JOURNAL OF COMPUTATIONAL PHYSICS 18, 326–341 (1975)

The Equivalence and Numerical Solution of Regression, Optimal Policy, and Boundary Value Problems Involving Differential Equations

WALTER J. WNEK AND WILLARD R. HAAS

IIT Research Institute, 10 West 35th Street, Chicago, Illinois 60616

Received May 6, 1974; revised February 10, 1975

It is shown that regression, optimal policy, and boundary value problems involving differential equations are mathematically equivalent and that the latter two are best solved numerically as regression problems by means of transformations. This technique results in a more efficient calculational procedure than is currently available. A statistical analysis is also presented which serves as a tool in deciding the appropriateness of a proposed model.

INTRODUCTION

Three problems that are commonly encountered in engineering analysis are regression, optimal policy, and boundary value problems involving differential equations. In regression problems, the values of parameters in a set of differential equations are to be estimated from experimental data. In optimal policy problems, a given quantity is to be optimized by determining the control variables where this quantity also depends on variables defined by a set of differential equations. In boundary value problems, a set of differential equations is to be solved subject to specification of the dependent variables at various values of the independent variable. In general, the solution of these problems requires numerical techniques. In this paper, it is shown that these three problems are mathematically equivalent and that the latter two can be formulated as a regression problem which can be solved by a more efficient numerical technique than is currently used. A statistical analysis is also presented which aids in deciding whether a given set of differential equations and its parameters constitute an appropriate model with respect to the data.

FORMULATION OF THE REGRESSION PROBLEM

It is assumed that the physical system is described by a set of first-order nonlinear ordinary differential equations of the form

$$\frac{dY_i}{d\tau} = f_i(Y_1, Y_2, \ldots, Y_N, k_1, k_2, \ldots, k_M, \tau), \qquad i = 1, 2, \ldots, N, \tag{1}$$

subject to the initial conditions

$$Y_i(0) = Y_{0i}, \qquad i = 1, 2,..., N, \tag{2}$$

where $Y_1, Y_2,..., Y_N$ = the dependent variables; τ = the independent variable; $f_1, f_2,..., f_N$ = arbitrary functions, $k_1, k_2,..., k_M$ = parameters characteristic of the system.

Since an nth-order differential equation can be reduced to a system of n first-order equations, this formulation is completely general.

The problem is to estimate the values of the parameters $k_1, k_2,..., k_M$ from data on the system which give the values of $Y_1, Y_2,..., Y_N$ at various values of $\tau(\tau_1, \tau_2,..., \tau_j,..., \tau_{N_D})$ when the set of differential equations and initial conditions are known. The criterion for selecting the values of these parameters is that the weighted sum of the square of the difference between the data values, $Y_i^D(\tau_j)$, and the corresponding computed values, $Y_i(\tau_j)$,

$$S^2 = \sum_{j=1}^{N_D} \sum_{i=1}^{N} w_{ij}[Y_i^D(\tau_j) - Y_i(k_1, k_2,..., k_M, \tau_j)]^2, \tag{3}$$

is a minimum. w_{ij} is the weighting factor associated with the data point $Y_i^D(\tau_j)$ and is usually taken as the reciprocal of the variance of that point.

A necessary condition for S^2 to be a minimum is that its derivatives with respect to $k_1, k_2,..., k_M$ are all zero, which requires

$$\frac{\partial S^2}{\partial k_l} = -2 \sum_{j=1}^{N_D} \sum_{i=1}^{N} w_{ij}[Y_i^D(\tau_j) - Y_i(k_1, k_2,..., k_M, \tau_j)] \cdot \frac{\partial Y_i(k_1, k_2,..., k_M, \tau_j)}{\partial k_l} = 0, \qquad l = 1, 2,..., M. \tag{4}$$

The term $\partial Y_i/\partial k_l$ may be evaluated by noting that

$$\frac{\partial}{\partial \tau}\left(\frac{\partial Y_i}{\partial k_l}\right) = \frac{\partial}{\partial k_l}\left(\frac{\partial Y_i}{\partial \tau}\right) = \frac{\partial f_i}{\partial k_l}, \tag{5}$$

which upon integration with respect to τ gives

$$\frac{\partial Y_i}{\partial k_l}\bigg|_0^{\tau_j} = \int_0^{\tau_j} \frac{\partial f_i}{\partial k_l}\, d\tau,$$

or since $(\partial Y_i/\partial k_l)|_{\tau=0} = 0$,

$$\frac{\partial Y_i(k_1, k_2,..., k_M, \tau_j)}{\partial k_l} = \int_0^{\tau_j} \frac{\partial f_i(Y_1, Y_2,..., Y_N, k_1, k_2,..., k_M, \tau)}{\partial k_l}\, d\tau. \tag{6}$$

