MAXIMUM LIKELIHOOD ESTIMATES

OF A STRUCTURAL RELATION

By

DON RICHARD BEER

Bachelor of Science
University of Texas
Austin, Texas
1960

Master of Arts
University of Texas
Austin, Texas
1963

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1972

MAXIMUM LIKELIHOOD ESTIMATES

OF A STRUCTURAL RELATION

Thesis Approved:

Robert D Morrison
Thesis Adviser

J. Leroy Folks

Richard L. Cummins

John E. Hoffman

D. Durham
Dean of the Graduate College

ACKNOWLEDGEMENTS

I wish to express my gratitude to Dr. Robert D. Morrison for serving as chairman of my advisory committee. I am also indebted to Dr. J. Leroy Folks for suggesting this thesis problem and for his guidance in the preparation of this thesis.

I, also, appreciate Professors John E. Hoffman and Richard L. Cummins serving on my advisory committee.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

PROBLEM DEFINITION

A statement of the general problem is presented first. The specific problem which is considered in this thesis is then defined and the results are discussed.

The General Problem

From theoretical considerations it is known that the mathematical variables $Y_1, Y_2, \ldots, Y_p$ and X are related by the equations

$$Y_i = \alpha_i + \beta_i X , \qquad i = 1, \ldots, p. \qquad (1.1)$$

However, Y_i and X can only be measured with error. Let the measurements be denoted by $y_i = Y_i + e_i$ and $x = X + e$ where the e_i's and e are the errors of measurement and $E(e_i) = E(e) = 0$. The parameters α_i and β_i are not known and are to be estimated. The relation given in equation 1.1 is classified into two categories. If the variables X and Y are fixed, the relation is said to be a functional relation. If X and Y are random, the relation is said to be a structural relation.

Assume that n sets of observations are recorded for x and y, say x_j and y_{ij}, $(i = 1, \ldots, p;\ j = 1, \ldots, n)$. If X_i is known, i.e. $\sigma_e^2 = 0$, then the classical least squares or weighted least squares gives the minimum variance, unbiased estimates of α_i and β_i. If X_i is not known and x_i is used to find a least squares estimate, it is known that the estimates of α_i and β_i are not only biased, they are not consistent. Methods other than least squares have been considered in an effort to obtain estimates which are at least consistent. One such method is maximum likelihood estimation.

Consider the problem where there is only one relation. Let $x = X + u$ and $y = Y + v$ where u and v are independent normal variables with zero means and variances equal to σ_u^2 and σ_v^2, respectively. If one knows 1) σ_v^2, 2) σ_u^2, 3) σ_v^2 and σ_u^2, or 4) $\lambda = \sigma_v^2/\sigma_u^2$, various consistent estimates of the parameters for the functional relation are available in the literature. Quite often it is known that if $X = 0$, then $Y_i = 0$. Hence, $\alpha_i = 0$. This added knowledge has a considerable effect on the types of consistent estimates which can be used.

For the single functional relation where α is known to be zero, a relative maximum likelihood estimate of the slope parameter has already been found. However, little work has been done on the single structural relation when α is known to be zero. Let $Y = \beta X$ and let X be a random variable with mean $\mu(\neq 0)$; then $E(X) = \mu$ and $E(Y) = \beta\mu$. Under regularity conditions $\hat{\beta} = \bar{y}/\bar{x}$ is a consistent estimate of β. Since $\text{cov}(x_i, y_i) = \beta\text{var}(X)$, information about β is also available in the sample deviations. It is desirable to combine the information in the sample means and covariance so as to obtain one consistent estimate of β.

The Specific Problem

If X and the measurement error have independent normal densities, the maximum likelihood equations do produce consistent estimates of all the parameters. In this paper relative maximum likelihood estimates for a single structural relation are found when there is the restriction that $\alpha = 0$ and $\lambda = \sigma_v^2/\sigma_u^2$ is known. The estimates are then derived for the case where both σ_v^2 and σ_u^2 are known. The results are then extended to p structural relations. Although results are obtained for specific problems, the method used to solve what have formerly been intractable equations is the major contribution of this paper.

The key to estimating all parameters is to first find the estimates of the slope parameters. All other estimates are readily expressed as functions of these slope estimates. For the single structural relation, $\hat{\beta}$ is found to be the root of a fourth degree polynomial. When there are p structural relations, the estimates of the slope parameters are functions of the roots of p fourth degree polynomials. The polynomials are derived for the situation where either $\lambda = \sigma_v^2/\sigma_u^2$ is known or both σ_v^2 and σ_u^2 are known. All of the derivations are done in Chapter III. In Chapter IV three problems are worked. The first is for a single structural relation and the second is for four structural relations. The last problem is an application in econometrics.

CHAPTER II

LITERATURE REVIEW

The following problem has a long history. Two mathematical variables, X and Y, are known to have the relation

$$Y = \alpha + \beta X \tag{2.1}$$

and we wish to estimate α and β. If X and Y are fixed, we call this a functional relation. If X and Y are random, this is called a structural relation. Unfortunately, we are not able to observe either X or Y; we observe only the values of the two random variables x and y, defined by $x = X + u$ and $y = Y + v$ where u and v are random variables. For our purposes we will assume that they are stochastically independent. Also, we will assume that $E(u) = E(v) = 0$, $\text{var}(u) = \sigma_u^2$, and $\text{var}(v) = \sigma_v^2$. In some way we wish to fit a straight line to the observed variables x and y, so as to obtain estimates of α and β, and, where necessary, estimates of σ_u^2 and σ_v^2.

Although this might appear to be a problem in regression analysis, in general it is not. It must be remembered that both variables are subject to error. It would be a regression model only if $\sigma_u^2 = 0$. However, if we were to find the regression of y on x, we would minimize the sum of squares of deviations about the fitted line vertical to the x-axis. Since it may be that $\sigma_u^2 = \sigma_v^2$, it would be just as reasonable to minimize the sum of squares of the deviations vertical to the y-axis. A compromise might be to minimize the sum of squares of deviations normal to the fitted line. This would not be unreasonable if it is known that $\sigma_u^2 = \sigma_v^2$, since we would not want to favor one direction over another. This procedure was suggested as far back as 1878.

Early History

C. H. Kummel (16) was the first person to find the best fitting line when the deviations were measured normal to the fitted line. Unfortunately, he published his results in an obscure journal called The Analyst, so his solution went unnoticed. Although his approach was purely mathematical in nature, he effectively assumed that $\sigma_u^2 = \sigma_v^2$.

Not knowing of this result, K. Pearson (21) tackled the problem of finding the best fitting lines and planes. Kummel's solution is a special case of Pearson's results. Pearson's approach also was not statistical in nature. Consequently, his solution was not considered to be invariant under a linear transformation. In a mathematical sense this was true, for he also assumed that the variances of the errors did not change under a linear transformation. This latter fact was later pointed out by D. V. Lindley (17). So, at the time of his publication, his results came under some criticism.

R. Frisch (8), in 1934, developed a general theory of regression analysis which employed new methods, but was not based on probability concepts.

Three years later a number of events happened. T. Koopmans (15) combined Frisch's theory with the classical one in a new general theory based on probability concepts. At the same time B. M. Dent (6) formulated the problem as one of maximum likelihood estimation, and, objecting to the requirement that $\lambda = \sigma_v^2/\sigma_u^2$ be known, she solved the maximum likelihood equations for the functional relation when the joint density is normal and nothing is known about σ_u^2 or σ_v^2. Her estimate of β was $\hat{\beta} = \pm\sqrt{\hat{\sigma}_v^2/\hat{\sigma}_u^2}$. It is consistent only if $\beta = \sqrt{\lambda}$. This estimate was referred to as the maximum likelihood estimate until 1969 when M. E. Solari (24) proved that it is a saddle point, not a maximum. In fact, the necessary and sufficient conditions for the existence of the maximum likelihood estimate is still an open question.

Also in 1937, C. F. Roos (23) attacked the problem from another angle in an attempt to make the estimates invariant under linear transformations. He showed that the function $U = \sum_k Q_k f(ax + by + c)$ must be minimum in order that the fitted line $ax + by + c$ is invariant under the transformation. Here, Q_k is a weighting factor and f is an arbitrary function. He chose $f = (ax + by + c)^2$ and assumed that λ was known. His estimate of β was

$$\hat{\beta} = \frac{S_{yy} - \sqrt{\lambda}\, S_{xy}}{S_{xy} - \sqrt{\lambda}\, S_{xx}} \tag{2.2}$$

where $S_{yy} = \sum_i (y_i - \bar{y})^2$, $S_{xx} = \sum_i (x_i - \bar{x})^2$, and $S_{xy} = \sum_i (y_i - \bar{y})(x_i - \bar{x})$. In general, this estimate is not consistent. There is also the problem that f is not unique and his choice was arbitrary.

Partly because of the fact that these estimates were not consistent, the emphasis was changed from trying to fit lines and, from them, estimate the slope parameter, to trying to obtain consistent estimates of the slope parameter. Five methods have been widely investigated which give consistent estimates of β. They are called grouping, instrumental variables, analysis of variance estimates, cumulants, and maximum likelihood estimates.

History of Consistent Estimates

Grouping

The method of grouping was introduced by Abraham Wald (26) in 1940. He divided the sets of observations into two groups using some criterion that is independent of the data. For N observations the first group consisted of the variables X_1 through X_m where $m = N/2$; the second group consisted of the remaining variables. He defined

$$\overline{a_1} = \frac{X_1 + X_2 + \dots + X_m - X_{m+1} - \dots - X_N}{N}$$

and

$$\overline{a_2} = \frac{Y_1 + Y_2 + \dots + Y_m - Y_{m+1} - \dots - Y_N}{N}.$$

He then defined

$$a_1 = \frac{x_1 + x_2 + \dots + x_m - x_{m+1} - \dots - x_N}{N}$$

and

$$a_2 = \frac{y_1 + y_2 + \dots + y_m - y_{m+1} - \dots - y_N}{N}.$$

Then he considered the estimate $\hat{\beta} = a_2/a_1$. He showed that $\hat{\beta}$ is a consistent estimate of β if $\lim \inf |a_1| > 0$. He pointed out that if the variables have a normal density, then the condition on the limit inferior is not satisfied and no consistent estimate of β exists unless more information is known about the parameters.

