TECHNOMETRICS© VOL. 15, NO. 1 FEBRUARY, 1973

ASQC Chemical Division Technical Conference
1971 Prize Winning Paper

Some Problems Associated with the Analysis of Multiresponse Data

G. E. P. Box*
W. G. HUNTER
J. F. MACGREGOR**
J. ERJAVEC***

The University of Wisconsin
Madison, Wisconsin

Experience has shown that unless special care is exercised in analyzing multiresponse data serious mistakes can be made. In this paper some problems associated with fitting multiresponse models are identified and discussed. In particular, three kinds of dependencies are considered: dependence among the errors, linear dependencies among the expected values of the responses, and linear dependencies in the data. Since ignoring such dependencies can lead to difficulties, a method is described for detecting and handling them. The concepts involved are illustrated with a chemical example.

KEY WORDS

Estimation
Multiresponse data
Linear dependencies
Eigenvalue-Eigenvector analysis
Nonlinear models
Chemical kinetics
Multivariate data

1. INTRODUCTION

Engineers and scientists frequently need to analyze multiresponse data. When studying a chemical reaction for instance, for each setting of a group of "input" variables determining the reaction conditions, not one but a number of "output" variables or responses (such as the concentrations of each of the chemical constituents) may be measured. The capability of making such multiple measurements has greatly increased with the advent of better analytical tools such as the gas chromatograph. This capability has, in turn, increased the potential information generated by a particular experimental run, making possible more precise discrimination among models, more adequate checking of models, and more accurate estimation of parameters.

But with the capability to measure with comparative ease all the substituents in the reaction mixture comes the necessity to take account of possible dependencies

* Supported by the Air Force Office of Scientific Research under Grant AF-AFOSR 72-2362.
** Present address: Dept. of Chemical Engineering, McMaster Univ., Hamilton, Ontario
*** Present address: American Cyanamid, Bound Brook, New Jersey
Received Sept. 1970; revised May 1972.

among them (arising, for example, as a consequence of the law of conservation of matter). If a particular chemical mechanism is given, it may follow, for example, that to maintain the carbon balance a certain linear relationship must exist among the amounts of the substituents in *all* the experimental runs. A second such relationship may exist to maintain the nitrogen balance and so on. Relationships of this kind are called stoichiometric. Now in practice the experimenter is *not* given the mechanism but must learn about it as a result of an iterative conversation between the generated data and the theoretical possibilities sparked off in his mind. He tentatively entertains possibilities and by suitably planned experiments and suitable analysis he allows the data to comment on these.

Various kinds of problems all associated with dependencies of one kind or another then arise:

(i) Whether or not stoichiometric type dependencies exist, to fit a tentative model to the data one must properly take account of *correlations* between the errors in different substituents.

(ii) If the nature of dependencies among the expected values of the various substituents is known, this knowledge might confirm or deny the validity of the model under study or suggest the appropriateness of some new mechanistic form.

(iii) The investigator frequently uses one or more stochiometric type relations to deduce the presumed values of substituents which are difficult to measure, or to adjust measured values to agree with these relations. He may be unaware of the implications of such practices on the estimation process and the data analyst may not always be aware of what the experimenter has done.

Since the advent of new methods of chemical analysis and increasing interest in mechanistic studies using multiresponse data, we have encountered a rash of problems which point to the need for discussing and distinguishing the different types of dependencies and for providing methods for use in practical data analysis whereby mistakes can be avoided, and forgotten or unknown relationships can be made manifest.

2. Three Kinds of Dependencies

Suppose that, in an experimental program, n sets of reaction conditions (not necessarily all different) are run, and at each set of conditions r responses $(y_1 , y_2 , \cdots , y_i , \cdots , y_r)$ are recorded. Suppose furthermore that we can write a mathematical model for the ith response at the uth set of reaction conditions

$$y_{iu} = \eta_i(\xi_u , \theta) + \epsilon_{iu} \qquad \begin{matrix} i = 1, \cdots , r \\ u = 1, \cdots , n \end{matrix} \tag{1}$$

where ϵ_{iu} is the error in the ith response for the uth run, θ are unknown parameters and ξ_u are the values of the input variables defining the reaction conditions for the uth run.

Three kinds of dependencies among the responses will be considered in this paper along with the effects of each on the fitting of multiresponse models.

2.1 *Dependence Among the Errors*

Consider the r errors committed in the uth run, $\epsilon_u' = (\epsilon_{1u} , \epsilon_{2u} , \cdots , \epsilon_{ru})$. It will usually be true that these errors are correlated. It is important that the statistical treatment of the data should take account of this correlation, and this has not always been done. For example, one technique for estimating parameters which

has been used (for example, see Ball (1966)) is to find those parameters which minimize the overall residual sum of squares from all the responses

$$\text{RSS} = \sum_{i=1}^{r} \sum_{u=1}^{n} [y_{iu} - \eta_i(\xi_u , \boldsymbol{\theta})]^2. \tag{2}$$

However, it is easily seen (Box & Draper (1965), Hunter (1967)) that this criterion is appropriate only if (a) the errors are all uncorrelated and (b) the errors all have equal variances. In practice neither of these circumstances is likely to be true, and analysis of data as if it were true can give incorrect results (Eakman (1969), Erjavec (1969)). A method which overcomes these difficulties was developed by Box and Draper (1965). Assuming that the errors were distributed according to a multivariate Normal distribution with unknown variance-covariance matrix $\boldsymbol{\Sigma} = E(\boldsymbol{\varepsilon}_u \boldsymbol{\varepsilon}_u')$ and using a "non-informative" prior distribution, they showed that, given the data, the posterior distribution for all the parameters, $\boldsymbol{\theta}$, is proportional to

$$|\mathbf{V}_r|^{-n/2} \tag{3}$$

where

$$V = \{v_{ij}\} \quad \text{and} \quad v_{ij} = \sum_{u=1}^{n} (y_{iu} - \eta_{iu})(y_{ju} - \eta_{ju}).$$

Estimates of $\boldsymbol{\theta}$ yielding maximum posterior density are those obtained by minimizing the determinant $|\mathbf{V}_r|$. In the particular case of a single response, this procedure leads to the method of least squares.

2.2 *Linear Dependence Among Expected Values of Responses*

In chemical systems stoichiometry, material and energy balances, or steady-state conditions will usually require that certain linear relationships exist among the expected values of the responses. For example, for every run, stoichiometry may dictate that the sum of the expected values of the number of moles of the r constituents in the system must be the same. That is,

$$\sum_{i=1}^{r} E(y_{iu}) = \sum_{i=1}^{r} \eta_i(\xi_u , \boldsymbol{\theta}) = a_0 \qquad u = 1, \cdots , n \tag{4}$$

More generally, there may be m independent linear relations among the expectations which must be satisfied for each run,

$$\sum_{i=1}^{r} a_{qi} E(y_{iu}) = \sum_{i=1}^{r} a_{qi} \eta_i(\xi_u , \boldsymbol{\theta}) = a_{q0} \qquad \begin{matrix} u = 1, \cdots , n \\ q = 1, \cdots , m \end{matrix} \tag{5}$$

Or, in matrix notation,

$$\mathbf{A}E(\mathbf{y}_u) = \mathbf{a}_0 , \qquad u = 1, \cdots , n \tag{6}$$

where $\mathbf{A} = \{a_{qi}\}$ is an $m \times r$ matrix and $\mathbf{y}_u = (y_{1u} , \cdots , y_{ru})'$ is the uth data vector.

