EMPIRICAL INVESTIGATION OF THE LIKELIHOOD SURFACE IN LINEAR FUNCTIONAL AND STRUCTURAL RELATIONSHIPS WHEN BOTH VARIABLES ARE SUBJECT TO ERROR

MUSTAFA AKAR

Çukurova University Faculty of Agriculture Department of Animal Science

ABSTRACT

The linear functional and structural relationships are of great importance in the physical, biological, economical and agricultural scinences but the estimation of these relation is not always saticfactory.

We have shown, by an empirical examination, that when the errors are uncorrelated, the likelihood surface contains a saddle point and that the maximum likelihood method leads to this saddle point and not to a real maximum. We have thus confirmed the work done by Solari (1969) and shown that MAXİMUM LİKELİHOOD METHOD fails to solve for this problem.

INTRODUCTION

We consider the situation when two properties are linearly related and we wish to estimate this relationship from measurements of these properties on a selected set of samples covering the range of interest of these properties. The measurements of both properties are subject to random errors. These properties will be denoted by variables X and Y, but the errors in X and Y are considered to be independent.

When both X and Y are random variables, then the relation is called a structural relation by Kendall (1951 and 1952), but according to Lindley (1947), this relation is called a functional relation. However, Kendall called this relation a functional relationship when both X and Y are not random.

The problem, which has described above, has attracted much interest over a long period. For example, Lindley (1947), Madansky (1959), Sprent (1970), Lindley and El-Sayyad (1968), Solari (1969), Bartlett (1949), Berkson (1950), Tukey (1951), Lindley (1953), Kendall and Stuart (1970), Copas (1972) and Taylor (1973).

Theory

Suppose that there is a functional relationship between true values X and Y, and that observations, denoted by x and y, have errors which are not correlated. Under these situations the following assumptions are made [Lindley, 1947; Lindley ve El-Seyyad (1968); Kendall and Stuart, (1970)]:

a) n pairs of observations are

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n).$$

b) n pairs of true values are

$$(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n).$$

c) $\left. \begin{array}{l} x_i = X_i + \delta_i \\ y_i = Y_i + \varepsilon_i \end{array} \right\}$ $i = 1, 2, \ldots, n$

and

$E(\delta_i) = E(\varepsilon_i) = 0$, $Var(\delta_i) = \sigma^2_x$, $Var(\varepsilon_i) = \sigma^2_y$, for all i;

$Cov(\delta_i, \delta_j) = Cov(\varepsilon_i, \varepsilon_j) = 0$, $i \neq j$;

$Cov(\delta_i, \varepsilon_j) = 0$ for all $i \neq j$

$E(x_i) = X_i$, $E(y_i) = Y_i$, for all i

and

$x_i \sim N(X_i, \sigma^2_x)$, $y_i \sim N(Y_i, \sigma^2_y)$

d) The functional relationship between true values is

$$Y_i = \alpha + \beta X_i, \quad i = 1, 2, \ldots, n \tag{1}$$

Under the above assumptions we may write the log–likelihood function (Akar, 1975) as;

$$logL = - nlog\, 2\Pi - nlog\sigma_x - nlog\sigma_y - \frac{1}{2} \left[\frac{\Sigma(x_i - X_i)^2}{\sigma^2_x} + \right.$$

$$\left. \frac{\Sigma((y_i - \bar{y}) - \beta(x_i - \bar{x}))^2}{\sigma^2_y} \right] \tag{2}$$

