A REPORT ON SOME ASPECTS OF MULTIVARIATE ANALYSIS

by

S. N. Roy
Institute of Statistics
University of North Carolina

This research was supported, in part, by the United States Air Force through the Office of Scientific Research of the Air Research and Development Command

Institute of Statistics
Mimeograph Series No. 121

December 30, 1954
PREFACE

This report does not by any means attempt to cover the entire area of multivariate analysis, or even a major part of it. Aside from certain basic notions and results due to Fisher, Hotelling, Mahalanobis, Karl Pearson, Wilks, Wishart, Yule and some of their predecessors, which have now become current coin, this report is primarily concerned with those developments in multivariate analysis in which the author has been specially interested and with which he and some of his collaborators have been associated over several years. Part of the material presented here, as far as the author is aware, has not been published before, while the rest has been collected from papers by various workers in this sector including the author and his collaborators. It will be seen that in this report the statistical approach to different problems and the mathematical treatment of all such problems are uniform and perhaps somewhat individual, and that this applies to all specific results, no matter whether they are due to the author and his collaborators, or to other workers in the field or to both groups simultaneously.

What has not been discussed in this report has been developed and adequately handled in important papers by Anderson, Bartlett, Bose, Hsu, Kendall, Mahalanobis, Mosteller, Narea, Rao, Votew, Wald and Brookner, Wilks and several other workers. Three excellent books touching upon but not primarily restricted to this sector, "Advanced Theory of Statistics" by M. G. Kendall, Vol. 2 §35, "Advanced Statistical Methods in Biometric Research" by C. R. Rao §14 and "Mathematical Statistics" by S. S. Wilks §28 have, between them, brought together and competently presented a substantial part of this material. For an adequate, unified and up-to-date presentation of this whole material the author, among others, is looking forward to the forthcoming book by T. W. Anderson, supposed to be dealing perhaps more or less exclusively with multivariate analysis.

The preparation of this report, within a relatively short period, has been made possible only through the active co-operation of the entire secretarial staff of the department of statistics at Chapel Hill including, in particular, Mrs. Bonnie Baker Fethman and Mrs. Anne Kiley who did most of the typing and of several students of the author including, in particular, K. V. Ramachandran and A. E. Serhan who rendered indispensable mechanical and critical help. This research was supported, in part, by the United States Air Force through the Office of Scientific Research of the Air Research and Development Command. To all these individuals and organizations are due the sincerest thanks of the author.

The author would be deeply grateful if errors were brought to his notice and suggestions were made for improvement in form no less than in content.
# TABLE OF CONTENTS

1. Notation and preliminaries. ........................................... 1

2. General scope and objectives of the report. ....................... 5

3. A class of tests including, in particular, the type I and type II tests. 6

4. Extended type I test and an obvious property of it. ............. 15

5. Random samples from p-variate normal populations. .............. 16

6. Statement of the specific problems to be discussed in the report. 22

7. Tests for the null hypothesis with respect to (i) one dispersion matrix, (ii) two dispersion matrices, (iii) multivariate analysis of variance of means and (iv) independence between two sets of variates. 25

8. Reduction of some distribution problems and some actual distributions. 33

9. On the c.d.f. of the largest or/and of the smallest root. ....... 51

10. General observations on the operating characteristics of the tests for (i)-(iv) of Chapter 7. 53

11. Lower bounds on the power functions of the four tests given in Chapter 7. 55

12. On the monotonic character of the lower bounds on the power functions of the four tests and hence the near monotonic character of the four power functions themselves. 59

13. On two other monotonic lower bound functions, one for the multivariate analysis of variance test and the other for the test of independence, and on some further properties of the powers of the tests relating to one dispersion matrix and two dispersion matrices. 63

14. Least squares and univariate analysis of variance and covariance with observations on extensions. 77

15. Some univariate and bivariate confidence bounds. ............... 84

16. Multivariate confidence bounds. .................................... 92

# APPENDIX

A.1. Some preliminary results in matrix theory. .................... A-1

A.2. Some results in quadratic forms. ............................... A-8
| A.3. | Some results in transformations. | A-12 |
| A.4. | Invariance of the characteristic roots under certain transformations. | A-32 |
| A.5. | Some general theorems on Jacobians. | A-35 |
| A.7. | Canonical reduction of certain distribution problems. | A-59 |
| A.8. | Some general results in integration. | A-66 |
| A.9. | Some results in integration connected with the distribution of the largest or/and of the smallest root. | A-70 |
1. Notation and preliminaries. As far as possible the following notation and convention will be used, all departures being clearly indicated at the proper places. Greek letters will stand for population parameters and Italic letters over the first half of the alphabet for given (non-stochastic) quantities and over the latter part from, say, \( r \) to the end for sample quantities. Matrices and vectors under consideration will consist of real elements (these will be called real matrices or vectors) except occasionally when they might have complex elements (these will be called complex matrices or vectors). Capital letters will stand for matrices, small letters for scalars, small letters underscored for column vectors and for row vectors if they are primed. The transpose of a matrix or a column vector will be denoted by priming such quantities, the conjugate complex transpose of a matrix \( M \) by \( \hat{M} \), the set of characteristic roots of \( M \) (if it is square) by \( c(M) \), its trace by \( \text{tr} M \), the modulus of such a determinant by \( |M| \), the modulus of a scalar \( m \) by \( |m| \) and the inverse of a matrix \( M \) (if it is square and non-singular) by \( M^{-1} \). A real square matrix \( M(p \times p) \) will be called \( \perp \) if it is orthogonal, i.e., if \( MM' = I(p) \) (= \( M'M \), necessarily) and if \( M(p \times q) \) \((q \leq p)\) is such that \( MM' = I(p) \) then \( M \) will be called semi-orthogonal.

To indicate the structure, a \( p \times q \) matrix, say \( M \), or a \( p \times 1 \) column vector, say \( m \), will sometimes be written respectively as \( M(p \times q) \) or \( m(p \times 1) \). A matrix \( M \) whose typical element is \( m_{ij} \) will sometimes be denoted by \( (m_{ij}) \). A diagonal matrix whose diagonal elements are, say, \( a_1, a_2, \ldots, a_p \) will be denoted by \( D_a \). A diagonal matrix with \( +1 \) for its diagonal elements will be denoted by \( D_k \). \( \hat{A}(p \times p) \) or sometimes simply \( \hat{A} \) will stand for the triangular matrix

\[
\begin{pmatrix}
a_{11} & 0 & \cdots & 0 \\
a_{21} & a_{22} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
a_{p1} & a_{p2} & \cdots & a_{pp}
\end{pmatrix}
\]  

We have also \( |\hat{A}| = \prod_{i=1}^{p} a_{ii} \), and it is easy to check that if \( \hat{A} \) is non-singular, then
\( a_{i1} \neq 0 \) and \( \tilde{\Lambda}^{-1} \) will also be a triangular matrix with the same configuration as \( \tilde{\Lambda} \). The product of two triangular matrices of the same configuration is a triangular matrix of the same configuration. \( \tilde{\Lambda}' \) is a triangular matrix of the opposite configuration to \( \tilde{\Lambda} \). If \( \Lambda(p \times q) = (a_{ij}) \), then \( d\Lambda \) will stand for \( \prod_{j=1}^{p} \prod_{i=1}^{q} da_{ij} \) and if \( a'(1 \times p) = (a_1, \ldots, a_p) \), then \( d\sigma \) will denote \( \prod_{i=1}^{p} da_i \). The Jacobian of the transformation from any independent set of variables, say, \( y \) (with of course the same number of elements as \( x \)) will be denoted by \( J(x : y) \), while a symbol, say, \( \frac{\partial(u_1, \ldots, u_k)}{\partial(\theta_1, \ldots, \theta_k)} \) (= \( \frac{\partial u}{\partial \theta} \)) will have the same well known meaning as in the calculus.

The terms "positive definite" and "positive semi-definite" will be abbreviated p.d. and p.s.d. respectively. "Almost everywhere", that is, "except for a set of (probability) measure zero" will be referred to as a.e. As usual, p.d.f. and c.d.f. will stand respectively for the probability density function and the cumulative distribution function (of a stochastic variate).

A stochastic variate \( x (-\infty < x < \infty) \) will as usual be called \( N(\xi, \sigma^2) \) if it has the p.d.f.

\[
(1.2) \quad \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{(x - \xi)^2}{2\sigma^2} \right],
\]

where \( -\infty < \xi < \infty \) and \( \sigma > 0 \). It is well known that \( E(x) = \xi \) (to be called the mean) and \( E(x - \xi)^2 = \sigma^2 \) (to be called the variance). A stochastic vector \( x(p \times 1) \) (-\( \infty < x_i < \infty \)) will be called \( N(\xi, \Sigma) \) if it has the p.d.f.

\[
(1.3) \quad \frac{1}{(2\pi)^{p/2}} \exp \left[ -\frac{1}{2} \text{tr} \Sigma^{-1}(x - \xi)(x' - \xi') \right],
\]

where \( -\infty < \xi_i < \infty \) and where \( \Sigma \) is a \( p \times p \) symmetric p.d. matrix. It is also well known that

\[
(1.4) \quad E(x) = \xi \text{ and } E(x - \xi)(x' - \xi') = \Sigma.
\]

\( \xi \) will be called the population mean vector and \( \Sigma \) the population dispersion matrix.

The symbols \( \in, \cup, \cap \), "A statement \( \iff \) another statement",
"A statement \(\rightarrow\) another statement", will all be taken over from the notation and terminology of set theory and measure theory and so also \(w\)' for the complement of a set \(w\) in a space \(x\). The most powerful critical region of size, say, \(\beta_H (\geq 1)\) (which, under fairly general conditions, will exist and which, under slightly less general conditions, will also be unique) of a simple hypothesis \(H_0\) against a simple alternative \(H\) (such that \(H \in \mathcal{S}\) where \(\mathcal{S}\) stands for the domain of possible alternatives) will be denoted by \(w(H_0, H, \beta_H)\) and its complement, the acceptance region by \(w'(H_0, H, \beta_H)\), to indicate that, in general, both will depend on \(\beta_H, H_0\) and \(H\). The union of regions \(w(H_0, H, \beta_H)\) over different \(H \in \mathcal{S}\) will be denoted by \(\bigcup_{H \in \mathcal{S}} w(H_0, H, \beta_H)\) or simply by \(\bigcup_{H \in \mathcal{S}} w\), and the intersection of regions \(w'(H_0, H, \beta_H)\) over \(H \in \mathcal{S}\) by \(\bigcap_{H \in \mathcal{S}} w'(H_0, H, \beta_H)\) or simply by \(\bigcap_{H \in \mathcal{S}} w'\). \(P(H_0, H, \beta_H)\) will stand for the power of the most powerful critical region of size \(\beta_H\) for \(H_0\) against \(H\). \(\phi_H\) will usually denote the p.d.f. under the hypothesis \(H\). It is well known that \(w(H_0, H, \beta_H)\) and \(w'(H_0, H, \beta_H)\) are given respectively by

(1.5) \[ w(H_0, H, \beta_H) : \phi_H \geq \lambda \phi_{H_0}, \quad w'(H_0, H, \beta_H) : \phi_H < \lambda \phi_{H_0}, \]

where \(\lambda\) is determined by \(P(x \in w(H_0, H, \beta_H) \setminus H_0) = \beta_H\). It can be shown that

(1.6) \[ P(H_0, H, \beta_H) > \beta_H \]

**Proof.** Assume that \(\phi\) is such that \(w\) defined by (1.5) is unique. Integrating the first inequality of (1.5) over \(w(H_0, H, \beta_H)\) and the second one over \(w'\), we have respectively, \(P(H_0, H, \beta_H) \geq \lambda \phi_{H} \) and \(1 - P(H_0, H, \beta_H) < \lambda(1 - \phi_{H})\), from which, after a slight reduction, we have (1.6). Note that in general \(\lambda\) will be of the form \(\lambda(H_0, H, \beta_H)\), depending on all the elements. Incidentally, any critical region of size \(\beta\) for \(H_0\), whose power with respect to an alternative \(H\) is greater than or equal to \(\beta\), will be called an unbiased critical region for \(H_0\) against \(H\).

Alongside the more common terminology, namely the most powerful test of \(H_0\) against \(H\), a locally most powerful test of \(H_0\) (in those situations where it is meaningful), a uniformly most powerful test of \(H_0\) (if it exists at all with respect
to the whole relevant class of alternatives) we shall use the less common terminology, namely, an unbiased test of $H_0$ against $H$, a locally unbiased test and a uniformly unbiased test. The result (1.6) shows that a most powerful test of $H_0$ against $H$, a locally most powerful test of $H_0$ and a uniformly most powerful test of $H_0$ are also respectively an unbiased test of $H_0$ against $H$, a locally unbiased test and a uniformly unbiased test. Of course, in general, unbiased tests will be a much larger class of which the most powerful test will be just a member.

The likelihood ratio critical region at a level, say $a$, of a simple $H_0$ against the whole class of simple $H \in \mathcal{T}$, provided that it exists, will be denoted by $w(a, H_0)$. As is well known it is given by

$$w(H_0, a) : \phi(x) \geq \mu(H_0, a) \phi_{H_0}(x),$$

where for a given $x$, $\phi(x)$ stands for the largest $\phi_H(x)$ (provided that it exists) with respect to variation of $H$ over $\mathcal{T}$, and where $\mu(H_0, a)$ is given by

$$P(x \in w(H_0, a) \mid H_0) = a.$$  

Notice that $\phi(x)$ is a function of $x$ only, being independent of $H \in \mathcal{T}$, but may depend on the total domain $\mathcal{T}$. The power of this test, against any particular alternative $H \in \mathcal{T}$, will be denoted by $P(H_0, H, a_H)$.

Assume now that $H_0$ is a composite hypothesis and $H \in \mathcal{T}$ a composite alternative. In earlier papers $\int 18, 19 \int$ the author gave a set of sufficient conditions on $\phi_{H_0}$ for the availability of similar regions for $H_0$, and a set of (further) restrictions on $\phi_H$ and $\phi_{H_0}$ for the availability, among those similar regions, of one which is the most powerful for $H_0$ against $H$ in the following sense. Suppose that $H_0$ and $H$ are composite hypotheses, each characterized by some specified and some unspecified elements, so that, if the unspecified elements were specified, both $H_0$ and $H$ would be simple hypotheses. Now suppose that, among the similar regions for $H_0$, there is one whose location in the sample space depends on the specified elements of $H_0$ and possibly on those of $H$, but not on the unspecified elements of $H_0$ or $H$, but which is nevertheless the most powerful critical region for any simple hypothesis.
within $H_0$ (obtained by specifying the unspecified elements) against any simple alternative within $H$ (obtained by specifying the unspecified elements). But this "most powerful" is "most powerful among similar regions". If we drop the restriction of similarity and set up in a straightforward manner the most powerful critical region for the simple hypothesis in question against the simple alternative in question, then we may get a (non-similar) region having a larger power than that of the most powerful similar critical region just referred to. Such a most powerful similar critical region may be conveniently called a bisimilar region for $H_0$ against $H$. The likelihood ratio critical region for composite $H_0$ against all composite $H \in \mathcal{C}$ (which we know how to construct, provided that it exists), can be shown to be a similar region for $H_0$, under the restrictions just referred to. In this situation the same notation will be used as introduced in the previous paragraph for the case of a simple hypothesis against simple alternatives, and the result (1.8) will also hold, it being noted that, while the regions will be independent of the unspecified elements in $H_0$ and $H$, $P(H_0, H, \beta_H)$ and $P(H_0, H, \alpha_H)$ however, might depend on the unspecified elements of $H$ though not on those of $H_0$.

2. General scope and objectives of this report. Throughout this report we shall restrict ourselves to very limited objectives, namely solution of certain non-sequential, i.e., fixed sample size two-decision problems, in which, for a preassigned level $\alpha$ or a confidence coefficient $1 - \alpha$, we are interested respectively in obtaining (i) a (similar) region test of $H_0$ which has some kind of reasonably 'good' property against the whole class of relevant (composite) alternatives $H (e \mathcal{C})$ or (ii) a set of simultaneous confidence bounds on deviations from $H_0$, naturally occurring in the problems to be considered (all to be explained later), the confidence bounds, again, having some kind of 'good' properties in terms of covering 'wrong' values of the deviations. The scope of the discussion is thus professedly quite narrow and by no means adequate for the needs of any possible user of statistics, but that is as far as we can get at the moment. It is hoped that, in the near future,
methods and techniques will develop perhaps in extension of those offered here, which can cope with the more recondite problems that are of real interest to the possible users of statistics.

Toward these limited objectives, a heuristic method of test construction will be offered which leads to a certain class of tests including, in particular, two members of special importance to be called respectively type I and type II tests and a modified form of type I test, to be called an extended type I test. Type II test will be identified with the widely known likelihood ratio criterion, but it is the type I and the extended type I test that will be used throughout this report, and, in the specific situations to be considered, it will be possible, in every case, to obtain, by inversion of these tests, suitable confidence bounds on certain deviations or measures of departure from the hypothesis that naturally arise in the case considered. As observed at the outset the general method is entirely heuristic and, therefore, the test or the set of confidence bounds that emerges as the end product, in any specific problem, has to be justified by its operating characteristics in that situation, no 'good' properties being guaranteed in advance by the general method of test construction itself.

3. A class of tests including, in particular, the type I and type II tests /21/.  

3.1. Definitions and some remarks. Consider, for simplicity but without any essential loss of generality (for the definitions could be immediately carried over into the case of composite hypothesis and alternative) a simple hypothesis \( H_0 \) against a simple alternative \( H \in (\) .

(1) Put \( H = H_0 \), and set up as the rejection and acceptance regions for \( H_0 \) \( \cap \) \( H \subseteq \) \( \cap \) \( W(H_0, H, \beta) \) and its complement \( \cap \) \( H \subseteq \) \( \cap \) \( W(H_0, H, \beta) \), to be called respectively \( U_H \) and \( \cap_H \). This is defined to be a type I test for \( H_0 \) against the whole class \( H \in (\) , the level of significance \( \alpha \) being given by

\[
(3.1.1) \quad P(x \in \cap \cap W(H_0, H, \beta) \mid H_0) = a(H_0, \beta) \quad (> \beta) .
\]

Let us for the moment assume non-triviality, that is, that, given \( \alpha < 1 \), we can find
\( \beta = \beta(H_0, \alpha) > 0 \), for which (3.1.1) will hold.

(ii) Put, in section 1, \( \lambda(H_0, H, \beta_H) = \mu \) (a preassigned constant) for all \( H \in \mathcal{L} \) and rewrite \( w(H_0, H, \beta_H) \) and \( w'(H_0, H, \beta_H) \) as \( w^*(H_0, H, \beta_H) \) and \( w'^*(H_0, H, \beta_H) \) respectively.

Now set up, as the rejection and acceptance regions for \( H_0, U_H^*w^*(H_0, H, \mu) \) and its complement \( \bigcap_H w'^*(H_0, H, \mu) \), to be called, respectively, \( U_H^* \) and \( \bigcap_H^* \), where the \( \beta_H \)'s \( (H \in \mathcal{L}) \) are subject to \( \lambda(H_0, H, \beta_H) = \mu \) (a preassigned constant). This is defined to be a Type II test for \( H_0 \) against the whole class \( H \in \mathcal{L} \) the level of significance \( \alpha^* \) being given by

\[
(3.1.2) \quad P(x \in U_H^*(H_0, H, \mu) \mid H_0) = \alpha^*(H_0, \mu).
\]

Here again let us, for the moment, assume non-triviality, that is, that given \( \alpha^* (< 1) \), we can find a \( \mu \) such that \( \alpha(H_0, H, \mu) = \beta_H \) \( (> 0) \) and that (3.1.2) will hold. This can be easily recognized as the likelihood ratio test by the following consideration.

Notice that \( w^*(H_0, H, \mu) \) (with a preassigned \( \mu \)) is given by

\[
(3.1.3) \quad w^*(H_0, H, \mu) : \varphi_H(x) \geq \mu \varphi_{H_0}(x).
\]

Any \( x \) would belong to \( U_H^*w^*(H_0, H, \mu) \) if for that \( x \), there were at least one \( H \in \mathcal{L} \) for which (3.1.3) holds. It is easy to see that this would be accomplished if for that \( x \) the largest \( \varphi_H(x) \) (under variation of \( H \) over \( \mathcal{L} \)) were \( \geq \mu \varphi_{H_0}(x) \). Hence it is obvious that

\[
(3.1.4) \quad U_H^*w^*(H_0, H, \mu) : \varphi(x) \geq \mu \varphi_{H_0}(x), \quad \bigcap_H^*w^*(H_0, H, \mu) : \varphi(x) \leq \mu \varphi_{H_0}(x).
\]

3.2. An obvious property of the two types of tests. Notice that \( U_H \) includes all \( w(H_0, H, \beta) \) and \( U_H^* \) all \( w^*(H_0, H, \mu) \). Now putting

\[
P(x \in U_H \mid H) = P(U_H, H, \alpha) \quad \text{and} \quad P(x \in U_H^* \mid H) = P(U_H^*, H, \alpha)
\]

we shall have from Section (3.1) for the two types of tests

\[
(3.2.1) \quad \beta(H_0, \alpha) = \beta < P(H_0, H, \beta) \leq P(U_H, H, \alpha) \leq P(H_0, H, \alpha) \leq 1
\]

\[
P(H_0, H, \alpha) > \alpha
\]

\[
(3.2.2) \quad \alpha^*(H_0, H, \alpha) = \alpha^* < P^*(H_0, H, \mu) \leq P(U_H^*, H, \alpha) \leq P(H_0, H, \alpha) \leq 1
\]

\[
P(H_0, H, \alpha) > \alpha.
\]
(3.2.1) and (3.2.2) give respectively, for all \( H \in \mathcal{H} \), the lower bounds \( P(H_0, H, \theta) \) and \( P^*(H_0, H, \mu) \) for \( P(U^H_H, H, \alpha) \) and \( P(U^*_H, H, \mu) \), which, however, in general, would be far from close except sometimes for large "deviations" from \( H_0 \). With more knowledge of the forms of \( \phi_{H_0} \) and \( \phi_H \), it is often possible to get far closer bounds; even the actual powers are often computable without much difficulty (and turn out to be pretty high) as for example in most of the classical tests on normal populations.

It is easy to see that the results of (3.1) and (3.2) could be easily generalized to cover the case of composite \( H_0 \) against composite \( H \in \mathcal{H} \) provided that we have similar regions for \( H_0 \) and a bisimilar region for \( H_0 \) against \( H \). This, therefore, need not be separately treated.

3.3. Display of two classical tests as Type I tests. (i) Almost all classical tests on univariate and multivariate normal populations, (ii) most classical tests on other types of populations and (iii) many tests on multivariate normal populations proposed in recent years are known to be derivable (and indeed many of them have, in fact, been derived) from the "likelihood ratio" principle, so that they belong to Type II. The author finds that all the customary tests in category (i), for example, the test of significance of (1) a mean, (2) a mean difference, (3) total or partial or multiple correlation, and (4) regressions, (5) the F-test in analysis of variance, (6) the test based on Hotelling's \( T^2 \), all belong to Type I as well. Those classical tests in category (ii) that the author has examined so far also all belong to Type I. Coming to those situations that are sought to be handled by tests proposed under category (iii), the author finds that the likelihood ratio tests offered so far, while they automatically belong to Type II, do not belong to Type I. On the other hand, if, in these situations, one carries out the spirit and method of discriminant analysis, one gets tests which belong to Type I in a sense slightly more general than we have indicated so far.

In this section we consider, for illustration, two well known classical tests and show that they belong to Type I.
(i) For \( N(\xi_1, \sigma^2) \) and \( N(\xi_2, \sigma^2) \) the classical test of \( H(\xi_1 = \xi_2) = H_0 \) against \( H(\xi_1 \neq \xi_2) = H \) at a level \( \alpha \) is based on a critical region given by

\[
t \geq t_0 \quad \text{or} \quad t \leq -t_0,
\]

where \( t = (n_1 + n_2 - 2)^{1/2} \left\{ \frac{n_1 n_2}{(n_1 + n_2)} \right\}^{1/2} \left( \frac{\bar{x}_1 - \bar{x}_2}{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2} \right)^{1/2}, \)

and \( t_0 \) is given by \( P(t \geq t_0 | H_0) = \alpha/2 \) and where \((\bar{x}_1, \bar{x}_2), (s_1, s_2)\) stand for the means and standard deviations of two random samples of sizes \( n_1 \) and \( n_2 \) drawn from \( N(\xi_1, \sigma^2) \) and \( N(\xi_2, \sigma^2) \), respectively. This is well known as a likelihood ratio test but it is easily checked as Type I as well, in the following way. It is well known that \( t \geq t_0 \) is a one-sided uniformly most powerful (bisimilar) region of size \( \alpha/2 \) for the composite \( H_0 \) against the composite \( H(\xi_1 > \xi_2) = H_1 \) and so also is \( t \leq -t_0 \) for \( H_0 \) against \( H(\xi_1 < \xi_2) = H_2 \); taking the union we have (3.3.1) of size \( \alpha \).

(ii) Consider the testing of a general linear hypothesis in analysis of variance which, as is well known, can be formally reduced to the following. Suppose we have random samples of sizes \( n_i \), means \( \bar{x}_i \) and standard deviations \( s_i \), drawn respectively from \( N(\xi_h, \sigma^2) \) \((h = 1, \ldots, k)\), and suppose we want to test \( H(\xi_1 = \xi_2 = \ldots = \xi_k) = H_0 \) against the whole class \( H \) of \((\xi_1, \ldots, \xi_k)\) violating \( H_0 \). Put \( n = \sum_{h=1}^{k} n_h \); \( \bar{x} = \sum_{h=1}^{k} \frac{n_h \bar{x}_h}{n_h} \); \( \bar{\xi} = \sum_{h=1}^{k} \frac{n_h \xi_h}{n_h} \). Now the classical F-test for \( H_0 \), which is well known to be a likelihood ratio of Type II test has at a level \( \alpha \) the critical region given by

\[
F \geq F_0
\]

where \( F = \sum_{h=1}^{k} \frac{n_h (\bar{x}_h - \bar{x})^2}{(k-1)} / \sum_{h=1}^{k} \frac{(n_h - 1)s_h^2}{(n_h - k)} \) and where \( F_0 \) is given by

\[
P(F \geq F_0 | H_0) = \alpha.
\]

To recognize this as a Type I test as well we proceed as follows. It was observed in earlier papers \( \int_{18}^{7}, \int_{8}^{7} \) that among similar regions for \( H_0 \) (which exist) there is a most powerful (bisimilar) region for \( H_0 \) against any specific \((\xi_1, \ldots, \xi_k) = \xi \) violating \( H_0 \), the region of size, say, \( \beta \) being given by

\[
t \geq t_0
\]
where \( t = \sqrt{n-2} \cot \theta \) and \( \cos \theta = \sum_{h=1}^{k} n_h (\bar{x}_h - \bar{x}) (e_{h}\xi^2) / \left\{ \sum_{h=1}^{k} n_h (\bar{x}_h - \bar{x})^2 + (n_h - 1)s_h^2 \right\}^{1/2} \times \sum_{h=1}^{k} n_h (e_{h}\xi^2)^{1/2} \) and where \( t_0 \) is given by \( P(t \geq t_0 / H_0) = \beta \). It was also noticed in those papers that this \( t \) has exactly the usual \( t \)-distribution with \( n-2 \) degrees of freedom. Notice that \( t_0 = t_0(n, \beta) \) and \( \beta = \beta(n, t_0) \). To obtain now the union of regions: \( t \geq t_0 \) over different sets of \( (\xi_1, \ldots, \xi_k) \) we note that a given set there was at least one \( t \) such that \( t \geq t_0 \). The union is thus easily checked to be given by: the largest \( t \) (by varying over \( \xi_1, \ldots, \xi_k \)) \( \geq t_0 \) (which is fixed). But by (3.3.3) the largest \( t \) would correspond to the largest value of \( \cos \theta \), and, given \( \bar{x}_i \)'s and \( s_i \)'s, the largest value of \( \cos \theta \) (under variation over \( \xi_1, \ldots, \xi_k \)) is easily seen to be given by:

\[
(3.3.4) \quad \cos \theta = \sum_{h=1}^{k} n_h (\bar{x}_h - \bar{x})^2 \sum_{h=1}^{k} (n_h - 1) s_h^2 + n_h (\bar{x}_h - \bar{x})^2 \sum_{h=1}^{k} \frac{1}{s_h} \sum_{h=1}^{k} \frac{1}{s_h}.
\]

so that the largest \( t \) is given by

\[
(3.3.5) \quad t = (n - 2)^{1/2} \sum_{h=1}^{k} n_h (\bar{x}_h - \bar{x})^2 \sum_{h=1}^{k} (n_h - 1) s_h^2 \sum_{h=1}^{k} \frac{1}{s_h}.
\]

Therefore the union of regions: \( t \geq t_0 \), is given exactly by (3.3.2), which is the critical region of the \( F \)-test. Notice that given the \( a \) of the \( F \)-test, \( F_0 \) is obtained from (3.3.2) in the form \( F_0(k - 1, n - k; a) \); and next by identifying the union of regions \( t \geq t_0 \), with \( F \geq F_0 \) we have

\[
t_0 = \sqrt{(k - 1)(n - 2)F_0/(n - k)} \sum_{h=1}^{k} (n - k) t_0(k - 1, n - k; a);
\]

and next from (3.3.3) we have

\[
\beta = \beta(n, t_0) = \beta(k - 1, n - k; a).
\]

3.4. Some further remarks on the two types of test. It may be noted (See Section 3.1) that by specializing the \( \beta \)-'s (the sizes of the most powerful critical regions against different alternatives) in two special ways we get in a heuristic manner the two types of test. By specializing the \( \beta \)-'s in other ways other heuristic principles could be set up, some of which, in special situations, might be "better" than
the Type I or Type II tests. It has already been observed that in many situations Type I and Type II tests would coincide. This does not mean, however, that in those situations, $\beta(H_0, H, \alpha)$ of the Type II test would be the $\beta$ of the Type I test.

Given $H_0$ and the $H$'s, it would be possible to find a $\beta$ for Type I and a $\mu$ for Type II such that the same critical region for $H_0$ against the whole class $H \in (\subseteq)$ could be looked upon as $U_H w(H_0, H, \beta)$ in relation to the first type and also as $U_H w^*(H_0, H, \mu)$ in relation to the second type.

The following theoretical question or group of questions now under investigation is extremely important. Under what general restrictions on the probability law of $x$ and on $H_0$ and $H \in (\subseteq)$ would either or both of the tests be nontrivial (in the sense discussed in Section 3.1) and usable (in the sense of having a distribution problem amenable to tabulation), and unbiased (against all relevant alternatives) and/or admissible and/or reasonably powerful (in the sense of having not too bad a power against all relevant alternatives)? So far as the author is aware, these questions have not yet been adequately discussed in a general manner (let alone being answered) even for the likelihood ratio or Type II test (which has so long been extensively used in practice), and no attempt will be made in this report to discuss these questions. The advantage, however, of having two such heuristic principles (with the possibility of having two different tests in many situations) is that it gives us more elbow room than we would have had with one such principle, in the matter of construction of nontrivial, usable and "pretty good" tests.

