

SOME NOTES ON LEAST SQUARES AND ANALYSIS OF VARIANCE - I

By

S. N. ROY

Institute of Statistics
University of North Carolina

Institute of Statistics
Mimeograph Series No. 81

August, 1953

1. Summary. This paper is purely expository and uses a particular type of matrix methods, believed to be new, to obtain, under the customary linear hypothesis model for "least squares" and analysis of variance and covariance, (i) the "unbiased minimum variance" estimate of any linear function of the unknown parameters, (ii) the "least squares" estimate of the same function (and then show, incidentally, that they are the same) and (iii) the usual F-test for any linear hypothesis on the parameters. The final results come out in matrix forms which can be identified with these occurring in previous work. In the derivations two matrix theorems are repeatedly used, the first of which, being very well known, is merely stated without proof, and the second, not being so well known, is both stated and proved.

2. Notation. Any capital letter, say M , will denote a matrix, M' will denote its transpose and $M(p \times q)$ will indicate that it consists of p rows and q columns, small letters underscored will denote column vectors, underscored and primed will denote row vectors, and plain small letters will denote scalars. As far as possible Roman letters in the latter half of the alphabet, say, u, v, w, x, y, z etc. will stand for stochastic variates and Greek letters will stand for population parameters and Roman letters in the beginning of the alphabet, say a, b, c, d , etc. for other types of (non-stochastic) quantities. A triangular matrix, with zero for the upper right hand triangle, will be denoted by, say, \widetilde{T} . In other words,

$$(2.1) \quad \widetilde{T}(p \times p) = \begin{pmatrix} t_{11} & 0 & \cdot & \cdot & \cdot & 0 \\ t_{21} & t_{22} & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ t_{p1} & t_{p2} & \cdot & \cdot & \cdot & t_{pp} \end{pmatrix}$$

As usual, $I(p)$ will stand for a $p \times p$ unit matrix and $O(p \times q)$ for a matrix

with zero elements in its p rows and q columns.

3. Statement of the problems. Let $\underline{x}(n \times 1)$ denote a set of n uncorrelated stochastic variates with the same (unknown) variance σ^2 and let $E(\underline{x})$ be subject to the constraint:

$$(3.1) \quad E(\underline{x}) = A(n \times p)\underline{\xi}(p \times 1),$$

where $p \leq n$ and $\underline{\xi}(p \times 1)$ is a set of unknown parameters (to be estimated) and A is a matrix of rank $r \leq p \leq n$, whose elements are given by the particular experimental design.

Problem I. Given a non-null $\underline{c}'(1 \times p)$ (subject to certain restrictions to be brought out in section 5) and given \underline{x} , it is required to obtain for $\underline{c}'\underline{\xi}$ a linear estimate $\underline{b}(1 \times n)\underline{x}(n \times 1)$ such that (i) $E(\underline{b}'\underline{x}) = \underline{c}'\underline{\xi}$ (for all $\underline{\xi}$) and (ii) $v(\underline{b}'\underline{x})$ is to be a minimum. $\underline{c}'\underline{\xi}$ will be said to be linearly estimable (or sometimes just "estimable") if and only if (i) is satisfied.

Problem II. Given \underline{c}' and \underline{x} as above, it is required to obtain $\hat{\underline{\xi}}$ so that $(\underline{x}' - \hat{\underline{\xi}}'A')(\underline{x} - A\hat{\underline{\xi}})$ is a minimum.

It will then be incidentally verified that $\underline{b}'\underline{x}$ of Problem I = $\underline{c}'\hat{\underline{\xi}}$ of Problem II.

Problem III. To the model of Problem I add the further condition that each x_i is $N(E(x_i), \sigma^2)$ ($i = 1, 2, \dots, n$). Let us now try to obtain (in terms of given elements) the customary F -test for the hypothesis

$$(3.2) \quad C(q \times p)\underline{\xi}(p \times 1) = \underline{0}(q \times 1),$$

where $r \leq p$ (r being the rank of the A -matrix of (3.1)) and C is a given matrix of rank $s \leq \min(r, q)$.

I and II, which are given below.

4. Auxiliary theorems.

Theorem 1. If $M_1(p \times n)$ ($p \leq n$) is such that $M_1 M_1' = I(p)$, then it is possible to adjoin an $M_2((n-p) \times n)$ to M_1 and complete it into M given by

$$(4.1) \quad M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix} \begin{matrix} p \\ n-p \\ n \end{matrix}, \text{ such that } MM' = I(n).$$

This M_2 is not of course unique. This is a very well-known result and no proof will be given.