We wish to estimate the $(n + 3)$ unknown parameters β, σ^2_x, σ^2_y and $X_1, X_2, \ldots, X_n$. For this reason, we have to obtain maximum–likelihood equations of (2) by differentiating w.r.t. (with respect to) β, σ_x, σ_y and X_i in turn and equating to zero, as follows:

i) $\dfrac{\partial \log L}{\partial \beta} = 0 \implies \Sigma\,(\hat{X}_i - \bar{x})\,((y_i - \bar{y}) - \hat{\beta}\,(\hat{X}_i - \bar{x})) = 0$ \hfill (3)

ii) $\dfrac{\partial \log L}{\partial \sigma^2{}_x} = 0 \implies n\hat{\sigma}^2{}_x = \Sigma\,(x_i - \hat{X}_i)^2$ \hfill (4)

iii) $\dfrac{\partial \log L}{\partial \sigma^2{}_y} = 0 \implies n\,\hat{\sigma}^2{}_y = \Sigma\,((y_i - \bar{y}) - \hat{\beta}\,(\hat{X}_i - \bar{x}))^2$ \hfill (5)

iv) $\dfrac{\partial \log L}{\partial X_i} = 0 \implies -\dfrac{(x_i - \hat{X}_i)}{\hat{\sigma}^2{}_x} -$

$\dfrac{\hat{\beta}\,((y_i - \bar{y})) - \hat{\beta}\,(\hat{X}_i - \bar{x}))}{\hat{\sigma}^2{}_y} = 0$ \hfill (6)

where a cirfumflex denotes the maximum–likelihood estimators. From (6), we obtain

$$\hat{X}_i = \frac{\hat{\beta}\,\hat{\sigma}^2{}_x\,(y_i - \bar{y}) + \hat{\sigma}^2{}_y x_i + \hat{\beta}^2\,\hat{\sigma}^2{}_x \bar{X}}{\hat{\sigma}^2{}_y + \hat{\beta}^2\,\hat{\sigma}^2{}_x}$$ \hfill (7)

Thus we obtain the result as

$$(x_i - \hat{X}_i) = \frac{\hat{\beta}\,\hat{\sigma}^2{}_x\,(\hat{\beta}\,(x_i - \bar{x}) - (y_i - \bar{y}))}{\hat{\sigma}^2{}_y + \hat{\beta}^2\,\hat{\sigma}^2{}_x}$$ \hfill (8)

and

$$(y_i - \bar{y}) - \hat{\beta}\,(\hat{X}_i - \bar{x}) = \frac{-\hat{\sigma}^2{}_y(\hat{\beta}(x_i - \bar{x})\ \ (y_i - \bar{y}))}{\hat{\sigma}^2{}_y + \hat{\beta}^2\hat{\sigma}^2{}_x}$$ \hfill (9)

Substituting (8) and (9) in (4) and (5), and then dividing (4) by (5), we obtain,

$$\hat{\beta}^2 = \frac{\hat{\sigma}^2{}_y}{\hat{\sigma}^2{}_x}$$ \hfill (10)

This is as a disturbing result since it implies that the maximum likelihood estimator of the slope is either minus or plus the square root of the variances ratio of the estimators of the variances of δ and ε.

In fact, the above result is due to Lindley (1947), but as he stated it will not generally hold between true values, so that the maximum likelihood method breaks down for this instance, and consequently, we can not estimate all three parameters β, σ_x and σ_y consistently.

However, Solari (1969) proved that non–existence of the maximum likelihood solution in this situation and demonstrated that (10) is a saddle point of the likelihood surface rather than a maximum. Futhermore, it might be useful here to mention the heuristic argument by Kendall and Stuart (1970). According to them, without prior knowledge of σ^2_x and σ^2_y *there is little* that can be said about the existence of a linear functional relationship; therefore, to make our problem definite we need only the eccentricity of the elipses, i.e. the ratio σ^2_y / σ^2_x.

Apart from these arguments it should become apparent, from the graphs which will be presented in the next section, why we need such an assumption that the ratio of the variances is known.

