AN ABSTRACT OF THE THESIS OF

WALLACE HENRY REED for the M. A. in Mathematics
(Name) (Degree) (Major)

Date thesis is presented May 4, 1965

Title LEAST SQUARES SOLUTIONS FOR NON-ORTHOGONAL TWO-WAY CLASSIFICATIONS

Abstract approved Redacted for Privacy
(Major professor)

One answer to the problem of missing observations in two-way classification experiments is to insert estimates of missing observations into deficient cells. Once missing observations have been estimated, the experimenter may proceed with his analysis using the familiar normal equations which apply to complete data. This paper discusses generally the problem of obtaining the desired estimates and provides explicit solutions in certain special cases, among them the case where there appears no more than a single deficient cell in any row and column, the case where deficient cells occur in a block at the intersections of certain rows and columns, and a comprehensive generalization to an arbitrary number of blocks having no rows or columns in common. Also presented is an iterative process providing approximate solutions to the normal equations of a two-way classification, which is especially useful in dealing with a large layout having a majority of cells empty.

LEAST SQUARES SOLUTIONS FOR NON-ORTHOGONAL TWO-WAY CLASSIFICATIONS

by

WALLACE HENRY REED

A THESIS

submitted to

OREGON STATE UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

MASTER OF ARTS

June 1965

APPROVED:

Redacted for Privacy

Professor of Mathematics

In Charge of Major

Redacted for Privacy

Chairman of the Department of Mathematics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented May 4, 1965

Typed by Carol Baker

ACKNOWLEDGMENT

Great thanks are owing my mentor, Dr. E. L. Kaplan, for his patience, understanding, and assistance throughout the preparation of this paper.

TABLE OF CONTENTS

LIST OF FIGURES

LEAST SQUARES SOLUTIONS FOR NON-ORTHOGONAL TWO-WAY CLASSIFICATIONS

I. INTRODUCTION

In the investigation of any natural phenomenon, one early attempts to discover what factors control the phenomenon and to what extent. Scientists in agriculture were supplied with the tools to answer these questions by R. A. Fisher in a book first published in 1925. Fisher's book was revolutionary, but his methods spread rapidly from agriculture to the biological sciences, later to the physical sciences, and today pervade almost every area where men conduct experiments.

In factorial design experiments certain parameters provide a measure of the influence each factor exerts over the phenomenon under investigation. Fisher has said that the estimation of these parameters is perhaps more fundamental than any other aspects of experimental analysis. Indeed, in many experiments the analysis stops with the estimation of the important parameters.

It is with point estimation that this paper is concerned.

1. Two-way Classifications

Traditionally this subject is introduced by an example, usually from agriculture, usually concerning the yield of varieties of corn under different fertilizer applications. We shall discuss a "subject crop" and the influence on yield of certain "factors." The description which follows is to be interpreted in the greatest generality.

Imagine that the yield of a certain crop is influenced by two factors. To investigate the effect of these two factors on yield we plant a number of plots of our subject crop, and imagine that a plot may be subjected to the first factor at any of I different levels and to the second factor at any of J different levels. There are IJ combinations of levels for the two factors, and we suppose that each possible treatment combination is assigned randomly to a plot.

When the crop is harvested the yield for each plot is measured. The yield from the plot subjected to the first factor at the i^{th} level and the second factor at the j^{th} level is denoted by x_{ij} and called the ij^{th} observation. The observations are arranged in a rectangular array of I rows and J columns with x_{ij} located in the i^{th} row and j^{th} columns. The position occupied by x_{ij} in this array is called the ij^{th} cell of the layout of observations.

By making certain assumptions regarding the population from which each observation has been drawn, it is possible to associate a paremeter with the effect of each level of a given factor. The set of assumptions which permit this is called the model for the experiment.

2. The Model

The assumptions which follow constitute a possible model for an experiment with two-way classification.

We assume that there are IJ different populations, each with the same population variance. Each observation represents a random sample of size one from one of these populations, one sample from each population. We assume the random variable x_{ij} may be written

$$x_{ij} = \alpha_i + \beta_j + \xi + e_{ij} \quad , \tag{1.1}$$

where α_i is a parameter associated with the i^{th} row of the layout (and hence the i^{th} level of the first factor) called the row effect for the i^{th} row, β_j is a parameter associated with the j^{th} column of the layout called the column effect for the j^{th} column, ξ is a parameter to be defined shortly, and e_{ij} is a random variable, called the error in x_{ij}, with zero mean and the same variance as x_{ij}.

If we denote the mean of x_{ij} by ξ_{ij} we have

$$\begin{aligned} E(x_{ij}) &= E(\alpha_i + \beta_j + \xi + e_{ij}) \\ &= \alpha_i + \beta_j + \xi + E(e_{ij}) \\ &= \alpha_i + \beta_j + \xi = \xi_{ij} \quad , \end{aligned} \tag{1.2}$$

and if we sum up the means of all the observations in the layout we have

$$\sum_{i=1}^{I} \sum_{j=1}^{J} (\alpha_i + \beta_j + \xi) = \sum_{i=1}^{I} \sum_{j=1}^{J} \xi_{ij}$$

or

$$J\sum_{i=1}^{I} \alpha_i + I\sum_{j=1}^{J} \beta_j + I J \xi = \sum_{i=1}^{I} \sum_{j=1}^{J} \xi_{ij} \ . \qquad (1.3)$$

The final assumption is that the sum of all row effects and the sum of all column effects are 0. Symbolically the assumption is that

$$\sum_{i=1}^{I} \alpha_i = \sum_{j=1}^{J} \beta_j = 0 \ . \qquad (1.4)$$

This is a mere convenience and involves no loss of generality. If these conditions are imposed on Equation (1.3) above, note that we have

$$\xi = \frac{1}{IJ}\sum_{i=1}^{I} \sum_{j=1}^{J} \xi_{ij} \ , \qquad (1.5)$$

so that the parameter ξ turns out to be the arithmetic average of all observation means.

The model which is described here is said to be "additive." This means that it has been assumed that rows and columns do not "interact," i.e., for example, if one of the factors tends to produce a certain response in the subject, the pattern of response will not be altered by the particular level of the remaining factor. This assumption is reflected in (1.2), which expresses the IJ parameters ξ_{ij} in terms of the $I+J+1$ parameters α_i, β_j, and ξ, of which only $I+J-1$ are algebraically independent, hence the need for (1.4) to make the parameters and their estimates uniquely determined.

The fundamental assumption (1.1) above is, by Equation (1.2), equivalent to

$$E(x_{ij}) = \alpha_i + \beta_j + \xi \quad , \tag{1.6}$$

i.e., the mean of an observation is made up of a row effect, a column effect, and a general mean.

If an experimenter were not satisfied with planting only one plot for each of the possible combinations of levels for the two factors, he might plant two plots for each combination and thereby certainly increase the precision of his experiment. In general, he might plant K_{ij} replicates of the plot subjected to the first factor at the i^{th} level and the second factor at the j^{th} level and hence fill the ij^{th} cell of the layout with K_{ij} observations, each a random sample of size one, all from the same population. The number K_{ij} will be called the cell frequency of the ij^{th} cell. The k^{th} observation $k = 1, \cdots, K_{ij}$ in the ij^{th} cell of the layout will be denoted by x_{ijk}. The experimental model has now been fully described.

With the description of the model complete we may rewrite our fundamental assumption (1.1) with some increased generality by assuming plural cell frequencies. A slight manipulation yields

$$e_{ijk} = x_{ijk} - \alpha_i - \beta_j - \xi \quad . \tag{1.7}$$

If we square the error e_{ijk} and sum over all ijk we obtain an expression Q called the sum of squared errors.

$$Q = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K_{ij}} (e_{ijk})^2$$

$$= \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K_{ij}} (x_{ijk} - \alpha_i - \beta_j - \xi)^2 \quad . \tag{1.8}$$

By the method of least squares estimation, one takes as the estimates of α_i, β_j, and ξ those values which produce a minimum in Q. Some attempt will be made to streamline the notation of least squares analysis.

3. Notation

The estimate of a parameter will be denoted by fixing a caret over the symbol for the parameter, e.g. (least squares estimate of ξ) $= \hat{\xi}$.

Where the limits of summation are clear we shall write

$$\sum_{ijk} \quad \text{for} \quad \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K_{ij}} \quad .$$

Throughout this work the range of i, j, and k shall be

$$i = 1, \cdots, I$$

$$j = 1, \cdots, J$$

$$k = 1, \cdots, K_{ij}$$

unless otherwise stated.

We make the following definitions of symbols:

$$K = \sum_{ij} K_{ij}$$

$$x_{\cdot\cdot\cdot} = \sum_{ijk} x_{ijk}$$

$$x_{ij.} = \sum_{k} x_{ijk}$$

$$x_{\cdot j \cdot} = \sum_{ik} x_{ijk}$$

$$x_{i\cdot\cdot} = \sum_{jk} x_{ijk}$$

$$\bar{x} = \frac{1}{K} \sum_{ijk} x_{ijk}$$

$$\bar{x}_{ij.} = \frac{1}{K_{ij}} \sum_{k} x_{ijk} \quad .$$

At times we shall be dealing with the case where $K_{ij} = 1$ for every cell and the k will then be dropped from the symbol x_{ijk}. In this case the notation above will be used with the obvious modifications.

4. Non-orthogonality

In section two it was pointed out that of the $I + J + 1$ parameters used to describe an $I \times J$ two-way classification model only $I + J - 1$ of the parameters are independent. That is to say, the model contains only $I + J - 1$ essential parameters. One is convinced of this by considering the introduction of two new arbitrary parameters into the model. For example, for arbitrary numbers α and β, let us write

$$\alpha_i = \alpha_i' + \alpha \ , \qquad i = 1, \cdots, I$$

$$\beta_j = \beta_j' + \beta \ , \qquad j = 1, \cdots, J$$

so that assumption (1.1) becomes

$$x_{ij} = \alpha_i' + \alpha + \beta_j' + \beta + \xi + e_{ij} \ .$$

One has now only to define

$$\xi = \xi' - \alpha - \beta$$

and we have

$$x_{ij} = \alpha_i' + \beta_j' + \xi' + e_{ij} \ ,$$

which is exactly the same model as before, described now, however, by a different set of parameter values. This shows that two of the $I + J + 1$ parameters α_i, β_j, and ξ are dependent, and points up again the need for a pair of additional constraints, such as are provided by (1. 4), if one hopes to determine these parameters uniquely.