2.3 *Linear Dependencies in the Data*

Suppose an experimenter knows that a formula such as (4), for example, expressing some material balance relationship must be true for each of his experimental runs and yet, because of experimental error, it is not exactly satisfied by his observed responses. As is commonly done in practice, the experimenter may force his observations to fit his relationship by some normalizing calculation to give

$$y_1 + y_2 + \cdots + y_r = a_0 \tag{7}$$

This is often done by multiplying each of the originally measured values of the

r observations y_{1u}^* , y_{2u}^* , $\cdots$, y_{ru}^* for a given run by the factor $a_0/\sum_{u=1}^{r} y_u^*$. In so doing he produces an exact linear dependence among the responses y_1 , y_2 , $\cdots$, y_r . Another way in which exact linear dependence in the data is introduced is when the investigator measures only $r - 1$ responses independently and calculates the rth response using the relationship (7). More generally, by placing absolute trust in the specific model he favors, the experimenter might make use of $m_1 \leq m$ of the relationships in (6). Thus he might make only $r - m_1$ independent determinations of the substituents and obtain the others by calculation, or, he might complete r independent determinations and then force the m_1 relationships to hold by normalizing or otherwise adjusting the data. Suppose the first $m_1 < m$ relations are used for normalizing, and the matrix $\mathbf{A}$ and the vector $\mathbf{a}_0$ are partitioned after the m_1th row so that (6) may be written

$$\begin{bmatrix} \mathbf{A}_1 \\ \hline \mathbf{A}_2 \end{bmatrix} E(\mathbf{y}_u) = \begin{bmatrix} \mathbf{a}_{10} \\ \hline \mathbf{a}_{20} \end{bmatrix}$$

If $\mathbf{y}_u$ is an $r \times 1$ vector of "data" values including those obtained by calculation, then there will be the following m_1 exact linear relationships connecting each data vector

$$\mathbf{A}_1\mathbf{y}_u = \mathbf{a}_{10} \tag{8}$$

Whenever possible, of course, normalizing or adjusting "data" in the manner described above should be avoided. The experimenter should be prepared to go to some trouble to determine each response independently, and having done so he should refrain from forcing observed responses to satisfy theoretical relationships that he believes to be true. He should do this because independent information on each of the m relationships can not only provide better estimates of the parameters but also make possible a more comprehensive check on the model which is currently being entertained. In some situations, however, the avoidance of linear dependencies in the data is not possible. In some instances the analytical procedures or equipment necessarily make use of such relationships. Such is the case with chemical composition data obtained from gas chromatographs, for example, where it is only possible to calculate relative percentages.

When it is impossible to determine all the substituents independently or to avoid "normalized" data, careful note should be made of which observations are independent, which were obtained by calculation, and which of the expectation relationships (6) have been employed in obtaining the data. Ignoring such dependencies can lead to serious mistakes in interpretation. Unfortunately, even if the investigator can identify the precise nature of dependencies in the data, he may be unaware of the importance of ensuring that they are taken account of in the subsequent statistical analysis.

Careful preliminary analysis of the system under study ought in principle to reveal all the dependencies in the expected values of the responses, and adequate inquiry ought to show which of the relationships may have been utilized, either consciously or otherwise, to normalize the results. Unfortunately, we have found from bitter experience that in practice dependencies are frequently overlooked and we therefore now regard it a practical necessity to look for such relationships empirically as a preliminary to the analysis of multiresponse data. Frequently, the results of such an analysis reveals unexpected but highly informative dependencies. Further analysis should not be proceeded with until these dependencies have been satisfactorily explained.

3. Eigenvalue-Eigenvector Analysis

Suppose the data analyst were a priori unaware of the nature of the possible linear relationships existing among the responses. He could then proceed as follows. Ordinarily, the vector of constants $\mathbf{a}_0$ on the right of (6) are unknown, but they can be eliminated by working with the matrix $\mathbf{D} = \{d_{iu}\} = \{y_{iu} - \bar{y}_i\}$ of deviations from the individual averages. The eigenvalues λ_k and the r-dimensional eigenvectors $\mathbf{z}_k$ of $\mathbf{DD}'$ are such that

$$\mathbf{z}_k'\mathbf{DD}' = \lambda_k\mathbf{z}_k' \tag{9}$$

and the eigenvectors can be normalized so that

$$\mathbf{z}_k'\mathbf{z}_k = 1 \qquad k = 1, \cdots, r. \tag{10}$$

If there are m_1 independent exact linear relations in the *data*

$$\mathbf{A}_1\mathbf{y}_u = \mathbf{a}_{10} \qquad u = 1, 2, \cdots, n \tag{11}$$

then

$$\mathbf{A}_1\mathbf{d}_u = \mathbf{0} \qquad u = 1, 2, \cdots, n. \tag{12}$$

In this case therefore there will be m_1 zero eigenvalues, $\lambda_1 = 0, \cdots, \lambda_{m_1} = 0$, and the associated m_1 eigenvectors $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{m_1}$ will define the same hyperplane as do the m_1 rows of $\mathbf{A}_1$. Thus, if $\mathbf{Z}_1$ is the $m_1 \times r$ matrix whose rows consist of these eigenvectors, then a non-singular transformation exists such that

$$\mathbf{A}_1 = \mathbf{T}\mathbf{Z}_1 \tag{13}$$

In general, in addition to the m_1 exact linear relations in the *data* we will have $m_2 = m - m_1$ further linear relations in the *expected values* so that

$$\mathbf{A}_2\mathbf{d}_u = \mathbf{e}_u \qquad u = 1, \cdots, n \tag{14}$$

where $\mathbf{A}_2$ is an $m_2 \times r$ matrix of coefficients and $\mathbf{e}_u$ is an $m_2 \times 1$ vector of errors all of whose elements have expected value zero. To correspond with these relations one can expect a further m_2 small eigenvalues whose expected size depends on the experimental errors via the following relation derived in the Appendix:

$$E(\lambda_k) = (n - 1)\mathbf{z}_k'\mathbf{\Sigma}\mathbf{z}_k \qquad k = m_1 + 1, \cdots, m_1 + m_2 \tag{15}$$

where $\mathbf{\Sigma} = E(\boldsymbol{\varepsilon}_u\boldsymbol{\varepsilon}_u')$ is the $r \times r$ variance-covariance matrix for the errors in the r responses. The corresponding eigenvectors (which we suppose form the rows of the $m_2 \times r$ matrix $\mathbf{Z}_2$) define a hyperplane which approximately coincides with the hyperplane given by those components of the m_2 vectors in $\mathbf{A}_2$ which are orthogonal to $\mathbf{A}_1$.

In the situation envisaged, then, there would be

(i) m_1 eigenvalues produced by adjustment of data which differ from zero, if at all, only because of rounding error*;

(ii) a further $m_2 = m - m_1$ eigenvalues produced by other relations among the expected values and whose magnitudes are determined by experimental error;

(iii) $r - m$ values which would typically be very much larger since they would be quadratic functions of the changes in response produced by changes in the experimental conditions.

These three different kinds of roots usually differ in size by orders of magnitude. To separate them into appropriate classes all that is usually needed is a rough estimate of the anticipated size of the λ's associated with the expectation rela-

* For a precise measure of how large these near zero eigenvalues could be a covariance matrix $\mathbf{\Sigma}_{re}$ for rounding error could be substituted in (15).

tionships given by (15). When runs have been replicated, then the sample estimate $\hat{\mathbf{\Sigma}}$ may be substituted in (15) to yield estimates of the λ_k's. When no replicated experiments exist sometimes previous experimentation will provide a rough value for $\mathbf{\Sigma}$. Since only a very approximate "order of magnitude" value is really needed, if high correlations between experimental errors in the various responses are not expected it will be enough to approximate $\mathbf{\Sigma}$ by $\mathbf{I}\bar{\sigma}^2$ where $\bar{\sigma}^2$ is an average value for the residual error variance of the r responses. Then

$$E(\lambda_k) \simeq (n-1)\mathbf{z}_k'\mathbf{I}\mathbf{z}_k\bar{\sigma}^2 = (n-1)\bar{\sigma}^2 \tag{16}$$

If the order of magnitude of the r responses is very different, for example because of different measurement units, then it would be better to approximate $\mathbf{\Sigma}$ by a diagonal matrix containing the appropriate estimates of the variances of the individual responses.