Let $\lambda = \sigma^2_y / \sigma^2_y$ and put it in (7), we obtain,

$$\widetilde{X}_i = \frac{\hat{\beta} (y_i - \bar{y}) + \lambda x_i + \bar{x} \hat{\beta}^2}{\lambda + \hat{\beta}^2} \tag{11}$$

If we substitute $\widetilde{X}_i$ from (11) into (3), we obtain

$$-\hat{\beta}^2 S_{xy} + \hat{\beta} (S_{yy} - \lambda S_{xx}) + \lambda S_{xx} = 0 \tag{12}$$

where,

$S_{yy} = \Sigma (y_i - \bar{y})^2$, $S_{xx} = \Sigma (x_i - \bar{x})^2$ and $S_{xy} = \Sigma (x_i - \bar{x}) (y_i - \bar{y})$.

The equation (12) is a quadratic in β, we can easily solve it, whence

$$\hat{\beta} = \varnothing \pm (\varnothing^2 + \lambda)^{1/2} \tag{13}$$

where

$$\varnothing = \frac{S_{yy} - \lambda S_{xx}}{2 S_{yy}}$$

The result (13) is due to Lindley (1947). Similarly, the maximum likelihood estimates of other parameters can be defined.

Empirical Investigation

The main purpose of this section is to show empirically why the maximum likelihood method fails to yield satisfactory estimators for a linear functional relationship when both variables are subject to normal independent error. Solari (1969), has already proved that an application of the method of maximum likelihood leads to a *saddle point* of the likelihood surface rather than a maximum; therefore it becomes apparent that our aim is to carry out a full examination of

the likelihood surface so that we may ve able to confirm her results and also see how satisfactory the assumptions which where made in the previous section, are in obtaining the maximum likelihood estimator of β. In order to do this different sets of graphs were drawn using the computer at Aberystwyth University and U.M.R.C.C. (University of Manchester Regional Computer Center).

Initially, we have supposed that there is a functional relationship between the true values X and Y as $Y = 8 + 1.6X$, and then, having chosen the true values of the X's to be between 10 and 80 inclusively, the random variables were generated by appliying the Box and Müller method, from the distribution $x_i \sim N (X_i, \sigma^2_x)$ and $y_i \sim N (Y_i, \sigma^2_y)$ for all i, where the values of σ_x and σ_y were 8 and 10, respectively.

After obtaining the data, the fitted regression lines of y on x and x on y are calculated and the results are, respectively, $\hat{y} = 14. 144 + 1. 39x$, with residual standart deviation $S_{(y)} = 1.33$ and $\hat{x} = - 2.61 + 0.61y$, with residual standart deviation $S_{(x)} = 8.77$. Then, to ensure that the true values of σ_x and σ_y were covered, we multiplied the standart errors by 1.5; we thus obtained the range of β, σ_x and σ_y for empirical investigation

$$1.39 \leq \beta \leq 1.65$$

$$0 < \sigma_x \leq 13.10$$

and $\quad\quad 0 < \sigma_y \leq 19.10$

We then proceeded to draw four sets of graphs to represent the likelihood surface as fallows:

(a) The First Set of Graphs

This set of graphs consists of σ_x against σ_y and $\ln\sigma_x$ against $\ln\sigma_y$ for a range of given values of β. To draw this set of graphs we used the following function,

$$L^* = -n\log\ \sigma_x - n\log\ \sigma_y - \frac{S}{2\ (\sigma^2_y + \beta\ \sigma^2_x)} \tag{14}$$

Where, $L^* = \log L + n\log 2\Pi$ and the sign of L^* is negative and $S = \Sigma\ (\beta\ (x_i - \bar{x}) - (y_i - \bar{y}))^2$.

The function (14) was obtained from (2) after making some changes on it. The equation (14) is non–linear form and so when we want to plot likelihood contours for either σ_x against σ_y or $\log \sigma_x$ against $\log\sigma_y$ for given values of β we first have to obtain the coordinates of σ_x and σ_y.

After obtaining the coordinates, this set of graphs was drawn. Two examples of this set of graphs were presented in Figure 1.