Theorem 2. For any $X(p \times n)$ ($p \leq n$; rank $r \leq p$) suppose (as we could without any loss of generality) that

$$(4.2) \quad X(p \times n) = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \begin{matrix} r \\ p-r \\ n \end{matrix},$$

such that X_1 is a matrix of r independent rows while X_2 consists of $p-r$ rows each of which is a linear combination of the rows of X_1 . Then there exists a transformation:

$$(4.3) \quad X(p \times n) = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \begin{matrix} r \\ p-r \\ n \end{matrix} = \begin{matrix} r \\ p-r \end{matrix} \begin{pmatrix} \tilde{T}_1 \\ T_2 \end{pmatrix} \begin{matrix} r \\ r \end{matrix} L(r \times n),$$

such that $LL' = I(r)$.

This transformation could be made unique by taking the diagonal elements of \tilde{T}_1 to be of fixed signs, say all positive. In what follows uniqueness, based

on all the diagonal elements of \widetilde{T}_1 being positive, will be assumed. This theorem is not so well known and a proof will be given in the appendix. Notice that by the conditions of (4.3)

$$(4.4) \quad \begin{pmatrix} X_1 X'_1 & X_1 X'_2 \\ X_2 X'_1 & X_2 X'_2 \end{pmatrix} = \begin{pmatrix} \widetilde{T}_1 \widetilde{T}'_1 & \widetilde{T}_1 \widetilde{T}'_2 \\ \widetilde{T}_2 \widetilde{T}'_1 & \widetilde{T}_2 \widetilde{T}'_2 \end{pmatrix},$$

so that $\widetilde{T}_1 \widetilde{T}'_1 = X_1 X'_1$ and so on.

5. Solution of Problem I. Let $A'(p \times n)$ of (3.1) be factorized into:

$$(5.1) \quad \begin{matrix} r \\ p-r \\ n \end{matrix} \begin{pmatrix} A'_1 \\ A'_2 \end{pmatrix} = \begin{matrix} r \\ p-r \\ r \end{matrix} \begin{pmatrix} \widetilde{T}_1 \\ \widetilde{T}_2 \end{pmatrix} L(r \times n),$$

such that $LL' = I(r)$,

and let L , by adjunction of $L_1((n - r) \times n)$, be completed into $L_2(n \times n)$ such that

$$(5.2) \quad L_2(n \times n) = \begin{pmatrix} L \\ L_1 \end{pmatrix} \begin{matrix} r \\ n-r \\ n \end{matrix} \quad \text{and} \quad L_2 L'_2 = I(n) = L'_2 L_2$$

Notice that L_1 is not unique. Also observe that

$$(5.3) \quad I(n) = \begin{pmatrix} L \\ L_1 \end{pmatrix} (L' \quad L'_1) = \begin{pmatrix} LL' & LL'_1 \\ L_1 L' & L_1 L'_1 \end{pmatrix} = (L' \quad L'_1) \begin{pmatrix} L \\ L_1 \end{pmatrix} \\ = L'L + L'_1 L_1 .$$

Furthermore, with an A having the structure (5.1), let (3.1) be rewritten as

$$(5.4) \quad E(\underline{x}) = \begin{matrix} n & (A_1 & A_2) & & \\ & r & p-r & \begin{pmatrix} \underline{\xi}_1 \\ \underline{\xi}_2 \end{pmatrix} & \begin{matrix} r \\ p-r \end{matrix} \end{matrix},$$

and let $\underline{c}\underline{\xi}$ be rewritten as

$$(5.5) \quad \begin{matrix} (\underline{c}'_1 & \underline{c}'_2) & & \\ & r & p-r & \begin{pmatrix} \underline{\xi}_1 \\ \underline{\xi}_2 \end{pmatrix} \end{matrix} \begin{matrix} r \\ p-r \end{matrix}$$