After proceeding with the subsequent main analysis in which the parameters $\boldsymbol{\theta}$ are estimated a value for $\mathbf{\Sigma}$ will be available from the residuals so that at this later stage we can return to the preliminary analysis and recheck for agreement.

4. A Simple Illustration

For illustration, suppose that $r = 3$ constituents can be measured, and that

$$\text{(i)}\quad E(y_1) + E(y_2) + E(y_3) = 6 \tag{17}$$

$$\text{(ii)}\quad E(y_1) - 2E(y_2) = 3. \tag{18}$$

Suppose, in fact, the experimenter chemically determines y_1 and y_2 separately but estimates y_3 "by difference" according to

$$y_1 + y_2 + y_3 = 6 \tag{19}$$

(This same linear dependency (19) would also result if the experimenter measured all three responses for a given run u, y_{1u}^*, y_{2u}^*, and y_{3u}^*, and then "normalized" the data by multiplying each measurement for that run by the factor $6/(y_{1u}^* + y_{2u}^* + y_{3u}^*)$). Then a typical set of data might be as follows:

$$\mathbf{Y}' = \begin{array}{c} \begin{matrix} y_1 & y_2 & y_3 \end{matrix} \\ \begin{bmatrix} 2 & 0 & 4 \\ -1 & -2 & 9 \\ 3 & 0 & 3 \\ 6 & 1 & -1 \\ 5 & 1 & 0 \end{bmatrix} \end{array} \tag{20}$$

with averages $\bar{y}_1 = 3$, $\bar{y}_2 = 0$, and $\bar{y}_3 = 3$. The relationship (19) is exactly true for every row, but because y_1 and y_2 are subject to error, the relationship (18) yields

$$(y_1 - \epsilon_1) - 2(y_2 - \epsilon_2) = 3 \tag{21}$$

so that

$$y_1 - 2y_2 = 3 + e \tag{22}$$

where

$$e = \epsilon_1 - 2\epsilon_2\,. \tag{23}$$

The matrix **D** of deviations from the individual averages is

$$\mathbf{D}' = \begin{bmatrix} -1 & 0 & 1 \\ -4 & -2 & 6 \\ 0 & 0 & 0 \\ 3 & 1 & -4 \\ 2 & 1 & -3 \end{bmatrix} \tag{24}$$

and

$$\mathbf{DD}' = \begin{bmatrix} 30 & 13 & -43 \\ 13 & 6 & -19 \\ -43 & -19 & 62 \end{bmatrix} \tag{25}$$

The eigenvalues and eigenvectors for **DD′** are shown in Table 1. Since λ_1 is zero, $\mathbf{z}_1$ corresponds to an exact linear dependence among the responses,

$$0.5774\, y_{1u} + 0.5774\, y_{2u} + 0.5774\, y_{3u} = \text{a constant} \tag{26}$$

Or, after multiplying both sides by 1/0.5774, we have

$$y_{1u} + y_{2u} + y_{3u} = \text{a constant} \tag{27}$$

TABLE 1

The Eigenvalues and Eigenvectors of **DD′** *for the Simple Illustration.*

i	λ_i	$\underset{\sim}{z}_i'$		
1	0	(0.5774	0.5774	0.5774)
2	0.338	(0.6006	-0.7793	0.1787)
3	97.66	(-0.5531	-0.2436	0.7967)

agreeing with (19). Suppose that the error variance for the substituents y_{iu} is known to be of the order of 0.1. Since λ_2 (= 0.338) is of the same order of magnitude as $(n - 1)\bar{\sigma}^2$ (= 4 × 0.1 = 0.4) we anticipate that $\mathbf{z}_2$ corresponds to a linear relationship among the expected values of the responses, or, more specifically, to that part of the relationship which is orthogonal to $\mathbf{z}_1$. To see that this is so, note that the component of $\mathbf{A}_2 = (1, -2, 0)$ which is orthogonal to $\mathbf{A}_1 = (1, 1, 1)$ is (0.617, −0.772, 0.154) when normalized. This vector is very close to the eigenvector $\mathbf{z}_2$ as expected. We see in this simple illustration how eigenvalue-eigenvector analysis helps in uncovering the linear relationships which exist among the observations and among the expected values.

5. Implications in Data Fitting

When there are r responses containing m_1 exact linear relations the $r \times r$ matrix $\mathbf{V}_r$ formed as in equation (3) has rank $r - m_1$. Attempts which have occasionally

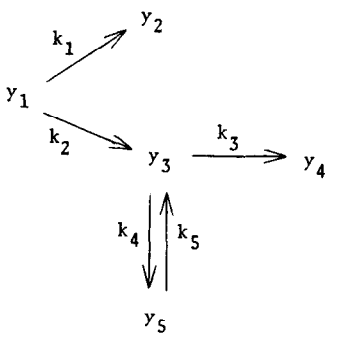
been made to make inferences about the parameters by studying $|\mathbf{V}_r|$ have led to nonsensical results since this determinant is equal to zero for all values of the parameters. Small rounding errors have further confused the picture since, when these are present, $|\mathbf{V}_r|$ will be not quite zero and will change as the parameters are changed. A correct analysis is obtained by studying $|\mathbf{V}_{r-m_1}|$. This $(r - m_1) \times (r - m_1)$ determinant should contain only those responses which have been independently determined (or more generally any $r - m_1$ independent linear combinations of them). In particular, on the assumptions made earlier the posterior distribution of the parameters $\boldsymbol{\theta}$ is proportional to $|\mathbf{V}_{r-m_1}|^{-n/2}$ and the estimates yielding maximum posterior density are obtained by minimizing this determinant.

The existence of m_2 further independent linear relationships in the *expected values* of the r responses should *not* lead to elimination of m_2 further responses. These relationships do not cause singularities in $|\mathbf{V}_{r-m_1}|$ and they contain valuable information allowing us to check the adequacy of the model and to obtain more precise estimates of the parameters $\boldsymbol{\theta}$. An eigenvalue-eigenvector analysis can draw attention to both kinds of relationships. The former kind must be allowed for in subsequent data fitting; the latter are of interest to the experimenter and in some cases may confirm or deny the adequacy of the model which is being fitted, but they will not directly affect the fitting of the data.

Each problem must be considered on its merits. The object of this paper is to point out that a preliminary eigenvalue-eigenvector analysis is a useful tool and, preferably, further analysis should not be proceeded with until the relationships it points to have been satisfactorily explained.

6. Chemical Example

The thermal isomerization of α-pinene to dipentene and allo-ocimene which in turn yields α- and β-pyronene and a dimer was studied by Fuguitt and Hawkins (1947). The proposed reaction scheme for this homogeneous chemical reaction is:



The concentrations of the reactant and the four products were reported by Fuguitt and Hawkins at eight time intervals, and these data are reporduced in Table 2. Mathematical models can be derived which give the concentration of the various species as a function of time if the chemical reaction orders are known. Hunter and MacGregor (1967), assuming first-order kinetics throughout, derived the following equations:

$$\eta_{1u} = y_{10}e^{-\phi t_u}, \qquad \eta_{2u} = \frac{\theta_1 y_{10}}{\phi}(1 - e^{-\phi t_u}), \qquad \eta_{3u} = C_1 e^{-\phi t_u} + C_2 e^{\beta t_u} + C_3 e^{\gamma t_u},$$

$$\eta_{4u} = \theta_3\left(\frac{C_1}{\phi}(1 - e^{-\phi t_u}) + \frac{C_2}{\beta}(e^{\beta t_u} - 1) + \frac{C_3}{\gamma}(e^{\gamma t_u} - 1)\right),$$

and

$$\eta_{5u} = \theta_4\left(\frac{C_1}{(\theta_5 - \phi)}e^{-\phi t_u} + \frac{C_2}{(\theta_5 + \beta)}e^{\beta t_u} + \frac{C_3}{(\theta_5 + \gamma)}e^{\gamma t_u}\right),$$

where y_{10} = the value of y_1 at $t = 0$, $\alpha = \theta_3+\theta_4+\theta_5$, $\beta = (-\alpha + \sqrt{\alpha^2 - 4\theta_3\theta_5})/2$, $\gamma = (-\alpha - \sqrt{\alpha^2 - 4\theta_3\theta_5})/2$, $\phi = \theta_1 + \theta_2$, $C_1 = \theta_2 y_{10}(\theta_5 - \phi)/((\phi + \beta)(\phi + \gamma))$, $C_2 = \theta_2 y_{10}(\theta_5 + \beta)/((\phi + \beta)(\beta - \gamma))$, $C_3 = \theta_2 y_{10}(\theta_5 + \gamma)/((\phi + \gamma)(\gamma - \beta))$.

