

A REPORT ON SOME RESULTS IN SIMULTANEOUS
OR JOINT LINEAR (POINT) ESTIMATION

S. N. ROY

Institute of Statistics
Mimeograph Series No. 394
June, 1964

UNIVERSITY OF NORTH CAROLINA
Department of Statistics
Chapel Hill, N. C.

A REPORT ON SOME RESULTS IN SIMULTANEOUS
OR JOINT LINEAR (POINT) ESTIMATION

by

S. N. Roy

June 1964

This research was supported by the National Science
Foundation Grant No. GP-1660

Institute of Statistics
Mimeo Series No. 394

A REPORT ON SOME RESULTS IN SIMULTANEOUS
OR JOINT LINEAR (POINT) ESTIMATION¹

by

S. N. Roy

University of North Carolina, Chapel Hill, N. C.

=====

1) Summary. The unbiased minimum variance linear estimate of an estimable linear function of parameters under a linear model has been long well known and has long figured prominently in statistical literature. The main results here have been expressed in different forms, among which is the one used by the author and his collaborators [2, 3]. The problem may be called one of single linear estimation for the univariate case; univariate in the physical sense because all experimental units are measured on one variate, and also in the mathematical sense because we have only one unknown variance² to deal with. As a follow-up on this problem three other developments suggest themselves to us, namely (1) the problem of estimating with linear functions of the observations several linear functions of the parameters under a linear model, in the univariate case, (2) the problem of a single linear estimation in the multivariate case where each experimental unit is measured on p variates and we have an unknown $p \times p$ dispersion matrix to deal with, and (3) the problem of simultaneous linear estimation under the model of problem (2).

In this paper by way of a necessary introduction to the following sections, section 1 reproduces in the form given in [2, 3], the main results of a single linear estimation for the univariate case. Section 2 discusses the problem of simultaneous linear estimation for the univariate case, section 3 that of single linear estimation for the multivariate case and section 4 that of simultaneous linear estimation for the multivariate case. The results offered

¹This research was supported by the National Science Foundation Grant No. GP-1660.

in sections 2, 3 and 4 have been obtained in the course of frequent discussion between the author and some of his collaborators in this particular area and, with the exception of one particular result, have not been published so far by our group and, as far as we are aware, by anybody else either. Of this particular result a proof is given here different from the one given in an earlier mimeographed paper [2]. Under the different models of linear estimation introduced in the following sections, linear functions of parameters are estimated with linear functions of the observations and this is what is traditionally called linear estimation. However, there are other principles or methods of estimation of the same linear functions including the important and interesting method of least squares, leading, in some situations, to the same estimates as in linear estimation, and, in some situations, to different estimates. Except in section 1 where an interesting tie-up with least-squares solution is just mentioned, the method of least-squares does not figure elsewhere in this paper. We shall discuss in a later communication such connection as we have noticed between the two approaches, for the estimates of linear parametric functions under the models of sections 2, 3 and 4.

1. Single linear estimation for the univariate case.

The model $\underline{x}' = (x_1, x_2, \dots, x_n)$ is a set of uncorrelated random variates with common variance σ^2 and expectation given by

$$(1.1) \quad E(\underline{x}) = \underline{A} \underline{\xi} = n \left[\begin{array}{c|c} \underline{A}_I & \underline{A}_D \\ \hline r & m-r \end{array} \right] \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_r \\ \xi_{r+1} \\ \vdots \\ \xi_m \end{pmatrix} \begin{matrix} r \\ m-r \end{matrix},$$

say, where \underline{A} is a given $n \times m$ matrix, to be called the model matrix, which depends partly upon the design of the experiments and partly upon what experimental statisticians call the "model", and $\underline{\xi}$ is a set of m unknown parameters. The x_i 's, of course, are measurements of a single character on

n experimental units. Without any loss of generality we assume that $m < n$; let $\text{rank } \underset{\sim}{A} = r \leq m < n$ and let $\underset{\sim}{A}_I$ be a basis and $\underset{\sim}{A}_D$ the so-called set of dependent column vectors.

The problem is to estimate $\underset{\sim}{c}' \underset{\sim}{\xi}$ with $\underset{\sim}{b}' \underset{\sim}{x}$, where the $l \times n$ $\underset{\sim}{c}'$ is given but $\underset{\sim}{\xi}$ is unknown, while $\underset{\sim}{x}$ is observed but the $l \times n$ $\underset{\sim}{b}'$ is to be determined.