The final remark preliminary to the defining of orthogonality is that one minimizes the sum of squared errors Q, Equation (1. 8), by solving the system

$$\frac{\partial Q}{\partial \alpha_i} = \frac{\partial Q}{\partial \beta_j} = \frac{\partial Q}{\partial \xi} = 0 \ .$$

These equations are known as the normal equations.

Let us suppose that one defines a set of essential parameters, $I - 1$ of which describe row effects and $J - 1$ of which describe column effects, and that the resulting normal equations fall into three sets of size $I-1$, $J-1$, and 1. The first set contains only the parameters describing row effects, the second only the parameters describing column effects, and the third only the general mean ξ. The three sets can be solved independently of one another and it can be shown that estimates in one set are uncorrelated with estimates in another set, if the errors of observation are independent of one another as we have assumed. Under these conditions a two-way classification model is said to be orthogonal. A model which lacks this convenient property is by definition non-orthogonal.

We will illustrate this definition by showing that a 3×3 layout of complete data with one observation per cell is orthogonal. Instead of the usual parameters α_1, α_2, α_3, β_1, β_2, β_3, and ξ we take our parameters to be α_1', α_2', β_1', β_2', and ξ' where

$$\alpha_i = \alpha_i', \quad i = 1, 2$$

$$\alpha_3 = -\alpha_1' - \alpha_2'$$

$$\beta_j = \beta_j', \quad j = 1, 2$$

$$\beta_3 = -\beta_1' - \beta_2'$$

$$\xi = \xi' \quad .$$

The normal equations become

$$
\begin{aligned}
6\alpha_1' + 3\alpha_2' &= x_{11} + x_{12} + x_{13} - x_{31} - x_{32} - x_{33} \\
3\alpha_1' + 6\alpha_2' &= x_{21} + x_{22} + x_{23} - x_{31} - x_{32} - x_{33} \\
6\beta_1' + 3\beta_2' &= x_{11} - x_{13} + x_{21} - x_{23} + x_{31} - x_{33} \\
3\beta_1' + 6\beta_2' &= x_{12} - x_{13} + x_{22} - x_{23} + x_{32} - x_{33} \\
\xi' &= \overline{x} \ ,
\end{aligned}
\tag{1.9}
$$

which clearly have the desired properties.

The solution of these equations is

$$
\begin{aligned}
\hat{\alpha}_1' &= \frac{1}{9}(2x_{11}+2x_{12}+2x_{13}-x_{21}-x_{22}-x_{23}-x_{31}-x_{32}-x_{33}) \\
\hat{\alpha}_2' &= \frac{1}{9}(-x_{11}-x_{12}-x_{13}+2x_{21}+2x_{22}+2x_{23}-x_{31}-x_{32}-x_{33}) \\
\hat{\beta}_1' &= \frac{1}{9}(2x_{11}-x_{12}-x_{13}+2x_{21}-x_{22}-x_{23}+2x_{31}-x_{32}-x_{33}) \\
\hat{\beta}_2' &= \frac{1}{9}(-x_{11}+2x_{12}-x_{13}-x_{21}+2x_{22}-x_{23}-x_{31}+2x_{32}-x_{33}) \\
\hat{\xi}' &= \frac{1}{9}(x_{11}+x_{12}+x_{13}+x_{21}+x_{22}+x_{23}+x_{13}+x_{23}+x_{33}) \ .
\end{aligned}
\tag{1.10}
$$

The three sets of parameters are $\{\alpha_1', \alpha_2'\}$, $\{\beta_1', \beta_2'\}$, and $\{\xi'\}$. We will show first that $\hat{\alpha}_1'$ and $\hat{\beta}_1'$ are uncorrelated. Recall that the observations all have the same variance, let us denote it by σ^2. We may, for convenience, and without loss of generality assume, for the moment, that the observations all have zero mean. Thus we must show that

$$\mathrm{cov}(\hat{\alpha}'_1, \hat{\beta}'_1) = E(\hat{\alpha}'_1 \; \hat{\beta}'_1) = 0 \;.$$

Well,

$$E(x_{ij} x_{pq}) = 0 \quad \text{if} \quad (i,j) \neq (p,q)$$
$$= \sigma^2 \quad \text{if} \quad (i,j) = (p,q) \;,$$

so

$$E(\hat{\alpha}'_1, \hat{\beta}') = E\left[\frac{1}{81}(2x_{11}+2x_{12}+2x_{13}-x_{21}-x_{22}-x_{23}-x_{31}-x_{32}-x_{33}) \right.$$
$$\left. (2x_{11}-x_{12}-x_{13}+2x_{21}-x_{22}-x_{23}+2x_{31}-x_{32}-x_{33}) \right]$$
$$= \frac{1}{81} E\,(4x_{11}^2 - 2x_{12}^2 - 2x_{13}^2 - 2x_{21}^2 + x_{22}^2 + x_{23}^2 - 2x_{31}^2 + x_{32}^2 + x_{33}^2)$$
$$= \frac{\sigma^2}{81}\,(4-2-2-2+1+1-2+1+1) = 0 \;.$$

There are seven other similar calculations which can be made by inspection from (1.10). It is from this property that the aptness of the word "orthogonal" becomes most clear.

II. MISSING DATA

In this chapter we shall restrict our model to the case where there is a single observation in each cell, except one, which shall be empty. There is no great difficulty in imagining conditions under which such a model might arise. Most often, of course, a missing observation is the result of one of the infinite variety of calamities which assail experimenters. It is an axiom of research that if, in a given experiment, anything can go wrong, it will. It is conceivable, on the other hand, that an empty cell might be an unavoidable feature of the design, though this would be rather at variance with the assumption of additivity.

It is generally true, I think, that the feeling persists among research workers that the abortion of even a small phase of an experiment produces damage out of proportion to the number of observations lost. There is no justification for this idea. When information is lost, according to Fisher, "...there is no reason to suppose that the loss of information suffered will be disproportionate to the value of the experiment as a whole" (2, p. 176).

The method proposed by Cochran and Cox for dealing with empty cells they have chosen to call the "correct least squares method" (1). The method is perhaps the most natural of all in that it consists simply in carrying out the usual least squares analysis on the observations that are not missing.

We shall apply the correct least squares method, hereafter abbreviated CLSM, to an experiment whose design calls for one observation per cell, but in fact has one cell empty.

1. Correct Least Squares Method With One Observation Per Cell

Take the IJ^{th} cell to be the empty one. The sum of squared errors is

$$Q = \sum_{ij \neq IJ} (x_{ij} - \alpha_i - \beta_j - \xi)^2 .$$

The normal equations are

$$\hat{\xi} : (x_{..} - x_{IJ}) - J\sum_{i=1}^{I-1} \hat{\alpha}_i - I\sum_{j=1}^{J-1} \hat{\beta}_j - (IJ - 1)\hat{\xi} - (J-1)\hat{\alpha}_I - (I-1)\hat{\beta}_J = 0$$

$$\hat{\alpha}_i : x_{i.} - J\hat{\alpha}_i - J\hat{\xi} = 0 , \qquad i \neq I$$

$$\hat{\beta}_j : x_{.j} - I\hat{\beta}_j - I\hat{\xi} = 0 , \qquad j \neq J$$

$$\hat{\alpha}_I : (x_{I.} - x_{IJ}) - (J-1)\hat{\alpha}_I - \sum_{j=1}^{J-1} \hat{\beta}_j - (J-1)\hat{\xi} = 0$$

$$\hat{\beta}_J : (x_{.J} - x_{IJ}) - \sum_{i=1}^{I-1} \hat{\alpha}_i - (I-1)\hat{\beta}_J - (I-1)\hat{\xi} = 0$$

$$\sum_i \hat{\alpha}_i = \sum_j \hat{\beta}_j = 0 .$$

These equations reduce immediately to

$$x_{..} - x_{IJ} + \hat{\alpha}_I + \hat{\beta}_J - (IJ - 1)\hat{\xi} = 0$$

$$\hat{\alpha}_i = \frac{1}{J} x_{i.} - \hat{\xi}, \qquad i \neq I$$

$$\hat{\beta}_j = \frac{1}{I} x_{.j} - \hat{\xi}, \qquad j \neq J \tag{2. 1}$$

$$(x_{I.} - x_{IJ}) - (J - 1)\hat{\alpha}_I + \hat{\beta}_J - (J - 1)\hat{\xi} = 0$$

$$(x_{.J} - x_{IJ}) + \hat{\alpha}_I - (I - 1)\hat{\beta}_J - (I - 1)\hat{\xi} = 0 .$$

The last two equations may be multiplied by I and J respectively and then added to give

$$I(x_{I.} - x_{IJ}) + J(x_{.J} - x_{IJ}) + [(1 - J)I + J]\,\hat{\alpha}_I + [I + (1 - I)J]\,\hat{\beta}_J$$

$$- [I(J - 1) + J(I - 1)]\,\hat{\xi} = 0 ,$$

or

$$(IJ - I - J)(\hat{\alpha}_I + \hat{\beta}_J) = I(x_{I.} - x_{IJ}) + J(x_{.J} - x_{IJ}) - (2IJ - I - J)\hat{\xi} .$$

We now use this result to eliminate $(\hat{\alpha}_I + \hat{\beta}_J)$ from the first equation of the system (2. 1) above. We obtain :

$$(IJ-I-J)(x_{..}-x_{IJ})+I(x_{I.}-x_{IJ})+J(x_{.J}-x_{IJ})$$

$$=[(IJ-I-J)(IJ-1)+(2IJ-I-J)]\hat{\xi}=IJ(I-1)(J-1)\hat{\xi},$$

or

$$\hat{\xi}=\frac{(IJ-I-J)}{IJ(I-1)(J-1)}(x_{..}-x_{IJ})+\frac{1}{J(I-1)(J-1)}(x_{I.}-x_{IJ})$$

$$+\frac{1}{I(I-1)(J-1)}(x_{.J}-x_{IJ}).$$

The last two equations of (2. 1) may now be solved for $\hat{\alpha}_I$ and $\hat{\beta}_J$:

$$(I-1)(x_{I.}-x_{IJ})+(x_{.J}-x_{IJ})-[(I-1)(J-1)-1]\hat{\alpha}_I-(I-1)J\hat{\xi}=0.$$

$$(x_{I.}-x_{IJ})+(J-1)(x_{.J}-x_{IJ})-[(I-1)(J-1)-1]\hat{\beta}_J-I(J-1)\hat{\xi}=0.$$

$$\hat{\alpha}_I=\frac{(I-1)}{IJ-I-J}(x_{I.}-x_{IJ})+\frac{1}{IJ-I-J}(x_{.J}-x_{IJ})-\frac{(I-1)J}{IJ-I-J}\hat{\xi}.$$

$$\hat{\beta}_J=\frac{1}{IJ-I-J}(x_{I.}-x_{IJ})+\frac{(J-1)}{IJ-I-J}(x_{.J}-x_{IJ})-\frac{I(J-1)}{IJ-I-J}\hat{\xi}.$$