TABLE 2

Concentration vs. Time Data for the Isomerization of α-pinene at 189.5°C.

Time (min.)	y_1 α-pinene	y_2 dipentene	y_3 allo-ocimene	y_4 pyronene	y_5 dimer
1230	88.35	7.3	2.3	0.4	1.75
3060	76.4	15.6	4.5	0.7	2.8
4920	65.1	23.1	5.3	1.1	5.8
7800	50.4	32.9	6.0	1.5	9.3
10680	37.5	42.7	6.0	1.9	12.0
15030	25.9	49.1	5.9	2.2	17.0
22620	14.0	57.4	5.1	2.6	21.0
36420	4.5	63.1	3.8	2.9	25.7

Assuming these models to be appropriate, we can obtain the posterior distribution of the parameters (the five rate constants) following Box and Draper (1965). In particular, we can find those parameter values which have the highest posterior density by minimizing the determinant criterion. If data dependencies are ignored, however, and parameter values which minimize the determinant criterion $|\mathbf{V}_5|$ (using all five responses) are found, the result is the unsatisfactory data fit shown in Figure 1. This example demonstrates how analysis of multiresponse data that ignores dependencies can lead to meaningless answers if linear dependencies are present.

From an examination of Fuguitt and Hawkins' paper, it can be found that y_4 (α- plus β-pyronene), because of experimental difficulties, was not measured independently but rather was assumed to constitute three percent of the total conversion of α-pinene (y_1). That is, it was assumed that

$$y_4 = 0.03(100 - y_1).$$

Thus, there is the following exact linear relationship among the observations:

$$(0.03)y_1 + (0)y_2 + (0)y_3 + (1)y_4 + (0)y_5 = 3. \tag{28}$$

Furthermore, by reducing the first order differential equations defining the kinetic

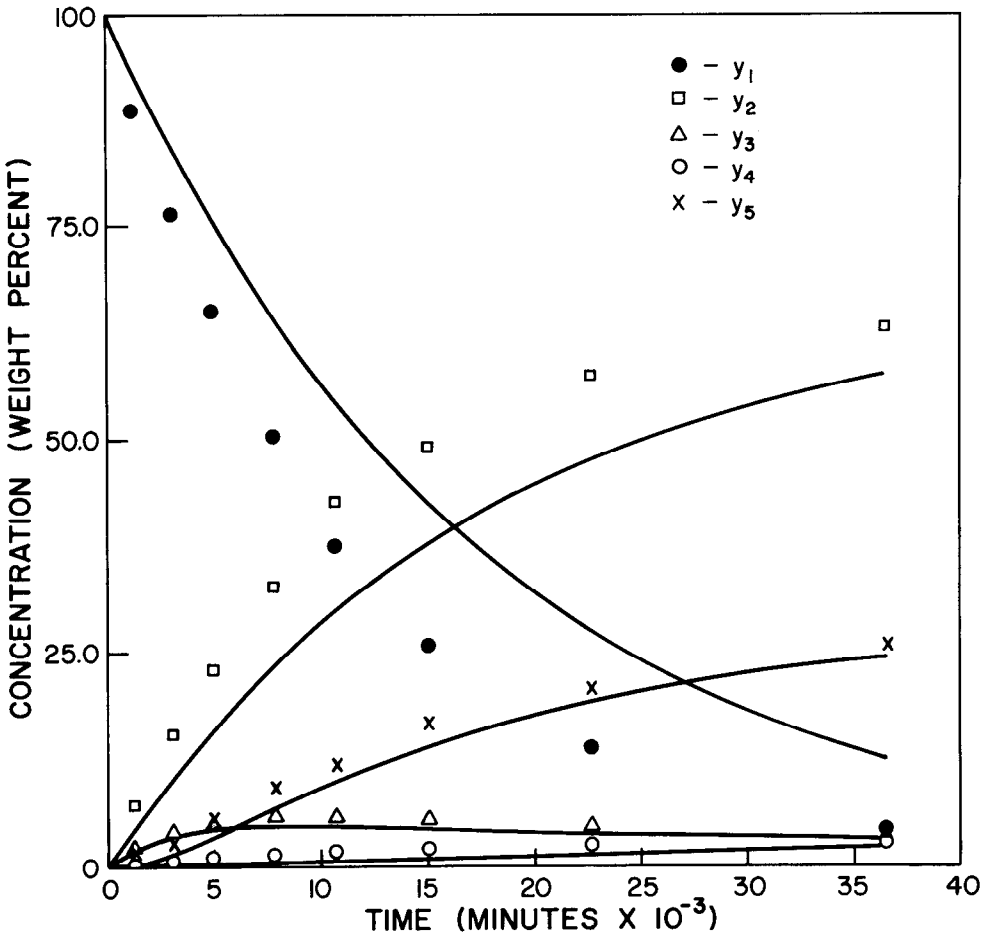


FIGURE 1—Data Fit with Parameter Values Which, Ignoring Dependencies, Minimize $|\mathbf{V}_5|$.

system to a canonical form following Ames (1962), it can be shown that the following two relationships must exist among the expectations of the y's:

$$E(y_1) + E(y_2) + E(y_3) + E(y_4) + E(y_5) = \text{const.} \tag{29}$$

and

$$E(y_1) + (1 + \theta_2/\theta_1)E(y_2) = \text{const.} \tag{30}$$

The first of these relationships (29) simply expresses an overall mass balance for the system while the second (30) results from the fact that the isomerizations of α-pinene (y_1) are assumed irreversible. The experimenters in this study chose to report their data in "normalized" weight percentage form, and so in effect, consciously or otherwise, used their knowledge of the expectation relationship (29) to force the following relationship among the observed responses:

$$y_1 + y_2 + y_3 + y_4 + y_5 = 100. \tag{31}$$

Thus there exist three linear relationships (28), (30), and (31) in this system, two of which, (28) and (31), are exact apart from rounding error. Because of the rounding error the value of the determinant was not exactly zero for all values of the parameters and this made it technically possible to obtain a minimum for $|\mathbf{V}_5|$ which however is meaningless.

6.1 *Eigenvalue-Eigenvector Analysis*

We now illustrate how empirical analysis could reveal these relationships. To implement the eigenvalue-eigenvector analysis, the individual means are first eliminated from the various responses to form a new data matrix, $\mathbf{D} = \{d_{iu}\}$, where $d_{iu} = y_{iu} - \bar{y}_i$. The eigenvalues and the associated eigenvectors of the $\mathbf{DD}'$ matrix are then obtained. These quantities are given in Table 3 for the present example.