This eliminates the need to integrate the differential equations with each parameter slightly perturbed in turn in order to calculate these derivatives as done by other investigators [1]. This result has also been noted by Snow [15] and Eakman [4]. It should be pointed out that $\partial f_i/\partial k_l$ is defined as

$$\left(\frac{\partial f_i}{\partial k_l}\right)_{k_j \neq k_l} = \left(\frac{\partial f_i}{\partial k_l}\right)_{y_i, k_j \neq k_l} + \sum_{j=1}^{N} \left(\frac{\partial f_i}{\partial y_j}\right)_{y_i \neq y_j, k_n} \left(\frac{\partial y_j}{\partial k_l}\right), \tag{7}$$

so that according to Eq. (5) the $\partial f_i/\partial k_l$'s are given by a set of ordinary differential equations which could be solved along with Eq. (1); however, as indicated by Eakman [4], the summation term represents second-order effects. Since very accurate values of $\partial f_i/\partial k_l$ are not required in the iteration scheme, the summation term may be neglected, and Eq. (6) integrated by a simple technique such as Euler's method.

If the set of differential equations could be integrated analytically, the expressions for the Y_i's could be substituted into Eq. (4) to give rise to M nonlinear algebraic equations to be solved for the parameters $k_1, k_2, \ldots, k_M$. This would then be a standard problem in regression analysis; however, such solutions are generally not available so that the solution of the differential equations must be obtained numerically. This means that a solution for arbitrary values of the parameters cannot be obtained, but only for specific values. Therefore, a procedure is needed which can generate improved values of the parameters from information obtained for an incorrect set of values.

This can be accomplished as follows. The dependent variables Y_i are approximated by a Taylor series expansion about the incorrect set $k_1^0, k_2^0, \ldots, k_M^0$ with only terms no higher than first order retained [2].

$$Y_i(k_1, k_2, \ldots, k_M, \tau) \approx Y_i(k_1^0, k_2^0, \ldots, k_M^0, \tau) = \sum_{h=1}^{M} (k_h - k_h^0) \frac{\partial Y_i(k_1^0, k_2^0, \ldots, k_M^0, \tau)}{\partial k_h}. \tag{8}$$

Substituting Eq. (8) into Eq. (4) for $Y_i(k_1, k_2, \ldots, k_M, \tau_j)$ gives

$$\frac{\partial S^2}{\partial k_l} \approx -2 \sum_{j=1}^{N_D} \sum_{i=1}^{N} w_{ij} \left[Y_i^D(\tau_j) - Y_i(k_1^0, k_2^0, \ldots, k_M^0, \tau_j) - \sum_{h=1}^{M} (k_h - k_h^0) \frac{\partial Y_i(k_1^0, k_2^0, \ldots, k_M^0, \tau_j)}{\partial k_h} \right] \cdot \frac{\partial Y_i(k_1^0, k_2^0, \ldots, k_M^0, \tau_j)}{\partial k_l} = 0, \qquad l = 1, 2, \ldots, M, \tag{9}$$

where

$$\frac{\partial Y_i(k_1^0, k_2^0,..., k_M^0, \tau_j)}{\partial k_l} = \int_0^{\tau_j} \frac{\partial f_i(Y_1, Y_2,..., Y_N, k_1^0, k_2^0,..., k_M^0, \tau)}{\partial k_l} d\tau.$$

Rewriting Eq. (9) as

$$\beta_l = \sum_{h=1}^{M} \alpha_{hl} \Delta k_h, \qquad l = 1, 2,..., M, \tag{10}$$

where

$$\beta_l = \sum_{j=1}^{N_D} \sum_{i=1}^{N} w_{ij}[Y_i^D(\tau_j) - Y_i(k_1^0, k_2^0,..., k_M^0, \tau_j)]$$

$$\cdot \frac{\partial Y_i(k_1^0, k_2^0,..., k_M^0, \tau_j)}{\partial k_l} = -\frac{1}{2} \frac{\partial S^2}{\partial k_l}\bigg|_{k^0},$$

$$\alpha_{hl} = \sum_{j=1}^{N_D} \sum_{i=1}^{N} w_{ij} \frac{\partial Y_i(k_1^0, k_2^0,..., k_M^0, \tau_j)}{\partial k_h}$$

$$\cdot \frac{\partial Y_i(k_1^0, k_2^0,..., k_M^0, \tau_j)}{\partial k_l},$$

$$\Delta k_h = k_h - k_h^0,$$

one sees that a set of M linear algebraic equations in unknowns $k_1, k_2,..., k_M$ is obtained. Thus, the solution of Eq. (10) allows improved values of the parameters to be determined from calculations based on an incorrect set of values. Because of the first-order expansion assumption, the first correction based on an initial guess of the parameters in general will not be sufficiently accurate that several iterations of this type will be required where the improved set is in turn treated as the incorrect set in order to generate a better set. This method is known as the Newton–Raphson technique.