The criteria used to determine $\underset{\sim}{b}$ are that

$$(1.2) \quad (i) \quad E(\underset{\sim}{b}' \underset{\sim}{x}) = \underset{\sim}{c}' \underset{\sim}{\xi} = [\underset{\sim}{c}'_I : \underset{\sim}{c}'_D] \begin{pmatrix} \underset{\sim}{\xi}_1 \\ \underset{\sim}{\xi}_2 \end{pmatrix}, \text{ say, (for all } \underset{\sim}{\xi}), \text{ and}$$

$$(ii) \quad V(\underset{\sim}{b}' \underset{\sim}{x}) \quad \text{is to be a minimum.}$$

(i) is called the unbiasedness condition and only partially determines $\underset{\sim}{b}'$, (ii) is called the minimum variance condition, and, together with (i), completely determines $\underset{\sim}{b}$. (i) also imposes a restriction on $\underset{\sim}{c}$, usually called the estimability condition. Linear estimation thus involves (i) a linear model, (ii) a linear function of $\underset{\sim}{\xi}$ to be estimated, and (iii) a linear estimator.

The main results of a single linear estimation for the univariate case are summarized below from (1.3)-(1.8).

The estimability condition is that

$$(1.3) \quad \text{rank } \underset{\sim}{A} = \text{rank} \begin{pmatrix} \underset{\sim}{A} \\ \underset{\sim}{c}' \end{pmatrix}$$

or equivalently

$$(1.3.1) \quad \underset{\sim}{c}'_D = \underset{\sim}{c}'_I (\underset{\sim}{A}'_I \underset{\sim}{A}_I)^{-1} \underset{\sim}{A}'_I \underset{\sim}{A}_D.$$

The unbiased minimum variance of an estimable $\underset{\sim}{c}' \underset{\sim}{\xi}$ is

$$(1.4) \quad \underset{\sim}{c}'_I (\underset{\sim}{A}'_I \underset{\sim}{A}_I)^{-1} \underset{\sim}{A}'_I \underset{\sim}{x},$$

which is often called the best unbiased linear estimate or BLUE estimate.

The variance of this estimate is

$$(1.5) \quad \underset{\sim}{c}'_I (\underset{\sim}{A}'_I \underset{\sim}{A}_I)^{-1} \underset{\sim}{c}_I \sigma^2$$

An unbiased estimate of σ^2 is

$$(1.6) \quad \tilde{x}' [I(n) - A_I (A_I' A_I)^{-1} A_I'] \tilde{x} / (n-r) .$$

(1.7) All these results are invariant under the choice of a basis A_I for the matrix A .

Observe that the choice of a basis A_I induces a partitioning of \tilde{x} , and a partitioning of $\tilde{\xi}$ induces a partitioning of \tilde{c}' .

(1.8) The least squares estimate of an estimable linear function $\tilde{c}'\tilde{\xi}$ is the same as the BLUE estimate given by (1.4).

1.1 Some remarks on the derivation of the results from (1.3)-(1.5).

Although the above results are derived elsewhere [2], it is worthwhile to sketch briefly the derivation of the results (1.3)-(1.5) in as much as the tools involved are the same as will be used in sections 2, 3 and 4. For the derivation the two following mathematical lemmas are needed, in particular.

Lemma I. There exists the factorizations

$$(1.9) \quad \begin{pmatrix} A_I' \\ \tilde{A}_D' \\ n \end{pmatrix} \begin{matrix} r \\ m-r \\ n \end{matrix} = \begin{matrix} r \\ m-r \\ r \end{matrix} \begin{pmatrix} I_1 \\ I_2 \\ r \end{pmatrix} \tilde{L}$$

where I_1 is a non-singular triangular matrix (say lower triangular) and \tilde{L} is $n \times r$ and orthonormal, i.e., $\tilde{L} \tilde{L}' = I(r)$. All that is needed is just this existence and not the expression for the right side matrices.

(1.9) will imply that

$$(1.9.1) \quad A_I' = I_1 \tilde{L} \quad \text{or} \quad \begin{matrix} A_I' \\ \tilde{A}_D' \end{matrix} \begin{matrix} A_I \\ \tilde{A}_I \end{matrix} = \begin{matrix} \tilde{T}_1' \\ \tilde{T}_2' \end{matrix} \begin{matrix} \tilde{L} \\ \tilde{L} \end{matrix} .$$

Lemma II. \tilde{L} could be completed by an $(n-r) \times n$ \tilde{L}'_1 into an orthogonal matrix, so that

$$(1.10) \quad \begin{matrix} r \\ n-r \end{matrix} \begin{pmatrix} \tilde{L} \\ \tilde{L}'_1 \end{pmatrix} \begin{pmatrix} \tilde{L}' & \tilde{L}'_1 \\ r & n-r \end{pmatrix} = \begin{pmatrix} \tilde{L} \tilde{L}' & \tilde{L} \tilde{L}'_1 \\ \tilde{L}'_1 \tilde{L}' & \tilde{L}'_1 \tilde{L}'_1 \end{pmatrix} = \begin{pmatrix} \tilde{I}(r) & 0 \\ 0 & \tilde{I}(n-r) \end{pmatrix} \\ = \begin{pmatrix} \tilde{L}' & \tilde{L}'_1 \\ \tilde{L} & \tilde{L}'_1 \end{pmatrix} \begin{pmatrix} \tilde{L} \\ \tilde{L}'_1 \end{pmatrix} = \tilde{L}' \tilde{L} + \tilde{L}'_1 \tilde{L}'_1 .$$