TABLE 3

Eigenvalues and Eigenvectors of $\mathbf{DD}'$

Eigenvalues of $\underset{\sim}{DD}'$				
λ_1	λ_2	λ_3	λ_4	λ_5
.0013	.0168	1.21	25.0	9660.
Eigenvectors of $\underset{\sim}{DD}'$				
$\underset{\sim}{z}_1$	$\underset{\sim}{z}_2$	$\underset{\sim}{z}_3$	$\underset{\sim}{z}_4$	$\underset{\sim}{z}_5$
-.169	.476	-.296	.057	.809
-.211	.490	-.611	-.224	-.540
-.161	.435	.640	-.612	-.013
.931	.364	-.010	.004	-.024
-.185	.459	.360	.756	-.231

By employing the overall residual sum of squares (RSS) obtained from minimizing (2), one can obtain a crude estimate of the average experimental error variance by calculating

$$\bar{\sigma}^2 \simeq \frac{\text{RSS}}{\text{d.f.}} = \frac{19.87}{(40 - 5)} = 0.6$$

where d.f. stands for residual degrees of freedom. Hence by equation (16) we can expect those eigenvalues arising from linear relationships among the expected values of the responses to be of the order of $(n - 1)\bar{\sigma}^2 = 4.2$. The eigenvalue λ_3 is seen to be of this order of magnitude and will be shown later to correspond to equation (30). The eigenvalues λ_4 and λ_5 are however much larger and we would therefore not expect them to be associated with any linear relationships. It is clear that the capability of estimating the parameters in our models is coming mostly from $\mathbf{z}_5$, the eigenvector associated with λ_5.

To see whether there are any exact linear relationships among the responses, one should look for eigenvalues which are zero. In this case, there are none which are exactly zero but both λ_1 and λ_2 are very small and much smaller than the value 4.2 which one would expect from a linear expectation relationship. Thus one might

suspect that eigenvalues $\mathbf{z}_1$ and $\mathbf{z}_2$ correspond to linear relationships among the recorded responses which are exact except for rounding error.

To obtain an estimate of the expected value of an eigenvalue when there is only rounding error present, we can assume that the rounding error is distributed uniformly with range -0.5 to $+0.5$ of the last digit reported. Rounding error variance σ^2_{RE} is then given by the range squared divided by 12, and since in this example all responses have been rounded to the same number of significant figures, the corresponding expected value of an eigenvalue using (3) may be approximated by

$$E(\lambda_k) = (n - 1)\mathbf{z}_k'\mathbf{I}\mathbf{z}_k\sigma^2_{\mathrm{re}} = (n - 1)\sigma^2_{\mathrm{re}} .$$

The concentration data here were reported to the nearest 0.1 percent and therefore the range is from -0.05 to $+0.05$ or 0.10. Thus, the expected value of an eigenvalue is $7(0.10)^2/12$ or approximately 0.006. Both eigenvalues λ_1 and λ_2 are of this order of magnitude, thus helping to confirm that $\mathbf{z}_1$ and $\mathbf{z}_2$ represent exact dependencies among the responses.

The question now is to determine, if possible, what are the true linear dependencies among the responses that are causing this two dimensional singularity plane. As shown in equation (13) the plane defined by $\mathbf{Z}_1 = [\mathbf{z}_1 , \mathbf{z}_2]'$ will only be some non-singular transformation of the true constraint matrix $\mathbf{A}_1$. We previously stated that the linear relationships given by equations (28) and (31) were expected to be present in the data. To test whether these account for the two dimensional singularity region represented by $\mathbf{Z}_1$ we need to test for the coplanarity of the regions defined by $\mathbf{A}_1 = [\mathbf{a}_1 , \mathbf{a}_2]'$ and $\mathbf{Z}_1$ where $\mathbf{a}_1'$ = (.03, 0, 0, 1.0, 0) and $\mathbf{a}_2'$ = (1., 1., 1., 1., 1.). A very simple check on this is to calculate the cosine of the angle made by each of the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ with the plane $\mathbf{Z}_1$. Doing this yielded cosines of .9999 and .9993 respectively, implying that $\mathbf{a}_1$ and $\mathbf{a}_2$ do indeed lie almost entirely within the plane of $\mathbf{Z}_1$.

If the true underlying relationships are not known it may be possible to use the empirical eigenvectors to provide some indication of what they may be. For instance, if the smallest of the "zero" eigenvalues is considerably smaller than the others, its corresponding eigenvector may correspond fairly closely to the most exact of the linear relationships. However, this probably would not be so if the "zero" eigenvalues happened to be of nearly equal magnitude. Looking at Table 3 in our example the first eigenvector $\mathbf{z}_1'$ corresponds reasonably closely to $\mathbf{a}_1'$ = (0.03, 0, 0, 1.0, 0), and $\mathbf{z}_2'$ to that component of $\mathbf{a}_2'$ = (1., 1., 1., 1., 1.) which is orthogonal to $\mathbf{z}_1'$, namely (.465, .468, .464, .363, .466).

We are also in a position now to check whether the third eigenvalue $\mathbf{z}_3$ corresponds to the relationship among the expected values of the responses given by (30). If we use the current estimates of θ_1 and θ_2 and take that component of the vector corresponding to the relation (31) which is orthogonal to both $\mathbf{z}_1$ and $\mathbf{z}_2$ and normalize it, we get (−.308, −.665, .482, .008, .482) which is indeed very similar to $\mathbf{z}_3'$ = (−.296, −.611, .640, −.010, .360).

The confirmation of an expectation relationship such as (30) provides a valuable check of the tentatively entertained model structure for the system, in this case of the chemical stochiometry. Had either of the isomerization reactions been reversible, then (30) would not have been true.

6.2 *Analysis of the Data:*

Before any meaningful analysis of the data can be conducted, the two-dimensional singularity resulting from the relationships (28) and (31) must be removed. Perhaps

the most natural approach to this problem is to think in terms of the responses themselves and ask the question: which responses should be dropped? It must be done in such a way as to leave an independent subset of three responses. In many cases, the structure of the problem and of the dependencies may well dictate a natural way for dropping responses. In our problem, if we look at the two dependencies which we know to exist, it is obvious that at least one of y_1 or y_4 must be eliminated, preferably y_4 since it is known to be the fabricated response. The second relationship (31) contains all five responses and if there were no rounding error, it would make absolutely no difference which of the additional four responses was dropped. Another approach to this problem is to conduct the final analysis of the data on three linearly independent combinations of all five responses. Again if there were no rounding error it would make no difference which three linearly independent combinations were used so long as they adequately defined the 3-dimensional subspace orthogonal to the 2-dimensional singularity plane defined by the vectors $\mathbf{a}_1' = (0.03, 0, 0, 1.0, 0)$ and $\mathbf{a}_2' = (1., 1., 1., 1., 1.)$. From an estimation point of view, it would be convenient to use three orthogonal vectors to define this space. However, since the singularity relationships are not exactly satisfied by the data, and since the likelihood surface for the parameters is poorly conditioned, the final parameter estimates and their variance-covariance matrix will be sensitive to some extent to how the singularities are removed.

6.3 *Use of Empirical Eigenvectors*

Were it not for the presence of roundoff error, the eigenvalues λ_1 and λ_2 would have been zero and their corresponding eigenvectors $\mathbf{z}_1$ and $\mathbf{z}_2$ would define the exact singularity plane defined by $\mathbf{a}_1$ and $\mathbf{a}_2$. A natural set of vectors to use in defining the remaining three dimensions would then be the remaining three empirical eigenvectors $\mathbf{z}_3$, $\mathbf{z}_4$, and $\mathbf{z}_5$ since these satisfy the requirements of being independent, orthogonal vectors, all orthogonal to the singularity plane. Therefore, in practice, when it has not been possible to pinpoint the true singularity relationships ($\mathbf{a}_1$ and $\mathbf{a}_2$) or if the roundoff error is considered to be negligible, then these eigenvectors corresponding to the non-zero eigenvalues can be used to form the three independent linear combinations $f_{3u} = \mathbf{z}_3'\mathbf{y}_u$, $f_{4u} = \mathbf{z}_4'\mathbf{y}_u$, and $f_{5u} = \mathbf{z}_5'\mathbf{y}_u$. By minimizing the determinant $|\mathbf{V}_3|$ where

$$\mathbf{V}_3 = \{(\mathbf{f}_i - E(\mathbf{f}_i))'(\mathbf{f}_j - E(\mathbf{f}_j))\}, \qquad i, j = 3, 4, 5 \tag{32}$$

where $E(f_{iu}) = \mathbf{z}_i'\boldsymbol{\eta}_u$, one obtains the point estimates of the rate constants given in the second row of Table 4. The first row contains those rate constant estimates obtained previously by minimizing the overall residual sum of squares of all five responses. Figure 2 shows the fits of the responses $\mathbf{y}_i$ obtained using these parameter estimates.