Summing up, now, CLSM yields the following estimates:

$$\hat{\alpha}_i=\frac{1}{J}x_{i.}-\hat{\xi}, \qquad i \neq I$$

$$\hat{\alpha}_I=\frac{(I-1)}{IJ-I-J}(x_{I.}-x_{IJ})+\frac{1}{IJ-I-J}(x_{.J}-x_{IJ})-\frac{(I-1)J}{IJ-I-J}\hat{\xi}$$

$$\hat{\beta}_j=\frac{1}{I}x_{.j}-\hat{\xi}, \qquad j \neq J \tag{2.2}$$

$$\hat{\beta}_J = \frac{1}{IJ-I-J}(x_{I.}-x_{IJ}) + \frac{(J-1)}{IJ-I-J}(x_{.J}-x_{IJ}) - \frac{I(J-1)}{IJ-I-J}\hat{\xi}$$

$$\hat{\xi} = \frac{IJ-I-J}{IJ(I-1)(J-1)}(x_{..}-x_{IJ}) + \frac{1}{J(I-1)(J-1)}(x_{I.}-x_{IJ})$$

$$+ \frac{1}{I(I-1)(J-1)}(x_{.J}-x_{IJ}) \ .$$

The following interesting comment by Cochran and Cox on missing data serves to end this section and aptly introduce the next section.

> ...missing data may be handled by applying the standard least squares procedure to all observations that are <u>not</u> missing.
>
> ... To the experimenter it may be a difficult business to carry out the construction and solution of a set of unfamiliar normal equations, even though he is quite competent to analyze a set of complete data. For this reason Yates (3.9), following a suggestion by Fisher, considered inserting values for the missing observations so as to obtain a set of complete data.
>
> ... If several observations are absent, a repeated application of the formula enables values to be substituted for each missing observation.
>
> This method is essentially an ingenious computational device whose purpose is to enable the easy computations that apply to complete data to be used even when data are incomplete. Substitution of estimates for missing data does not in any way recover the information that is lost through loss of data, as some experimenters have suggested, usually facetiously; it merely attempts to reproduce the results obtained by an application of the least squares method to the data that are present. The only complete solution of the 'missing data' problem is not to have them (1, p. 72-74).

2. Method of Insertion With One Observation Per Cell

Though the method of insertion is due to Yates, he did not call it that himself. In applying the method, one calculates an estimate of the missing observation, inserts this estimate into the empty cell of the layout, and proceeds as in the analysis of orthogonal data (4).

There are two equivalent ways of looking at the procedure for obtaining an estimate of the missing observation. First we can suppose that we are in possession of the estimates $\hat{\alpha}_i$, $\hat{\beta}_j$, and $\hat{\xi}$. A reasonable prediction for a missing observation x_{ij} is just the mean of the random variable x_{ij}, $E(x_{ij})$. While we do not actually know the value of $E(x_{ij})$ we have an estimate of it. That is, we know that

$$x_{ij} = \alpha_i + \beta_j + \xi + e_{ij} \quad ,$$

and

$$E(x_{ij}) = \alpha_i + \beta_j + \xi \quad ,$$

so

$$\hat{E}(x_{ij}) = \hat{\alpha}_i + \hat{\beta}_j + \hat{\xi} \quad .$$

Accordingly we take as our estimate of x_{ij}

$$\hat{x}_{ij} = \hat{\alpha}_i + \hat{\beta}_j + \hat{\xi} \quad .$$

Second we might set up the sum of squared errors Q, symbolically, as if there were no missing observation, and regard the missing observation, x_{pq} say, as another unknown parameter. We then minimize Q as a function of α_i, β_j, ξ, and x_{pq} thus obtaining the same least squares estimate of x_{pq} as just indicated above.

The latter view will be taken in the work to follow. As before we take x_{IJ} to be the missing observation.

There is a small notational difficulty here which we shall circumvent as follows: Define the symbols $x_{..}^{-}$, $x_{.J}^{-}$, and $x_{I.}^{-}$ by

$$x_{..}^{-} = \sum_{ij \neq IJ} x_{ij}$$

$$x_{.J}^{-} = \sum_{i=1}^{I-1} x_{ij}$$

$$x_{I.}^{-} = \sum_{j=1}^{J-1} x_{ij} \ .$$

This notation permits the use of the symbol $\hat{x}_{IJ}$ for the estimate of x_{IJ} after the normal equations have been solved. Prior to that time x_{IJ} will be used to denote both the missing observation and its estimate.

Formally the normal equations are

$$\frac{\partial Q}{\partial \alpha_i} = \frac{\partial Q}{\partial \beta_j} = \frac{\partial Q}{\partial \xi} = \frac{\partial Q}{\partial x_{IJ}} = 0 \ ,$$

where

$$Q = \sum_{ij} (x_{ij} - \alpha_i - \beta_j - \xi)^2 \quad .$$

In detail they become :

$$\begin{aligned}
x_{i.} - J\hat{\alpha}_i - J\hat{\xi} &= 0 \\
x_{.j} - I\hat{\beta}_j - I\hat{\xi} &= 0 \\
x_{..} - IJ\hat{\xi} &= 0 \\
x_{IJ} - \hat{\alpha}_I - \hat{\beta}_J - \hat{\xi} &= 0 \ ,
\end{aligned} \tag{2.3}$$

where the conditions

$$\sum_i \alpha_i = \sum_j \beta_j = 0$$

have been invoked to simplify the normal equations as they were written down.

From (2.3) we have

$$\begin{aligned}
\hat{\alpha}_I &= \frac{1}{J} x_{I.} - \hat{\xi} \\
\hat{\beta}_J &= \frac{1}{I} x_{.J} - \hat{\xi} \\
\hat{\xi} &= \frac{1}{IJ} x_{..} \quad ,
\end{aligned}$$

so

$$x_{IJ} - \hat{\alpha}_I - \hat{\beta}_J - \hat{\xi} = x_{IJ} - \frac{1}{J}x_{I.} + \hat{\xi} - \frac{1}{I}x_{.J} + \hat{\xi} - \hat{\xi} = 0 ,$$

or

$$x_{IJ} - \frac{1}{J}x_{I.} - \frac{1}{I}x_{.J} + \hat{\xi} = 0$$

$$x_{IJ} - \frac{1}{J}x_{I.} - \frac{1}{I}x_{.J} + \frac{1}{IJ}x_{..} = 0$$

$$x_{IJ} - \frac{1}{J}x_{IJ} - \frac{1}{I}x_{IJ} + \frac{1}{IJ}x_{IJ} = \frac{1}{J}x^-_{I.} + \frac{1}{I}x^-_{.J} - \frac{1}{IJ}x^-_{..}$$

$$x_{IJ}\left(1 - \frac{1}{J} - \frac{1}{I} + \frac{1}{IJ}\right) = \frac{1}{J}x^-_{I.} + \frac{1}{I}x^-_{.J} - \frac{1}{IJ}x^-_{..} ,$$

i. e. ,

$$\hat{x}_{IJ} = \frac{I}{(I-1)(J-1)}x^-_{I.} + \frac{J}{(I-1)(J-1)}x^-_{.J} - \frac{x^-_{..}}{(I-1)(J-1)} .$$

Thus the estimates provided by the method of insertion are:

$$\hat{\alpha}_i = \frac{1}{J}x_{i.} - \hat{\xi} , \quad i \neq I$$

$$\hat{\alpha}_I = \frac{1}{J}(x^-_{I.} + \hat{x}_{IJ}) - \hat{\xi}$$

$$\hat{\beta}_j = \frac{1}{I}x_{.j} - \hat{\xi} , \quad j \neq J$$

$$\hat{\beta}_J = \frac{1}{I}(x^-_{.J} + \hat{x}_{IJ}) - \hat{\xi} \tag{2.4}$$

$$\hat{\xi} = \frac{1}{IJ}(x^-_{..} + \hat{x}_{IJ})$$

$$\hat{x}_{IJ} = \frac{I}{(I-1)(J-1)}x^-_{I.} + \frac{J}{(I-1)(J-1)}x^-_{.J} - \frac{x^-_{..}}{(I-1)(J-1)} .$$

To conclude this section we remark that of the two methods of estimation, the correct least squares method and the method of insertion, the latter is certainly superior from a computational point of view, and hence generally the superior method in view of the fact that both methods produce exactly the same estimates. Compare (2. 4) and (2. 2). Notice, for example, that by the method of insertion we have

$$
\begin{aligned}
\hat{\xi} &= \frac{1}{IJ}(x^-_{..} + \hat{x}_{IJ}) \\
&= \frac{1}{IJ}\left[x^-_{..} + \frac{I}{(I-1)(J-1)}\, x^-_{I.} + \frac{J}{(I-1)(J-1)}\, x^-_{.J} - \frac{x^-_{..}}{(I-1)(J-1)}\right] \\
&= \frac{x^-_{..}}{IJ} + \frac{x^-_{I.}}{J(I-1)(J-1)} + \frac{x^-_{.J}}{I(I-1)(J-1)} - \frac{x^-_{..}}{IJ(I-1)(J-1)} \\
&= \frac{(I-1)(J-1)-1}{IJ(I-1)(J-1)}\, x^-_{..} + \frac{x^-_{I.}}{J(I-1)(J-1)} + \frac{x^-_{.J}}{I(I-1)(J-1)} \\
&= \frac{IJ-I-J}{IJ(I-1)(J-1)}(x_{..} - x_{IJ}) + \frac{1}{J(I-1)(J-1)}(x_{I.} - x_{IJ}) + \frac{1}{I(I-1)(J-1)}(x_{.J} - x_{IJ}),
\end{aligned}
$$

exactly the expression given by CLSM in (2. 2).

Non-orthogonality arising from empty cells in a model designed to have no more than a single observation per cell is, in theory, a special case of the non-orthogonality arising from unequal cell frequencies in models of more general design.

