Department of Statistics

The University of Wisconsin
Madison, Wisconsin

TECHNICAL REPORT No. 101
February, 1967

ON A TEST FOR SEVERAL LINEAR RELATIONS

by

A. P. Basu\textsuperscript{1, 2}

\textsuperscript{1} This work started while the author was at the University of Minnesota.

\textsuperscript{2} This research has been supported in part by the Office of Naval Research under contract no. 1202(17) Project No. NR 042-222 at the University of Wisconsin.
ON A TEST FOR SEVERAL LINEAR RELATIONS

by

A. P. Basu¹,²

University of Wisconsin
Madison, Wisconsin

1. Summary and Introduction.

The problem of inference concerning the parameters of several functional relationships arises in many different contexts. Various methods have been proposed of attacking the above problem from time to time. For a description of some of these methods see Halperin [5], Johnston [7], Madansky [9], Kendall and Stuart [8] and Tintner [13]. Recently Villegas [14], [15] has considered a method which is based on the assumption that the data were provided by replicated experiments. The above assumption is quite realistic in controlled experiments. In some econometric investigations also, particularly for the study of "error models" of Anderson and Hurwicz [2] replication is possible by pooling cross-section and time series data as shown below. In microeconomic analysis, suppose we are interested in studying the consumption pattern of families belonging to the same income group (or with almost same per capita income, which can be ensured by stratified sampling). Once we establish the microequations, we can derive the corresponding macro equations by using the methods described in Theil[11].

¹ This work started while the author was at the University of Minnesota. The author would like to thank Prof. John Chipman of the University of Minnesota for providing some financial assistance. Thanks are also due to Prof. Raol Lepage of Columbia University and Profs. George Box, John Gurland and Richard Johnson of the University of Wisconsin for some helpful discussions. However, all responsibilities for errors are mine.

² This research has been supported in part by the Office of Naval Research under contract no. 1202(17) Project No. NR 042-222 at the University of Wisconsin.
The object of this paper is to propose a test for the parameters of the econometric "error model" where there are $k$ linear relations ($k \geq 1$). Villegas' problem becomes a special case of our problem. The corresponding estimation problem has been treated by Basu [3] elsewhere. In Section 2 we describe the model to be studied and in Section 4 we propose the test statistic $R$ for testing the parameters of the model. To derive the distribution of the statistic $R$ we need a result which we derive in Section 3. The distribution of $R$ is given in Section 6.

Previously Tintner [12] proposed a test for the above situation. However, he assumed that the covariance matrix $\Sigma$ of errors (to be defined later) is known.

2. Notation and Model.

Let

$$
\begin{align*}
\alpha_1 + \beta_{11} \xi_1 + \cdots + \beta_{1p} \xi_p &= 0 \\
\alpha_2 + \beta_{21} \xi_1 + \cdots + \beta_{2p} \xi_p &= 0 \\
& \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \\
\alpha_k + \beta_{k1} \xi_1 + \cdots + \beta_{kp} \xi_p &= 0
\end{align*}
\tag{2.1}
$$

be $k$ linear relations among the $p$ variables $\xi_1, \xi_2, \ldots, \xi_p$ where $p > k$. Let us assume that the $\xi$'s are not observable in the sense that they are all subject to errors or fluctuations. Rewriting (2.1) in matrix notation, we have

$$
\bar{\alpha} + \bar{B}' \bar{\xi} = 0
\tag{2.2}
$$

Where

$$
\bar{B} = \begin{pmatrix}
\beta_{11} & \beta_{12} & \cdots & \beta_{1k} \\
\beta_{21} & \beta_{22} & \cdots & \beta_{2k} \\
\vdots & \vdots & \ddots & \vdots \\
\beta_{p1} & \beta_{p2} & \cdots & \beta_{pk}
\end{pmatrix} = \begin{pmatrix}
\bar{\alpha}_1 & \bar{\alpha}_2 & \cdots & \bar{\alpha}_k
\end{pmatrix}, \text{ say}
\tag{2.3}
$$

$\bar{B}'$ is the transpose of $\bar{B}$.
\[
\xi = \left( \begin{array}{c}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_p 
\end{array} \right) \quad \text{and} \quad \alpha = \left( \begin{array}{c}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_k 
\end{array} \right).
\]