Now condition (i) (of unbiasedness) of Problem I (section 3) becomes

$$(5.6) \quad (\underline{c}'_1 \quad \underline{c}'_2) \begin{pmatrix} \underline{\xi}_1 \\ \underline{\xi}_2 \end{pmatrix} = E(\underline{b}'\underline{x}) = \underline{b}'E(\underline{x}) = \underline{b}'(A_1 \quad A_2) \begin{pmatrix} \underline{\xi}_1 \\ \underline{\xi}_2 \end{pmatrix} = \underline{b}'(A_1\underline{\xi}_1 + A_2\underline{\xi}_2),$$

and, since this is to be true of all $\underline{\xi}_1$ and $\underline{\xi}_2$, we should have

$$(5.7) \quad \underline{b}'A_1 = \underline{c}'_1 \quad \text{and} \quad \underline{b}'A_2 = \underline{c}'_2,$$

which imposes a number of restrictions ($\leq p$) on \underline{b}' ($1 \times n$) but by no means fully determines \underline{b}' (which has to be determined). Substituting in (5.7) for A_1 and A_2 from (5.1) we have

$$(5.8) \quad \underline{b}'L'\tilde{T}'_1 = \underline{c}'_1 \quad \text{or} \quad \underline{b}'L' = \underline{c}'_1(\tilde{T}'_1)^{-1}, \quad \text{and} \quad \underline{b}'L'T'_2 = \underline{c}'_2.$$

Now to minimize $V(\underline{b}'\underline{x})$ subject to (5.8) we proceed as follows:

$$(5.9) \quad V(\underline{b}'\underline{x}) = \sigma^2 \underline{b}'\underline{b} \quad (\text{since } \underline{x} \text{ is an uncorrelated set with a common variance } \sigma^2)$$

$$= \sigma^2 \underline{b}' \begin{pmatrix} L' & \\ & L'_2 \end{pmatrix} \begin{pmatrix} L \\ L_1 \end{pmatrix} \underline{b} \quad (\text{by (5.2)})$$

$$= \sigma^2 \sqrt{\underline{b}'L'L\underline{b} + \underline{b}'L'_1L_1\underline{b}}$$

$$= \sigma^2 [\underline{c}'_1 (\tilde{T}'_1)^{-1} (\tilde{T}_1)^{-1} \underline{c}_1 + \underline{b}' L'_1 L_1 \underline{b}] \quad (\text{by (5.8)})$$

The minimum $V(\underline{b}'\underline{x})$ is thus reached when

$$(5.10) \quad \underline{b}' L'_1 = \underline{0} ,$$

so that, combining (5.2), (5.8) and (5.10), we have

$$(5.11) \quad \underline{b}' = \underline{c}'_1 (\tilde{T}'_1)^{-1} L ,$$

and hence

$$(5.12) \quad \begin{aligned} \underline{b}'\underline{x} &= \underline{c}'_1 (\tilde{T}'_1)^{-1} L \underline{x} = \underline{c}'_1 (\tilde{T}'_1)^{-1} (\tilde{T}_1)^{-1} A'_1 \underline{x} \quad (\text{using (5.1)}) \\ &= \underline{c}'_1 (\tilde{T}_1 \tilde{T}'_1)^{-1} A'_1 \underline{x} = \underline{c}'_1 (A'_1 A_1)^{-1} A'_1 \underline{x} \quad (\text{using 4.4}) \end{aligned}$$

This gives the "unbiased minimum variance" estimate of $\underline{c}'\underline{\xi}$.

Restriction on \underline{c}' . Now, substituting in (5.7) for \underline{b}' from (5.11) we have

$$(5.13) \quad \begin{aligned} \underline{c}'_2 &= \underline{b}' A_2 = \underline{c}'_1 (\tilde{T}'_1)^{-1} L A_2 = \underline{c}'_1 (\tilde{T}'_1)^{-1} (\tilde{T}_1)^{-1} A'_1 A_2 \quad (\text{using (5.1)}) \\ &= \underline{c}'_1 (A'_1 A_1)^{-1} A'_1 A_2 \quad (\text{using (4.4)}). \end{aligned}$$

We have thus that, in order that $\underline{c}'\underline{\xi}$, i.e. $(\underline{c}'_1 \quad \underline{c}'_2) \begin{pmatrix} \text{or} \\ \text{or} \\ 1 \\ 2 \end{pmatrix}$ may be "esti-

mable" (in the sense indicated), \underline{c}'_2 must be related to \underline{c}'_1 by (5.13), which can be expressed in another form that is more suggestive. From (5.1) we have

$$(5.14) \quad A_2 = L' T'_2 = A_1 (\tilde{T}'_1)^{-1} T'_2 \quad \text{or} \quad A'_2 = T_2 (\tilde{T}_1)^{-1} A'_1 ,$$

which on substitution into (5.13) yields

$$(5.15) \quad \underline{c}'_2 = \underline{c}'_1 (A'_1 A_1)^{-1} A'_1 A_1 (\tilde{T}'_1)^{-1} T'_2 = \underline{c}'_1 (\tilde{T}'_1)^{-1} T'_2$$

Thus \underline{c}'_2 is related to \underline{c}'_1 by the same post factor by which A_2 is related to A_1 .