It is to this more general problem that we now turn our attention.

III. UNEQUAL CELL FREQUENCIES

The problem of finding solutions to the normal equations for models with unequal cell frequencies is a difficult one -- at least to the extent that one desires a simple explicit expression for the estimate of each parameter in terms of the observations. Of course, the solution to any linear system of equations, with unique solutions, can be found, at least in principle, using determinants and in practice with modern digital computers. In many instances approximate solutions may be had which provide, with small loss of elegance, results which are, practically, as satisfactory as exact solutions.

We must concern ourselves, then, with less than general models whose cell frequencies exhibit some particular pattern which makes the normal equations easy to solve. Such a pattern exists when cell frequencies are proportional.

1. Cell Frequencies in Rows and Columns Proportional

What we mean by proportional cell frequencies in rows is that the row vectors of cell frequencies in any two rows are proportional. If all row vectors of cell frequencies are proportional then the column vectors of cell frequencies must also be proportional. To be precise, the assumptions to be made in this section are as follows: We shall assume that the cell frequency K_{ij} of the ij^{th} cell can be written

$$K_{ij} = K'_i \cdot K''_j \ , \tag{3.1}$$

and K_{ij} is never zero.

The initial step is to define new parameters a_i, b_j, and x as follows:

$$\begin{aligned} a_i &= \alpha_i - \frac{1}{K'} \sum_{i=1}^{I} K'_i \alpha_i \\ b_j &= \beta_j - \frac{1}{K''} \sum_{j=1}^{J} K''_j \beta_j \\ x &= \xi + \frac{1}{K'} \sum_{i=1}^{I} K'_i \alpha_i + \frac{1}{K''} \sum_{j=1}^{J} K''_j \beta_j \ , \end{aligned} \tag{3.2}$$

where

$$K' = \sum_{i=1}^{I} K'_i \ ,$$

and

$$K'' = \sum_{j=1}^{J} K''_j \ .$$

Under these definitions we have

$$\sum_{i=1}^{I} K_i' a_i = \sum_{i=1}^{I} K_i' (\alpha_i - \frac{1}{K'} \sum_{r=1}^{I} K_r' \alpha_r)$$

$$= \sum_{i=1}^{I} K_i' \alpha_i - \frac{1}{K'} \sum_{i=1}^{I} K_i' \sum_{r=1}^{I} K_r' \alpha_r$$

$$= \sum_{i=1}^{I} K_i' \alpha_i - \sum_{r=1}^{I} K_r' \alpha_r = 0 , \tag{3.3}$$

and similarly

$$\sum_{j=1}^{J} K_j'' b_j = 0 . \tag{3.4}$$

Now

$$a_i + b_j + x = \alpha_i - \frac{1}{K'} \sum_i K_i' \alpha_i + \beta_j - \frac{1}{K''} \sum_j K_j'' \beta_j + \xi$$

$$+ \frac{1}{K'} \sum_i K_i' \alpha_i + \frac{1}{K''} \sum_j K_j'' \beta_j = \alpha_i + \beta_j + \xi , \tag{3.5}$$

so that

$$Q = \sum_{ijk} (x_{ijk} - \alpha_i - \beta_j - \xi)^2 = \sum_{ijk} (x_{ijk} - a_i - b_j - x)^2 .$$

Q will be minimized with respect to the new parameters. It is clear that the minimun so obtained will have exactly the same value as if Q were minimized with respect to the original parameters.

Minimizing with respect to the new parameters now, the normal equations are :

$$x_{...} - \sum_{ijk} \hat{a}_i - \sum_{ijk} \hat{b}_j - K\hat{x} = 0$$

$$x_{i..} - K'' K_i' \hat{a}_i - K_i' \sum_j K_j'' \hat{b}_j - K'' K_i' \hat{x} = 0$$

$$x_{.j.} - K_j'' \sum_i K_i' \hat{a}_i - K' K_j'' \hat{b}_j - K' K_j'' \hat{x} = 0 \ .$$

This system may be simplified by noting that

$$\sum_{ijk} a_i + \sum_{ijk} b_j = \sum_{ij} K_i' K_j'' a_i + \sum_{ij} K_i' K_j'' b_j$$

$$= K'' \sum_i K_i' a_i + K' \sum_j K_j'' b_j$$

$$= K'' \cdot 0 + K' \cdot 0 = 0 \ .$$

The result is :

$$\hat{x} = \frac{1}{K} x_{...} = \bar{x}$$

$$\hat{a}_i = \frac{1}{K'' K_i'} x_{i..} - \bar{x} \tag{3.5}$$

$$\hat{b}_j = \frac{1}{K' K_j''} x_{.j.} - \bar{x} \ .$$

The next step will be to replace the new parameters in (3.5) by their equivalents in terms of the original parameters, but first notice that

$$0 = \sum_{i=1}^{I} \alpha_i = \sum_{i=1}^{I} (a_i + \frac{1}{K'} \sum_{r=1}^{I} K'_r \alpha_r)$$

$$= \sum_{i=1}^{I} a_i + \frac{I}{K'} \sum_{i=1}^{I} K'_i \alpha_i \ ,$$

that is that

$$\frac{1}{K'} \sum_i K'_i \alpha_i = - \frac{1}{I} \sum_i a_i \ . \tag{3.6}$$

Similarly

$$\frac{1}{K''} \sum_j K''_j \beta_j = - \frac{1}{J} \sum_j b_j \ . \tag{3.7}$$

By the definitions (3.2) the system (3.5) becomes:

$$\hat{\xi} = -\frac{1}{K'} \sum_i K'_i \hat{\alpha}_i - \frac{1}{K''} \sum_j K''_j \hat{\beta}_j + \overline{x}$$

$$\hat{\alpha}_i = \frac{1}{K'} \sum_i K'_i \hat{\alpha}_i + \frac{1}{K'' K'_i} x_{i..} - \overline{x}$$

$$\hat{\beta}_j = \frac{1}{K''} \sum_j K''_j \hat{\beta}_j + \frac{1}{K' K''_j} x_{.j.} - \overline{x} \ ,$$

which in the light of (3.6) and (3.7) becomes:

$$\hat{\xi} = \frac{1}{I}\sum_i \hat{a}_i + \frac{1}{J}\sum_j \hat{b}_j + \bar{x}$$

$$\hat{\alpha}_i = -\frac{1}{I}\sum_i \hat{a}_i + \frac{1}{K''K_i'} x_{i..} - \bar{x}$$

$$\hat{\beta}_j = -\frac{1}{J}\sum_j \hat{b}_j + \frac{1}{K'K_j''} x_{.j.} - \bar{x} .$$

If we introduce the notation

$$\bar{x}_{i..} = \frac{1}{K''K_i'} x_{i..}$$

$$\bar{x}_{.j.} = \frac{1}{K'K_j''} x_{.j.} ,$$

and substitute the estimates (3.5) in this last system we arrive at :

$$\hat{\xi} = \frac{1}{I}\sum_i (\bar{x}_{i..} - \bar{x}) + \frac{1}{J}\sum_j (\bar{x}_{.j.} - \bar{x}) + \bar{x} ,$$

$$\hat{\alpha}_i = \bar{x}_{i..} - \frac{1}{I}\sum_i (\bar{x}_{i..} - \bar{x}) - \bar{x} ,$$

$$\hat{\beta}_j = \bar{x}_{.j.} - \frac{1}{J}\sum_j (\bar{x}_{.j.} - \bar{x}) - \bar{x} ,$$

which is easily manipulated to give the following least squares estimates:

$$\hat{\xi} = \frac{1}{I}\sum_{i} \overline{x}_{i..} + \frac{1}{J}\sum_{j} \overline{x}_{.j.} - \overline{x}$$

$$\hat{\alpha}_i = \overline{x}_{i..} - \frac{1}{I}\sum_{i} \overline{x}_{i..} \tag{3.8}$$

$$\hat{\beta}_j = \overline{x}_{.j.} - \frac{1}{J}\sum_{j} \overline{x}_{.j.} \quad .$$

These differ from (3.5) only by constants.

The interpretation of these formulas is, in the case of $\hat{\alpha}_i$ for example, as follows: The estimate of the row effect α_i for the i^{th} row is the difference between the average of observations in the i^{th} row and the average of all row averages.

It is doubtful that there exists any other pattern of cell frequencies that will produce expressions for the estimates as simple and compact as have been had under the assumption of proportional cell frequencies. Still there is a small class of models that may be dealt with efficiently. This special case will be considered next.

2. Cell Frequencies Unequal in Fewer Than $\min\{I-1, J-1\}$ Cells

Henry Scheffé has shown that for a model with a layout of I rows and J columns and completely arbitrary cell frequencies K_{ij}, the problem of solving the normal equations under given linear constraints may be reduced to the solving of a system of r

equations in r unknowns, where

$$r = \min\{I-1, J-1\}$$

(3, p. 114-115). In certain special cases of unequal cell frequencies the estimates may be calculated with a considerable saving in labor over that required for the solving of a general system of r equations in r unknowns, viz. in the case where unequal frequencies occur in fewer than r cells.

Let

$$m = \max_{ij} \{K_{ij}\} .$$

We take the view that if $K_{ij} < m$, then the ij^{th} cell has $m-K_{ij}$ missing observations. We propose to fill each deficient cell with estimates of the missing observations, obtained by the method of insertion. If n is the number of deficient cells the estimates of the missing observations may be had by solving a system of n equations in n unknowns. If $n > r$ it will generally be simpler to solve the system produced by Scheffé's simplification. When $n < r$ the proposed method provides the simpler procedure, and if n is much smaller than r the saving in labor is significant.