One remark on the admissibility of a test (in the Neyman-Pearson set-up) is especially important. In this set-up suppose we have a hypothesis $H_0$ and a class of alternatives $H \in (\subseteq)$. Assume, for simplicity of discussion, that $H_0$ and each $H$ are simple hypotheses. Now suppose that there is any critical region of size, say $\alpha$, for $H_0$. $w_0$ will be said to be inadmissible (or admissible) against the whole class $H \in (\subseteq)$ according as we can find (or fail to find) another critical region of size $\alpha$, say $w_1$, such that
(3.4.1) \[ P(x \in w_1 \mid H) \geq P(x \in w_0 \mid H) \quad \text{for all } H \in \{\cdot\}, \text{ and} \]
\[ P(x \in w_1 \mid H) > P(x \in w_0 \mid H) \quad \text{for at least one } H \in \{\cdot\}. \]

Suppose now that \( w_0 \) is an inadmissible critical region in that we can find a \( w_1 \) satisfying (3.4.1) and, assume for simplicity of discussion that \( w_1 \) itself is admissible. It is easy to satisfy oneself that from any physical point of view \( w_1 \) is better than \( w_0 \). Suppose now that \( w_2 \) is another critical region for \( H_0 \) of size \( \alpha \), which is admissible against all \( H \in \{\cdot\} \). It does not follow from the definition of admissibility that \( w_2 \) will necessarily have the property (3.4.1) in relation to \( w_0 \). On the contrary it may well be that

(3.4.2) \[ P(x \in w_2 \mid H) < P(x \in w_0 \mid H) \quad \text{for most } H \in \{\cdot\}, \text{ and} \]
\[ P(x \in w_2 \mid H) \geq P(x \in w_0 \mid H) \quad \text{for some } H \in \{\cdot\}, \text{ and} \]
\[ P(x \in w_2 \mid H) > P(x \in w_0 \mid H) \quad \text{for some } H \in \{\cdot\}. \]

A precise definition of 'most' need not detain us here. In fact, if a most powerful critical region of \( H_0 \) against a specific \( H \in \{\cdot\} \) is most powerful in the strict sense of having a power against \( H \), which is \( > \) and not just \( \geq \) that of any other rival, then this critical region will be, by definition, an admissible one against the whole class of \( H \in \{\cdot\} \). But it may have a poor power against most other alternatives. In other words, it is easy to convince ourselves that a particular inadmissible region may, from any physical point of view, be much better than many admissible regions, although there must be at least one admissible test (and usually a whole subclass of such tests) which satisfies (3.4.1) with respect to \( w_0 \) and is thus better than \( w_0 \) from any physical point of view. This is a point which is apt to be missed by the statistician, especially the theoretical statistician.

3.5. On the operating characteristics of certain specific tests. It turns out that

in many specific situations (as in the cases to be discussed herein) it is possible
to obtain a class of admissible critical regions for \( H_0 \) against all \( H \in \{\cdot\} \), each
region having a power which is a function of certain parameters which are naturally interpreted as measures of deviation from \( H_0 \). This admissible class may not of
course constitute the totality of all admissible critical regions. Now among this class, if there is a subclass which is not only unbiased against all \( H \in \mathcal{S} \) but is such that the power of each is a monotonically increasing function of each of the 'deviations', then this subclass is, from any physical point of view, the really valuable subset and will be said to be an admissible, unbiased subset having the monotonicity property. In situations where this is available and all that we know about \( H \) is that \( H \in \mathcal{S} \), the rest of the admissible class may, for most purposes, be thrown out. It seems to the author that in such situations, this subset or subclass of critical regions is the best that we can obtain as a whole and any further attempt at any choice among this subclass, on the basis of some stronger optimum property or principle would be open to controversy in that the selection principle would be likely to be artificial and not universally convincing. The author is aware of the asymptotic optimum properties of the likelihood ratio criterion for simple hypotheses, but even aside from the question as to how far they go over into the case of composite hypotheses, there are strong reasons to suppose that these asymptotic optimum properties are not peculiar to the likelihood ratio criterion but must be shared by a large class of criteria or critical regions. Where \( H_0 \) is composite there is the further restriction of similarity which, of course, can be relaxed by just requiring that any critical region should have a size \( \leq \alpha \ (\leq 1) \) under variation of the unspecified elements of \( H_0 \), in which case the region will be said to be a valid one, a special case of a valid region being a similar region. In any actual situation (usually involving a composite \( H_0 \)), if we can find a similar (or valid) critical region which is (i) unbiased against all \( H \in \mathcal{S} \), (ii) has the monotonicity property and is also (iii) admissible, then we shall consider this to be a satisfactory region and any attempt at getting a region with a stronger optimum property would, in most practical situations, be futile for reasons already indicated. However, if, as in most of the situations to be discussed herein, we have a number of rival regions available satisfying (i)-(iii), then it is no doubt an interesting and useful question as to how the powers of the different rivals
compare over the whole range of \( H \in \mathcal{L} \), one rival being better than another over some part of the range with a reversal in another part of the range and so on. In most of the rather complex situations to be discussed in this course this would not be possible, because not only are the actual powers not available, but we do not even have, at the moment, methods and techniques of comparing powers (in the sense of greater or less) of two rivals without actually obtaining the powers. It is hoped that such techniques will be available in the near future. It may be noticed here that quite often it is possible to assert properties (i) and (ii) and sometimes also (iii) without explicitly obtaining the power functions. It may also be observed that among similar (or valid) regions satisfying (i)-(iii) an additional consideration for recommendation might be (iv) reasonable simplicity of the null distribution problem, i.e., the distribution problem under \( H_0 \). If we are also interested, as we shall be in all the problems herein, in simultaneous confidence statements on deviation parameters (or functions thereof) then another additional consideration would be (v) the possibility of inverting the test to obtain (without running into excessively difficult distribution problems) such simultaneous confidence bounds (preferably intervals).

It will be seen that the tests offered in this report are similar region tests (in fact, they will be shown to be stronger than that, in a sense to be explained hereafter) having properties (i), (ii), (iv) and (v). There are strong grounds for believing (although we do not yet have a rigorous proof except for the degenerate special cases which will be indicated as we get along) that the tests also satisfy (iii). Furthermore the tests that are being offered for the different situations are such that it has been possible to obtain for each test a 'pretty good' (easily available) lower bound to the power function (and consequently a lower bound to the shortness, i.e., the probability of covering wrong values of the parameters or parametric functions, of the associated set of simultaneous confidence intervals), 'pretty good' in the sense that the lower bound itself is reasonably large and
rapidly goes up as the deviations increase. To the tests considered herein there are certain rivals (better known but not discussed in this report for reasons indicated at the proper places) for which some of the above properties are well known to be true and some of the others are also conjectured by the author to be true, but have not yet been proved.

4. **Extended Type I test (and an obvious property of it)**. Consider a composite hypothesis $H_0$ against a set of composite alternatives $H, i \in \mathbb{C}$, $(i \in \text{continuum})$. It often happens, as for example in the broad situations discussed in Section 6, that, while there are similar regions for $H_0$, there is among these no most powerful (b(similar) region for $H_0$ against any $H_i$ $(i \in \text{continuum})$, but that we have instead the following situation. Suppose we have composite hypotheses $H_{0j}$ $(j \in \text{continuum})$ such that $\bigcap_j H_{0j} = H_0$ and composite alternatives $H_{ij} (i \in \text{continuum}; j \in \text{continuum})$ such that $\bigcap_i H_{ij} = H_i$. Notice that $H_{0j}$ and $H_{ij}$ have more unspecified elements than $H_0$ and $H_i$ respectively. It may well be that we have (as in the cases discussed in Section 6) not only similar regions for $H_{0j}$ but also, among these, a most powerful (b(similar) region for $H_{0j}$ against any $H_{ij}$ (one for each $i$ with $j \in \text{continuum}$; and $i \in \text{continuum}$). Consider critical regions $w(H_{0j}, H_{ij}, \theta)$ of size $\theta$ each. Then by our test procedure, over $\bigcap_j \bigcap_i$ of $w(H_{0j}, H_{ij}, \theta)$ (which we call $\bigcap_{ji}$ for simplicity), we are anyway accepting $\bigcap_j H_{0j}$, that is, $H_0$ and over its complement $U_j U_i w(H_{0j}, H_{ij}, \theta)$ we are rejecting at least one $H_{0j}$ and therefore $H_0$ itself. Suppose we set this up as a heuristic test for $H_0$ against the whole class $H_i \in \mathbb{C}$. Then the critical region will be $U_j U_i w(H_{0j}, H_{ij}, \theta)$ or $U_{ji}$ of size $\alpha$, given by

\[ P(\mathbf{x} \in U_{ji} \mid H_0) = \alpha \]

so that $\alpha = \alpha(H_0, \theta)$ and $\theta = \beta(H_0, \alpha)$. As before, nontriviality will be assumed, and it is easy to check that we shall have for all $i$ and $j$ the following inequality:

\[ \beta < P(H_{0j}, H_{ij}, \theta) \leq P(U_{ji}, H_i, \alpha) \leq 1. \]

It may be noted that while $w(H_{0j}, H_{ij}, \theta)$, a bisimilar region of size $\theta$ for $H_{0j}$
against $H_{ij}$ is independent of the unspecified elements of $H_{0j}$ and $H_{ij}$ and while the location of $V_{ji}$ must be and its size might be (as indeed it is for all the cases considered in Section 6) independent of the unspecified elements of $H_{0j}$ and $H_{ij}$, the power $P(H_{0j}, H_{ij}, a)$ might involve the unspecified elements of $H_{ij}$ and $P(H_{0j}, H_{1j}, a)$ involve those of $H_{1j}$. As observed in Section 3.2, the lower bound to the power of the test, given by (4.2), while it is in general easily available, is, at the same time, much too crude. With more knowledge of the probability law a much closer lower bound can often be found as will be exemplified in later sections.

5. Random samples from $p$-variate normal populations. If $X(p \times m) = (x_1 \ldots x_m)^T$, where $x_\lambda$'s ($\lambda = 1, \ldots, m$) are an independent set and each $x_\lambda$ is $N(\xi, \Sigma)$ and $p < m$, then denoting by $\xi(p \times m)$ the matrix $(\xi \xi \ldots \xi)^T$, we have the following probability law for $X$:

$$
(5.1) \quad \int \frac{1}{(2\pi)^{m/2}} |\Sigma|^{m/2} \exp \left(-\frac{1}{2} \text{tr} \Sigma^{-1}(X - \xi)(X - \xi)^T\right) dX.
$$

The elements of $X$, of course, lie between $-\infty$ and $\infty$. If now we pass over from $X(p \times m)$ to $X_1(p \times m)$ by a transformation: $X(p \times m) = X_1(p \times m)\Lambda(m \times m)$, where $\Lambda$ is non-stochastic, then, by (A.1.6), $J(X : X_1) = 1$ and we have also $XX' = X_1X_1'$ (since $\Lambda$ is non-stochastic). Putting now $\bar{x}_1(1 \times p) = (\bar{x}_1, \ldots, \bar{x}_m)^T$, where $\bar{x}_i = \Sigma x_{i\lambda}/m$ ($i = 1, \ldots, m$), it is easy to see that

$$
(5.2) \quad \xi(\xi)(X' = X_1X_1').
$$

Hence

$$
(5.3) \quad (X - \xi)(X' - \xi') = XX' - \text{max}_{i = 1, \ldots, m} \xi_iX_1' \xi_i.
$$

If we now choose the transformation matrix (from $X$ to $X_1$) such that

$$
(5.4) \quad A = \begin{pmatrix}
\frac{1}{m} & \frac{1}{m} & \cdots & \frac{1}{m} \\
a_{21} & a_{22} & \cdots & a_{2m} \\
. & . & \cdots & . \\
a_{m1} & a_{m2} & \cdots & a_{mm}
\end{pmatrix} = \begin{pmatrix}
\frac{1}{m} & \frac{1}{m} & \cdots & \frac{1}{m} \\
. & . & \cdots & . \\
B(m - 1 \times m)
\end{pmatrix} (\text{say})
$$
we have

\[
(5.5) \quad X_1 = XA' = \begin{pmatrix}
\sqrt{m} x_1 \\
\sqrt{m} x_2 \\
\vdots \\
\sqrt{m} x_p
\end{pmatrix} A' = (A' X (n \times m) B' (m \times m-1)) = (\sqrt{m} x (n \times 1) Y (n \times m-1))
\]

(say). Thus

\[
(5.6) \quad XX' = X_1 X_1' = mxx' + YY',
\]

and hence substituting for \(XX'\) the right hand side of (5.3) becomes

\[
(5.7) \quad YY' + \frac{mxx'}{m} - \frac{mxx'}{m} + \frac{mxx'}{m} - \frac{mxx'}{m} \quad \text{or} \quad YY' + m(\bar{x} - \bar{x})(\bar{x}' - \bar{x}').
\]

Remembering now that \(J(X : X_1) = 1\) and transforming from \(X\) to \(X_1\) we have for \(X_1\) the probability law:

\[
(5.8) \quad \left(\frac{1}{2n}\right)^\frac{n}{2} |\Sigma|^\frac{n}{2} \exp \left( - \frac{1}{2} \text{tr} \Sigma^{-1} \left\{ YY' + m(\bar{x} - \bar{x})(\bar{x}' - \bar{x}') \right\} \right) d(\sqrt{m} x) dY.
\]

Let us put

\[
(5.9) \quad m-1 = n \quad \text{and} \quad \bar{X}(n \times m) = \begin{pmatrix}
\bar{x}_1 \\
\bar{X}_2 \\
\vdots \\
\bar{X}_p
\end{pmatrix} = \begin{pmatrix}
\bar{x}_1 \\
\bar{x}_2 \\
\vdots \\
\bar{x}_p
\end{pmatrix} = \begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_p
\end{pmatrix}
\]

and recall that for a sample \(X(n \times m)\), the sample dispersion matrix \(S\) or \(s_{ij}\) is defined by

\[
(5.10) \quad nS = n(s_{ij}) = (X - \bar{X})(X' - \bar{X}').
\]

It is easy to check that

\[
(5.11) \quad XX' = XX' = mxx', \quad \text{or} \quad (X - \bar{X})(X' - \bar{X}') = XX' - mxx',
\]

so that, using (5.6), (5.10), and (5.11) we have

\[
(5.12) \quad YY' = XX' - mxx' = (X - \bar{X})(X' - \bar{X}').
\]

We note that if, as in this case, the elements of \(X\) vary from \(-\infty\) to \(\infty\), so do those of \(X\) and \(Y\), to make the transformation one to one. Now integrating out (5.8) over \(X (-\infty \to \infty)\) we have for \(Y\) the probability law:

\[
(5.13) \quad \left(\frac{1}{2n}\right)^\frac{n}{2} |\Sigma|^\frac{n}{2} \exp \left( - \frac{1}{2} \text{tr} \Sigma^{-1} YY' \right) dY,
\]
and integrating out over $Y$ we have for $X$ the probability law:

$$(5.14) \quad (1/(2\pi)^{n/2} |\Sigma|^2) \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1}(X - \xi)(X' - \xi')^T \right\} d\Sigma(X)$$

or

$$(1/(2\pi)^{n/2} |\Sigma|^2) \exp \left\{ -\frac{m}{2}(X' - \xi')^T \Sigma^{-1}(X - \xi) \right\} d\Sigma(X),$$

which shows that $X$ is $N(\xi, 1/m \Sigma)$.

For the purpose of any study of the sample or population dispersion matrices, we could, without any loss of generality, start right off (as we will quite often do) from (5.13) replacing $Y$ by $X$, but with the understanding that now $X$ is not the original matrix of $p \times m$ observations, but is a part of the transformed matrix (considered under (5.5)), being $n \times n$ in structure. We shall customarily call this the reduced matrix.

For an $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ consisting of $m$ independent $((p+q) \times 1)$ column vectors, the reduced matrix $Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ ($n \times (p+q) \leq n$) will have the probability law:

$$
(5.15) \quad \begin{pmatrix} (p+q)n \\ n \end{pmatrix} \left( \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{pmatrix} \right)^{-1} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} dY_1 dY_2,
$$

where $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{pmatrix}$ is the partitioned population dispersion matrix (symmetric p.d.) for the $(p+q)$ normal variates.

For $k$ random samples of sizes $m_h$ from $\mathcal{N}(\xi_h, \Sigma)$, we have for $Y_h$'s and $\Sigma$, the joint probability law:

$$
(5.16) \quad \begin{pmatrix} m \\ m \end{pmatrix} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} \sum_{h=1}^k \left( \frac{Y_h Y_h^T}{m_h} + \sum_{i=h}^k m_i (X_i - \xi_h)(X_i - \xi_h') \right) \right\} \prod_{h=1}^k dY_h \prod_{h=1}^k d\Sigma_h(Xh),
$$
where $\bar{x}_{i} = \sum_{h=1}^{k} \bar{x}_{i h} / \bar{m}_{i}$, $n = \sum_{h=1}^{k} n_{h} = m - k$, $Y(p \times m)$

$$= (Y_1, Y_2, \ldots, Y_k)^T; \bar{x} = \sum_{h=1}^{k} m_{i h} \bar{x}_{i h} / \bar{m}_{i}, \xi = \sum_{h=1}^{k} m_{i h} \xi_{i h} / \bar{m}_{i},$$

with components $\bar{x}_{i} = \sum_{h=1}^{k} m_{i h} \bar{x}_{i h} / \bar{m}_{i}$,

$$\xi_{i} = \sum_{h=1}^{k} m_{i h} \xi_{i h} / \bar{m}_{i} (i = 1, \ldots, p),$$

and finally set

$$\bar{x}(p \times k) = \left( \begin{array}{c}
\bar{x}_{11} \cdots \bar{x}_{1k} \\
\cdots \cdots \cdots \\
\bar{x}_{p1} \cdots \bar{x}_{pk}
\end{array} \right) = \left( \begin{array}{c}
\bar{x}_{11} \cdots \bar{x}_{1k} \\
\bar{x}_{21} \cdots \bar{x}_{2k} \\
\vdots \cdots \vdots \\
\bar{x}_{pk} \cdots \bar{x}_{pk}
\end{array} \right)$$

and

$$\xi(p \times k) = \left( \begin{array}{c}
\xi_{11} \cdots \xi_{1k} \\
\cdots \cdots \cdots \\
\xi_{p1} \cdots \xi_{pk}
\end{array} \right) = \left( \begin{array}{c}
\xi_{11} \cdots \xi_{1k} \\
\xi_{21} \cdots \xi_{2k} \\
\vdots \cdots \vdots \\
\xi_{pk} \cdots \xi_{pk}
\end{array} \right).$$

Using now an $A(k \times k)$ of the structure

$$\left( \begin{array}{cccc}
\bar{m}_{11} & \bar{m}_{12} & \cdots & \bar{m}_{1k} \\
\bar{m}_{21} & \bar{m}_{22} & \cdots & \bar{m}_{2k} \\
\vdots & \vdots & \ddots & \vdots \\
\bar{m}_{k1} & \bar{m}_{k2} & \cdots & \bar{m}_{kk}
\end{array} \right) \quad \text{or} \quad \left( \begin{array}{cccc}
\bar{m}_{11} & \cdots & \bar{m}_{1k} \\
\bar{m}_{21} & \cdots & \bar{m}_{2k} \\
\vdots & \ddots & \vdots \\
\bar{m}_{k1} & \cdots & \bar{m}_{kk}
\end{array} \right)$$

(say)

and transforming from $X$ and $\xi$ to $X_1$ and $\xi_1$ such that

$$X_1 = \bar{X} A = (\bar{x} \bar{X} p \times 1) I(p \times k-1) \quad \text{and} \quad \xi_1 = (\bar{m} \xi) p \times 1) I(p \times k-1),$$

reminding that $\sum_{h=1}^{k} Y_{h} Y_{h}^T = YY^T$, and substituting in (5.16), it is easy to check for $X(p \times 1)$, $\xi(p \times k-1)$ and $Y(p \times n)$ the following probability law:

$$\left( 1/2 \pi \right)^{\frac{m}{2}} \left| \Sigma \right|^{-\frac{1}{2}} \int \exp \left( - \frac{1}{2} \Sigma^{-1} \{ Y Y^T + (Z-\xi)(Z-\xi)^T + m(X-\bar{x})(X-\bar{x})^T \} \right) dY dZ (\bar{m} \bar{x}).$$

As before, all elements of $(Y, Z, \xi)$ vary from $-\infty$ to $\infty$, and now integrating out over $X$, we have for $(Y, Z)$ the joint probability law

$$\left( 1/(2\pi) \right)^{\frac{m-1}{2}} \left| \Sigma \right|^{-\frac{m-1}{2}} \exp \left( - \frac{1}{2} \Sigma^{-1} \{ Y Y^T + (Z-\xi)(Z-\xi)^T \} \right) \int dY dZ.$$

Denoting by $(s_{ij})_{h, h}$, $\bar{x}_{i h}$, $\xi_{i h}$ the dispersion matrix of the $h$th sample, the mean of the $h$th sample for the $i$th variate and the mean of the $h$th population for the $i$th variate ($i, j = 1, 2, \ldots, p; h = 1, 2, \ldots, k$), we note that
\[ (5.22) \quad \mathbf{Y}' = \left( \sum_{h=1}^{k} n_h(s_{ij}) \right), \quad \mathbf{Z}' = \left( \sum_{h=1}^{k} m_h(\bar{x}_{ij} - \bar{x}_i)(\bar{x}_{jh} - \bar{x}_j) \right) \text{ and} \]

\[ \zeta \zeta' = \left( \sum_{h=1}^{k} m_h(\xi_{ih} - \bar{x}_i)(\xi_{jh} - \bar{x}_j) \right), \]

where all the elements of the right hand side are either defined explicitly in terms of the original set of observations or parameters or are directly calculable in terms of that set. We shall denote \( \left( \sum_{h=1}^{k} n_h(s_{ij}) \right)/n-k \) by \( S \) (to be called the sample \text{{"}within\text{"} dispersion matrix}), \( \left( \sum_{h=1}^{k} m_h(\bar{x}_{ik} - \bar{x}_i)(\bar{x}_{jk} - \bar{x}_j) \right)/k-1 \) by \( S^* \) (to be called the sample \text{{"}between\text{"} dispersion matrix}), \( \left( \sum_{h=1}^{k} m_h(\xi_{ir} - \bar{x}_i)(\xi_{jr} - \bar{x}_j) \right)/k-1 \) by \( \bar{S}^* \) (to be called the population \text{{"}between\text{"} dispersion matrix}) and the vectors \( \mathbf{x} \) and \( \bar{x} \) defined by (5.16) will be called respectively the sample and the population grand mean vector.

For \( k = 2 \) it is easy to check that \( \mathbf{Z(p x k-1)} \) and \( \zeta (p x k-1) \) become respectively the column vectors, say, \( z(p x 1) \) and \( \zeta (p x 1) \) given by

\[ (5.23) \quad z(p x 1) = n_{12}^{1/2} (\bar{x}_1 - \bar{x}_2), \quad \zeta (p x 1) = n_{12}^{1/2} (\bar{x}_1 - \bar{x}_2), \quad \text{where} \quad n_{12} = \frac{n_1 n_2}{n_1 + n_2} \]

and we have for \( \mathbf{X(p x (n_1+n_2))} \) and \( (\bar{x}_1 - \bar{x}_2) \) the probability law:

\[ (5.24) \quad \sqrt{1/(2n) \int \exp - \frac{1}{2} \text{tr} \Sigma^{-1} \left( \mathbf{Y}' + n_{12}^{1/2} (\bar{x}_1 - \bar{x}_2) \right) \left( \bar{x}_1 - \bar{x}_2 \right)^\prime \right) \frac{dY}{d\Sigma}, \]

\[ x \in \mathbb{R} \|

The regression set up we have an \( \mathbf{X(p x n)} \) \( (p < n) \) with a probability law

\[ (5.25) \quad \int 1/(2n) \| \Sigma \| \frac{n}{2} \exp - \frac{1}{2} \text{tr} \Sigma^{-1} (X - \xi)(X - \xi)^\prime \right) dX, \]

where \( \Sigma(p x p) \) is symmetric p.d. and \( \xi(p x n) = \mu(p x q)U(q x n) \), (when \( q < n \) but might be \( \geq p \) or \( < p \) and where \( \mu(p x q) \) is a given matrix of parameters and \( U \), a given non-stochastic matrix of rank \( q \)). As in (5.1)-(5.13), denote the column vectors of \( X \) and \( \xi \) by \( \mathbf{x}_j \) and \( \xi_j \) \( (j = 1, 2, \ldots, n) \), put \( \bar{x}(p x 1) = \frac{1}{n} \sum_{j=1}^{n} x_j / n \xi(p x 1) \).
\[ = \sum_{j=1}^{n} \xi_j / n, \text{ use the transformation } X_1 = XA = (\sqrt{n} \, \bar{X} : Y) \, p \text{ and } \xi_1 = (\sqrt{n} \, \bar{X} : \xi) \, p, \]

where \( A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \) is \( \perp \) and integrate out over \( \bar{X} \) and obtain for \( Y \)

the distribution

\[
\int_{\mathbb{R}^n} \sum_{j=1}^{n} \frac{n-1}{2} \left| \Sigma \right|^{-\frac{n-1}{2}} \exp \left\{-\frac{1}{2} \text{ tr } \Sigma^{-1} (Y - \xi)(Y - \xi)^\top \right\} \, dY,
\]

where we notice that \( YY' = XX' - n \, \bar{X} \bar{X}' \) and \( \xi' = \xi - n \, \bar{X} \, \bar{X}' \). Now denoting the column vector of \( U \) by \( u_j(q \times 1) \), putting \( \bar{u}(q \times 1) = \sum_{j=1}^{n} \frac{n}{n} u_j / n \), using the transformation \( U_1 = UA \) (where \( A \) is the same matrix as before) \( (\sqrt{n} \, \bar{u} : v) \, q \), we have

\[
\xi_j(p \times 1) = \mu(p \times q) u_j(q \times 1), \quad \xi(p \times 1) = \mu(p \times q) \bar{u}(q \times 1), \quad \xi = \mu V, \quad VV' = \bar{U} \bar{u} \bar{u}' \bar{u} \bar{u}',
\]

so that (5.26) can be rewritten as

\[
\int_{\mathbb{R}^n} \sum_{j=1}^{n} \frac{n-1}{2} \left| \Sigma \right|^{-\frac{n-1}{2}} \exp \left\{-\frac{1}{2} \text{ tr } \Sigma^{-1} (YY' - 2YV'\mu' + \mu VV'\mu') \right\} \, dY.
\]

Using (A.3.11) put \( V(q \times n-1) = \bar{f}(q \times q) L_1(q \times n-1) \) subject to \( LL' = I(q) \), using

(A.1.7) complete \( L_1(q \times n-1) \) by \( L_2(n-1-q \times n-1) \) so that

\[
\begin{pmatrix} L_1 \\ L_2 \end{pmatrix}_{n-q-1}^{n-1} \text{ is } \perp \text{. Next put}
\]

\[ Y(p \times n-1)(L_1^\top L_1^{n-1}) = (Z_1 \, Z_2)_{p} = Z \text{ (say) so that, by (A.1.6), } J(Y; Z) = 1. \]

Also note that \( VV' = \frac{\mu' \mu}{\bar{u} \bar{u}' \bar{u} \bar{u}'} \). We have now for \( Z_1(p \times q) \) and \( Z_2(p \times n-1-q) \) the distribution

\[
\int_{\mathbb{R}^n} \sum_{j=1}^{n} \frac{n-1}{2} \left| \Sigma \right|^{-\frac{n-1}{2}} \exp \left\{-\frac{1}{2} \text{ tr } \Sigma^{-1} (Z_1^2 + Z_2^2 + 2Z_1 Z_2 \bar{u}' \bar{u} + \mu' \mu) \right\} \, dZ_1 \, dZ_2,
\]

which shows that the joint distribution of \( Z_1 \) and \( Z_2 \) is exactly of the same form as of \( Z \) and \( Y \) in (5.21), \( \bar{u} \) there being replaced by \( n \) here and \( k-1 \) there being replaced by \( q \) here.
6. Statement of the specific problems to be discussed in this report. The problems will be formulated in terms of testing of hypothesis and in each case the associated problem in terms of simultaneous confidence interval estimation will also be indicated, although the latter will be discussed in full in sections 16.1 - 16.11. For each hypothesis to be considered here, the associated (set of) simultaneous confidence bounds will be referred to as A.S.C.B. It will be seen later that corresponding to each hypothesis and its class of alternatives (to be presently stated) there is a 'natural' and 'physically meaningful' set of parameters (or rather functions of the primitive population parameters) which can be easily interpreted as measures of deviations from the hypothesis. It will be also seen that the tests of hypotheses going to be offered here are such that, for each test, it is possible to obtain by inversion (and without running into any very difficult distribution problems) a set of simultaneous confidence bounds on these 'deviations' from the hypotheses. In this section, for most (though not for all) of the hypotheses stated, the structure of the corresponding 'deviations' are also stated without any attempt to show just why and how they are 'appropriate' or 'natural'; this is done later. The following arc the problems: (i) For \( N(\xi(p \times 1), \Sigma(p \times p)) \) (where \( \Sigma \) is symmetric p.d.), to test \( H_0: \Sigma = \Sigma_0 \) against \( H: \Sigma \neq \Sigma_0 \); the associated simultaneous confidence bounds, as will be seen later, will be bounds on characteristic roots of \( \Sigma \), i.e., on all \( \text{c}(\Sigma) \), or by using (A.2.5), bounds on \( a'(1 \times p) \Sigma(p \times p) a(p \times 1) \) (for all arbitrary vectors \( a'(1 \times p) \)) of unit length each; (ii) for \( N(\xi_h(p \times 1), \Sigma_h(p \times p)) \) \( (h = 1, 2) \) and \( \Sigma_1 \) and \( \Sigma_2 \) are both symmetric p.d., to test \( H_0: \Sigma_1 = \Sigma_2 \) against \( H: \Sigma_1 \neq \Sigma_2 \); the A.S.C.B. will be bounds on all \( \text{c}(\Sigma_1 \Sigma_2^{-1}) \), or using (A.2.6), on \( a'(1 \times p) \Sigma_1(p \times p) a(p \times 1)/a'(1 x p) \Sigma_2(p \times p) a(p \times 1) \) (for all arbitrary non-null \( a'(1 \times p) \)); (iii) for \( N(\xi_p(p \times 1), \Sigma(p \times p)) \) \( (r = 1, 2, \ldots, k) \); \( \Sigma \) is symmetric p.d., to test \( H_0: \xi_1 = \xi_2 = \ldots = \xi_k \) against \( H: \) not \( H_0 \), i.e., violation of at least one equality; if the random samples from the \( k \) populations are of sizes \( n_h \) \( (h = 1, 2, \ldots, k) \), then the A.S.C.B. will be on \( a'(1 \times p) \eta(p \times k) \eta(k \times 1) \) (for all arbitrary non-null \( a'(1 \times p) \)) and
arbitrary \( \mathbf{b}(k \times 1) \) of unit length, where \( \eta \) stands for the \( p \times (k-1) \) population matrix with \( k \) column vectors (each \( p \times 1 \)) \( \sqrt{n_1}(\xi_1-\xi), \sqrt{n_2}(\xi_2-\xi), \ldots, \sqrt{n_k}(\xi_k-\xi) \) and 

\[
\xi = \sum_{h=1}^{k} \frac{\eta_h \xi_h}{\eta_h \eta} \Sigma_{n_h} \eta_h. \quad \text{Notice that } \eta \text{ will be of rank } \leq \min (p,k-1); \quad \text{(iv) for} 
\]

\( N(\xi((p+q) \times 1), \Sigma((p+q) \times (p+q))) \), where \( \Sigma \) is symmetric p.d. of the form

\[
\begin{pmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{12}^T & \Sigma_{22}
\end{pmatrix}^{p \times q}, \quad \text{to test } H_0: \Sigma_{12}^{p \times q} = 0 \text{ against } \Sigma_{12} \neq 0; \quad \text{the A.S.C.B.}
\]

will be on \( a^T(1 \times p)\Sigma_{12}(1 \times q)\Sigma_{22}^{-1}(q \times q)\mathbf{b}(q \times 1) \) (for all arbitrary unit-length vec-

\( a^T(1 \times p) \) and \( \mathbf{b}(q \times 1) \) \( \sim \chi^2 \).