Invariance of the linear estimate (5.13) under choice of A'_1 . If, instead of A'_1 and \underline{c}'_1 , we choose another set of independent row vectors, say A'_3 and the \underline{c}'_3 (say) to match it, then in place of the right hand side of (5.12) we should have the linear estimate given by replacing the subscript 1 by 3. But remembering that

$$(5.16) \quad A'_3 (r \times n) = T_3 (r \times r) L (r \times n),$$

where T_3 is obtained by picking out from the right hand side of (5.1) the rows corresponding to A'_3 and is necessarily non-singular (since A'_3 is of rank r), and using (5.15) and (5.16), we have

$$\begin{aligned} (5.17) \quad \underline{c}'_2 (A'_3 A_3)^{-1} A'_3 \underline{x} &= \underline{c}'_1 (\tilde{T}'_1)^{-1} T'_3 (T_3 (\tilde{T}'_1)^{-1} A'_1 A_1 (\tilde{T}'_1)^{-1} T'_3)^{-1} T_3 (\tilde{T}'_1)^{-1} A'_1 \underline{x} \\ &= \underline{c}'_1 (\tilde{T}'_1)^{-1} T'_3 (T'_3)^{-1} \tilde{T}'_1 (A'_1 A_1)^{-1} \tilde{T}'_1 (T_3)^{-1} T_3 (\tilde{T}'_1)^{-1} A'_1 \underline{x} \\ &= \underline{c}'_1 (A'_1 A_1)^{-1} A'_1 \underline{x}, \end{aligned}$$

which proves the invariance.

Variance of the "unbiased minimum variance" estimate. From (5.9), (5.11)

and (4.4) this variance is given by

$$(5.18) \quad V(\underline{b}' \underline{x}) = \sigma^2 \underline{b}' \underline{b} = \sigma^2 \underline{c}'_1 (\tilde{T}'_1)^{-1} L L' (\tilde{T}'_1)^{-1} \underline{c}_1$$

$$= \sigma^2 \underline{c}'_1 (\tilde{T}'_1 \tilde{T}'_1)^{-1} \underline{c}_1 = \sigma^2 \underline{c}'_1 (A'_1 A_1)^{-1} \underline{c}'_1,$$

which again, by the method of the previous paragraph, can be shown to be invariant under choice of A'_1 .

6. Solution of Problem II or the "least squares solution". To minimize

$(\underline{x}' - \underline{\xi}' A') (x - A \underline{\xi})$ with respect to $\underline{\xi}$, we differentiate it with regard to $\underline{\xi}$ and equate to zero, and end up with the "least squares" solution $\hat{\underline{\xi}}$ for $\underline{\xi}$:

$$(6.1) \quad A' \underline{x} = A' A \hat{\underline{\xi}}, \text{ or } \begin{pmatrix} \tilde{T}'_1 \\ T'_2 \end{pmatrix} L \underline{x} = \begin{pmatrix} \tilde{T}'_1 \\ T'_2 \end{pmatrix} (T'_1 \ T'_2) \hat{\underline{\xi}},$$

from which we have

$$(6.2) \quad \tilde{T}'_1 L \underline{x} = \tilde{T}'_1 (T'_1 \ T'_2) \hat{\underline{\xi}},$$

or, since \tilde{T}'_1 is non-singular,

$$(6.3) \quad L \underline{x} = (T'_1 \ T'_2) \hat{\underline{\xi}} = (\tilde{T}'_1 \hat{\underline{\xi}}_1 + T'_2 \hat{\underline{\xi}}_2).$$