To demonstrate the details of the procedure imagine first an experimental design model where the cell frequency K_{ij} is equal to m in all but n cells, $n < r$. Let us suppose, for convenience in notation, that the k^{th} observation is missing in cells $i_1j_1, \cdots, i_nj_n$. The missing observations will be denoted accordingly by

$$x_{i_1j_1k}, \cdots, x_{i_nj_nk} .$$

Now the normal equations for a layout with m observations in every cell may be written :

$$\begin{aligned} \hat{\alpha}_i &= \frac{1}{mJ} x_{i..} - \bar{x} \\ \hat{\beta}_j &= \frac{1}{mI} x_{.j.} - \bar{x} \\ \hat{\xi} &= \bar{x} , \end{aligned} \tag{3.11}$$

using the conditions

$$\sum_i \alpha_i = \sum_j \beta_j = 0 ,$$

and the estimates of the missing observations are had, according to the method of insertion, by solving the equations :

$$\begin{aligned} & x_{i_1j_1k} - \hat{\alpha}_{i_1} - \hat{\beta}_{j_1} - \hat{\xi} = 0 \\ & \cdot \\ & \cdot \\ & \cdot \\ & x_{i_nj_nk} - \hat{\alpha}_{i_n} - \hat{\beta}_{j_n} - \hat{\xi} = 0 , \end{aligned}$$

subject to these normal equations. Thus the system which must be solved to obtain the estimates of the missing observations is

$$
\begin{aligned}
&x_{i_1 j_1 k} - \frac{1}{mJ} x_{i_1 \cdot\cdot} - \frac{1}{mI} x_{\cdot j_1 \cdot} + \bar{x} = 0 \\
&\cdot \\
&\cdot \\
&\cdot \\
&x_{i_n j_n k} - \frac{1}{mJ} x_{i_n \cdot\cdot} - \frac{1}{mI} x_{\cdot j_n \cdot} + \bar{x} = 0 \ .
\end{aligned}
\tag{3.12}
$$

We have supposed that there is only one observation missing from each deficient cell, but such a restriction is unnecessary -- in fact, the analysis is not complicated by supposing that all of the deficient cells are completely empty. To see this let x_{ijr} and x_{ijs} be any two observations missing from the ij^{th} cell. Then, according to the method of insertion, denoting both the observations and their estimates, as we have given ourselves license to do, by x_{ijr} and x_{ijs}, the estimates must satisfy the two equations

$$x_{ijr} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\xi} = 0$$

$$x_{ijs} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\xi} = 0 \ .$$

Thus

$$x_{ijr} = x_{ijs} \ , \tag{3.13}$$

that is, any two observations missing from the same cell have the same estimate.

The method is highly efficient when there are only one or two deficient cells. For example suppose that the t observations $x_{ij1}, \cdots, x_{ijt}$ are the only missing observations. Then, if we denote the sums of all observations present in the i^{th} row and in the j^{th} column by $x^-_{i..}$ and $x^-_{.j.}$ respectively, and the sum of all observations present in the whole layout by $x^-_{...}$ we have

$$x_{ij1} - \frac{1}{mJ} x_{i..} - \frac{1}{mI} x_{.j.} + \overline{x} = 0 ,$$

$$x_{ij1} - \frac{t}{mJ} x_{ij1} - \frac{1}{mJ} x^-_{i..} - \frac{t}{mI} x_{ij1} - \frac{1}{mI} x^-_{.j.} + \frac{t}{mIJ} x_{ij.} + \frac{1}{mIJ} x^-_{...} = 0,$$

$$(1 - \frac{t}{mJ} - \frac{t}{mI} + \frac{t}{mIJ}) x_{ij1} = \frac{x^-_{i..}}{mJ} + \frac{x^-_{.j.}}{mI} - \frac{x^-_{...}}{mIJ} ,$$

$$\hat{x}_{ij1} = \frac{1}{(mIJ - tI - tJ + t)} (Ix^-_{i..} + Jx^-_{.j.} - x^-_{...}) , \qquad (3.14)$$

and

$$\hat{x}_{ij1} = \hat{x}_{ij2} = \cdots = \hat{x}_{ijt}.$$

One has now only to substitute these estimates for the missing observations and read off the estimates of the parameters from Equations (3.11).

This method becomes progressively complicated as deficient cells become more numerous. However it is possible to obtain explicit expressions for the estimates of the missing observations for the case of as many as $r+1$ deficit cells, if the pattern of deficient cells in the layout has a special character.

The requisite character is possessed by a model in which there is no more than one deficient cell in any row or column.

3. No More than a Single Deficient Cell in Each Row and Column of an $I \times J$ Layout

For notational convenience we shall assume that the deficient cells lie on the main diagonal of the layout in the first n rows, where the main diagonal is defined in such a way as to include the first cell in the first row. There is clearly no loss of generality in this assumption since if such a pattern does not exist, we have merely to effect a rearrangement of the original array to bring the deficient cells onto the main diagonal and onto the first n rows. We take the designations of the deficient cells to be

$$i_1 i_1 \quad , \cdots , \quad i_n i_n$$

where

$$n \leq \min\{I, J\} \ .$$

We shall assume that there are m observations in each non-deficient cell, that there are t observations missing from each deficient cell, and that it is the first t observations missing in each case. Thus the observations missing from cell $i_1 i_1$ are denoted by

$$x_{i_1 i_1 1}, \ldots, x_{i_1 i_1 t} .$$

To obtain estimates of the missing observations we must minimize

$$Q = \sum_{ijk} (x_{ijk} - \alpha_i - \beta_j - \xi)^2 .$$

Thus we consider the following system of equations :

$$\begin{aligned} & x_{i_1 i_1 1} - \hat{\alpha}_{i_1} - \hat{\beta}_{i_1} - \hat{\xi} = 0 \\ & \cdot \\ & \cdot \\ & \cdot \\ & x_{i_n i_n 1} - \hat{\alpha}_{i_n} - \hat{\beta}_{i_n} - \hat{\xi} = 0 \end{aligned} \tag{3.15}$$

$$\begin{aligned} \hat{\alpha}_i &= \frac{1}{mJ} x_{i..} - \overline{x} \\ \hat{\beta}_j &= \frac{1}{mI} x_{.j.} - \overline{x} \\ \hat{\xi} &= \overline{x} \end{aligned} \tag{3.16}$$

$$\sum_i \alpha_i = \sum_j \beta_j = 0 . \tag{3.17}$$

We substitute from (3.16) into (3.15) and obtain

$$
\begin{aligned}
&x_{i_1 i_1 1} - \frac{1}{mJ} x_{i_1 \cdot\cdot} - \frac{1}{mI} x_{\cdot i_1 \cdot} + \bar{x} = 0 \\
&\cdot \\
&\cdot \\
&\cdot \\
&x_{i_n i_n 1} - \frac{1}{mJ} x_{i_n \cdot\cdot} - \frac{1}{mI} x_{\cdot i_n \cdot} + \bar{x} = 0 .
\end{aligned} \tag{3.18}
$$

Recalling that $x_{i_1 i_1 1}$ may be used to denote any of the observations missing from the $i_1 i_1{}^{th}$ cell and similarly for the other deficient cells, we may write (3.18) in the form

$$
\begin{aligned}
&\left(1+\frac{t}{mIJ}-\frac{t}{mI}-\frac{t}{mJ}\right)x_{i_1 i_1 1}+\frac{t}{mIJ}x_{i_2 i_2 1}+\cdots+\frac{t}{mIJ}x_{i_n i_n 1}=\frac{1}{mJ}\bar{x}_{i_1 \cdot\cdot}+\frac{1}{mI}\bar{x}_{\cdot i_1 \cdot}-\frac{1}{mIJ}\bar{x}_{\cdots} \\
&\cdot \\
&\cdot \\
&\cdot \\
&\frac{t}{mIJ}x_{i_1 i_1 1}+\cdots+\frac{t}{mIJ}x_{i_{n-1} i_{n-1} 1}+\left(1+\frac{t}{mIJ}-\frac{t}{mI}-\frac{t}{mJ}\right)x_{i_n i_n 1}=\frac{1}{mJ}\bar{x}_{i_n \cdot\cdot}+\frac{1}{mI}\bar{x}_{\cdot i_n \cdot}-\frac{1}{mIJ}\bar{x}_{\cdots} .
\end{aligned} \tag{3.19}
$$

If we multiply each of the n equations in (3.19) on both sides by mIJ we obtain

$$
\begin{aligned}
&(mIJ+t-tJ-tI)x_{i_1 i_1 1}+tx_{i_2 i_2 1}+\cdots+tx_{i_n i_n 1} = I\bar{x}_{i_1 \cdot\cdot}+J\bar{x}_{\cdot i_1 \cdot}-\bar{x}_{\cdots} \\
&\cdot \\
&\cdot \\
&\cdot \\
&tx_{i_1 i_1 1}+\cdots+tx_{i_{n-1} i_{n-1} 1}+(mIJ+t-tJ-tI)x_{i_n i_n 1} = I\bar{x}_{i_n \cdot\cdot}+J\bar{x}_{\cdot i_n \cdot}-\bar{x}_{\cdots} .
\end{aligned} \tag{3.20}
$$

Adding all these equations gives

$$Mx_{i_1 i_1 1} + \cdots + Mx_{i_n i_n 1} = I \sum_{r=1}^{n} \bar{x}_{i_r ..} + J \sum_{r=1}^{n} \bar{x}_{. i_r .} - n\bar{x}_{...} \quad , \tag{3.21}$$

where

$$M = mIJ + nt - tJ - tI \ . \tag{3.22}$$

Divide both sides of (3.21) by M and denote the right hand side of the resulting equation by X. This gives

$$x_{i_1 i_1 1} + \cdots + x_{i_n i_n 1} = X \ . \tag{3.23}$$

Divide each equation in (3.20) on both sides by t, then subtract (3.23) from each resulting equation. The system obtained by this procedure is

$$\begin{aligned} \frac{mIJ-tJ-tI}{t} \hat{x}_{i_1 i_1 1} &= \frac{I}{t} \bar{x}_{i_1 ..} + \frac{J}{t} \bar{x}_{. i_1 .} - \frac{\bar{x}_{...}}{t} - X \\ &\ \ \vdots \\ \frac{mIJ-tJ-tI}{t} \hat{x}_{i_n i_n 1} &= \frac{I}{t} \bar{x}_{i_n ..} + \frac{J}{t} \bar{x}_{. i_n .} - \frac{\bar{x}_{...}}{t} - X \ . \end{aligned} \tag{3.24}$$

From (3.24) one obtains as the estimate of $x_{i_s i_s 1}$,

$$\hat{x}_{i_s i_s 1} = \left\{\frac{t}{mIJ-tJ-tI}\right\}\left\{\frac{I}{t}\bar{x}_{i_s..} + \frac{J}{t}\bar{x}_{.i_s.} - \frac{\bar{x}_{...}}{t} - X\right\} =$$

$$\left\{\frac{t}{mIJ-tJ-tI}\right\}\left\{\frac{I}{t}\bar{x}_{i_s..} + \frac{J}{t}\bar{x}_{.i_s.} - \frac{\bar{x}_{...}}{t}\right.$$

$$\left. - \left(\frac{1}{mIJ+nt-tJ-tI}\right)\left(I\sum_{r=1}^{n}\bar{x}_{i_r..} + J\sum_{r=1}^{n}\bar{x}_{.i_r.} - n\bar{x}_{...}\right)\right\} . \qquad (3.25)$$

The estimates given by (3.25) are substituted for the missing observations in Equations (3.16), and the estimates $\hat{\alpha}_i$, $\hat{\beta}_j$, and $\hat{\xi}$ take their usual form.

