent simplicity of measurements and of the processing of measurement results. Many authors, are, however, skeptic about the utility of these techniques for studying reaction mechanisms (see, e.g. [1, 2]). There are various reasons for this attitude, both of general nature [3] and based on the consideration of thermoanalytical experimental characteristics [4]. On the other hand, little attention has been given to the calculation procedures for treating thermoanalytical data. Thus, very often, the function $f(\alpha)$ used to describe kinetic curves is chosen solely using the linearization criterion. However, because of a strong Arrhenius dependence, the possibility of the linearization depends only weakly on the type of that function [5]. The influence of measurement errors and the calculation procedure on the quality of the solution has practically been ignored in the literature.

In this work, we examine certain techniques that can be used to determine the kinetic parameters from non-isothermal thermogravimetric experiment data.

The classical reaction of thermal dehydration of calcium oxalate monohydrate, $CaC_2O_4 \cdot H_2O$, has been studied. The kinetic model has been chosen in the form

$$\frac{d\alpha}{dT} = \frac{A}{q} \cdot \exp\left(-\frac{E}{RT}\right) \alpha^n (1-\alpha)^m \qquad (1)$$

John Wiley & Sons, Limited, Chichester
Akadémiai Kiadó, Budapest

where α is the experimental conversion. T the temperature, q the heating rate, $q = dT/dt$ =const, A the preexponential factor, E the activation energy, and m and n are kinetic equation parameters. There are certain reasons for the choice of the kinetic equation in the form $f(\alpha) = \alpha^n (1-\alpha)^m$. First, this empirical equation has successfully been used to describe various heterogeneous processes [6]. Secondly, the exponents m and n in Eq. (1) are determined by the reaction mechanism. Extracting their values from the experiment, therefore, enables one to draw conclusions concerning the kinetic model without recourse to *a priori* considerations.

Equation (1) can easily be linearized as follows:

$$\ln \frac{d\alpha}{dT} = \ln \frac{A}{q} - \frac{E}{RT} + n \cdot \ln(\alpha) + m \cdot \ln(1-\alpha) \tag{2}$$

that is it reduces to a linear LSQ problem

$$\underline{U}x = b \tag{3}$$

Here and below, the notation is as follows: $\underline{U}$ is the matrix of the coefficients of Eq. (1) ($\underline{a}_{i1} = 1/q$, $\underline{a}_{i2} = -1/T_i$, $\underline{a}_{i3} = \ln(\alpha_i)$ and $\underline{a}_{i4} = \ln(1 - \underline{a}_i)$, i is the index of a kinetic curve point), x is the vector of the sought parameters ($x_1 = \ln(A)$, $x_2 = E/R$, $x_3 = n$ and $x_4 = m$) and b is the vector of the quantities $b_i = \ln(d\alpha/dT|_{\alpha=\alpha_i})$.. The underlined letters are used to denote matrices and their elements, and italics to denote vectors and vector components. The dimensionality of the matrix $\underline{U}(1 \times 4)$ and $b(l)$ is determined by the number of measurements along the kinetic curve.

Solving problem (3) involves certain difficulties, for the quantity $1/T$ varies only insignificantly within the temperature range of a chemical transformation. The first and second columns of the matrix $\underline{U}$ are therefore almost identical (to within a constant factor). For this reason, multiplying both sides of Eq. (3) by the matrix $\underline{U}^T$ on the left (the usual scheme for solving linear LSQ problems) yields an almost degenerate set of equations:

$$\underline{U}^T \underline{U}x = \underline{U}^T b \tag{4}$$

Actually, the transition to problem (4) alters the situation for the worse substantially even compared with Eq. (3) where as mentioned above, we have an almost linear dependence between matrix columns. In fact, the relative error in the sought parameters can be estimated as

$$\frac{|\Delta x|}{|x|} \leq \text{cond}(\underline{U}) \cdot \frac{|\Delta b|}{|b|} \tag{5}$$