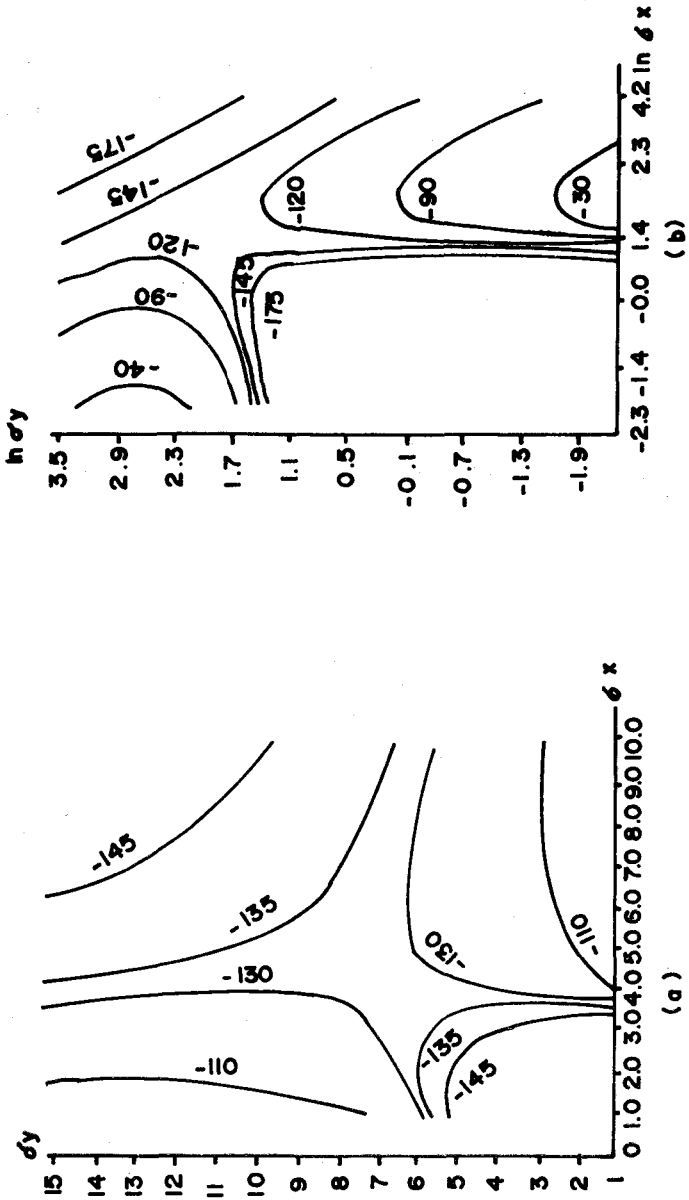Figure 1. Contours of L^* for, (a)σ_x against σ_y and (b)$\ln\sigma_x$ against $\ln\sigma_v$ when the slope is 1.6.

On examination of graphs in Figure 1, the following conclusions may be made.

First, passing from the bottom left hand corner to the top right hand corner would involve the crossing of a 'maximum' of the likelihood surface. On the other hand, passing from the bottom right hand corner to the left hand corner would involve the traversal of a 'minimum'. Therefore, we conclude that the maximum likelihood surface contains a saddle point. The so–called 'maximum likelihood' solution given by.

$$\beta = \pm \left[\frac{S_{yy}}{S_{xx}} \right]^{1/2}, \ 2n\sigma^2_x = S_{xx} - \frac{S_{xy}}{\beta} \ , \ 2n\sigma^2_y = S_{yy} - \beta S_{xx}$$

Where $S_{xx} = \Sigma(x_i - \bar{x})^2$; $S_{yy} = \Sigma \ (y_i - \bar{y})^2$ and $S_{xy} = \Sigma \ (x_i - \bar{x})$ $(y_i - \bar{y})$, was shown by Solari (1969) to be a saddle point rather than a maximum. In order to demonstrate that our saddle point is the same one which she predicted, we calculated the β, σ_x and σ_y as shown in below table.