The pattern considered in this section does not permit more than one deficient cell in any row or column. We consider next a generalization of this pattern which permits a plurality of deficient cells in both rows and columns.

4. The Indices of Deficient Cells Form a Cartesian Product Set

The assumptions for this section are as follows: The layout is $I \times J$, there are m observations in each non-deficient cell, and the first t observations are missing from each deficient cell. The indices of deficient cells form a Cartesian product set, that is, if rows $i_1, \cdots, i_p$ contain deficient cells and columns $j_1, \cdots, j_q$ contain deficient cells, then cell ij is deficient if

$$i \epsilon S = \{i_1, \cdots, i_p\}$$

and

$$j \in T = \{j_1, \cdots, j_q\} \quad .$$

We achieve convenience in notation and suffer no loss in generality if we assume that it is the first p rows and first q columns which contain the deficient cells. Thus we assume that

$$S = \{1, \cdots, p\}$$

and

$$T = \{1, \cdots, q\} \quad ,$$

so that the deficient cells form a block in the upper left hand corner of the layout (Figure 1). In Figure 1, deficient cells are marked with an $\times$; non-deficient cells are unmarked. If $t = m$, we assume $p < I$ and $q < J$, as otherwise some rows or columns would contain no observations, and the corresponding effects could not be estimated, while the non-deficient cells would continue to form a (smaller) rectangular array.

From the minimizing of the sum of squared errors Q as a function of the missing observations, we obtain the system of equations

$$x_{ij1} - \alpha_i - \beta_j - \xi = 0, \quad i \in S, \quad j \in T. \tag{3.26}$$

As usual we also have

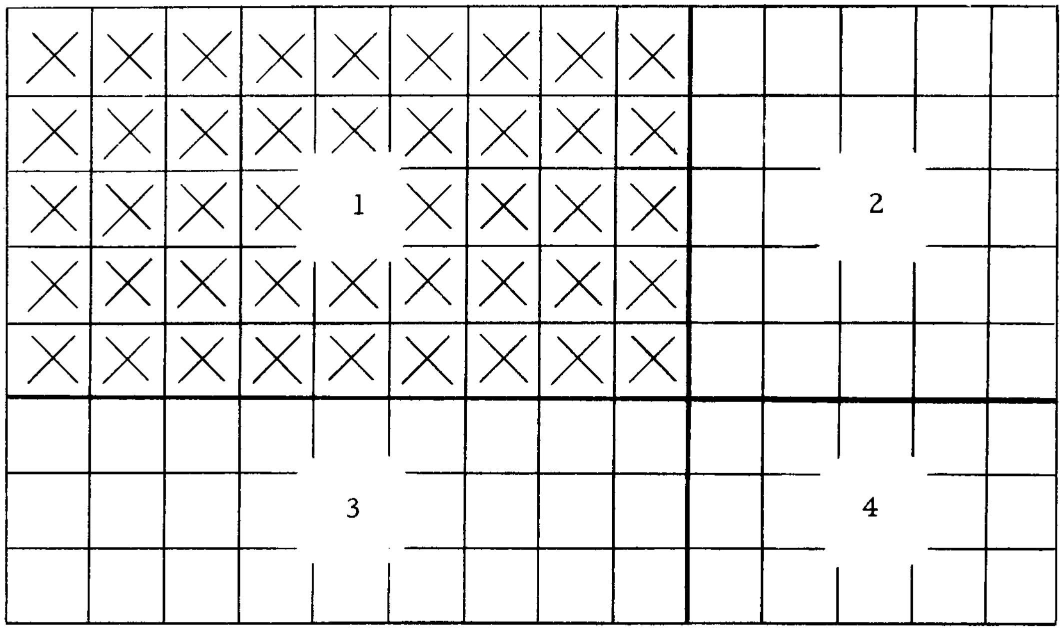

Figure 1. Indices of deficient cells form a Cartesian product set.

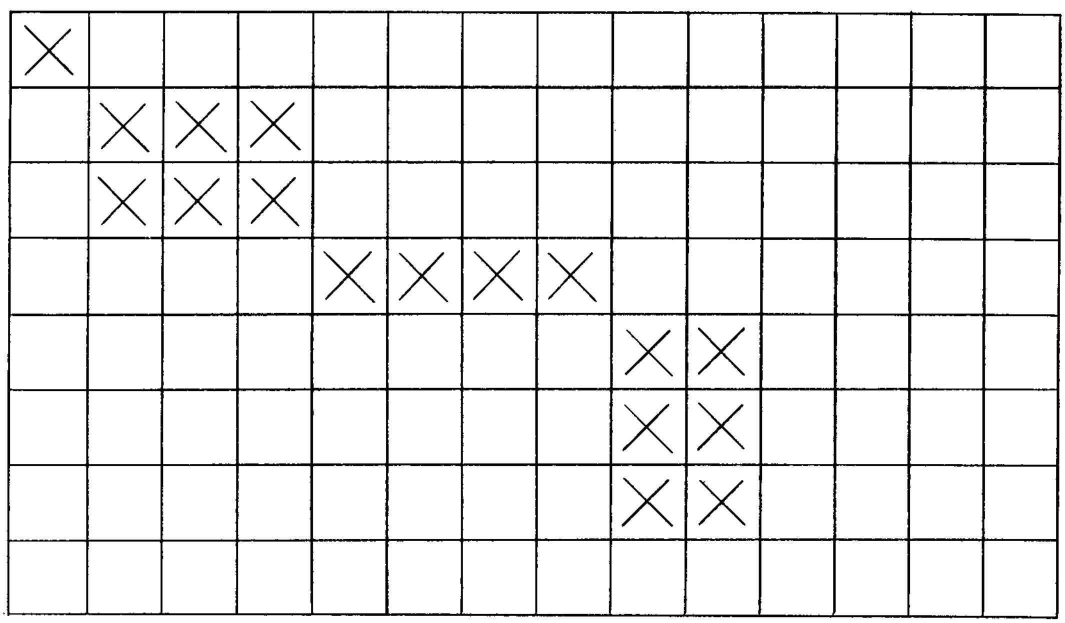

Figure 2. Indices of deficient cells form disjoint sets of Cartesian product sets.

$$\hat{\alpha}_i = \frac{1}{mJ} x_{i..} - \bar{x}$$

$$\hat{\beta}_j = \frac{1}{mI} x_{.j.} - \bar{x} \qquad (3.27)$$

$$\hat{\xi} = \bar{x}$$

$$\sum_i \alpha_i = \sum_j \beta_j = 0 \ .$$

Substituting from (3.27) into (3.26) we have, for $i \epsilon S$, $j \epsilon T$,

$$x_{ij1} - \frac{1}{mJ} x_{i..} - \frac{1}{mI} x_{.j.} + \bar{x} = 0 , \qquad (3.28)$$

where x_{ij1} is used to represent any of the symbols $x_{ij1}, \cdots, x_{ijt}$, since they all have the same estimate. Referring to (3.27) one sees that it is not actually necessary to possess an estimate of each missing observation. It clearly would suffice to have estimates of the sum of all observations missing from any row, from any column, and from the whole layout. Thus it would suffice to have estimates of $x^{+}_{i..}$, $x^{+}_{.j.}$, and $x^{+}_{...}$, where $x^{+}_{i..}$ denotes the sum of observations missing from the i^{th} row, $1 \leq i \leq p$, $x^{+}_{.j.}$ the sum of observations missing from the j^{th} column, $1 \leq j \leq q$, and $x^{+}_{...}$ the sum of all missing observations. In the calculation to follow the symbols $x^{-}_{i..}$, $x^{-}_{.j.}$, and $x^{-}_{...}$ are to be defined as in the preceding section. Notice that

$$x^+_{i..} + x^-_{i..} = x_{i..}$$

$$x^+_{.j.} + x^-_{.j.} = x_{.j.}$$

$$x^+_{...} + x^-_{...} = x_{...} \quad .$$

There are pq equations represented by (3.28). We multiply each equation by t and sum first over i, and then over j, and finally over all ij, $1 \leq i \leq p$, $1 \leq j \leq q$, to obtain

$$\begin{aligned}
&x^+_{.j.} - \frac{t}{mJ} \sum_{i=1}^{p} x_{i..} - \frac{tp}{mI} x_{.j.} + tp\,\bar{\bar{x}} = 0 \\
&x^+_{i..} - \frac{tq}{mJ} x_{i..} - \frac{t}{mI} \sum_{j=1}^{q} x_{.j.} + tq\,\bar{\bar{x}} = 0 \\
&x^+_{...} - \frac{tq}{mJ} \sum_{i=1}^{p} x_{i..} - \frac{tp}{mI} \sum_{j=1}^{q} x_{.j.} + tpq\,\bar{\bar{x}} = 0 \quad .
\end{aligned} \tag{3.29}$$

This may be rewritten in the form

$$\begin{aligned}
&x^+_{.j.} - \frac{t}{mJ} x^+_{...} - \frac{tp}{mI} x^+_{.j.} + \frac{tp}{mIJ} x^+_{...} = \frac{t}{mJ} \sum_{i=1}^{p} x^-_{i..} + \frac{tp}{mI} x^-_{.j.} - \frac{tp}{mIJ} x^-_{...} \\
&x^+_{i..} - \frac{tq}{mJ} x^+_{i..} - \frac{t}{mI} x^+_{...} + \frac{tq}{mIJ} x^+_{...} = \frac{tq}{mJ} x^-_{i..} + \frac{t}{mI} \sum_{j=1}^{q} x^-_{.j.} - \frac{tq}{mIJ} x^-_{...} \\
&x^+_{...} - \frac{tq}{mJ} x^+_{...} - \frac{tp}{mI} x^+_{...} + \frac{tpq}{mIJ} x^+_{...} = \frac{tq}{mJ} \sum_{i=1}^{p} x^-_{i..} + \frac{tp}{mI} \sum_{j=1}^{q} x^-_{.j.} - \frac{tpq}{mIJ} x^-_{...} \quad .
\end{aligned}$$

(3.30)