Applying (1.9) to the expectation model (1.1) and the unbiasedness requirement (1.2) (i), we have

$$(1.11) \quad \tilde{b}' \tilde{L}' = \tilde{c}'_I \tilde{T}'_1^{-1}$$

and

$$(1.12) \quad \tilde{c}'_D = \tilde{c}'_I \begin{pmatrix} \tilde{T}'_1^{-1} & \tilde{T}'_2 \end{pmatrix} = \tilde{c}'_I (\tilde{A}'_I \tilde{A}_I)^{-1} \tilde{A}'_I \tilde{A}_D .$$

The condition (1.12) is the same as (1.3.1) which is equivalent to (1.3). The condition (1.11), as observed earlier, only partially determines \tilde{b} . For the minimum variance estimate we proceed as follows .

$$(1.13) \quad V(\tilde{b}' \tilde{x}) = \sigma^2 (\tilde{b}' \tilde{b}) = \sigma^2 \tilde{b}' [\tilde{L}' \tilde{L} + \tilde{L}'_1 \tilde{L}'_1] \tilde{b} = \sigma^2 (\tilde{c}'_I \tilde{T}'_1^{-1} \tilde{T}'_1^{-1} \tilde{c}'_I) \\ + \sigma^2 \tilde{b}' [\tilde{L}'_1 \tilde{L}'_1] \tilde{b}$$

and the extreme side of the equations is minimized by putting

$$(1.14) \quad \tilde{b}' \tilde{L}'_1 = 0 ,$$

and this together with (1.11) and (1.9.1), gives

$$(1.15) \quad \tilde{b}' = \tilde{c}'_I \tilde{T}'_1^{-1} \tilde{L} = \tilde{c}'_I \tilde{T}'_1^{-1} \tilde{T}'_1^{-1} \tilde{L} = \tilde{c}'_I (\tilde{A}'_I \tilde{A}_I)^{-1} \tilde{A}'_I ,$$

whence follows (1.4). It is easy to check that (1.5) has been already obtained enroute.

1.2 Comparison between two different designs, in terms of a single linear estimation.

For simplicity of discussion, let us consider a connected block-treatment type of design, i.e., a two-way classification with blocks and treatments. The n will be the same in both, the m will usually be different and so also the A matrix and the σ^2 . The part of ξ that relate to the treatment contributions will stay the same and the linear function to be estimated will involve only this part of ξ . That design will be the better one for which $c_I'(A_I' A_I)^{-1} c_I \sigma_i^2$ ($i = 1, 2$) is less. It may be observed that, for connected designs, $\text{rank} = b+v-1$, and hence we have the same c_I for both, and also that, since we are considering only the treatment part of ξ , the block part of ξ will be absent, and hence also the corresponding part of A_I . We note, in addition, that if both designs are equal block side designs, then σ_i^2 will be less for the design whose block side k is less.

1.3 The quasi-multivariate model [2]. When x has the symmetric $p \times p$ positive definite dispersion matrix Σ which is supposed to be known except for a scalar multiplier, the expectation set-up is the same as (1.1), and, under this general model, the linear estimation problem is the same as before. To handle this we use the matrix factorization $\Sigma = \tau \tau'$, where τ is a lower triangular non-singular matrix, put $\tau^{-1} x = y$, note that y has the dispersion matrix I , which practically throws back the problem on the previous case, work in terms of y and then transform back to x . The second form of this estimability condition (1.3.1) now becomes

$$\begin{aligned}
 (1.16) \quad c_D' &= c_I'(A_I' \Sigma^{-1} A_I)^{-1} A_I' \Sigma^{-1} A_D \\
 &= c_I' \tilde{T}_1^{-1} (L \Sigma^{-1} L')^{-1} \tilde{T}_1^{-1} \tilde{T}_1 (L \Sigma^{-1} L') T_2' \\
 &= c_I' \tilde{T}_1^{-1} T_2' = c_I'(A_I' A_I)^{-1} A_I' A_D.
 \end{aligned}$$

Thus the estimability condition stays the same ^{as} before, which can also be checked more directly.

The unbiased minimum variance estimate now becomes

$$(1.17) \quad \underset{\sim}{c}_I' (\underset{\sim}{A}_I' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{A}_I)^{-1} \underset{\sim}{A}_I' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{x} ,$$

and the variance of this estimate becomes

$$(1.18) \quad \underset{\sim}{c}_I' (\underset{\sim}{A}_I' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{A}_I)^{-1} \underset{\sim}{c}_I .$$