Substituting in the "least squares estimate" $\hat{\underline{\xi}}$, we have

$$(6.4) \quad \underline{c}' \hat{\underline{\xi}} = \underline{c}'_1 \hat{\underline{\xi}}_1 + \underline{c}'_2 \hat{\underline{\xi}}_2 = \underline{c}'_1 \hat{\underline{\xi}}_1 + \underline{c}'_1 (\tilde{T}'_1)^{-1} T'_2 \hat{\underline{\xi}}_2 \quad (\text{from (5.16)})$$

$$= \underline{c}'_1 (\tilde{T}'_1)^{-1} (\tilde{T}'_1 \hat{\underline{\xi}}_1 + T'_2 \hat{\underline{\xi}}_2) = \underline{c}'_1 (\tilde{T}'_1)^{-1} L \underline{x} \quad (\text{from (6.3)})$$

$$= \underline{c}'_1 (\tilde{T}'_1)^{-1} (\tilde{T}'_1)^{-1} A'_1 \underline{x} \quad (\text{from (5.1)}) = \underline{c}'_1 (A'_1 A_1)^{-1} A'_1 \underline{x} \quad (\text{from (4.4)})$$

which proves the identity of the "least squares solution" with the "unbiased minimum variance solution" (5.12).

7. Solution of Problem III. It is well known that

(7.1) if \underline{x} ($n \times 1$) is a set of n uncorrelated $N(E(\underline{x}), \sigma^2)$ (and thus also independent) variates and if L ($p \times n$) ($p \leq n$) is subject to $LL' = I(p)$, then

$L(p \times n)\underline{x}(n \times 1)$ is a set of p uncorrelated $N(LE(\underline{x}), \sigma^2)$ variates.

It is also well known that

(7.2) if $\underline{u}(p \times 1)$ is an uncorrelated $N(0, \sigma^2)$ set and so is $\underline{v}(q \times 1)$ and if \underline{u} and \underline{v} are mutually uncorrelated, then $\underline{u}'\underline{u}/\sigma^2$ is a χ^2 with p degrees of freedom, $\underline{v}'\underline{v}/\sigma^2$ is a χ^2 with q degrees of freedom and $\underline{qu}'\underline{u}/\underline{pv}'\underline{v}$ is an F with degrees of freedom p and q .

Going back to the model of Problem III in section 3 and to (5.1)-(5.3) we observe that

(7.3) if $\underline{x}(n \times 1)$ is an uncorrelated $N(E(\underline{x}), \sigma^2)$ set, then $L(r \times n)\underline{x}(n \times 1)$ is an uncorrelated $N(LE(\underline{x}), \sigma^2)$ set and $L_1((n-r) \times n)\underline{x}(n \times 1)$ is an uncorrelated $N(L_1E(\underline{x}), \sigma^2)$ set which is also uncorrelated with the $L\underline{x}$ set, since $LL_1 = 0$.

Now from (3.1) and (5.1) - (5.3) we have $E(\underline{x}) = L'(T'_1 T'_2)\underline{\xi}$, so that

$$L_1E(\underline{x}) = L_1L'(T'_1 T'_2)\underline{\xi} = \underline{0}.$$

Thus $L_1\underline{x}$ is an uncorrelated $N(0, \sigma^2)$ set, whence it follows that we have a χ^2 (with $n-p$ degrees of freedom) given by:

$$(7.4) \quad \underline{x}'L'_1L_1\underline{x}/\sigma^2 \quad \text{or} \quad \underline{x}'(I(n) - L'L)\underline{x}/\sigma^2 \quad \text{or}$$

$$\underline{[x'x - x'A_1(T_1T'_1)^{-1}A'_1x]}/\sigma^2 \quad \text{or} \quad \underline{[x'x - x'A_1(A'_1A_1)^{-1}A'_1x]}/\sigma^2.$$

Consider now the hypothesis $C(q \times p)\underline{\xi}(p \times 1) = \underline{0}$, where C is of rank $s \leq \min(q, r)$, r being the rank of the A -matrix and thus being $\leq p < n$. Let us rewrite the hypothesis as

$$(7.5) \quad \begin{matrix} s \\ q-s \\ r \end{matrix} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{matrix} r \\ p-r \\ 1 \end{matrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \underline{0},$$

where $(C_{11} \ C_{12})$ are a set of s independent row vectors and $\begin{pmatrix} C_{11} \\ C_{21} \end{pmatrix}$ a matrix.

each row of which is of the nature of \underline{c}'_1 of section 5.