6.4 *Use of Theoretical Eigenvectors*

In this example, however, the true, or theoretical linear dependencies ($\mathbf{a}_1'y_u$ and $\mathbf{a}_2'y_u$) have in fact been uncovered and it is therefore better to use three independent linear combinations of the five responses which are orthogonal to this true singularity plane rather than the approximate one represented by the eigenvectors $\mathbf{z}_1$ and $\mathbf{z}_2$ (although these will differ only very slightly since the roundoff error is relatively small). For this purpose, we used as the basis for our three linear combinations the vector components ($\mathbf{a}_3$, $\mathbf{a}_4$, and $\mathbf{a}_5$) of the eigenvectors $\mathbf{z}_3$, $\mathbf{z}_4$, and $\mathbf{z}_5$ which are orthogonal to $\mathbf{a}_1$ and $\mathbf{a}_2$. By minimizing the determinant of the form (32) where now $f_{iu} = \mathbf{a}_i'\mathbf{y}_u$ and $E(f_{iu}) = \mathbf{a}_i'\boldsymbol{\eta}_u$ we obtained the point estimates of the

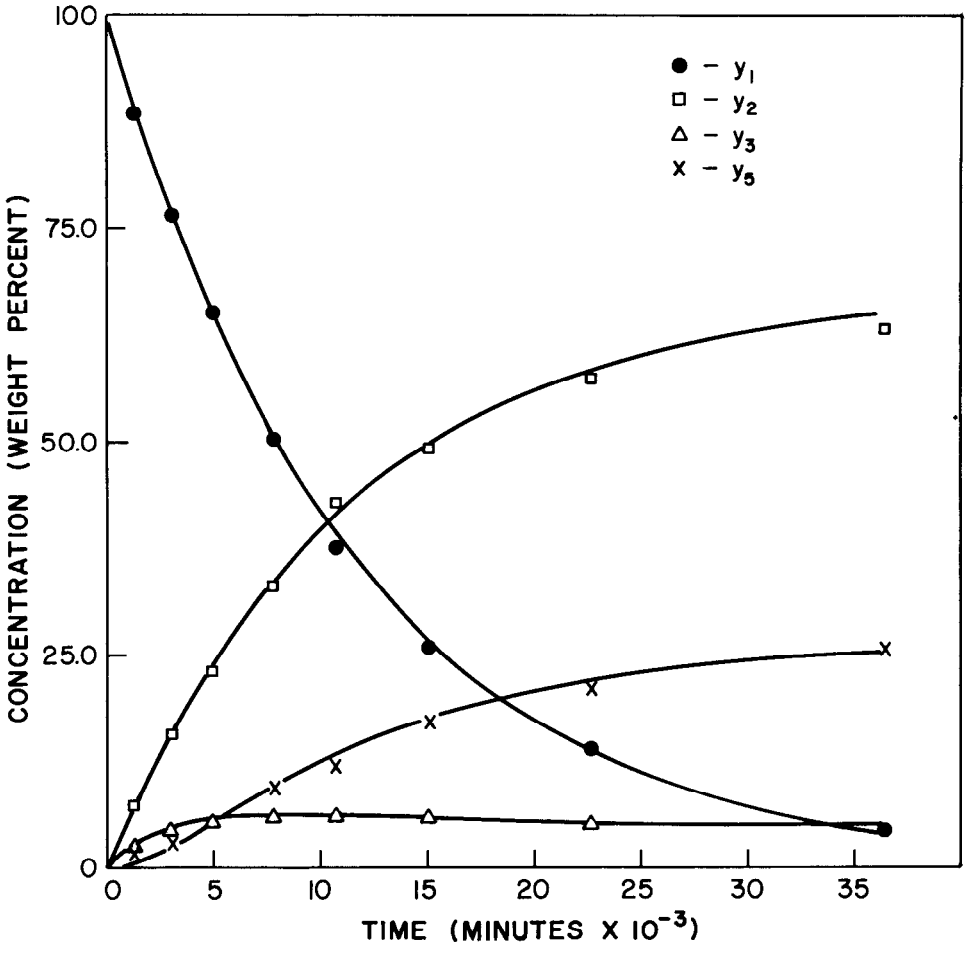


FIGURE 2—Fit of the Responses Using Parameter Values Which Minimize $|\mathbf{V}_3|$ Based on the Three Linear Combinations ($\mathbf{f}_i = \mathbf{z}'\mathbf{y}_u$, $i = 3, 4, 5$).

rate constants shown in the third row of Table 4. These are the values which give the highest posterior probability density. It can be seen that, as expected, they differ very little from those values in the second row. The resulting fit of the data is obviously very much better now that the singularities have been removed (see Figures 1 and 2).

The estimate of θ_3 is imprecise as is reflected by an extremely flat surface of the determinant function in the θ_3 direction. This was to be expected since most of the information on the θ_3 rate constant is contained in the singularity plane of the response space and, in particular, in y_4 .

The confidence region for the parameters can be computed using the general formula (Box and Draper, 1965)

$$|\mathbf{V}|_{(1-\alpha)} \simeq |\mathbf{V}|_{\min} \exp \{\chi_p^2(1 - \alpha)/n\} \tag{33}$$

where p = number of parameters
n = number of observations
$\chi_p^2(1 - \alpha)$ = chi-square value for p degrees of freedom and $(1 - \alpha) \times 100\%$ probability level.

For our example $p = 5$ and $n = 8$.

TABLE 4

Estimates of Rate Constants.

Minimization of	Rate Constants, 10^{-5} Min.$^{-1}$				
	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$
RSS	5.93	2.96	2.05	27.5	4.00
$\lvert \underset{\sim}{V}_3 \rvert$ (empirical eigenvectors)	5.95	2.84	0.43	31.3	5.74
$\lvert \underset{\sim}{V}_3 \rvert$ (theoretical eigenvectors)	5.95	2.85	0.50	31.5	5.89

7. SUMMARY

In the process of model-building using multiresponse data, one should always be alert for possible linear relationships among the responses. When there is doubt as to what relationships, if any, are present, an empirical eigenvalue-eigenvector analysis should be used. When dependencies are found, there must be good reason for them and considerable effort should be made to uncover their causes. Such studies can confirm or deny the validity of the model under study or suggest the appropriateness of a new and previously unsuspected mechanistic form. If m_1 exact linear relationships are known to exist, then m_1 dependent responses or m_1 linear combinations of them must be deleted before the data are analyzed; otherwise parameter estimation procedures yield meaningless results. So far as the actual analysis of the data at hand is concerned, however, knowing these true or theoretical dependencies is not absolutely necessary. One can proceed by making use of the empirically determined eigenvectors which do not necessarily represent true dependencies. A simplified version of the proposed procedure is shown in flow diagram form in Figure 3. Sections of this paper in which there is a fuller explanation are indicated.