In practice he recommended that even though the grouping often has to be based on the data, acceptable estimates of β can be obtained if we order the data and divide the data into three groups: 1) those below some interval (a,b), 2) those within the interval, and 3) those above the interval. The first and third groups can be used to obtain an estimate of β in a form similar to the two group estimate.

In 1950 J. Neyman and E. Scott (20) proved that there have to be intervals about a and b such that the probability of a value of x falling in either interval is zero in order to satisfy the condition on the limit inferior. So this method is not applicable when u and v are normally distributed.

In 1949 Bartlett (3) extended Wald's two groups to three groups, and using the line joining the center of gravity of the two end groups, he showed that the asymptotic variance of the new estimate of β was approximately 15% less than that for the two group estimate. However, using sampling experiments, A. Madansky reached the opposite conclusion.

A method similar to the above, where multiple observations are taken for each value of X_i, was presented by G. Housner and J. Brennan (11) in 1948. Their estimate is of the form

$$\hat{\beta} = \frac{\sum_i \sum_j N_i N_j (\overline{y}_i - \overline{y}_j)}{\sum_i \sum_j N_i N_j (\overline{x}_i - \overline{x}_j)} \tag{2.3}$$

where N_i is the number of observations of the i^{th} replicate.

When the variates X, u, and v have independent normal densities, all of the above estimates have the disadvantage of not being consistent. They also have the important property that they are not functions of the sufficient statistics. Hence, their relative efficiency may suffer.

Instrumental Variables

We will now introduce an estimate which is a function of the sufficient statistics. It is due to R. C. Geary (10). Assume that we have the supplementary variable z, such that $z = f(X) + w$ where w has an independent normal density and X is fixed. Thus, z is related to x and y only through X; and x, y, and z have a joint normal density. Consider the linear combinations of x_i and y_i, $L_1 = \sum_i z_i x_i$ and $L_2 = \sum_i z_i y_i$; then

$$\text{plim } L_1/N = \text{plim } \frac{\sum_i X_i f(X_i)}{N}$$

and

$$\text{plim } L_2/N = \beta \text{ plim } \frac{\sum_i X_i f(X_i)}{N}$$

where plim is the probability limit as N increases without bound. Therefore, if the limits exist and do not equal zero, then plim $L_2/L_1 = \beta$. Notice that L_1 and L_2 are functions of at least part of the sufficient statistics.

This method has the disadvantage that if more than one instrumental variable is available, the problem of combining such estimates still exists. However, if all of the instrumental variables are of about the same magnitude, one can take their average, thus reducing the variance while maintaining approximately the same expected value.

The error w in z does not have to be normally distributed for the estimate to be consistent. Let half of the values of z be +1 and the other half be -1; the resulting estimate is Wald's grouping estimate. The Housner and Brennan estimate is also a special case of the instrumental variable estimate.

Cumulants

In 1942 R. C. Geary (9) suggested the use of product cumulants to obtain estimates of the relation parameters. He pointed out that, under regularity conditions, the product cumulant of x is the same as the product cumulant of X. He showed that if $Y = \beta X$, then the r^{th} cumulant for Y was equal to β^r times the r^{th} cumulant of X. Hence, the r^{th} cumulant of y is equal to β^r times the r^{th} cumulant of x. If consistent estimates of the cumulants (of order ≥ 3) could be found, then β could be estimated by the r^{th} root of the ratio of the cumulant estimates. Unfortunately, for the normal density the cumulants of order greater than or equal to three are all zero. Hence, his method is not always appropriate.

Analysis of Variance Estimates

J. W. Tukey (26) has indicated how to obtain consistent estimates of the relation parameters, when there is replication of the unobserved values, by the use of analysis of variance tables. However, M. Dorff and J. Gurland (7) have shown that the replicated values can be put to better use by using them to estimate λ and then using the estimate given in equation 2.3.

Maximum Likelihood Estimates

For the bivariate normal distribution we can be certain that the estimate of the slope parameter is a function of the sufficient statistics if we use the maximum likelihood estimates (MLE) for those cases where the joint density uniquely determines the relation parameters. D. V. Lindley (17) was the first to find the relative maximum likelihood estimates for this kind of problem. He considered the structural and functional relations where all of the random variables have independent normal densities and either λ or σ_v^2 is known. His results are

$$\hat{\beta} = \frac{S_{yy} - \lambda S_{xx} + \sqrt{(S_{yy} - \lambda S_{xx})^2 + 4S_{xy}^2}}{2S_{xy}}, \quad \lambda \text{ known}, \qquad (2.4)$$

and

$$\hat{\beta} = (S_{yy} - \sigma_v^2)/S_{xy}, \qquad \sigma_v^2 \text{ known}. \quad (2.5)$$

In all cases he assumed that α was not known.

It should be noted at this point that the likelihood equations may not yield the global maximum likelihood estimates, Dent's estimate being such an example. However, if one showed that the matrix of second order derivatives, evaluated using the estimated parameters, is negative definite, he would verify that the estimates are, indeed, relative maximum likelihood estimates.

In 1967 V. D. Barnett (2) found the relative maximum likelihood estimates when both σ_v^2 and σ_u^2 are known, the overidentified case. His estimate of the slope parameter is a root of the fourth degree polynomial

$$S_{yy}S_{xy}\beta^4 - (S_{yy}^2 - \lambda S_{xx}S_{yy} - 2\lambda S_{xy}^2)\beta^3 - 3\lambda(S_{yy}S_{xy} - \lambda S_{xy}S_{xx})\beta^2$$
$$+ \lambda^2(\lambda S_{xx}^2 - S_{xx}S_{yy} - 2S_{xy}^2) - \lambda^3 S_{xx}S_{xy} = 0. \qquad (2.6)$$

The solution is the same as in equation 2.4 and is true for the functional and structural relation.

In 1956 M. A. Creasy (5) added the result that if the relation is functional, all the densities are independent normal, $\lambda = 1$, and $\alpha = 0$, then the relative maximum likelihood estimate of β is

$$\hat{\beta} = \frac{S^*_{yy} - S^*_{xx} + \sqrt{(S^*_{yy} - S^*_{xx})^2 + 4S^{*2}_{xy}}}{2S^*_{xy}} \tag{2.7}$$

where S^*_{ij} are the sum of squares and crossproducts measured about zero, not the mean. It will be shown in the first example in Chapter IV that this estimate has a very interesting property.

He also showed that if $\alpha = 0$ and no information is available on the error variances, then $\hat{\beta} = \bar{y}/\bar{x}$ is a relative maximum likelihood estimate. If λ is known, he stated that the problem is intractable. This is the problem that is solved in this thesis.

Problems Encountered in Estimation

There are various problems associated with the estimation of β. An article of great importance to this subject was published by D. V. Lindley in 1947. He considered the question of regressing y on x when the relation $Y = \alpha + \beta X$ is true. He defined y to have a linear regression on x if, for any values of x, y is distributed about a mean which is a linear function of x. He found the necessary and sufficient condition to be that the cumulative function of x is a multiple of the cumulative function of u. His most important result, however, is that, even if the regression is linear, the parameters are _not_ the same as those in the original relation unless $\sigma_u^2 = 0$. It is partly because of this result that he derived some relative maximum likelihood estimates in the same article.

Another property with which we are concerned is the identification of the parameters in the relation. We say that the parameters are identifiable if, given the joint distribution of the observed variables, there is one and only one set of parameters which will produce that distribution. Assume the structural relation given in equation 2.1 is true and the variables X, u, and v have independent normal densities. Then both sets of values of the parameters given in Table I produce the same mean and covariance matrix, and, hence, the same multivariate normal distribution. So, if no restrictions are put on the parameters, they are not identifiable only if the joint density is multivariate normal. In all other cases they are identifiable.

TABLE I

A TABLE OF UNIDENTIFIABLE PARAMETERS

α	μ	β	σ_X^2	σ_u^2	σ_v^2
0	1	1	2	4	4
-1	1	2	1	5	2

Neyman (19) proved the existence of consistent estimates of β for a structural relation. In 1956 J. Kiefer and J. Wolfowitz (14) showed that for the structural relation, if the parameters are identifiable, then the relative maximum likelihood estimate of β is consistent under regularity conditions. In 1950 J. Berkson (4) pointed out that if x is controlled, that is, predetermined and repeatable, then standard regression techniques are applicable, a most important result.

Econometrics

Models similar to the one that has been discussed so far have wide application in the field of econometrics. Econometrics deals with the investigation of the mathematical and statistical relations between economic variables. For example, it may be postulated that there is a linear relation between the economic variables Y_1 and Y_2, such that $Y_2 = \alpha + \beta Y_1 + w$ where w is a disturbance term and Y_1 and Y_2 are observed values. In general, Y_1 and Y_2 may be correlated with w. This model can be considered as equivalent to the following one. There is a linear mathematical relation between the unobserved variables X_1 and X_2, such that $X_2 = \alpha + \beta X_1$, but the observed values of the X's are $Y_1 = X_1 + u$ and $Y_2 = X_2 + v$ where u and v are random disturbances of the observed values around the unobserved values. This can be interpreted as meaning that we <u>should</u> have the relation $Y_2 = \alpha + \beta Y_1$, but, because of the random nature of the process, we actually observe $Y_2 = \alpha + \beta Y_1 + (v - \beta u)$, so that $(v - \beta u)$ is the disturbance term and it is correlated with Y_1 and Y_2.

The economics models are further complicated by the fact that there is usually more than one relation and they involve more than two variables. The variables are divided into two categories. Endogenous variables are those measured with error. Exogenous variables are those measured without error. Various methods of estimating the parameters in the relations have been proposed by economists. If only one variable is endogenous, least squares estimation may be appropriate. But when many variables are present, the problem is usually suboptimized in some manner. The methods that are most frequently considered by econometricians are indirect least squares, two stage least squares, three stage least squares, limited information maximum likelihood estimation (LIML), and full information maximum likelihood estimation (FIMLE).