where Δb is the vector of errors on the right hand side of Eq. (3), and the symbol $|\cdot|$ denotes the norm (length) of a vector. The errors $\Delta b_i = \Delta(\ln[\mathrm{d}\alpha/\mathrm{d}T|_{\alpha=\alpha_i}])$, are derivative from conversion, α_i, measurement errors and errors involved in the analytical approximation to the observed dependence $\alpha(T)$. The quantity $\mathrm{cond}(\underline{U})$ is the conditionality of matrix $\underline{U}$ (Matrix $\underline{U}$ elements also include errors and therefore, Eq. (5) in fact underestimates $|\Delta x|$). If we use the Euclidean norm, $|x| = (\sum x_i^2)^{1/2}$, $\mathrm{cond}(\underline{U})$ is merely equal to the ratio of the largest matrix $\underline{U}$ singular value to the smallest one (see below). If matrix $\underline{U}$ has an incomplete rank (the matrix is degenerate) one or more of its singular values reduce to zero, and problem (3) becomes undetermined along the corresponding directions, that is vectors y such that $\underline{U}y \equiv 0$. Although in actual kinetic problems matrix $\underline{U}$ is not degenerate, its conditionality amounts to several hundred, and calculating kinetic parameters with an accuracy of the order of 10% requires that uncertainties in vector b components be not larger than 0.05%. To develop a clearer idea of the situations, imagine an ellipsoid of solutions, that is a region in the space of parameters whose all points are compatible with experimental data to within measurement errors. The longest to the shortest principal axis ratio is just equal to the matrix conditionality. The larger this value (the greater the matrix degeneracy) the more prolate is the ellipsoid which in the limit of $\mathrm{cond}(\underline{U}) \to \infty$ becomes an infinite cylinder. $\mathrm{Cond}(\underline{U}^T\underline{U}) = (\mathrm{cond}(\underline{U}))^2 \cong 10^5 - 10^6$ and therefore, using scheme (4) appears to be not the best way to solve problem (3), for as is clear from Eq. (5) even computer arithmetics rounding errors can then noticeably affect the calculation results.

The most effective technique for solving poorly conditioned linear LSQ problems is to use a factorization of the type

$$\underline{U} = \underline{uSB} \tag{6}$$

known as SVD (singular value decomposition) [7]. Here $\underline{u}$ and $\underline{B}$ are unitary matrices with dimensions $l \times l$ and $k \times k$ (in problem (3), $k = 4$) and $\underline{S}$ is an $l \times k$ matrix with nonzero elements (singular values of matrix $\underline{U}$ in the principal diagonal. The LSQ problem reduces to minimizing the discrepancy

$$|r| = |\underline{U}x - b| \tag{7}$$

Using factorization (6) enables Eq. (7) to be rewritten as

$$|\underline{u}^T r| = |\underline{u}^T\underline{U}x - \underline{u}^T b| = |\underline{SB}x - \underline{u}^T b| \tag{8}$$

for $\underline{u}^T\underline{u} = \underline{J}$, a unit matrix, by definition. Actually, the problems of minimizing the discrepancies (7) and (8) are equivalent: multiplying vectors by unitary matrices does not change their length. Moreover, this operation does not affect the problem conditionality, for the conditionality of any unitary matrix is equal to unity. The

minimization problem now has a trivial solution. As k upper rows are only nonzero in matrix $\underline{S}$ the vector $\underline{S}x$ has only k, that is four, nonzero components. Thus the problem is reduced to a set of four simultaneous equations. Using the notation $\underline{u}^T b = d$ yields

$$|r| = |\underline{u}^T r| = \left(\sum_{i=5}^{l} d_i^2\right)^{1/2} \tag{9}$$

and, provided the initial matrix $\underline{U}$ is not degenerate (there is no zero singular numbers),

$$x = \underline{V}^T \underline{S}^{-1} \underline{u}^T b \tag{10}$$

Here $\underline{S}^{-1}$ is an $l \times k$ matrix whose main diagonal contains reciprocals of singular values and whose all off-diagonal elements are zero. If one of the singular values, eg. $\underline{s}_i$, is equal to zero the corresponding matrix $\underline{S}^{-1}$ element is also set to zero to find the shortest vector x minimizing discrepancies (7) [7]. The general solution can then be obtained as the sum of this vector and the i^{th} row of matrix $\underline{V}$ taken with an indefinite (the ellipsoid of solutions turns into a cylinder). In actual calculations, singular values are set to zero if they are smaller than a threshold value depending on the accuracy of the determination of vector b components.

The advantages of the SVD technique are as follows.

1. It provides an easy means of estimating problem sensitivity to the accuracy of measurements by calculating the conditionality of matrix $\underline{U}$ equal to the ratio of its largest singular value to the smallest one,

$$\text{cond}(\underline{U}) = \underline{s}_{\max}/\underline{s}_{\min} \tag{11}$$

This is difficult to do using the Gauss algorithm (the $\underline{Lu}$ factorization of matrix $\underline{U}^T\underline{U}$).