Let us assume, without any loss of generality, that \( R \) is of full rank. Consider \( r \) points \( \xi_1, \xi_2, \ldots, \xi_r \) on the linear manifold (2.2) and let us assume that \( n_1 \) measurements \( x_{ij} \) \((j = 1, 2, \ldots, n_1)\) are available for the \( i \)-th point \((i = 1, 2, \ldots, r)\). Let

\[(2.4) \quad x_{ij} = \bar{\xi}_i + \epsilon_{ij}\]

where the \( \epsilon_{ij} \)'s are independent and identically distributed random vectors, each following the \( p \)-variate normal distribution with mean vector \( \bar{\xi}_i \) and covariance matrix \( \Sigma \). Let

\[(2.5) \quad n = \sum_{i=1}^{r} n_i\]

\[(2.6) \quad \bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}\]

and

\[(2.7) \quad S = \frac{1}{n-r} \sum_{i,j} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)'\]

If we further assume that \( n-r \geq p \) then it is well known (see Anderson [1]) that \( S \) follows the Wishart distribution with \( \nu = n-r \) degrees of freedom. We shall denote this by writing

\[
S \sim W(\Sigma, p, \nu).
\]

We want to test for the \( \beta \)'s and \( \alpha \)'s.
3. **A geometrical result.**

For our subsequent analysis we need a result which we shall derive below. For some of the concepts used see Halmos [4]. Consider the Hilbert (inner product) space $H$ generated by the $p$-dimensional vectors $\beta_1, \beta_2, \ldots, \beta_k$. Let $x$ be any point in the $p$-dimensional Euclidean space $\mathbb{R}^p$ and $d(x, M)$ is the minimum distance of $x$ from the $(p-k)$-dimensional manifold $M$ defined by

$$
M = \{ \xi \mid \alpha + B^T \xi = 0 \}.
$$

That is, there exists a $\xi_0 \in M$ such that

$$
d^2(x, M) = \rho(x, \xi_0) \leq \rho(x, \xi) \text{ for all } \xi \in M
$$

and $\rho$ is the metric here defined as

$$
\rho(x, \xi) = (x - \xi)'(x - \xi).
$$

We prove the following

**Theorem 3.1.** $d(x, M)$ is given by

$$
d^2(x, M) = (\alpha + B^T x)'(B'B)^{-1}(\alpha + B^T x).
$$

**Proof.** The proof of Theorem 3.1 will follow in several steps. First of all we note, because of the invariance property of the distance function, that

$$
d(x, M) = d(0, M - x)
$$

where

$$
M - x = \{ \xi \mid \alpha + B^T(\xi + x) = 0 \}.
$$

Therefore, we need to find a $\xi_1 \in M - x$ such that

$$
d^2(0, M - x) = \rho(0, \xi).
$$

We obtain the following obvious
Lemma 3.1 \( \xi_j \) satisfies the relation

\[ (\xi_j - \xi_j)^\prime \xi_j = 0 \quad \text{for all} \quad \xi_j \in M - \chi, \]

since the two vectors are orthogonal.

Since \( \xi_j \in M - \chi \), we obtain from (3.6)

\[ \alpha + B'(\xi_j + \chi) = 0. \]

Next we prove

Lemma 3.2. \( \xi_j \) belongs to the space \( \mathcal{H} \).

Proof If not, let \( \xi_j^* \) be the projection of \( \xi_j \) on \( \mathcal{H} \). Then

\[ \rho(\xi_j, \xi_j^*) = (\xi_j - \xi_j^*)^\prime (\xi_j - \xi_j^*) = (\xi_j - \xi_j^*)^\prime \xi_j = 0. \]

Since \( \xi_j \in \mathcal{H} \), we may express \( \xi_j \) as

\[ \xi_j = \sum_{i=1}^{k} \gamma_i \beta_i = \beta \chi. \]

where \( \chi = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_k \end{pmatrix} \) is a suitable coefficient vector, and \( \beta \) is defined in (2.3). Using (3.8) and (3.10)

\[ \alpha + B' \chi = -B' \xi_j = -B' \beta \chi. \]

Therefore \( \chi \) is given by

\[ \chi = - (B' \beta)^{-1} (\alpha + B' \chi), \]

and hence
(3.13) \[ \xi_i = -B(B'B)^{-1}(\bar{a} + B'\bar{x}) \].

Finally, \( d(\bar{a}, M - \bar{x}) \) is given by

(3.14) \[ d'(\bar{a}, M - \bar{x}) = \bar{\xi}_i' \xi_i = (\bar{a} + B'\bar{x})(B'B)^{-1}(\bar{a} + B'\bar{x}), \]

which proves Theorem 3.1.

4. A test statistic for \( \bar{a} \) and \( \bar{B} \).

In this section we propose a statistic which may be used to test the hypothesis

(4.1) \[ H_0: \bar{a} = \bar{a}_0, \quad \bar{B} = \bar{B}_0. \]

The result derived may be also useful for obtaining confidence region for \( \bar{a} \) and \( \bar{B} \).