If some of the variables are exogenous, it is often possible to rearrange the equations so that one can find the regression of one or more of the endogenous variable on some of the exogenous variables. This will give unbiased estimates of functions of the parameters. These estimates can then be combined to give consistent estimates of all the parameters. This method is called indirect least squares and is applicable only in special circumstances.

A second use of exogenous variables is in the method of two stage least squares. Part of the trouble encountered in trying to estimate α and β is the fact that Y_1 and w are correlated. If Y_1 has a linear relation with an exogenous variable, say $Y_1 = a + bZ + e$, then it can be regressed against the variable Z to get an estimate of the expected value of Y_1, namely $\widehat{Y_1} = \hat{a} + \hat{b}Z$. If this is used in the relation $Y_2 = \alpha + \beta Y_1 + w$, with Y_1 replaced by $\widehat{Y_1}$, then $\widehat{Y_1}$, which varies only with Z, is independent of w. So least squares can be used on this new equation to obtain unbiased estimates of the remaining parameters. Three stage least squares is an extension of this basic idea.

The limited information maximum likelihood estimators are found by maximizing the likelihood function on a subset of the entire set of equations. If one considers the normal density, this method is equivalent, say, to using only the sample deviations to estimate a parameter when information on the parameter is also available in the distribution of the sample means. All of the above methods of estimation are suboptimal procedures.

It would seem desirable to incorporate all of the information available in the sample into the estimates. By using the full information maximum likelihood estimation procedure we are assured of doing this, since the estimates will be functions of the sufficient statistics. But the FIML estimates are, in general, difficult to evaluate since the differential equations which produce them are nonlinear and rather awkward. However, when they can be found they are desirable, since, under general conditions, they are consistent and asymptotically efficient, and they are in some sense optimal.

One object of this paper is to develop a procedure by which the FIMLE equations can be written in a form that is amenable to hand calculations, even for multiple relations. The results of this thesis indicate that the complexity of the problem of finding the FIMLE is dependent more on the form of the relations than on the number of relations.

CHAPTER III

DERIVATION OF THE ESTIMATES

A Single Structural Relation

Let there be the following structural relation between the two random variables X_1 and X_2:

$$X_2 = bX_1 \tag{3.1}$$

where b is unknown. To ease the problem of typing, the Greek letters that were used in the first two chapters will be changed to Roman letters. Thus, in equation 3.1, the letter b takes the place of β. The problem, then, is stated as follows.

Suppose that X_1 and X_2 can only be observed with error and let these observed values be $Y_1 = X_1 + e_1$ and $Y_2 = X_2 + e_2$. Also, let $X_1 \sim NID(m,v^*)$, $e_1 \sim NID(0,v_1)$, and $e_2 \sim NID(0,v_2)$. Taking pairs of observations, $Y_i = (Y_{1i}\ Y_{2i})$, we find

$$Y_i \sim MVN\left[\begin{pmatrix} m \\ bm \end{pmatrix}, \begin{pmatrix} v^* + v & bv^* \\ bv^* & b^2v^* + v \end{pmatrix}\right]$$

Reiersol has shown that all parameters are identifiable in the above distribution if there exists a relation between v_1 and v_2 of the form $v_2 = \lambda v_1$ where λ is known. We shall assume $\lambda = 1$. There is no loss in generality by doing this since one can make the transformations $Y_2^* = Y_2/\sqrt{\lambda}$ and $b^* = b/\sqrt{\lambda}$. The relative maximum likelihood estimates in the original problem will then be known functions of the relative maximum likelihood estimates of the new problem.

At this point we wish to find the relative maximum likelihood estimates of v^*, v_1, m, and b. Define

$$B = \begin{pmatrix} 1 \\ b \end{pmatrix}, \qquad v_1 = v,$$

and

$$A = \begin{pmatrix} v^* + v & bv^* \\ bv^* & b^2v^* + v \end{pmatrix}$$

$$= v^*BB' + vI. \tag{3.2}$$

Then

$$Y_i \sim MVN(mB,\ v^*BB' + vI). \tag{3.3}$$

We also have $|A| = v(v^*B'B + v)$. To find the inverse of A, consider the following. Assume A^- is of the form $A^- = cBB' + dI$. Then

$$I = AA^- = (v^*BB' + vI)(cBB' + dI).$$

Equating coefficients of BB' and I, we get

$$cv^*B'B + dv^* + cv = 0$$

and

$$vd = 1.$$

So, $d = 1/v$ and $c = (-v^*/v)/(v^*B'B + v) = -v^*/|A|$. Hence,

$$A^- = \frac{-v^*}{|A|} BB' + vI. \tag{3.4}$$

Note that $B'B = 1 + b^2$. Let the elements of A and A^- be functions of a parameter, say t. Then

$$\frac{d(AA^-)}{dt} = \frac{dA}{dt}A^- + A\frac{dA^-}{dt} = \frac{dI}{dt} = \emptyset$$

where d/dt is defined throughout this thesis to be the partial derivative with respect to t. The above equation can be rearranged to give

$$\frac{d(A^-)}{dt} = -A^-\frac{dA}{dt}A^-. \tag{3.5}$$

From the preceding equations we can derive some preliminary results.

$$\begin{aligned}
&\frac{dA}{db} = \begin{pmatrix} 0 & v^* \\ v^* & 2bv^* \end{pmatrix} && \frac{d|A|}{db} = 2bvv^* \\
&\frac{dA}{dm} = \emptyset && \frac{d|A|}{dm} = 0 \\
&\frac{dA}{dv^*} = BB' && \frac{d|A|}{dv^*} = vB'B \\
&\frac{dA}{dv} = I && \frac{d|A|}{dv} = v^*B'B + 2v
\end{aligned} \tag{3.6}$$

The log likelihood function is

$$\begin{aligned}
f &= k + \tfrac{1}{2}N \ln |A^-| - \tfrac{1}{2}\mathrm{tr}\, FA^- \\
&= k + \tfrac{1}{2}N\ln |A^-| - \tfrac{1}{2}\mathrm{tr}\, DA^- - \tfrac{1}{2}N\mathrm{tr}\, EA^-
\end{aligned} \tag{3.7}$$

where

$$k = \text{a constant}$$

$$D = \sum_i (Y_i - \overline{Y})(Y_i - \overline{Y})'$$

$$E = (\overline{Y} - mB)(\overline{Y} - mB)'$$

$$S = D/N$$

and

$$F = D + NE.$$

So,

$$\frac{df}{dt} = -\tfrac{1}{2}N\frac{d\ln|A|}{dt} + \tfrac{1}{2}\text{tr } FA^{-}\frac{dA}{dt}A^{-} - \tfrac{1}{2}\text{tr}\frac{dF}{dt}A^{-}. \tag{3.8}$$

Let $t = v*$. Then $d(\ln|A|)/dv* = vB'B/|A|$ and $dF/dv* = \emptyset$. Also,

$$\frac{dA}{dv*}A^{-} = BB'\left(\frac{-v*}{|A|}BB' + \frac{1}{v}I\right)$$

$$= \left(\frac{-B'Bv* + B'Bv* + v}{|A|}\right)BB'$$

$$= \frac{v}{|A|}BB'. \tag{3.9}$$

So,

$$\text{tr } FA^{-}\frac{dA}{dv*}A^{-} = \text{tr } F\left(\frac{-v*}{|A|}BB' + \frac{1}{v}I\right)\frac{v}{|A|}BB''$$

$$= \text{tr } F\left(\frac{-v*vB'B}{|A|^2} + \frac{1}{|A|}I\right)BB'$$

$$= B'FB\frac{v^2}{|A|^2}. \tag{3.10}$$

Thus, from equation 3.8

$$\frac{-\tfrac{1}{2}NvB'B}{|A|} + \tfrac{1}{2}B'FB\frac{v^2}{|A|^2} = 0 \tag{3.11}$$

which implies

$$|A| = \frac{B'FB}{B'B}\frac{v}{N}. \tag{3.12}$$

From equation 3.8 we get

$$\frac{df}{dm} = -\tfrac{1}{2}N\text{ tr }\frac{dE}{dm}A^{-} = 0 \tag{3.13}$$

since A^{-} is not a function of m. Now define $z = \bar{Y} - mB$. Then $E = zz'$ and

$$\frac{dE}{dm} = \frac{dz}{dm} z' + z \frac{dz'}{dm} = -Bz' - zB'. \tag{3.14}$$

Thus,

$$\text{tr} \frac{dE}{dm} A^- = \text{tr } (-Bz' - zB') \left(\frac{-v*}{|A|} BB' + \frac{1}{v} I\right)$$

$$= \frac{2v*(B'B)(B'z)}{|A|} - \frac{2B'z}{v}$$

$$= 2(B'z) \left[\frac{B'Bv* - v*B'B - v}{|A|}\right]$$

$$= 2(B'z) \left(\frac{-v}{|A|}\right)$$

$$= 0. \tag{3.15}$$

So,

$$B'z = B'(\overline{Y} - mB)$$

$$= B'\overline{Y} - mB'B$$

$$= 0 \tag{3.16}$$

and

$$m = \frac{B'\overline{Y}}{B'B}. \tag{3.17}$$

If we substitute equation 3.16 into equation 3.12, we get

$$|A| = \frac{vB'(D + Nzz')B}{NB'B}$$

$$= v \frac{B'SB}{B'B}. \tag{3.18}$$

Taking the partial derivative of f with respect to v, we get

$$\frac{df}{dv} = -\tfrac{1}{2}N\text{tr} \frac{dA}{dv} A^- - \tfrac{1}{2}\text{tr } F \frac{dA^-}{dv} - \tfrac{1}{2}\text{tr} \frac{dF}{dv} A^-$$

$$\frac{df}{dv} = 0. \tag{3.19}$$

Now,

$$\text{tr } A^- = \text{tr}\left(\frac{-v^*}{|A|} BB' + \frac{1}{v} I\right)$$

$$= \frac{-v^*B'B + 2(v^*B'B + v)}{|A|}$$

$$= \frac{\text{tr } A}{|A|}. \tag{3.20}$$

Also, from equation 3.5

$$\frac{dA^-}{dv} = -A^- \frac{dA}{dv} A^- = -A^-IA^- = -A^-A^-. \tag{3.21}$$

Now,

$$\text{tr } F \frac{dA^-}{dv} = -\text{tr } FA^-A^-$$

$$= -\text{tr } A^-FA^-$$

$$= -\text{tr}\left(\frac{-v^*}{|A|} BB' + \frac{1}{v} I\right) F \left(\frac{-v^*}{|A|} BB' + \frac{1}{v} I\right)$$

$$= \frac{-v^*(B'B)(B'FB)}{|A|^2} + \frac{2v^*(B'FB)}{|A|v} - \frac{\text{tr}F}{v^2}$$

$$= -\frac{B'FB}{|A|^2}\left[v^{*2}B'B - 2v^*(v^*B'B + v)\right] - \frac{\text{tr}F}{v^2}$$

$$= -\frac{NB'B}{|A|v}\left[-v^*\text{tr } A\right] - \frac{\text{tr } F}{v^2}. \tag{3.22}$$

