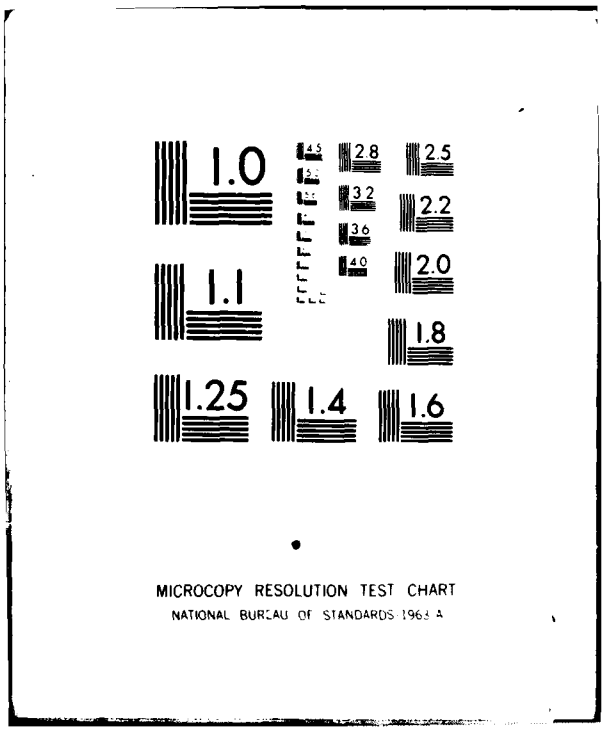
AD-A083 819 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
NEW ALGORITHMS FOR NONLINEAR LEAST SQUARES AND BAYESIAN PARAMET--ETC(U)
FEB 80 W E STEWART, J P SORENSEN DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-2037 NL

END
DATE FILMED
6-80
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

LEVEL IV (2)

AD A083819

MRC Technical Summary Report #2037

NEW ALGORITHMS FOR NONLINEAR LEAST SQUARES AND BAYESIAN PARAMETER ESTIMATION

Warren E. Stewart and Jan P. Sørensen

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

February 1980

Received January 4, 1980

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, DC 20550

Wisconsin Alumni
Research Foundation
University of Wisconsin
Graduate School
Madison, WI 53706

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

NEW ALGORITHMS FOR NONLINEAR LEAST SQUARES AND BAYESIAN PARAMETER ESTIMATION

Warren E. Stewart and Jan P. Sørensen

Technical Summary Report #2037
February 1980

ABSTRACT

New algorithms are described for Bayesian estimation of parameters in nonlinear models of multiple-response systems. Modal and interval estimates are provided for the parameter vector $\underset{\sim}{\theta}$ of the predictor model, and for the variance-covariance matrix $\underset{\sim}{C}$ of a Normal error distribution. Allowance is made for gaps (missing values of responses), such as commonly occur in practice. Two chemical examples are analyzed.

AMS(MOS) Subject Classification - 62F10, 62F15, 62F25

Key Words: Parameter estimation, Bayesian estimation, Least squares, Nonlinear estimation, Covariance estimation, Multiple-response systems, Missing values, Statistics.

Work Unit No. 2 - Physical Mathematics

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024; the National Science Foundation Grant ENG76-24368; the Wisconsin Alumni Research Foundation through funds from the Graduate School, University of Wisconsin.

SIGNIFICANCE AND EXPLANATION

Some new algorithms are presented for fitting mathematical models to multiple-response experiments. These algorithms give estimates of the parameters in a user-defined predictor model, and also estimate the parameters of a Gaussian model of the observational error distribution. The development is based on Bayes' theorem, and provides a natural extension of known least-squares estimation methods. Allowance is made for missing values of responses, which occur frequently in practical work.

NEW ALGORITHMS FOR NONLINEAR LEAST SQUARES AND BAYESIAN PARAMETER ESTIMATION

Warren E. Stewart and Jan P. Sørensen

New algorithms are described for Bayesian estimation of parameters in nonlinear models of multiple-response systems. Modal and interval estimates are provided for the parameter vector $\underset{\sim}{\theta}$ of the predictor model, and for the variance-covariance matrix $\underset{\sim}{\sigma}$ of a Normal error distribution. Allowance is made for gaps (missing values of responses), such as commonly occur in practice. Two chemical examples are analyzed.

INTRODUCTION

Realistic models of multivariate phenomena often relate several predicted responses to a common set of parameters. Multiresponse experiments are required to establish such models, but frequently yield irregular data which are difficult to analyze by classical methods.

Bayes' theorem is a good starting point for parameter estimation in these situations. The multivariate error distribution can be estimated concurrently, whereas it has to be prescribed when least-squares methods are used. Thus, the Bayesian approach allows more objective parameter estimates, if sufficient data are provided. An excellent general account of this approach is given by Box and Tiao (1973).

Bayesian inference deals with a data array $\{y_{ui}\} \equiv \underset{\sim}{y}$, a model for $E(\underset{\sim}{y})$ with parameter vector $\underset{\sim}{\theta}$, and an error distribution model. If a Normal error model is used, with variance-covariance matrix $\underset{\sim}{\sigma}$, the unknown elements of $\underset{\sim}{\sigma}$ will appear as additional parameters. The full set of parameters can be estimated optimally by maximizing the posterior density $p(\underset{\sim}{\theta},\underset{\sim}{\sigma}|\underset{\sim}{y})$; confidence regions can also be calculated from this function.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024; the National Science Foundation Grant ENG76-24368; the Wisconsin Alumni Research Foundation through funds from the Graduate School, University of Wisconsin.

In certain cases, the posterior density can be integrated analytically to obtain the marginal density $p(\underset{\sim}{\theta}|\underset{\sim}{y})$. Box and Draper (1965) accomplished this for multivariate Normal error distributions and rectangular data structures (Table 1a). For block-rectangular structures (Table 1b), $p(\underset{\sim}{\theta}|\underset{\sim}{y})$ is the product of the Box-Draper densities for the individual rectangles. More complicated data structures often occur, however, such as that in Table 1c, for which $p(\underset{\sim}{\theta}|\underset{\sim}{y})$ cannot be expressed in closed form. Therefore, in this paper we use the full posterior density $p(\underset{\sim}{\theta},\underset{\sim}{\sigma}|\underset{\sim}{y})$, which has a closed form for any finite data structure.

Inspection of the parameter estimates and residuals often suggests alternatives to the postulated model. Therefore, parameter estimation should not be viewed as an end in itself, but should be followed by critical examination of the model and investigation of any promising alternatives. Interesting predictions or unresolved differences between models will naturally lead to further experiments.