Let \( d_i = d(\bar{a}, M - \bar{x}_i) \) be the distance of the point \( \bar{x}_i \) from the space \( M \) given in (2.1) and (3.1). From Theorem 3.1 we obtain, under the null hypothesis,

(4.2) \[ d_i' = (\bar{a}_0 + B_0' \bar{x}_i)'(B_0'B_0)^{-1}(\bar{a}_0 + B_0' \bar{x}_i) \] (i = 1, 2, \ldots, r).

From (2.4) and (2.6) the vectors \( \bar{x}_i \)'s are independently and identically distributed with each

(4.3) \[ \bar{x}_i \sim N(\bar{\xi}_i, \Sigma_i / n_i). \]

Hence

(4.4) \[ \bar{a}_0 + B_0' \bar{x}_i \sim \text{NID}(\bar{a}_0, B_0'B_0^{-1} \Sigma_0 / n_i) \] (i = 1, 2, \ldots, r).

Where \( u_i \sim \text{NID}(\mu_i, \Sigma_i) \) (i = 1, 2, \ldots, r)

means that the vectors \( u_i \)'s are independently distributed each following the multivariate normal distribution with mean vector \( \mu_i \) and covariance matrix \( \Sigma_i \).
Since \((\widetilde{B}_o^t \widetilde{B}_o)^{-1}\) is a symmetric nonsingular matrix, there exists a nonsingular matrix \(P\) which simultaneously diagonalize \((\widetilde{B}_o^t \widetilde{B}_o)^{-1}\) and \((\widetilde{B}_o^t \widetilde{B}_o)^{-1}\). Letting

\[(4.5) \quad \widetilde{\alpha}_o + \widetilde{B}_o^t \widetilde{\xi}_i = P \gamma_i \quad (i = 1, 2, \ldots, r)\]

we obtain from (4.2),

\[d_i^2 = \sum_{j=1}^{k} \gamma_{ij}^2\]

where

\[(4.6) \quad P^t (\widetilde{B}_o^t \widetilde{B}_o)^{-1} P = L = \begin{pmatrix} \ell_1 & 0 & \cdots & 0 \\ 0 & \ell_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \ell_k \end{pmatrix}\]

It also follows that

\[(4.7) \quad \widetilde{\eta}_i \sim \text{NID}(0, \text{I}/n_i)\]

where

\[(4.8) \quad P^t (\widetilde{B}_o^t \widetilde{B}_o)^{-1} P = \text{I} .\]

Thus

\[n_i d_i^2 \sim \sum_{j=1}^{k} \gamma_{ij}^2 (1)\]

where \(\chi^2(1)\) is the \(\chi^2\)-distribution with one degree of freedom. Finally,

\[(4.9) \quad d^2 = \sum_{i=1}^{r} n_i d_i^2 \sim \sum_{j=1}^{k} \chi^2(1) .\]

Now \(\text{Cov}(\widetilde{\alpha}_o + \widetilde{B}_o^t \widetilde{\xi}_i) = \widetilde{B}_o^t \widetilde{B}_o / n_i = \widetilde{S} / n_i\) (say) can be estimated by \(\widetilde{B}_o^t \widetilde{S} \widetilde{B}_o / n_i\). But it is well known (see Anderson [1]) that \(\widetilde{S}\) can be expressed as

\[\frac{1}{n} \sum_{j=1}^{v} \tilde{z}_j \tilde{z}_j^t \quad \text{where} \quad \tilde{z}_j \sim \text{NID}(0, \widetilde{\Sigma}).\]

Thus \(\Sigma\) can be estimated by

\[\Sigma = \frac{1}{n} \sum_{j=1}^{v} \tilde{z}_j \tilde{z}_j^t\]

\[(4.10) \quad \Sigma = \frac{1}{n} \sum_{j=1}^{v} \tilde{z}_j \tilde{z}_j^t\]

Now

\[\text{Cov}(\tilde{\alpha}_o + \widetilde{B}_o^t \tilde{\xi}_i) = \widetilde{B}_o^t \widetilde{B}_o / n_i = \widetilde{\Sigma} / n_i\] (say) can be estimated by

\[\widetilde{B}_o^t \tilde{S} \widetilde{B}_o / n_i\]. But it is well known (see Anderson [1]) that \(\tilde{S}\) can be expressed as

\[\frac{1}{n} \sum_{j=1}^{v} \tilde{z}_j \tilde{z}_j^t \quad \text{where} \quad \tilde{z}_j \sim \text{NID}(0, \widetilde{\Sigma}).\]

Thus \(\tilde{\Sigma}\) can be estimated by

\[\tilde{\Sigma} = \frac{1}{n} \sum_{j=1}^{v} \tilde{z}_j \tilde{z}_j^t\]
where \( \mathbf{v}_j \sim \mathbf{B}' \mathbf{z}_j \sim \text{NID}(0, \mathbf{B}' \mathbf{z}_j \mathbf{B}) \)

and \( \mathbf{v}_j \) follows the Wishart distribution independently of \( d^2 \). A statistic that may be used for testing the hypothesis (4.1) is therefore given by

\[
R = \frac{d^2}{|V|^{1/k}}
\]

Where \( |V| \) is the determinant of \( V \) and the distribution of \( |V| \) is well-known to be the product of \( k \) independent \( \chi^2 \)-distributions.