One disadvantage of this method is that while it converges quite rapidly in the vicinity of the correct set of values, it converges slowly and even diverges otherwise. In order to circumvent this problem, a grid or gradient search method which converges rapidly from afar, but not in the immediate neighborhood of the solution, can be used in conjunction with the Newton–Raphson method. An excellent algorithm which combines the best features of the gradient search and Newton–Raphson methods has been given by Marquardt [13, 2]. It consists of modifying Eq. 10 to read

$$\beta_l = \sum_{h=1}^{M} \alpha'_{hl} \Delta k_h, \qquad l = 1, 2,..., M, \tag{11}$$

where

$$\alpha'_{hl} = \begin{cases} \alpha_{hl}(1 + \lambda), & \text{for } h = l, \\ \alpha_{hl}, & \text{for } h \neq l. \end{cases}$$

For very small values of λ, Eq. (11) is identical with the Newton–Raphson method; on the other hand, for very large values the α_{ll} terms are dominate so that the uncoupled equations

$$\beta_l \approx \lambda \alpha_{ll} \Delta k_l, \qquad l = 1, 2, \ldots, M, \tag{12}$$

are obtained. Thus, as in the gradient search, steps are taken in the direction of the gradient of S^2 with respect to the parameters since β is proportional to $\partial S^2/\partial k$. That convergence is assured by this algorithm is shown as follows. A small initial value of λ such as 0.001 is selected. If the improved set of parameters increases S^2, this means that the Newton–Raphson method diverged because one is far from the correct solution. As a result, λ is increased (for example, by a factor of 10) in order to weight the gradient method more heavily. If S^2 decreases, this means that the solution is being approached, so that λ is decreased (for example, by a factor of 10) in order to weight the Newton–Raphson method more heavily. Thus, in this way, convergence can be assured.

Conversion of a Boundary Value Problem into a Regression Problem

Consider the boundary value problem as defined by the set of ordinary differential equations

$$\frac{dz_i}{d\tau} = f_i(z_1, z_2, \ldots, z_N, \tau), \qquad i = 1, 2, \ldots, N, \tag{13}$$

subject to the boundary conditions that the value of z_i is specified at $\tau = \tau_i$,

$$z_i(\tau_i) = z_{0i}, \qquad i = 1, 2, \ldots, N, \tag{14}$$

where z_i is the ith dependent variable.

This problem can also be formulated as Eq. (13) subject to the unknown initial conditions at $\tau = 0$,

$$z_i(0) = k_i, \qquad i = 1, 2, \ldots, N, \tag{15}$$

where the k_i's must take on values such that the boundary conditions given by Eq. (14) are satisfied. This type of problem has been solved previously by guessing

values for the k_i's and integrating the resulting initial value problem numerically, after which the calculated values of the z_i's at τ_i are compared with the specified ones in order to correct the initial guess [12, 5, 9]. The Newton–Raphson method is used to obtain this correction which necessitates the system to be integrated N times with each k_i slightly perturbed in turn in order to obtain the $\partial y_k/\partial k_i$'s at each τ_i. Such a procedure can be quite time consuming and is needlessly inefficient as will be now shown.

The boundary value problem can be converted to the regression problem, which can be solved very efficiently by means of the change of variables

$$Y_i = (z_i - k_i)/(z_{0i} - k_i) \quad \text{or} \quad z_i = k_i + (z_{0i} - k_i)\, Y_i\,, \qquad i = 1, 2,..., N. \tag{16}$$

If a τ_i is zero, Y_i is not transformed.

This transforms Eqs. (13) and (14) into the system

$$\frac{dY_i}{d\tau} = \frac{1}{(z_{0i} - k_i)} f_i(Y_1\,, Y_2\,,..., Y_N\,, k_1\,, k_2\,,..., k_N\,, \tau), \qquad i = 1, 2,..., N, \tag{17}$$

subject to the initial conditions

$$Y_i(0) = 0, \qquad i = 1, 2,..., N, \tag{18}$$

where the k_i's are to be selected such that

$$Y_i(\tau_i) = 1, \qquad i = 1, 2,..., N. \tag{19}$$

Comparing Eqs. (17)–(19) with Eqs. (1) and (2) for the regression problem, one sees that they are equivalent by interpreting Eq. (19) as specifying the data to be fitted. Since the regression problem calculates $\partial Y_i/\partial Y_k$'s during the numerical integration, it is not necessary to repeat the integration several times with slightly perturbed values of the k_i's. Thus, the boundary value problem can be solved more efficiently than is currently done by converting it to a regression problem.

The authors have applied this calculational technique to the adiabatic fixed-bed reactor problem [9]. Consider that the chemical reaction $2A \rightarrow B$ is occurring in an adiabatic fixed-bed reactor. The differential equations which result from a material and energy balance for steady-state using Lee's [9] nomenclature are

$$\begin{aligned} \frac{1}{M}\frac{d^2x}{dt^2} - \frac{dx}{dt} &= \beta x^2 \exp\left[-\frac{E_1}{T}\right], \\ \frac{1}{M}\frac{d^2T}{dt^2} - \frac{dT}{dt} &= -Q\beta x^2 \exp\left[-\frac{E_1}{T}\right], \end{aligned} \tag{20}$$

subject to the boundary conditions

$$\begin{aligned} x_e &= x(0) - \frac{1}{M}\frac{dx(0)}{dt}, \\ \frac{dx(t_f)}{dt} &= 0, \\ T_e &= T(0) - \frac{1}{M}\frac{dT(0)}{dt}, \\ \frac{dT(t_f)}{dt} &= 0, \end{aligned} \tag{21}$$

where x = concentration, T = temperature, t = dimensionless reactor length variable with respect to diameter of packing particle, t_f = length of reactor, M = Peclet number, E_1 = activation energy group, β = reaction rate group, Q = heat of reaction group, and e = the condition of the reactant before entering the reactor.