It is desirable to have a check of some kind for data dependencies (such as the eigenvalue-eigenvector analysis described in this paper) built into general purpose computer programs for multiresponse fitting problems in the same way it is desirable to have a check on possible singularity or near-singularity incorporated in standard regression programs so that the user is given warning that answers produced by the program may not be meaningful. The eigenvalue-eigenvector analysis can be used to distinguish between dependencies in the data $\mathbf{y}$ on the one hand and in the expectations $E(\mathbf{y})$ on the other. In the first case, the eigenvalues will be of the order of magnitude of the mean square rounding error while in the second they are of the order of magnutide of the mean square residual error. Usually the latter is several orders of magnitude larger than the former so that it is possible to distinguish between the two kinds of dependencies so they can be handled accordingly.

We are indebted to a referee for pointing out that one way in which dependencies

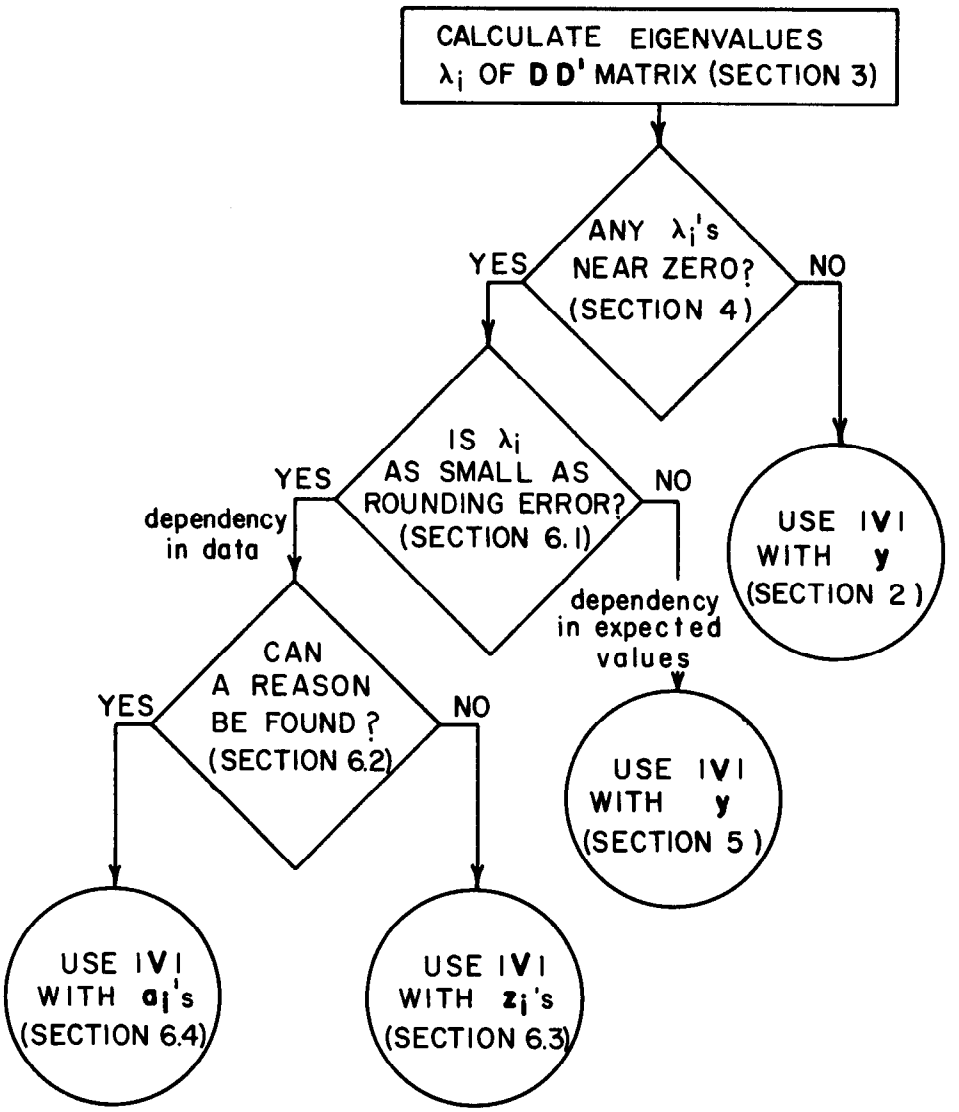


FIGURE 3—Simplified Version of Procedure for Fitting a Multiresponse Models.

among the expected values may enter a problem is through a steady-state relationship. This situation will usually be a local one in that it will depend upon the range of the experimental conditions. For instance, if in our example the data had been collected only over the restricted time range of 5000 to 15000 seconds where the rate of formation of alloocimene (y_3) is relatively constant, the following kinetic relationship would nearly hold:

$$\frac{d\eta_3}{dt} = \theta_2\eta_1 - \theta_3\eta_3 - \theta_4\eta_3 + \theta_5\eta_5 = 0$$

yielding a linear relationship among the expected values of several responses which would need to be appropriately explained.

8. Appendix

Expected Value of an Eigenvalue

The expected value of an eigenvalue, λ, whose associated eigenvector, $\mathbf{z}$, corresponds to a linear relationship in the expected values of the responses, can be found as follows:

Since $\mathbf{z}$ is an eigenvector of $\mathbf{DD}'$

$$\mathbf{DD}'\mathbf{z} = \lambda\mathbf{z} \tag{A.1}$$

Premultiplying equation (A.1) by $\mathbf{z}'$ we get

$$\mathbf{z}'\mathbf{DD}'\mathbf{z} = \mathbf{z}'\lambda\mathbf{z} = \lambda\mathbf{z}'\mathbf{z} \tag{A.2}$$

But since the $\mathbf{z}$ vector is scaled such that $\mathbf{z}'\mathbf{z} = 1$,

$$\lambda = \mathbf{z}'\mathbf{DD}'\mathbf{z} \tag{A.3}$$

By definition, $\mathbf{D} = \{d_{iu}\} = \{y_{iu} - \bar{y}_i\}$. We can further write $y_{iu} = \eta_{iu} + \epsilon_{iu}$ and $\bar{y}_i = \bar{\eta}_i + \bar{\epsilon}_i$ so that

$$\mathbf{D} = \{\eta_{iu} + \epsilon_{iu} - \bar{\eta}_i - \bar{\epsilon}_i\} \tag{A.4}$$

The ijth element of $\mathbf{DD}'$ is then

$$(\mathbf{DD}')_{ij} = \sum_{u=1}^{n} (\eta_{iu} + \epsilon_{iu} - \bar{\eta}_i - \bar{\epsilon}_i)(\eta_{ju} + \epsilon_{ju} - \bar{\eta}_j - \bar{\epsilon}_j) \tag{A.5}$$

which can be expanded to

$$\begin{aligned}(\mathbf{DD}')_{ij} = {} & \sum_{u=1}^{n} (\eta_{iu}\eta_{ju} - \eta_{iu}\bar{\eta}_j - \bar{\eta}_i\eta_{ju} + \bar{\eta}_i\bar{\eta}_j) \\ & + \sum_{u=1}^{n} (\eta_{iu}\epsilon_{ju} - \eta_{iu}\bar{\epsilon}_j - \bar{\eta}_i\epsilon_{ju} + \bar{\eta}_i\bar{\epsilon}_j + \epsilon_{iu}\eta_{ju} - \epsilon_{ju}\bar{\eta}_j - \bar{\epsilon}_i\eta_{ju} + \bar{\epsilon}_i\bar{\eta}_j) \\ & + \sum_{u=1}^{n} (\epsilon_{iu}\epsilon_{ju} - \epsilon_{iu}\bar{\epsilon}_j - \bar{\epsilon}_i\epsilon_{ju} + \bar{\epsilon}_i\bar{\epsilon}_j)\end{aligned} \tag{A.6}$$

The last three terms in the first summation of equation (A.6) can be seen to be equivalent,

$$\sum_{u=1}^{n} \eta_{iu}\bar{\eta}_j = \sum_{u=1}^{n} \bar{\eta}_i\eta_{ju} = \sum_{u=1}^{n} \bar{\eta}_i\bar{\eta}_j = n\bar{\eta}_i\bar{\eta}_j \tag{A.7}$$

Assuming that the models, η_{iu} , are correct so that $E(\epsilon_{iu}) = 0$, it follows that $E(\bar{\epsilon}_i) = 0$ also. Then, when the expected value of $(\mathbf{DD}')_{ij}$ is taken, all the terms in the second summation of (A.6) become zero.