So,

$$\frac{df}{dv} = \tfrac{1}{2}\left[\frac{-N\text{tr } A}{|A|} - \frac{v^*(\text{tr } A)N(B'B)}{|A|v} + \frac{\text{tr } F}{v^2}\right] = 0. \tag{3.23}$$

But, from the first two terms

$$\frac{-Nv(\text{tr } A) - v^*N(\text{tr } A)B'B}{|A|v} = \frac{(\text{tr } A)N(-v-v^*B'B)v}{|A|v^2}$$

$$= \frac{-N(\text{tr } A)}{v^2}.$$

Hence,

$$\text{tr } A = \frac{\text{tr } F}{N}. \tag{3.24}$$

Now consider equations 3.18 and 3.24. We find

$$B'Bv^* + v = \frac{B'SB}{B'B} \tag{3.25}$$

and

$$B'Bv^* + 2v = \frac{\text{tr } F}{N}. \tag{3.26}$$

Rearranging terms we get

$$v = \frac{\text{tr } F}{N} - \frac{B'SB}{B'B} \tag{3.27}$$

and

$$v^* = \frac{2\frac{B'SB}{B'B} - \frac{\text{tr } F}{N}}{B'B}. \tag{3.28}$$

It is easy to see that b is the key to estimating all parameters since m, v, and v* are functions of b.

To find b we must solve

$$\frac{df}{db} = -\tfrac{1}{2}N\,\frac{d(\ln|A|)}{db} - \tfrac{1}{2}\text{tr } F\,\frac{dA^-}{db} - \tfrac{1}{2}\,\text{tr}\,\frac{dF}{db}\,A^- = 0. \tag{3.29}$$

From equations 3.6 we have

$$\frac{d(\ln|A|)}{db} = \frac{2bv^*v}{|A|}.$$

Define

$$J_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} ;$$

then

$$\frac{dBB'}{db} = \begin{pmatrix} 0 & 1 \\ 1 & 2b \end{pmatrix} = \frac{BB'J_1 + J_1BB'}{b} \tag{3.30}$$

since

$$BB'J_1 = \begin{pmatrix} 0 & b \\ 0 & b^2 \end{pmatrix}.$$

Also,

$$\frac{dA}{db} = \frac{d(v^*BB' + vI)}{db} = v^* \frac{dBB'}{db}. \tag{3.31}$$

Hence,

$$\text{tr } F \frac{dA^-}{db} = -\text{tr } FA^- \frac{dA}{db} A^-$$

$$= -v^*\text{tr } FA^- \frac{dBB'}{db} A^-$$

$$= \frac{-v^*}{b} \text{ tr } FA^-(BB'J_1 + J_1BB')A^-$$

$$= \frac{-v^*}{b} (B'J_1A^-FA^-B + B'A^-FA^-J_1B). \tag{3.32}$$

Remembering that $B'z = 0$, so $B'FB = B'DB$, we get

$$B'J_1A^-FA^-B = B'J_1 \left(\frac{-v^*}{|A|} BB' + \frac{1}{v} I\right) F \left(\frac{-v^*}{|A|} BB' + \frac{1}{v} I\right)B$$

$$= \frac{v^{*2}}{|A|^2} b^2(B'DB)(B'B) - \frac{v^*}{|A|v} b^2(B'DB)$$

$$- \frac{v^*}{|A|v} (B'J_1DB)(B'B) + \frac{B'J_1DB}{v^2}$$

$$= \frac{(B'DB)b}{|A|^2 v}\left[v^*v(B'B)v^* - v^*v(B'Bv^* + v)\right]$$

$$- \frac{B'J_1DB}{|A|v^2}\left[v^*vB'B - v(B'Bv^* + v)\right]$$

$$= \frac{N(B'B)b^2(-v^*)}{|A|} + \frac{B'J_1DB}{|A|} \qquad (3.33)$$

where

$$B'J_1DB = (1\ \ b)\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}d_{11} & d_{12}\\ d_{21} & d_{22}\end{pmatrix}\begin{pmatrix}1\\ b\end{pmatrix}$$

$$= b(d_{21} + bd_{22}).$$

In a similar manner

$$B'A^-FA^-J_1B = \frac{N(B'B)b^2(-v^*)}{|A|} + \frac{B'DJ_1B}{|A|} \qquad (3.34)$$

where

$$B'DJ_1B = b(d_{12} + bd_{22}).$$

So,

$$\text{tr } F\,\frac{dA^-}{db} = \frac{-v^*}{b}\left[\frac{2N(B'B)b^2(-v^*) + 2b(d_{12} + bd_{22})}{|A|}\right]. \qquad (3.35)$$

Hence, the first two terms of equation 3.29 give us

$$N\,\frac{d(\ln|A|)}{db} + \text{tr } F\,\frac{dA^-}{db} = \frac{2Nbv^*v + 2v^{*2}(B'B)bN - 2v^*(d_{12} + bd_{22})}{|A|}$$

$$= \frac{2Nbv^*(v + B'Bv^*)}{|A|} - \frac{2v^*(d_{12} + bd_{22})}{|A|}$$

$$= \frac{2Nbv^*}{v} - \frac{2v^*(d_{12} + bd_{22})}{|A|}$$

$$= \frac{2v^*}{v}\left[Nb - \frac{NB'B}{B'SB}(S_{12} + bS_{22})\right]. \qquad (3.36)$$

Now,

$$\frac{dF}{db} = \frac{d(D + Nzz')}{db} = N\frac{dzz'}{db}$$

$$= N\frac{d(\bar{Y}\bar{Y}' - mB\bar{Y}' - m\bar{Y}B' + m^2BB')}{db}$$

$$= N\left[-m\begin{pmatrix}0\\1\end{pmatrix}\bar{Y}' - m\bar{Y}\begin{pmatrix}0\\1\end{pmatrix}' + m^2\left(\frac{BB'J_1 + J_1BB'}{b}\right)\right].$$

So,

$$\text{tr}\,\frac{dF}{db}A^- = N\text{tr}\left[-m\begin{pmatrix}0\\1\end{pmatrix}\bar{Y}' - m\bar{Y}\begin{pmatrix}0\\1\end{pmatrix}' + m^2\left(\frac{BB'J_1 + J_1BB'}{b}\right)\right]\left[\frac{-v^*}{|A|}BB' + \frac{1}{v}I\right]$$

$$= \frac{-Nv^*}{|A|}\left[-2mb(B'\bar{Y}) + \frac{m^2 2(B'B)b^2}{b}\right] + \frac{N}{v}\left[-2m\bar{Y}_2 + \frac{m^2 2b^2}{b}\right]$$

$$= \frac{2N}{v}(m^2b - m\bar{Y}_2)$$

since $B'\bar{Y} = mB'B$. Thus,

$$\frac{df}{db} = \frac{-v^*}{v}\left[b - \frac{(B'B)}{B'SB}(S_{12} + bS_{22})\right] - \frac{1}{v}\left[\left(\frac{B'\bar{Y}}{B'B}\right)^2 b - \left(\frac{B'\bar{Y}}{B'B}\right)\bar{Y}_2\right] = 0. \qquad (3.37)$$

We also have the relations

$$z'z = (\bar{Y} - mB)'(\bar{Y} - mB) = \bar{Y}'\bar{Y} - 2mB'\bar{Y} + m^2B'B$$

$$= \bar{Y}'\bar{Y} - \frac{(B'\bar{Y})^2}{B'B} \qquad (3.38)$$

and

$$\text{tr}\,F = \text{tr}\,D + Nz'z. \qquad (3.39)$$

Rearranging equation 3.28 and making the proper substitutions, we get

$$B'Bv^* = 2\,\frac{B'SB}{B'B} - (S_{11} + S_{22})\left(\frac{B'B}{B'B}\right) - \left(\frac{(B'B)(\bar{Y}'\bar{Y}) - (B'\bar{Y})^2}{B'B}\right). \qquad (3.40)$$

Combining equations 3.40 and 3.37, and finding a common denominator for equation 3.37, we find

$$\frac{\left[2B'SB - B'B(S_{11} + S_{22}) - (B'B)(\bar{Y}'\bar{Y}) + (B'\bar{Y})^2\right]\left[b(B'SB) - B'B(S_{12}+S_{22})\right]}{v(B'B)^2B'SB} +$$

$$\frac{b(B'\bar{Y})^2(B'SB) - (B'\bar{Y})\bar{Y}_2(B'B)(B'SB)}{v(B'B)^2(B'SB)} = 0. \qquad (3.41)$$

The numerator of this equation is at most a fifth degree polynomial in b. We now write the numerator of equation 3.41 in the form

$$a_5b^5 + a_4b^4 + a_3b^3 + a_2b^2 + a_1b^1 + a_0 = 0, \qquad (3.42)$$

and define

$$S^*_{12} = S_{12} + \bar{Y}_1\bar{Y}_2$$

$$S^*_{11} = S_{11} + \bar{Y}_1^2$$

and

$$S^*_{22} = S_{22} + \bar{Y}_2^2. \qquad (3.43)$$

Then,

$$a_0 = S_{12}S^*_{22} - S^*_{12}S_{11}$$

$$a_1 = (S^*_{11} - S^*_{22})S_{11} - S^*_{22}(S_{11} - S_{22}) - 4S_{12}S^*_{12}$$

$$a_2 = 3S^*_{12}(S_{11} - S_{22}) + 3S_{12}(S^*_{11} - S^*_{22})$$

$$a_3 = (S^*_{11} - S^*_{22})S_{22} - (S_{11} - S_{22})S^*_{11} + 4S_{12}S^*_{12}$$

$$a_4 = S_{22}S^*_{12} - S_{12}S^*_{11}$$

$$a_5 = 0 \tag{3.44}$$

where the S^*_{ij} are the sums of squares and crossproducts of the deviations measured about zero, not about the mean.