Equations (20) and (21) can be put into the form of Eqs. (13) and (14) by means of the transformations

$$\begin{aligned} z_1 &= x - \frac{1}{M}\frac{dx}{dt}, \\ z_2 &= \frac{1}{M}\frac{dx}{dt}, \\ z_3 &= T - \frac{1}{M}\frac{dT}{dt}, \\ z_4 &= \frac{1}{M}\frac{dT}{dt}, \end{aligned} \tag{22}$$

with the result

$$\begin{aligned} \frac{dz_1}{dt} &= -\beta(z_1 + z_2)^2 \exp\left[-\frac{E_1}{z_3 + z_4}\right], \\ \frac{dz_2}{dt} &= \frac{1}{M}\frac{d^2x}{dt} = \frac{dx}{dt} - \frac{dz_1}{dt} = Mz_2 - \frac{dz_1}{dt}, \\ \frac{dz_3}{dt} &= Q\beta(z_1 + z_2)^2 \exp\left[-\frac{E_1}{z_3 + z_4}\right], \\ \frac{dz_4}{dt} &= -\frac{dz_3}{dt} + Mz_4, \end{aligned} \tag{23}$$

subject to

$$\begin{aligned} z_1(0) &= x_e\,, \\ z_3(0) &= T_e\,, \\ z_2(t_f) &= 0, \\ z_4(t_f) &= 0. \end{aligned} \tag{24}$$

It is known that the integration of Eqs. (20) and (21) leads to an inherent instability [3]. This can be explained by noting that if the transformation according to Eq. (16) were carried out, dY_2/dt and dY_4/dt in Eq. (17) would be inversely proportional to $dx(0)/dt$ and $dT(0)/dt$, respectively, which are extremely small. It was found that this problem could be overcome by integrating backward. Thus, let $\tau = t - t_f$ so that Eq. (23) is only changed by the replacement of t by τ and Eq. (24) becomes

$$\begin{aligned} z_1(-t_f) &= x_e\,, \\ z_3(-t_f) &= T_e\,, \\ z_2(0) &= 0, \\ z_4(0) &= 0. \end{aligned} \tag{25}$$

Introducing the values of x and T at the end of the reactor

$$\begin{aligned} z_1(0) &= k_1\,, \\ z_3(0) &= k_2\,, \end{aligned}$$

and making the change of variables according to Eq. (16) transforms the system to

$$\begin{aligned} \frac{dY_1}{d\tau} &= \frac{-\beta}{x_e - k_1}\,[(x_e - k_1)\,Y_1 + k_1 + z_2]^2 \exp\left[\frac{-E_1}{(T_e - k_2)\,Y_3 + k_2 + z_4}\right], \\ \frac{dz_2}{d\tau} &= Mz_2 - (x_e - k_1)\,\frac{dY_1}{d\tau}\,, \\ \frac{dY_3}{d\tau} &= \frac{Q\beta}{T_e - k_2}\,[(x_e - k_1)\,Y_1 + k_1 + z_2]^2 \exp\left[\frac{-E_1}{(T_e - k_2)\,Y_3 + k_2 + z_4}\right], \\ \frac{dz_4}{d\tau} &= Mz_4 - (T_e - k_2)\,\frac{dY_3}{d\tau}\,, \end{aligned} \tag{26}$$

subject to the initial conditions

$$\begin{aligned} Y_1(0) &= 0, \\ z_2(0) &= 0, \\ Y_3(0) &= 0, \\ z_4(0) &= 0, \end{aligned}$$

and the data to be fitted,

$$Y_1(-t_f) = 1,$$
$$Y_3(-t_f) = 1,$$

where

$$Y_1 = (z_1 - k_1)/(x_e - k_1),$$
$$Y_3 = (z_3 - k_2)/(T_e - k_2).$$

It should be noted that the original variables x and T are recovered according to

$$\begin{aligned} x &= z_1 + z_2 = (x_e - k_1)\ Y_1 + k_1 + z_2\,, \\ T &= z_3 + z_4 = (T_e - k_2)\ Y_3 + k_2 + z_4\,. \end{aligned} \tag{27}$$

Also, it is seen from Eq. (26) that since k_1 and k_2 are the concentration and temperature at the end of the reactor, no stability problem is encountered as before because $dY_1/d\tau$ and $dY_4/d\tau$ are inversely proportional to $(x_e - k_1)$ and $(T_e - k_2)$, respectively.