A number of useful problems can be formally tied up with problem (iii), of

which the more important are the following: (iii a) For \( N(\xi_h, \Sigma) \) (\( h = 1,2 \)), to test

\( H_0: \xi_1 = \xi_2 \text{ against } H_1: \xi_1 \neq \xi_2 \) the A.S.C.B. being now on \( a^T(1 \times p)(\xi_1-\xi_2)(p \times 1) \)

(for all arbitrary non-null \( a^T(1 \times p) \)); (iii b) for \( N(\xi, \Sigma) \), to test \( H_0: \xi = \xi_0 \)

against \( H_1: \xi \neq \xi_0 \) the A.S.C.B. being on \( a^T(1 \times p)\xi(p \times 1) \) (for all arbitrary non-

null \( a^T(1 \times p) \)); (iii c) given an observation matrix \( X(p \times n) \) of \( (p,p(p+1)/2 < n) \) of

stochastic variates with independent \( p \times 1 \) column vectors \( x_n \) (\( h = 1,2,\ldots,n \)) having

p.d.'s \( N(\mathbf{E}(x), \Sigma) \), where \( \mathbf{E}(X^T)(n \times p) = A(n \times m)\xi(m \times p) \) (\( m \leq n; \xi \) is a matrix of

unknown population parameters and \( A \) is a non-stochastic matrix of rank \( r \leq m \leq n 

whose elements are supposed to be given by the particular experimental situation),

to test \( H_0: C(q \times m)\xi(m \times p) = 0 \), where \( C \) is such that \( H_0 \) is testable (see \( (14.6.5) \))

against \( H_1: \eta \neq H_0 \); the A.S.C.B. will be given in section (16.6); this \( H_0 \) is called

the general multivariate linear hypothesis which includes the usual problems of

multivariate analysis of variance and covariance as particular cases and also of

course the problem (iii) as a very particular case; next (iii d) for \( N(\xi_h, \Sigma) \) (\( h = 1, \)

\( 2,\ldots,k) \), where \( \xi_h = \begin{pmatrix} \xi_{1h} \\ \xi_{2h} \end{pmatrix} 

\)

(say) and \( \Sigma \) is symmetric p.d. of the structure

\[
\begin{pmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{21} & \Sigma_{22}
\end{pmatrix}^{p \times q}, \quad \text{to test } H_0: \Sigma_{12} \text{ against } \Sigma_{12} \neq 0; \quad \text{the A.S.C.B.}
\]

will be on \( a^T(1 \times p)\Sigma_{12}(1 \times q)\Sigma_{22}^{-1}(q \times q)\mathbf{b}(q \times 1) \) (for all arbitrary unit-length vec-

\( a^T(1 \times p) \) and \( \mathbf{b}(q \times 1) \) \( \sim \chi^2 \).

A number of useful problems can be formally tied up with problem (iii), of

which the more important are the following: (iii a) For \( N(\xi_h, \Sigma) \) (\( h = 1,2 \)), to test

\( H_0: \xi_1 = \xi_2 \text{ against } H_1: \xi_1 \neq \xi_2 \) the A.S.C.B. being now on \( a^T(1 \times p)(\xi_1-\xi_2)(p \times 1) \)

(for all arbitrary non-null \( a^T(1 \times p) \)); (iii b) for \( N(\xi, \Sigma) \), to test \( H_0: \xi = \xi_0 \)

against \( H_1: \xi \neq \xi_0 \) the A.S.C.B. being on \( a^T(1 \times p)\xi(p \times 1) \) (for all arbitrary non-

null \( a^T(1 \times p) \)); (iii c) given an observation matrix \( X(p \times n) \) of \( (p,p(p+1)/2 < n) \) of

stochastic variates with independent \( p \times 1 \) column vectors \( x_n \) (\( h = 1,2,\ldots,n \)) having

p.d.'s \( N(\mathbf{E}(x), \Sigma) \), where \( \mathbf{E}(X^T)(n \times p) = A(n \times m)\xi(m \times p) \) (\( m \leq n; \xi \) is a matrix of

unknown population parameters and \( A \) is a non-stochastic matrix of rank \( r \leq m \leq n 

whose elements are supposed to be given by the particular experimental situation),

to test \( H_0: C(q \times m)\xi(m \times p) = 0 \), where \( C \) is such that \( H_0 \) is testable (see \( (14.6.5) \))

against \( H_1: \eta \neq H_0 \); the A.S.C.B. will be given in section (16.6); this \( H_0 \) is called

the general multivariate linear hypothesis which includes the usual problems of

multivariate analysis of variance and covariance as particular cases and also of

course the problem (iii) as a very particular case; next (iii d) for \( N(\xi_h, \Sigma) \) (\( h = 1, \)

\( 2,\ldots,k) \), where \( \xi_h = \begin{pmatrix} \xi_{1h} \\ \xi_{2h} \end{pmatrix} 

\)

(say) and \( \Sigma \) is symmetric p.d. of the structure
\[
\begin{pmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{12} & \Sigma_{22}
\end{pmatrix}^p, \text{ to test } H_0: \Sigma_{1h}(p \times 1) = \Sigma_{12}(p \times q)\Sigma_{22}^{-1}(q \times q)\Sigma_{2h}(q \times 1) \text{ (} h = 1, 2, \ldots, k \text{) against } H: \Sigma_{1h} \neq 0 \text{; not } H_0; \text{ the A.S.C.B. will be given in section (16.7); this } H_0 \text{ is called the hypothesis of (a particular kind of) multicollinearity of the means; and finally (iii c) for the linear regression model of (5.25) to test } H_0: \mu(p \times q) = 0 \text{ (or, say } \mu = \mu_0) \text{ against } H: \mu \neq 0 \text{ (or, say } \mu \neq \mu_0); \text{ the A.S.C.B. will be given in section (16.8).}
\]

Formally tied up with (iv) is the following: (iv a) for \( N(\xi, \Sigma) \) where \( \Sigma \) is symmetric p.d. of the form \( \Sigma = \begin{pmatrix}
\Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\
\Sigma_{12} & \Sigma_{22} & \Sigma_{23} \\
\Sigma_{13} & \Sigma_{23} & \Sigma_{33}
\end{pmatrix}^p, \) to test \( H_0: \Sigma_{12,3} = 0, \)

where \( \Sigma_{12,3}(p \times q) = \Sigma_{12}(p \times q) - \Sigma_{13}(p \times r)\Sigma_{33}^{-1}(r \times r)\Sigma_{23}(r \times q); \) the A.S.C.B. will be on \( \hat{a}'(1 \times p)\Sigma_{12,3}(p \times q)\Sigma_{22,3}^{-1}(q \times q)\hat{b}(q \times 1) \) (for all arbitrary unit length vectors \( \hat{a}'(1 \times p) \) and \( \hat{b}(q \times 1) \)), where \( \Sigma_{22,3}(q \times q) = \Sigma_{22}(q \times q) - \Sigma_{23}(q \times r)\Sigma_{33}^{-1}(r \times r)\times \Sigma_{23}(r \times q). \)

In addition to those considered in the two previous paragraphs there are several other problems whose solutions can be formally thrown back upon those of (i)-(v) and these need not be discussed or even stated separately here. But even within the very restricted set-up (considered in this report) of non-sequential, one stage, fixed-sample-size, two-decision problems of the classical type there are several problems of great practical and theoretical interest which have had to be excluded, because of the fact that (so far as the author is aware) no suitable and reasonably easy techniques are known at the moment. Among such problems (unfortunately to be omitted) a particularly important one is the following: for \( N(\xi_h, \Sigma_h) \) \( (h = 1, 2, \ldots, k > 2), \) to test \( H_0: \Sigma_1 = \Sigma_2 = \ldots = \Sigma_k \) against \( H: \not \Sigma_0 \); and of course the A.S.C.B. on 'appropriate deviations' from \( H_0. \)

In what follows section 7 will give the derivation of the proposed tests for
\( H_0 \) in the situations (i)-(v) and make the formal identification of (iii a)-(iii e) with (iii) and of (iv a) with (iv), sections 10 - 13 will give the operating characteristics of the proposed tests, section 16 will deal with all the set of simultaneous confidence bounds associated with each test of section 7, the operating characteristics of the proposed set of simultaneous confidence bounds in each case being easily available from sections 10-13.

7. Tests for \( H_0 \) in problems (i)-(iv) \[12.7\].

7.1. Direct type I construction not possible. It is well known \[18, 19\] that for each composite \( H_0 \) above there are infinitely many similar regions but no most powerful (bisimilar) region against any specific composite alternative, i.e., any composite alternative in which the specifiable elements are given special values. Thus direct type I construction will not work here.

7.2. Reduction to pseudo-univariate and pseudo-bivariate problems. At this point suppose that, starting from an \( \mathbf{x}(p \times 1) \) which is \( \mathcal{N}(\mathbf{\xi}, \Sigma) \) we consider a linear compound \( a'\mathbf{x} \) (with an arbitrary constant, i.e., non-stochastic \( a'(1 \times p) \) of non-zero modulus). This \( a'\mathbf{x} \) is a scalar well known to be \( \mathcal{N}(a'\mathbf{\xi}, a'\Sigma a) \). Notice that \( a'\mathbf{\xi} \) and \( a'\Sigma a \) are also scalars. Suppose also that given \( \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \)

\[
= \mathcal{N} \begin{bmatrix} (\xi_1)^	op p \\ (\xi_2)^	op q \\ \Sigma_{11} p \\ \Sigma_{21} q \\ \Sigma_{12} p \\ \Sigma_{22} q \\ 1 \\ p \\ q \end{bmatrix}
\]

we consider linear compounds \( a'x_1 \) and \( b'x_2 \) (where \( a(p \times 1) \) and \( b(q \times 1) \) are each non-null and non-stochastic); then these two scalars \( a'x_1 \) and \( b'x_2 \) are well known to be distributed as a bivariate normal with a correlation coefficient

\[
(7.2.1) \quad \rho(a, b) = \rho_{12} = \frac{a'\Sigma_{12}b}{\sqrt{(a'\Sigma_{11}a)^{1/2}(b'\Sigma_{22}b)^{1/2}}}.
\]

Now suppose that, in place of \( H_0 \) of (i)-(iv) of (7.1), we consider respectively

(v) \( H(a'\Sigma a = a'\Sigma a) (= H_0) \) against \( \forall \) \( H(a'\Sigma a \neq a'\Sigma a) (= H_0) \), (a fixed),
(vi) \( H(a\Sigma_2 = a\Sigma_0) (= H_{a_0}) \) against all \( H(a\Sigma_2 \neq a\Sigma_0) (= H_a) \), (a fixed),
(vii) \( H(a\Sigma_1 = a\Sigma_2 = \ldots = a\Sigma_k) (= H_{a_0}) \) against all \( H_a (\neq H_{a_0}) \), (a fixed),
(viii) \( H(a\Sigma_1b = 0) (= H_{0ab}) \) against all \( H(a\Sigma_1b \neq 0) (= H_{ab}) \), (a, b fixed).

We now consider the totality of all non-null \( a \) for (v)-(viii) and all non-null \( a \) and \( b \) for (viii). Notice that (a) \( \bigcap_{a} H(a\Sigma_2 = a\Sigma_0) = H(\Sigma = \Sigma_0) \),
(b) \( \bigcap_{a} H(a\Sigma_1a = a\Sigma_2a) = H(\Sigma_1 = \Sigma_2) \), (c) \( \bigcap_{a} H(a\Sigma_1 = a\Sigma_2 = \ldots = a\Sigma_k) = H(\Sigma_1 = \Sigma_2 = \ldots = \Sigma_k) \), and (d) \( \bigcap_{a,b} H(a\Sigma_1b = 0) = H(\Sigma_1 = 0) \). We could have worked in terms of any subset of \( a \)'s which led by intersection to the same \( H_0 \), but this we do not do here. It may be noted that by the procedure to be used here, apart from set-theoretic difficulties which, however, do not arise in these applications, the total set of \( a \)'s or any subset of it (of the kind considered) will uniquely define an extended Type I test associated with the total set or with that particular subset.

Next suppose that, in the alternative, under (v)-(viii), we substitute "specific" for "all" and thus have four new situations (ix)-(xii). It is well known that for each of the situations (ix)-(xii) we have one most powerful (bisimilar) region, so that from these we can construct respective (modified in a sense to be explained in section 7.3) type I tests for the pseudo-univariate situations (v) and (vi), Straight Type I tests for the pseudo-univariate situation (vii) and the pseudo-bivariate situation (viii). From these modified Type I and straight Type I tests we can try to construct the respective extended type I tests for the situations (i)-(iv). This ties up (see section 4) the \( p \)-variate problems (i)-(iii) with the pseudo-univariate problems (v)-(vii), the \( (p+q) \)-variate problem (iv) with the pseudo-bivariate problem (viii).

7.3. Modified type I tests for (v)-(vi) and type I tests for (vii)-(viii). We now take over the notation and symbols from section 5.

(v) Starting from (5.11) put \( x_a^2 = na\Sigma_a / a\Sigma_0a \) and notice that, at a level \( \theta_1 \), for \( H(a\Sigma_2 = a\Sigma_0a) (= H_{0a}) \) against all \( H(a\Sigma_2 > a\Sigma_0a) \) we have the one-sided uniformly most powerful (bisimilar) region: \( x_a^2 > x_{\theta_1}^2 \) (n), and for \( H_{0a} \) against all \( H(a\Sigma_2 < a\Sigma_0a) \)
we have the one-sided uniformly most powerful (bisimilar) region: \( x^2 \leq x^{'2}(n) \), where

\( x^2 \) and \( x^{'2}(n) \) are the upper \( a_1 \) and \( a_2 \) points of the \( x^2 \)-distribution with D.F. \( n \).

Notice that \( x^2 \) has the ordinary \( x^2 \)-distribution with D.F. \( n \). Now consider the union

\[ \bigcup_{a_1} x^2 < x^{'2}(n) \cup \bigcup_{a_2} x^2 > x^{'2}(n) = U(a) \]

say, which, if we decide to call it a new critical region, will be one of size \( a_1 + a_2 = \beta \) (say). Notice that given \( \beta \), we can regard \( a_1 \) and \( a_2 \) now as flexible, subject to \( a_1 + a_2 = \beta \). At this point, using (12.4), we can choose \( a_1 \) and \( a_2 \), i.e., the tail ends \( x^{'2}_{a_1} \) and \( x^{'2}_{a_2} \) as to make \( U(a) \) a locally unbiased (here it will turn out to be also locally most powerful critical region (in the neighborhood of \( H_{0_2} \)). It will be seen that the condition of unbiasedness imposes a relation between \( x^{'2}_{a_1} \) and \( x^{'2}_{a_2} \) which involves only \( n \) but is independent of the total size of the region \( \beta \). We now call these tail ends \( x^{'2}_{a_1}(p,n) \) and \( x^{'2}_{a_2}(p,n) \). We now recall from (16) that this \( U(a) \) is also a uniformly unbiased region (having, in fact, the stronger property of monotonicity) and is also admissible.

With this choice of \( x^{'2}_{a_1}(n) \) and \( x^{'2}_{a_2}(a) \) we have now for \( H(a^1z_a = a^1z_{0_a}) \) against all \( H(a^1z_a \neq a^1z_a) \) a modified type I critical region of size \( \beta \)

\[
\frac{x^2}{\beta} = \frac{a^1z_a}{a^1z_{0_2}} > x^{'2}_{a_1}(n) \text{ or } x^{'2}_{a_2}(n)
\]

which is uniformly unbiased, monotonic and is also admissible.

(vi) Starting from the product of two distributions like (5.1.3), put \( F_a = a^1S_{a^2}/a^1S_{0^2} \)

and notice as in the previous case that, at a level \( \alpha_1 \), for \( H(a^1z_a = a^1z_{0_a}) = H_{0_a} \)

against all \( H(a^1z_a > a^1z_a) \) we have the one-sided uniformly most powerful (bisimilar) region: \( F \geq F_{a_1}(n_1,n_1) \), and, for \( H_{0_2} \) against all \( H(a^1z_a < a^1z_{a_2}) \), the one-sided uniformly most powerful region: \( F \leq F_{a_2}(n_2,n_2) \), where \( F_{a_1}(n_1,n_1) \) and \( F_{a_2}(n_2,n_2) \) are the upper \( a_1 \) and lower \( a_2 \) points of the \( F \)-distribution with D.F. \( n_1 \) and \( n_2 \). Notice that \( F \) has the ordinary \( F \)-distribution with D.F. \( n_1 \) and \( n_2 \). Take the union of the two regions and as in the previous case call it a new critical region, say \( U(a) \) of
size $\beta_1 + \beta_2 = \beta$ (say), and given $\beta$, pick out (see 12.4) the tails $F_{\beta_1}(n_1, n_2)$ and $F_{\beta_2}(n_1, n_2)$ so as to make $U(a)$ a locally unbiased region (in the neighborhood of $H_{02}$), notice that this imposes an extra relation between $F_{\beta_1}(n_1, n_2)$ and $F_{\beta_2}(n_1, n_2)$ which involves only $n_1, n_2$ and not the total size of the region $\beta$. Recall also from (1.6) that this is a uniformly unbiased region (also having the monotonicity property) and also admissible. As before, with this choice of $F_{\beta_1}$ and $F_{\beta_2}$ to be called $F_{\beta_1}(n_1, n_2)$ and $F_{\beta_2}(n_1, n_2)$ we have now for $H(a'\Sigma a = a'\Sigma a)$ against all $H(a'\Sigma a \neq a'\Sigma a)$ a modified type I critical region of size $\beta$ (uniformly unbiased, monotonic and admissible)

\[(7.3.7) \quad F_{\beta} = \frac{a'S_{1}a}{a'S_{2}a} \geq F_{\beta}(n_1, n_2) \text{ or } \leq F_{\beta}(n_1, n_2).\]

(vii) Start from (5.16)-(5.22) and recall from (ii) of section 3.3 that for $H(a'\xi_1 = a'\xi_2 = \ldots = a'\xi_k) = H_{02}$ against any specific $H_{02} \neq H_{02}$, there is the most powerful (bisimilar) critical region (of size, say $\gamma$) which is a one-sided t-region, and by taking the union of these regions (for fixed $\gamma$ but by variation over $\xi_1, \xi_2, \ldots, \xi_k$), we have the straight type I region of size, say $\beta$, given by (notice that $F_{\beta}$ has the ordinary F-distribution with $D.F. n_1$ and $n_2$)

\[(7.3.3) \quad F_{\beta} = \frac{a'S_{a}^*}{a'S_{a}} \geq F_{\beta}(n_1, n_2),\]

where $F_{\beta}(n_1, n_2) = F_{\beta}$ (say) is obtained from $P(F_{\beta} \geq F_{\beta} \mid H_{02}) = \beta$.

This is well known to be a type II or likelihood ratio region as well and is also well known to have a number of desirable properties (including uniform unbiasedness, the stronger property of monotonicity and also admissibility).

(viii) Start from (5.15) and put

\[(7.3.4) \quad \left( \begin{array}{cc} S_{11} & S_{12} \\ S_{12} & S_{22} \end{array} \right)^p q = \frac{1}{n} \left( \begin{array}{cc} Y_1 \\ Y_2 \end{array} \right)^p q \cdot \begin{array}{c} p \\ q \end{array} \]

Next put

\[(7.3.5) \quad r_{ab} = \frac{a'S_{12}b}{(a'S_{11}a)^{1/2}(b'S_{22}b)^{1/2}}\]
and notice that, at a level \( \beta \), for \( H(a^1\Sigma_{12}b = 0) (= H_{0ab}) \) against all \( H(a^2\Sigma_{12}b > 0) \) we have the one-sided uniformly most powerful (bisimilar) region: \( r_{ab} \geq r_{\beta}(n-1) \) and for \( H_{0ab} \) against all \( H(a^2\Sigma_{12}b < 0) \) the one-sided uniformly most powerful (bisimilar) region: \( r_{ab} \leq -r_{\beta}(n-1) \), where \( r_{\beta}(n-1) (= r_{\beta}, \text{say}) \) is given by

\[
P(r_{ab} \geq r_{\beta} \mid H_{0ab}) = \beta.
\]

Notice also that this \( r_{ab} \) has the distribution of the central correlation coefficient with \( D.F. \) (n-1). Taking the union of the two regions we shall have a straight type I critical region of size 2\( \beta \) given by

\[
\bigg\{ r_{ab} \geq r_{\beta}(n-1) \bigg\} \cup \bigg\{ r_{ab} \leq -r_{\beta}(n-1) \bigg\},
\]

that is, \( r \geq r_0 \) or \( r^2 \geq r_0^2 \).

This is well known to be a type II or likelihood ratio region as well and it is also well known that this has a number of desirable properties (including uniform unbiasedness, the stronger property of monotonicity and also admissibility).

7.4. Actual construction of extended type I regions for the situations (i)-(iv).

(i) By the test procedure (7.3.1) over \( \chi^2_{1\beta}(n) \leq \chi^2_{a} \leq \chi^2_{2\beta}(n) \) we accept

\( H(a^1\Sigma_2 = a^1\Sigma_0a) \), so that, by using the heuristic principle of section 4, over

\[
\bigg\{ \chi^2_{1\beta}(n) \leq \chi^2_{a} \leq \chi^2_{2\beta}(n) \bigg\} \text{ we accept } \bigg\{ \chi^2_{a} \leq \chi^2_{1\beta}(n) \bigg\} \text{ we accept } \bigg\{ \chi^2_{a} \leq \chi^2_{2\beta}(n) \bigg\}
\]

we note that a particular \( S \) would belong to the intersection if, for that \( S \),

\[
\chi^2_{1\beta} \leq \frac{a^1Sa}{a^1\Sigma_0a} \leq \chi^2_{2\beta} \text{ for all non-null } a.
\]

This statement \( \iff \chi^2_{1\beta} \leq \text{ smallest } \frac{a^1Sa}{a^1\Sigma_0a} \leq \frac{a^1Sa}{a^1\Sigma_0a} \leq \chi^2_{2\beta} \), the "largest" and "smallest" being under variation of \( a \) (for a given \( S \)). Now, given \( S \), and of course \( \Sigma_0 \), the largest and smallest values of \( \frac{a^1Sa}{a^1\Sigma_0a} \) are easily seen from (A.2.5) to be the largest and smallest roots, say \( c_1 \) and \( c_p \) of the \( p \)th degree equation in \( c \):
(7.4.1) \[ |S - c\Sigma_0| = 0, \]

all the p roots \(c_1, c_2, \ldots, c_p\) being in this situation, a.c., positive, since \(\Sigma_0\) is
given to be symmetric p.d. and \(S\) is, by definition and the assumptions, a.c., p.d.
Starting out from the (modified) type I test (7.3.1) for \(H_{02}\), we have for \(H_0\), i.e.,
\(H(\Sigma = \Sigma_0)\) the extended type I critical region

\[(7.4.2) \quad c_p \geq x^2_{2p}(n) \text{ and/or } c_1 \leq x^2_{18}(n).\]

To find \(x^2_{2p}\) and \(x^2_{18}\) we make use of the condition of local unbiasedness (which
involves only \(n\)) (see (vi) of (7.3) and also 12.4) and write down the further condi-
tion (which now completely determines \(x^2_{2p}\) and \(x^2_{18}\))

\[(7.4.3) \quad F(x^2_{2p} \leq c_1 \leq c_p \leq x^2_{23} \mid H_0) = 1 - a.\]

Notice from (A.7.1.1) that under \(H_0\) the distribution of \(c_1, \ldots, c_p\) and thus also of
\(c_1\) and \(c_p\) turn out to be independent of \(\xi\), depending only on \(p\) and \(n\) and thus the
c.d.f. (7.4.3) depends only on \(a\), \(p\) and \(n\), so that it will now be proper to write
the tail ends more properly as \(c_{1a}(p,n)\) and \(c_{2a}(p,n)\). The actual distribution prob-
lem on which depends the evaluation of the left hand side of (7.4.3) is solved in
(A.11.2) and (A.12.9).

(ii) The general nature of the arguments will be exactly the same as in the pre-
vvious case. Starting from (7.3.2), over \(F_{18} \leq F_{a} \leq F_{2p}\) we accept \(H(a'\Sigma_{1a} = a'\Sigma_{2a})\),
so that, by using the principle of section 4, over \(\{\exists F_{18} \leq a'S_{1a}a / a'S_{2a}a \leq F_{2p}\}\) we
accept \(H(a'\Sigma_{1a} = a'\Sigma_{2a}) = H(\Sigma_1 = \Sigma_2) = H_0\), and thus over its complement

\[U = \sum \frac{a'S_{1a}a}{a'S_{2a}a} \leq F_{28} \text{ or } U \leq F_{18} \text{ we reject } H_0.\]

As before we set it up as the extended
type I test and, using (A.2.6), notice that the statement \(F_{18} \leq a'S_{1a}a / a'S_{2a}a \leq F_{28}\)
(where \(S_1, S_2\) and \(F_{18}\) and \(F_{28}\) are held fixed and \(a\) alone is varied) \(\iff\) \(F_{18} \leq c_1 \leq c_p \leq F_{28}\), where \(c_1\) and \(c_p\) are the smallest and largest roots of the p'th degree
equation in \(c\):
all the p roots being here, a.e., positive, since \( S_1 \) and \( S_2 \) are by the conditions of the problem, a.e., p.d. Starting out from the (modified) type I test (7.3.2) for \( H_{0,a} \) we have thus for \( H(\Sigma_1 = \Sigma_2) \) the extended type I region

\[
(7.4.5) \quad c_p \geq F_{2\beta}(n_1, n_2) \quad \text{and/or} \quad c_1 \leq F_{1\beta}(n_1, n_2).
\]

As in the previous case, given \( \alpha \), to determine \( F_{2\beta} \) and \( F_{1\beta} \) we first take over (see (vii) of (7.3) and also 12.4) the relation (involving only \( n_1 \) and \( n_2 \)) between \( F_{2\beta} \) and \( F_{1\beta} \) imposed by the condition of local unbiasedness and write down the further condition (which now completely determines \( F_{2\beta} \) and \( F_{1\beta} \))

\[
(7.4.6) \quad P(F_{1\beta} \leq c_1 \leq c_p \leq F_{2\beta} \mid H_0) = 1 - \alpha.
\]

Notice from (A.7.1) that under \( H_0 \) the distribution of \( c_1, \ldots, c_p \) and thus also of \( c_1 \) and \( c_p \) happen to be independent of the common value of \( \Sigma_1 \) and \( \Sigma_2 \) and also of \( \xi_1, \ldots, \xi_k \), depending only on \( p, n_1 \) and \( n_2 \) and thus the c.d.f. (7.4.6) depends only on \( \alpha, p, n_1 \) and \( n_2 \), so that the tail ends \( F_{2\beta} \) and \( F_{1\beta} \) can be more appropriately written as \( c_{1\alpha}(p, n_1, n_2) \) and \( c_{2\alpha}(p, n_1, n_2) \). The actual distribution problem on which depends the evaluation of the left side of (7.4.6) is solved in section (A.8.3).

(iii) By the test procedure (7.3.3), over \( \frac{a^tS^*a}{a^tSa} < F_{\beta}(n_1, n_2) \) we accept \( H(a^t\xi_1 = \ldots = a^t\xi_k) \) ( = \( H_{0,a} \)), so that using the principle of section 4, over

\[
\bigcap \frac{a^tF_a}{a^tSa} \leq F_{\beta} \quad \text{we accept} \quad \bigcap \frac{a^t\xi_1 = \ldots = a^t\xi_k}{\xi_1 = \xi_2 = \ldots = \xi_k} = H_0 \quad \text{and over its complement} \quad \frac{a^tF_a}{a^tSa} > F_{\beta} \quad \text{we reject} \quad H_0.
\]

We set it up as the extended type I test and, using (A.2.6), notice that the statement \( a^tS^*a / a^tSa \leq F_{\beta} \) (where \( S^* \) and \( S \) and \( F_{\beta} \) are held fixed and \( a \) alone varied) \( \iff \) \( c_r \leq F_{\beta} \) where \( c_r \) is the largest root of the \( p \)th degree equation in \( c \)

\[
(7.4.7) \quad |S^* - cs|_1 = 0.
\]

From the definitions and assumptions of section (7.2) and section (5) it is easy to check that \( S \) is, a.e., p.d. while \( S^* \) is, a.e., at least p.s.d. of rank \( r = \min(p, k-1) \).
It will of course be, a.e., p.d. if \( p \leq k-1 \). In any case, we can say that, out of the \( p \) roots of (7.4.7), \( p-r \) will be always zero, while \( r \) roots, to be called \( c_1 \leq c_2 \leq \ldots \leq c_r \), will be, a.e., positive when \( r = \min(p,k-1) \). Starting out from the straight type I test (7.3.3) for \( H_{0a} \) we have thus for \( H(\Sigma_1 = \Sigma_2) \) the extended type I region:

\[
(7.4.8) \quad c_r \geq F_\beta(n_1,n_2),
\]

where, given the size of (7.4.8) \( \alpha \), \( F_\beta \) is to be determined by

\[
(7.4.9) \quad P(c_r \geq F_\beta / H_0) = \alpha.
\]

Notice from (4.7.5) that under \( H_0 \) the distribution of \( c_1, \ldots, c_r \) and thus also of \( c_r \) happen to depend only on \( p, n_1 \) and \( n_2 \), i.e., on \( p, k-1, n-k \) (where \( n \) is the total number of observations and \( k \) the total number of samples or populations), being independent of all other nuisance parameters. Also the c.d.f. (7.4.9) depends only on \( \alpha, p, k-1, n-k \). Thus the tail end \( F_\beta \) can now be more appropriately written as \( c_\alpha(p,k-1,n-k) \). The actual distribution problem on which the evaluation of the left side of (7.4.9) depends is solved in (8.7) and (9.1)-(9.8).

(iv) By the test procedure (7.3.7), over \( r_{ab}^2 = \frac{(a'S_{12}b)^2}{(a'S_{11}a)(b'S_{22}b)} \leq r_\beta^2(n-1) \) we accept

\( H(a'\Sigma_{12}b = 0) \), so that, using the principle of section 4, over \( \bigcap_{ab} r_{ab}^2 \leq r_\beta^2(n-1) \)

we accept \( \bigcap_{ab} (a'\Sigma_{12}b = 0) = H(\Sigma_{12} = 0) = H_0 \), and over its complement

\( U_{ab} r_{ab}^2 > r_\beta^2(n-1) \) we reject \( H_0 \). As before, we set this up as an extended type I test and, using (A.2.3), notice that the statement \( \frac{(a'S_{12}b)^2}{(a'S_{11}a)(b'S_{22}b)} \leq r_\beta^2(n-1) \) (where \( S_{11}, S_{12}, S_{22} \) and \( r_\beta \) are held fixed and \( a, b \) alone varied) \( \iff \)

\( c_p \leq r_\beta^2(n-1) \), where \( c_p \) is the largest root of the \( p \)th degree equation in \( c \):

\[
(7.4.10) \quad \left| c_{S_{11}} - S_{12} S_{22}^{-1} S_{12} \right| = 0.
\]

From the definitions and assumptions of section 5 and (7.2) it is easy to see that, a.e., \( S_{11} \) is p.d. and so also \( S_{12} S_{22}^{-1} S_{12} \), so that, a.e., all roots will be positive.
Under these conditions it is known (from (4.1.16)) that the $p$ roots will all, i.e., lie between 0 and 1, satisfying, say, $0 < c_1 < c_2 < \ldots < c_p < 1$. Starting out from the straight type I test (7.3.7) for $H_{0\beta}$ we have thus for $H(E_p = 0)$ the extended type I region:

(7.4.11) \[ c_p \geq r_\beta^2(n - 1), \]

where, given that $a$ is the size of (7.4.11), $r_\beta$ is to be determined by

(7.4.12) \[ P(c_p \geq r_\beta^2 \mid H_0) = a. \]

Notice from (A.7.3) that under $H_0$ the distribution of $c_1, \ldots, c_p$ and thus also of $c_p$ happen to depend only on $p$, $q$, and $n$, being independent of all other nuisance parameters. Thus the c.d.f. (7.4.12) also depends only on $a$, $p$, $q$, $n$ and hence the tail and $r_\beta^2$ can be now more appropriately written as $c_a(p,q,n)$. The actual distribution on which the evaluation of the left side of (7.4.12) is solved in (8.4) and (9.1)-(9.8).