2. Simultaneous linear estimation for the univariate case.

With the univariate model of the previous section the problem now is one of estimating the vector $\underset{\sim}{C} \underset{\sim}{\xi}$ with the vector $\underset{\sim}{B} \underset{\sim}{\xi}$, where $\underset{\sim}{C}$ is a given $s \times m$ matrix of rank $s \leq r$ and $\underset{\sim}{B}$ is to be determined. As to the total criterion to be used, we hold off for a moment, while we observe that a part of the total criterion could well be unbiasedness, as before, leading to the estimability condition

$$(2.1) \quad \text{rank} \begin{pmatrix} \underset{\sim}{A} \\ \underset{\sim}{C} \end{pmatrix} \begin{matrix} n \\ s \\ m \end{matrix} = \text{rank } \underset{\sim}{A} ,$$

or equivalently,

$$(2.2) \quad \underset{\sim}{C}_D = \underset{\sim}{C}_I (\underset{\sim}{A}_I' \underset{\sim}{A}_I)^{-1} \underset{\sim}{A}_I' \underset{\sim}{A}_D .$$

At the back of this we have, as before, from unbiasedness, a condition similar to (1.11), being now given by

$$(2.3) \quad \underset{\sim}{B} \underset{\sim}{L}' = \underset{\sim}{C}_I \underset{\sim}{T}_I^{-1} .$$

Let us now denote by $\underset{\sim}{c}_i'$ and $\underset{\sim}{b}_i'$ ($i = 1, 2, \dots, s$) the row vectors of $\underset{\sim}{C}$ and $\underset{\sim}{B}$, pick up from (1.4) the BLUE for each $\underset{\sim}{c}_i' \underset{\sim}{\xi}$ and set up and study the properties of the following estimator for $\underset{\sim}{C} \underset{\sim}{\xi}$:

$$(2.4) \quad \underset{\sim}{C}_I (\underset{\sim}{A}_I' \underset{\sim}{A}_I)^{-1} \underset{\sim}{A}_I' \underset{\sim}{x} ,$$

which we shall just call, for shortness, the BLUE estimator. Its dispersion matrix is easily seen to be given by

$$(2.5) \quad \sigma^2 \underset{\sim}{C}_I (A'_I \underset{\sim}{A}_I)^{-1} \underset{\sim}{C}'_I = \sigma^2 \underset{\sim}{M}_O \text{ (say) } .$$

For any competitor $\underset{\sim}{B}^* \underset{\sim}{x}$, satisfying the unbiasedness condition (2.3), the dispersion matrix will be given by

$$(2.6) \quad \begin{aligned} \sigma^2 \underset{\sim}{B}^* \underset{\sim}{B}^{*'} &= \sigma^2 \underset{\sim}{B}^* [\underset{\sim}{L}' \underset{\sim}{L} + \underset{\sim}{L}'_1 \underset{\sim}{L}_1] \underset{\sim}{B}^{*'} \\ &= \sigma^2 [\underset{\sim}{C}_I (A'_I \underset{\sim}{A}_I)^{-1} \underset{\sim}{C}'_I + \underset{\sim}{B}^* (\underset{\sim}{L}'_1 \underset{\sim}{L}_1) \underset{\sim}{B}^{*'}] = \sigma^2 [\underset{\sim}{M}_O + \underset{\sim}{M}^*], \text{ say.} \end{aligned}$$

We observe that $\underset{\sim}{M}_O$ is an $s \times s$ symmetric positive definite and $\underset{\sim}{M}^*$ is an $s \times s$ symmetric and at least positive semi-definite matrix of the same rank as $\underset{\sim}{B}^* \underset{\sim}{L}'_1$. This implies, in terms of characteristic roots, that

(2.7) any ch $[\underset{\sim}{M}_O + \underset{\sim}{M}^*] \geq$ the corresponding ch $\underset{\sim}{M}_O$, the inequality holding for at least one root, unless $\underset{\sim}{M}^*$ is zero, i.e., $\underset{\sim}{B}^* \underset{\sim}{L}'_1 = \underset{\sim}{O}$, in which case we are back in the BLUE (2.4). As a consequence, among other results we have, in particular, the following

$$(2.8) \quad \begin{aligned} \text{(i)} \quad \text{ch}_{\max} \underset{\sim}{M}_O &\leq \text{ch}_{\max} [\underset{\sim}{M}_O + \underset{\sim}{M}^*] \quad , \\ \text{(ii)} \quad \text{tr} \underset{\sim}{M}_O &< \text{tr} [\underset{\sim}{M}_O + \underset{\sim}{M}^*] \end{aligned}$$

and

$$\text{(iii)} \quad |\underset{\sim}{M}_O| < |\underset{\sim}{M}_O + \underset{\sim}{M}^*| .$$

In fact, a strict inequality like (ii) or (iii) would hold for all the symmetric functions of the roots (in the sense of theory of equations) and, in general, for any function of the roots which is such that it decreases if any root is decreased without decreasing the other roots. Thus, the BLUE defined by (2.4), is a unique solution in terms of minimization of any such function of the roots, including the trace and the determinant, but is not a unique solution (though no doubt a solution) in terms of minimizing the root of

any order, including the largest root.