The numerical example given by Lee [9] was worked out by this procedure. No stability or convergence problems were encountered, and the results were in agreement with those of Lee [9], who used a quasilinearization technique; however, as Gidaspow [6] has pointed out, the condition of zero concentration and temperature gradient at the end of the reactor was not satisfied by Lee's solution. The present method forces this condition to be satisfied.

The authors have also applied this technique to the solution of problems in laminar boundary layer theory. By means of similarity transformations, the governing set of partial differential equations can be reduced to a boundary value problem formulation [10, 5]. Infinity is represented by a finite value of the independent variable such that a larger value does not change the solution.

Conversion of an Optimum Policy Problem into a Regression Problem

The object of an optimum policy problem is to maximize or minimize some quantity y which is defined in terms of an integral over the domain $0 \leqslant l \leqslant L$ as

$$y = \int_0^L Y[x_1(l), x_2(l),..., x_N(l), T_1(l), T_2(l),..., T_M(l)]\, dl \tag{28}$$

by appropriately selecting the control variables $T_j(l)$ where the x_i's are variables specified according to the set of differential equations

$$\frac{dx_i}{dl} = f_i(x_1\,, x_2\,,..., x_N\,, T_1\,, T_2\,,..., T_M), \qquad i = 1, 2,..., N, \tag{29}$$

subject to the boundary conditions

$$x_i(0) = x_{0i}, \qquad i = 1, 2,..., p,$$
$$x_i(L) = x_{0i}, \qquad i = p + 1, p + 2,..., N.$$

The set of conditions that the T_j's must satisfy for an optimum are obtained from the Euler–LaGrange equations [14]:

$$\frac{d\lambda_i}{dl} = -\sum_{j=1}^{n} \frac{\partial f_j}{\partial x_i} \lambda_j - \frac{\partial Y}{\partial x_i}, \qquad i = 1, 2,..., N, \tag{30}$$

subject to

$$\lambda_i(L) = 0, \qquad i = 1, 2,..., p,$$
$$\lambda_i(0) = 0, \qquad i = p + 1, p + 2,..., N,$$

and

$$\frac{\partial Y}{\partial T_j} + \sum_{i=1}^{n} \lambda_i \frac{\partial f_i}{\partial T_j} = 0, \qquad j = 1, 2,..., M. \tag{31}$$

Equations (29) and (30) constitute a boundary value problem with Eq. (31) acting as a constraint which gives an algebraic relationship between the T_j's and the x_i's and λ_i's. If each T_j can be solved for explicitly in terms of the x_i's and λ_i's, they can be eliminated in Eqs. (29) and (30) to yield the type of boundary value problem which was treated previously and shown to be reducible to a regression problem. If this is not possible, the situation is somewhat complicated by the fact that at each integration step, the Runge–Kutta method evaluates the right-hand sides of Eqs. (29) and (30) for several increments of the x_i's and λ_i's. In order to compute the corresponding values of the T_j's, the nonlinear set of algebraic equations given by Eq. (31) must be solved. Since this has to be done several times for each step, such a procedure would be quite time consuming. However, this problem can be overcome by expanding Eq. (31) about the values of the T_j's for the previous step and retaining only the linear terms in the T_j's. Such an approximation should be adequate since it is not expected that the T_j's would change greatly over each step. This linear system can then be solved analytically to give explicit expressions of the T_j's.

One finds application of this type of problem to the determination of the optimal temperature distribution in a reaction system [14]. Equation (29) would represent the reaction mechanism while Y in Eq. (28) could be the amount of an intermediate product to be maximized. By including boundary conditions at $l = 0$ and $l = L$, the design of tubular reactors with or without axial dispersion can be considered. The authors have applied this calculational method to the problem of optimal temperature policies for enzymatic reactors subject to catalyst deactivation [7].

Estimation of Uncertainties in the Parameters

It is possible to estimate the uncertainties in the parameters following Bevington [2]. The general comments given in [8] are also of interest. Each parameter k_h can be considered as a function of the data points $Y_i{}^D(\tau_j) = Y_{ij}^D$ which are random variables so that the k_h's are also random variables.

$$k_h = F_h(Y_{11}^D, \ldots, Y_{NN_D}^D). \tag{32}$$

Expanding Eq. (32) about the expected values of the Y_{ij}^D's and retaining terms no higher than first order gives

$$k_h \approx F_h|_{\bar{Y}_{ij}} + \sum_{j=1}^{N_D} \sum_{i=1}^{N} \left(\frac{\partial F_h}{\partial Y_{ij}^D}\right)_{\bar{Y}_{ij}} (Y_{ij}^D - \bar{Y}_{ij}^D), \tag{33}$$

where the bar denotes and expected value. It is assumed that the expected value of k_h is approximately equal to $F_h|_{\bar{Y}_{ij}}$. Then

$$(k_h - \bar{k}_h)(k_r - \bar{k}_r) = \sum_{i,j}^{N,N_D} \sum_{m,n}^{N,N_D} \left(\frac{\partial F_h}{\partial Y_{ij}^D}\right)_{\bar{Y}_{ij}} \left(\frac{\partial F_r}{\partial Y_{mn}^D}\right)_{\bar{Y}_{mn}} (Y_{ij}^D - \bar{Y}_{ij}^D)(Y_{mn}^D - \bar{Y}_{mn}^D). \tag{34}$$