8. Reduction of some distribution problems and some actual distributions.

(8.1) Distribution of $\mathcal{T}(p \times p)$, where $\mathcal{T} = X(p \times n)X'(n \times p)$ ($p < n$) and where $X$ has the probability law (5.13): \[ \int 1/(2\pi)^{n/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1}X'(n \times p)X(1 \times n) \right\} dX, \]

symmetric p.d. As in (A.8.6), put $X(p \times n) = \mathcal{T}(p \times p)L(p \times n)$, subject to $LL' = I(p)$, observe that $\mathcal{T}$ and $L_I$ have the distribution

(8.1.1) \[ 2^p \int 1/(2\pi)^{p/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1}L_{i=1}^n \mathcal{T} \right\} f(t_{ii}^n) d\mathcal{T} / \left| \frac{\partial(\mathcal{L}' \mathcal{L})}{\partial(L_I)} \right|_{L_I}. \]

Now, using (A.8.6.3) to integrate out over $L_I$, we have the following distribution for $\mathcal{T}$:

(8.1.2) \[ \int 1/(2\pi)^p \prod_{i=1}^n \Gamma(n-i+1/2) \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1}L_{i=1}^n \right\} f(t_{ii}^n) d\mathcal{T}. \]

From (8.1.2), by using (A.6.1.12) and the fact that $|nS| = |\mathcal{T}|^2 = \prod_{i=1}^n t_{ii}^2$, we have
the following distribution for $S$ (usually known as the Wishart distribution):

\[
\begin{align*}
\text{(8.1.3) } & \frac{\Gamma(p+1)}{\pi^\frac{p(p-1)}{2}} \prod_{i=1}^{n} \Gamma\left(\frac{n-i+1}{2}\right) \exp \left(-\frac{1}{2} \text{tr } nZ^{-1}S \right) \frac{1}{S^{\frac{n-p-1}{2}}} \text{d}S.
\end{align*}
\]

(8.2) Distribution of $c(X')$'s where $X(p \times n)$ ($p \leq n$) has the distribution (5.13), $\Sigma$ being symmetric p.d. Using the results of (A.7.1) we start, without any loss of generality, from the canonical form (A.7.1.1), use (A.3.6) to set $X(p \times n)$

\[
= M(p \times p)D_{\gamma}(p \times p)L(p \times n)
\]

where $LL' = I(p)$ and $\gamma$ is $\perp$ with a positive first row, take over from (A.6.3.1) the Jacobian $J(k; 1, \gamma, c)$'s, $L$'s) the distribution:

\[
\text{(8.2.2) } \int \frac{\Gamma(p_{n/2})}{\Gamma(p_{n/2})} \exp \left(-\frac{1}{2} \text{tr } D_{1/\gamma}^{MD} M' \gamma \right) \prod_{i=1}^{n} \left( c_{i} - c_{j} \right) \gamma \text{d}c_{i} \mod \left( \prod_{i<j=1}^{n} \left( c_{i} - c_{j} \right) \right) \frac{\partial M_{1}^{1}}{\partial (M_{D}^{1})} \frac{\partial L_{1}^{1}}{\partial (L_{D}^{1})} \text{M}_{1}^{1} \text{L}_{1}^{1}.
\]

Using (A.8.6.3) to integrate out over $L_{1}$, we have for $c$ and $M_{1}$ the distribution

\[
\text{(8.2.2) } F(p,n) \int \frac{\Gamma(p_{n/2})}{\Gamma(p_{n/2})} \exp \left(-\frac{1}{2} \text{tr } D_{1/\gamma}^{MD} M' \gamma \right) \prod_{i=1}^{n} \left( c_{i} - c_{j} \right) \gamma \text{d}c_{i} \mod \left( \prod_{i<j=1}^{n} \left( c_{i} - c_{j} \right) \right) \frac{\partial M_{1}^{1}}{\partial (M_{D}^{1})} \text{M}_{1}^{1} \text{L}_{1}^{1}.
\]

This is the point to which the reduction of the distribution problem for the general case can be conveniently carried out. If, however, all $\gamma_{i} = \text{a constant } 1 \text{ (without any loss of generality), then }\text{tr } D_{1/\gamma}^{MD} M' = \text{tr } MD M' = \text{tr } c_{i} = \frac{p}{2} \Sigma c_{i}$, and now using (A.8.6.3) again to integrate out over $M_{1}$ and remember that

\[
\int_{MM'=I(p)} dM_{1}^{1} \left| \frac{\partial (MM')}{\partial (M_{D})} \right| = \frac{1}{2^{p}} \int_{MM'=I(p)} dM_{1}^{1} \left| \frac{\partial (MM')}{\partial (M_{D})} \right| M_{1}^{1}.
\]

and (M to have a positive first row) we have for $c$'s the distribution on the null hypothesis all $\gamma_{i}$'s = const. = 1):
\[(8.3.3) \sum \frac{n_p}{n_2} \sum \frac{n_p}{2} -p \int F(n, p) F(p, p) \exp \left\{ -\frac{1}{2} \sum_{i=1}^{p} \left( c_i \right) \sum_{i=1}^{n_p} \frac{n_p-1}{2} \right\} \text{mod} \int_{\sum_{i<j=1}^{p} (c_i - c_j)} \frac{p-1}{2} \text{d} \frac{p}{2} \frac{(c_i - c_j)}{2} \]

where \( F(n, p) \) and \( F(p, p) \) are given by \((A.8.6.3) \sum_{10} \).

\((8.3)\) Distribution of \( cX_1X_2^{-1} \) where \( X_1(p \times n_1), X_2(p \times n_2) \) \((p < n_1, n_2)\) have the probability law of \((A.7.2)\). As in \((8.2)\), using the results of \((A.7.2)\), we start without any loss of generality, from the canonical form \((A.7.2.1)\), use \((A.3.3)\) to set \( X_1(p \times n_1) = A(p \times p)D_{\text{c}}(p \times p)L_1(p \times n_1) \) and \( X_2(p \times n_2) = A(p \times p)L_2(p \times n_2) \), where \( A \) is non-singular and \( L_{11}L_{21} = L_{22}L_{12} = I(p) \), take over from \((A.6.2.11)\) the Jacobian \( J(X_1X_2; A, c, I_{11}, I_{21}) \) and obtain for \( A, c, I_{11}, I_{21} \) the distribution

\[ \sum \frac{1}{(2\pi)^{n_2}} \frac{p}{n_1/2} \int \exp \left\{ -\frac{1}{2} \text{tr} \left( D_{1/2} \text{c}A^2 + AA^2 \right) \right\} \text{d}L_{11} \left( \left| \frac{\partial (L_{11}/L_{1D})}{\partial (L_{2D})} \right| \right)_{12} \]

Using \((A.8.6.3)\) to integrate out over \( L_{11} \) and \( L_{21} \) we have for \( A \) and \( c \)'s the distribution

\[ \sum \frac{1}{(2\pi)^{n_2}} \frac{p}{n_1/2} \int \exp \left\{ -\frac{1}{2} \text{tr} \left( D_1 \text{DA}A^2 + AA^2 \right) \right\} \text{d}A \text{d}c \text{mod} \int_{\sum_{i<j=1}^{p} (c_i - c_j)} \frac{p-1}{2} \text{d} \frac{p}{2} \frac{(c_i - c_j)}{2} \]

As before, this is the point to which, for the general case, the reduction of the distribution problem can be carried out. If, however, all \( \gamma \)'s, i.e., all \( c(\Sigma_1^{E_1}) = 1 \) (which, by \((A.1.13)\), happens if and only if \( \Sigma_1 = \Sigma_2 \)), then further reduction is possible and \((8.3.2)\) reduces to

\[ \sum \frac{1}{(2\pi)^{n_2}} \frac{p}{n_1} \frac{p}{n_2} \int \exp \left\{ -\frac{1}{2} \text{tr} (AD_1 A^2 + AA^2) \right\} _{A} \frac{p}{2} \frac{(n_1 - p - 1)}{2} \text{d}A \text{d}c \text{mod} \int_{\sum_{i<j=1}^{p} (c_i - c_j)} \frac{p-1}{2} \text{d} \frac{p}{2} \frac{(c_i - c_j)}{2} \]

involving the \( c \)'s taken over from \((8.3.2)\).
Now putting $\frac{AD}{1+c} = B$ and using (A.5.2) we have for B and c's the distribution

$$
\begin{aligned}
\int_{1/2n}^{\frac{p}{2n}} F(p,n_1)F(p,n_2) \exp \left[ -\frac{1}{2} \operatorname{tr} BB' \mathcal{J} |B|^{n_1+n_2-p} \right]
\times \prod_{i=1}^{p} \frac{n_{1-p-1}}{c_i} \cdot \frac{n_{1+n_2}}{(1+c_i)^2} \mod \prod_{i<j} (c_i - c_j) .
\end{aligned}
$$

Now using (A.8.7.1) and remembering that

$$
\int \exp \left[ -\frac{1}{2} \operatorname{tr} BB' \mathcal{J} |B|^{q} \right] d\mathcal{B}
$$

over B with a positive first row

$$
= \frac{1}{2^p} \int \exp \left[ -\frac{1}{2} \operatorname{tr} BB' \mathcal{J} |B|^{q} \right] d\mathcal{B},
$$

we have for c's the distribution (on the null hypothesis $\Sigma_1 = \Sigma_2$) $\mathcal{J}_{3,4,7,15,167}$:

$$
\begin{aligned}
\int \left[ \prod_{i=1}^{p} \frac{n_{1+n_2-1}}{\Gamma(\frac{1+n_2-1}{2})} \prod_{i=1}^{p} \frac{n_{1-1+1}}{\Gamma(\frac{1-1+1}{2})} \prod_{i=1}^{p} \frac{n_{2-1+1}}{\Gamma(\frac{2-1+1}{2})} \prod_{i=1}^{p} \frac{p-1}{\Gamma(\frac{p-1}{2})} \prod_{i<j} (c_i - c_j) \right] .
\end{aligned}
$$

(8.4) Distribution of $c_{i}(X_{11}X_{11})^{-1}(X_{12}X_{12})^{-1}(X_{22}X_{22})^{-1}(X_{33}X_{33})^{-1} \mathcal{J}$ (to be here called just c's), when $X_1(p \times n)$, $X_2(q \times n)(p \leq q; p \neq q \leq n)$ has the probability law of (A.7.3).

As in the two previous cases, start, without any loss of generality, from the canonical distribution form (A.7.3.5) and let $c_{i}(\Sigma_{11}^{-1}\Sigma_{22}^{-1}\Sigma_{22}^{-1}\Sigma_{11}^{-1}) = x_i$ (say) ($i=1,2,\ldots,p$). Next use (A.3.17) to set $X_2(q \times n) = \mathcal{N}(q \times q)L_2(q \times n)$ and $X_1(p \times n)$

$= U(p \times p)\mathcal{M}_1(p \times n-q) \mathcal{D}_2(p \times p)\mathcal{M}_2(p \times q)\mathcal{J}(L_{11}^{n-q}, L_{22}^{n-q}, L_{21}^{n-q}, L_{12}^{n-q})$, where $\mathcal{N}$, $U$, $L_1$ and $L_2$ are characterized in (A.3.17) and $e_{i} = (1-c_{i})/c_{i}$. Now take over from (A.6.5.1) the Jacobian $J(x_1, x_2, \mathcal{N}, U, c's, \mathcal{M}_{11}^{I}, \mathcal{M}_{21}^{I}, L_{11})$ and obtain for $\mathcal{N}, U, c's, \mathcal{M}_{11}^{I}, \mathcal{M}_{21}^{I}$ and $L_{11}$ the distribution
\[(8.4.1) \quad 2^{p+q} \sum_{l/(2n)} \frac{n(p+q)}{2} \prod_{i=1}^{p} (1-\gamma_i) \exp \left\{ -\frac{1}{2} \text{tr} \begin{bmatrix} D_{1/1-\gamma} & -(D_{/1-\gamma})^2 \hline -(D_{/1-\gamma})^3 & D_{1/1-\gamma} \\ \hline 0 & 0 \end{bmatrix} \right\} \]

\[
x \left[ \begin{bmatrix} UD_{1+e} U' & U M_{2I} \hline \tilde{T}, M_{1I} U' & \tilde{T}_{2I} \end{bmatrix} \right] \left[ \begin{bmatrix} u \n^{-p} \text{d}U \frac{A}{t} \prod_{i=1}^{n} \frac{t}{t_i} \text{d}T \frac{P}{P_i} \exp \left\{ -\frac{p-1}{2} \prod_{i=1}^{n-1} \text{d}e_i \text{mod} \sum_{i<j=1}^{n} \prod_{i=1}^{p-1} (e_i - e_j) \right\} \hline \begin{bmatrix} \partial (M_{1I})/\partial (M_{1I}) \hline \partial (M_{2I})/\partial (M_{2I}) \end{bmatrix} \end{bmatrix} \cdot \right]

Using (A.8.6.3) to integrate out over \( M_{1I} \) and \( L_{2I} \), we have for \( \tilde{T}, U, M_{2I} \) and \( e \)'s the distribution

\[(8.4.2) \quad 2^{p+q} \sum_{l/(2n)} \frac{n(p+q)}{2} \prod_{i=1}^{p} (1-\gamma_i) \exp \left\{ -\frac{1}{2} \text{tr} \begin{bmatrix} D_{1/1-\gamma} & -(D_{/1-\gamma})^2 \hline -(D_{/1-\gamma})^3 & D_{1/1-\gamma} \\ \hline 0 & 0 \end{bmatrix} \right\} \]

\[
x \left[ \begin{bmatrix} UD_{1+e} U' & U M_{2I} \hline \tilde{T}, M_{1I} U' & \tilde{T}_{2I} \end{bmatrix} \right] \left[ \begin{bmatrix} u \n^{-p} \text{d}U \frac{A}{t} \prod_{i=1}^{n} \frac{t}{t_i} \text{d}T \frac{P}{P_i} \exp \left\{ -\frac{p-1}{2} \prod_{i=1}^{n-1} \text{d}e_i \text{mod} \sum_{i<j=1}^{n} \prod_{i=1}^{p-1} (e_i - e_j) \right\} \hline \begin{bmatrix} \partial (M_{2I})/\partial (M_{2I}) \end{bmatrix} \end{bmatrix} \cdot \right]

This is the point to which, in the general case, the reduction of the distribution problem can be conveniently carried out. However, if all \( c(\Sigma 1_{11} 1_{22} 1_{12}) = 0 \), i.e., all \( \gamma_i \)'s = 0, (which, according to (A.1.17), happens if and only if \( \Sigma 1_{12} = 0 \)), then further reduction is possible and (8.4.2) reduces to

\[(8.4.3) \quad 2^{p+q} \sum_{l/(2n)} \frac{n(p+q)}{2} \exp \left\{ -\frac{1}{2} \text{tr} UF_{D_{1+e}} U' + \text{tr} \tilde{T}_{2I} \right\} \left[ \begin{bmatrix} u \n^{-p} \text{d}U \frac{A}{t} \prod_{i=1}^{n} \frac{t}{t_i} \text{d}T \frac{P}{P_i} \exp \left\{ -\frac{p-1}{2} \prod_{i=1}^{n-1} \text{d}e_i \text{mod} \sum_{i<j=1}^{n} \prod_{i=1}^{p-1} (e_i - e_j) \right\} \hline \begin{bmatrix} \partial (M_{2I})/\partial (M_{2I}) \end{bmatrix} \end{bmatrix} \cdot \right]

Note that \( \text{tr} \tilde{T}_{2I} = \sum_{i=1}^{q} t_{i1}^2 \) and
\[
\int \exp \left( - \frac{1}{2} \text{tr} \, T \right) \left[ \sum_{i=1}^{n} t_{i}^{n-1} \sum_{i \geq j=1}^{q} \text{det}_{ij} \right] = 2^{-q+n/2} \pi^{(q-1)/4} \frac{\Gamma(n+1)}{\Gamma(n+1/2)} \] and hence
integrating out over \( \hat{T} \) and \( M_{21} \) obtain for \( U \) and \( \varepsilon \)'s the distribution
\[
(8.4.4) \quad 2^{p} \left( \frac{n+1}{2} \right)^{\frac{p+q-1}{2}} \pi \left[ \frac{\Gamma(n+1)}{\Gamma(n+1/2)} \right] F(p,n-q) F(p,q) F(q,n) \]
\[\times \exp \left( - \frac{1}{2} \text{tr} \, U D_{1+c} U^\dagger \right) \left| U \right|^{n-p} \frac{p}{\pi} \sum_{i=1}^{n-p-q-1} \left( \frac{c_{i}}{c_{i+1}} \right)^{n/2} \text{det}_{i} \mod \frac{p-1}{\pi} \left( c_{i} - c_{j} \right)^{n/2}.\]

Now put \( UD_{1+c} = V \), use (A.5.3) and (A.8.7) to integrate out over \( V \) and obtain for \( \varepsilon \)'s the distribution
\[
(8.4.5) \quad \text{Const.} \left[ \prod_{i=1}^{p} \frac{1}{(n-p-q-1/2) c_{i}^{n/1+c_{i}}^{(n/2)}} \right] \mod \left[ \prod_{i<j=1}^{p-1} \left( c_{i} - c_{j} \right)^{n/2} \right].\]

Putting \( c_{i} = (1-c_{i})/c_{i} \) we have for \( c_{i} \)'s, on the null hypothesis \( \Sigma_{12} = 0 \), i.e.,
\( c(\Sigma_{11}^{12} \Sigma_{22}^{12}) = 0 \), i.e., \( \gamma = 0 \), the distribution, \( \int_{3,4,7,15,16} \)
\[
(8.4.6) \quad \text{Const.} \left[ \prod_{i=1}^{p} \frac{(1-c_{i})^{n-p-q-1}}{c_{i}^{n/2} \text{det}_{i} \mod \frac{p-1}{\pi} \left( c_{i} - c_{j} \right)^{n/2}} \right].\]

An important special case is when \( p = 1 \) and this we shall consider both on the null and on the non-null hypothesis. In this case there is only one non-zero (and hence positive) \( c \) or \( c \) and only one possible non-zero (and hence positive) \( \gamma \). Thus
\[
\mod \left[ \prod_{i<j=1}^{p-1} \left( c_{i} - c_{j} \right)^{n/2} \right] \text{or} \mod \left[ \prod_{i<j=1}^{p-1} \left( c_{i} - c_{j} \right)^{n/2} \right] \text{will drop out. On the null hypothesis} \gamma = 0, \text{the distribution of} \ c \text{and} \ c, \text{as special cases of (8.6.5) and (8.4.6), will respectively be}\]
\[
(8.4.7) \quad \frac{\Gamma(n/2)}{\Gamma(q/2) \Gamma(n-q/2)} \begin{bmatrix} \left( n-q-2 \right)^{c_{i}} / (1+c_{i})^{n/2} \end{bmatrix},\]
\[
(8.4.8) \quad \frac{\Gamma(n/2)}{\Gamma(q/2) \Gamma(n-q/2)} \begin{bmatrix} (1-c)^{c_{i}} (n-q-2)^{c_{i}} (q-2)^{c_{i}} \end{bmatrix}.\]

For the distribution on the non-null hypothesis \( \gamma \neq 0 \), we start from (8.4.2), put
\( U(1 \times 1) = u \) (a scalar) so that \( UD_{1+e} U^\dagger = (1+e)u^{2} \), \( M_{2}(1 \times q) = m_{i}^{1}(1 \times q) = (m_{1}, m_{2}, \ldots, m_{q}) \) (say), and take \( m_{2}, \ldots, m_{q} \) to be the so-called independent elements of \( m_{2}^{1} \),
so that

\[
\frac{dN_{2I}}{d(M_{2D})} = \frac{q}{\frac{1}{i=2} (1-i)} \frac{dm_i/(1-\Sigma m_i^2)}{m_{2I}}
\]

and obtain for \(e\) the distribution

\[
(8.4.9) \quad 2^{q+1} \int \frac{n(q+1)}{\Gamma(2)} \frac{n}{(1-\gamma)^2} \int F(1,n-q) \cdot (q,n) e^{-\frac{n-q-2}{2}} \frac{1}{2} \sum_{i=2}^{p} \sum_{j=1}^{i} t_{ij}^2 \frac{u}{m_{11}} \frac{v}{m_{1i}} \frac{w}{m_{1i}} \frac{1}{2} \frac{d^2}{d^2 t_{ij} dt_{ii}} \frac{d^2}{d^2 t_{ii} dt_{ij}} \frac{d}{d^2 t_{ij} dt_{ii}} \frac{d}{d^2 t_{ii} dt_{ij}} \frac{d}{d^2 t_{ij} dt_{ii}} \frac{d}{d^2 t_{ii} dt_{ij}}
\]

Now putting \(m_1 = \cos \Theta\) so that \(\sum_{i=2}^{q} m_i^2 = \sin^2 \Theta\), we note from (A.8.4) that

\[
(8.4.10) \quad \int \exp \left[ -\frac{1}{2(1-\gamma)} \left( \sum_{i=2}^{q} \frac{m_i^2}{1} \right)^{1/2} \right] \sin \Theta \leq \left( \sum_{i=2}^{q} \frac{m_i^2}{1} \right)^{1/2} \leq \sin \Theta + d \sin \Theta
\]

\[
x \int \frac{n(q+1)}{\Gamma(2)} \frac{n}{(1-\gamma)^2} \int F(1,n-q) \cdot (q,n) e^{-\frac{n-q-2}{2}} \frac{1}{2} \sum_{i=2}^{p} \sum_{j=1}^{i} t_{ij}^2 \frac{u}{m_{11}} \frac{v}{m_{1i}} \frac{w}{m_{1i}} \frac{1}{2} \frac{d^2}{d^2 t_{ij} dt_{ii}} \frac{d^2}{d^2 t_{ii} dt_{ij}} \frac{d}{d^2 t_{ij} dt_{ii}} \frac{d}{d^2 t_{ii} dt_{ij}}
\]

Using (8.4.10) and also integrating out over \(t_{ij}\)'s \((j = 1,2,\ldots,i\) and \(i = 2,3,\ldots,p\)), and setting \(v^2 = (1+e)u^2\), we have (A.8.4.9) reducing to

\[
(8.4.11) \quad \text{Const} \ (\text{which is easily obtained}) \quad \int \frac{n-q-2}{2} \frac{1}{(1+e)^2} \frac{d^2}{d^2 \Theta} \sin \Theta \geq \left( \sum_{i=2}^{q} \frac{m_i^2}{1} \right)^{1/2} \leq \sin \Theta + d \sin \Theta
\]

\[
x \int \exp \left[ -\frac{1}{2(1-\gamma)} \left( v^2 + t_{11}^2 \right) \right] \cosh \left( \frac{v}{1-\gamma} \right) \left( v^2 + t_{11}^2 \right) \frac{d^2}{d^2 v dt_{11}} \frac{d^2}{d^2 v dt_{11}} \frac{d}{d^2 v dt_{11}} \frac{d}{d^2 v dt_{11}} \frac{d}{d^2 v dt_{11}} \frac{d}{d^2 v dt_{11}}
\]

the limits of \(v\) and \(t_{11}\) being from \(0\) to \(\infty\) and of \(\Theta\) from \(0\) to \(\pi\). To evaluate the integral.
\[
\int_{\theta=0}^{\pi} \int_{x,y=0}^{\infty} \exp \left( -a(x^2 + y^2) - 2bxy \cos \theta \right) \mathcal{J}(xy)^n \ dx \ dy \ (\sin \theta)^m \ d\theta
\]

we proceed as follows

\[(8.1.12) \quad \mathcal{I} = \int_{x,y=0}^{\pi} \exp \left( 2abxy \cos \theta \right) \mathcal{J} \ (\sin \theta)^m \ d\theta = 2 \int_{x,y=0}^{\pi} \cosh(2abxy \cos \theta) \sin^n \theta \ d\theta \]

\[= 2\Gamma(\frac{m+1}{2}) \Gamma(\frac{1}{2}) \frac{1}{2} I_{m/2} (2abxy), \] where \( I \) stands for the Bessel I function in the usual notation (Watson's Bessel functions, p. 79, formula (9)). Thus we have

\[(8.1.13) \quad \mathcal{I} = 2\Gamma(\frac{m+1}{2}) \Gamma(\frac{1}{2}) (ab) \frac{1}{2} \int_{x,y=0}^{\infty} \frac{1}{2} \cosh(2abxy) \ dx \ dy.
\]

To evaluate this put \( xy = z \) and \( x/y = e^v \), so that \( J(x,y;z,v) = 1/2 \) and the range of \( z \) and \( v \) could be taken as \( 0 \leq z < \infty \) and \( -\infty < v < \infty \) and from symmetry the resulting integral has to be divided by 2. Thus we have

\[(8.1.14) \quad \mathcal{I} = \frac{1}{2} \Gamma(\frac{1}{2}) \Gamma(\frac{m+1}{2}) (ab) \frac{1}{2} \int_{z=0}^{\infty} \int_{v=0}^{\infty} \exp \left( -2az \cosh v \right) \mathcal{J} \ a^{-\frac{m}{2}} I_{m/2} (2abz) \ dz \ dv.
\]

But putting \( v = 0 \) in formula (9), p. 181, Watson's Bessel functions and noting that \( K \) stands for the Bessel K-function in the usual notation we have

\[(8.1.15) \quad \int_{v=0}^{\infty} \exp \left( -2az \cosh v \right) \mathcal{J} \ dv = K_0 (2az).
\]

Hence

\[(8.1.16) \quad \mathcal{I} = \frac{1}{2} \Gamma(\frac{m+1}{2}) \Gamma(\frac{1}{2}) (ab) \frac{1}{2} \int_{z=0}^{\infty} \frac{1}{2} I_{m/2} (2abz) K_0 (2az) \ a^{-\frac{m}{2}} \ dz.
\]

Now putting \( \mu = 0 \), \( v = m/2 \) and \( \lambda = n - \frac{m}{2} \) in (1) of (13.45), p. 410, Watson's Bessel functions and checking up on the validity conditions indicated there, we have
\[ (8.4.17) \quad \int_0^\infty K_0(2az)J_m(2acz)z^{-m/2}dz = \int (2ac)^2 \Gamma^2(\frac{n+1}{2}) \frac{n-1-m}{2}/(2a)^n \Gamma(\frac{m+2}{2}) \frac{x_2^{F_1}(\frac{n+1}{2}, \frac{n+1}{2}; \frac{m+2}{2}; -c^2)}{\Gamma^2(\frac{n+1}{2})}. \]

If now we put \( c = ib \), we shall have \( J_{m/2}(2abz) = (i)^m \Gamma_{m/2}(2abz) \), so that substituting in (8.4.16) we should have

\[ (8.4.18) \quad \text{Integral} = \int x^{-3} \Gamma^2(\frac{n+1}{2}) \Gamma(\frac{m+1}{2})/\Gamma^2(\frac{m+2}{2}) x^{F_1}(\frac{n+1}{2}, \frac{n+1}{2}; \frac{m+2}{2}; b^2). \]

Substituting in (8.4.11) we have for the distribution proportional to

\[ (8.4.19) \quad \frac{\Gamma^2(\frac{n}{2})/\Gamma^2(\frac{m+2}{2})}{\Gamma^2(\frac{n+1}{2})} x^{F_1}(\frac{n+1}{2}, \frac{n+1}{2}; \frac{q}{2}, \frac{q}{1+c}) c^{\frac{n-2}{2}} \frac{n}{2} dc/(1+c)^2. \]

Now putting \( c = 1/(1+c) \), we have for the distribution, \( z_2 \),

\[ (8.4.20) \quad \frac{\Gamma^2(\frac{n}{2})/\Gamma(\frac{m+2}{2})}{\Gamma^2(\frac{n+1}{2})} x^{F_1}(\frac{n+1}{2}, \frac{n+1}{2}; \frac{q}{2}, \frac{q}{1+c}) (1+c)^{-\frac{n-2}{2}} \frac{n}{2} dc. \]

\[ (8.5) \quad \text{Distribution of} \quad c\sum_{(x_1, x_1', x_1') x_1 x_1'} x_1 x_1' x_1' x_1' x_1 x_1' x_1' x_1' x_1' \frac{x_2 x_2' x_2' x_2' x_2 x_2'}{x_3 x_3' x_3' x_3' x_3 x_3' x_3' x_3'} \quad \text{where} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad p \leq q, \quad q, \quad r \quad (p+q+r \leq n) \quad \text{has the probability law (5.15) of (A.7.4)} \quad \text{or (without any loss of generality) just the canonical form (A.7.4.5).} \quad \text{Notice that (A.7.4.5) can be rewritten as}

\[ (8.5.1) \quad \int 1/(2\pi) \prod_{i=1}^n \left( I(p) \right) \exp \left\{ -\frac{1}{2} \text{tr} \left[ I(p) \right] - \frac{D}{\Gamma(1-\gamma)} \right\} \left[ I(p) \right]^{0} \left[ 0 \right]^{0} \left[ D/\Gamma(1-\gamma) \right]^{0} \left[ 0 \right]^{I(q-p)} \left\{ \frac{1}{2} \text{tr} x_3 x_3' \right\} dx_1 dx_2 dx_3. \]
Now using (A.3.19) make the transformation $X_3(r \times n) = \vec{T}(r \times r)L_3(r \times n)$ subject to

$L_3L_3^\dagger = I(r)$ and

\[
\begin{pmatrix}
X_1 \\
X_2
\end{pmatrix}
\sim
\begin{pmatrix}
Z_{11} \\
Z_{21}
\end{pmatrix}
(\begin{pmatrix}
L_{11} & Z_{12} \\
L_{21} & Z_{22}
\end{pmatrix})^{n-r}
\begin{pmatrix}
n \\
n-r \\
r
\end{pmatrix}
\]

where $L$ is a completion of $L_3$ to $n$. We have, by (A.6.6),

\[
J(X;Z,\vec{T},L_3I) = 2^r \prod_{i=1}^{n}\int \frac{e^{-(A_iL_3)}}{\sqrt{(2\pi)^n}} dt_i dL_{3i} \frac{\partial L_{3i}}{\partial (L_3^{*1})}
\]

We have also $X_3^{*1} = \vec{T}\vec{T}^\dagger$. Hence from (8.5.1) we have for $Z, \vec{T}, L_3I$, the distribution

\[
n(2\pi)^{n-r} \prod_{i=1}^{n} (1-Y_i)^{\frac{n}{2}} \exp\left\{-\frac{1}{2} \text{tr} \begin{pmatrix}
D_1/1-Y & -D_{\vec{T}/1-Y} \\
-D_{\vec{T}/1-Y} & D_1/1-Y
\end{pmatrix}
\right\}
\]

\[
x \begin{pmatrix}
Z_{11} \\
Z_{21}
\end{pmatrix}
\left(\begin{pmatrix}
Z_{11} \\
Z_{21}
\end{pmatrix}
\right)^* - \frac{1}{2} \text{tr} \begin{pmatrix}
D_1/1-Y & -D_{\vec{T}/1-Y} \\
-D_{\vec{T}/1-Y} & D_1/1-Y
\end{pmatrix}
\begin{pmatrix}
Z_{12} \\
Z_{22}
\end{pmatrix}
\begin{pmatrix}
Z_{12}^* \\
Z_{22}^*
\end{pmatrix}
\right.
\]

\[-\frac{1}{2} \text{tr} \vec{T}\vec{T}^\dagger dt_{11} dt_{12} dt_{21} dt_{22} \prod_{i=1}^{n-r} t_i^{n-i} \frac{\partial L_{3i}}{\partial (L_3^{*1})}
\]

It is thus seen that $(Z_{11}, Z_{21}),(Z_{12}, Z_{22})$ and $(\vec{T}, L_3I)$ are distributed as three independent sets. Therefore integrating out over $(\vec{T}, L_3I)$ and $(Z_{12}, Z_{22})$ and noting from (A.3.19) that the $c_i$s of this section are exactly the same as

\[
c_{(Z_{11}Z_{12})^{-1}(Z_{11}Z_{21})(Z_{21}Z_{22})^{-1}(Z_{21}Z_{11})^{-1}}
\]

it becomes evident that both on the null and on the non-null hypothesis the distribution of these $c_i$s are exactly the same as the $c_i$s in section (8.4) with $n$ of that section being replaced here by $n-r$.