Table 1. Examples of Data Structures with $m = 4$ and $n = 8$

	1a. Rectangular				1b. Block-rectangular				1c. Irregular			
u	y_{u1}	y_{u2}	y_{u3}	y_{u4}	y_{u1}	y_{u2}	y_{u3}	y_{u4}	y_{u1}	y_{u2}	y_{u3}	y_{u4}
1	+	+	+	+	+	+			+		+	+
2	+	+	+	+	+	+			+	+	+	+
3	+	+	+	+	+	+				+	+	
4	+	+	+	+			+	+	+			
5	+	+	+	+			+	+		+		+
6	+	+	+	+			+	+				+
7	+	+	+	+			+	+	+	+		+
8	+	+	+	+			+	+			+	+

PROBLEM FORMULATION

Consider a set of independent experiments, $u = 1,\ldots,n$, in which a table $\{y_{ui}\}$ of observed responses have been obtained at known settings $\{\underset{\sim}{x}_u\}$ of the independent variables. There are m linearly independent kinds of observations; thus the index i ranges from 1 to m, but in each experiment some values may be missing as in Tables 1b and 1c.

The observations in the uth experiment are regarded as a sample from a population of the form

$$y_{ui} = f_i(\underset{\sim}{x}_u, \underset{\sim}{\theta}) + \varepsilon_{ui} \ . \tag{1}$$

The functions $f_i(\underset{\sim}{x}_u, \underset{\sim}{\theta})$ are models for the expected responses $E(y_{ui}|\underset{\sim}{\theta})$. The residuals ε_{ui} in the uth experiment are treated as a random sample from an m-variate Normal distribution; this gives the probability density (Wilks, 1962)

$$p(\underset{\sim}{\varepsilon}_u|\underset{\sim}{\sigma}) = (2\pi)^{-m_u/2} |\underset{\sim}{\sigma}_u|^{-1/2} \exp(-\tfrac{1}{2}\, \underset{\sim}{\varepsilon}_u^T\, \underset{\sim}{\sigma}_u^{-1}\, \underset{\sim}{\varepsilon}_u). \tag{2}$$

Here $\underset{\sim}{\varepsilon}_u$ is the column vector of error variables $\varepsilon_{u1},\ldots,\varepsilon_{um}$ with dummy zeroes inserted where observations are missing. Correspondingly, $\underset{\sim}{\sigma}_u$ is obtained from the full variance-covariance matrix, $\underset{\sim}{\sigma} = \{\sigma_{ij}\}$, by substituting dummy elements δ_{ij} whenever observation y_{ui} or y_{uj} is missing. Here δ_{ij} is unity when $i=j$, and zero otherwise.

The joint error density model for the set of n experiments follows directly from Equation (2):

$$p(\underset{\sim}{\varepsilon}|\underset{\sim}{\theta},\underset{\sim}{\sigma}) = \prod_{u=1}^{n} (2\pi)^{-m_u/2} |\underset{\sim}{\sigma}_u|^{-1/2} \exp(-\tfrac{1}{2}\, \underset{\sim}{\varepsilon}_u^T\, \underset{\sim}{\sigma}_u^{-1}\, \underset{\sim}{\varepsilon}_u). \tag{3}$$

Insertion of Equation (1) gives the corresponding density in observation space:

$$p(\underset{\sim}{y}|\underset{\sim}{\theta},\underset{\sim}{\sigma}) = \Big[\prod_{u=1}^{n} (2\pi)^{-m_u/2} |\underset{\sim}{\sigma}_u|^{-1/2}\Big] \cdot \exp\Big\{-\tfrac{1}{2} \sum_{u=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{m} \sigma_u^{ij} [y_{ui} - f_{ui}(\underset{\sim}{\theta})][y_{uj} - f_{uj}(\underset{\sim}{\theta})]\Big\}. \tag{4}$$

Here the functions $f_{ui}(\underset{\sim}{\theta})$ stand for $f_i(\underset{\sim}{x}_u, \underset{\sim}{\theta})$ evaluated at the known settings $\underset{\sim}{x}_u$ of the independent variables. The σ_u^{ij} are the elements of the precision matrices $\underset{\sim}{\sigma}_u^{-1}$. The right-hand term may also be regarded, by Bayes' theorem, as the likelihood function for $\underset{\sim}{\theta}$ and $\underset{\sim}{\sigma}$ when evaluated with given observations $\underset{\sim}{y}$.

The usual factorization of the prior density $p(\underset{\sim}{\theta}, \underset{\sim}{\sigma})$ is assumed,

$$p(\underset{\sim}{\theta}, \underset{\sim}{\sigma}) = p(\underset{\sim}{\theta})\, p(\underset{\sim}{\sigma}) \tag{5}$$

and a locally uniform density $p(\underset{\sim}{\theta})$ is assumed in the region of appreciable likelihood. The latter assumption requires some care in the parametrization of the model. The prior density of $\underset{\sim}{\sigma}$ is taken from Box and Draper (1965):

$$p(\underset{\sim}{\sigma}) \propto |\underset{\sim}{\sigma}|^{-(m+1)/2} . \tag{6}$$

Bayes' theorem then gives the posterior density

$$\begin{aligned} p(\underset{\sim}{\theta}, \underset{\sim}{\sigma} | \underset{\sim}{y}) &= p(\underset{\sim}{\theta}, \underset{\sim}{\sigma})\, p(\underset{\sim}{y} | \underset{\sim}{\theta}, \underset{\sim}{\sigma}) \\ &= c|\underset{\sim}{\sigma}|^{-(m+1)/2} \left[\prod_{u=1}^{n} |\underset{\sim}{\sigma}_u|^{-1/2} \right] \\ &\cdot \exp\left\{ -\tfrac{1}{2} \sum_{u=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{m} \sigma_u^{ij} [y_{ui} - f_{ui}(\underset{\sim}{\theta})][y_{uj} - f_{uj}(\underset{\sim}{\theta})] \right\} \end{aligned} \tag{7}$$

in which c is a proportionality constant. All that the data reveal about the parameters $\underset{\sim}{\theta}$ and $\underset{\sim}{\sigma}$ is contained in this density function.

Point estimates of $\underset{\sim}{\theta}$ and $\underset{\sim}{\sigma}$ are obtainable by maximizing the posterior density just described, or by minimizing the function

$$\begin{aligned} S(\underset{\sim}{\psi}) \equiv S(\underset{\sim}{\theta}, \underset{\sim}{\sigma}) &= -2 \ln p(\underset{\sim}{\theta}, \underset{\sim}{\sigma} | \underset{\sim}{y}) + 2 \ln c \\ &= (m+1) \ln |\underset{\sim}{\sigma}| + \sum_{u=1}^{n} \ln |\underset{\sim}{\sigma}_u| \\ &+ \sum_{u=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{m} \sigma_u^{ij} [y_{ui} - f_{ui}(\underset{\sim}{\theta})][y_{uj} - f_{uj}(\underset{\sim}{\theta})] \end{aligned} \tag{8}$$

over the permitted region of $\underset{\sim}{\theta}$ and $\underset{\sim}{\sigma}$. Here $\underset{\sim}{\psi}$ is a column array of the model parameters $\theta_1, \ldots, \theta_p$ and the independent elements of $\underset{\sim}{\sigma}$.