Operating on Eq. (34) with the expectation

$$E\{(k_h - \bar{k}_h)(k_r - \bar{k}_r)\} = \operatorname{cov}(k_h, k_r) = \sum_{i,j}^{N,N_D} \sum_{m,n}^{N,N_D} \left(\frac{\partial F_h}{\partial Y_{ij}^D}\right)_{\bar{Y}_{ij}} \left(\frac{\partial F_r}{\partial Y_{mn}^D}\right)_{\bar{Y}_{mn}} \operatorname{cov}(Y_{ij}^D, Y_{mn}^D). \tag{35}$$

Since it is not expected that there is a correlation between the measurement errors of any two different data points, the covariance of any two different points is zero while that for the same point is the variance $\operatorname{Var}(Y_{ij}^D)$. Thus, in Eq. (35) only the terms in which i, j equals m, n are nonzero, and Eq. 35 becomes

$$\operatorname{cov}(k_h, k_r) = \sum_{j=1}^{N_D} \sum_{i=1}^{N} \left(\frac{\partial F_h}{\partial Y_{ij}^D}\right)\left(\frac{\partial F_r}{\partial Y_{ij}^D}\right) \operatorname{var}(Y_{ij}^D). \tag{36}$$

For the case where $h = r$, one obtains

$$\operatorname{var}(k_h) = \sum_{j=1}^{N_D} \sum_{i=1}^{N} \left(\frac{\partial F_h}{\partial Y_{ij}^D}\right)^2 \operatorname{var}(Y_{ij}^D). \tag{37}$$

In order to evaluate F_h, Eq. (11) is used in matrix form

$$\beta = \Delta k\, \alpha'$$

and is solved for Δk:

$$\Delta k = \beta(\alpha')^{-1} = \beta\epsilon',$$

where ϵ' is the inverse of α' or

$$k_h = k_h^0 + \sum_{l=1}^{M} \beta_l \epsilon'_{hl}, \tag{38}$$

where ϵ' is inverse matrix of α'.

For convenience, β_l and α_{hl} are written in the following form, with $w_{ij}=1/\mathrm{var}(Y_{ij}^D)$.

$$\beta_l = \sum_{j=1}^{N_D} \sum_{i=1}^{N} \frac{1}{\mathrm{var}(Y_{ij}^D)} (Y_{ij}^D - Y_{ij}\,|_{k^0})\, X_{il}(\tau_j),$$

$$\alpha_{hl} = \sum_{j=1}^{N_D} \sum_{i=1}^{N} \frac{1}{\mathrm{var}(Y_{ij}^D)} X_{ih}(\tau_j)\, X_{il}(\tau_j),$$

where

$$X_{il} = \frac{\partial Y_i}{\partial k_l}.$$

Since

$$F_h = k_h^0 + \sum_{l=1}^{M} \beta_l \epsilon'_{hl},$$

then

$$\frac{\partial F_h}{\partial Y_{ij}^D} = \sum_{l=1}^{M} \frac{\partial \beta_l}{\partial Y_{ij}^D} \epsilon'_{hl} = \sum_{l=1}^{M} \frac{1}{\mathrm{var}(Y_{ij}^D)} X_{il}(\tau_j)\, \epsilon'_{hl}. \tag{39}$$

Substituting Eq. (39) into (36) gives

$$\begin{aligned}
\mathrm{cov}(k_h, k_r) &= \sum_{j=1}^{N_D} \sum_{i=1}^{N} \frac{1}{\mathrm{var}(Y_{ij}^D)} \sum_{l,m=1}^{M} \epsilon_{hl}\epsilon_{rm} X_{il}(\tau_j)\, X_{im}(\tau_j) \\
&= \sum_{l,m=1}^{M} \epsilon_{hl}\epsilon_{rm} \sum_{j=1}^{N_D} \sum_{i=1}^{N} \frac{1}{\mathrm{var}(Y_{ij}^D)} X_{il}(\tau_j)\, X_{im}(\tau_j) \\
&= \sum_{l,m=1}^{M} \epsilon_{hl}\epsilon_{rm}\alpha_{lm} = \sum_{m=1}^{M} \epsilon_{rm} \sum_{l=1}^{M} \epsilon_{hl}\alpha_{lm} \\
&= \sum_{m=1}^{M} \epsilon_{rm}\delta_{hm} = \epsilon_{rh} = \epsilon_{hr},
\end{aligned} \tag{40}$$

where δ_{hm} is the Kronecker delta. Also

$$\text{var}(k_h) = \epsilon_{hh} . \tag{41}$$

Thus, the variance of the parameters is given by the diagonal terms of the inverse matrix of α while the other terms give the covariance of any two parameters. For this reason, the system of simultaneous linear equations as given by Eq. (11) is solved by matrix inversion.