Her notation	β	$\emptyset$	$\sqrt{\Psi}$
Our notation	β	σ_x	σ_x
By calculating	1.516*	4.325*	6.556*
Given	1.600	5.466	3.986

The result (*) in the above table coincides with our saddle point; consequently we have shown empirically that she is right.

Secondly, there is very litte difference between the graphs in Figure 1.

We notice that when σ_x and σ_y are both small the contours are more closely grouped together, indicating a steeper slope to the surface than that in places where σ_x and σ_y both take large values.

Thirdly, for a given value of either σ_x or σ_y we always get a maximum by drawing a line through the given point and perpendicular to either σ_x or σ_y respectively. This means that for a given value of either σ_x or σ_y we may estimate either σ_y or σ_x and β by the method of maximum likelihood.

Lastly, if we rotate a line passing through the origin between the σ_x and σ_y axes, then we always obtain a maximum along our line. This

implies that for any given ratio of variances, we could estimate β (slope) and the variances.

Consequently, we conclude that we cannot estimate all three parameters β, σ_x and σ_y consistently by applying the metod of maximum likelihood unless given either σ_x or σ_y or their ratio.

(b) The Second and Third Sets of Graphs

The second set of graphs consists of σ_x against β and $\log\sigma_x$ against β for a range of given values of σ_y but the third set of graphs consists of σ_y against β and $\log\sigma_y$ against β for a range of given values of σ_x.

The following formula was used for drawing these two sets of graphs:

$$\psi \, \beta^2 + 2 \, Q \, \beta + \varnothing = 0 \tag{15}$$

This equation was obtained from the function (14). It is quadratic in β and so the roots are;

$$\beta_{2,2} \;=\; \frac{-Q \pm (Q^2 - \psi \, \varnothing \,)^{1/2}}{2}$$

where $\psi \;=\; 2\sigma_x^2 \, (L^+ - n\log \, \sigma_x - n\log\sigma_y) - S_{xx}$

$Q \;=\; S_{xy}$

$\varnothing \;=\; 2\sigma_y^2 \, (L^+ - n\log \, \sigma_x - n\log \, \sigma_y) - S_{yy}$

$L^+ \;=\; -\, L^*.$

One example of the second set of graphs was given in Figure 2.

On examination of graph in Figure 2, there is an evidence of a saddle point whose approximate coordinates and height are as given (+) in Table 1, which presents a summary of the main features of the surface.

If we examine L^* along the line paralel to the β–axis which passes through the saddle point, we find that as we increase or decrease β from saddle point, the surface goes down–hill for both cases and the saddle point is therefore a 'maximum' in the β direction. However, the saddle point is a 'minimum' in the $\ln\sigma_x$ direction.

One example of the third set of graphs was given in Figure 3.

On examination of the graph in Figure 3, there is evidence of a saddle point whose approximate coordinates and height are as given (+) in Table 2, which presents a summary of the main features of the surface.
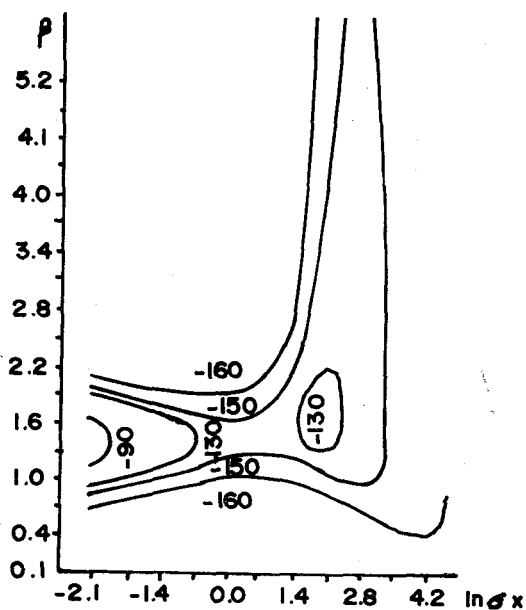
Figure 2. Contours of L* for $\ln\sigma_x$ against slope when σ_y is 5.