It is now easy to see that the hypothesis $C \underline{\xi} = \underline{0}$ is equivalent to $C_{11}\underline{\xi}_1 + C_{12}\underline{\xi}_2 = \underline{0}$, so that we shall work in terms of this latter. Going back to (5.12), (5.13) and (5.15) we note that

$$(7.6) \quad C_{12} = C_{11}(\widetilde{T}'_1)^{-1}T'_2 \text{ and}$$

$$(7.7) \quad \underline{0} = C_{11}\underline{\xi}_1 + C_{12}\underline{\xi}_2 = E[C_{11}(A'_1A_1)^{-1}A'_1\underline{x}] \\ = E[C_{11}(\widetilde{T}_1\widetilde{T}'_1)^{-1}\widetilde{T}_1L\underline{x}].$$

Now $(\widetilde{T}_1\widetilde{T}'_1)^{-1}\widetilde{T}_1L$ is a $r \times n$ matrix of rank r and C_{11} is a $s \times r$ matrix

($s \leq r$) of rank s . Let $C_{11}(\widetilde{T}_1\widetilde{T}'_1)^{-1}\widetilde{T}_1L$, which is a $s \times n$ matrix, be of rank,

say $t \leq s \leq \min(q, r)$ (note that $r \leq p \leq n$). In practically all the designs, to be discussed, $t = s$. We shall, therefore, first consider this case.

Case $t = s$. From (4.3) let

$$(7.8) \quad C_{11}(\widetilde{T}_1\widetilde{T}'_1)^{-1}\widetilde{T}_1L = \widetilde{V}(s \times s) M (s \times n), \text{ where } MM' = I(s), \text{ and } \widetilde{V} \text{ of}$$

course is non-singular.

Then we have

$$(7.9) \quad E(\underline{Mx}) = (\tilde{V})^{-1} E \left[C_{11} (\tilde{T}_1 \tilde{T}'_1)^{-1} \tilde{T}_1' L \underline{x} \right] = \underline{0} \quad (\text{from (7.7)})$$

and furthermore

$$(7.10) \quad \underline{M} L' \underline{x} = (\tilde{V})^{-1} C_{11} (\tilde{T}_1 \tilde{T}'_1)^{-1} \tilde{T}_1' L L' \underline{x} = 0, \text{ so that}$$

$$(7.11) \quad \underline{Mx} \text{ is a } s\text{-set of uncorrelated } N(0, \sigma^2), \underline{L}_1 \underline{x} \text{ (of (7.9)) is a } (n - r)$$

set of uncorrelated $N(0, \sigma^2)$, \underline{Mx} and $\underline{L}_1 \underline{x}$ are mutually uncorrelated, and

hence

$$(7.12) \quad (n - r) \underline{x}' \underline{M}' \underline{Mx} / s \underline{x}' L' \underline{x} \text{ is an } F \text{ with degrees of freedom } s \text{ and } n - r.$$

Using (7.4), (7.8) and of course (5.1) and (4.4), we can reduce (7.12)

to

$$(7.13) \quad \frac{(n - r) \underline{x}' A_1 (A_1' A_1)^{-1} C'_{11} \left[C_{11} (A_1' A_1)^{-1} C'_{11} \right]^{-1} C_{11} (A_1' A_1)^{-1} A_1' \underline{x}}{r \left[\underline{x}' \underline{x} - \underline{x}' A_1 (A_1' A_1)^{-1} A_1' \underline{x} \right]}$$

which is an F (with degrees of freedom r and $n - r$) for testing the hypothesis $C \underline{\xi} = \underline{0}$ and which is expressed in terms of quantities directly observed or given by the experimental design and the hypothesis to be tested. The form (7.13) can be shown to be invariant under the kind of choice indicated in section 5, in much the same way as there.

Case $t < s$. From (4.3), (7.4), (7.8), (5.1) and (4.4) it is possible to construct the corresponding F -statistic which will be essentially of the same form as (7.13) but more complicated in detail. This is not given here, since it is not of much practical interest.

The optimum properties of the usual F -test with respect to the linear

hypothesis $C\underline{\xi} = \underline{0}$ are well known and need not be discussed here.

8. Concluding remarks. This paper which is Part I gives the abstract set-up. Corresponding to different experimental designs and different sample survey plans or different regression problems, there will be different sets of A_1 and C_{11} or A_1 and \underline{c}_1 . For the main types, these matrices (and vectors) will be given in Part II in which will also be given various short cut methods of actual calculation of the linear estimate (5.17) or the F-statistic (7.13), together with the final forms (5.17) or (7.13) in each case.

APPENDIX

9. Proof of theorem 2 of section 4. To establish the existence (with real values on both sides) of the one-to-one transformation (4.3) (under the conditions stated there) we first observe that, given the right hand side, X is unique and of rank r. Next, given X, let us assume that

(9.1) there is a unique (real) non-singular \widetilde{T}_1 (r x r) in terms of X_1 such that $X_1 X'_1 = \widetilde{T}_1 \widetilde{T}'_1$.