The significance of the variance is well known in that it gives a measure of the possible error in the parameters; however, that of the covariance requires some explanation. By using the correlation coefficient defined as

$$r_{ij} = [\text{cov}(k_i , k_j)]/[\text{var}(k_i)\ \text{var}(k_j)]^{1/2}, \tag{42}$$

where $r_{ii} = 1$, information can be obtained regarding whether a given parameter belongs in the model. The value of r_{ij} ranges from zero when there is no correlation to ± 1 when there is complete linear correlation. Since according to Eq. (11), each iteration attempts to fit the data to a linear function of the parameters, each correlation coefficient should ideally be unity if all the parameters belong in the model; however, in practice, values less than unity are obtained so that the correlation coefficient cannot be used directly.

Nevertheless, an indication of the degree of correlation between any two parameters can be obtained by comparing the probability distribution of k_i and k_j with one in which k_i and k_j are uncorrelated or $r_{ij} = 0$. Such a distribution is a two-dimensional normal distribution in k_i and k_j . Tables are available which give the probability that a random sample with ν degrees of freedom obtained from an uncorrelated population yields a value of $| r_{ij} |$ as large or larger than some given value of r_{ij} [2]. Using the calculated values of r_{ij} and $\nu = NN_D - M$, one obtains the probability that the observed value of r_{ij} could have been obtained from a population in which k_i and k_j are uncorrelated. A small value of this probability implies that k_i and k_j are correlated and belong in the model.

In order to characterize the overall goodness of fit of the model to the data, the chi-square test may be used [2]. This is justified if the covariance of any two data points is zero, which as indicated before is a reasonable assumption. The reduced chi-square statistic χ_ν^2 is defined as

$$\chi_\nu^2 = \frac{1}{\nu} \sum_{j=1}^{N_D} \sum_{i=1}^{N} \frac{1}{\text{var}(Y_{ij}^D)} [Y_{ij}^D - Y_{ij}]^2, \tag{43}$$

where $\nu = NN_D - M$ and Y_{ij} is the value as predicted from the model. The variance of the data σ^2 is

$$\sigma^2 = \left[\frac{1}{NN_D} \sum_{j=1}^{N_D} \sum_{i=1}^{N} \frac{1}{\text{var}(Y_{ij}^D)} \right]^{-1} \tag{44}$$

and an estimate of the variance of the fit S^2 is

$$S^2 = \left\{\frac{1}{\nu}\sum_{j=1}^{N_D}\sum_{i=1}^{N}\frac{1}{\text{var}(Y_{ij}^D)}[Y_{ij}^D - Y_{ij}]^2\right\}\Big/\left[\frac{1}{NN_D}\sum_{j=1}^{N_D}\sum_{i=1}^{N}\frac{1}{\text{var}(Y_{ij}^D)}\right]. \quad (45)$$

Thus,

$$\chi_\nu^2 = S^2/\sigma^2. \quad (46)$$

Since S^2 depends on the accuracy of both the data and the fit while σ^2 depends only on that of the data, χ_ν^2 should be near unity if the model is appropriate. The probability distribution of χ_ν^2 is tabulated and gives the probability that the correct model would yield a value of χ_ν^2 as large or larger than the calculated value. Since $\chi_\infty^2 = 1$ gives a probability of 0.5, the model is appropriate if this probability is near 0.5.

In cases where the uncertainties in the data are not known or cannot be estimated, they can be approximated from

$$\sigma^2 \approx \frac{1}{NN_D - M}\sum_{j=1}^{N_D}\sum_{i=1}^{N}(Y_{ij}^D - Y_{ij})^2, \quad (47)$$

which is S^2 or χ_ν^2 with all the $\text{var}(Y_{ij}^D)$'s set equal. From Eqs. (40) and (41) it is seen that

$$\text{cov}(k_h, k_r) \cong S^2\epsilon_{hr}[\text{var}(Y_{ij}^D) = 1], \quad (48)$$

$$\text{var}(k_h) \cong S^2\epsilon_{hh}[\text{var}(Y_{ij}^D) = 1], \quad (49)$$

where $\epsilon_{hr}[\text{var}(Y_{ij}^D) = 1]$ is the inverse matrix of α evaluated with $\text{var}(Y_{ij}^D) = 1$. It should be noted that χ_ν^2 has no meaning in this case since it is identically unity.

The authors have utilized these statistical calculations as an aid in the determination of chemical reaction mechanisms from experimental concentration data. The correct mechanism is characterized by (1) a value of χ_ν^2 near unity, (2) standard deviations of the regressed rate constants smaller than the constants themselves, and (3) correlation coefficients which give probabilities that all the rate constants belong in the mechanism. For an incorrect mechanism, one can deduce from the statistical analysis what portions of the mechanism are in error and need improvement. Cases have also been solved where data are not available for every species in the mechanism.