The author has been unable to discover a solution for equation 3.42 which is in a simple form. However, for specific values of the coefficients, it can be solved by methods available in standard reference books. It should be noted that $a_2 = -3(a_0 + a_4)$.

To get an insight into the structure of this equation, let us look at the limiting values of the coefficients. Since the observations are from a normal distribution, it follows that

$$\text{plim } E(S_{ij}S_{kl}) = \text{plim } E(S_{ij})\text{plim } E(S_{kl}), \ (i,j) \neq (k,l).$$

So,

$$\begin{aligned} \text{plim } a_4 &= (b^2v^* + v)b(v^* + m^2) - bv^*(v^* + m^2 + v) \\ &= v^*(v^* + m^2)b(1 - b^2) + bvm^2. \end{aligned}$$

Similarly,

$$\text{plim } a_3 = -v^*(v^* + m^2)\left[(1 - b^2)^2 - 4b^2\right] + (1 - b^2)m^2v$$

$$\text{plim } a_2 = 6v^*(v^* + m^2)b(1 - b^2)$$

$$\text{plim } a_1 = v^*(v^* + m^2)\left[(1 - b^2)^2 - 4b^2\right] + (1 - b^2)m^2v$$

$$\text{plim } a_0 = v^*(v^* + m^2)b(1 - b^2) - bvm^2.$$

Since there are only four independent coefficients, they can be combined to get

$$\text{plim } (a_4 + a_0) = -2v^*(v^* + m^2)b(1 - b^2)$$

$$\text{plim } (a_1 - a_3) = 2v^*(v^* + m^2)\left[(1 - b^2)^2 - 4b^2\right]$$

$$\text{plim } (a_4 - a_0) = 2bvm^2$$

$$\text{plim } (a_1 + a_3) = 2(1 - b^2)vm^2. \qquad (3.45)$$

Since b is a consistent estimate of the slope parameter, it would appear that, in order to eliminate $v^*(v^* + m^2)$ and vm^2 from any solution of the equation, $(a_4 + a_0)$ and $(a_1 - a_3)$ would be grouped, as would $(a_4 - a_0)$ and $(a_1 + a_3)$.

To find the relative maximum likelihood estimates when v_1 and v_2 are known, let $v_2 = \lambda v_1$ and $v_1 = v$. Then, as before, we let $b^* = b/\sqrt{\lambda}$. The only difference between the former derivation and the present one is that df/dv is not needed. Rearrange equation 3.25 to get

$$v^* = \frac{B'SB - B'Bv}{(B'B)^2} \qquad (3.46)$$

The estimate of m is the same as in equation 3.17. In fact, equation 3.37 is still valid, since there was no use of equation 3.24 in its derivation. When equations 3.46 and 3.17 are substituted into equation 3.37 and the terms are rearranged, we once again obtain a fourth degree polynomial, but the new coefficients are

$$a_0 = S_{12}v - S^*_{12}S_{11}$$

$$a_1 = (S^*_{11} - S^*_{22})S_{11} - (S_{11} - S_{22})v - 2S_{12}S^*_{12}$$

$$a_2 = 2S_{12}(S^*_{11} - S^*_{22}) + S^*_{12}(S_{11} - S_{22})$$

$$a_3 = (S^*_{11} - S^*_{22})S_{22} - (S_{11} - S_{22})v + 2S_{12}S^*_{12}$$

$$a_4 = S_{22}S^*_{12} - S_{12}v. \qquad (3.47)$$

Equation 3.47 can be solved by methods available in standard reference books.

Multiple Structural Relations

The previous results will now be extended to the problem with p structural relations and $a_i = 0$. There are p structural relations of the form

$$X_i = b_i X_0, \quad i = 1, \ldots, p; \quad b_0 = 1.$$

Assume that each X_i is measured with error. Let $Y_i = X_i + e_i$ be the observed values where $e_i \sim NID(0,v)$. It is assumed that X_0 is also known to have a normal distribution independent of the measurement errors. Define $Y_j' = (Y_{0j}\ Y_{1j} \ldots Y_{pj})$ and $B' = (1\ b_1\ b_2 \ldots b_p)$. Then $Y_j \sim MVN(mB,\ v^*BB' + vI)$, where $X_0 \sim NID(m,v^*)$. As before, let $A = v^*BB' + vI$. However, the determinant of A is now changed.

Define $|A|^* = v(v^*B'B + v)$. It is immediately seen that the derivation of A^- can be performed as done in equation 3.6. Hence,

$$A^- = \frac{-v^*}{|A|^*} BB' + \frac{1}{v} I. \tag{3.48}$$

We will now prove that

$$|A| = v^p (v^*B'B + v). \tag{3.49}$$

The proof is by induction. It has already been shown that equation 3.49 is true when p = 1. Let A_p be the covariance matrix when there are p relations. Let $B^{*\prime} = (1\ b_1\ b_2 \ldots b_{p-1})$. Now assume that equation 3.49 is true for p − 1. Then, $|A_{p-1}| = v^{p-1}(v^*B^{*\prime}B^* + v)$. Also, A_p can be written as

$$A_p = \left[\begin{array}{c|c} A_{p-1} & v^*b_pB^* \\ \hline v^*b_pB^{*\prime} & v^*b_p^2 + v \end{array}\right].$$

So,

$$|A_p| = |A_{p-1}| \cdot \left|(b_p^2v^* + v) - v^{*2}b_p^2B^{*\prime}A_{p-1}^{-}B^*\right|$$

$$= |A_{p-1}| \cdot \left|(b_p^2v^* + v) - v^{*2}b_p^2B^{*\prime}\left(\frac{-v^*}{|A|^*}B^*B^{*\prime} + \frac{1}{v}I\right)B^*\right|$$

$$= \frac{|A_{p-1}|}{|A_{p-1}|^*}\left[(b_p^2v^* + v)v(v^*B^{*\prime}B^* + v) - v^{*2}b_p^2(B^{*\prime}B^*)v\right]$$

$$= v^p(v^*B'B + v).$$

Note that trace $A_p = v^*B'B + (p + 1)v$.

Some preliminary results follow.

$$\frac{d|A|}{dm} = 0 \qquad \frac{dA}{dm} = \emptyset$$

$$\frac{d|A|}{dv^*} = v^pB'B \qquad \frac{dA}{dv^*} = BB'$$

$$\frac{d|A|}{dv} = pv^{p-1}v^*B'B + (p + 1)v^p \qquad \frac{dA}{dv} = I$$

$$\frac{d|A|}{db_i} = 2b_iv^pv^* \qquad \frac{dA}{db_i} = (a_{ijk})$$

where (a_{ijk}) is a $(p + 1)$ by $(p + 1)$ matrix and

$$a_{ijk} = \begin{cases} b_jv^*, & j \neq k = i + 1 \\ b_kv^*, & k \neq j = i + 1 \\ 2b_iv^*, & j = k = i + 1 \\ 0, & \text{otherwise} \end{cases},$$

We proceed in the same way that we did in the first part of the chapter. We find

$$\frac{d(\ln|A|)}{dv^*} = \frac{v^p B'B}{|A|} = \frac{vB'B}{|A|^*},$$

$$\frac{dF}{dv^*} = \emptyset$$

and

$$\frac{dA}{dv^*} A^- = \frac{vBB'}{|A|^*}.$$

These are the same results that were derived previously except that $|A|$ is replaced with $|A|^*$. So, equations 3.10 and 3.12 are the same with $|A|$ replaced by $|A|^*$. Likewise, df/dm does not change form, so $B'z = 0$ and $m = B'\bar{Y}/B'B$. Equation 3.18 is rewritten

$$|A|^* = v\,\frac{B'SB}{B'B}. \tag{3.50}$$

There is a change in trace A^-, however, for

$$\text{tr } A^- = \frac{-v^*B'B + (p + 1)(v^*B'B + v)}{|A|^*}$$

$$= \frac{(p - 1)v^*B'B + \text{tr } A}{|A|^*}. \tag{3.51}$$

But we still have $dA^-/dv = -A^-A^-$, so

$$\text{tr } F\,\frac{dA^-}{dv} = \frac{-B'FB}{|A|^{*2}}\left[-v^{*2}B'B - 2v^*v\right] - \frac{\text{tr } F}{v^2}.$$

Hence,

$$\frac{df}{dv} = \tfrac{1}{2}\left[\frac{Npv^*B'B + (p + 1)v}{|A|^*} - \frac{NB'B}{|A|^*v}\,(v^{*2}B'B + 2v^*v) + \frac{\text{tr } F}{v^2}\right]$$

$$= \tfrac{1}{2}\left[\frac{-N(p + 2)v^*vB'B + v^{*2}(B'B)^2 + (p + 1)v^2}{|A|^*v} + \frac{\text{tr } F}{v^2}\right]$$

$$= \tfrac{1}{2}\left[-N\,\frac{(p + 1)v}{v^2} + \frac{B'Bv^*}{v^2} + \frac{\text{tr } F}{v^2}\right] = 0.$$

Thus,

$$\text{tr } A = \text{tr } F/N, \tag{3.52}$$

as before. Combine equations 3.50 and 3.52 to obtain

$$v^*B'B + (p + 1)v = \text{tr } F/N$$

$$v^*B'B + \qquad v = B'SB/B'B,$$

and

$$v = \frac{\text{tr } F/N - B'SB/B'B}{p}, \tag{3.53}$$

$$v^* = \frac{(p + 1)B'SB/B'B - \text{tr } F/N}{pB'B}. \tag{3.54}$$

Now, define J_i to be a (p + 1) dimensional square matrix with all elements zero except the (i + 1)st diagonal element, which is 1. Consequently,

$$\frac{dBB'}{db_i} = \frac{BB'J_i + J_iBB'}{b_i}.$$

The previous derivations hold for df/db_i, except that

$$B'J_iDB = B'DJ_iB = b_i \sum_{k=0}^{p} d_{ik}b_k \; .$$

So, we have p polynomials of the form

$$\frac{df}{db} = \frac{-v^*}{v}\left(b_i - \frac{B'B}{B'SB} \sum_{k=0}^{p} S_{ik}b_k\right) - \frac{1}{v}\left(m^2b_i - m\bar{Y}_i\right) = 0. \tag{3.55}$$

For computational purposes the equation will not be expanded as a power series, but will be left in the above form.