(8.6) Distribution of $c(X_1^{-1}X_2^{-1})$'s (to be called just $c_i$s), where $X_1(p \times n_1), X_2(p \times n_2)$ (p ≤ n_2 but might be ≤ n_1 or > n_1) have the distribution law of (A.7.5) with $\gamma_i$'s standing for $c(\xi \xi'\Sigma^{-1})$'s or, in another notation, $c(\Sigma_1^{-1})$. Without any
loss of generality we start from the canonical form (A.7.4.5) and consider two cases separately, namely where (i) \( p < n_1 \) and (ii) \( p > n_1 \) having respectively, a.e., \( p \) non-zero and \( n_1 \) non-zero c's. (i) For case (i) use (A.3.8) to set \( X_1(p \times n_1) = A(p \times p)D_c(p \times p)L_1(p \times n_1) \) and \( X_2(p \times n_2) = A(p \times p)L_2(p \times n_2) \), where \( A, L_1, L_2 \) satisfy the conditions of (A.3.9), take over from (A.6.2.11) the Jacobian
\[
\left( X_1, X_2, A, c_i's, L_{1I}, L_{2I} \right)
\]
and obtain for \( A, c_i's \) and \( L_{1I} \) and \( L_{2I} \) the distribution
\[
\frac{p(n_1+n_2)}{2} \exp \left\{ \frac{1}{2} \left\{ \text{tr} \left( AD_1+cA^t \right) + \sum_{i=1}^{s} \gamma_i - 2 \sum_{i=1}^{s} (AD_c^tL_1)_{1i} \right\} \right\} \\
\times \left( \begin{array}{c}
A \\
\left( \begin{array}{c}
\frac{1}{2} \\
c_i \\
\end{array} \right)^2 \\
\end{array} \right) \prod_{i=1}^{n_1-p-1} \frac{p-1}{(c_i-c_j)} \\
\delta(L_{1I}) \\
\delta(L_{2I}) \\
\end{array} \right).
\]

Use (A.8.6.3) to integrate out over \( L_{2I} \) and obtain for \( A, c_i's \) and \( L_{1I} \) the distribution
\[
\frac{p(n_1+n_2)}{2} \exp \left\{ \frac{1}{2} \left\{ \text{tr} \left( AD_1+cA^t \right) + \sum_{i=1}^{s} \gamma_i - 2 \sum_{i=1}^{s} (AD_c^tL_1)_{1i} \right\} \right\} \\
\times \left( \begin{array}{c}
A \\
\left( \begin{array}{c}
\frac{1}{2} \\
c_i \\
\end{array} \right)^2 \\
\end{array} \right) \prod_{i=1}^{n_1-p-1} \frac{p-1}{(c_i-c_j)} \\
\delta(L_{1I}) \\
\delta(L_{1D}) \\
\end{array} \right).
\]

This is the point to which, for the general case, the distribution problem can be conveniently reduced. If \( \gamma_i's = 0 \), i.e., all \( c(\Sigma_1^{-1})'s = 0 \) (which, by (A.1.13), happens if and only if \( \Sigma_1 = 0 \), i.e., \( \xi = 0 \)), then further reduction is possible and using (A.8.6.3) to integrate out over \( L_{1I} \) we have for \( A \) and \( c_i's \) the distribution
\[
\frac{p(n_1+n_2)}{2} \exp \left\{ \frac{1}{2} \text{tr} \left( AD_1+cA^t \right) \right\} \\
\left( \begin{array}{c}
\frac{p}{2} \\
c_i \\
\end{array} \right)^2 \prod_{i=1}^{n_1-p-1} \frac{p-1}{(c_i-c_j)} \\
\end{array} \right).
\]

Now as in (8.3), integrating out over \( A \) we have for \( c_i's \), i.e., for \( c(X_1X_1^tX_2X_2^t)^{-1} \)'s
the following distribution on the null hypothesis $\xi = 0$.

\[
(8.6.4) \quad \frac{p}{2} \prod_{i=1}^{n_1} \frac{\Gamma\left(\frac{n_1+n_2-i+1}{2}\right)}{\Gamma\left(\frac{n_1+i-1}{2}\right)\Gamma\left(\frac{n_2-i+1}{2}\right)\Gamma\left(\frac{p-i+1}{2}\right)} \\
\times \left[ \prod_{i=1}^{p-1} \frac{n_1-p-1}{c_i} \frac{n_1+n_2}{2} \right] \mod \left\{ \prod_{i<j=1}^{p-1} (c_i-c_j) \right\},
\]

which is exactly the same form as (8.3.6). (ii) For the case (ii) use (A.3.14) to set $X_1(p \times n_1) = p-n_1 \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right)$, $D_{c}(n_1 \times n_1)L_1(n_1 \times n_1)$ and $X_2(p \times n_2)$

\[
n_1 \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right)
\]

= $U(p \times p)L_2(p \times n_2)$, where $U$, $L_1$, $L_2$ and $c$ satisfy the conditions of (A.3.14), take over from (A.5.7.8) the Jacobian $J(X_1X_2; c's, U, L_1, L_2)$, and obtain for $C, U_I, L_{1I}, L_{2I}$ the distribution

\[
(8.6.5) \quad \frac{p}{2} \prod_{i=1}^{1/2n_j} \frac{1}{2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^{s} \Sigma_i \gamma_i^2 \right\} \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right) \\
\times \left[ \prod_{i=1}^{n_1+n_2-p} \frac{n_1+i-1}{c_i} \frac{n_1+n_2}{2} \right] \mod \left\{ \prod_{i<j=1}^{p-1} (c_i-c_j) \right\},
\]

where $D_{1c}(p) = \left( \begin{array}{cc} \Sigma & \Sigma^T \\ \Sigma^T & \Sigma \\ 0 & 0 \end{array} \right)^{n_1} + I(p)$. Using (A.8.6.3) to integrate out over $L_{2I}$, we have for $C, U, L_{1I}$ the distribution

\[
(8.6.6) \quad \frac{p}{2} \prod_{i=1}^{1/2n_j} \frac{1}{2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^{s} \Sigma_i \gamma_i^2 \right\} \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right) \\
\times \left[ \prod_{i=1}^{n_1+n_2-p} \frac{n_1+i-1}{c_i} \frac{n_1+n_2}{2} \right] \mod \left\{ \prod_{i<j=1}^{p-1} (c_i-c_j) \right\},
\]
As in case (i) we shall stop here so far as the non-central distribution is concerned. For the central case, i.e., for the null hypothesis that \( \gamma_1 \)'s = 0, i.e., all \( c(x_2^{-1}) \)'s = 0 (which happens if and only if \( \xi = 0 \)) further reduction is possible as in case (i) and, using (A.8.6.3) to integrate out over \( L_{12} \), we have for \( U \) and \( c \)'s the distribution

\[
2^{p-1/2} \sqrt{\frac{n_1}{2}} \prod F(n_1, n_1) F(p, n_2) \exp \left( -\frac{1}{2} \text{tr} \left[ U_{1+c} U_1^{-1} \right] U_1^{n_1+n_2-p} \right) dU
\]

\[
\times \prod_{i=1}^{n_1} (\xi_i^3)^{p-n_1-i} \prod_{i=1}^{n_1} \frac{1}{c_i+2} \frac{1}{d_i} \mod \sum_{i<j}^{n_1} (c_i - c_j).
\]

Now putting \( U_{1+c} = V \) and using (A.8.6.2) we integrate out over \( V \) and obtain for \( c \)'s the distribution

\[
\sum_{3, 4, 7, 15, 16} \prod_{i=1}^{n_1} \frac{1}{\Gamma\left(\frac{1}{2}\right)} \frac{1}{\Gamma\left(\frac{n_1+n_2+l}{2}\right)} \prod_{i=1}^{p-n_1-l} \frac{1}{\Gamma\left(\frac{1}{2}\right)} \frac{1}{\Gamma\left(\frac{n_2+l-l}{2}\right)} \prod_{i=1}^{n_1} \frac{1}{\Gamma\left(\frac{1}{2}\right)} \frac{1}{\Gamma\left(\frac{n_1+l-l}{2}\right)}
\]

\[
\times \prod_{i=1}^{n_1} \frac{1}{c_i+2} \frac{1}{d_i} \mod \sum_{i<j}^{n_1} (c_i - c_j).
\]

A special case of (ii), namely where \( n_1 = 1 \), is of considerable importance and in this case not only the central but also the non-central distribution is available. Notice that in this case \( n_1 = 1 \), \( X_1(p \times 1) = X(p \times 1) \) (say). \( \xi(p \times n_1) = \xi(p \times 1) \) (say), \( L_1(1 \times 1) \), subject to \( L_{11} = I(1) \), is equal to \( +1 \) so that \( L_{11} \) drops out and

\[
\left. \frac{\partial\left( L_{11} \right)}{\partial \left( L_{11} \right)} \right|_{L_{11}} = 1/2 \text{ which we absorb in the constant. There is only one non-zero (and positive) } c \text{ which is equal to } x' \left( X_2^{-1} X_2 \right) x \text{ since, in general, } \tau D \left( \sum_{i=1}^{n_1} c_i \right)
\]

\[
= \text{tr} \left( X_1^{-1} X_2 \right) = \text{tr} \left( X_2^{-1} X_1 \right) \text{ and in this case } \sum_{i=1}^{n_1} c_i = c \text{ and tr } X_1^{-1} X_2 \text{ drop out. Substituting in (8.6.6) and (8.6.8) the factor } \prod_{i<j}^{n_1} (c_i - c_j) \text{ will also drop out. Substituting in (8.6.8) we have thus in this case, for } c, \text{ the central}
\]
distribution, \( i, j, k \),

\[
(8.6.9) \quad \frac{\Gamma\left(\frac{2n_2+1}{2}\right)}{\Gamma\left(\frac{n_2-p+1}{2}\right)} \frac{p-2}{2} \frac{\Gamma\left(\frac{n_2+1}{2}\right)}{dc/(1+c)^2}.
\]

For the non-central distribution, i.e., when \( \gamma \neq 0 \), we have, in this case, \( \sum_{i=1}^{n_1} \gamma_i = \gamma \),

\[
\sum_{i=1}^{s} \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right) \frac{dL}{\sqrt{L_1}} = u_{11}\sqrt{c}\gamma \text{ and thus it is easily checked that by substituting}
\]

in (8.6.6) and remembering that \( L_1 \) could take just two values \( \pm 1 \), we shall have, for \( U \) and \( c \), the joint distribution

\[
(8.6.10) \quad \text{const.} \exp \left[ -\frac{1}{2} \text{tr} \begin{pmatrix} UD_{1+c}(p) & U^* \gamma + 2u_{11}\sqrt{c}\gamma \end{pmatrix} \right] \prod_{i=1}^{n_2+1-p} \frac{dU_3}{(U_3)^{p-1-i}} \prod_{i=1}^{n_2+1} \frac{dU_1}{(U_1)^{p-2}} \frac{dU_2}{(U_2)^{p-2}}.
\]

or

\[
(8.6.11) \quad \text{const.} \exp \left[ -\frac{1}{2} \text{tr} \begin{pmatrix} UD_{1+c}(p) & U^* \gamma \cosh(u_{11}\sqrt{c}\gamma) \end{pmatrix} \right] \prod_{i=1}^{n_2+1-p} \frac{dU_3}{(U_3)^{p-1-i}} \prod_{i=1}^{n_2+1} \frac{dU_1}{(U_1)^{p-2}} \frac{dU_2}{(U_2)^{p-2}}.
\]

Now putting \( UD_{1+c} = \begin{pmatrix} U_1 & U_2 & U_3 \end{pmatrix} \begin{pmatrix} 1+c & 0 & 1 \\ 0 & 1 & p-1 \\ 1 & p-1 & 1 \end{pmatrix} = \begin{pmatrix} V_1 & V_2 & V_3 \end{pmatrix} \begin{pmatrix} 1+c & 0 & 1 \\ 0 & 1 & p-1 \\ 1 & p-1 & 1 \end{pmatrix} = V \text{ (say), we}
\]

have for \( V \) and \( c \) the joint distribution

\[
(8.6.12) \quad \text{const.} \exp \left[ -\frac{1}{2} \text{tr} \begin{pmatrix} VV^* \gamma \cosh(v_{11}\sqrt{c}\gamma/(1+c)) \end{pmatrix} \right] \prod_{i=1}^{n_2+1-p} \frac{dV_3}{(V_3)^{p-1-i}} \prod_{i=1}^{n_2+1} \frac{dV_1}{(V_1)^{p-2}} \frac{dV_2}{(V_2)^{p-2}} \frac{dV_3}{(V_3)^{p-2}}.
\]

To obtain the distribution of \( c \) by integrating over \( V \), we use the same dodge as in (A.8.8), change over to a solid matrix \( W \) and obtain for \( c \) the distribution

\[
(8.6.13) \quad \text{const.} \frac{\Gamma\left(\frac{p-2}{2}\right)}{F(p-1, p-1)} \int \frac{dW_1}{W_1^{p-1}} \left( \begin{array}{c} n_2+1 \end{array} \right) \int \exp \left( -\frac{1}{2} \text{tr} \begin{pmatrix} WW^* \gamma \cosh(w_{11}\sqrt{c}\gamma/(1+c)) \end{pmatrix} \right) \frac{dW}{(W)^{p-2}}.
\]

where \( F(p-1, p-1) \) is given by (A.8.6.3). To affect the integration over \( W \), denote the row vectors of \( W \) by \( w_i \) \( (i = 1, 2, \ldots, p) \) and make the transformation
\[
\begin{pmatrix}
\begin{pmatrix}
1 & 1 \\
\vdots & \vdots \\
1 & 1 \\
\end{pmatrix}
\end{pmatrix}
\begin{pmatrix}
\begin{pmatrix}
\begin{pmatrix}
\begin{pmatrix}
1 & 1 \\
\vdots & \vdots \\
1 & 1 \\
\end{pmatrix}
\end{pmatrix}
\end{pmatrix}
\] 
\(L(p \times p)\) (where \(L\) is arbitrary such that the first row vector of \(L\) is along \(w_1\)). Thus, though \(L\) will depend on \(w_1\), the Jacobian of this transformation is easily checked to be unity. In this case the new matrix is

\[
\begin{pmatrix}
w_{11} & w_{12} & \cdots & w_{1p} \\
g_{21} & g_{22} & \cdots & g_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
g_{p1} & g_{p2} & \cdots & g_{pp} \\
\end{pmatrix}
= 
\begin{pmatrix}
w_{11} & w_{12} & \cdots & w_{1p} \\
g_{21} & g_{22} & \cdots & g_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
g_{p1} & g_{p2} & \cdots & g_{pp} \\
\end{pmatrix}
\] 
\(G\) (say).

It is easy to see that \(|\tilde{w}| = |w_{11}|^{1/2} = (w_{11})^{1/2} = |G|^{|1/2|}. Also tr \(w w^t\) goes over into tr \(w w^t + \sum_{i=2}^{p} g_{i1}^2 + \text{tr } G G^t\). Now by using (A.8.7.1) it is easy to integrate out

\[
\int \exp \left[ -\frac{1}{2} \left( \sum_{i=2}^{p} g_{i1}^2 + \text{tr } G G^t \right) \right] \prod_{i=2}^{p} dg_{i1} |G|^{|n_2-p+1|} \text{d}G,
\]
over \(G\) and \(g_{i1}(i=2, 3, \ldots, p)\)

and obtain a constant which we absorb into the constant factor and obtain for \(c\) the distribution

\[
(8.6.14) \text{ const.} \left[ \frac{p-2}{2} \frac{\sqrt{n_2+1}}{c/(1+c)} \right]^{n_2+1} \exp \left[ -\frac{1}{2} \sum_{j=1}^{p} w_{i1}^2 \right] \cosh (w_{11} \sqrt{c/(1+c)}) \prod_{j=1}^{p} \sum_{i=1}^{n_2+1} \frac{w_{ij}^2}{j} \prod_{j=1}^{p} \text{d}w_{ij}.
\]

To evaluate this integral we put \(\sum_{j=1}^{p} w_{ij}^2 = r^2\) and \(w_{11} = r \cos \theta\), so that, using (A.8.5.1), we have the integral = constant (which is easily obtained)

\[
x \int_{r=0}^{\infty} \int_{\theta=0}^{\pi/2} \exp \left[ -\frac{1}{2} r^2 \right] \cosh (r \cos \theta \sqrt{c/(1+c)}) r^{n_2-2} \text{dr} \text{d}\theta.
\]

We note
that the integral over $\theta$ could be taken from 0 to $\pi/2$ and the result multiplied by 2. Using now the formula (9) p. 79 and formula (2) p. 393 of Watson's Bessel Functions and remembering the relation between Bessel I and Bessel J functions, this integral reduces to

$$\text{const.} \int_1^{n_2} \frac{n_2!}{n_2!} \left( \frac{p-2}{2} c \gamma \frac{n_2+1}{1+c} \right) \exp \left( -\frac{p-2}{2} c^2 / (1+c) \right) \frac{\Gamma(\frac{n_2+1}{2})}{\Gamma(\frac{p+n_2+1}{2})} \text{I}_1 \left( \frac{p-2}{2} c \gamma \frac{n_2+1}{1+c} \right) \int_1^\gamma J_1 \left( \frac{p-2}{2} c, \frac{1}{2} \frac{p}{(1+c)} \right).$$

(8.6.15)

It is of interest to note that on the null hypothesis $\gamma = 0$, the confluent hypergeometric function reduces to a constant and (8.6.15) goes over, as it should, into

(8.6.9).

(8.7) The distribution of the roots connected with the multivariate regression model of (5.24)-(5.27). Going back to (5.27) and noting the identity of this distribution form with that given by (5.21), it is easy to check from section (4.7.5) that the distribution of the characteristic roots of $(Z_1Z_1')(Z_2Z_2')^{-1}$ (or $c_i$'s, say) could not involve as parameters anything except the characteristic roots of $(\mu TT'\mu')\Sigma^{-1} = \gamma_1$'s (say). The problem of obtaining the distribution of

$$c\int(Z_1Z_1')(Z_2Z_2')^{-1}$$

can thus be thrown back, when $p \leq q$, on the case (i) of (8.6) and, when $p > q$, on the case (ii) of (8.6) in both cases putting $n_1 = q, n_2 = n-q$.

The complete reduction of the distribution problem, i.e., the derivation of the joint distribution of the roots on the null hypothesis $\gamma = 0$ (i.e. 1, 2, ..., p) (which, in this case, happens if and only if $\mu = 0$), can be affected in exactly the same manner as for the distribution on the null hypothesis in the cases (i) and (ii) of section (8.6). Turning now to $(Z_1Z_1')(Z_2Z_2')^{-1}$ and checking with (5.24)-(5.27), we see that

$$(Z_1Z_1')(Z_2Z_2')^{-1} = (Y_{11}'L_1Y')(Y_{22}'L_2Y')^{-1} = \int_{Y'}(W')^{-1}W_{YY'}W_{YY'}^{-1}W_{Y'}^{-1} \left( Y_{11}' \right)^{-1} \left( Y_{22}' \right)^{-1}$$

so that

$$c\int(Z_1Z_1')(Z_2Z_2')^{-1} = c \left( 1 - e \right) \text{where e's are the roots of the equation in e:}$$

$$\left| eY_{YY'}(W_{YY'})^{-1}Y_{YY'} \right| = 0.$$  Going back to (5.24)-(5.27) notice further that $Y_{YY'} = XX' - nxx', W_{YY'} = UU'-nmu'$, $Y_{YY'} = XU'-nxu'$. Also with regard to the population roots
\( \gamma \)'s (\( = c(\mu T\mu')\Sigma^{-1} \)), notice that \( \frac{\mu T\mu}{\Sigma} = \nu\nu' = \nu\nu - \mu\mu' \) so that
\[ \gamma = c\int \mu(\nu\nu' - \mu\mu')\mu' \Sigma^{-1} \nu. \]

An important special case is that of \( p = 1 \) which we can handle by putting, in case (i) of (8.6), \( p = 1, n_1 = q \) and \( n_2 = n-q, L_1(p \times n_1) = L_1(1 \times q) \) and \( A(1 \times 1) = a \) (say), a scalar. Substituting in (8.6.2) we have for \( c \) (note that there is only one non-zero \( c \) here) the distribution
\[ (8.7.1) \quad 2^{\Gamma(1/2m^2)}F(1, n-q) c^{\frac{a-2}{2}} dc \]

\[ x \int_{a=0}^{\infty} \sum_{q_2, \ldots, q_q} \exp \left( -\frac{1}{2}(a^2(1+c) + \gamma + 2a\sqrt{cy} l_1) - \frac{q}{2} s^{n-1} d\nu_1/(1- \sum_{i=2}^q y_i^2)^{1/2} \right). \]

To evaluate the integral proceed as follows
\[ \int \exp \left( +a\sqrt{cy} l_1 \right) \left( \sum_{i=2}^q y_i^2 \right)^{1/2} = \int \cosh \left( a\sqrt{cy} \cos \theta \right) \sin^q \theta \, d\theta, \]
so that substituting in (8.7.1) we have the integral under (8.7.1) reducing to
\[ \text{const.} \int_{a=0}^{\infty} \int_{\theta=0}^{\pi/2} \exp \left( -\frac{1}{2}a^2(1+c) \right) \cosh \left( a\sqrt{cy} \cos \theta \right) \, da \, (\sin \theta)^{q-2} \, d\theta. \]

Taking account of the discussion after (A.11.6.1b) this integral reduces to
\[ \text{const.} \Gamma(n/2) \left( \frac{\Gamma(q/2)\Gamma(n-q/2)}{\Gamma(n/2)} \right)^{\frac{a-2}{2}} \Gamma(q/2) \Gamma(n-q/2) \quad \text{F}(n/2; q/2, 1/2), \]
the const. factor being easily evaluated (since the constant factor at each stage is known and carried over to the next stage).

If \( \gamma = 0 \), this reduces to
\[ (8.7.3) \quad \frac{\Gamma(n/2)}{\Gamma(q/2)\Gamma(n-q/2)} \left[ \frac{\frac{q-2}{2}}{\frac{c}{2} \Gamma(1+c)^2} \right]. \]

For \( \alpha \) given by \( c = \alpha/(1-\alpha) \) we have on the non-null and the null hypothesis the
respective distributions

\[
\frac{\Gamma(n/2)}{\Gamma(q/2)\Gamma(n-q/2)} \exp(-\frac{J}{2}) (1-c) \frac{n-q-2}{2} e^{-\frac{q-2}{2}} dc \Gamma\left(\frac{n}{2} ; \frac{q}{2} ; \frac{1}{2} \gamma\right)
\]

and

\[
\frac{\Gamma(q/2)}{\Gamma(q/2)\Gamma(n-q/2)} \left[ \frac{n-q-2}{2} e^{-\frac{q-2}{2}} dc \right].
\]

(8.8) The reduction of the various joint distributions of the roots (on the respective null hypothesis) to a common standard form. On the respective null hypotheses consider the distributions (8.3.5), (8.4.5), (8.5.4) and (8.5.8) and check that they can all be reduced to the following common standard form (expressed in terms of (8.4.5).

\[
\frac{1}{s!} \int \frac{n^{s/2}}{\prod \Gamma(\frac{1+2i}{2})} \frac{\prod \Gamma(\frac{1+2i}{2})\Gamma(\frac{2m+i+1}{2})}{\prod \Gamma(\frac{1}{2}-\frac{i}{2})\Gamma(\frac{2m+i+1}{2})} \mathcal{J} \prod x_i^m (1-x_i)^{m_2}
\]

\[
\int \int \int \prod_{i<j} (x_i-x_j) \int \prod dx_i,
\]

where \(0 \leq x_1, \ldots, x_s \leq 1\) and, a.e., \(0 < x_1, \ldots, x_s < 1\). If, however, we order the \(x_i\)'s as \(0 \leq x_1 < x_2 \leq \ldots \leq x_s < 1\), or, a.e., \(0 < x_1 < x_2 < \ldots < x_s < 1\), then the above distribution can be rewritten as

\[
\int \frac{n^{s/2}}{\prod \Gamma(\frac{1+2i}{2})} \frac{\prod \Gamma(\frac{1+2i}{2})\Gamma(\frac{2m+i+1}{2})}{\prod \Gamma(\frac{1}{2}-\frac{i}{2})\Gamma(\frac{2m+i+1}{2})} \mathcal{J} \prod x_i^m (1-x_i)^{m_2}
\]

\[
\int \prod_{i>j} (x_i-x_j) \int \prod dx_i.
\]

It is well known that \(\frac{1}{\prod (x_i-x_j)}\) can be written in another form, namely, as a Vandermonde determinant

\[
\begin{vmatrix}
\frac{1}{s!} \prod (x_i-x_j) \int \prod dx_i
\end{vmatrix}
\]

\[
\begin{vmatrix}
\begin{array}{ccc}
x_1^{s-1} & x_2^{s-1} & \ldots & x_s^{s-1} \\
x_1^{s-2} & x_2^{s-2} & \ldots & x_s^{s-2} \\
\vdots & \vdots & \ddots & \vdots \\
1 & 1 & \ldots & 1
\end{array}
\end{vmatrix}
\]

(8.8.3)
Denoting now the constant factor (within the square brackets) of (8.8.1) or (8.8.2) by \( k(s, m_1, m_2) \), we can rewrite (8.8.1) and (8.8.2) respectively as

\[
(8.8.4) \quad \frac{1}{s!} k(s, m_1, m_2) \prod_{i=1}^{s} \left( x_i^1 (1-x_i^1)^m_2 dx_i \right) \text{mod } \begin{vmatrix}
& x_{s-1}^1 & \cdots & x_{s-1}^s \\
& x_{s-2}^1 & \cdots & x_{s-2}^s \\
& \vdots & \ddots & \vdots \\
& 1 & \cdots & 1
\end{vmatrix},
\]

and

\[
(8.8.5) \quad k(s, m_1, m_2) \prod_{i=1}^{s} x_i^1 (1-x_i^1)^m_2 dx_i \begin{vmatrix}
x_{s-1}^1 & \cdots & x_{s-1}^s \\
x_s^1 & \cdots & x_s^s \\
1 & \cdots & 1
\end{vmatrix}.
\]

or

\[
(8.8.6) \quad k(s, m_1, m_2) \begin{vmatrix}
x_s^1 (1-x_s^1) & \cdots & x_s^1 (1-x_s^1)^m_2 \\
x_1^1 (1-x_1^1) & \cdots & x_1^1 (1-x_1^1)^m_2 \\
\vdots & \ddots & \vdots \\
x_s^1 (1-x_s^1)^m_2 & \cdots & x_s^1 (1-x_s^1)^m_2 \\
x_1^1 (1-x_1^1)^m_2 & \cdots & x_1^1 (1-x_1^1)^m_2 \\
\end{vmatrix} \prod_{i=1}^{s} dx_i.
\]

The following substitutions are to be made in (8.3.1) or (8.8.4) in order to obtain (8.3.5), (8.4.5), (8.5.4) and (8.5.8). For (8.3.5) put \( x_i = c_i / (1+c_i) \), \( s = p, \)
\[
m_1 = \frac{n_1-p-1}{2}, \quad m_2 = \frac{n_2-p-1}{2}, \quad \text{for (8.4.5) put } x_i = c_i, \quad s = p, \quad m_1 = \frac{p-q-1}{2}, \quad m_2 = \frac{n_2-p-1}{2},
\]
for (8.5.4) put \( x_i = c_i / (1+c_i) \), \( s = n_1, \quad m_1 = \frac{n_1-p-1}{2}, \quad m_2 = \frac{n_2-p-1}{2}, \) and for (8.5.8) put \( x_i = c_i / (1+c) \), \( s = n_1, \quad m_1 = \frac{n_1-1}{2}, \quad m_2 = \frac{n_2-1}{2}. \)

It is of interest to note that ordering the \( x_i \)'s is exactly equivalent to ordering the \( c_i \)'s of (8.3.5), (8.5.4) and (8.5.8); in other words, \( 0 \leq x_1 \leq \ldots \leq x_s \leq 1 \iff 0 \leq c_1 \leq \ldots \leq c_s < \infty \) and \( 0 < x_1 < \ldots < x_s < 1 \iff 0 < c_1 < \ldots < c_s < \infty. \)

9. On the c.d.f. of the largest or of the smallest root. In this section starting from (8.8.2) or (8.8.5) we shall obtain the c.d.f. of \( x_s \), i.e., \( P(x_s \leq x_0) \),
where \( x_0 \) is a given constant \( \leq 1 \) (from which it is easy to check that one can obtain the c.d.f. of \( x_1 \) by merely interchanging \( m_1 \) and \( m_2 \)), and also obtain \( P(x_0^i \leq x_1 \leq x_s \leq x_0) \), where \( x_0^i \) and \( x_0 \) are also given constants subject to \( 0 \leq x_0^i \leq x_0 \leq 1 \). Starting from (8.8.5) and putting in (A.9.6.13) \( m_1 = m_1 + i - 1 \) and \( n = m_2 \) (i = 1, 2, \ldots, s) we have

\[
P(0 \leq x_1 \leq \ldots \leq x_s \leq x) = P(x_s \leq x) = k(s, m_1, m_2)
\]

\[
x \in \left\{ x_1; \begin{pmatrix} m_1 + s - 1, m_2 & m_1 + s - 2, m_2 & \cdots & m_1 m_2 \\ \vdots & \ddots & \ddots & \vdots \\ m_1 + s - 1, m_2 & m_1 + s - 2, m_2 & \cdots & m_1 m_2 \end{pmatrix} \right\},
\]

where \( \theta \) is to be successively and completely reduced with the help of the fundamental formula (A.9.6:13).