2.1 A physical interpretation of the largest root, the trace and the determinant minimization criterion. If we attach, on economic or utilization considerations, different economic weights $(a_1, a_2, \dots, a_s) = \underline{a}'$ normalized by $\underline{a}' \underline{a} = 1$ to the components of $\underline{C} \underline{\xi}$ and thus to those of any estimator $\underline{B} \underline{x}$, we come out with a single linear function $\underline{a}' \underline{C} \underline{\xi}$ estimated by $\underline{a}' \underline{B} \underline{\xi}$, the variance of the estimator being given by $\sigma^2(\underline{a}' \underline{B} \underline{B}' \underline{a})$.

(i) The normalized weight system \underline{a} for which $(\underline{a}' \underline{B} \underline{B}' \underline{a})$ is a maximum ($= \text{ch} [\underline{B} \underline{B}']$), is thus the most unfavorable weight system, producing the largest variance for the estimator. Choosing \underline{B} , subject to (2.3), so as to minimize $\text{ch} [\underline{B} \underline{B}']$, which is the largest root minimization criterion, is, therefore, a minimax criterion, in this sense.

(ii) For the trace criterion, we observe first that $\underline{B} \underline{B}' = \underline{\Gamma} \underline{D} \underline{\Gamma}'$, where $\underline{\Gamma}$ is an orthogonal matrix and $\underline{\gamma}$'s are the roots (all positive) of $\underline{B} \underline{B}'$. Consider now

$$(2.9) \quad \int_{\underline{a}' \underline{a} = 1} \underline{a}' [\underline{B} \underline{B}'] \underline{a} \, d(\underline{a}) = \int_{\underline{a}' \underline{a} = 1} \underline{a}' [\underline{\Gamma} \underline{D} \underline{\Gamma}'] \underline{a} \, d(\underline{a})$$

$$= \int_{\underline{a}^* \underline{a}^* = 1} [\underline{a}^* \underline{\gamma} \underline{D} \underline{\gamma} \underline{a}^*] d(\underline{a}^*) = \sum_{i=1}^s \gamma_i \int_{\underline{a}^* \underline{a}^* = 1} a_i^{*2} d(\underline{a}^*) = \text{const. tr} [\underline{B} \underline{B}'] ,$$

Thus the trace is proportional to the average variance, the averaging being over this particular measure on \underline{a} , and hence the trace minimization (over \underline{B} , subject to (2.3)), is minimizing this average variance.

(iii) If the \underline{x} of section 1 is, in addition $N[E(\underline{x}), \sigma^2 \underline{I}(n)]$, then any unbiased vector estimator $\underline{B} \underline{x}$ of $\underline{C} \underline{\xi}$ is $N[\underline{C} \underline{\xi}, \sigma^2 \underline{B} \underline{B}']$, and in this case, it is well known that $|\underline{B} \underline{B}'|$, the determinant of the dispersion matrix is

proportional to just the square of the volume of the ellipsoid given by $[\underline{B} \underline{x} - \underline{C} \underline{\xi}]' [\underline{B} \underline{B}']^{-1} [\underline{B} \underline{x} - \underline{C} \underline{\xi}] = \text{constant}$, which has a probability content independent of \underline{B} and \underline{C} . Thus minimizing the determinant merely means minimizing the volume for this given probability content. However, if \underline{x} is not multinormal, this kind of interpretation gets a lot fuzzier.

(iv) We also have incidentally the important result that the BLUE estimator vector (2.4) produces, for any single linear function $\underline{a}' \underline{C} \underline{\xi}$, an estimator which could not be improved by any other estimator vector.

2.2 Comparison between two different designs, in terms of simultaneous linear estimation [1]. As in section 1, for each design we take the corresponding BLUE estimator given by (2.4). Then, as in section 1, aside from the differential in σ_1^2 and σ_2^2 , the comparison is now between the matrices $\underline{C}_I' (\underline{A}'_{\sim I} \underline{A}_{\sim I})^{-1} \underline{C}_I$, ($i = 1, 2$), and the two designs might and, in general, would compare differently according as we use the largest root or trace or determinant criterion on the matrix given above.

3. Single linear estimation for the multivariate case [4].

The model. We have here $p \times n$ matrix \underline{X} consisting of n uncorrelated p -dimensional column vectors with a common but unknown $p \times p$ dispersion matrix $\underline{\Sigma}$. There are experimental units and on each p different variates are measured. If we roll out \underline{X} into a $(1 \times pn)$ row vector $[\underline{x}'_1, \underline{x}'_2, \dots, \underline{x}'_p] = \underline{x}^*$ (say), then \underline{x}^* has a dispersion matrix $[\underline{I}(n) \otimes \underline{\Sigma}]$, in the notation of Kronecker product. For the expectation, let us first consider the following model which is less than the most general one (and \underline{x} which is based on a common $n \times m$ model matrix \underline{A} for the different variates and an $m \times p$ matrix of parameters $\underline{\xi}$):

$$(3.1) \quad E(\underline{X}') = \underline{A} \underline{\xi}.$$

If we now roll out ξ into a vector $[\xi_1', \xi_2', \dots, \xi_p'] = \xi^{*'} (say)$, then (3.1) can be rewritten as

$$(3.2) \quad E(x^*) = [A \otimes I(p)] \xi^{*}$$

Suppose now that (3.2) we want the BLUE estimate of a linear function $c^{*'} \xi^{*} = \sum_{i=1}^p c_i' \xi_i$ (say). From section 1.3 the estimability condition is now

$$(3.3) \quad \text{rank} \begin{pmatrix} A \otimes I(p) \\ c^{*'} \end{pmatrix} = \text{rank} [A \otimes I(p)]$$

or

$$\text{rank} \begin{pmatrix} A \\ c_i' \end{pmatrix} = \text{rank} A, \quad (i = 1, 2, \dots, p),$$

which means that each $c_i' \xi_i$, ($i = 1, 2, \dots, p$), should be separately estimable.