CHAPTER IV

APPLICATIONS

Three problems will be considered in this chapter. In the first a review will be made of a comparison of various estimates which was initially done by Albert Madansky in 1959. The relative maximum likelihood estimate will be found for his example and its asymptotic variance will be compared to those of the other estimates.

In the second problem it will be shown that a consistent estimate of v is possible even though all parameters are not identifiable. The full information relative maximum likelihood estimate for a Keynesian economic problem will be obtained in the last example.

Madansky's Comparison

In 1959 A. Madansky published an article in the Journal of the American Statistical Society in which he made a comparison of the estimates and the estimated standard deviation of the slope estimators which were then available. The example he used can be stated as follows. We wish to estimate the linear relation between the Brinell hardness and the yield strength of artillery shells. Shells were manufactured from two different heats of steel, a random sample of 25 shells manufactured from each of the two heats of steel was taken, and the Brinell hardness and yield strength of the 50 shells were measured. We have been told that v_1 and v_2 are approximately 50 and 7500, respec-

TABLE II

MADANSKY'S DATA

Low Heat		High Heat	
Yield Strength	Brinell Hardness	Yield Strength	Brinell Hardness
x	y	x	y
229	845	277	900
230	810	285	800
235	750	285	815
235	750	285	815
235	755	285	925
235	755	285	965
235	765	285	970
235	795	285	970
235	795	285	975
235	930	285	975
239	905	285	975
241	760	285	1,010
241	760	285	1,030
241	760	285	1,045
241	795	285	1,150
241	800	285	1,160
241	805	293	940
241	815	293	1,005
241	825	293	1,005
241	825	293	1,015
241	835	293	1,040
241	870	302	935
241	875	302	1,075
241	960	302	1,095
241	1,050	321	1,140

TABLE III

MADANSKY'S RESULTS

Method	Estimate of β	Standard Deviation of Estimate
Least squares y on x	3.288	.47
Least squares knowing σ_u^2	3.536	.41
Least squares knowing σ_v^2	2.738	.62
Least squares knowing $\lambda = \sigma_u^2/\sigma_v^2$	3.475	.49
Least squares knowing σ_u^2 and σ_v^2	3.112	.25
Grouping, $p_1 = p_2 = \frac{1}{2}$	3.204	.22
Grouping, $p_1 = p_2 = 17/50$	4.256	.60
Instrumental variable $Z_i = i$	3.969	.29
Components of variance 1	3.202	—
Components of variance 2	3.142	—
Components of variance 3	3.172	—
Components of variance 4	3.427	—
Cumulants--3rd order	5.660	3.1
Cumulants--4th order (1)	4.538	—
Cumulants--4th order (2)	3.326	—
$\bar{Y}_2/\bar{Y}_1$	3.4343	.04

tively. The values of the estimates and their estimated standard deviations are given in Table III. None of the first 15 estimates use the assumption that a = 0. Only the last estimate, $\hat{b} = \bar{Y}_2/\bar{Y}_1$, uses this assumption. But this assumption would appear to be valid since one would not expect any yield strength if the Brinell hardness was zero.

If this is considered to be a functional relation and the Y_2's are divided by $\sqrt{150}$ so as to make $\lambda = 1$, the relative maximum likelihood estimate of the transformed slope parameter is

$$\hat{b} = \frac{(S^*_{22} - S^*_{11}) + \sqrt{(S^*_{22} - S^*_{11})^2 + 4S^{*2}_{12}}}{2S^*_{12}}, \quad \lambda = 1, \qquad (4.1)$$

where S^*_{ij} measures the deviation about zero, not about the mean. Hence, $S^*_{11} = S_{11} + \bar{Y}_1^2$, $S^*_{22} = S_{22} + \bar{Y}_2^2$, and $S^*_{12} = S_{12} + \bar{Y}_1\bar{Y}_2$. So, equation 4.1 can be written

$$\hat{b} = \frac{(S_{22} - S_{11}) + (\bar{Y}_2^2 - \bar{Y}_1^2) + \sqrt{\left[(S_{22} - S_{11}) + (\bar{Y}_2^2 - \bar{Y}_1^2)\right]^2 + 4(S_{12} + Y_1Y_2)^2}}{2(S_{12} + \bar{Y}_1\bar{Y}_2)}.$$

From the data that was presented in his article, we have $S_{11} = 714.1936$, $S_{22} = 92.8736$, $S_{12} = 191.7462$, $\bar{Y}_1 = 263.92$, and $\bar{Y}_2 = 74.00694$. It can be shown that the ratio $(\bar{Y}_2^2 - \bar{Y}_1^2)/(S_{22} - S_{11})$ is greater than 100. Hence, $\bar{Y}_2^2 - \bar{Y}_1^2$ is the dominant term. If the deviation terms are, in fact, neglected in the expression, the estimate reduces to

$$\hat{b} = \frac{\bar{Y}_2^2 - \bar{Y}_1^2 + \sqrt{(\bar{Y}_2^2 - \bar{Y}_1^2)^2 + 4(\bar{Y}_1\bar{Y}_2)^2}}{2\bar{Y}_1\bar{Y}_2}$$

$$= \bar{Y}_2/\bar{Y}_1.$$

Upon substitution of the proper values in equation 4.1 and multiplication by $\sqrt{150}$, we find that $\hat{b} = 3.4346$. The estimated asymptotic standard deviation is also approximately .04, so there seems to be little to choose between the two estimates.

It is interesting now to consider the case where the relation is structural. Equation 3.37 can be rewritten as

$$v^*\left[b - \frac{B'B}{B'SB}(S_{12} + bS_{22})\right] + m^2\left[b - \frac{\bar{Y}_2}{m}\right] = 0. \tag{4.2}$$

Consider the coefficient of v^* in the above equation and equate it to zero. If we used this equation to estimate b, the estimate would be the same as in equation 4.1. This is the relative maximum likelihood estimate if only the deviations are considered. If we equated the coefficient of m^2 to zero and used this equation to estimate b, we would find $\hat{b} = \bar{Y}_2/m$ which is the relative maximum likelihood estimate of b if only the sample means are considered. Since we have two consistent estimates of b, we might consider combining them in some way, say by taking a weighted average. However, equation 4.2 indicates a possibly more interesting way to "combine estimates". Don't combine the estimates directly, but instead combine the equations which give rise to the estimates, so as to obtain new estimating equations. But why would we want to weight them proportional to v^* and m^2? Intuitively, if the mean was near zero and the deviation of X was large, (m^2/v^* small), we would want to increase the effect that deviations had on the estimate. Similarly, if the deviations were small and the mean large (m^2/v^* large) the effect of the sample means should be more pronounced. For an idea as to the value of the ratio m^2/v^* in this example, we again look at

TABLE IV

COEFFICIENTS OF EQUATION 3.42

$$a_0 = -13{,}018{,}504$$
$$a_1 = 27{,}689{,}940$$
$$a_2 = 74{,}038{,}392$$
$$a_3 = -22{,}575{,}280$$
$$a_4 = -11{,}660{,}952$$

TABLE V

SECOND ORDER DERIVATIVES

r	s	$d^2f/drds$
m	m	$-NB'Bv/\lvert A\rvert$
v*	v*	$-N(vB'B)^2/(2\lvert A\rvert^2)$
v	v	$\frac{-N}{2}\left[\frac{2}{\lvert A\rvert} - \frac{2\text{tr } A}{\lvert A\rvert^2} - \frac{2v^*(B'FB)(\text{tr } A)^2}{\lvert A\rvert^3} \frac{2v^*B'FB}{\lvert A\rvert^2} + \frac{2\text{tr } F}{v^3}\right]$
b	b	$\frac{-N}{2}\left[\frac{2v^*}{v} + \frac{2v^*(F_{12} + bF_{22})v^*v2b}{\lvert A\rvert^2} - \frac{2v^*d(z_1z_2 - bz_2^2)/db}{\lvert A\rvert}\right.$ $\left.\frac{-v^*}{\lvert A\rvert} -2m(\bar{Y}_1 + 2b\bar{Y}_2) + m^2(1 + 3b^2) + \frac{2m^2}{v}\right]$
b	v	$\frac{-N}{2}\left[\frac{-2bv^*}{v^2} + \frac{2v^*(F_{12} + bF_{22})\text{tr } A}{\lvert A\rvert 2} + \frac{2m\bar{Y}_2 - 2bm^2}{v^2}\right]$
b	m	$Nv(Y_2 - 2mb)/\lvert A\rvert$
b	v*	$\frac{-N}{2}\left[\frac{2b}{v} - \frac{2v^2(F_{12} + bF_{22}}{\lvert A\rvert^2}\right]$
m	v	0
m	v*	0
v	v*	$-Nv^2B'B/(2\lvert A\rvert^2)$

the ratio $R = (\bar{Y}_2^2 - \bar{Y}_1^2)/(S_{22} - S_{11})$ which is a consistent estimate of m^2/v^*. We find R = 100. This again indicates that we should put more emphasis on the means. Since the estimate of b using the deviations is 3.475 and the estimate of b using only the means is 3.4343, the true MLE should not differ much from 3.4343. Upon solving equation 4.2 we find $\hat{b} = 3.4345$, a difference of less that .01% from $\bar{Y}_2/\bar{Y}_1$.