The latter are taken from the lower triangle of $\underset{\sim}{\sigma}$ in row order, i.e. $\psi_{p+k} = \sigma_{ij}$ with $i \geq j$ and $k = j + i(i-1)/2$. Thus, the total number of parameters is $q = p + m(m+1)/2$.

If the matrix $\underset{\sim}{\sigma}$ were believed to be known, i.e., if a sharply focussed prior density $p(\underset{\sim}{\sigma})$ were assumed, then $S(\underset{\sim}{\psi})$ would reduce to $S(\underset{\sim}{\theta})$ and we would have a least-squares estimation problem with just p parameters. In practice, one seldom knows $\underset{\sim}{\sigma}$ accurately; hence, the full Bayesian solution is recommended.

PARAMETER ESTIMATION ALGORITHMS

Several algorithms are described here for obtaining summary information from Equation (8). These algorithms are part of a Fortran IV package available from the authors.

1. Counting Algorithm

Before analyzing S we count Equations (1) to see which parameters can plausibly be estimated from the data. We first try to match each parameter σ_{kj} in $\underset{\sim}{\psi}$ with an observation pair (y_{uk}, y_{uj}) of a replicate experiment (i.e., an experiment which has the same expected response values as a prior experiment in the data set). If this process cannot be completed for a given k, we then try to match each remaining error parameter σ_{kj}, and each model parameter θ_r in the function pairs $[f_{uk}(\underset{\sim}{\theta}), f_{uj}(\underset{\sim}{\theta})]$, with a non-replicate observation pair (y_{uk}, y_{uj}). Finally, any remaining model parameters θ_r are matched with remaining non-replicate observations. If the matching can be completed for all elements of $\underset{\sim}{\psi}$, we proceed with the estimation. Otherwise, the full set of parameters cannot be estimated from the data.

The counting algorithm is a logical Gaussian elimination. This test is a useful diagnostic, but is not infallible, since the actual rank of the estimation equations depends on the numerical values of $\underset{\sim}{x}$, $\underset{\sim}{y}$, and $\underset{\sim}{\psi}$.

2. Minimization Algorithm

A modified Newton method is used to find a minimum of $S(\underset{\sim}{\psi})$. Let $\underset{\sim}{\psi}_0$ be the value of $\underset{\sim}{\psi}$ at the start of an iteration. A correction vector $(\underset{\sim}{\psi}_1 - \underset{\sim}{\psi}_0)$ is computed by minimizing the local quadratic expansion (see Appendix A for derivative expressions)

$$\tilde{S}(\underset{\sim}{\psi}) = S(\underset{\sim}{\psi}_0) + \left.\frac{\partial S}{\partial \underset{\sim}{\psi}}\right|_0 (\underset{\sim}{\psi} - \underset{\sim}{\psi}_0) + \tfrac{1}{2} (\underset{\sim}{\psi} - \underset{\sim}{\psi}_0)^T \left.\frac{\partial^2 S}{\partial \underset{\sim}{\psi} \partial \underset{\sim}{\psi}}\right|_0 (\underset{\sim}{\psi} - \underset{\sim}{\psi}_0) \qquad (9)$$

over a user-specified rectangular region around $\underset{\sim}{\psi}_0$. The region is chosen small enough to ensure that $\tilde{S}(\underset{\sim}{\psi})$ is a good approximation to the function $S(\underset{\sim}{\psi})$ of Equation (9). A search is then made for a minimum of S in the interval of positive definite σ on the line from $\underset{\sim}{\psi}_0$ through $\underset{\sim}{\psi}_1$; this gives the starting point for the next iteration. The calculation continues until two successive line-minima agree within confidence intervals calculated from Equation (14) for each parameter.

3. Response-Independence Test

Box and co-workers (1973) have pointed out the need to test the responses for linear independence. Preferably, one should perform this test on the residuals $[y_{ui} - f_{ui}(\theta)]$, which might become linearly dependent in certain regions of $\underset{\sim}{\theta}$. In the present procedure, such linear dependence is readily detected during the inversion of σ at the start of each iteration. The calculation can continue if all pivot elements (Stewart, 1973) found in this inversion are greater than a specified fraction, say 0.1, of the corresponding elements σ_{ii}.

4. Confidence Regions

Equation (8) gives the simple form

$$p(\underset{\sim}{\psi}|\underset{\sim}{y}) \propto \exp[-\tfrac{1}{2} S(\underset{\sim}{\psi})] \tag{10}$$

for the posterior density function, or "confidence density". Use of Equation (9) gives the approximation

$$p(\underset{\sim}{\psi}|\underset{\sim}{y}) \propto \exp[-\tfrac{1}{2}(\underset{\sim}{\psi} - \hat{\underset{\sim}{\psi}})^T \hat{\underset{\sim}{A}}(\underset{\sim}{\psi} - \hat{\underset{\sim}{\psi}})], \tag{11}$$

valid in the neighborhood of the minimum point $\hat{\underset{\sim}{\psi}}$. Here $\hat{\underset{\sim}{A}}$ is the $q \times q$ matrix (positive definite since S is at a minimum) with elements

$$\hat{A}_{km} = \tfrac{1}{2} \frac{\partial^2 S}{\partial \psi_k \partial \psi_m}\bigg|_{\hat{\underset{\sim}{\psi}}} \tag{12}$$

computed as described in the Appendix. Thus, near the optimum, the parameters are Normally distributed with variance-covariance matrix $\hat{\underset{\sim}{A}}^{-1}$. If Equation (11) is used as an approximation for all values of $\underset{\sim}{\psi}$, then the confidence intervals for Normal distributions can be applied. For example, the ellipsoidal region

$$(\underset{\sim}{\psi} - \hat{\underset{\sim}{\psi}})^T \hat{\underset{\sim}{A}}(\underset{\sim}{\psi} - \hat{\underset{\sim}{\psi}}) \leq \chi^2(q,\alpha) \tag{13}$$

roughly approximates the $100(1 - \alpha)$ percent highest-posterior-density region, or joint confidence region, for $\underset{\sim}{\psi}$ based on the given data. The intervals

$$(|\psi_k - \hat{\psi}_k| / \sqrt{2\, A^{kk}}) \leq \text{erfc}^{-1}(\alpha) \tag{14}$$

roughly approximate the $100(1 - \alpha)$ percent confidence intervals for the individual parameters. For symmetric 95 percent confidence intervals ($\alpha = 0.05$), $\text{erfc}^{-1}(\alpha)$ has the value 1.96.

Equation (14) is more reliable than (13), since the integration used to obtain it is less affected by the tails of the posterior density function. More accurate intervals can be obtained, but with greater effort, by numerical integration of Equation (7) or (10).