Table 1. A summary of the main features of the L*– surface for $\sigma_y = 5$.

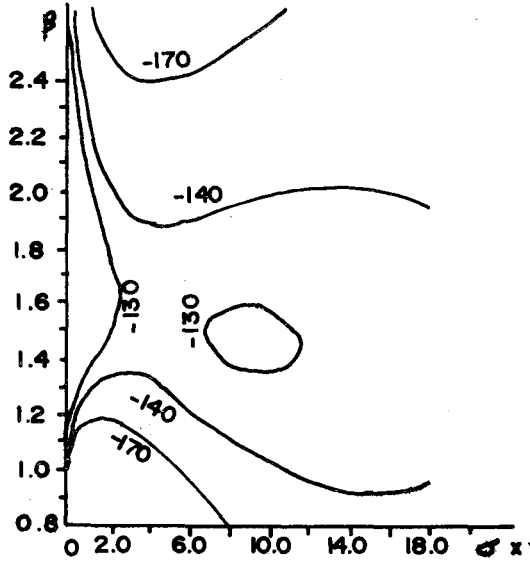| β	$\ln\sigma_x$	L*	For	Varying
1.46	0.48	–144.71	Saddle point (+)	
0.01	0.48	–693.71	Approaching minimum	
1.46	0.48	–144.71	Max. = Saddle point	β
7.00	0.48	–1504.75	Approaching minimum	
1.46	– 3.00	–58.55	Approaching global maximum	
1.46	0.48	–144.71	Min. = Saddle point	σ_y
1.46	1.98	–126.01	Approaching local maximum	
1.60	2.08	–125.34	Local maximum	

Figure 3. Contours of L* for σ_y against slope when σ_x is 4.

Table 2. A summary of the main features of the L*-surface for $\sigma_{xy} = 4$.

β	σ_x	L*	For	Varying
1.56	4	−131.73	Saddle point (+)	
0.01	4	−1067.95	Approaching minimum	
1.56	4	−131.73	Max. = Saddle point	β
2.91	4	−204.89	Approaching minimum	
1.56	0.01	28.07	Approaching global maximum	
1.56	4	−131.73	Min. = Saddle point	σy
1.56	9.01	−129.74	Approaching local maximum	
1.48	9.4	−129.27	Local maximum	

(c) The Fourth Set of graphs

The fourth set of graphs consists of λ ($\lambda = \sigma^2_y / \sigma^2_x$) against β for a range of given values of σ_x.

The following formula was used for drawing this set of graphs:

$$A\beta^2 + 2B\beta + C = 0 \tag{16}$$

This equation was obtained from the function (14). It is quatratic in β and so the roots are:

$$\beta_{1,2} \;=\; \frac{-B \pm (B^2 - AC)^{1/2}}{A}$$

and where

$$A \;=\; 2\sigma^2_x \; (L^+ - 2n\log\sigma_x - \frac{n}{2} \log\lambda) - S_{xx}$$

$$B \;=\; S_{xy}$$

$$C \;=\; 2\lambda \, \sigma^2_x \; (L^+ - 2n\log\sigma_x - \frac{n}{2} \log\lambda) - S_{yy}$$