5. Distribution of \( R \).

The exact distribution of \( R \) may be obtained if the distribution of \( |V|^{1/k} \) is known. The exact distribution of \( R \) can be easily found for \( k = 1 \) or \( 2 \). If \( k = 1 \) \( R \) clearly reduces to the statistic given by Villegas [15] whose exact distribution has been given by him. For \( k = 2 \) it is well known that \( |V|^{1/2} \) follows the multiple of a \( \chi^2 \)-distribution. Hence the exact distribution of \( R \) is seen to be a linear combination of two \( \Gamma \)-distributed variables. However, for higher values of \( k \) it is very difficult to find the distribution of \( R \) in closed form. Instead, we make use of the following approximations. First we note that \( d^2 \) is a linear combination of \( k \) independent and identically distributed random variables, each distributed as a \( \chi^2 \) with \( r \) degrees of freedom (the coefficients \( \ell_j \)'s being all positive). Following the line of approach of Satterthwaite [10] we approximate the distribution of \( d^2 \) by \( g \chi^2(h) \) (that is, \( g \) times a \( \chi^2 \) distribution with \( h \) degrees of freedom) where

\[
g = \frac{\sum_{j=1}^{k} \ell_j^2 / \sum_{j=1}^{k} \ell_j}{\sum_{j=1}^{k} \ell_j^2} \quad \text{and} \quad h = r \left( \frac{\sum_{j=1}^{k} \ell_j}{\sum_{j=1}^{k} \ell_j^2} \right)^{\frac{1}{2}} \left( \frac{\sum_{j=1}^{k} \ell_j^2}{\sum_{j=1}^{k} \ell_j} \right)^{\frac{k}{2}}.
\]
Finally, following Hoel [6], we approximate the density of $u = |\underline{y}|^{1/k}$ by

\[ f(u) = \frac{u^{(k-1)/2} \Gamma(k \frac{1}{k})}{\Gamma(k \frac{1}{k} + 1)} e^{-c u^{1/k}} u^{(k-1)/2 - 1} \]

where

\[ c = \frac{k}{2} \left( 1 - \frac{(k-1)(k-2)}{2N} \right)^{1/k} \]

That is \( |\underline{y}|^{1/k} \sim \frac{1}{c} \frac{1}{\sqrt{\Gamma(k \frac{1}{k})}} \frac{\chi^2(k \frac{1}{k} + 1)}{k \frac{1}{k}} \).

An approximate distribution of $R$ is therefore given by

\[ R \sim \frac{q \cdot h \cdot c \cdot u}{|\underline{y}|^{1/k} \cdot k \frac{1}{k} \cdot \Gamma(k \frac{1}{k} + 1)} \]

where $\Gamma_{m,n}$ is the gamma distribution with $(m,n)$ degrees of freedom.

6. **Concluding Remarks.**

If the equations in (2.2) are homogeneous linear equations with $\underline{\varphi} = 0$, we can repeat our arguments taking $\underline{\varphi} = 0$ throughout. Thus equation (3.4) will be reduced to

\[ d^2(\underline{z}, M) = \underline{z}'B(\underline{z}'B)^{-1} \underline{z}' \underline{z} \quad \text{and so forth.} \]

Above results may also be used to find confidence regions for $\underline{\varphi}$ and $B$. However, the computation may be quite involved.

Finally, a second statistic based on $d^2$ alone may be proposed if we use the metric...
(6.2) \[ r(x, \xi) = (x - \xi)' \Sigma^{-1} (x - \xi) \]

instead of (4.4). The distribution of this statistic is not known to the author, however, he conjectures it will also follow the F distribution.
REFERENCES


ON A TEST FOR SEVERAL LINEAR RELATIONS

A. P. Basu

February, 1967

12 pp.

Technical Report #101

Distribution of this document is unlimited

13. Abstract — In this paper we have proposed a test for the parameters of \( k \) linear relations among the \( p \) variables \( \xi_1, \xi_2, \ldots, \xi_p \), where the \( \xi \)'s are not observable and \( p > k \). This extends the work of Villegas (1967) who considered the above problem for the special case \( k=1 \). The applicability of the above model in an econometric problem is also considered.

14. Key Words
   1. Regression Analysis
   2. Econometric error model
   3. Multivariate Analysis