We further assume that the errors in the responses have a variance-covariance matrix $\mathbf{\Sigma} = \{\sigma_{ij}^2\}$ so that $E(\epsilon_{iu}\epsilon_{ju}) = \sigma_{ij}^2$, and that these errors are independent from run to run. That is, $E(\epsilon_{iu}\epsilon_{iv}) = 0$ for all $u \neq v$. Now the expectations of the terms in the third summation of (A.6) can be evaluated.

$$E\left(\sum_{u=1}^{n} \epsilon_{iu}\epsilon_{ju}\right) = \sum_{u=1}^{n} E(\epsilon_{iu}\epsilon_{ju}) = \sum_{u=1}^{n} \sigma_{ij}^2 = n\sigma_{ij}^2 \tag{A.8}$$

$$E\left(\sum_{u=1}^{n} \epsilon_{iu}\bar{\epsilon}_j\right) = nE\left(\epsilon_{iu}\sum_{v=1}^{n} \epsilon_{jv}/n\right) = nE(\epsilon_{iu}\epsilon_{ju}/n) = \sigma_{ij}^2 \tag{A.9}$$

$$E\left(\sum_{u=1}^{n} \bar{\epsilon}_i \epsilon_{ju}\right) = nE\left[\sum_{v=1}^{n} (\epsilon_{iv}/n)\epsilon_{ju}\right] = nE[(\epsilon_{iu}/n)\epsilon_{ju}] = \sigma_{ij}^2 \tag{A.10}$$

$$E\left(\sum_{u=1}^{n} \bar{\epsilon}_i \bar{\epsilon}_j\right) = nE\left(\sum_{u=1}^{n} \frac{\epsilon_{iu}}{n}\right)\left(\sum_{v=1}^{n} \frac{\epsilon_{jv}}{n}\right) = \frac{1}{n} E\left(\sum_{u=1}^{n} \epsilon_{iu}\epsilon_{ju}\right) = \sigma_{ij}^2 \tag{A.11}$$

When the above information is incorporated in equation (A.6) it becomes

$$E(\mathbf{DD}')_{ij} = \sum_{u=1}^{n} \eta_{iu}\eta_{ju} - n\bar{\eta}_i\bar{\eta}_j - n\bar{\eta}_i\bar{\eta}_j + n\bar{\eta}_i\bar{\eta}_j + n\sigma_{ij}^2 - \sigma_{ij}^2 - \sigma_{ij}^2 + \sigma_{ij}^2 \tag{A.12}$$

which, upon simplification, reduces to

$$E(\mathbf{DD}')_{ij} = \sum_{u=1}^{n} \eta_{iu}\eta_{ju} - n\bar{\eta}_i\bar{\eta}_j + (n-1)\sigma_{ij}^2 \tag{A.13}$$

The expected value of λ can now be found by taking expectations of both sides of equation (A.3).

$$E(\lambda) = E(\mathbf{z}'\mathbf{DD}'\mathbf{z}) = \mathbf{z}'E(\mathbf{DD}')\mathbf{z} = \sum_{i=1}^{r}\sum_{j=1}^{r} z_i z_j E(\mathbf{DD}')_{ij} \tag{A.14}$$

Substituting (A.13) into (A.14) we obtain

$$E(\lambda) = \sum_{i=1}^{r}\sum_{j=1}^{r} z_i z_j \left[\sum_{u=1}^{n} \eta_{iu}\eta_{ju} - n\bar{\eta}_i\bar{\eta}_j + (n-1)\sigma_{ij}^2\right] \tag{A.15}$$

Since z represents a linear relationship in the expected values of the responses, and the expected value of y_{iu} is η_{iu}, we know that

$$\sum_{i=1}^{r} z_i \eta_{iu} = a_0 \tag{A.16}$$

which is the same constant, a_0, for every run, u. It follows that

$$\sum_{i=1}^{r} z_i \bar{\eta}_i = \sum_{i=1}^{r} z_i \sum_{u=1}^{n} \frac{\eta_{iu}}{n} = \sum_{u=1}^{n}\sum_{i=1}^{r} \frac{z_i \eta_{iu}}{n} = \sum_{u=1}^{n} \frac{a_0}{n} = a_0 \tag{A.17}$$

By expanding equation (A.15) we obtain

$$E(\lambda) = \sum_{u=1}^{n}\sum_{i=1}^{r} z_i \eta_{iu} \sum_{j=1}^{r} z_j \eta_{ju} - n \sum_{i=1}^{r} z_i \bar{\eta}_i \sum_{j=1}^{r} z_j \bar{\eta}_j + (n-1) \sum_{i=1}^{r}\sum_{j=1}^{r} z_i z_j \sigma_{ij}^2 \tag{A.18}$$

Then, when we insert (A.16) and (A.17) into (A.18) we get

$$E(\lambda) = \sum_{u=1}^{n} (a_0)(a_0) - n(a_0)(a_0) + (n-1)\mathbf{z}'\boldsymbol{\Sigma}\mathbf{z} \tag{A.19}$$

which simplifies to

$$E(\lambda) = (n-1)\mathbf{z}'\boldsymbol{\Sigma}\mathbf{z} \tag{A.20}$$

If $\boldsymbol{\Sigma}$ is known, the expected value of λ can be found directly from equation (A.20). And even when $\boldsymbol{\Sigma}$ is unknown, it may be possible to estimate it, for example, there may be some replication in the data. The best estimate for the expected value of λ in this case, too, is found from equation (A.20).

If $\boldsymbol{\Sigma}$ is unknown and cannot be estimated, we can get a very crude estimate of $E(\lambda)$ by approximating $\boldsymbol{\Sigma}$ by $\mathbf{I}\bar{\sigma}^2$. Then $E(\lambda) \simeq (n-1)\bar{\sigma}^2$, where the average variance, $\bar{\sigma}^2$, may be estimated from the residual sum of squares.

9. Acknowledgment

The authors would like to thank Paul W. Tidwell of the Monsanto Company for valuable and enjoyable discussion on ideas presented in this paper.

References

[1] Ames, W. F. (1962). "Canonical Forms for Nonlinear Kinetic Differential Equations", *Ind. Eng. Chem. Fundamentals, 1*, 214.

[2] Ball, W. E. and Groenweghe, L. C. D. (1966). "Determination of Best-Fit Rate Constants in Chemical Kinetics", *Ind. Eng. Chem. Fundamentals, 5*, 181.

[3] Box, G. E. P. and Draper, N. R. (1965). "Bayesian Estimation of Common Parameters from Several Responses", *Biometrika, 52*, 355.

[4] Eakman, J. M. (1969). "Strategy for Estimation of Rate Constants from Isothermal Reaction Data", *Ind. & Eng. Chem. Fundamentals, 8*, 53.

[5] Erjavec, J. (1970). "Strategy for Estimation of Rate Constants from Isothermal Reaction Data", *Ind. & Eng. Chem. Fundamentals, 9*, 187.

[6] Fuguitt, R. E. and Hawkins, J. E. (1947). "Rate of Thermal Isomerization of α-Pinene in the Liquid Phase", *J.A.C.S., 69*, 319.

[7] Hunter, W. G. (1967). "Estimation of Unknown Constants from Multiresponse Data", *Ind. & Eng. Chem. Fundamentals, 6*, 461.

[8] Hunter, W. G. and McGregor, J. F. (1967). "The Estimation of Common Parameters from Several Responses: Some Actual Examples", Unpublished Report, The Department of Statistics, The University of Wisconsin.