We denote $[c_{i1}', c_{i2}', \dots, c_{ip}']$ by $c_i^{*'}.$ The BLUE estimate, using section 1.3, will be given by

$$\begin{aligned} (3.4) \quad & c_i^{*'} [(A_i \otimes I(p))' (I(n) \otimes \Sigma)^{-1} (A_i \otimes I(p))]^{-1} [A_i \otimes I(n)]' [I(n) \otimes \Sigma]^{-1} x^* \\ &= c_i^{*'} [(A_i' \otimes \Sigma^{-1}) (A_i \otimes I(p))]^{-1} [A_i' \otimes I(p)] [I(n) \otimes \Sigma^{-1}] x^* \\ &= c_i^{*'} [A_i' A_i \otimes \Sigma^{-1}]^{-1} [A_i' \otimes \Sigma^{-1}] x^* \\ &= c_i^{*'} [(A_i' A_i)^{-1} \otimes \Sigma] [A_i' \otimes \Sigma^{-1}] x^* \\ &= \sum_{i=1}^p c_{ii}' (A_i' A_i)^{-1} A_i' x_i, \end{aligned}$$

which is the sum of the BLUE estimates for the individual $c_i' \xi_i$'s ($i = 1, 2, \dots, p$)

The variance of this estimate, using again section 1.3, is given by

$$\begin{aligned}
(3.5) \quad & c_{\sim I}^{*'} [(A_{\sim I}' \otimes I(p))' (I(n) \otimes \Sigma)^{-1} (A_{\sim I}' \otimes I(p))^{-1}] c_{\sim I}^* \\
& = c_{\sim I}^{*'} [(A_{\sim I}' \otimes \Sigma^{-1}) (A_{\sim I}' \otimes I(p))]^{-1} c_{\sim I}^* \\
& = c_{\sim I}^{*'} [(A_{\sim I}' A_{\sim I}) \otimes \Sigma^{-1}]^{-1} c_{\sim I}^* = c_{\sim I}^{*'} [(A_{\sim I}' A_{\sim I})^{-1} \otimes \Sigma] c_{\sim I}^* \\
& = \sum_{i,j=1}^p c_{\sim I i}^i (A_{\sim I}' A_{\sim I})^{-1} c_{\sim I j} \quad \sigma_{ij}
\end{aligned}$$

where $\Sigma = [\sigma_{ij}]$.

For an unbiased estimate of the dispersion matrix we have, as an analogue of (1.6), the expression

$$(3.6) \quad X [I(n) - A_{\sim I} (A_{\sim I}' A_{\sim I})^{-1} A_{\sim I}'] X' / (n-r).$$

In the development of (3.4) - (3.6) we have tacitly assumed that Σ is a symmetric positive definite matrix. If it were symmetric positive semi-definite (as in the study of certain nonparametric problems), then, while (3.4) is a BLUE estimate, it is not unique, but there could be others, depending now on the "orthogonal completion" of Σ . This will be discussed in a later paper.

Let us now go back to the case we were considering, namely when Σ is symmetric positive definite. The most significant fact is the unknown Σ drops out of the BLUE estimate. This is possible only because the model matrix is of the form $A \otimes I(p)$, while the dispersion matrix is of the form $I(n) \otimes \Sigma$. This would not happen if we had different A_i 's for the different variates, or in other words, a pseudo-diagonal model matrix with diagonal submatrices of the form A_i ($i = 1, 2, \dots, p$). However, the estimability condition would still be the simple form

$$(3.7) \quad \text{rank } A_i = \text{rank} \begin{pmatrix} A_i \\ c_i^i \end{pmatrix}, \quad (i = 1, 2, \dots, p)$$

For some time now we have, in fact, been working in terms of what might be called multivariate designs consisting of $(n_i \times m_i) \underline{A}_i$ ($i = 1, 2, \dots, p$) where even n_i and m_i might be different for the different variates. The possibility of linear estimation under such models will be discussed in subsequent papers. For special patterns of the different \underline{A} 's and/or for special patterns of $\underline{\Sigma}$, something useful (from which the unknown $\underline{\Sigma}$ drops out) is still available.

Going back to the univariate model matrix \underline{A} we have been considering so far, we now observe that because of (3.5) the comparison between two different designs would now be more complicated and uncertain than in sections 1 and 2. Whatever we have been able to say meaningfully about comparison of designs in this set up will be discussed later.