For the c.d.f. of the smallest root \( x_1 \) we note that \( P(x_1 \leq x) = 1 - P(x_1 > x) = 1 - P(x \leq x_1 \leq \ldots \leq x_s \leq 1) \). Going back to the c.d.f. of \( (x_1, \ldots, x_s) \) and using the transformation \( x_i = 1 - z_i \) (i = 1, 2, \ldots, s) we have

\[
k(s, m_1, m_2) \int_{x_1}^{1} dx_1 \int_{x_1}^{1} dx_2 \cdots \int_{x_1}^{1} dx_s \prod_{i=1}^{s} \frac{m_1}{x_i(1-x_i)} \prod_{i,j=1}^{s} (x_i-x_j)
\]

\[
= k(s, m_1, m_2) \int_{0}^{1-x} dz_1 \int_{0}^{z_1} dz_2 \cdots \int_{0}^{z_{s-1}} dz_s \prod_{i=1}^{s} \frac{m_2}{z_i(1-z_i)} \prod_{i,j=1}^{s} (z_i-z_j).
\]

It is now easy to see that the integral on the right hand side of (9.2) is exactly the same as that on the right hand side of (9.1) with just the interchange of \( m_1 \) and \( m_2 \), which shows that the c.d.f. of the smallest root can be thrown back on that of the largest root and vice versa. The final reduction of the exact c.d.f. of the largest or the smallest root is necessarily lengthy and need not be given here. When \( m_1 + m_2 \) is large, which is the case in most practical applications, there is a good approximation in relatively far simpler terms (especially when percentage points, up to, say, 10 o/o are needed) which will be given in the next
In this section the final reduction for the exact c.d.f. of the largest root in the case of \( s = 2, 3, 4 \) will be given. This is as follows:

(9.3) For \( s = 2 \),

\[
P(x_2 \leq x) = \frac{k(2, m_1, m_2)}{m_1 + m_2 + 2} \left\{ -\phi_0(x; m_1 + 1, m_2 + 1) \beta(x; m_1, m_2) + 2\phi(x; 2m_1 + 1, 2m_2 + 1) \right\}.
\]

(9.4) For \( s = 3 \),

\[
P(x_3 \leq x) = \frac{k(3, m_1, m_2)}{m_1 + m_2 + 3} \left\{ -\phi_0(x; 2m_1 + 3, 2m_2 + 1) \beta(x; m_1, m_2) + 2\phi(x; 2m_1 + 2, 2m_2 + 1) \beta(x; m_1 + 1, m_2) \right. \\
- \frac{1}{k(3, m_1, m_2)} \beta_0(x; m_1 + 2, m_2 + 1) P(x_2 \leq x). \]

(9.5) For \( s = 4 \),

\[
P(x_4 \leq x) = \frac{k(4, m_1, m_2)}{m_1 + m_2 + 4} \left\{ -\phi_0(x; m_1 + 3, m_2 + 1) \frac{1}{k(3, m_1, m_2)} P(x_3 \leq x) + 2\phi(x; 2m_1 + 5, 2m_2 + 1) \right\}
\]

\[
x \left( \frac{1}{k(2, m_1, m_2)} P(x_2 \leq x) + \frac{2}{m_1 + m_2 + 3} \beta(x; 2m_1 + 3, 2m_2 + 1) \right) - \frac{2}{m_1 + m_2 + 3} \beta(x; 2m_1 + 3, 2m_2 + 1) \left\{ -\phi_0(x; m_1 + 2, m_2 + 1) \beta(x; m_1 + 1, m_2) \right. \\
+ 2\beta(x; 2m_1 + 3, 2m_2 + 1) \right\} \left. - \frac{m_1 + 2}{k(2, m_1, m_2)} P(x_2 \leq x) \right\}.
\]

Again starting from (6.8.5) and putting in (A.9.7.2) \( m_i = m_1 + i - 1 \) and \( n = m_2 \) \( i = 1, 2, \ldots, s \) we have:

(9.6) \( P(x_0 \leq x_1 \leq \cdots \leq x_s \leq x) = P(x_0 \leq x_1 \leq x_s \leq x) \)

\[
= \frac{k(s, m_1, m_2)}{x, x_0; \left( \begin{array}{ccc}
m_1 + s - 1, m_2 & m_1 + s - 2, m_2 & \cdots & m_1, m_2 \\
m_1 + s - 1, m_2 & m_1 + s - 2, m_2 & \cdots & m_1, m_2 \\
& \cdots & \cdots & \cdots \\
m_1 + s - 1, m_2 & m_1 + s - 2, m_2 & \cdots & m_1, m_2 \\
\end{array} \right)}
\]

where \( \beta \) is to be successively and completely reduced with the help of the fundamental formula (A.9.7.2). Below is given the complete reduction of the left side of (9.6) for \( s = 2, 3, \) and 4, the left side of (9.6) being conveniently denoted by
$$P_2, P_3, P_4, \text{ etc.}$$

(9.7) $P_2 = \frac{k(2,m_1,m_2)}{(m_1+m_2+2)} \beta(x,x_0;2m_1+1,2m_2+1) - \beta(x,x_0;m_1,m_2) \beta_0(x;m_1+1,m_2+1)$

$$+ \beta_0(x_0;m_1+1,m_2+1) \beta_0(x_0;m_1+1,m_2+1)$$

(9.8) $P_3 = \frac{k(3,m_1,m_2)}{(m_1+m_2+3)} \beta(x,x_0;3m_1,m_2) \beta(x,x_0;2m_1+3,2m_2+1)$

$$- 2\beta(x,x_0;2m_1+1,m_2) \beta(x,x_0;2m_1+2,2m_2+1)$$

$$- \frac{P_2}{k(2,m_1,m_2)} \left\{ \beta_0(x;m_1+2,m_2+1) - \beta_0(x_0;m_1+2,m_2+1) \right\}$$

and

(9.9) $P_4 = \frac{k(l,m_1,m_2)}{(m_1+m_2+4)} \beta(x,x_0;2m_1+5,2m_2+1) \frac{P_2}{k(2,m_1,m_2)}$

$$- \frac{P_3}{k(3,m_1,m_2)} \left\{ \beta_0(x;m_1+3,m_2+1) + \beta_0(x_0;m_1+3,m_2+1) \right\}$$

$$+ \frac{2\beta(x,x_0;2m_1+3,2m_2+1)}{(m_1+m_2+3)} \left\{ -\beta_0(x;m_1+2,m_2+1) \beta(x,x_0;m_1+1,m_2) \right\}$$

$$- \beta_0(x_0;m_1+2,m_2+1) \beta(x,x_0;m_1+1,m_2) \beta_0(x,x_0;2m_1+3,2m_2+1)$$

$$- \frac{2\beta(x,x_0;2m_1+2,m_2+1)}{(m_1+m_2+3)} \left\{ -\beta_0(x;m_1+2,m_2+1) \beta(x,x_0;m_1+1,m_2) \right\}$$

$$- \beta_0(x_0;m_1+2,m_2+1) \beta(x,x_0;m_1+1,m_2) + 2\beta(x,x_0;2m_1+2,2m_2+1) + \frac{m_1+2}{k(2,m_1,m_2)} P_2 \right\}$$

For larger values of s the reduction of $P_s$ to exact expressions like those given by the right side of (9.7)-(9.9) will no doubt be increasingly lengthy but the remarks made after (9.7) will apply to this situation as well, so that if $m_1+m_2$ is reasonably large, as would be the case in most practical situations, we can use much shorter expressions as good approximations.

10. General observations on the operating characteristics of the test regions (7.4.2), (7.4.5), (7.4.8) and (7.4.11) corresponding to the situations (i)-(iv) of section 7. As of the moment the exact (small sample) power functions of the regions
(7.4.2), (7.4.5), (7.4.8) and (7.4.11) seem to be, in the general cases, quite intractable. At any rate, so far as the author is aware, no method is known at the moment by which the requisite distribution problems could be solved and the final c.d.f.'s be given except in very symbolic (and, for practical purposes, quite useless) forms. However, it is possible, even without exact expressions for the c.d.f.'s, to obtain a number of useful semi-qualitative and semi-quantitative properties of the power functions, which, as will be presently seen, are about all that
would really matter for most practical purposes.

We observe from (A.7.1), (A.7.2), (A.7.5) and (A.7.3) respectively that the powers of the critical regions (7.4.2), (7.4.5), (7.4.8) and (7.4.11) depend only on the corresponding sets of populations roots, namely, \( c(\Sigma_{0}^{-1})'s \) (to be called \( \gamma'\)s) for the first case, \( c(\Sigma_{1}^{-1})'s \) (to be called \( \gamma'\)s) for the second case, \( c(\Sigma^{*}_{11}^{-1})'s \) (to be called \( \gamma'\)s) for the third case and \( c(\Sigma_{11}^{-1},\Sigma_{22}^{-1})'s \) (to be called \( \gamma'\)s) for the fourth case. For convenience we write down the respective powers for the four cases as

\[
(10.1) \quad P_{c_{p}} \geq c_{a}(p,n) \quad \text{and/or} \quad c_{1} \leq c_{1a}(p,n) / H_{r} = P(a,p,n; \gamma_{1}, \gamma_{2}, \ldots, \gamma_{p}),
\]

\[
(10.2) \quad P_{c_{p}} \geq c_{a}(p,n_{1},n_{2}) \quad \text{and/or} \quad c_{1} \leq c_{1a}(p,n_{1},n_{2}) / H_{r} = P(a,p,n_{1},n_{2}; \gamma_{1}, \gamma_{2}, \ldots, \gamma_{p}),
\]

\[
(10.3) \quad P_{c_{p}} \geq c_{a}(p,k-1,n-k) / H_{r} = P(a,p,k-1,n-k; \gamma_{1}, \gamma_{2}, \ldots, \gamma_{r}), \quad \text{and}
\]

\[
(10.4) \quad P_{c_{p}} \geq c_{a}(p,q,n) / H_{r} = P(a,p,q,n; \gamma_{1}, \gamma_{2}, \ldots, \gamma_{p}).
\]

Notice that, depending on the rank of \( \Sigma^{*} \) and \( \Sigma_{12} \), some of the \( \gamma'\)s of (10.3) and (10.4) might be zero but the most general case will be one in which as many as are set down will be positive. Notice also that in (10.3), \( r = \min(p,k-1) \). Recall now from (A.3.3) that for (10.1), \( 0 \leq all \gamma'\)s < \( \infty \), from (A.1.9) that for (10.2) and (10.3), \( 0 \leq all \gamma'\)s < \( \infty \) and from (A.1.14) that for (10.4) \( 0 \leq all \gamma'\)s < \( 1 \).

With this introduction we shall consider the power functions (10.1) and (10.2) for the problems of one dispersion matrix and two dispersion matrices and the power function (10.3) for the problem of multivariate analysis of variance and the power function (10.4) for the problem of independence between two sets of variates. In section 11, to each power function a lower bound will be obtained which will be called a lower bound function and which will be seen to involve (aside from the degrees of freedom) just those deviation parameters that occur in the power function itself. In section 12 it will be shown that for each power function the lower bound function monotonically increases as each deviation parameter separately tends away from its value under the null hypothesis. Although, under the null hypothesis,
the lower bound function does not assume the value a which is the significance level of the test, this value is attained soon enough under deviations from the hypothesis. Thus the power function stays greater than a monotonically increasing function of each deviation parameter and is also shown to be unbiased against all deviations from the hypothesis for which the lower bound function is greater than or equal to the size a of the test. In section 13, for each of the power functions (10.3) and (10.4) another such monotonic lower bound function is obtained which is believed to be closer than the lower bound functions of section 11; also for each of the power functions (10.1) and (10.2) some near monotonic properties are proved.

11. Lower bounds on the power functions (10.1)-(10.4) of the test regions (7.4.2), (7.4.5), (7.4.8) and (7.4.11). The lower bounds are obtained in three different stages to be called (11 a), (11 b) and (11 c), \( \sqrt{21 \cdot 7} \).

11 a. Reduction to canonical forms. Without any loss of generality we can, for the case of (7.4.2), start right from the canonical form (A.7.1.1); for the case of (7.4.5) from the canonical form (A.7.2.1); for the case of (7.4.8) from the canonical form (A.7.4.5); and for the case of (7.4.11) from the canonical form (A.7.3.5). For the case of (7.4.2) there is an additional point to be noted. Putting together (A.7.1) and (7.4.1) it is clear that it will be appropriate, instead of using as we did in (A.7.1) the transformation \( X(p \times n) = \mu(p \times p)Y(p \times n) \), where \( \Sigma = \mu D^T \mu' \) (\( \gamma \)'s being the roots of \( \Sigma \)), to use the transformation \( X(p \times n) = \mu(p \times p)A_\Sigma (p \times p)Y(p \times n) \), where \( \Sigma_0 = \Delta_0 \Delta_0^T \) and \( \Delta_0^{-1} \Sigma(\Delta_0^T)^{-1} = \mu D^T \mu' \), \( \gamma \)'s being the roots of \( \Delta_0^{-1} \Sigma(\Delta_0^T)^{-1} \), i.e., of \( \Sigma(\Delta_0 \Delta_0^T)^{-1} \), i.e., of \( \Sigma \Sigma_0^{-1} \). Under this transformation the \( \gamma \)'s of the canonical form (7.1.1) will really be the roots of \( \Sigma \Sigma_0^{-1} \) and the roots of \( (XY')/n \) will really be the roots of the equation (7.4.1) and thus we have an exact tie-up with the problem involving the power function of (7.4.2).

11 b. The inclusion within the test regions (7.4.2), (7.4.5), (7.4.8) and (7.4.11) of regions having simpler probability measures (under the respective non-null hypothesis).
(i) We recall from (7.4) and the canonical form (A.7.1.1) that the test region (7.4.2) is really $U_{a} \sqrt{a} Y^{a}_{1}/n_{a} a \geq c_{2a}(p,n)$ or $\leq c_{1a}(p,n)$, where $Y$ has the distribution (A.7.1.1). We also notice from the canonical form (A.7.1.1) that the $p$ functions $a_{i}^{Y^{a}_{i}/n_{a} a}$ $(i = 1, 2, \ldots, p)$ and $a_{i}$ is an $1 \times n$ row vectors having 1 for the $i^{th}$ element and 0 for all other elements) are distributed as $p$ independent $X^{2}$'s with D.F. $n$ each. Putting these two facts together we have that the test region (7.4.2) includes the union of $p$ regions, each composed of the tail ends of an ordinary $X^{2}$-region, all the $p$ $X^{2}$'s being independent.

(ii) We recall from (7.4) and the canonical form (A.7.2.1) that the test region (7.4.5) is really $U_{a} \sqrt{n_{a} a} Y^{a}_{i}/n_{a} a \geq c_{2a}(p,n_{1},n_{2})$ or $\leq c_{1a}(p,n_{1},n_{2})$, where $Y$ has the distribution (A.7.3.1). We also notice from the canonical form (A.7.3.1) that the $p$ functions $n_{a} a_{i}^{Y^{a}_{i}/n_{a} a}$ $(i = 1, 2, \ldots, p)$ and $a_{i}$ is an $1 \times p$ row vector having 1 for the $i^{th}$ element and 0 for all other elements) are distributed as $p$ independent $F$'s with D.F. $n_{1}$ and $n_{2}$ each. Putting these facts together it is easy to see that the test region (7.4.5) includes the union of $p$ regions, each composed of the tail ends of an ordinary $F$-region, all the $p$ $F$'s being independent.

(iii) From the canonical form (A.7.4.5) and from (7.4) we notice that the test region (7.4.8) will really be $U_{a} \sqrt{(n-k)a} Y^{*} a^{*}_{a} /((k-1)a^{*}_{a} a \geq c_{a}(p,k-1,n-k)$, where $Y^{*}$ and $Y$ have the distribution (A.7.4.5). We also notice from the canonical form (A.7.4.5) that the $p$-functions $(n-k)a_{i}^{Y^{*} a^{*}_{a} a}$ $/(k-1)a^{*}_{a} a_{i}$ $(i = 1, 2, \ldots, p)$ and $a_{i}$ is an $1 \times p$ row vector having 1 for the $i^{th}$ element and 0 for all other elements) are distributed as $p$ independent $F$'s, out of which at least $p-r$ are central and at the most $r$ are non-central with non-contralility parameters $(\gamma_{1}, \gamma_{2}, \ldots, \gamma_{r})$. (notice that if $s \leq r = \min(p,k-1)$, then, out of these $\gamma$'s, $s$ will be positive and the rest, i.e., $r-s$ will be zero). Putting these together we observe that the test region (7.4.8) includes the union of $p$ regions, out of which at least $p-r$ are central $F$-regions and at the most $r$ are non-central $F$-regions with non-central
parameters $\gamma_i$'s ($i = 1, 2, \ldots, r$), all $F$'s being independent and each being based on D.F. $n_1$ and $n_2$.

(iv) We notice from the canonical form (A.7.3.5) and from (7.41) that the test region (7.4.11) will really be

$$\frac{1}{2c^2} \left( \frac{a'_1 Y_1 b}{a'_1 Y_1 a} - \frac{b'_1 Y_2 b}{b'_1 Y_2 b} \right) \geq c^2 (p, q, n),$$

where $Y_1$ and $Y_2$ have the distribution (A.7.3.5). We notice further that there are $p$ functions $\left( \frac{a'_1 Y_1 b}{a'_1 Y_1 a} \right)^2 / \left( \frac{a'_1 Y_1 a}{a'_1 Y_1 a} \right)$, with $l$ for the $i$th element and $0$ for all other elements and $b'(1 \times q)$ is a row vector with $l$ for the $i$th element and $0$ for all other elements) which are distributed as the squares of $p$ independent correlation coefficients (some of them central and some non-central, the respective non-centrality or deviation parameters being $\gamma_i$ (notice that out of these $p$ $\gamma_i$'s, $t$ are positive and the rest, i.e., $p-t$ are zero, where $t \leq p \leq q$ is the rank of $\Sigma_{12}$, i.e., of $\Sigma_{12}$ and all lie between $0$ and $1$; the positive $\gamma$'s can be conveniently arranged as $0 < \gamma_1 < \ldots < \gamma_t < 1$). Putting these together it is easy to check that the test region (7.4.11) includes the union of $p$ regions out of which $(p-t)$ are central correlation (square) regions and $t$ are non-central ones, all being independent and each based on D.F. $(n-1)$.

When $q > p$ it is possible to improve on this in the following manner. Pick out linked $a'_i$ and $b'_i$ ($i = 1, 2, \ldots, p-1$) and at the last stage an $a'_p$ with a set of $b'_i$'s ($i = p, p+1, \ldots, q$) such that there are $p$ independently distributed correlation squares, of which $(p-1)$ are total correlation (squares) and the last one is a multiple correlation (square) between the $p$th variate of the $Y_1$-set and the $(p, p+1, \ldots, q)$ variates of the $Y_2$-set. The deviation parameters being $\gamma_i$'s ($0 < \gamma_1 < \ldots < \gamma_t < 1$), we could so arrange that the first $p-t$ sample (total) correlation (squares) had zero deviation parameters to go with, the next $t-1$ sample (total) correlation (squares) had respective (one each) deviation parameters ($\gamma_1, \gamma_2, \ldots, \gamma_{t-1}$) to go with and the last sample (multiple) correlation (square) had $\gamma_t$ to go with. Notice that the distributions of the square of a correlation (central and non-central) are easily available from those of the central and non-central multiple correlation.
aside from the central, i.e., ordinary, \( x^2 \) and \( F \) and the total correlation (squared) distributions (the last one being transformable to an \( F \)-distribution), which are all well known and have their percentage points tabulated, we need, in addition, tables for the c.d.f. of non-central \( F \) and non-central multiple correlation (connected with the multivariate normal population). These tables are available in part [26,29,30] and could be easily extended with modern computing facilities.

It may be noted that if in (11.3) we put \( k = 2 \), i.e., \( s = 0 \) or 1, then each side of (11.3) is computationally accessible, the left side being the power function of Hotelling's \( T^2 \), while the right side is also easily available (in this as in all other cases).

It is of considerable importance at this stage to ask how "good" the lower bounds indicated in (11.1), (11.2) and (11.3) or (11.4) are. A lower bound to the power could be said to be "good" if it were (i) close to the actual power, and/or (ii) if it were itself pretty large, being greater than the level of significance \( \alpha \) for reasonably large values of the deviation parameters and possibly getting larger as those parameters increase. For all the three tests condition (ii) has been numerically checked to be true over a fairly wide range of test values of the several parameters involved. With regard to condition (i), in general, that is, for small samples, not only do we not know the actual power (in which case the search for a lower bound would have been redundant) but at the moment we do not even know an upper bound of the expression: \( \text{actual power} - \text{given lower bound to it} \) ÷ actual power. In large samples, however, the situation improves and it turns out that the relative error is "small", so that the given lower bounds are "good" also in the sense (i).

12. On the monotonic character of the lower bounds on the power functions of the four tests and hence the near monotonic character of the four power functions themselves.

12.1. The problem of one dispersion matrix. For convenience we rewrite (10.1) as
(12.1.1) \( P(\alpha, p, n; \gamma_1, \gamma_2, \ldots, \gamma_p) > 1 - \prod_{i=1}^{P} \frac{c_{1a}(p, n)}{\gamma_i} \leq x^2 \leq \frac{c_{2a}(p, n)}{\gamma_i} \),

each \( x^2 \) being based on D.F. \( n \). Now denoting, for shortness, the factors in the product on the right side of (12.1.1) by \( P_1, \ldots, P_p \), we shall show that \( \frac{\partial P_i}{\partial (\frac{1}{\gamma_i})} \) is positive or negative according as \( \gamma_i \) is > or < 1, or in other words, \( P_i \) decreases as \( \gamma_i \) tends away from 1, provided that \( c_{1a} \) and \( c_{2a} \) are so chosen that

\[ \sum \frac{\partial P_i}{\partial (\frac{1}{\gamma_i})} = 0. \]

**Proof.** Aside from a constant and positive factor of proportionality which is free from \( \gamma_i \), we have

(12.1.2) \[ P_i = \int_{c_{1a}/\gamma_i}^{c_{2a}/\gamma_i} e^{-\frac{1}{2}x^2} \frac{x^{n-2}}{(2\pi)^{n/2}} d(x^2), \]

and thus

(12.1.3) \[ \frac{\partial P_i}{\partial (\frac{1}{\gamma_i})} = e^{-\frac{c_{2a}/\gamma_i}{2}} \frac{(\frac{c_{2a}}{\gamma_i})^{n/2}}{2} c_{2a} - e^{-\frac{c_{1a}/\gamma_i}{2}} \frac{(\frac{c_{1a}}{\gamma_i})^{n/2}}{2} c_{1a}, \]

where \( e^{-\frac{c_{2a}/\gamma_i}{2}} \frac{n}{2} = e^{-\frac{c_{1a}/\gamma_i}{2}} \frac{n}{2} = 0. \) It is easy to check from (12.1.3) that

\[ \frac{\partial P_i}{\partial (\frac{1}{\gamma_i})} \] is positive if \( \gamma_i > 1 \) and negative if \( \gamma_i < 1 \) and also that \( P_i \to 0 \) as \( \gamma_i \to \infty \) or \( \to 0. \)

Thus the right side of (12.1.1) monotonically increases as each \( \gamma_i \), separately, tends away from unity and the left side, which is the power function of the test, always stays greater than this monotonic function. Furthermore although at \( H_0 \), i.e., when all \( \gamma_i's = 1 \), this monotonic function is < \( \alpha \), it becomes greater than or equal to \( \alpha \) for all \( \gamma_i's \) satisfying

(12.1.4) \[ \prod_{i=1}^{P} \frac{c_{1a}}{\gamma_i} \leq x^2 \leq \frac{c_{2a}}{\gamma_i} \leq 1 - \alpha. \]

This means that the test itself is unbiased at least against all alternatives \( \gamma_i's \)
satisfying (12.1.1).

12.2. The problem of two dispersion matrices. As in the previous case we rewrite (10.2) as

\[ P(\gamma_1, n_1, n_2) > 1 - \sum_{i=1}^{P} \frac{c_{1a}(p, n_1, n_2)}{\gamma_i} \leq F \leq \frac{c_{2a}(p, n_1, n_2)}{\gamma_1} \]

each \( F \) being based on D.F. \( n_1 \) and \( n_2 \). Now denoting, for shortness, the factors in the product on the right side of (12.2.1) by \( P_1, P_2, \ldots, P_p \) we can show exactly as in the previous case that \( \partial P_i / \partial (1/\gamma_1) \) is positive or negative according as \( \gamma_1 > 1 \) or \( \gamma_1 < 1 \), provided that \( c_{1a} \) and \( c_{2a} \) are so chosen that \( \sum_{i=1}^{P} \partial P_i / \partial (1/\gamma_1) \gamma_1 = 0 \). It is also easy to check that \( P_i \to 0 \) as \( \gamma_1 \to 0 \) or \( \gamma_1 \to \infty \). Thus, as before, the right side of (12.2.1) monotonically increases to 1 as each \( \gamma_1 \) separately, tends away from unity and the expression on the left side of (12.2.1) which is the power function of the test always stays greater than this monotonic function. As in the previous case it follows here also that the test itself is unbiased at least against all alternative \( \gamma_1 \)'s satisfying

\[ \sum_{i=1}^{P} P_i \frac{c_{1a}}{\gamma_1} \leq F \leq \frac{c_{2a}}{\gamma_1} \leq 1 - \alpha. \]

12.3. The problem of multivariate analysis of variance. We rewrite (10.3) as

\[ P(\alpha, p, k-1, n-k; \gamma_1, \gamma_2, \ldots, \gamma_s) > 1 - \sum P(\text{central } F < c_{\alpha}(p, k-1, n-k)) \frac{1}{\gamma_i} \]

each \( F \) being based on D.F. \( k-1 \) and \( n-k \) and \( s = \min(p, k-1) \). It is well known that \( P(\text{non-central } F < c_{\alpha} | \gamma) \) is a monotonically decreasing function of \( \sqrt{\gamma} \), which has been and can be proved in various ways, perhaps the simplest proof being the following. It is well known that with a canonical p.d.f., we can, except for a constant and positive factor of proportionality not involving \( \gamma \), write
12.4. The problem of test of independence between two sets of variates. We rewrite (10.5) as

\[(12.4.1) \quad P(a,p,q,n;\gamma_1,\gamma_2,\ldots,\gamma_p) > 1 - \prod_{i=1}^{p-1} P(\text{non-central } r^2 < c_a(p,q,n) \mid \gamma_i) \times P(\text{non-central } R^2 < c_a(p,q,n) \mid \gamma_p),\]

where all \( r^2 \)'s are based on D.F. n-1 and \( R^2 \) is the square of a multiple correlation based on D.F. n-1 and q-p and where \( \gamma_p \) is the largest population canonical correlation coefficient. Notice that for a particular alternative some of the \( \gamma \)'s might be zero and in any case the \( \gamma \)'s vary from zero to 1. As in the previous case it is well known and can be proved in various ways that both \( P(\text{non-central } r^2 < c_a \mid \gamma) \) and \( P(\text{non-central } R^2 < c_a \mid \gamma_p) \) are each a monotonically decreasing function of \( \sqrt{\gamma} \), which \( \rightarrow 0 \) as \( \sqrt{\gamma} \rightarrow 1 \). The simplest proof of this theorem can be developed exactly on the same lines as in the previous case. But this need not be spelled out here.

Thus the right side of (12.4.1) is a monotonically increasing function of each \( \sqrt{\gamma_i} \) separately, tending to unity as each \( \sqrt{\gamma_i} \rightarrow 1 \), and the left side of (12.4.1), which is the power function of the test, stays greater than this monotonic function. As in the previous case, this test is unbiased at least against all alternatives satisfying

\[(12.4.2) \quad \prod_{i=1}^{p-1} P(\text{non-central } r^2 < c_a \mid \gamma_i) \times P(\text{non-central } R^2 < c_a \mid \gamma_p) \leq 1 - \alpha.\]

13. On two other monotonic lower bound functions, one for the multivariate analysis of variance test and the other for the test of independence, and some further properties of the powers of the test on one dispersion matrix and on two dispersion matrices.

(13.1) Another lower bound function for the power of the multivariate analysis of variance test. We start from the canonical form (A.7.5.5) and denote by \( c_s \) the largest characteristic root of \( (Y_1'Y_1)(Y_2'Y_2)^{-1} \), by \( H_0 \) the \( H(\gamma_i = 0) \) (\( i = 1,2,\ldots,s \))
and by \( K \) its complement, and observe that for a given positive \( c_0 \), \( P(c_s \leq c_0 \mid H) \) is a function of \( \gamma'_1, \gamma'_2, \ldots, \gamma'_s = \psi'(\gamma_1, \gamma_2, \ldots, \gamma_s) \), say. We shall prove that

\[
(13.1.1) \quad P(c_s \leq c_0 \mid H), \text{ i.e., } \psi'(\gamma_1, \gamma_2, \ldots, \gamma_s) \text{ stays less than a monotonically decreasing function of each } |\sqrt{\gamma_1}| \text{ separately (notice that each } \gamma_i \geq 0), \text{ which is different from the decreasing function on the right side of (12.3.1)}.
\]

**Proof.** We recall from (A.2.2) that the largest characteristic root \( c_r \) of 

\[
(Y_1Y_1')(Y_2Y_2')^{-1}
\]

can be written as \( \operatorname{Sup}_a \frac{(a'y_1y_1a)}{(a'y_2y_2a)} \) and the domain \( c_r \leq c_0 \) can be rewritten as

\[
(13.1.2) \quad \operatorname{Sup}_a \frac{(a'y_1y_1a)}{(a'y_2y_2a)} \leq c_0, \text{ or } \bigcap_a \sqrt{\frac{(a'y_1y_1a)}{(a'y_2y_2a)}} \leq c_0.
\]

It is easy to see now that the canonical p.d.f. based on (A.7.5.5) can be rewritten as

\[
(13.1.3) \quad \text{Constant } \exp \int \frac{1}{2} \left( \sum_{i=1}^p (\sum_{j=1}^{n_1} x_{ij}^2) + \sum_{i=1}^p (\sum_{j=1}^{n_2} y_{ij}^2) \right) d
\]

and the region (12.1.2) can be rewritten as

\[
(13.1.4) \quad \operatorname{Sup}_a \frac{a'(x+\delta)(x'+\delta')a}{a'y'a} \leq c_0, \text{ or } \bigcap_a \frac{a'(x+\delta)(x'+\delta')a}{a'y'a} \leq c_0
\]

where \( \delta(p \times n_1) \) is such that \( \delta_{ij} = |\sqrt{\gamma_i}| \) (if \( i = j = 1, 2, \ldots, s \)) and = 0 otherwise, and where

\[
(13.1.5) \quad Y_1(p \times n_1) = X(p \times n_1) + \delta(p \times n_1) \text{ and } Y_2(p \times n_2) = X(p \times n_2).
\]

Notice that all the components of \( X \) and \( Y \) will vary from \(-\infty\) to \( \infty \). Notice also that \( r = \min(p, n_1) \) and \( s, \text{ i.e., the non-zero population roots might go up to } r \).

Observe further that the constant factor in (13.1.3) does not, in this case, involve the \( \gamma_i \)'s. The problem is now one of integrating (13.1.3) over the domain (13.1.4) (which let us call \( \bigcap_a \), for shortness) and showing that the integral stays greater than a monotonically decreasing function of each \( \gamma_i \) on \( |\sqrt{\gamma_1}| \), separately. It will suffice to show the monotonic character of this integral with respect to variation of, say, \( |\sqrt{\gamma_1}| \). To this end, remembering that \( a'_1 \) is a non-null row vector \( (a'_1, a'_2, \ldots, a'_p) \), we might, without any loss of generality, put \( a'_1 = 1 \) and rewrite (13.1.4) as
\[(13.1.6) \quad \bigvee_{p=1}^{n_1} \left( x_{11} + \frac{1}{y_1} \right)^{\frac{1}{2}} + a x_{21} + \ldots + a x_{pl} \bigvee_{j=2}^{n_2} \left( \sum_{i=1}^{p} a_i (x_{ij} + \delta_{ij}) \right)^2 \leq c_0 \sum_{j=1}^{n_2} \left( \sum_{i=1}^{p} a_i y_{ij} \right)^2 \right] \]

where \( \delta_{ij} = \frac{x_{ij}}{y_1} \) (if \( i = j = 1, 2, \ldots, s \)) and \( = 0 \) otherwise, and where \( a_1 = 1 \). To carry out the integration of (13.1.3) over (13.1.6), we first integrate out over \( x_{11} \) and then check the total integral, which we call \( I_1 \), is proportional to

\[(13.1.7) \quad I_1 = \int \left[ \inf_{\mathcal{A} \subseteq \mathcal{A}_1} \mathcal{J} \right] \int_{\mathcal{A}_1} \exp(-\frac{1}{2} x_{11}^2) dx_{11} \mathcal{J} \exp(-\frac{1}{2} (z_{x_{21}}^2 + z_{y_{ij}}^2 + z_{y_{ij}}^2)) dx_{21} d y_{ij}, \]

the symbols being defined in the following way. For \( y_{ij} \)'s, as in (12.1.3), \( i = 1, 2, \ldots, p \) and \( j = 1, 2, \ldots, n_2 \), but for \( x_{11} \)'s, \( i = 2, 3, \ldots, p \). Also \( z_{ij} = x_{ij} \) with \( i = 1, 2, \ldots, p \) and \( j = 2, 3, \ldots, k-1 \), and

\[(13.1.8) \quad \mathcal{J} = \int_{\mathcal{A}_1}^{\mathcal{A}_2} f_{1a} + f_{2a} \mathcal{J} \quad \text{and} \quad \mathcal{J} = \int_{\mathcal{A}_1}^{\mathcal{A}_2} f_{1a} - f_{2a} \mathcal{J}, \]

where

\[(13.1.9) \quad f_{1a} = \sum_{i=1}^{p} a_i x_{1i} \quad \text{and} \quad f_{2a} = \sum_{j=1}^{n_2} c_0 \mathcal{J} (\sum_{i=1}^{p} a_i y_{ij})^2 - \sum_{j=1}^{n_2} c_0 (z_{ij} + \delta_{ij})^2 \mathcal{J}^{1/2}. \]

Furthermore, (i) the constant of proportionality in (13.1.7) is free from \( \gamma_1 \)'s, (ii) \( x_{21}, \ldots, x_{pl} \) vary from \( -\infty \) to \( \infty \) while \( y_{ij} \)'s and \( z_{ij} \)'s from \( -\infty \) to \( \infty \) subject to \( f_{2a} \) always staying real, (iii) for \( f_{2a} \) only the positive square root is to be taken, (iv) \( f_{1a} \) and \( f_{2a} \) are free from \( \gamma_1 \). Now with \( a_1 = 1 \), let \( a_1^* \) denote the value of \( a \) for which \( f_{2a} \) is a minimum. Then it is clear that this \( a_1^* \) is free from \( \gamma_1 \) and \( x_{1i} \)'s and is a function of \( z_{ij} \)'s, \( y_{ij} \)'s, \( c_0 \) and possibly also of \( \delta_{ij} \)'s. Notice that \( a_1^* = 1 \). Also let \( \mathcal{J}^* \) and \( \mathcal{J}^* \) stand for the values of \( \mathcal{J}_{1a} \) and \( \mathcal{J}_{2a} \) on substitution of \( a_1^* \) for \( a_1 \). It is now clear that \( \inf_{\mathcal{A}_1} \mathcal{J}_{1a} < \mathcal{J}_{1a}^* \), \( \sup_{\mathcal{A}_1} \mathcal{J}_{2a} > \mathcal{J}_{2a}^* \), so that

\[\text{Interval} \sup_{\mathcal{A}_1} \mathcal{J}_{2a} \text{Inf}_{\mathcal{A}_1} \mathcal{J}_{1a} \leq \text{Interval} \sup_{\mathcal{A}_1} \mathcal{J}_{2a} \mathcal{J}_{1a}^*. \]

Let us now introduce an \( I_1^* \) such that, aside from a constant and positive factor of proportionality (the same as for \( I_1 \)) it
is defined by

\[(13.1.10) \quad I_1^* = \int \int \int \int_{2a^*} dx_{il} dz_{ij} dy_{ij} \exp \left\{ -\frac{1}{2} \left( \Sigma x_{il}^2 + \Sigma z_{ij}^2 + \Sigma y_{ij}^2 \right) \right\} \exp \left\{ -\frac{1}{2} \left( \Sigma x_{il}^2 + \Sigma z_{ij}^2 + \Sigma y_{ij}^2 \right) \right\} \]

It will be seen that, while \( I_1 \) is the integral of the, a.e., positive function (13.1.3) over the domain (13.1.6) which is the intersection of a class of domains, \( I_1^* \) is the integral of the same, a.e., positive function (13.1.3) over the intersection of a subclass of the previous class. In fact, the subclass is formed by excluding from (13.1.6) all \( a \)'s for which \( \inf_{2a^*} \frac{\gamma_{1a}}{\gamma_{2a}} < \frac{\gamma_{1a}}{\gamma_{2a}^*} \) and/or \( \frac{\gamma_{2a}^*}{\gamma_{2a}} < \frac{\gamma_{2a}}{\gamma_{2a}^*} \). This shows that \( I_1 < I_1^* \).