Then, assuming (9.1), we have

(9.2) $L = (\widetilde{T}_1)^{-1} X_1$, so that, given X_1 (and thus \widetilde{T}_1), L is easily seen to exist and to be unique. Next,

$$(9.3) \quad LL' = (\widetilde{T}_1)^{-1} X_1 X'_1 (\widetilde{T}'_1)^{-1} = (\widetilde{T}_1)^{-1} (\widetilde{T}_1 \widetilde{T}'_1) (\widetilde{T}'_1)^{-1} = I(r).$$

Next, we note that

$$(9.4) \quad T_2 L = X_2 \text{ or } T_2 LL' = X_2 L' \text{ or } T_2 = X_2 X'_1 (\widetilde{T}'_1)^{-1},$$

so that given X_1, X_2 (and thus \widetilde{T}_1), T_2 exists and is unique.

Now, to prove (9.1), let X_1 ($r \times n$) (rank r) consist of row vectors \underline{u}'_i ($1 \times n$) ($i = 1, 2, \dots, r$), so that the equation (9.1) will yield

$$(9.5) \quad t_{11}^2 = \underline{u}'_1 \underline{u}_1, \quad t_{11}^2 t_{22}^2 = \left| \begin{pmatrix} \underline{u}'_1 \\ \underline{u}'_2 \end{pmatrix} (\underline{u}_1 \underline{u}_2) \right|,$$

$$t_{11}^2 t_{22}^2 t_{33}^2 = \left| \begin{pmatrix} \underline{u}'_1 \\ \underline{u}'_2 \\ \underline{u}'_3 \end{pmatrix} (\underline{u}_1 \underline{u}_2 \underline{u}_3) \right|, \text{ and so on and finally}$$

$$t_{11}^2 t_{22}^2 \dots t_{rr}^2 = X_1 X_1'.$$

The right hand expression being all positive, real solutions for t_{ii} 's

($i = 1, 2, \dots, r$) exist and can be made unique by taking, say, all t_{ii} 's

(the diagonal elements of \widetilde{T}_1) to be positive. Next, we have from (9.1),

for the non-diagonal elements, the equations

$$(9.6) \quad t_{11} t_{21} = \underline{u}'_1 \underline{u}_2, \quad t_{11} t_{31} = \underline{u}'_1 \underline{u}_3, \text{ and so on and in general}$$

$$\sum_{k=1}^{\min(i,j)} t_{ik} t_{jk} = \underline{u}'_i \underline{u}_j \quad (i \neq j = 1, 2, \dots, r).$$

The diagonal elements t_{ii} 's being already uniquely determined by (9.5),

it is easy to check that (9.6) gives the appropriate set of linear equations

for (unique) real solutions for t_{ij} 's ($i \neq j = 1, 2, \dots, r$). This com-

pletes the proof of theorem 2.

Adlenda to "Some Notes on Least squares and Analysis of variance - I".

The references are to the sections and subsections of the previous paper, unless otherwise indicated.

6. Solution of problem II or the "Least squares solution".

The following proof, which does not use the differential calculus, seems to be preferable:

$$\begin{aligned}
 (\underline{x}' - \underline{\xi}' A')(\underline{x} - A\underline{\xi}) &= (\underline{x}' - \underline{\xi}' A')(L' \quad L_1') \begin{pmatrix} L \\ L_1 \end{pmatrix} (\underline{x} - A\underline{\xi}) \\
 &= \sqrt{\underline{x}' - \underline{\xi}' \begin{pmatrix} \tilde{T}_1 \\ T_2 \end{pmatrix} L} \sqrt{(L' \quad L_1') \begin{pmatrix} L \\ L_1 \end{pmatrix}} \sqrt{\underline{x} - L' \begin{pmatrix} \tilde{T}_1 \\ T_2 \end{pmatrix} \underline{\xi}} \\
 &= \sqrt{\underline{x}' L' - \underline{\xi}' \begin{pmatrix} \tilde{T}_1 \\ T_2 \end{pmatrix}} \sqrt{L' \underline{x} - \begin{pmatrix} \tilde{T}_1 \\ T_2 \end{pmatrix} \underline{\xi}} + \underline{x}' L_1' L_1 \underline{x}
 \end{aligned}$$