2. Singular value decomposition enables a general solution to be found when matrix $\underline{U}$ is degenerate or nearly degenerate:

$$x = \beta(\underline{v}^i)^T + \underline{V}^T \underline{S}^{-1} \underline{u}^T b \tag{12}$$

where β is an indefinite factor, i is the number of the zero (or nearly zero) singular value, and $\underline{v}^i$ is the i^{th} row of matrix $\underline{V}$. With β set to zero we obtain the shortest vector x solving an LSQ problem (3). The Gauss algorithm is unfit to handle degenerate problems.

3. Lastly, the SVD technique makes it possible to characterize the ellipsoid of solutions in detail. For this purpose, the initial problem (3) should first be rescaled using a diagonal matrix $\underline{W}$ with the elements $\underline{w}_{ii}$ equal to the reciprocals of

the root mean square deviations for the corresponding vector b components, $\underline{w}_{ii} = 1/\sigma_{b_i}$:

$$\underline{WU}x = \underline{W}b \tag{13}$$

The SV decomposition of matrix $\underline{WU}$ results in matrix $\underline{V}$ whose rows determine the directions of the principal axes of the ellipsoid of solutions. The semiaxes lengths are equal to the reciprocals of the corresponding singular values.

The shortcoming of the SVD procedure is an increase in computer time expenditures compared with the standard technique. As far as problem (3) is concerned this increase is almost unnoticeable.

One more problem is that of the determination of uncertainties in the parameters calculated from experimental data. As the hypothesis of a Gauss or at least symmetrical distribution can by no means be accepted for the logarithm of the quantity $\ln(\mathrm{d}\alpha/\mathrm{d}T)$ (Eq. (2)), the standard procedure for calculating σ_{LSQ} values can only give very rough estimates of confidence intervals. Actually, errors involved in matrix $\underline{U}$ elements should also be taken into account

$$\Delta x \cong \underline{U}^{-1}\Delta b + \Delta(\underline{U}^{-1})\, b \tag{14}$$

Using the notation $\underline{a}_{ij}$ and $\underline{d}_{ij}$ for the matrix $\underline{U}^{-1}$ and $\Delta(\underline{U}^{-1})$ elements, respectively $\Delta(\underline{U}^{-1})$ can be calculated as $\Delta(\underline{U}^{-1}) = (\underline{U} + \Delta\underline{U})^{-1} - \underline{U}^{-1} \approx \underline{U}^{-1}\Delta\underline{U}\underline{U}^{-1}$ see Eq. (19), the i^{th} component of vector Δx can be written as

$$\Delta x_i = \sum_{j=1}^{l} (\underline{a}'_{ij}\Delta b_j + \underline{d}_{ij} b_j) \tag{15}$$

and

$$(\Delta x_i)^2 = \sum_{j=1}^{l} (\underline{a}'_{ij}\Delta b_j)^2 + 2 \sum_{k>j=1}^{j=l-1} \underline{a}'_{ij}\underline{a}'_{ik}\Delta b_j \Delta b_k + \left(\sum_{j=1}^{l} \underline{d}_{ij} b_j\right)^2$$

$$+ 2\left(\sum_{j=1}^{l} \underline{a}_{ij}\Delta b_j\right)\left(\sum_{k=1}^{l} \underline{d}_{ik} b_k\right) \tag{16}$$

Hence the variance is

$$Dx_i = \sum_{j=1}^{l} (\underline{a}'_{ij})^2 Db_j + 2 \sum_{k>j=1}^{j=l-1} \underline{a}_{ij}\underline{a}_{ik} M[\Delta b_j \Delta b_i] + M\left[\left(\sum_{j=1}^{l} \underline{d}_{ij} b_j\right)^2\right]$$

$$+ 2 \sum_{j,k=1}^{l} \underline{a}'_{ij} b_k M[\Delta b_j \underline{d}_{ik}] \tag{17}$$

where $M[x]$ is the mathematical expectation for x. If the problem is scaled according to Eq. (13), we have $Db_j = 1$ for all j. Assuming the distributions for Δb_j and $\underline{d}_{ij}$ to be symmetrical and statistically independent yields $M[\Delta b_j \Delta b_k] = \underline{d}_{jk}$ (the Kronecker symbol) and $M[\Delta b_j \underline{d}_{ik}] = 0$ for all j, k. The third term in Eq. (17) can easily be calculated using the approximation to $\underline{d}_{ij}$ given above.

Below, confidence intervals calculated using Eq. (14) are compared with more accurate Monte–Carlo estimates.