3.1 The same problem under a somewhat more general model. While staying with the same \underline{A} for the different variates and the same problem of a single linear estimation of $\underline{c}^* \underline{\xi}^*$, let us change the expectation model into

$$(3.7) \quad E(\underline{X}') = \underline{A} \underline{\xi} \underline{G} ,$$

where \underline{X}' is $n \times p$, \underline{A} is $n \times m$, $\underline{\xi}$ is $m \times q$, \underline{G} is $q \times p$, $\text{rank } \underline{G} = q \leq p$.

One way (though by no means a unique way) to go at the problem of a single linear estimation in this set up is to transform (3.7) to

$$(3.8) \quad \underline{A} \underline{\xi} = E(\underline{X}') \underline{G}' (\underline{G} \underline{G}')^{-1} = E(\underline{Y}') \quad (\text{say}) ,$$

where $\underline{X}' \underline{G}' (\underline{G} \underline{G}')^{-1} = \underline{X}' \underline{G}^* \quad (\text{say}) = \underline{Y}'$.

Note that $V(\underline{Y}') = (\underline{G} \underline{G}')^{-1} \underline{G} \underline{\Sigma} \underline{G}' (\underline{G} \underline{G}')^{-1} = \underline{G}^* \underline{\Sigma} \underline{G}^*$.

Next, we seek, as before, to estimate $\underline{c}^* \underline{\xi}^*$ with $\underline{b}^* \underline{y}^*$, where \underline{y}^* is rolled out as a vector from \underline{Y} in the same way as \underline{x}^* is rolled out as a vector from \underline{X} . The estimability condition will stay the same as before, while the BLUE estimate in terms of \underline{y}^* or \underline{y}_i 's will come out in the same form as (3.4). This, transformed to \underline{x}_i 's, will come out as

$$(3.9) \quad \sum_{i=1}^q \underset{\sim}{c}'_{Ii} (\underset{\sim}{A}'_I \underset{\sim}{A}_I)^{-1} [\underset{\sim}{G}^* \underset{\sim}{X}]_i ,$$

where $[\underset{\sim}{G}^* \underset{\sim}{X}]_i$ denotes the i -th row of the matrix. Likewise the variance of this estimate comes out as

$$(3.10) \quad \sum_{i,j=1}^q \underset{\sim}{c}'_{Ii} (\underset{\sim}{A}'_I \underset{\sim}{A}_I)^{-1} \underset{\sim}{c}_{Ij} \sigma^*_{ij}$$

where

$$[\sigma^*_{ij}] = \underset{\sim}{\Sigma}^* = \underset{\sim}{G}^* \underset{\sim}{\Sigma} \underset{\sim}{G}^{*'} .$$

Even within the set up of a linear estimation there are other ways of going at the problem (in the sense of both formulation and execution) which will be discussed later.

4. Simultaneous linear estimation for the multivariate case: Under the same model as in section 3 the problem now is to estimate $\underset{\sim}{C}^* \underset{\sim}{\xi}^*$ with $\underset{\sim}{B}^* \underset{\sim}{x}^*$, where $\underset{\sim}{C}^*$ is $s \times mp$ and $\underset{\sim}{B}^*$ is $s \times np$. Let us write $\underset{\sim}{C}^*$ as $[\underset{\sim}{C}_1, \underset{\sim}{C}_2, \dots, \underset{\sim}{C}_p]$, combining (2.1) with (3.3), we have now the estimability condition

$$(4.1) \quad \text{rank} \begin{pmatrix} \underset{\sim}{A} \\ \underset{\sim}{C}_i \end{pmatrix} = \text{rank} \underset{\sim}{A} , \quad (i = 1, 2, \dots, p) .$$

The BLUE estimate vector (from which, as in the BLUE estimate scalar of (3.4), the unknown dispersion matrix $\underset{\sim}{\Sigma}$ will drop out) is in a form analogous to (3.4) and is given by

$$(4.2) \quad \sum_{i=1}^p \underset{\sim}{C}_{Ii} (\underset{\sim}{A}'_I \underset{\sim}{A}_I)^{-1} \underset{\sim}{A}'_I \underset{\sim}{x}_i ,$$

and the dispersion matrix of this estimate, in a form analogous to (3.5), is given by the $s \times s$ matrix

$$(4.3) \quad \sum_{i,j=1}^p \underset{\sim}{C}_{Ii} (\underset{\sim}{A}'_I \underset{\sim}{A}_I)^{-1} \underset{\sim}{C}'_{Ij} \sigma_{ij} .$$

As to comparison of two different designs in terms of this dispersion matrix of the BLUE estimate vector, we can go back to the considerations of section 2 and use the largest root or the trace or the determinant criterion on this matrix for the different designs and then notice the same kind of difficulty that we commented upon while discussing the problem of similar comparison at the end of section 3.