RESULTS FOR RECTANGULAR DATA STRUCTURES

If every experiment gives a full set of observations $y_{u1},\ldots,y_{um}$, then Equation (7) takes the form

$$p(\underset{\sim}{\theta},\underset{\sim}{\sigma}|\underset{\sim}{y}) \propto |\underset{\sim}{\sigma}|^{-(m+n+1)/2} \exp[-\tfrac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \sigma^{ij} v_{ij}(\underset{\sim}{\theta})] \tag{15}$$

in which

$$\underset{\sim}{v}_{ij}(\underset{\sim}{\theta}) = \sum_{u=1}^{n} [y_{ui} - f_{ui}(\underset{\sim}{\theta})][y_{uj} - f_{uj}(\underset{\sim}{\theta})] \; . \tag{16}$$

Integration of Equation (15) over the region of positive definite $\underset{\sim}{\sigma}$ gives the marginal density function

$$p(\underset{\sim}{\theta}|\underset{\sim}{y}) \propto |\underset{\sim}{v}(\underset{\sim}{\theta})|^{-n/2} \tag{17}$$

as shown by Box and Draper (1965). We wish to compare the estimates based on this function with those obtained from the full density function of Equation (15).

Setting $p(\underset{\sim}{\theta}|\underset{\sim}{y})$ stationary with respect to its parameters gives

$$\frac{\partial \ln |\underset{\sim}{v}(\underset{\sim}{\theta})|}{\partial \theta_k} \equiv \sum_i \sum_j v^{ij} \frac{\partial v_{ij}}{\partial \theta_k} = 0 \qquad k = 1,\ldots,p \tag{18}$$

when use is made of the Laplace expansion of $|\underset{\sim}{v} + \underset{\sim}{dv}|$. Here the v^{ij} are the elements of the matrix $\underset{\sim}{v}^{-1}$.

Setting $p(\underset{\sim}{\theta},\underset{\sim}{\sigma}|\underset{\sim}{y})$ stationary with respect to its parameters gives, after use of Equation (15),

$$-2 \frac{\partial \ln p(\underset{\sim}{\theta},\underset{\sim}{\sigma}|\underset{\sim}{y})}{\partial \theta_k} \equiv \sum_i \sum_j \sigma^{ij} \frac{\partial v_{ij}(\underset{\sim}{\theta})}{\partial \theta_k} = 0 \qquad k = 1,\ldots,p \tag{19}$$

$$-2 \frac{\partial \ln p(\underset{\sim}{\theta},\underset{\sim}{\sigma}|\underset{\sim}{y})}{\partial \sigma^{rs}} \equiv (m+n+1) \frac{\partial}{\partial \sigma^{rs}} (\ln |\underset{\sim}{\sigma}|) + \frac{\partial}{\partial \sigma^{rs}} \sum_i \sum_j \sigma^{ij} v_{ij}(\underset{\sim}{\theta})$$

$$\equiv (2 - \delta_{rs})[-(m+n+1)\, \sigma_{rs} + v_{rs}(\underset{\sim}{\theta})] = 0 \tag{20}$$

$$r = 1,\ldots,m \qquad s = 1,\ldots,r \; .$$

Equation (20) gives, at the stationary point,

$$\hat{\sigma}_{rs} = \frac{v_{rs}(\hat{\theta})}{m+n+1} \quad . \tag{21}$$

Hence,

$$\hat{\sigma}^{rs} = (m + n + 1)\, v^{rs}(\hat{\theta}) \quad . \tag{22}$$

Insertion of Equation (22) into (19) gives Equation (18) at the stationary point of $p(\theta,\sigma|y)$. Hence, for rectangular data structures, the same values of $\hat{\theta}$ and $\hat{\sigma}$ are obtained whether one maximizes $p(\theta,\sigma|y)$ or $p(\theta|y)$. Of course, the marginal confidence regions for θ can be estimated more directly in the latter case. The normal equations based on $p(\theta|y)$, given by Stewart and Sørensen (1976), are convenient for this purpose.

The covariance estimates in Equation (21) are maximum-density values, and thus differ from the expectation values $E(\sigma_{rs}|y)$ unless $n-m-p$ is very large. If expectation estimates of the σ_{rs} are desired, one can compute them as the corresponding moments of the normalized posterior density $p(\theta,\sigma|y)$.

EXAMPLE 1. Kinetics of a Three-Component System

Consider the chemical conversion of initially pure species 1 to species 2 and 3 in a batch isothermal reactor. Simulated data for the system are given in Table 1, reproduced from Box and Draper (1965); here y_{ui} is the yield of species i in experiment u. The system is modelled by the differential equations

$$\frac{df_1}{dt} = -k_1 f_1$$

$$\frac{df_2}{dt} = k_1 f_1 - k_2 f_2$$

$$\frac{df_3}{dt} = k_2 f_2$$

which have the solution

$$f_1 = \exp(-k_1 t)$$

$$f_2 = [\exp(-k_1 t) - \exp(-k_2 t)]k_1/(k_2 - k_1)$$

$$f_3 = 1 - f_1 - f_2$$

under the indicated initial conditions. As noted by Box and Draper, it is natural to regard the parameters $\theta_i = \ln k_i$ as uniformly distributed a priori.

There are three responses y_{ui} per experiment. Only two would be linearly independent if the yields were mass-balanced (i.e., if the yields in each row added up to unity). The data in Table 2 are clearly not mass-balanced, so we use all three columns of responses.

The replicates in Table 2 allow preliminary estimation of the parameters σ_{ij}, by the relation

$$s_{ij} = \frac{1}{2n_R} \sum_{r=1}^{n_R} (y_{ri} - y'_{ri})(y_{rj} - y'_{rj}) .$$

Here y_{ri} and y'_{ri} are the observations of response i in the first and second tests of replicate pair r, and n_R is the number of such pairs. This procedure gives

$$\{s_{ij}\} = \begin{Bmatrix} 0.00102 & -0.00128 & 0.00025 \\ -0.00128 & 0.00351 & 0.00024 \\ 0.00025 & 0.00024 & 0.00101 \end{Bmatrix}$$

as a preliminary expectation estimate of $\underset{\sim}{\sigma}$. This is a well-conditioned matrix, so our choice $m = 3$ was correct.

The parameter vector $\underset{\sim}{\psi}$ for the present example consists of θ_1, θ_2, and the six elements on and below the diagonal of $\underset{\sim}{\sigma}$. To test the convergence of the estimation from a poor initial guess, the calculation was started from the initial value shown in Table 3. Convergence was obtained in eight iterations, to the point estimates and 95 percent confidence intervals given there.