This technique is based on superimposing random Gauss noise on measurement results (experimental conversion values). The half-width of the Gauss distribution can be determined experimentally. Using this approach implies generating a desired number of 'experimental' kinetic curves that are treated as described above to obtain a series of parameter sets. We thus can determine not only the mathematical expectation values for the squares of parameter deviations from their means, $M[\Delta x_i^2]$ but also the coefficients of skewness and excess $\chi_{1i} = M[\Delta x_i^3]/(M[\Delta x_i^2])^{3/2}$ and $\chi_{2i} = M[\Delta x_i^4]/(M[\Delta x_i^2])^2 - 3$; $M[\Delta x_i^n] = \sum_{j=1}^{N} \Delta x_j^n / N$, where N is the number of generated kinetic curves), that is we can fully characterize the distribution functions for the sought parameters. The results cease to vary after generating some 1000 curves. The numerical data cited below have been obtained with $N = 5000$.

Applying the SVD procedure to treat each generated curve would require too large computation time expenditures. However as $|\Delta \underline{U}| << |\underline{U}|$ we can use the approximation

$$(\underline{U} + \Delta \underline{U})^{-1} = (\underline{J} + \underline{U}^{-1}\Delta \underline{U})^{-1}\underline{U}^{-1} \approx [\underline{J} - \underline{U}^{-1}\Delta \underline{U} + (\underline{U}^{-1}\Delta \underline{U})^2]\,\underline{U}^{-1} \tag{18}$$

whence

$$\Delta x = [\underline{J} - \underline{U}^{-1}\Delta \underline{U} + (\underline{U}^{-1}\Delta \underline{U})^2]\,\underline{U}^{-1}(b + \Delta b) - \underline{U}^{-1}b \tag{19}$$

SV decomposition can therefore be performed but once to calculate the inverse (or pseudoinverse [7]) matrix $\underline{U}^{-1} = \underline{V}^T \underline{S}^{-1} \underline{u}^T$ (see above). As far as cubic spline and reaction rate calculations, random noise generation, and matrix multiplications are concerned, these operations do not take much time.

One more point should be mentioned. Equations (3) include conversion and conversion rate values as independent quantities, although in fact, they are not so ($\alpha(t) = \int_0^t (d\alpha/dt)\,d\tau$). The parameters obtained by solving LSQ problem (3) should therefore be verified by a numerical integration of Eq. (1) and a comparison of the calculated and experimental dependences $\alpha(T)$.

The calculation scheme described above was realized on DVK–3M (a Soviet analog of PDP–11) and IBM AT microcomputers. The algorithms of key procedures were borrowed from the book by Forsythe, Malcolm, and Moler [7]. At the first stage, experimental conversions, $\alpha_{exp.}$, are introduced and the experimental $\alpha(T)$ dependence is approximated by a cubic spline for calculating the derivatives, $d\alpha/dT|_{\alpha=\alpha_i}$, analytically. The α_i, $d\alpha/dT|_{\alpha=\alpha_i}$, and T_i values are used to calculate matrix $\underline{U}$ and vector b components (Eq. (3)). After introducing weight factors (the transition to Eq. (13)) the problem is solved using the SVD procedure and the σ_{LSQ} quantities are calculated (Eq. (14)). Next a certain number of kinetic curves are generated by superimposing random noise on the experimental curve to obtain reliable estimates of confidence intervals for the sought parameters. Lastly, the parameters are substituted into Eq. (1) and numerical integration is performed to determine the $\alpha_{theor.}$ values. At the stage of solving Eq. (13), certain parameters can be fixed which makes it possible to test various kinetic models.

The weight factors, Eq. (13), should be chosen especially carefully, for they strongly affect both parameter and confidence interval values. The weight factors are determined by the variance of conversion measurements, $D\alpha$, which is instrument dependent and should be investigated very thoroughly. We have found that $D\alpha$ varies with the degree of conversion and can be described by the sum of a sine and Gauss functions, the latter with a maximum near the highest reaction rate. The problem of measurements errors in the thermogravimetric experiment will be discussed in more detail elsewhere.