The matrix of second order derivatives with the appropriate restrictions on the parameter estimates is given in Table V. Upon substitution of the estimated values of the parameters, we obtain

$$\text{INVERSE ASYMPTOTIC VARIANCE MATRIX} = \begin{bmatrix} -141{,}125 & -4.817 & -.3461 & -12.98 \\ -.4817 & -.0702 & 0 & 0 \\ -.3461 & 0 & -.000049 & -.000046 \\ -12.98 & 0 & -.000046 & -26.543 \end{bmatrix}$$

where the order of the parameters is b, m, v*, and v. The upper left corner of the asymptotic variance matrix, then, is $-.000090231/12.4815 = -7.22 \times 10^6$. Hence, the estimated asymptotic standard deviation of the estimate of b in the original problem is $\widehat{\text{Var}(\hat{b})} = \sqrt{150} \cdot \sqrt{7.22} \times 10^3 = .0324$, about the same as that for $\bar{Y}_2/\bar{Y}_1$. Since the variance matrix is negative definite, the solution is a relative maximum.

The Bottle Problem

The investigation of the topic considered in this paper was originally suggested by the following problem. In 1964 Dr. J. Leroy Folks was approached by a chemistry student and was asked for assistance in solving a particular statistics problem. The student had weighed four bottles on seven days. It was realized that there was error present in the observations not only as measurement error, but also as day-effect variation. The model was postulated as being additive in measurement error, but multiplicative in day-to-day variation. So the model could be written

$$Y_{ij} = b_i X_j + e_{ij}, \quad i = 1, \ldots, 4, \quad j = 1, \ldots, 7 \tag{4.3}$$

where Y_{ij} = observed weight

b_i = true weight of bottle i, i.e. the weight of the bottle in a vacuum

X_j = random multiplicative factor due to atmospheric conditions during the j^{th} day

e_{ij} = measurement error.

The student was not concerned with the true weight of the bottles. He only wanted an estimate of the measurement error standard deviation. After Dr. Folks presented this problem to the students in one of his classes, the author, at Dr. Folks suggestion, was able to furnish the student with a consistent estimate of $\text{var}(e_{ij})$. However, since the distribution of the estimator was not known, no indication of the precision of the estimate could be given. Since the asymptotic variance of the MLE is often available, it was decided to consider the MLE for this model. However, no information for this model was available in the literature. Hence, this paper.

Let equation 4.3 hold true; then $a_i = 0$ for all i. Since the bottles are similar in weight, we will assume that the measurement error variance is the same for each bottle. We will further assume that $X_j \sim \text{NID}(m, v^*)$ and $e_{ij} \sim \text{NID}(0, v)$. Then $Y_j \sim \text{MVN}(mB, v^*BB' + vI)$

where $B' = (b_1 b_2 b_3 b_4)$. Unfortunately, there exists a problem of identification here, for if we double the values of all the b's, and take half the values of m and $\sqrt{v^*}$, we would still have the same multivariate normal density for the Y_j's. However, v is not effected by this process and so is unique. To get around this problem, define $b_1 = 1$ and $\lambda = 1$. Then the parameter space is well identified. Since the b's are not sought in this problem, this procedure is permissible.

The data for the problem which will now be worked is presented in Table VI. It is not the original data, but was devised so as to be useful for this discussion. The original data, which has long since disappeared, was given to seven significant digits and only the last three varied in the observed values. This made m^2/v^* on the order of 10^7, so that the terms involving the mean were dominant. Hence, if $\bar{Y}_2/\bar{Y}_1$ were substituted into the polynomials, the values of the functions would be on the order of .0001, showing that $\bar{Y}_2/\bar{Y}_1$ gives a good approximation to the true MLE since $|f'| \simeq |-m^2/v| > 10^7$. In order to show the method used to solve the equations, new data was created.

The method used to derive the estimates from the equations will be a modification of the Newton-Ralphson method. The Newton-Ralphson procedure is an iterative process. Let $b_{i,k}$ be the k^{th} estimate of b_i. Then we create new and hopefully better estimates by the relation

$$b_{i,k+1} = b_{i,k} - \frac{f_i^*(B)}{f_i^{*\prime}(B)}.$$

For f_i^* we use the function associated with $df/db_i = 0$. In place of evaluating $f^{*\prime}$ at each iteration, however, we will use the fact that $-m^2/v$ is the dominant term in $f^{*\prime}$. As our initial estimates of b_i, we

TABLE VI

WEIGHT OF BOTTLES

Day	Bottle 1	2	3	4
1	12.7	15.3	17.8	14.0
2	12.5	15.0	17.5	13.7
3	10.8	13.0	15.2	11.9
4	9.8	11.8	13.8	10.8
5	10.6	12.7	14.8	11.6
6	9.4	11.3	13.2	10.3
7	11.1	13.3	15.5	12.2

TABLE VII

THE MEAN AND COVARIANCE MATRIX

$$\bar{Y} = \begin{bmatrix} 10.9857 \\ 13.2000 \\ 15.4000 \\ 12.0714 \end{bmatrix} \qquad S = \begin{bmatrix} 1.3355 & 1.6057 & 1.8514 & 1.4739 \\ 1.6057 & 1.9314 & 2.2271 & 1.7728 \\ 1.8514 & 2.2271 & 2.5686 & 2.0443 \\ 1.4739 & 1.7728 & 2.0443 & 1.6278 \end{bmatrix}$$

TABLE VIII

SAMPLE CALCUALTIONS

Original Estimates

$\hat{b}_1 = 1.20156$ $\hat{b}_2 = 1.40180$ $\hat{b}_3 = 1.09883$

$\hat{m} = 10.9857$

$$\widehat{\frac{m^2}{v^*}} = \frac{(13.2)^2 - (10.9857)^2}{1.9314 - 1.3355} \approx 92$$

$\widehat{B'B} = 5.616279$

$\widehat{B'SB} = 41.906871$

$\widehat{\frac{B'SB}{B'B}} = 7.4616$

$\frac{tr\ F}{N} = tr\ S = 7.463265$

$\sum_{k=0}^{3} \widehat{S_{2k}b_k} = 8.996416$

$\sum_{k=0}^{3} \widehat{S_{3k}b_k} = 10.374800$

$\sum_{k=0}^{3} \widehat{S_{3k}b_k} = 8.258320$

$\hat{v} = .00052867$ $\widehat{v^*} = 1.328486$

$\widehat{\frac{v^*}{v}} = 2512.8833$ $\widehat{\frac{m^2}{v}} = 228,281.55$

$\hat{f}_1 = 10.3581$ $\hat{f}_2 = -28.6199$ $\hat{f}_3 = 19.9430$

New Estimates

$\hat{b}_1 = 1.201605$ $\hat{b}_2 = 1.401675$ $\hat{b}_3 = 1.098915$

will use $b_i = \bar{Y}_i/\bar{Y}$. So, our initial estimates are given in Table VIII. In one iteration we find that $\hat{b}_i$ changes by less than .01% and $\hat{\sqrt{v}}$ = .0230.

An Economic Problem

A simple Keynesian problem in economics can be stated as follows. Let

C = consumption expenditure
Y = income
Z = nonconsumption expenditure
u = a stochastic disturbance term
t = time period;

then it is surmised that the following relations hold:

$$C_t = a + bY_t$$

and

$$Y_t = C_t + dZ_t$$

These equations can be rewritten in reduced form as

$$C_t = \frac{a}{1-b} + \frac{bd}{1-b} Z_t$$

and

$$Y_t = \frac{a}{1-b} + \frac{d}{1-b} Z_t.$$

It is assumed that C, Y, and Z are endogenous variables. Thus, we only observe $c = C + e_1$, $y = Y + e_2$, and $z = Z + e_3$ where the errors are independent normal with zero means and $\text{var}(e_1) = \text{var}(e_2) = 9\ \text{var}(e_3)$. Although this might not seem to fit the case where a is assumed to be zero, it can be transformed so that it does since the intercept terms are the same in the reduced equations. Let us make the following definitions:

$$a^* = a/(1 - b), \quad b_1 = d/(3 - 3b), \quad b_2 = bd/(3 - 3b),$$

$$c_t^* = c_t - a^*, \quad y_t^* = y_t - a^*, \text{ and } \quad z_t^* = 3z_t.$$

If we further define $Y_t^{*\prime} = (z_t^* \; y_t^* \; c_t^*)$ and $B' = (1 \; b_1 \; b_2)$, then $Y_t^* \sim MVN(mB, v^*BB' + vI)$ where $Z \sim NID(m, v^*)$. Then all of the equations used in Chapter III are appropriate with $\bar{Y}$ replaced by $\bar{Y}^*$. Consequently, a^* is still present in equation 3.37. To eliminate it we must set $df/da^* = 0$. Since only z (as used in equation 3.38) depends on a^*, we get from equation 3.8

$$\frac{df}{da^*} = -N\left[-2(\bar{y} - a^* - mb_1) - 2(\bar{c} - a^* - mb_2)\right] = 0. \tag{4.4}$$

We can rewrite equations 4.4 and 3.17 as

$$2a^* + (b_1 + b_2)m = \bar{y} + \bar{c} \tag{4.5}$$

and

$$(b_1 + b_2)a^* + (1 + b_1^2 + b_2^2)m = \bar{z}^* + b_1\bar{y} + b_2\bar{c}. \tag{4.6}$$

The procedure is as in the bottle problem except that $\hat{a}^*$ and $\hat{m}$ are calculated at the same time. For an initial estimate of a^*, we treat z^* as though it were an exogenous variable and regress y and c against it, using the same intercept parameter. So, our initial estimate is $\hat{a}^* = 5682$. Since the initial estimate of m^2/v^* is 2400, the sample deviations can be ignored in the initial estimates. The procedure is to use equations 4.5 and 4.6, and the estimates $\hat{b}_1 = (\bar{y} - \hat{a}^*)/\hat{m}$ and $\hat{b}_2 = (\bar{c} - \hat{a}^*)/\hat{m}$ to satisfy $(b_i - \bar{Y}_i/m) = 0$ in equation 4.2. The sequence of estimated values of b_1 and b_2 is given in Table XI.