Table 2. Data for Example 1, from Box and Draper (1973)

t_u	y_{u1}	y_{u2}	y_{u3}
0.5	0.959	0.025	0.028
0.5	0.914	0.061	0.000
1.	0.855	0.152	0.068
1.	0.785	0.197	0.096
2.	0.628	0.130	0.090
2.	0.617	0.249	0.118
4.	0.480	0.184	0.374
4.	0.423	0.298	0.358
8.	0.166	0.147	0.651
8.	0.205	0.050	0.684
16.	0.034	0.000	0.899
16.	0.054	0.047	0.991

Table 3. Parameter Values for Example 1

Parameter	Initial Value	Solution 1 Eqs. (8,14)*	Solution 2 Eqs. (8,14)*	Solution 3 Eqs. (18,21)*
θ_1	-2.3026	-1.5723±0.0567	-1.5723±0.0558	-1.5723±0.0800
θ_2	0.	-0.7023±0.1374	-0.7023±0.1346	-0.7023±0.1931
σ_{11}	0.01	$(0.76\pm0.52)\ 10^{-3}$	$(0.76\pm0.53)\ 10^{-3}$	$0.76\ 10^{-3}$
σ_{21}	0.	$-(0.50\pm0.63)\ 10^{-3}$	$-(0.50\pm0.63)\ 10^{-3}$	$-0.50\ 10^{-3}$
σ_{22}	0.01	$(1.86\pm1.28)\ 10^{-3}$	$(1.86\pm1.29)\ 10^{-3}$	$1.86\ 10^{-3}$
σ_{31}	0.	$(0.32\pm0.41)\ 10^{-3}$	$(0.32\pm0.41)\ 10^{-3}$	$0.32\ 10^{-3}$
σ_{32}	0.	$(0.40\pm0.62)\ 10^{-3}$	$(0.40\pm0.62)\ 10^{-3}$	$0.40\ 10^{-3}$
σ_{33}	0.01	$(0.77\pm0.54)\ 10^{-3}$	$(0.77\pm0.54)\ 10^{-3}$	$0.77\ 10^{-3}$

*All intervals are 95% highest posterior density regions. In Solution 3, the intervals are computed from the normal equations with "residual mean square" $|\underset{\sim}{v}(\hat{\theta})|/(n-2)$ and $n-2 = 10$ residual degrees of freedom. In Solution 1, the second-derivative terms of Equation (A10) are included.

A second calculation was made with the same initial values, but with second-order θ-derivatives neglected. Convergence was obtained to the same point estimates in nine iterations. The confidence intervals differed slightly, as shown in Table 3.

A third calculation was made by minimizing the determinant $|\underset{\sim}{v}(\theta)|$. Box and Draper (1965) did this by a search procedure; we used the modified Newton algorithm of Stewart and Sørensen (1976), but neglected the second-order θ-derivatives of the functions $f_{ui}(\underset{\sim}{\theta})$. Convergence was obtained in seven iterations, to the same point estimates $\hat{\theta}_i$. The point estimates for the σ_{ij}, computed from Equation (21), also agreed exactly with the two preceding solutions. The one-parameter confidence intervals (computed in this case only for θ_1 and θ_2) are wider than before, and are considered more accurate since in this case the σ_{ij} have been integrated out exactly (Box and Draper, 1965).

EXAMPLE 2. Kinetics of a Five-Component System

Fuguitt and Hawkins (1945, 1947) did extensive experiments on the liquid-phase thermal reactions of α-pinene and its decomposition products. The following products, in order of boiling point, were identified.

A.	α-Pinene	$C_{10}H_{16}$
B.	α - and β-Pyronene	$C_{10}H_{16}$
C.	Dipentene	$C_{10}H_{16}$
D.	allo-Ocimene	$C_{10}H_{16}$
E.	Dimer	$C_{20}H_{32}$

The reaction conditions and yields are reported in Table 3.

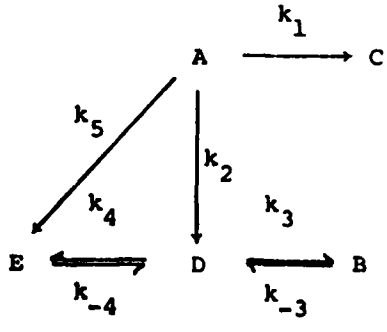
We have normalized the yields to obtain exact mass balances; this makes the yields linearly dependent, and accordingly we have omitted species D. The remaining species are grouped as cumulative distillation fractions:

A, A+B, A+B+C, and E. Each of these responses represents essentially the total mass fraction distilling above or below a particular temperature. The yields of B originally reported in tests 1-15 have been deleted, since they were interpolated values rather than observations (Fuguitt and Hawkins, 1947; Box and co-workers, 1973).

There are numerous gaps in the data. α-Pinene (A) was reported in experiments 1-16, but was considered negligible in the remaining experiments. Pyronenes (B) were reported only in experiments 16-31; they proved difficult to isolate except at small concentrations of α-pinene. Only the dimer fraction (E) was reported in the experiments with allo-ocimene (D) or dimer (E) as feed. The simplified reaction scheme proposed by Fuguitt and Hawkins (1947) implies that α-pinene (A) and dipentene (C) would not be formed in the latter experiments, but that the other three species would be present.

The first eight experiments were used for parameter estimation according to Equation (17) with $m = 3$ by Box and co-workers (1973), and by the present authors (1976). The full 41 experiments could not be so analyzed because of their irregular structure; therefore, only rough estimates were obtainable for several of the reaction parameters. With Equation (8), on the other hand, all 41 experiments can be analyzed.

We postulate the following reaction scheme,



with the following differential equations for the concentrations:

$$\frac{d\phi_A}{dt} = -(k_1 + k_2)\ \phi_A - 2k_5\ \phi_A^2$$

$$\frac{d\phi_B}{dt} = -k_{-3}\ \phi_B + k_3\ \phi_D$$

$$\frac{d\phi_C}{dt} = k_1\ \phi_A$$

$$\frac{d\phi_D}{dt} = k_2\ \phi_A + k_{-3}\ \phi_B - k_3\ \phi_D - 2k_4\ \phi_D^2 + 2k_{-4}\ \phi_E$$

$$\frac{d\phi_E}{dt} = k_5\ \phi_A^2 + k_4\ \phi_D^2 - k_{-4}\ \phi_E$$

Here we have assumed equal densities for the reaction mixture and all species. The ϕ_i are molar concentrations relative to the molar density of pure liquid α-pinene at the reaction temperature. The resulting initial ϕ_i values for the pure reactants are: 1.0 for α-pinene, 1.0 for allo-ocimene, and 0.5 for dimer. The rate coefficients are represented as Arrhenius functions,

$$\ln (k_i) = \theta_i - (1/T - 1/T_B)\ \theta_{i+5} \qquad i = 1,\ldots,5$$

$$\ln (k_3/k_{-3}) = -\theta_{11}/T_B + (1/T - 1/T_B)\ \theta_{13}$$

$$\ln (k_4/k_{-4}) = -\theta_{12}/T_B + (1/T - 1/T_B)\ \theta_{14}$$

with k_i values in min^{-1}, T in Kelvins, and a base temperature T_B of 478.5 K.