These equations reduce to

$$\hat{x}^{+}_{...} = \left\{\frac{t}{mIJ-tqI-tpJ+tpq}\right\}\left\{qI\sum_{i=1}^{p} x^{-}_{i..} + pJ\sum_{j=1}^{q} x^{-}_{.j.} - pq\,x^{-}_{...}\right\}$$

$$\hat{x}^{+}_{i..} = \left\{\frac{tJ}{mJ-tq}\right\}\left\{\frac{q}{J}x^{-}_{i..} + \frac{1}{I}\sum_{j=1}^{q} x^{-}_{.j.} - \frac{q}{IJ}x^{-}_{...} + \frac{J-q}{IJ}\hat{x}^{+}_{...}\right\} \qquad (3.31)$$

$$\hat{x}^{+}_{.j.} = \left\{\frac{tI}{mI-tp}\right\}\left\{\frac{1}{J}\sum_{i=1}^{p} x^{-}_{i..} + \frac{p}{I}x^{-}_{.j.} - \frac{p}{IJ}x^{-}_{...} + \frac{I-p}{IJ}\hat{x}^{+}_{...}\right\},$$

and we have obtained the estimates we seek. With (3.31) the estimates $\hat{\alpha}_i$, $\hat{\beta}_j$, and $\hat{\xi}$ are immediate from (3.27). From a computational point of view the formulas here are probably in their most suitable form. When deficient cells are empty, i.e. when $t = m$, the Equations (3.31) can be rewritten in a form that makes them easy to remember by considering the natural division of the layout induced by the rectangular block of empty cells (Figure 1).

$$\frac{1}{mpq}\hat{x}^{+}_{...} = \frac{1}{mp(J-q)}\sum_{i=1}^{p} x^{-}_{i..} + \frac{1}{mq(I-p)}\sum_{j=1}^{q} x^{-}_{.j.} - \frac{1}{m(I-p)(J-q)}\sum_{i=p+1}^{I}\sum_{j=q+1}^{J} x_{ij.}$$

$$\frac{1}{mq}\hat{x}^{+}_{i..} = \frac{1}{m(J-q)}x^{-}_{i..} + \frac{1}{mqI}\left(\hat{x}^{+}_{...} + \sum_{j=1}^{q} x^{-}_{.j.}\right) - \frac{1}{mI(J-q)}\sum_{i=1}^{I}\sum_{j=q+1}^{J} x_{ij.} \qquad (3.31a)$$

$$\frac{1}{mp}\hat{x}^{+}_{.j.} = \frac{1}{m(I-p)}x^{-}_{.j.} + \frac{1}{mpJ}\left(\sum_{i=1}^{p} x^{-}_{i..} + \hat{x}^{+}_{...}\right) - \frac{1}{mJ(I-p)}\sum_{i=p+1}^{I}\sum_{j=1}^{J} x_{ij.} \quad .$$

Thus the average of observations missing from block 1 is equal to the average of observations in block 2 plus the average of observations in block 3 minus the average of observations in block 4. An equally appealing description applies to the second equation in (3.31a), except that one must bear in mind that there are, in fact, no observations present in block 1, and one must use the estimate for the sum of these missing observations as given by the first equation of (3.31a). With this reservation then, the average of observations missing from the i^{th} row is equal to the average of observations present in the i^{th} row plus the average of observations in blocks 1 and 3 minus the average of observations in blocks 2 and 4. The third equation is described analogously.

It is possible to extend this result to a yet more general model, whose description is as follows: The layout is $I \times J$, non-deficient cells contain m observations, and deficient cells are missing the first t observations. The set of all deficient cells can be decomposed into n disjoint subsets $N_1, \cdots, N_n$ in such a way that the indices of deficient cells in N_r, $r = 1, \cdots, n$ form a Cartesian product set, and no row or column contains deficient cells from more than one N_r. Without loss of generality, it may be assumed that the cells in N_1 form a block in the upper right hand corner of the layout, the cells in N_2 form a block situated immediately below and to the right of the first block, and so forth, so that the deficient cells form a chain of blocks extending downward and to the right across the layout from the upper left hand corner. An example of such a layout is given in Figure 2. The position of each deficient cell is made precise by specifying the product set associated with each N_r.

$$N_1 : S_1 \times T_1, \quad S_1 = \{1, \cdots, p_1\}, \quad T_1 = \{1, \cdots, q_1\}$$

$$\vdots$$

$$N_r : S_r \times T_r, \quad S_r = \{p_{r-1}+1, \cdots, p_r\}, \quad T_r = \{q_{r-1}+1, \cdots, q_r\}$$

$$\vdots$$

$$N_n : S_n \times T_n, \quad S_n = \{p_{n-1}+1, \cdots, p_n\}, \quad T_n = \{q_{n-1}+1, \cdots, q_n\} \quad .$$

Thus, for example, cell ij belongs to N_r if $i \epsilon S_r$ and $j \epsilon T_r$. Clearly $p_n \leq I$, $q_n \leq J$.

In minimizing Q as a function of the missing observations we are led to n sets of equations, each like the system (3. 26), corresponding to the n sets $N_1, \cdots, N_n$. The equations associated with N_r are

$$x_{ij1} - \alpha_i - \beta_j - \xi = 0, \quad i \epsilon S_r, \quad j \epsilon T_r. \tag{3. 32}$$

In exactly the same way as before, we substitute from (3. 27) into (3. 32) to obtain, for $i \epsilon S_r$, $j \epsilon T_r$

$$x_{ij1} - \frac{1}{mJ} x_{i..} - \frac{1}{mI} x_{.j.} + \overline{x} = 0 \quad . \tag{3. 33}$$

Before proceeding, we pause to define the following symbols :

$x^+_{i..}$ = sum of observations missing from the i^{th} row,

$x^+_{.j.}$ = sum of observations missing from the j^{th} column,

y_r = sum of observations missing from cells belonging to N_r ,

$x^+_{...}$ = sum of all missing observations.

Here, as earlier, we take the term "missing observation" to mean the symbol which is used to represent the observation when it is not missing.

As in the earlier model, we do not require an estimate of each individual missing observation, but only estimates of the quantities represented by the symbol just defined. Following the procedure by which (3. 29) was obtained from (3. 28), we multiply each equation in (3. 33) by t, then sum the equations, first on i, then on j, and finally over all ij, $p_{r-1}+1 \leq i \leq p_r$, $q_{r-1}+1 \leq j \leq q_r$. We adopt the convention $p_0 = q_0 = 0$ and arrive at

$$x^+_{.j.} - \frac{t}{mJ}\sum_{i=p_{r-1}+1}^{p_r} x_{i..} - \frac{t(p_r-p_{r-1})}{mI} x_{.j.} + t(p_r-p_{r-1})\,\overline{x} = 0$$

$$x^+_{i..} - \frac{t(q_r-q_{r-1})}{mJ} x_{i..} - \frac{t}{mI}\sum_{j=q_{r-1}+1}^{q_r} x_{.j.} + t(q_r-q_{r-1})\,\overline{x} = 0 \qquad (3.34)$$

$$y_r - \frac{t(q_r-q_{r-1})}{mJ}\sum_{i=p_{r-1}+1}^{p_r} x_{i..} - \frac{t(p_r-p_{r-1})}{mI}\sum_{j=q_{r-1}+1}^{q_r} x_{.j.} + t(p_r-p_{r-1})(q_r-q_{r-1})\overline{x} = 0.$$

Rewriting (3.34) we have

$$[1-\frac{t(p_r-p_{r-1})}{mI}]\, x^+_{.j.} = \frac{t}{mJ} y_r - \frac{t(p_r-p_{r-1})}{mIJ} x^+_{...} + \frac{t}{mJ}\sum_{i=p_{r-1}+1}^{p_r} x^-_{i..} + $$

$$+ \frac{t(p_r-p_{r-1})}{mI} x^-_{.j.} - \frac{t(p_r-p_{r-1})}{mIJ} x^-_{...} \quad ,$$

$$[1-\frac{t(q_r-q_{r-1})}{mJ}]\, x^+_{i..} = \frac{t}{mI} y_r - \frac{t(q_r-q_{r-1})}{mIJ} x^+_{...} + \frac{t}{mI}\sum_{j=q_{r-1}+1}^{q_r} x^-_{.j.} + $$

$$+ \frac{t(q_r-q_{r-1})}{mJ} x^-_{i..} - \frac{t(q_r-q_{r-1})}{mIJ} x^-_{...} \quad , \tag{3.35}$$

$$[1-\frac{t(q_r-q_{r-1})}{mJ} - \frac{t(p_r-p_{r-1})}{mI}]\, y_r = -\frac{t(p_r-p_{r-1})(q_r-q_{r-1})}{mIJ} x^+_{...} + $$

$$+ \frac{t(q_r-q_{r-1})}{mJ}\sum_{i=p_{r-1}+1}^{p_r} x^-_{i..} + \frac{t(p_r-p_{r-1})}{mI}\sum_{j=q_{r-1}+1}^{q_r} x^-_{.j.}$$

$$- \frac{t(p_r-p_{r-1})(q_r-q_{r-1})}{mIJ} x^-_{...} \quad .$$

Rewriting once more we have finally

$$\hat{x}^{+}_{.j.} = \left\{\frac{tI}{mI-t(p_r-p_{r-1})}\right\} \left\{ \frac{1}{J}\sum_{i=p_{r-1}+1}^{p_r} x^{-}_{i..} + \frac{1}{I}(p_r-p_{r-1})x^{-}_{.j.} - \frac{(p_r-p_{r-1})}{IJ} x^{-}_{...} + \frac{1}{J}\hat{y}_r - \frac{(p_r-p_{r-1})}{IJ}\hat{x}^{+}_{...} \right\},$$

$$\hat{x}^{+}_{i..} = \left\{\frac{tJ}{mJ-t(q_r-q_{r-1})}\right\} \left\{\frac{(q_r-q_{r-1})}{J} x^{-}_{i..} + \frac{1}{I}\sum_{j=q_{r-1}+1}^{q_r} x^{-}_{.j.} - \frac{(q_r-q_{r-1})}{IJ} x^{-}_{...} + \frac{1}{I}\hat{y}_r - \frac{(q_r-q_{r-1})}{IJ}\hat{x}^{+}_{...} \right\}, \tag{3.36}$$

$$\hat{y}_r = \left\{\frac{tIJ}{mIJ-tI(q_r-q_{r-1})-tJ(p_r-p_{r-1})}\right\} \left\{\frac{(q_r-q_{r-1})}{J}\sum_{i=p_{r-1}+1}^{p_r} x^{-}_{i..} + \frac{(p_r-p_{r-1})}{I}\sum_{j=q_{r-1}+1}^{q_r} x^{-}_{.j.} - \frac{(p_r-p_{r-1})(q_r-q_{r-1})}{IJ} x^{-}_{...} - \frac{(p_r-p_{r-1})(q_r-q_{r-1})}{IJ}\hat{x}^{+}_{...} \right\}.$$