(using (5.3)). It is now quite easy to see that given \underline{x} and A the minimum value $(\underline{x}' - \underline{\xi}' A')(\underline{x} - A\underline{\xi})$, under variation of $\underline{\xi}$, will be attained if

$$L\underline{x} = \begin{pmatrix} \tilde{T}_1 \\ T_2 \end{pmatrix} \underline{\xi}.$$

If we now want the "least squares estimate" $\hat{c}' \underline{\xi}$ of an "estimable linear function" $c' \underline{\xi}$, we have from the above:

$$\begin{aligned}
 \underline{c}'\hat{\underline{x}} &= \underline{c}'_1\hat{\underline{x}}_1 + \underline{c}'_2\hat{\underline{x}}_2 = \underline{c}'_1\hat{\underline{x}}_1 + \underline{c}'_1(\tilde{T}'_1)^{-1} T'_{22}\hat{\underline{x}}_2 \quad (\text{from (5.16)}) \\
 &= \underline{c}'_1(\tilde{T}'_1)^{-1}(\tilde{T}'_1\hat{\underline{x}}_1 + T'_{22}\hat{\underline{x}}_2) = \underline{c}'_1(\tilde{T}'_1)^{-1} L\underline{x} \quad (\text{from 6.3}) \\
 &= \underline{c}'_1(\tilde{T}'_1)^{-1}(\tilde{T}'_1)^{-1} A'_1\underline{x} \quad (\text{from (5.1)}) = \underline{c}'_1(A'_1A_1)^{-1}A_1\underline{x} \quad (\text{from (4.6)}) ,
 \end{aligned}$$

which proves the identity of the "least squares solution" of an "estimable linear function" with the "unbiased minimum variance solution"

(7.5)-(7.6) Solution of Problem III - the hypothesis to be tested.

In the matrix C of (7.5)-(7.6), the submatrices C_{12} and C_{22} are related to C_{11} and C_{21} by the same matrix factor by which A_2 is related to A_1 . This is a restriction on the C -matrix of the hypothesis which will thus be called "testable" if the C -matrix satisfies this restriction, besides being of rank $s \leq \min (q,r)$, where r is the rank of the A -matrix.

(7.7). Solution of Problem III.

Consider the observations made about the rank of $C_{11}(\tilde{T}'_1\tilde{T}'_1)^{-1}\tilde{T}'_1L$, after (7.7) of the previous paper. Remembering that C_{11} is an $s \times r$ ($s \leq r$) matrix of rank s and $(\tilde{T}'_1\tilde{T}'_1)^{-1}\tilde{T}'_1L$ is an $r \times n$ ($r \leq n$) matrix of rank r , it can be shown that the rank of $C_{11}(\tilde{T}'_1\tilde{T}'_1)^{-1}\tilde{T}'_1L$ must also be s . In other words the possibility of this rank being less than s would not

arise at all. This is based on the following theorem.

Theorem 3. If $\rho(M)$ denotes the rank of a matrix M , and if $B(q \times r)$ ($q \leq r$) and $C(s \times p)$ ($p \leq s$) are of ranks q and p respectively, then

$$\rho [A(p \times q)] = \rho [A(p \times q)B(q \times r)] = \rho [C(s \times p)A(p \times q)].$$

Proof. The following proof, based upon certain well-known matrix theorems, is due to Donald L. Burkholder, a student of the author, and is about the simplest proof known to the author.

Assuming

$$(i) \rho(M_1 M_2) \leq \min [\rho(M_1), \rho(M_2)], \quad (ii) \rho(M) = \rho(M^t) = \rho(MM^t) \text{ and}$$

$$(iii) \rho [M(p \times q)] = \rho [M(p \times q) D(q \times q)]$$

$$= \rho [E(p \times p) M(p \times q)] \quad \text{if } D \text{ and } E \text{ are non-singular,}$$

we have

$$\rho [A(p \times q)] = \rho [A(p \times q) B(q \times r) B^t(r \times q)]$$

$$\leq \min [\rho(AB), \rho(B^t)], \text{ i.e., } \leq \rho(AB)$$

But $\rho(AB) \leq \rho(A)$, whence $\rho(A) = \rho(AB)$.

Likewise, starting with $A'C'$ and noting that $\rho(CA) = \rho(A'C')$, we should have, in an exactly similar manner, $\rho(CA) = \rho(A)$.

This proves theorem 3 under the conditions stated there. It is easy to see that (iii), given in all matrix text books, is a special case of theorem 3.