It is now easy to check that, aside from a constant and positive factor of proportionality, we have

\[(13.1.11) \quad \frac{\partial I_1^*}{\partial (\sqrt{\gamma_1})} = \int \int \int \int \int \int \int \int dx_{il} dz_{ij} dy_{ij} \exp \left\{ -\frac{1}{2} \left( \Sigma x_{il}^2 + \Sigma z_{ij}^2 + \Sigma y_{ij}^2 \right) \right\} \exp \left\{ -\frac{1}{2} \left( \Sigma x_{il}^2 + \Sigma z_{ij}^2 + \Sigma y_{ij}^2 \right) \right\} \]

by using (13.1.9). The domain of variation of \( x_{il} \)'s, \( z_{ij} \)'s and \( y_{ij} \)'s has been already defined immediately after (13.1.9). It will be proved that the expression on the right side of (13.1.11) is negative for positive values of \( \sqrt{\gamma_1} \) and positive for negative values of \( \sqrt{\gamma_1} \), or, in other words, \( I_1^* \) is a monotonically decreasing function of \( |\sqrt{\gamma_1}| \). To prove this we proceed as follows.

We recall from the remarks preceding (13.1.9) that \( f_{2a^*} \) is a function of \( z_{ij} \)'s, \( y_{ij} \)'s, \( c_0 \) and also of \( \delta_{ij} \)'s, while \( f_{la^*} \) is just a linear function of \( x_{il} \)'s
with a coefficient vector \( a^* \) which is a function of the same quantities that occur in \( f_{a^*} \). Thus, since \( x_{i1} \)'s are each \( \text{N}(0,1) \), therefore, the conditional distribution of \( f_{a^*} \), given \( a^* \), that is, given \( z_{ij} \)'s and \( y_{ij} \)'s, is normal with zero mean and variance \( \sigma^2 = \sum_{i=1}^{p} a_i^2 \). Therefore, aside from a constant and positive factor of proportionality, we can rewrite (13.1.11) as

\[
(13.1.12) \quad \frac{dI_1^*}{d(\sqrt{\gamma_1})} = \int \exp \left\{ -\frac{1}{2} \left( \frac{f}{l_{a^*} + \gamma_1} + \frac{f}{2a^*} \right)^2 \right\} \exp \left\{ -\frac{1}{2} \left( \frac{f}{l_{a^*} - \gamma_1} - \frac{f}{2a^*} \right)^2 \right\} \prod_i \prod_j dz_{ij} dy_{ij}.
\]

Integrating out over \( f_{a^*} \) it is easy to check that the right side reduces to

\[
(13.1.13) \quad \int \exp \left\{ -\frac{1}{2(1+\sigma^2)} \left( \frac{\gamma_1}{\sigma^*} + \frac{f}{2a^*} \right)^2 \right\} \exp \left\{ -\frac{1}{2(1+\sigma^2)} \left( \frac{\gamma_1 - f}{2a^*} \right)^2 \right\} \prod_i \prod_j dz_{ij} dy_{ij}.
\]

Remembering that \( f_{a^*} \) is, a.e., positive, it is now easy to check that, according as \( \gamma_1 \) is positive or negative, we have, a.e., \( (\frac{\gamma_1}{\sigma^*} + \frac{f}{2a^*})^2 > (\frac{\gamma_1 - f}{2a^*})^2 \), that is, a.e.,

\[
(13.1.14) \quad \exp \left\{ -\frac{1}{2(1+\sigma^2)} \left( \frac{\gamma_1}{\sigma^*} + \frac{f}{2a^*} \right)^2 \right\} < \exp \left\{ -\frac{1}{2(1+\sigma^2)} \left( \frac{\gamma_1 - f}{2a^*} \right)^2 \right\}.
\]

Thus, the integral (13.1.13) is negative or positive according as \( \gamma_1 \) is positive or negative, which proves that \( I_1^* \) is a monotonically decreasing function of each \( \gamma_1 \) separately. Thus \( I_1 \) stays less than a monotonically decreasing function of each \( \gamma_1 \) separately, so that the power of the test stays greater than a monotonically increasing function of each \( \gamma_1 \) separately and is unbiased at least against all alternatives \( \gamma_1 \)'s for which \( I_1^* \leq 1-a \).
(13.2) Another lower bound function for the power of the test of independence between two sets of variates. We start from the canonical form (A.7.3.5), denote by $c_p$ the largest characteristic root of $(Y_1Y_1)^{-1}(Y_1Y_2)(Y_2Y_1)^{-1}$, by $H_0$ the $H(Y_i = 0)$ ($i = 1, 2, \ldots, p$) and by $H$ its complement and then observe that, for a given $c_0 (< 1)$, $P(c_p \leq c_0 \mid H)$ is a function of $\gamma_1, \ldots, \gamma_p = \psi_2(\gamma_1, \ldots, \gamma_p)$, say. We shall prove that

(13.2.1) $P(c_p \leq c_0 \mid H)$, i.e., $\psi_2(\gamma_1, \ldots, \gamma_p)$ stays less than a monotonically decreasing function of each $\sqrt{\gamma_i}$ separately (notice that each $\gamma_i \geq 0$ and $< 1$), which is different from the decreasing function on the right side of (12.4.1).

Proof. It will suffice to prove this monotonicity with respect to any one parameter, say $\gamma_1$. Toward this end we proceed as follows. We first rewrite the canonical p.d.f. based on (A.7.3.5) in the expanded form

(13.2.2) $\text{Const. } \exp \left( -\frac{1}{2} \sum_{i=1}^{p} \frac{1}{\gamma_i} \sum_{k=1}^{n} (x_{ik}^2 + y_{ik}^2 - 2\sqrt{\gamma_i} x_{ik} y_{ik}) + \sum_{k=p+1}^{q} \sum_{i=1}^{n} y_{ik}^2 \right)$,

by putting $Y_1 = X$ and $Y_2 = Y$ (the elements of the latter matrices being $x_{ik}$ and $y_{ik}$) and letting all new variates also vary from $-\infty$ to $\infty$. We next use (A.3.8) to find a triangular $\tilde{U}(q \times q)$ such that

(13.2.3) $YY' = \tilde{U} \tilde{U}'$ and $u_{ij} = 0$ if $j > i$ ($j = 2, 3, \ldots, q$).

We recall from (A.3.8) that given $Y$, the elements of $\tilde{U}$ can be uniquely determined by adopting a convention, say, that $u_{ii} > 0$ ($i = 1, 2, \ldots, q$), provided that $Y$ is of rank $q$, as it will, almost everywhere, be. Now (see (A.3.15)) it is possible to choose an orthogonal transformation: $X(p \times n)\Gamma(n \times n) = X^\times$ and $Y = Y$ (notice that although $\Gamma$ might involve $Y$, yet the Jacobian is $1$), such that

(13.2.4) $XX' = X^\times X'^\times$, $YY' = YY'$ and $XY' = X^\times \tilde{U}'$.

This is easily seen if we put $Y(q \times n) = \tilde{U}(q \times q)L(q \times n)$ (where $LL' = I(q)$, complete $L(q \times n)$ into an orthogonal matrix $egin{pmatrix} L \\ M \end{pmatrix}$ $q = I'$, say and then put $X = X^\times \Gamma'$,
so that \( XX' = X^* X^* ' \) and \( YY' = X^* \left( \begin{array}{l} L \\ M \end{array} \right) \). The p.d.f. of \( X^*, Y \) can now be conveniently written as

\[
(13.2.5) \quad \text{Const.} \exp \left( - \frac{1}{2} \sum_{k=1}^{n} \sum_{i=1}^{p} \left( x_{ik}^* - \rho_{ik} u_{ik} \right)^2 / (l - \rho_{ik}^2) - \frac{1}{2} \sum_{i>j=1}^{q} u_{ij}^2 \right),
\]

where \( \rho_{ik} = 0 \) if \( k > i \) and \( = \sqrt{y_{ik}} \) if \( k \leq i \). We now put

\[
(13.2.6) \quad \left( \begin{array}{l} x_{ik}^* - \rho_{ik} u_{ik} \\ (l - \rho_{ik}^2)^{1/2} 
\end{array} \right) = z_{ik} \quad \text{(being the elements of a matrix } z(p \times n) \text{)}
\]

and \( \beta_{ik} = \rho_{ik} / (l - \rho_{ik}^2)^{1/2} \), and obtain the p.d.f. of \( z_{ik} \) and \( y_{ik} \) (which vary from \(-\infty\) to \(\infty\)) in the form

\[
(13.2.7) \quad \text{Const.} \exp \left( - \frac{1}{2} \left( \sum_{k=1}^{n} \sum_{i=1}^{p} z_{ik}^2 + \sum_{i>j=1}^{q} u_{ij}^2 \right) \right),
\]

where \( u_{ij} \)'s are given in terms of \( y_{ik} \)'s by \((13.2.3)\). Notice from \((13.2.2), (13.2.4)\) and \((13.2.6)\) that finally

\[
(13.2.8) \quad \left( \begin{array}{l} Y_{11} \end{array} \right)_{11} = \sum_{k=1}^{n} (z_{ik} + \beta_{ik} u_{ik}) (z_{ik} + \beta_{ik} u_{ik}) (l - \rho_{ik}^2)^{1/2} (l - \rho_{ik}^2)^{1/2},
\]

\[
(13.2.9) \quad (Y_{11}^2)_{jj} = \sum_{k=1}^{n} (z_{ik} + \beta_{ik} u_{ik}) (l - \rho_{ik}^2)^{1/2} u_{jk}^2 \quad \text{and } (Y_{22}^{'1})_{jj} = \sum_{k=1}^{n} u_{jk}^2 \quad \text{where } (i,i') = 1,2,\ldots,p; \quad (j,j') = 1,2,\ldots,q). \quad \text{Next we recall from (A.4.3) that the largest characteristic root } \lambda_p \text{ of } (Y_{11} Y_{11}^{-1} (Y_{12} Y_{22}^{'1})^{-1} (Y_{22} Y_{22}^{-1}) \text{ can be written as}
\]

\[
\text{Sup}_{a,b} \left( a' Y_{12} b / (a' Y_{11} a) (b' Y_{22} b) \right) \text{ and the domain } \lambda_p = \lambda_0 \text{ as}
\]

\[
(13.2.10) \quad \left( a' Y_{12} b / (a' Y_{11} a) (b' Y_{22} b) \right) \leq \lambda_0 \text{ or alternatively as}
\]

\[
\lambda_0 \sum_{a,b} \left( a' Y_{12} b / (a' Y_{11} a) (b' Y_{22} b) \right) \leq \lambda_0 \text{ or alternatively as}
\]

\[
(13.2.11) \quad \left( a' Y_{12} b / (a' Y_{11} a) (b' Y_{22} b) \right) \leq \lambda_0 \text{ (say)} \quad \text{or, using (13.2.8), as}
\]

\[
\text{expression on the left of } \lambda_p \text{ in (13.2.11).}
Now, by using certain standard inequalities, and taking the intersection over \( b \), it is easy to check that (13.2.11) reduces to

\[
(13.2.12) \quad \bigcap_a \left\{ \sum_{k=1}^{q} \sum_{i=1}^{p} a_i^k (z_{ik} + \beta_{ik} u_{ik})^2 \leq c_i^k \sum_{k=q+1}^{n} \sum_{i=1}^{p} a_i z_{ik}^2 \right\},
\]

in which, without any loss of generality, we can take \( a_i = 1 \). The problem now is one of integrating out (13.2.7) over (13.2.12), the \( u_{ij} \)'s being given by (13.2.3). To carry out the integration of (13.2.7) over (13.2.12) we proceed exactly as in the previous case, integrate out over \( z_{ii} \) and then check that, aside from a constant and positive factor of proportionality, the total integral which we call \( I_2 \) is given by \( \bigcap \) see (13.1.7).

\[
(13.2.13) \quad I_2 = \int \int_{\bigcap} \exp\left(-\frac{1}{2} \sum_{i=1}^{p} z_{ii}^2 \right) dz_{ii} \exp\left(-\frac{1}{2} \sum_{i=1}^{p} (z_{ik}^2 + \beta_{ik} u_{ik}) \right) dz_{ik} du_{ik}.
\]

In (13.2.13), \( i \) \( z_{ii} \) is omitted from \( z_{ik} \)'s, \( ii \) \( u_{ik} \)'s (for \( i \neq k \)) and \( z_{ik} \)'s vary from -oo to oo and \( u_{ii} \)'s from 0 to oo, all subject to \( \bigcap \) and \( \bigcap \) staying real and \( \bigcap \) \( \bigcap \) and \( \bigcap \) are given by

\[
(13.2.14) \quad \bigcap_{1a} = \bigcap_{-\sqrt{\gamma_1/(1-\gamma_1)}} u_{ii} - f_{1a} \quad \bigcap_{2a} \quad \bigcap_{-\sqrt{\gamma_1/(1-\gamma_1)}} u_{ii} - f_{1a} - f_{2a},
\]

in which \( f_{1a} \) and \( f_{2a} \) are defined by

\[
(13.2.15) \quad f_{1a} = \sum_{i=2}^{p} a_i (z_{il} + \beta_{il} u_{il}) \quad \text{and} \quad f_{2a} = \sum_{i=k+1}^{n} c_i \sum_{i=1}^{p} a_i z_{ik}^2 - \sum_{k=q+1}^{n} \sum_{i=1}^{p} (z_{ik} + \beta_{ik} u_{ik})^2 \bigcap_{1a}^{1/2}.
\]

Arguing now exactly in the same manner as in subsection (13.1) we can establish that \( I_2 \) stays less than a monotonically decreasing function \( I_2^* \) and thus the power of the test stays greater than a monotonically increasing function of \( \sqrt{\gamma_1/(1-\gamma_1)} \), that is, of \( \gamma_1^{1/2} \), that is, of any \( \gamma_1^{1/2} \), from considerations of symmetry. Also the test is unbiased against at least all alternatives \( \gamma_1 \)'s for which \( I_2^* \leq 1-a \).

There are reasons to believe that the lower bounds indicated in sections 13.1
and 13.2 are closer than the lower bounds for the corresponding problems, indicated in sections 12.3 and 12.4.

13.3. Test of independence between two sets of variates under the regression model of (5.25)-(5.28). It will be observed from section (8.7) that the distribution of the roots (and therefore that of the largest root) in this case can be identified with that of case (i) of (8.6) when \( p \leq q \) and with that of case (ii) of (8.6) when \( p > q \), in both cases, by putting \( n_1 = q \) and \( n_2 = n \). It will be also observed that this identification holds for the distributions on both the null and the non-null hypothesis. We have, therefore, exactly the same kind of monotonicity property in this situation as in the case (12.2); and no separate proof need, therefore, be given for this case.

13.4. On the monotonic character of the power function of a modified test for the equality of two dispersion matrices against a special class of alternatives\(^{137}\). We take over from (7.4.5) the acceptance region for the hypothesis \( H_0: \Sigma_1 = \Sigma_2 \) and re-write it as

\[
(13.4.1) \quad c_{1\alpha}(p, n_1, n_2) \leq \text{all } c_i's \leq c_{2\alpha}(p, n_1, n_2),
\]

where \( c_{1\alpha} \) and \( c_{2\alpha} \) are so chosen as to satisfy

\[
(13.4.2) \quad P(c_{1\alpha} \leq \text{all } c_i's \leq c_{2\alpha} \mid \Sigma_1 = \Sigma_2) = 1-\alpha \quad \text{and}
\]

\[
(13.4.3) \quad \left[ \frac{\partial P(c_{1\alpha} \leq \text{all } c_i's \leq c_{2\alpha} \mid \Sigma_1 = \neq \Sigma_2)}{\partial \gamma_i} \right]_{\gamma_1 = \gamma_2 = \ldots = \gamma_p = 1} = 0 \quad (i = 1, 2, \ldots, p), \quad \text{or}
\]

\[
\left[ \frac{\partial P(c_{1\alpha}, c_{2\alpha}, \gamma_1, \ldots, \gamma_p)}{\partial \gamma_i} \right]_{\gamma_1 = \gamma_2 = \ldots = \gamma_p = 1} = 0 \quad (i = 1, 2, \ldots, p),
\]

remembering that if \( \Sigma_1 \neq \Sigma_2 \), the probability \( P \) is, aside from the degrees of freedom \( n_1 \) and \( n_2 \) and the limits \( c_{1\alpha} \) and \( c_{2\alpha} \), purely a function of \( \gamma_i's \), the characteristic roots of \( \Sigma_1 \Sigma_2^{-1} \). It will be shown here that if \( \gamma_1 = \ldots = \gamma_p = \gamma \) (say) (which means that \( \Sigma_1 \Sigma_2^{-1} \) itself is equal to \( \gamma \), i.e., \( \Sigma_1 = \gamma \Sigma_2 \)), then \( P \) monotonically decreases, i.e.,
the power of the test monotonically increases as this common \( \gamma \) tends away from 1, which is the value of \( \gamma \) on the null hypothesis \( \Sigma_1 = \Sigma_2 \).

**Proof.** We start from the canonical probability

\[
\int \text{const} \prod_{i=1}^{n_1} \frac{1}{\gamma_i} \exp \left( -\frac{1}{2} \text{tr} (D_{i/\gamma_i} X_1 X_1^T + X_2 X_2^T) \right) \ dX_1 \ dX_2 ,
\]

where the constant factor is a pure constant not involving the parameters, \( D_{i/\gamma_i} \) stands for a diagonal matrix whose diagonal elements are \( 1/\gamma_i, \ldots, 1/\gamma_p \), \( X_1 \) and \( X_2 \) are \( p \times n_1 \) and \( p \times n_2 \) (\( p < n_1, n_2 \)) and where the \( c_i \)'s of (13.4.1) are the roots of \( X_1 X_1^T(X_2 X_2^T)^{-1} \). We first show that the \( p \) equations under (13.4.3) are really equivalent to one equation. To prove this we note that aside from the constant factor,

(13.4.4) \[ P = \int \prod_{i=1}^{n_1} \frac{1}{(1/\gamma_i)^2} \exp \left( -\frac{1}{2} \text{tr} (D_{i/\gamma_i} X_1 X_1^T + X_2 X_2^T) \right) \]

\[ c_{1a} \leq \text{all} \ c \sqrt{X_1 X_1^T(X_2 X_2^T)^{-1}} \leq c_{2a} \]

\[ x \ dX_1 \ dX_2 . \]

Hence

(13.4.5) \[ \frac{\partial P}{\partial (\frac{1}{\gamma_i})} = \int \prod_{j=1}^{n_1} \frac{1}{(1/\gamma_j)^2} \exp \left( \frac{1}{2} \gamma_j - \frac{1}{2} \text{tr} (D_{j/\gamma_j} X_1 X_1^T) \right) \]

\[ c_{1a} \leq \text{all} \ c \sqrt{X_1 X_1^T(X_2 X_2^T)^{-1}} \leq c_{2a} \]

\[ x \ \exp \left( -\frac{1}{2} \text{tr} (D_{i/\gamma_i} X_1 X_1^T + X_2 X_2^T) \right) \ dX_1 \ dX_2 . \]

Now using the transformation

(13.4.6) \[ X_1 (p \times n_1) = U(p \times p) D(p \times n_1) \]

and

\[ X_2 (p \times n_2) = U(p \times p) L(p \times n_2) , \]

where \( U \) is non-singular and \( L_{11} = L_{21} = I(p) \), and integrating out over \( L_{11} \) and \( L_{21} \), we observe that aside from a positive and constant factor of proportionality (13.4.5) reduces to
\[
\frac{\partial P}{\partial (1/\gamma_1^i)} = \int \prod_{j=1}^{p} \left( \frac{1}{\gamma_j} \right)^{n_1/2} \gamma_1^{\frac{n_1}{2}} \left( \frac{d_1}{\gamma_j} \right)^{n_2-p} \exp \left\{ -\frac{1}{2} \left( U \cdot U' \right) \right\} \frac{dU}{\mathcal{D}} \frac{1}{\gamma_1^i} \gamma_1^i \gamma_j \gamma_{j'} \gamma_{j''} \gamma_{j'''} \gamma_{j''''} \gamma_{j'''''} \gamma_{j''''''} \gamma_{j'''''''} \gamma_{j''''''''} \gamma_{j'''''''''} \gamma_{j'''''''''}. \]
where \( \mathcal{D} \) stands for a diagonal matrix with diagonal elements \( c_1, \ldots, c_p \) and the domain \( \mathcal{D} \) is

\[
\gamma_1 \leq \ldots \gamma_p \leq \infty \quad \text{all } c_i \text{'s } \leq c_{i+1} \quad \text{all } U_{ij} \text{'s } \leq \infty. \quad \text{We have thus}
\]

\[
\int \prod_{j=1}^{p} \left( \frac{1}{\gamma_j} \right)^{n_1/2} \gamma_1^{\frac{n_1}{2}} \left( \frac{d_1}{\gamma_j} \right)^{n_2-p} \exp \left\{ -\frac{1}{2} \left( U \cdot U' \right) \right\} \frac{dU}{\mathcal{D}} \frac{1}{\gamma_1^i} \gamma_1^i \gamma_j \gamma_{j'} \gamma_{j''} \gamma_{j'''} \gamma_{j''''} \gamma_{j''''} \gamma_{j'''''} \gamma_{j'''''}. \]

Having regard to the definition of the domain given by \((13.4.8)\) and the structure of the integral on the right side of \((13.4.9)\) it is easy to check that this integral is invariant under a change of the subscript \( i \), so that the expression on the left side of \((13.4.9)\) is the same for \( i = 1, 2, \ldots, p \) and hence the \( p \) equations \((13.4.3)\) are equivalent to really one equation. Now adding \( p \) formally different looking integrals like the right side of \((13.4.9)\) over \( i = 1, 2, \ldots, p \) and cancelling a factor we note that \((13.4.3)\) is really equivalent to

\[
\int \prod_{j=1}^{p} \left( \frac{1}{\gamma_j} \right)^{n_1/2} \gamma_1^{\frac{n_1}{2}} \left( \frac{d_1}{\gamma_j} \right)^{n_2-p} \exp \left\{ -\frac{1}{2} \left( U \cdot U' \right) \right\} \frac{dU}{\mathcal{D}} \frac{1}{\gamma_1^i} \gamma_1^i \gamma_j \gamma_{j'} \gamma_{j''} \gamma_{j'''} \gamma_{j''''} \gamma_{j''''} \gamma_{j'''''} \gamma_{j'''''}. \]

It is easy to check that the left side of \((13.4.10)\) is the same as if we had put

all \( \gamma_j \text{'s } = \gamma \) in \((13.4.4)\) and then differentiated the integral with respect to \( \gamma \) and then put \( \gamma = 1 \). As will be presently seen this will enable us to rewrite \((13.4.10)\) in a simpler form (which can also be derived in a straightforward though rather
lengthier manner). At this point, remembering the definition of $D$ by (13.4.8), we merely observe that (13.4.10) gives one relation between $c_{1\alpha}$ and $c_{2\alpha}$ which we call the condition of local unbiasedness and then (13.4.2) added to this determines $c_{1\alpha}$ and $c_{2\alpha}$ completely.

Going back now to the problem of proving the monotonicity of the integral on the right of (13.4.4) under the special assumption that $\gamma_1 = \ldots = \gamma_p = \gamma$ (say), we proceed as follows. Putting $D_{1/\gamma} X_1 = Y_1$ and $X_2 = Y_2$ we note that (13.4.4) reduces to

$$\int \exp \left[ -\frac{1}{2} \text{tr} (Y_1 Y_1^t + Y_2 Y_2^t) \right] dY_1 dY_2.$$  

Now putting all $\gamma_i = \gamma$ it is easy to check that this reduces to

$$\int \exp \left[ -\frac{1}{2} \text{tr} (Y_1 Y_1^t + Y_2 Y_2^t) \right] dY_1 dY_2.$$  

We are thus back in the problem of the distribution of the characteristic roots on the null hypothesis and we have, therefore, aside from a constant and positive factor of proportionality not involving the $\gamma$,

$$\int \int \cdots \int f(c_1, \ldots, c_p) \prod_{i=1}^p \frac{dc_i}{(1+c_i)^{n_i^2}}$$  

where

$$f(c_1, \ldots, c_p) = \prod_{i=1}^p \frac{c_i^{n_i^2}}{(1+c_i)^{n_i^2}} = \prod_{i=1}^m c_i^{m_i (1+c_i)^{-n_i}}$$  

It is now easy to check that

\[(13.4.15) \quad \frac{\partial p}{\partial (1/\gamma)} = \sum c_{2a} \left( \begin{array}{c} c_{p-1} \\ 2 \end{array} \right) \cdots \left( \begin{array}{c} c_2 \\ 2a \end{array} \right) \frac{1}{1/\gamma} \frac{dc_i}{c_{2a}/\gamma} \quad \text{for } i=1 \]

\[= c_{2a}(c_{2a}/\gamma)^m \left( \begin{array}{c} c_{2a}/\gamma \\ \cdots \\ c_{2a}/\gamma \end{array} \right) \left( \begin{array}{c} 1+c_{-1} \\ i=1 \end{array} \right) \frac{dc_i}{c_{2a}/\gamma} \quad \text{for } i=1 \]

\[= k_1(\gamma)I_1(\gamma) - k_2(\gamma)I_2(\gamma), \text{ say.} \]

The condition of local unbiasedness is that

\[(13.4.16) \quad k_1(1)I_1(1) = k_2(1)I_2(1). \]

We shall now show that subject to \((13.4.16)\) the last expression on the right of \((13.4.15)\) > 0 if \(\gamma > 1\) and < 0 if \(\gamma < 1\). The proof will thus be complete if we can show that according as \(\gamma > 1\) or < 1 we have

\[(13.4.17) \quad \frac{(1+c_{2a}/\gamma)^n}{(1+c_{2a}/\gamma)^n} \frac{I_1(\gamma)}{I_2(\gamma)} > (c_{2a}/c_{2a})^{m+1}, \text{ i.e., } > (1+c_{2a})^{m+1} I_1(1) \frac{1+c_{2a}}{1+c_{2a}} I_2(1), \text{ or } < (1+c_{2a})^n I_1(1) \frac{1+c_{2a}}{1+c_{2a}} I_2(1). \]

Now according as \(\gamma > 1\) or < 1 we have

\[(13.4.18) \quad (1 + \frac{c_{2a}}{\gamma})/(1 + \frac{c_{2a}}{\gamma}) > (1+c_{2a})/(1+c_{2a}). \]

Thus \((13.4.17)\) will be proved if we show that \(I_1(\gamma)\) is an increasing function of \(\gamma\) and \(I_2(\gamma)\) a decreasing function of \(\gamma\).
Now

\[
\frac{\partial I_1(\gamma)}{\partial (1/\gamma)} = -\frac{c_{1a}(c_{1a}/\gamma)^m}{(1 + c_{1a}/\gamma)^n} I \quad \text{and} \quad \frac{\partial I_2(\gamma)}{\partial (1/\gamma)} = \frac{c_{2a}(c_{2a}/\gamma)^m}{(1 + c_{2a}/\gamma)^n} I,
\]

where \(I\) stands for the positive quantity

\[
\frac{c_{2a}/\gamma}{c_{1a}/\gamma} \cdot \frac{c_{p-1}}{c_{1a}/\gamma} \cdot \ldots \cdot \frac{c_{2a}}{c_{1a}/\gamma} \cdot \frac{c_{p-1}}{c_{1a}/\gamma} = \frac{c_{2a}}{c_{1a}} \cdot \frac{c_{p-1}}{c_{1a}} \cdot \ldots \cdot \frac{c_{2a}}{c_{1a}} \cdot \frac{c_{p-1}}{c_{1a}}.
\]

It is thus easy to see that \(\frac{\partial I_1(\gamma)}{\partial (1/\gamma)} < 0\) and \(\frac{\partial I_2(\gamma)}{\partial (1/\gamma)} > 0\), so that \(F\) of (13.4.13) monotonically decreases or the power of the test (7.4.5) monotonically increases as \(\gamma\) tends away from 1.

13.5. On the monotonic character of the power function of a modified test of the hypothesis that a population dispersion matrix has a given (matrix) value against a special class of alternatives. We take over from (7.4.4) the acceptance region for the hypothesis \(H_0: \Sigma = \Sigma_0\) and rewrite it as

\[
(13.5.1) \quad c_{1a}(p,n) \leq c_1's \leq c_{2a}(p,n),
\]

where \(c_{1a}\) and \(c_{2a}\) are so chosen as to satisfy

\[
(13.5.2) \quad P(c_{1a} \leq c_1's \leq c_{2a} | \Sigma = \Sigma_0) = 1-\alpha
\]

\[
\int_{\gamma_1 = \ldots = \gamma_p = 1} \frac{\partial P(c_{1a} \leq c_1's \leq c_{2a} | \Sigma = \Sigma_0)}{\partial \gamma_i \gamma_1 = \ldots = \gamma_p = 1} = 0 \quad (i = 1, 2, \ldots, p), \text{ or}
\]

\[
\int_{\gamma_1 = \ldots = \gamma_p = 1} \frac{\partial P(c_{1a}'c_{2a}'\gamma_1, \ldots, \gamma_p)}{\partial \gamma_i \gamma_1 = \ldots = \gamma_p = 1} = 0 \quad (i = 1, 2, \ldots, p).
\]

Here the \(c_1's\) are the characteristic roots of \(\frac{1}{n}(XX')\Sigma_0^{-1}\), \(\gamma_1's\) are the characteristic roots of \(\Sigma_0^{-1}\) and \(X(p \times n)(p \leq n)\) is the reduced observation matrix. Exactly along the same lines as in the previous case it can be proved that (i) the \(p\) equations
(13.5.3) are really equivalent to one equation and that (ii) if $\gamma_1 = \gamma_2 = \ldots = \gamma_p = \gamma$ (say), in other words if $\Sigma_0^{-1} = \gamma$, i.e., $\Sigma = \gamma \Sigma_0$, then the $P$ of (13.5.3) monotonically decreases, i.e., the power of the test monotonically increases as $\gamma$ tends away from 1, which is the value on the null hypothesis.

13.6. It can be shown by very lengthy and tedious calculations that the two tests considered in 13.1 and 13.2 for multivariate analysis of variance and for independence between two sets of variates as also the modified tests considered in 13.4 and 13.5 for one and two dispersion matrices have each of them the monotonicity property, and not just the near monotonocity property which has been proved in chapters 12 and 13. But this lengthy proof is not being offered, in the hope that a much simpler and more elegant proof may be forthcoming in the near future.