It is clear that (3.36) will give us the estimates we need if we can obtain the estimate $\hat{x}^{+}_{...}$. This estimate can be had in the following way: First note that

$$\sum_{r=1}^{n} y_r = x^+_{...} \ .$$

Thus if we sum the last equation in (3. 36) over r we may solve the resulting equation for $x^+_{...}$. The result is

$$\hat{x}^+_{...} = \{1 + t\sum_{r=1}^{n} [\frac{(p_r - p_{r-1})(q_r - q_{r-1})}{mIJ - tI(q_r - q_{r-1}) - tJ(p_r - p_{r-1})}]\}^{-1}$$

$$\cdot \sum_{r=1}^{n} \{[\frac{tIJ}{mIJ - tI(q_r - q_{r-1}) - tJ(p_r - p_{r-1})}]$$

$$\cdot [\frac{(q_r - q_{r-1})}{J} \sum_{i=p_{r-1}+1}^{p_r} \bar{x}_{i..} + \frac{(p_r - p_{r-1})}{I} \sum_{j=q_{r-1}+1}^{q_r} \bar{x}_{.j.} \qquad (3.37)$$

$$- \frac{(p_r - p_{r-1})(q_r - q_{r-1})}{IJ} \bar{x}_{...}]\} \ .$$

The work is now complete. To use these formulas one first calculates $\hat{x}^+_{...}$ from (3. 37) and substitutes this result into the last equation of (3. 36) to obtain the n estimates $\hat{y}_1, \cdots, \hat{y}_n$. The estimates provided by the first two equations of (3. 36) can then be calculated and all estimates substituted as appropriate into (3. 27) to obtain, finally, the estimates $\hat{\alpha}_i$, $\hat{\beta}_j$, and $\hat{\xi}$.

IV. ITERATIVE SOLUTION

In the preceding chapters we have considered the problem of obtaining exact solutions to the normal equations for certain non-orthogonal layouts. We conclude here with a brief description of a method for obtaining approximate solutions to the normal equations when data are incomplete. In illustrating the method we confine our attention to a layout with a single observation in each non-empty cell.

Let us suppose that we have an $I \times J$ layout with numerous empty cells, but at least one observation in each row and column. We set up the normal equations in accordance with the correct least squares method discussed in Chapter II, i.e., we set up normal equations utilizing the data that are present in the layout. The equations corresponding to α_i and β_j are

$$\begin{aligned} m\hat{\alpha}_i + m\hat{\xi} &= \sum_{j=1}^{m} x_{ip_j} - \sum_{j=1}^{m} \hat{\beta}_{p_j} \\ n\hat{\beta}_j + n\hat{\xi} &= \sum_{i=1}^{n} x_{q_i j} - \sum_{i=1}^{n} \hat{\alpha}_{q_i} \end{aligned} \tag{4.1}$$

respectively, where cells $ip_1, \cdots, ip_m$ are the m non-deficient cells in row i, and cells $q_1 j, \cdots, q_n j$ are the n non-deficient cells in column j. Notice that these equations can be written

$$\hat{\alpha}_i + \hat{\xi} = \frac{1}{m}\sum_{j=1}^{m} x_{ip_j} - \frac{1}{m}\sum_{j=1}^{m} \hat{\beta}_{p_j}$$

$$\hat{\beta}_j = \frac{1}{m}\sum_{i=1}^{n} x_{q_i j} - \frac{1}{n}\sum_{i=1}^{n} (\hat{\alpha}_{q_i} + \hat{\xi}) \ . \tag{4. 2}$$

For reasons of convenience we will work with the parameter α_i' ,

$$\alpha_i' = \alpha_i + \xi \ ,$$

rather than α_i itself. As our first approximation to the estimate $\hat{\alpha}_i'$ (we will denote it by $\alpha_i'(0)$) we take

$$\alpha_i'(0) = \frac{1}{m}\sum_{j=1}^{m} x_{ip_j} \ , \tag{4. 3}$$

i. e., just the average of observations present in the i^{th} row. Since every row contains at least one non-deficient cell, we may compute first approximations for all of the estimates $\hat{\alpha}_1', \cdots, \hat{\alpha}_I'$. In the second of Equations (4. 2) we replace each $\hat{\alpha}_{q_i}'$ by its first approximation $\alpha_{q_i}'(0)$, and obtain a first approximation to $\hat{\beta}_j$, which we shall denote by $\beta_j(1)$. We then replace the $\hat{\beta}_{p_j}$ in the first of Equations (4. 2) by their first approximations $\beta_{p_j}(1)$, and let the resulting equation define the second approximation $\alpha_i'(1)$ to $\hat{\alpha}_i'$. Repetition of these steps gives an iterative process for obtaining the successive approximations $\alpha'_i(1)$, $\beta_j(2)$, $a'_i(2)$, $\beta_j(3)$, $\alpha'_i(3), \cdots$.

One may obtain an approximation to $\hat{\xi}$ at any stage by imposing the condition

$$\sum_{i=1}^{I} \hat{\alpha}_i = 0 \ ,$$

since we then have

$$\frac{1}{I}\sum_{i=1}^{I} \alpha'_i = \frac{1}{I}\sum_{i=1}^{I} (\hat{\alpha}_i + \hat{\xi}) = \hat{\xi} \ .$$

Given an approximation to $\hat{\xi}$ one may calculate approximations for the individual $\hat{\alpha}_i$ from

$$\hat{\alpha}'_i = \hat{\alpha}_i + \hat{\xi} \ .$$

The details of this procedure are demonstrated by setting up the normal equations and performing the first iteration using the data in Figure 3.

		2	1
		1	2
1	5		6

Figure 3. Sample Data .

The normal equations are

$$
\begin{aligned}
\hat{\alpha}'_1 &= \tfrac{1}{2}(2+1) - \tfrac{1}{2}(\hat{\beta}_3 + \hat{\beta}_4) \\
\hat{\alpha}'_2 &= \tfrac{1}{2}(1+2) - \tfrac{1}{2}(\hat{\beta}_3 + \hat{\beta}_4) \\
\hat{\alpha}'_3 &= \tfrac{1}{3}(1+5+6) - \tfrac{1}{3}(\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_4) \\
\hat{\beta}_1 &= \tfrac{1}{1}(1) - \tfrac{1}{1}(\hat{\alpha}'_3) \\
\hat{\beta}_2 &= \tfrac{1}{1}(5) - \tfrac{1}{1}(\hat{\alpha}'_3) \\
\hat{\beta}_3 &= \tfrac{1}{2}(2+1) - \tfrac{1}{2}(\hat{\alpha}'_1 + \hat{\alpha}'_2) \\
\hat{\beta}_4 &= \tfrac{1}{3}(1+2+6) - \tfrac{1}{3}(\hat{\alpha}'_1 + \hat{\alpha}'_2 + \hat{\alpha}'_3) \ .
\end{aligned}
\tag{4.4}
$$

Now, according to (4.3)

$$
\begin{aligned}
\alpha'_1(0) &= \tfrac{3}{2} \\
\alpha'_2(0) &= \tfrac{3}{2} \\
\alpha'_3(0) &= 4 \ ,
\end{aligned}
$$

and then from the last four equations of (4.4)

$$
\begin{aligned}
\beta_1(1) &= 1 - 4 = -3 \\
\beta_2(1) &= 5 - 4 = 1 \\
\beta_3(1) &= \tfrac{3}{2} - \tfrac{3}{2} = 0 \\
\beta_4(1) &= 3 - \tfrac{7}{3} = \tfrac{2}{3} \doteq .67 \ ,
\end{aligned}
$$

and, to complete the first iteration, we use the first three of Equations (4. 4)

$$\alpha'_1(1) = \frac{3}{2} - \frac{1}{3} = \frac{7}{6} \doteq 1.17$$

$$\alpha'_2(1) = \frac{3}{2} - \frac{1}{3} = \frac{7}{6} \doteq 1.17$$

$$\alpha'_3(1) = 4 + \frac{4}{9} = \frac{40}{9} \doteq 4.44 \ .$$

From the last three approximations we obtain

$$\xi(1) = \frac{1}{3}\left(\frac{7}{6} + \frac{7}{6} + \frac{40}{9}\right) = \frac{61}{27} \doteq 2.26 \ ,$$

and hence

$$\alpha_1(1) = \frac{7}{6} - \frac{61}{27} = \frac{-59}{54} \doteq -1.09$$

$$\alpha_2(1) = \frac{7}{6} - \frac{61}{27} = \frac{-59}{54} \doteq -1.09$$

$$\alpha_3(1) = \frac{40}{9} - \frac{61}{27} \doteq 2.19 \ .$$

The least squares estimates of α_i, β_j, and ξ were calculated exactly using formulas (3. 36) and (3. 37) of Chapter III. They are

$$\hat{\xi} = 1.50 \qquad \hat{\beta}_1 = -3.50$$

$$\hat{\alpha}_1 = -1.50 \qquad \hat{\beta}_2 = .50$$

$$\hat{\alpha}_2 = -1.50 \qquad \hat{\beta}_3 = 1.50$$

$$\hat{\alpha}_3 = 3.00 \qquad \hat{\beta}_4 = 1.50 \ .$$

Comparison with the approximations reveals that the approximations are rather crude after only one iteration, but it is to be expected that the convergence of the approximations will be slow if the layout is small or if the observations are scanty or poorly interrelated. The utility of the process lies in its application to large layouts having a majority of cells empty.

BIBLIOGRAPHY

1. Cochran, William G. and Gertrude M. Cox. Experimental designs. New York, Wiley, 1950. 454 p.

2. Fisher, Ronald A. The design of experiments. 6th ed. New York, Hafner, 1951. 244 p.

3. Scheffé, Henry. The analysis of variance. New York, Wiley, 1959. 477 p.

4. Yates, F. The analysis of replicated experiments when the field results are incomplete. Empire Journal of Experimental Agriculture 1:129-142. 1933.