Computer Program

A computer program whose flow chart is shown in Fig. 1 was prepared in FORTRAN to carry out the operations previously described. It can be constructed easily by combining available library subroutines. The differential equations are integrated by means of an IBM subroutine which uses Gill's modification of the

Runge–Kutta method. A convenient feature of this routine is that it changes the step size as it proceeds so that neither too large nor too small a step size is used in order to assure an accurate and as rapid a solution as possible. This subroutine was also modified to perform simultaneously the integration required to evaluate the $\partial Y_i/\partial k$'s. The iterative process of improving the initial guess of the parameters according to Marquardt's algorithm was carried out by modifying a set of subroutines given by Bevington [2] which makes a least-squares fit to a nonlinear function and performs a statistical analysis of the regreasion. The total program is implemented by a driver routine which reads in the initial conditions, the initial guess of the parameters, and the data to be fitted as well as produces the resulting output in both tabular and graphical form and by a subroutine which lists the set of differential equations.

The program is quite efficient in that it has never required a computation time greater than one minute on a UNIVAC 1108. It has run for cases where the initial guesses were off by as much as a factor of 100 and up to eight parameters and fifteen equations.

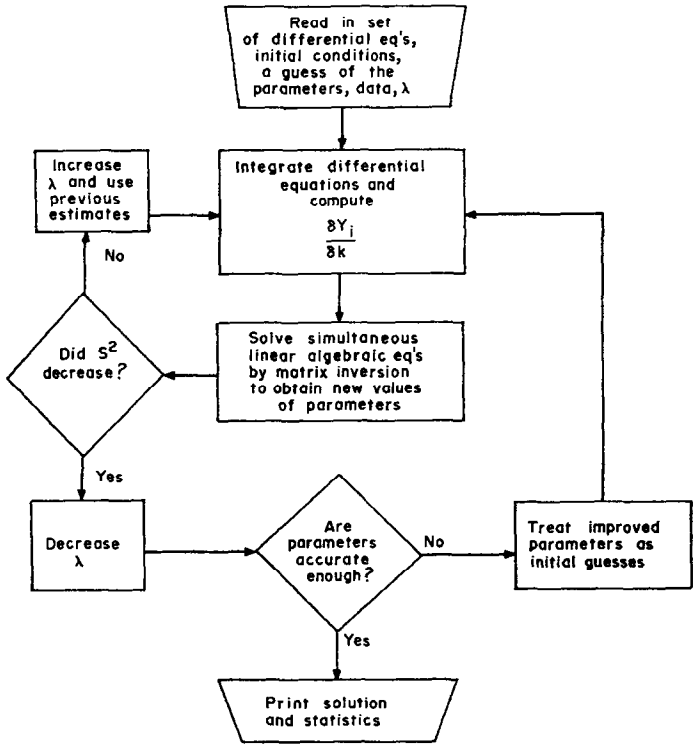

FIG. 1. Flow chart for computer program.

REFERENCES

1. W. E. BALL AND L. C. D. GROENWEGHE, *Ind. Eng. Chem. Fundamentals* **5** (1966), 181.
2. P. R. BEVINGTON, "Data Reduction and Error Analysis for the Physical Sciences," McGraw–Hill, New York, 1969.
3. J. COSTE, D. RUDD, AND N. R. AMUNDSON, *Can. J. Chem. Eng.* **39** (1961), 149.
4. J. M. EAKMAN, *Ind. Eng. Chem. Fundamentals* **8** (1969), 53.
5. D. GIDASPOW, W. TOCZYCK, AND C. W. SOLBRIG, Adiabatic Humidification in Boundary Layer Flow with a General Computer Program for Nonlinear Boundary Value Problems, in "Transport" (R. Gordon, Ed.), pp. 147–176, Atzec, Austin, Tex., 1973.
6. D. GIDASPOW, private communication, 1972.
7. W. R. HAAS, W. J. WNEK, AND L. L. TAVLARIDES, Optimal Temperature Policy for (First-Order) Reversible Reactions with Deactivations: Applied to Enzyme Reactors, 66th Annual Meeting A.I.Ch.E., Philadelphia, Pa., 1973; *A.I.Ch.E. J.* **20** (1974), 707–12.
8. D. M. HIMMELBLAU, C. R. JONES, AND K. B. BISCHOFF, *Ind. Eng. Chem. Fundamentals* **6** (1967), 539.
9. E. S. LEE, "Quasilinearization and Invariant Imbedding," Academic Press, New York, 1968.
10. V. G. LEVICH, "Physicochemical Hydrodynamics," Prentice–Hall, Englewood Cliffs, N.J., 1962.
11. A. L. LOEB, J. T. G. OVERBEEK, AND P. H. WIERSEMA, "The Electric Double Layer around a Spherical Colloid Particle," M.I.T. Press, Cambridge, Mass., 1960.
12. M. MAREK AND V. HLAVACEK, *Chem. Eng. Sci.* **21** (1966), 493 (German).
13. D. W. MARQUARDT, *J. Soc. Ind. Appl. Math.* **11** (1963), 431–441.
14. R. S. SCHECHTER, "The Variational Method in Engineering," McGraw–Hill, New York, 1967.
15. R. H. SNOW, *J. Phys. Chem.* **70** (1966), 2780.