The data and parameters were paired to check the feasibility of the estimation. This indicated a sufficient amount of data for estimation of all parameters except σ_{21}. However, the replicate comparisons (u = 18-19,20-21, 22-23,24-25) involving y_{u2} all give duplication of y_{u3}; furthermore each of these comparisons gives a duplication of either y_{u2} or y_{u4}. With these results, we find that neither σ_{32} nor σ_{42} can be estimated; indeed, an attempt to estimate them was terminated by the linear independence test described above. Thereafter, σ_{21}, σ_{32}, and σ_{42} were all fixed at zero, and the remaining parameters were estimated by minimization of S.

Initial values of the θ-parameters were chosen from the results of Fuguitt and Hawkins (1945, 1947), Box and co-workers (1973), and the present authors (1976). Initial variance estimates σ_{ii} were calculated from replicate data available in Table 4, and zeros were inserted initially as covariances.

The model was integrated, for each experiment, by the method of Guertin et al (1977) with 6 mesh points. The coefficients in Equation (9) were computed as described in the Appendix, with first-order sensitivities $\partial\phi_{ui}/\partial\theta_k$ computed by the method of Stewart and Sørensen (1976).

A first minimization, with reaction 5 omitted, converged within 20 iterations. This gave $\hat{S} = 41.06$ with parameter estimates as shown in Table 5. The confidence intervals show the θ's to be estimated quite precisely. The σ_{ij} are estimated less precisely, as anticipated from the limited number of data on several combinations of responses. The deviations of the data from the fitted model are shown in Table 6.

A second minimization of S was done with the full 5-reaction model. This calculation converged to a very flat minimum at $\hat{S} = 34.09$, with parameter estimates as shown in Table 5. The deviations of the data from this fitted model are also shown in Table 6.

The 5-reaction model is better able to describe the polymer yields from α-pinene at short times, as can be seen in Table 6. We can also test the significance of the added parameter θ_5 by use of the confidence intervals. Table 5 gives $\theta_5 = -11.945 \pm 0.698$, based on Equation (14); this implies the limits $(1 \pm 0.698)\exp(-11.945)$ for k_5 with the alternate prior $p(k_5) = c$. Hence, the 95% confidence interval for k_5 does not include zero.

On the other hand, Equations (9) and (13) give the following approximate expression for the 95% confidence region of the 20 fitted parameters of the 5-reaction model:

Table 4. Data for Example 2, from Fuguitt and Hawkins (1945,1947)

Expt. u	Feed	T, C.	t_u, min	w_u**	Normalized yields, weight percent			
					y_{u1} (A)	y_{u2} (A+B)	y_{u3} (A+B+C)	y_{u4} (E)
1	A	189.5	1230.	1	88.3	***	96.2	2.2
2*	A	189.5	1230.	1	88.2	***	95.7	1.3
3	A	189.5	3060.	2	76.4	***	92.7	2.8
4	A	189.5	4920.	2	64.8	***	88.9	5.8
5	A	189.5	7800.	2	50.3	***	84.7	9.3
6	A	189.5	10680.	2	37.5	***	82.0	12.0
7	A	189.5	15030.	2	25.9	***	77.1	17.0
8	A	189.5	22620.	2	14.0	***	73.9	21.0
9	A	204.5	440.	2	86.6	***	95.3	.6
10	A	204.5	825.	2	75.0	***	91.5	1.6
11	A	204.5	1200.	2	66.0	***	88.8	3.4
12	A	204.5	1500.	2	59.4	***	86.4	5.1
13	A	204.5	2040.	2	48.9	***	83.0	8.3
14	A	204.5	3060.	2	32.8	***	77.8	13.8
15	A	204.5	6060.	2	11.5	***	70.4	22.5
16	A	189.5	36420.	2	4.5	7.4	70.5	25.7
17	A	204.5	16020.	2	-	3.1	66.2	28.6
18	A	225.0	3000.	1	-	3.0	66.0	28.0
19*	A	225.0	3000.	1	-	4.0	66.0	28.0
20	A	245.0	630.	1	-	4.0	65.0	27.0
21*	A	245.0	630.	1	-	5.0	65.0	27.0
22	A	265.0	120.	1	-	7.0	65.0	23.0
23*	A	265.0	120.	1	-	7.0	65.0	24.0
24	A	285.0	30.	1	-	11.0	66.0	19.0
25*	A	285.0	30.	1	-	9.0	66.0	19.0
26	D	189.5	1020.	1	-	-	-	80.0
27	D	189.5	3990.	1	-	-	-	87.3
28*	D	189.5	3990.	1	-	-	-	87.3
29	D	189.5	6780.	1	-	-	-	87.5
30	D	189.5	8220.	1	-	-	-	86.5
31	D	189.5	13260.	1	-	-	-	88.5
32	D	189.5	14760.	1	-	-	-	89.8
33	D	204.5	3480.	1	-	-	-	87.5
34	D	204.5	5700.	1	-	-	-	86.8
35	E	189.5	8880.	1	-	-	-	91.9
36*	E	189.5	8880.	1	-	-	-	92.0
37	E	189.5	14340.	1	-	-	-	89.8
38	E	189.5	23400.	1	-	-	-	89.7
39*	E	189.5	23400.	1	-	-	-	88.5
40	E	204.5	5700.	1	-	-	-	88.4
41	E	204.5	8100.	1	-	-	-	87.9

* Replicate of the preceding test.

** w_u is the number of independent tests combined to obtain each observation y_{ui}.

*** Originally reported but not observed; see text.

- No value reported.

Table 5. Parameters for α-Pinene Conversion

Parameter	Estimates* for 4-Reaction Model	Estimates* for 5-Reaction Model
θ_1	-8.331 ± .024	-8.333 ± .025
θ_2	-8.898 ± .029	-8.961 ± .054
θ_3	-8.242 ± .341	-8.196 ± .325
θ_4	-5.389 ± .081	-5.438 ± .087
θ_5		-11.945 ± .698
θ_6	19814. ± 428.	19785. ± 457.
θ_7	20828. ± 474.	20890. ± 536.
θ_8	17336. ± 4079.	17212. ± 4203.
θ_9	10321. ± 915.	10322. ± 918.
θ_{10}		19957. **
θ_{11}	269. ± 83.	279. ± 83.
θ_{12}	-1976. ± 64.	-1985. ± 63.
θ_{13}	-336. ± 950.	-259. ± 958.
θ_{14}	-3873. ± 1624.	-3781. ± 1555.
σ_{11}	.696 ± .419	.784 ± .492
σ_{21}	.000 **	.000 **
σ_{22}	.391 ± .359	.376 ± .348
σ_{31}	.358 ± .412	.426 ± .456
σ_{32}	.000 **	.000 **
σ_{33}	.706 ± .426	.732 ± .444
σ_{41}	-.248 ± .344	-.294 ± .354
σ_{42}	.000 **	.000 **
σ_{43}	-.504 ± .317	-.493 ± .314
σ_{44}	.744 ± .304	.654 ± .282

*95% highest posterior density intervals calculated from Equation (14).