REFERENCES

- [1] Krishniah, P. R., "Simultaneous tests and the efficiency of generalized balanced incomplete block designs", Doctorial thesis, 1963, University of Minnesota, Minneapolis, Minn.
- [2] Roy, S. N., Some Aspects of Multivariate Analysis, Chapter 12, John Wiley and Sons, 1957.
- [3] Roy, S. N. and Gnanadesikan, R., "Some contributions to ANOVA in one or more dimensions", Ann. Math. Stat., vol. 30, pp. 304-317, 1959.
- [4] Srivastava, J. N., "A note on the best linear unbiased estimates for multivariate populations", Institute of Statistics, University of North Carolina mimeograph series no. 339 (1962).

NORTH CAROLINA STATE
OF THE UNIVERSITY OF NORTH CAROLINA
AT RALEIGH

INSTITUTE OF STATISTICS

(Mimeo Series available for distribution at cost)

364. Brosh, I. and P. Naor. On optimal disciplines in priority queueing. 1963.
365. Yniguez, A. D'B., R. J. Monroe, and T. D. Wallace. Some estimation problems of a non-linear supply model implied by a minimum decision rule. 1963.
366. Dall'Aglio, G. Present value of a renewal process. 1963.
367. Kotz, S. and J. W. Adams. On a distribution of sum of identically distributed correlated gamma variables. 1963.
368. Sethuraman, J. On the probability of large deviations of families of sample means. 1963.
369. Ray, S. N. Some sequential Bayes procedures for comparing two binomial parameters when observations are taken in pairs. 1963.
370. Shimi, I. N. Inventory problems concerning compound products. 1963.
371. Heathcote, C. R. A second renewal theorem for nonidentically distributed variables and its application to the theory of queues. 1963.
372. Portman, R. M., H. L. Lucas, R. C. Elston and B. G. Greenberg. Estimation of time, age, and cohort effects. 1963.
373. Bose, R. C. and J. N. Srivastava. Analysis of irregular factorial fractions. 1963.
374. Potthoff, R. F. Some Scheffe-type tests for some Behrens-Fisher type regression problems. 1963.
375. Mesner, D. M. A note on parameters of PBIB association schemes. 1963.
376. Bose, R. C. and J. N. Srivastava. Multidimensional partially balanced designs and their analysis, with applications to partially balanced factorial fractions. 1963.
377. Potthoff, R. F. Illustrations of some Scheffe-type tests for some Behrens-Fisher-type regression problems. 1963.
378. Bruck, R. H. and R. C. Bose. The construction of translation planes from projective spaces. 1963.
379. Hogan, M. D. Several test procedures (sequential and non-sequential) for normal means. M. S. Thesis. 1963.
380. van Zyl, G. J. J. Inventory control for perishable commodities. 1963.
381. Smith, W. L. Some remarks on a distribution occurring in neural studies. 1964.
382. Orense, M. and A. Grandage. Estimation of the exponentials in a double exponential regression curve. Ph.D. Thesis. 1963.
383. Rhyne, A. L. Some multiple comparison sign tests. Ph.D. Thesis. 1964.
384. Kuehl, R. O. and J. O. Rawlings. Estimators for genetic parameters of population derived from parents of a diallel mating. 1964.
385. Khatri, C. G. Distribution of the largest or smallest characteristic root. 1964.
386. Johnson, N. L. Cumulative sum control charts and the Weibull distribution. 1964.
387. Khatri, C. G. A note on the inequality of the characteristic roots of the product and the sum of two symmetric matrices. 1964.
388. Khatri, C. G. Monotonicity of the power functions of some tests for a particular kind of multicollinearity and unbiasedness and restricted monotonicity of the power functions of the step down procedures for MANOVA and independence. 1964.
389. Scott, McKinley. Study of some single counter queueing processes. 1964.
390. Overton, W. S. On estimating the mean ordinate of a continuous function over a specified interval. Ph.D. Thesis. 1964.
391. White, Kathleen. An investigation of small sample properties of Zahl's estimate of a general Markov process model for follow up studies. Ph.D. Thesis. 1964.
392. Rohde, C. A. Contributions to the theory, computation and application of generalized inverses. 1964.
393. Burton, R. C. An application of convex sets to the construction of error correcting codes and factorial designs. 1964.
394. Roy, S. N. A report on some results in simultaneous or joint linear (point) estimation. 1964.
395. Potthoff, Richard F. and Whittinghill, M. Testing for homogeneity in binomial, multinomial, and Poisson distributions. 1964.
396. Hoeffding, Wassily. Asymptotically optimal tests for multinomial distribution. 1964.
397. Magistad, J. and R. L. Anderson. Some Steady-state and transient solutions for sampled queues. Ph.D. Thesis. 1964.
398. Linder, Arthur. Design and analysis of experiments. 1964.
399. Khatri, C. G. On a MANOVA model applied to problems in growth curve. 1964.
400. Naor, P. On a first-passage time problem in chemical engineering. 1964.
401. Basu, D. On the minimality of a boundedly complete sufficient sub-field. 1964.