TABLE IX

UNITED KINGDOM NATIONAL-INCOME DATA

(Thousands of Pounds in 1954 Prices)

Year	y_t	c_t	z_t
1948	13,895	10,706	3,165
1949	14,377	10,940	3,359
1950	14,843	11,250	3,470
1951	15,307	11,089	4,166
1952	15,360	11,023	4,357
1953	15,951	11,474	4,404
1954	16,680	12,023	4,376
1955	17,237	12,443	4,320
1956	17,547	12,548	4,408
1957	17,788	12,802	4,318
1958	17,699	13,096	3,894

TABLE X

TABULATED VALUES

$$\bar{Y} = \begin{vmatrix} \bar{z}^* \\ \bar{y}^* \\ \bar{c} \end{vmatrix} = \begin{vmatrix} 12{,}064.6 \\ 16{,}062.2 \\ 11{,}763.1 \end{vmatrix} \qquad S = \begin{vmatrix} 1{,}820{,}226 & 1{,}283{,}008 & 549{,}253 \\ 1{,}283{,}008 & 1{,}792{,}985 & 1{,}042{,}172 \\ 549{,}253 & 1{,}042{,}112 & 694{,}616 \end{vmatrix}$$

TABLE XI

SEQUENCE OF ESTIMATES

$\widehat{a^*}$	$\hat{b}_1$	$\hat{b}_2$
5682	.86041	.50406
2212.15	1.0861	.76022
2724.8	1.1007	.74561
2773.71	1.10128	.74499
2771.55	1.101309	.74437

$$\text{Initial } \widehat{m^2/v} = \frac{(16062.2 - 5682)^2 - (12064.6)^2}{1792985 - 1820226} \simeq 2400$$

$$\text{Initial } \hat{m} = \bar{z}^* = 12064.6$$

The last iteration, using the sample means and its covariance matrix, give $\hat{b}_1 = 1.1013089$ and $\hat{b}_2 = .744366$, accurate to three decimal places. Thus, $\hat{b} = \hat{b}_2/\hat{b}_1 = .6759$ and $\hat{a} = (1 - \hat{b})\hat{a}^* = 897.8$.

CHAPTER V

SUMMARY

The variables X_i, $(i = 0, \ldots, p)$, are related by the set of equations

$$X_i = b_i X_0, \qquad i = 1, \ldots, p;$$

and $X_0 \sim NID(m, v^*)$. However, all S_i are observed with error, the observed values being $Y_i = X_i + e_i$, where $e_i \sim NID(0, v)$. The relative maximum likelihood estimate of b_i is determined by

$$f^* = \frac{-v^*}{v}\left[b_i - \frac{B'B}{B'SB}\left(\sum_{k=0}^{p} S_{ik} b_k\right)\right] - \frac{m^2}{v}\left(b_i - \frac{\bar{Y}_i}{m}\right) = 0 \qquad (5.1)$$

where

$$m = B'\bar{Y}/B'B \qquad (5.2)$$

$$v = \frac{\text{tr } F/N - B'SB/B'B}{p} \qquad (5.3)$$

and

$$v^* = \frac{(p + 1)B'SB/B'B - \text{tr } F/N}{pB'B}. \qquad (5.4)$$

To solve for $\widehat{b}_i$, first obtain initial estimates of all b_i, then evaluate $\hat{m}$, $\hat{v}$, and $\widehat{v^*}$. Next, use these estimated values to find a new estimate of b_i by using a modified Newton-Ralphson iterative process. Repeat this procedure until the estimates of all the b_i become stable.

When there is only one relation and $\lambda = v_2/v_1 = 1$, equations 5.1, 5.2, 5.3, and 5.4 can be combined to give a fourth degree polynomial as the estimating equation for b. The coefficients are given in equation 3.44, page 27.

When v_2 and v_1 are both known and there is one relation, b is found from a fourth degree polynomial in b with the coefficients given in equation 3.47, page 29.

In the first example it is demonstrated that the asymptotic standard deviation may increase considerably if it is known that the intercept parameter is zero, but this information is not used.

If all variables are measured with error, so that, in the classical regression sense, there are no "independent" variables, there may be a problem of identification of some of the parameters. In the second example it is shown that a consistent estimate of v exists even though the other parameters are not identifiable.

A Keynesian economics problem, where intercept parameters are present, can be solved by a modification of the above technique, if it is known that the intercepts are the same, as is demonstrated in the third application.

BIBLIOGRAPHY

(1) Anderson, T. W. *An Introduction to Multivariate Statistical Analysis*. New York: John Wiley and Sons, 1958.

(2) Barnett, V. D. "A Note on Linear Structural Relationships When Both Residual Variances Are Known." *Journal of the Royal Statistical Society, Series B*, 11, (1967), 670-671.

(3) Bartlett, M. S. "Fitting a Straight Line When Both Variables are Subject to Error." *Biometrics*, 5, (1949), 207-212.

(4) Berkson, J. "Are There Two Regressions?" *Journal of the American Statistical Association*, 45, (1950), 164-180.

(5) Creasy, M. A. "Confidence Limits for the Gradient in the Linear Functional Relationship." *Journal of the Royal Statistical Society, Series B*, 18, (1956), 65-69.

(6) Dent, B. M. *Proceedings of the Physical Society*, 47, (1935), 92.

(7) Dorff, M. and Gurland, J. "Estimation of the Parameters of a Linear Functional Relationship." *Journal of the Royal Statistical Society, Series B*, 23, (1961), 160-170.

(8) Frisch, R. *Statistical Confluence Analysis*. Oslo: Institute of Economics, 1934.

(9) Geary, R. C. "Inherent Relations Between Random Variables." *Proceedings of the Royal Irish Academy*, 47, (1942), 63-76.

(10) Geary, R. C. "Relations Between Statistics: The General and the Sampling Problem When the Samples are Large." *Proceedings of the Royal Irish Academy*, 49, (1943), 177-196.

(11) Housner, G. W. and Brennan, J. F. "Estimation of Linear Trends." *Annals of Mathematical Statistics*, 19, (1948), 380-388.

(12) Johnston, J. *Econometric Methods*. New York: McGraw-Hill, 1963.

(13) Kendall, M. G. and Stuart, A. *The Advanced Theory of Statistics*, Volume II. London: Griffin, 1961.

(14) Kiefer, J. and Wolfowitz, J. "Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters." Annals of Mathematical Statistics, 27, (1956), 887-906.

(15) Koopmans, T. C. Linear Regression Analysis of Economic Time Series. Haarlem, Netherlands: DeErven F. Bohn, 1937.

(16) Kummel, C. H. "Reduction of Observed Equations Which Contain More than One Observed Quantity." The Analyst, 6, (1879), 97-105.

(17) Lindley, D. V. "Regression Lines and the Linear Functional Relationship." Journal of the Royal Statistical Society, Supplement 9, (1947), 219-244.

(18) Madansky, A. "The Fitting of Straight Lines When Both Variables are Subject to Error." Journal of the American Statistical Association, 54, (1959), 173-205.

(19) Neyman, J. "Existence of Consistent Estimate of the Directional Parameter in a Linear Structural Relation Between Two Variables." Annals of Mathematical Statistics, 22, (1951), 496-512.

(20) Neyman, J. and Scott, E. L. "On Certain Methods of Estimating the Linear Structural Relation." Annals of Mathematical Statistics, 22, (1951), 352-361.

(21) Pearson, K. "On Lines and Planes of Closest Fit to Systems of Points in Space." Philosophical Magazine, 2, (1901), 559-572.

(22) Reiersol, O. "Identifiability of a Linear Relation Between Variables Which Are Subject to Error." Econometrica, 18, (1950), 375-389.

(23) Roos, C. F. Metron, 13, (1937), 3.

(24) Solari, M. E. "The Maximum Likelihood Solution of the Problem of Estimating a Linear Functional Relation." Journal of the Royal Statistical Society, Series B, 13, (1969), 372-375.

(25) Tukey, J. W. "Components in Regression." Biometrics, 7, (1951), 33-70.

(26) Wald, A. "Fitting a Straight Line if Both Variables Are Subject to Error." Annals of Mathematical Statistics, 11, (1940), 284-300.

VITA

Don Richard Beer

Candidate for the Degree of

Doctor of Philosophy

Thesis: MAXIMUM LIKELIHOOD ESTIMATES OF A STRUCTURAL RELATION

Major Field: Statistics

Biographical:

Personal Data: Born in Fairfield, Iowa, January, 31, 1935, the son of Mr. and Mrs. Joe F. Beer.

Education: Graduated from McAllen High School, McAllen, Texas, in June, 1953; received the Bachelor of Science degree with a major in physics, from the University of Texas, Austin, Texas, in June, 1960; received the Master of Arts degree, with a major in mathematics, from the University of Texas, Austin, Texas, in June, 1963; completed the requirements for the Doctor of Philosophy degree in May, 1972.

Professional Experience: Graduate Teaching Assistant in the Department of Mathematics and Statistics, Oklahoma State University, Stillwater, Oklahoma, 1963-1967; Operations Analyst, General Dynamics/Fort Worth, Fort Worth, Texas, 1967-1968; Professor in the Industrial Engineering Department, University of Texas at Arlington, Arlington, Texas, 1968-1972.