**Posterior estimates were not obtained for these parameters.

Table 6. Final Residuals $\varepsilon_{ui}(\hat{\theta})$ for Example 2.

	4-Reaction Model				5-Reaction Model			
Expt., u	ε_{u1} (A)	ε_{u2} (A+B)	ε_{u3} (A+B+C)	ε_{u4} (E)	ε_{u1} (A)	ε_{u2} (A+B)	ε_{u3} (A+B+C)	ε_{u4} (E)
1	-1.32	-	-.37	2.00	-1.22	-	-.26	1.69
2	-1.42	-	-.87	1.10	-1.32	-	-.76	.79
3	.26	-	.24	.88	.43	-	.43	.45
4	.28	-	-.15	1.10	.45	-	.06	.72
5	.38	-	-.04	.22	.48	-	.12	-.04
6	-1.13	-	.70	-.81	-1.12	-	.78	-.96
7	-.32	-	-.26	-.17	-.43	-	-.29	-.18
8	.66	-	.89	-1.06	.47	-	.74	-.92
9	.88	-	.21	.30	1.00	-	.35	-.11
10	.10	-	-.17	.14	.24	-	.04	-.38
11	.31	-	-.07	.16	.45	-	.15	-.34
12	.27	-	-.51	.23	.38	-	-.29	-.23
13	-.04	-	-.86	.42	.01	-	-.69	.08
14	-1.44	-	-1.56	.90	-1.52	-	-1.49	.75
15	-.47	-	-1.47	.63	-.70	-	-1.61	.77
16	.60	.78	.98	-.36	.44	.72	.74	-.14
17	-	-.12	-.67	.34	-	-.07	-.87	.50
18	-	-.81	.51	-.48	-	-.76	.38	-.39
19	-	.19	.51	-.48	-	.24	.38	-.39
20	-	-.89	.29	-.56	-	-.88	.22	-.47
21	-	.11	.29	-.56	-	.13	.22	-.47
22	-	-.54	-.31	-.37	-	-.58	-.32	-.28
23	-	-.54	-.31	.63	-	-.58	-.32	.72
24	-	1.54	.49	-.15	-	1.51	.58	-.20
25	-	-.46	.49	-.15	-	-.49	.58	-.20
26	-	-	-	1.12	-	-	-	1.95
27	-	-	-	-.92	-	-	-	-.61
28	-	-	-	-.92	-	-	-	-.61
29	-	-	-	-1.31	-	-	-	-1.16
30	-	-	-	-2.37	-	-	-	-2.27
31	-	-	-	-.41	-	-	-	-.42
32	-	-	-	.90	-	-	-	.86
33	-	-	-	.67	-	-	-	.72
34	-	-	-	-.31	-	-	-	-.40
35	-	-	-	1.26	-	-	-	1.17
36	-	-	-	1.36	-	-	-	1.27
37	-	-	-	.24	-	-	-	.16
38	-	-	-	.91	-	-	-	.80
39	-	-	-	-.29	-	-	-	-.40
40	-	-	-	.42	-	-	-	.27
41	-	-	-	.51	-	-	-	.35

$$S(\underset{\sim}{\psi}) - 34.09 < \chi^2_{20}(0.05) = 31.41.$$

All $\underset{\sim}{\psi}$ values such that $S(\underset{\sim}{\psi}) < 65.50$ lie within this estimated 95% joint confidence region. By this criterion, the model with $k_5 = 0$ is acceptable. However, as indicated earlier, Equation (14) is more reliable than (13). For this, and a study of the residuals, we conclude that the 5-reaction model is to be preferred.

ACKNOWLEDGMENT

This work was sponsored by the National Science Foundation through Grant ENG76-24368; by the United States Army under Contract DAAG29-75-C-0024; and by the Wisconsin Alumni Research Foundation through funds administrated by the Graduate School of the University of Wisconsin.

BIBLIOGRAPHY

[1] Box, G. E. P. and Draper, N. R. (1965).
The Bayesian estimation of common parameters from several responses.
Biometrika, 52, 355-365.

[2] Box, G. E. P., Hunter, W. G., MacGregor, J. F., and Erjavec, J. (1973).
Some problems associated with the analysis of multiresponse data.
Technometrics, 15, 33-51.

[3] Box, G. E. P. and Tiao, G. C. (1973).
Bayesian Inference in Statistical Analysis. Reading, Massachusetts:
Addison-Wesley, Inc.

[4] Fuguitt, R. E. and Hawkins, J. E. (1945).
The liquid-phase thermal isomerization of α-pinene.
J. A. C. S., 67 , 242-245.

[5] Fuguitt, R. E. and Hawkins, J. E. (1947).
Rate of the thermal isomerization of α-pinene in the liquid phase.
J. A. C. S., 69 , 319-322.

[6] Guertin, E. W., Sørensen, J. P., and Stewart, W. E. (1977).
Exponential collocation of stiff reactor models.
Computers and Chemical Engineering, 1 , 197-203.

[7] Stewart, G. W. (1973)
Introduction to Matrix Computations. New York: Academic Press, Inc.

[8] Stewart, W. E. and Sørensen, J. P. (1976).
Sensitivity and regression of multicomponent reactor models.
Fourth International Symposium on Chemical Reaction Engineering.
Frankfurt: DECHEMA, pp. I-12 to I-20.

[9] Wilks, S. S. (1962).
Mathematical Statistics. New York: John Wiley and Sons, Inc.

WES:JPS/db

APPENDIX: DERIVATIVES OF S .

The matrices $\underset{\sim}{\sigma}_u$ are real and symmetric; furthermore, S is defined only when these matrices are positive definite. The following derivative relations then hold:

$$\frac{\partial \ln |\underset{\sim}{\sigma}_u|}{\partial \sigma_{uij}} = (2 - \delta_{ij})\sigma_u^{ij} \qquad j \leq i \tag{A1}$$

$$\frac{\partial \sigma_u^{ij}}{\partial \sigma_{uk\ell}} = -\tfrac{1}{2}(2 - \delta_{k\ell})[\sigma_u^{ik}\sigma_u^{\ell j} + \sigma_u^{i\ell}\sigma_u^{kj}] \qquad \ell \leq k. \tag{A2}$$

The relations for second derivatives follow by combination of (A1) and (A2):

$$\frac{\partial^2 \ln |\underset{\sim}{\sigma}_u|}{\partial \sigma_{uk\ell}\partial \sigma_{uij}} = -\tfrac{1}{2}(2 - \delta_{ij})(2 - \delta_{k\ell})[\sigma_u^{ik}\sigma_u^{\ell j} + \sigma_u^{i\ell}\sigma_u^{kj}] \qquad j \leq i,\ \ell \leq k \tag{A3}$$

$$\frac{\partial^2 \sigma_u^{ij}}{\partial \sigma_{ust}\partial \sigma_{uk\ell}} = \tfrac{1}{2}(2 - \delta_{k\ell})(2 - \delta_{st})$$
$$[(\sigma_u^{is}\sigma_u^{tk} + \sigma_u^{it}\sigma_u^{sk})\sigma_u^{\ell j} + \sigma_u^{ik}(\sigma_u^{\ell s}\sigma_u^{tj} + \sigma_u^{\ell t}\sigma_u^{sj})$$
$$+ (\sigma_u^{is}\sigma_u^{t\ell} + \sigma_u^{it}\sigma_u^{s\ell})\sigma_u^{kj} + \sigma_u^{i\ell}(\sigma_u^{ks}\sigma_u^{tj} + \sigma_u^{kt}\sigma_u^{sj})]$$
$$\ell \leq k\ ,\quad t \leq s\ . \tag{A4}$$

As indicated earlier, if response h is absent from experiment u , the elements σ_{uhj} and σ_{ujh} are replaced by the constant dummy values δ_{hj} . Note also that the symmetry of $\underset{\sim}{\sigma}_u$ has been used to express these derivatives in terms of elements on and below the diagonal.