$$L^+ \;=\; - L^*$$

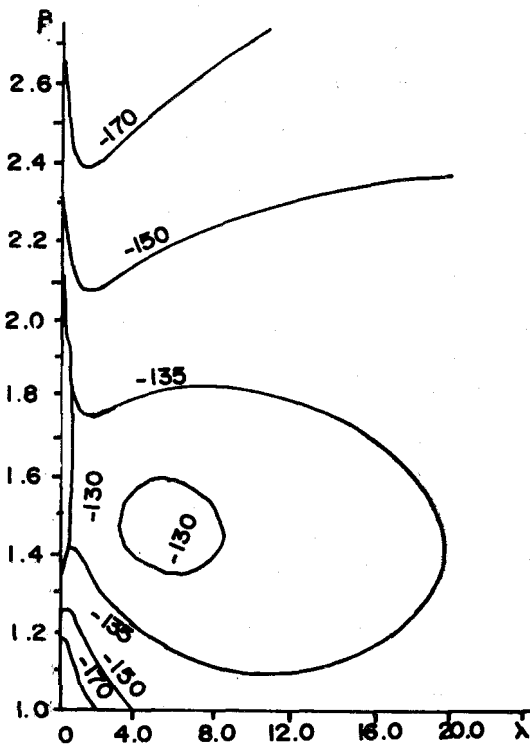One example of the fourth set of graphs was given in Figure 4.

Figure 4. Contours of L^* for lamda against slope when σ_x is 4.

On examination of the graph in Figure 4, there is an evidence of a saddle point whose approximate coordinates and height are as given (+) in Table 3, which presents a summary of the main features of the surface.

Table 4. A summary of main features of the L*-surface for $\sigma_x = 4$.

β	λ	L*	For	Varying
1.54	1.80	−131.30	Saddle point (+)	
0.01	1.80	−639.12	Approaching minimimu	
1.54	1.80	−131.30	Max. = Saddle point	β
2.91	1.80	−202.93	Approaching minimum	
1.54	0.01	−82.96	Approaching global maximum	
1.54	1.08	−131.30	Min. = Saddle point	λ
1.54	5.01	−129.56	Approaching local maximum	
1.46	5.40	−129.26	Local maximum	

From the second, third and fourth sets of graphs described above, we conclude that we cannot consistently estimate either β and σ_x for a given value of σ_y or β and σ_y for a given value of σ_x; therefore, the maximum likelihood method fails to yield any satisfactory estimators for this instance.

REFERENCES

AKAR, M. 1975. Empirical investigation of the likelihood surface in linear functional and structural relationships (MSc Thesis, Unpublished).

BARTLETT, M.S. 1949. Fitting a straight line when both variables are subject to error. *Biometrics*, 5, 207–212.

BERKSON, J. 1950. Are there two regressions. *Journal of the American Statistical Association*, 45, 164–180.

COPAS, J.B. 1972. The likelihood surface in the linear functional relationship problem, *Journal of the Royal Statistical Society*, Series, B, 34, 274–278.

KENDALL, M.G. 1951. Regression, structure and functional relationship, Part I, *Biometrika*, 38, 11–25.

KENDALL, M.G., 1952. Regression, structure and functional relationship, Part II, *Biometrika*, 39, 36–108.

KENDALL, M.G. and A. STUART, 1970. The Advanced Theory of Statistics, Vol. II. London: Griffin.

LÎNDLEY, D.V., 1947. Regression lines and the linear functional relationship, *Journal of the Royal Statistical Society*, Supplements 9, 219–244.

LÎNDLEY, D.V., 1953. Estimation of a structural relationship, *Biometrika*, 40, 47–49.

LÎNDLEY, D.V. and G.M. EL-SAYYAD, 1968. The bayesian estimation of a linear of the functional relationship, *Journal of the Royal Statistical Society*, Series B, 30, 100–202

MADANSKY, A., 1959. The fitting of straight lines when both variables are subject to error, *Journal of the American Statistical Association* 54, 1723–205.

SPRENT, P., 1970. The saddle point of the likelihood surface for a linear functional relationship, *Journal of the Royal Statistical Society*, Series B, 32, 432–434.

SOLARI, M.E., 1969. The 'maximum likelihood solution' of the problem of estimating a linear functional relationship, *Journal of the Royal Statistical Society*, Series B, 31, 372–375.

TAYLOR, J., 1973. A method of fitting several functional relations and of testing for differences between them, *Applied Statistics* 22, 239–248.

TUKEY, J.W., 1951. Components in regression *Biometrics*, 7, 33–70.