To exemplify the procedure, let us consider thermogravimetric data on $CaC_2O_4 \cdot H_2O$ obtained with the help of an OD–103 instrument. The input consisted of conversions measured at ten points in the temperature range of 348 to 433 K. In spite of the comparatively low temperatures, the conditionality of the problem was cond($\underline{U}$) =725 (at higher temperatures, the quantity $1/T$ varies within a still narrower range; thus for a similar process at T_{av} of 1000 K, the conditionality increases to 1500). Clearly, the determination of the kinetic parameters from this experiment can hardly make sense. In fact, a formal solution gave meaningless results: $\ln A = -54.9$, $E/R = -19839$, $m = 0.162$, and $n = 2.574$. The general solution was obtained by setting the smallest singular value to zero: $\ln A = -3.067 + \beta(-0.105)$, $E/R = 1287.41 + \beta(-42.843)$, $m = 0.604 + \beta(-9 \cdot 10^{-4})$ and $n = 0.741 + \beta(3.7 \cdot 10^{-3})$. The confidence intervals were gigantic ($\pm 6.023 \cdot 10^4$ for E/R. Attempts to improve the situation by the introduction of more measurement data corresponding to intermediate kinetic curve points did not give the desired result, for this only added dependent equations to system (3), and the conditionality remained practically the same when the number of experimental points was increased twofold.

Varying the heating rate rotates the ellipsoid of solutions. Combined treatment of data obtained for the same substance at various heating rates yields results that are determined by an intersection of ellipsoids of solution for individual experi-

ments. As a result, the conditionality decreases substantially, and far narrower confidence intervals are obtained. Applying the calculation procedure described above to three kinetic curves obtained at heating rates of 0.028, 0.076, and 0.167 deg/s yields cond($\underline{U}$) = 85, $\ln A = 18.9 \pm 3.1$, $E/R = 10062 \pm 1180$, $m = 0.981 \pm 0.405$, and $n = 0.154 \pm 0.427$. Typical coefficient of skewness and excess values are $\chi_1 = -1-2$ and $\chi_2 = 0.7-4.5$. The new parameter values correspond approximately to the general solution given above with $\beta \approx -205$, which shows that solving even very poorly conditioned problems makes sense if the results obtained are regarded properly.

The Monte–Carlo uncertainties have been given above. The corresponding σ_{LSQ} values are 5.9, 2490, 0.243, and 0.353, respectively. One can see that they give only a vague idea of the actual situation.

It is also clear that the results for the dehydration of $CaC_2O_4 \cdot H_2O$ are close to the kinetic model with $m = 1$ and $n = 0$. Assuming this model yields $\ln A = 21.2 \pm 1.8$ and $E/R = 11057 \pm 738$. The quality of this solution (the agreement with the experiment) is practically the same as with the parameter values given above. The root mean square deviation of the numerical integration results from experimental conversions is equal to 0.0144, the mean standard deviation of measurement data being 0.0261.

It seems likely that treating even a single kinetic curve can be justified if a variable heating rate is used. Unfortunately, the advantages of the techniques based on heating rate variations are only limited, for at high rates, instrumental function errors increase sharply. Besides, the reaction mechanism can change with the rate of heating. Instrumental function errors also make it impossible to resolve the problems of non-isothermal kinetics by increasing the precision of measurements.

To sum up, the choice of the calculation procedure is very important for handling non-isothermal kinetics problems. The data cited above demonstrate that it is practically impossible to obtain reliable estimates of kinetic parameters form a single thermogravimetric curve. Nevertheless the problem is not quite hopeless, if only the reaction mechanism does not change within a reasonable range of heating rate variations.

We hope that this work will to a certain extent reduce skepticism towards non-isothermal kinetics techniques.

References

1 J. Šestak and J. Kratochvil, Anal. Chem., 5 (1973) 193.
2 A. Ferre, M. Murat and C. Comel, J. Thermal Anal., 12 (1977) 429.
3 J. R. MacCallum and J. Tanner, Nature, 225 (1970) 1127.
4 J. Meindl, I. V. Arkhangel'skii and N. A. Chernova, J. Thermal Anal., 20 (1981) 39.
5 I. V. Arkhangel'skii, L. N. Komissarova and N. A. Chernova, J. Thermal Anal., 18 (1986) 347.

6 J. Šestak, Thermophysical properties of solids, Academia, Prague 1984.
7 G. E. Forsythe, M. A. Malcolm and C. B. Moler, Computer Methods for Mathematical Computations, Prentice-Hall, Inc. Englewood Cliffs, N. J. 1977.

Zusammenfassung — Ein auf dem SVD-Algorithmus beruhendes Rechenverfahren zur Lösung nichtisothermer kinetischer Probleme wir vorgeschlagen. Die Unbestimmtheit im Rauschparameterwert wurde in Abhängigkeit von den experimentellen Bedingungen durch die Überlagerung eines Random- (Gaußschen) Rauschens ermittelt.