The derivatives required for Equation (9) are obtained as follows:

$$-\tfrac{1}{2}\frac{\partial S}{\partial \theta_r} = -\sum_u \sum_i \sum_{j \leq i} \tfrac{1}{2}(2 - \delta_{ij})\sigma_u^{ij}\frac{\partial}{\partial \theta_r}(\varepsilon_{ui}\varepsilon_{uj}) \tag{A5}$$

$$\tfrac{1}{2}\frac{\partial^2 S}{\partial \theta_r \partial \theta_v} = \sum_u \sum_i \sum_{j \leq i} \tfrac{1}{2}(2 - \delta_{ij})\sigma_u^{ij}\frac{\partial^2}{\partial \theta_r \partial \theta_v}(\varepsilon_{ui}\varepsilon_{uj}) \tag{A6}$$

$$\tfrac{1}{2}\frac{\partial^2 S}{\partial\theta_r \partial\sigma_{k\ell}} = \sum_u \sum_i \sum_{j\le i} \tfrac{1}{2}(2-\delta_{ij})\frac{\partial\sigma_u^{ij}}{\partial\sigma_{ukl}}\frac{\partial}{\partial\theta_r}(\varepsilon_{ui}\,\varepsilon_{uj}) \qquad \text{(A7)}$$

$$-\tfrac{1}{2}\frac{\partial S}{\partial\sigma_{st}} = -\tfrac{1}{2}(m+1)\frac{\partial \ln|\underset{\sim}{\sigma}|}{\partial\sigma_{st}} - \sum_{u=1}^{n}\tfrac{1}{2}\frac{\partial \ln|\underset{\sim}{\sigma}_u|}{\partial\sigma_{st}}$$

$$-\sum_u\sum_i\sum_{j\le i}\tfrac{1}{2}(2-\delta_{ij})\frac{\partial\sigma_u^{ij}}{\partial\sigma_{st}}\,\varepsilon_{ui}\,\varepsilon_{uj} \qquad \text{(A8)}$$

$$\tfrac{1}{2}\frac{\partial^2 S}{\partial\sigma_{st}\partial\sigma_{k\ell}} = \tfrac{1}{2}(m+1)\frac{\partial^2 \ln|\underset{\sim}{\sigma}|}{\partial\sigma_{st}\partial\sigma_{k\ell}} + \sum_u \tfrac{1}{2}\frac{\partial^2 \ln|\underset{\sim}{\sigma}_u|}{\partial\sigma_{ust}\partial\sigma_{uk\ell}}$$

$$+\sum_u\sum_i\sum_{j\le i}\tfrac{1}{2}(2-\delta_{ij})\frac{\partial^2\sigma_u^{ij}}{\partial\sigma_{ust}\partial\sigma_{uk\ell}}\,\varepsilon_{ui}\,\varepsilon_{uj} \qquad \text{(A9)}$$

Equations (A6), (A7), and (A9) evaluated at $\underset{\sim}{\theta}_0$ and $\underset{\sim}{\sigma}_0$ provide the coefficient matrix $\underset{\sim}{A}$ of the normal equations. Equations (A5) and (A8) give the right-hand column vector.

The residuals ε_{ui} and ε_{uj} are expressed as functions of $\underset{\sim}{\theta}$ by use of Equation (1). The θ-derivative in Equation (A6) is expanded to give:

$$\frac{\partial^2}{\partial\theta_r\partial\theta_v}(\varepsilon_{ui}\,\varepsilon_{uj}) = \frac{\partial\varepsilon_{ui}}{\partial\theta_r}\frac{\partial\varepsilon_{uj}}{\partial\theta_v} + \frac{\partial\varepsilon_{uj}}{\partial\theta_r}\frac{\partial\varepsilon_{ui}}{\partial\theta_v}$$

$$+\varepsilon_{ui}\frac{\partial^2\varepsilon_{uj}}{\partial\theta_r\partial\theta_v} + \varepsilon_{uj}\frac{\partial^2\varepsilon_{ui}}{\partial\theta_r\partial\theta_v} \qquad \text{(A10)}$$

The second-derivative terms are unimportant if the data are well fitted; compare Solutions 1 and 2 in Table 3.

If the experiments have different weights w_u as in Table 4, then $\varepsilon_{ui}\,\varepsilon_{uj}$ and its derivatives should be multiplied by w_u throughout the development. As usual, the matrix $\underset{\sim}{\sigma}$ is defined for experiments of unit weight.

(14) MRC-TSR-2037

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
2037	AD-A083 819	(9) Technical

4. TITLE (and Subtitle)

(6) NEW ALGORITHMS FOR NONLINEAR LEAST SQUARES AND BAYESIAN PARAMETER ESTIMATION.

5. TYPE OF REPORT & PERIOD COVERED
Summary Report - no specific reporting period

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

(10) Warren E. Stewart and Jan P. Sørensen

8. CONTRACT OR GRANT NUMBER(s)

(15) DAAG29-75-C-0024, NSF-ENG76-24368

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Mathematics Research Center, University of Wisconsin
610 Walnut Street
Madison, Wisconsin 53706

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
#2 - Physical Mathematics

11. CONTROLLING OFFICE NAME AND ADDRESS
See Item 18

12. REPORT DATE
(11) February 1980

13. NUMBER OF PAGES
(12) 27

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, DC 20550

Wisconsin Alumni Research Foundation
University of Wisconsin
Graduate School
Madison, WI 53706

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Parameter estimation, Bayesian estimation, Least squares, Nonlinear estimation, Covariance estimation, Multiple-response systems, Missing values, Statistics.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

New algorithms are described for Bayesian estimation of parameters in non-linear models of multiple-response systems. Modal and interval estimates are provided for the parameter vector θ of the predictor model, and for the variance-covariance matrix σ of a Normal error distribution. Allowance is made for gaps (missing values of responses), such as commonly occur in practice. Two chemical examples are analyzed.

221 200

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)