14. Least squares and univariate analysis of variance and covariance with observations on extensions.

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

(14.1.1) $E(x) = A(n \times p)\xi(p \times 1)$,

where $p \leq n$ and $\xi(p \times 1)$ 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 $c'(1 \times p)$ (subject to certain restrictions to be brought out in (14.2)) and given $x$, it is required to obtain for $c'\xi$ a linear estimate $b'(1 \times n)x(n \times 1)$ such that (i) $E(b'x) = c'\xi$ (for all $\xi$) and (ii) $v(b'x)$ is to be a minimum. $c'\xi$ will be said to be linearly estimable (or sometimes just "estimable") if and only if (i) is satisfied.

**Problem II.** Given $c'$ and $x$ as above, it is required to obtain $\xi$ so that $(x' - c'A')(x - A\xi)$ is a minimum. It will then be incidentally verified that $b'x$ of
Problem I = c₁ of Problem II.

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

\[(14.1.2) \quad \Omega(q \times p) \xi(p \times l) = \Omega(q \times l),\]

where $r \leq p$ ($r$ being the rank of the $A$-matrix of (14.1.1)) and $C$ is a given matrix of rank $s \leq \min(r, q)$.

**(14.2)** Solution of Problem I. Assume that $A'(p \times n)$ is such that $A_1(r \times n)$ can be taken as a basis and let $A'(p \times n)$ of (14.1.1) be factorized into:

\[(14.2.1) \quad \begin{pmatrix} r \\ p-r \end{pmatrix} \begin{pmatrix} A' \\ A'_2 \end{pmatrix} = \begin{pmatrix} T' \\ T_2 \end{pmatrix} L(r \times n),\]

such that $LL' = I(r)$, and let $L_1((n-r) \times n)$ be an arbitrary completion of $L$ in the sense of (A.1.15), so that

\[(14.2.2) \quad L_2(n \times n) = \begin{pmatrix} L \\ L_1 \end{pmatrix}^{r}_{n-r} \quad \text{is } L\text{.}\]

Notice that $L_1$ is not unique. Also observe that

\[(14.2.3) \quad I(n) = \begin{pmatrix} L \\ L_1 \end{pmatrix} (L' \quad L_1) = \begin{pmatrix} LL' \\ L_1L' \quad L_1L_1 \end{pmatrix} = (L' \quad L_1) \begin{pmatrix} L \\ L_1 \end{pmatrix} = L'L + L_1L_1.\]

Furthermore, with an $A$ having the structure (14.2.1), let (14.1.1) be rewritten as

\[(14.2.4) \quad E(x) = n(A_1 \quad A_2) \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}^{r}_{p-r},\]

and let $c'\xi$ be rewritten as

\[(14.2.5) \quad l(c_1' \quad c_2') \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}^{r}_{p-r}.\]

Now condition (i) (of unbiasedness) of Problem I of (14.1) becomes
\[(14.2.6) \quad \begin{pmatrix} c_1' & c_2' \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = b'E(x) = \begin{pmatrix} b_1' & A_2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = b'(A_1 \xi_1 + A_2 \xi_2), \]

and, since this is to be true of all $\xi_1$ and $\xi_2$, we should have

\[(14.2.7) \quad b'A_1 = c_1 \quad \text{and} \quad b'A_2 = c_2', \]

which imposes a number of restrictions ($\leq p$) on $b'(1 \times n)$ but by no means fully determines $b'$ (which has to be determined). Substituting in (14.2.7) for $A_1$ and $A_2$ from (14.2.1), we have

\[(14.2.8) \quad b'L'_1 = c_1' \quad \text{or} \quad b'L' = c_1'(T_1')^{-1}, \quad \text{and} \quad b'L'_2 = c_2'. \]

Now to minimize $V(b'x)$ subject to (14.2.8) we proceed as follows:

\[(14.2.9) \quad V(b'x) = s^2 b'b^{-1} (\text{since } x \text{ is an uncorrelated set with a common variance } s^2) \]

\[= s^2 b'(L'_1 \begin{pmatrix} L \\ L_1 \end{pmatrix}) b \quad \text{(using (14.2.2))} = s^2 b'L'_1 L b + b'L_1 L_1 b \quad \text{(using (14.2.8)).} \]

The minimum $V(b'x)$ is thus reached when

\[(14.2.10) \quad b'L'_1 = 0, \]

so that, combining (14.2.2), (14.2.8) and (14.2.10), we have

\[(14.2.11) \quad b' = c_1'(T_1')^{-1} L, \quad \text{and hence} \]

\[(14.2.12) \quad b'x = c_1'(T_1')^{-1} L x = c_1'(T_1')^{-1} (T_1')^{-1} A_1 x \quad \text{(using (14.2.1))} = c_1'(T_1 T_1')^{-1} A_1 x \]

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

Restriction on $c'$. Now, substituting in (14.2.7) for $b'$ from (14.2.11), we have

\[(14.2.13) \quad c'_2 = b'A_2 = c_1'(T_1')^{-1} A_2 = c_1'(T_1')^{-1} (T_1')^{-1} A_1 A_2 \quad \text{(using (14.2.1))} \]

\[= c_1'(A_1 A_1')^{-1} A_1 A_2 \]

We have thus that, in order that $c'\xi$, i.e., $(c_1' \quad c_2') \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$ may be "estimable" (in the sense indicated), $c_2'$ must be related to $c_1'$ by (14.2.13), which can be expressed
in another form that is more suggestive. From (14.2.11) we have

\[
(14.2.11) \quad A_2 = L_1 T_2 = A_1 \left( \tilde{T}_1 \right)^{-1} T_1 \quad \text{or} \quad A_2 = T_2 \left( \tilde{T}_1 \right)^{-1} A_1,
\]

which on substitution into (14.2.13) yields

\[
(14.2.15) \quad c_2^1 = c_1^1 (A_1 A_1)^{-1} A_1 \left( \tilde{T}_1 \right)^{-1} T_2 = c_1^1 \left( \tilde{T}_1 \right)^{-1} T_2.
\]

Thus \( c_2 \) is related to \( c_1 \) by the same post factor by which \( A_2 \) is related to \( A_1 \).

Invariance of the linear estimate (14.2.13) under choice of \( A_1 \). If, instead of \( A_1 \) and \( c_1 \), we choose another set of independent row vectors, say \( A_3 \) and the \( c_3 \) to match it, then in place of the right hand side of (14.2.12) we should have the linear estimate given by replacing the subscript 1 by 3. But remembering that

\[
(14.2.16) \quad A_3 = T_3 (r \times n),
\]

where \( T_3 \) is obtained by picking out from the right hand side of (14.2.1) the rows corresponding to \( A_3 \) and is necessarily non-singular (since \( A_3 \) is of rank \( r \)), and using (14.2.15) and (14.2.16), we have

\[
(14.2.17) \quad c_2^1 (A_1 A_3)^{-1} A_1 X = c_1^1 \left( \tilde{T}_1 \right)^{-1} T_3 (T_3) T_3 \left( \tilde{T}_1 \right)^{-1} A_1 (A_1 A_3)^{-1} A_1 X
\]

\[
= c_1^1 \left( \tilde{T}_1 \right)^{-1} T_3 (T_3) T_3 \left( \tilde{T}_1 \right)^{-1} A_1 X = c_1^1 (A_1 A_3)^{-1} A_1 X,
\]

which proves the invariance.

Variance of the "unbiased minimum variance" estimate. From (14.2.9), (14.2.11)

this variance is given by

\[
(14.2.18) \quad V(b'X) = \sigma^2 b' b = \sigma^2 c_1^1 (\tilde{T}_1)^{-1} LL' (\tilde{T}_1)^{-1} c_1 = \sigma^2 c_1^1 (\tilde{T}_1)^{-1} c_1 = \sigma^2 c_1 (A_1 A_1)^{-1} c_1,
\]

which again by the method of the previous paragraph can be shown to be invariant under choice of \( A_1 \).

(14.3) Solution of problem II or the "Least squares solution".

\[
(14.3.1) \quad (X' X' A') (X - A X) = (X' \tilde{X} A') (L L') (X - A X).
\]

\[
= \left( X' \tilde{X} \right) \left( \left( T_1 L L' \left( L_1 \right)^{-1} \right)^{-1} \right) \left( T_2 \right) \left( T_1 \right)^{-1} \left( T_2 \right)^{-1} X.
\]
\[ \sum_{p=1}^{P} \sum_{q=1}^{Q} \left( \frac{r_{pq}}{s_{pq}} \right) \frac{t_{pq}}{u_{pq}} \]

(11.2.3). It is now quite easy to see that given \( x \) and \( A \) the minimum value \((x' - \tilde{A}')(x - \tilde{A})\), under variation of \( \hat{A} \), will be attained if

\[ Lx = \left( T_1' \right) \left( T_2' \right) \hat{A} \]

If we now want the "least squares estimate" \( c'(T) \) of an " estimable linear function" \( \hat{c}(T) \), we have from the above:

\[ c'(T) = c'_1 \hat{e}_1 + c'_2 \hat{e}_2 = c'_1 \hat{e}_1 + c'_1 (T')^{-1} \hat{e}_2 \]  
(from \((11.2.16)) \]

\[ = c'_1 (T')^{-1} (T_1' \tilde{e}_1 + T_2' \tilde{e}_2) = c'_1 (T')^{-1} Lx \]  
(from \((11.3.2)) \]

\[ = c'_1 (T')^{-1} (T_1')^{-1} A_1 x \]  
(from \((11.2.1)) \]

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

(11.4) Solution of problem III. It is well known that

(11.4.1) if \( x(n x 1) \) is a set of \( n \) uncorrelated \( N(E(x), \sigma^2) \) (and thus also independent) variates and if \( L(p x n) \) \((p < n)\) is subject to \( LL' = I(p) \), then \( L(p x n)x(n x 1) \) is a set of \( p \) uncorrelated \( N(LE(x), \sigma^2) \) variates.

It is also well known that

(11.4.2) if \( u(p x 1) \) is an uncorrelated \( N(0, \sigma^2) \) set and so is \( v(q x 1) \) and if \( u \) and \( v \) are mutually uncorrelated, then \( v'u/\sigma^2 \) is a \( \chi^2 \) with \( p \) degrees of freedom, \( v'v/\sigma^2 \) is a \( \chi^2 \) with \( q \) degrees of freedom and \( q'u/pv'v \) is an \( F \) with degrees of freedom \( p \) and \( q \).

Going back to the model of Problem III in \((11.4)\) and to \((11.2.1)-(11.2.3)\) we observe that

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

Now from \((11.1.1)\) and \((11.2.1)-(11.2.3)\) we have \( E(x) = L'(\tilde{T}_1'T_1) \tilde{e} \), so that \( L_1E(x) = L_1L'(\tilde{T}_1'T_1) \tilde{e} = 0 \). Thus \( L_1x \) is an uncorrelated \( N(0, \sigma^2) \) set, whence it
follows that we have a \( x^2 \) (with \( n-r \) degrees of freedom) given by:

\[
(14.4.4) \quad \frac{x' L_1 L_1 x}{\sigma^2} \text{ or } \frac{x'(I(n)-L'L)x}{\sigma^2} \text{ or } \frac{\sum x'x - x'A_1 (T_1 T_1')^{-1} A_1 x}{\sigma^2} \text{ or } \frac{\sum x'x - x'A_1 (A_1 A_1')^{-1} A_1 x}{\sigma^2}.
\]

Consider now the hypothesis \( C(q \times p)^{E}(p \times 1) = 0 \), 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

\[
(14.4.5) \quad s \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}^r = 0,
\]

where \( (C_{11} \ C_{12}) \) are a set of \( s \) independent row vectors and \( \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \) a matrix, each row of which is of the nature of \( c_1 \) of (14.2). In this case the hypothesis (14.4.5) will be said to be 'testable'.

It is now easy to see that the hypothesis \( C^E = 0 \) is equivalent to \( C_{11} \xi_1 + C_{12} \xi_2 = 0 \), so that we shall work in terms of this latter. Going back to (14.2.12), (14.2.13) and (14.2.15) we note that

\[
(14.4.6) \quad C_{12} = C_{11} (T_1 T_1')^{-1} T_2^T
\]

and

\[
(14.4.7) \quad 0 = C_{11} \xi_1 + C_{12} \xi_2 = E \sum C_{11} (A_1 A_1')^{-1} A_1 x = E \sum C_{11} (T_1 T_1')^{-1} T_1 x.
\]

Now \( (T_1 T_1')^{-1} T_1 \) is a \( r \times n \) matrix of rank \( r \) and \( C_{11} \) is a \( s \times r \) matrix (\( s \leq r \)) of rank \( s \). Then using (4.1.6) we note that \( C_{11} (T_1 T_1')^{-1} T_1 L \), which is a \( s \times n \) matrix, must be of rank \( s \leq \min(q,r) \) (note that \( r \leq p < n \)). Let

\[
(14.4.8) \quad C_{11} (T_1 T_1')^{-1} T_1 L = \tilde{V}(s \times s)(s \times n), \text{ where } iM^T = I(s), \text{ and } \tilde{V} \text{ of course is non-singular. Then we have}
\]

\[
(14.4.9) \quad E(Mx) = (\tilde{V})^{-1} E \sum C_{11} (T_1 T_1')^{-1} T_1 L x = 0 \quad \text{(from (14.4.7)) and furthermore}
\]

\[
(14.4.10) \quad LM_1 = (\tilde{V})^{-1} C_{11} (T_1 T_1')^{-1} T_1 LM_1 = 0, \text{ so that}
\]

\[
(14.4.11) \quad Mx \text{ is a } s \text{-set of uncorrelated } N(0,\sigma^2), \text{ } L_1 x \text{ (of (14.4.9)) is a } (n-r) \text{-set of uncorrelated } N(0,\sigma^2), \text{ } Mx \text{ and } L_1 x \text{ are mutually uncorrelated, and hence}
\]
(14.1.12) \((n-r)x'X^{-1}x/sx'X^{-1}x\) is an F with degrees of freedom \(s\) and \((n-r)\).

Using (14.1.4), (14.1.8) and of course (14.2.1) and (A.3.11), we can reduce (14.1.4) to

\[
(14.1.13) \quad \frac{(n-r)x'X^{-1}x(A_1A_1)^{-1}C_{11}(A_1A_1)^{-1}C_{11}^{-1}C_{11}(A_1A_1)^{-1}A_1x}{s\sqrt{x'x - x'A_1(A_1A_1)^{-1}A_1x}}
\]

which is an F (with degrees of freedom \(s\) and \((n-r)\)) for testing the hypothesis \(Q_x = 0\) and which is expressed in terms of quantities directly observed or given by the experimental design and the hypothesis to be tested. The form (14.1.13) can be shown to be invariant under the kind of choice indicated in (14.2), in much the same way as there.

14.5. A pseudo-multivariate generalization of the problems considered in (14.1).

Suppose that in (14.1) we assume \(\mathbf{x}(n \times 1)\) denote a correlated set with a p.d. dispersion matrix \(\sigma^2\Sigma(n \times n)\) where \(\sigma^2\) is unknown but \(\Sigma\) is supposed to be known and suppose that (14.1.1) is left unchanged. Also in problem III let us assume that \(\mathbf{x}\) is \(N(E(x),\sigma^2\Sigma)\) but let us leave (14.1.2) unchanged. Then putting

\[
(14.5.1) \quad \Sigma(n \times n) = \tilde{T}(n \times n)\tilde{T}'(n \times n) \quad \text{and} \quad \tilde{T}^{-1}(n \times n)\mathbf{x}(n \times 1) = \mathbf{y}(n \times 1)
\]

it is easy to check that \(\mathbf{y}(n \times 1)\) is a set of uncorrelated variates with a common variance \(\sigma^2\), and also that if \(\mathbf{x}\) is \(N(E(x),\sigma^2\Sigma)\), \(\mathbf{y}\) is \(N(E(y),\sigma^2I(n))\). Also in terms of \(\mathbf{y}\), (14.1.1) reduces to

\[
(14.5.2) \quad E(\mathbf{y}) = \tilde{T}^{-1}A_1\mathbf{x}
\]

It is now easy to check that (14.2.12) reduces to

\[
(14.5.3) \quad c_1(A_1^{-1}A_1)^{-1}A_1\tilde{T}^{-1}y = c_1(A_1^{-1}A_1)^{-1}A_1\tilde{T}^{-1}y = c_1(A_1^{-1}A_1)^{-1}A_1\Sigma^{-1}x.
\]

Also (14.1.13) similarly reduces to

\[
(14.5.4) \quad \frac{(n-r)x'\Sigma^{-1}A_1(A_1^{-1}A_1)^{-1}C_{11}(A_1^{-1}A_1)^{-1}C_{11}^{-1}C_{11}(A_1^{-1}A_1)^{-1}A_1\Sigma^{-1}x}{s\sqrt{x'\Sigma^{-1}x - x'A_1(A_1^{-1}A_1)^{-1}A_1\Sigma^{-1}x}}
\]

14.6. Multivariate generalization. The set-up for multivariate analysis of variance and covariance, i.e., for a test of the general multivariate linear
**Hypothesis** is an easy and direct extension of what has been considered so far in this section. This is given in the next section from (15.2 9) to (15.2.13).

15. Some univariate and bivariate confidence bounds.

15.1. Some general observations. The general theory (to which nothing is added here) of confidence bounds like the general theory of testing of hypotheses and tests of significance (with a part of which the previous sections have been concerned) has been worked out in a series of papers, now classic. This is readily available not only in papers but in standard books as well and need not be explained here. However, except for some comparatively recent work, most of the earlier applications have been concerned with confidence bounds on a single parameter or a single function of the parameters. Simultaneous confidence bounds on several parameters or parametric functions offer nothing new in principle, being already inherent in the general theory and will not, therefore, be discussed here from the point of view of the general theory. In this section several examples from univariate normal populations and one from a bivariate normal population will be discussed (some of the simultaneous and some of them "single") which will prepare the ground for the multivariate examples (all of them simultaneous) to be discussed in section 17. In this section, in every case except one we shall start from a current test of the corresponding hypothesis (having a number of optimum properties in respect of power) and obtain by inversion "single" or "simultaneous" confidence bounds which, therefore, by the general theory, will have similar optimum properties in respect of _shortness_, i.e., the probability of covering wrong values of the parameters or parametric functions.

15.2. Applications to means from normal populations.

(i) For \( N(\mu, \sigma^2) \) we have, in terms of a sample of size \( n \) with sample mean \( \bar{X} \) and sample standard deviation \( s \), the following well known confidence interval for \( \mu \) (with a confidence coefficient \( 1-\alpha \))
(15.2.1) \[ \bar{x} - st_{a/2}(n-1)/\sqrt{n} \leq \xi \leq \bar{x} + st_{a/2}(n-1)/\sqrt{n} , \]

where \( t_{a/2}(n-1) \) is the upper \( a/2 \) point of the ordinary \( t \)-distribution with D.F. \( n-1 \).

(ii) For \( N(\bar{\xi}_h, \sigma^2) \) (\( h = 1,2 \)) we have, in terms of two samples of sizes \( n_h \) with sample means and sample standard deviations \( \bar{x}_h \) and \( s_h \) (\( h = 1,2 \)), the following well known confidence interval for \( \bar{\xi}_1 - \bar{\xi}_2 \) (with a confidence coefficient \( 1-a/2 \))

\[ (\bar{x}_1 - \bar{x}_2) - st_{a/2}(n-2)/\sqrt{n_{12}} \leq \bar{\xi}_1 - \bar{\xi}_2 \leq (\bar{x}_1 - \bar{x}_2) + st_{a/2}(n-2)/\sqrt{n_{12}} , \]

where \( n = n_1 + n_2 \), \( s^2 = \sum (n_i - 1)s_i^2 + (n_2 - 1)s_2^2 )/n - 2 \), \( n_{12} = n_1 n_2 / n \) and \( t_{a/2}(n-2) \) is the upper \( a/2 \) point of the ordinary \( t \)-distribution with D.F. \( n-2 \), i.e., \( n_1 + n_2 - 2 \).

(iii) For confidence bounds relating to \( \bar{\xi}_h \)'s of \( N(\bar{\xi}_h, \sigma^2) \) (\( h = 1,2, \ldots, k \), where \( k > 2 \)) we proceed as follows. Suppose we have random samples of sizes \( n_h \), sample means \( \bar{x}_h \), sample standard deviations \( s_h \) (\( h = 1,2, \ldots, k \)). Put

\[ \sum_{h=1}^{k} n_h \bar{x}_h / n, \quad \sum_{h=1}^{k} n_h s_h^2 / (n-k), \quad \bar{x} = \sum_{h=1}^{k} n_h \bar{x}_h / n, \quad s^2 = \sum_{h=1}^{k} n_h (\bar{x}_h - \bar{x})^2 / k-1, \]

where \( H_0 : \xi_1 = \xi_2 = \ldots = \xi_k \), i.e., \( \xi_h = \bar{x} \) (\( h = 1, \ldots, k \)), we have at a level of significance, say \( \alpha \), the current F-test with a critical region

\[ F = s^2 / s_h^2 > F_{\alpha}(k-1, n-k) , \]

where \( F_{\alpha}(k-1, n-k) \) stands for the upper \( \alpha \) point of the central \( F \)-distribution with D.F. \( k-1 \) and \( n-k \) (we recall the well known result that when \( H_0 \) is true \( s^2 / s_h^2 \) is distributed as the central \( F \)). When \( H_0 \) is not true, it is easy enough to see that \( s^2 / s_h^2 \) is distributed as the central \( F \), where \( s^2 \) is given by

\[ s^2 = \sum_{h=1}^{k} n_h (\bar{x}_h - \bar{x} + \xi_h + \xi)^2 / k-1. \]

Suppose that we now start from a statement with probability \( 1-\alpha \), namely

\[ s^2 / s_h^2 \leq F_{\alpha}(k-1, n-k) \), i.e., \( \sum_{h=1}^{k} n_h (\bar{x}_h - \bar{x} + \xi_h + \xi)^2 / (k-1)s^2 \leq F_{\alpha}(k-1, n-k) . \]
It is easy to check (see (A.2.7)) that the statement (15.2.6) is equivalent to the following statement, \( \text{I}^{22} \),

\[
(15.2.7) \quad s \left( k-1 \right) F_{a}^{2}(k-1, n-k) \frac{1}{2} \leq \sum_{h=1}^{k} a_{h} \frac{1}{2} h_{h}^{2} \left( \bar{x}_{h} - \bar{x} \right)^{2} + \sum_{h=1}^{k} a_{h} \frac{1}{2} h_{h}^{2} \left( \bar{\xi}_{h} - \xi \right)^{2} \leq \sum_{h=1}^{k} a_{h} \frac{1}{2} h_{h}^{2} \left( \bar{x}_{h} - \bar{x} \right)^{2} + \sum_{h=1}^{k} a_{h} \frac{1}{2} h_{h}^{2} \left( \bar{\xi}_{h} - \xi \right)^{2} \]

\[
\text{or } \sum_{h=1}^{k} a_{h} \frac{1}{2} h_{h}^{2} \left( \bar{x}_{h} - \bar{x} \right)^{2} - s \left( k-1 \right) F_{a}^{2}(k-1, n-k) \frac{1}{2} \leq \sum_{h=1}^{k} a_{h} \frac{1}{2} h_{h}^{2} \left( \bar{\xi}_{h} - \xi \right)^{2} \leq \sum_{h=1}^{k} a_{h} \frac{1}{2} h_{h}^{2} \left( \bar{x}_{h} - \bar{x} \right)^{2} + \sum_{h=1}^{k} a_{h} \frac{1}{2} h_{h}^{2} \left( \bar{\xi}_{h} - \xi \right)^{2} \]

for all arbitrary \( a_{h} \)'s subject to \( \sum_{h=1}^{k} a_{h}^{2} = 1 \). (15.2.7) is obviously a set of simultaneous confidence bounds on all arbitrary linear compounds of \( h_{h}^{2} (\bar{x}_{h} - \bar{x}) \) (\( h = 1, 2, \ldots, k \)), the compounding coefficients \( a_{h} \)'s being subject to \( \sum_{h=1}^{k} a_{h}^{2} = 1 \). It is also easy to verify that the set of such linear compounds could be otherwise written as

\[
(15.2.8) \quad \sum_{h=1}^{k} a_{h} \frac{1}{2} h_{h}^{2} c_{h}, \text{ for all } a_{h} \text{'s subject to } \sum_{h=1}^{k} a_{h}^{2} = 1 \text{ and } \sum_{h=1}^{k} a_{h} \frac{1}{2} h_{h}^{2} = 0. \]

(iv) For confidence bounds in the case of the general linear hypotheses we proceed as follows from the set-up of section (14). Suppose we have \( x_{h} \)'s (\( h = 1, 2, \ldots, n \)) which are \( n \) independent \( N(\mu_{h}, \sigma^{2}) \), such that, putting \( x'(1 \times n) = (x_{1}, x_{2}, \ldots, x_{n}) \), we have

\[
(15.2.9) \quad E(x)(n \times 1) = A(n \times m) \zeta(m \times 1),
\]

where \( m < n \), \( A \) is a matrix of rank, say \( r \geq m < n \), given by the experimental situation and \( \zeta(m \times 1) \) is a set of unknown parameters. Putting \( A'(m \times n) = \left( \begin{array}{c} A_{1}^{1} \\ A_{1}^{2} \\ \vdots \\ A_{n} \end{array} \right)^{r} \), let us assume, as we can without any loss of generality, that \( A_{1}(r \times n) \) is a set of independent row vectors which might be taken to be the basis of \( A'(m \times n) \). Suppose now that it is required to test a "testable" hypothesis

\[
(15.2.10) \quad C(q \times m) \zeta(m \times 1) = 0,
\]

where \( C \) is of rank \( s \leq \min(q, p) \leq m < n \). Putting
\[ (15.2.11) \quad C(q \times m)x(m \times 1) = s \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \begin{pmatrix} r \\ m-r \end{pmatrix}, \]

assume, without any loss of generality, that \( \begin{pmatrix} C_{11} & C_{12} \end{pmatrix} \) can be taken as the basis of \( C \) and notice also from (14) that for 'testability' we should have the further condition

\[ (15.2.12) \quad \begin{pmatrix} C_{12} \\ C_{22} \end{pmatrix} \begin{pmatrix} s \\ q-r \end{pmatrix} = \begin{pmatrix} C_{11} \\ C_{21} \end{pmatrix} \begin{pmatrix} r \end{pmatrix} \begin{pmatrix} A_1(r \times n)A_1(n \times r) \end{pmatrix}^{-1} A_1(r \times n)A_2(n \times m-r). \]

We recall from (14, 14.13) that the current F-test for (15.2.10) (at a level, say \( \alpha \)) has a critical region given by

\[ (15.2.13) \quad \frac{(n-r)x' A_1(A_1^t A_1)^{-1} x - C_{11}^{-1} C_{11} (A_1^t A_1)^{-1} A_1^t x}{s \sqrt{x' x - x' A_1(A_1^t A_1)^{-1} A_1^t x}} \geq F_{\alpha}(s, n-r). \]

Recall that when (15.2.10) is true, the left hand side of (15.2.13) has the central F-distribution with D.F. \( s \) and \( n-r \). Assume next that (15.2.10) is not true, but what is true is

\[ (15.2.14) \quad C(q \times m)x(m \times 1) = \eta(q \times 1) \quad (\eta \text{ being given}), \quad \text{or, say} \]

\[ s \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \begin{pmatrix} r \\ m-r \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \begin{pmatrix} s \\ q-s \end{pmatrix}. \]

Then proceeding exactly as in section (14) we check that if in the left side of (15.2.13) we replace \( x(n \times 1) \) by \( x(n \times 1) - \eta(n \times s) \eta_1^{-1} (s \times 1) \), the resulting expression is distributed as ordinary F with D.F. \( s \) and \( n-r \), \( B \) being given by

\[ (15.2.15) \quad B(n \times s) = A_1(n \times r)(A_1^t A_1)^{-1} (r \times r) C_{11}^{-1} C_{11} (A_1^t A_1)^{-1} C_{11} (s \times s). \]

If in (15.2.13) we now replace \( x \) by \( x - B \eta_1 \) and also put

\[ (15.2.16) \quad \sum C_{11}^{-1} C_{11}^{-1} = \tilde{U}(s \times s) \tilde{U}^t(s \times s) \quad (\text{notice that by (4.3.7) \( \tilde{U} \) is determinate and also unique}), \quad \text{then it is easy to see exactly in the same way as in} \]
the previous case that the resulting statement \(\iff\) the following

\[
\begin{align*}
(15.2.17) \quad & x' A_1 (A_1 A_1)^{-1} C_{11} U_a (r \times 1) - (E.V. )^{1/2} \sqrt{s F_a (s, n-r)} \frac{1}{\sqrt{2}} \\
& \leq \eta (1 \times s) B' (s \times n) A_1 (A_1 A_1)^{-1} C_{11} U_a \leq x' A_1 (A_1 A_1)^{-1} C_{11} U_a + \frac{1}{\sqrt{2}} \sqrt{s F_a (s, n-s)} \frac{1}{\sqrt{2}},
\end{align*}
\]

for all \(a\) subject to \(a' (1 \times r) a (r \times 1) = 1\), where \(B\) is given by (15.2.15) and \(\bar{U}\) by (15.2.16) and the error variance \(E.V.\) by

\[
(15.2.18) \quad E.V. = \sqrt{x' x - x' A_1 (A_1 A_1)^{-1} A_1 x} / (n-r).
\]

This, therefore, is a set of simultaneous confidence bounds (with a joint confidence coefficient \(1-\alpha\)) on all arbitrary linear functions of \(\eta\). It is easy to see that (15.2.17) subsumes as special cases, the confidence statements (15.2.1), (15.2.2) and (15.2.7). Nevertheless, for expository purposes, it is worthwhile to discuss separately the simpler cases first. Two other particular examples of (15.2.17), of special practical interest are also discussed, separately, in (v) and (vi).

(v) Suppose we have \(y_h\)'s (\(h = 1, 2, \ldots, n\)) each being an \(N(\theta_h, \sigma^2)\) such that

\[
\text{cov}(y_h, y_{h'}) = \rho \sigma^2 \quad (h \neq h' = 1, 2, \ldots, n),
\]

where \(\rho\) is known, but \(\theta_h\) and \(\sigma^2\) are unknown, but an independent estimate \(s^2\) of \(\sigma^2\) based on \(n\) degrees of freedom is available. It is required to obtain a set of simultaneous confidence bounds on the mean differences

\[
(15.2.19) \quad \theta_h - \theta_{h'}, \quad h, h' = 1, 2, \ldots, n, h \neq h'.
\]

We have now a finite set of parametric functions. Let \(z_h + x \bar{y} = y_h + x \bar{y}\)

where \(\bar{y} = (y_1 + y_2 + \ldots + y_n) / n, \bar{\theta} = (\theta_1 + \theta_2 + \ldots + \theta_n) / n\) and the disposable constant \(x\) is so adjusted that the \(z_h\)'s are uncorrelated. Then

\[
(15.2.20) \quad E(z_h) = \theta_h, \quad \text{var}(z_h) = \sigma^2 (1 - \rho) \quad h = 1, 2, \ldots, n.
\]

Let

\[
(15.2.21) \quad \psi_{hh'} = \frac{(z_h - \theta_h) - (z_{h'} - \theta_{h'})}{\sqrt{s / (1 - \rho)}} \quad h, h' = 1, 2, \ldots, n, h \neq h'.
\]

Then

\[
(15.2.22) \quad \psi_{hh'} \leq d
\]

implies

\[
(15.2.23) \quad y_h - y_{h'} - sd / (1 - \rho) \leq \theta_h - \theta_{h'} \leq y_h - y_{h'} + sd / (1 - \rho).
\]
Let $W_\varphi$ be the intersection of the regions (15.2.22). Then clearly the necessary and sufficient condition for the sample point to lie in $W_\varphi$ is that

\[(15.2.21)\quad q = \frac{w}{s/1-\rho}\]

where

\[(15.2.25)\quad w = \sup_{h,h'} \left( z_h - \varphi_h \right) - \left( z_{h'} - \varphi_{h'} \right), \quad h,h' = 1,2,\ldots,n; \ h \neq h'.\]

Thus if we set $d = q_a(n,n')$, where $q_a(n,n')$ is the upper $\alpha$-point of the distribution of the studentized range with $n, n'$ degrees of freedom, that is the ratio of the range of $n$ independent normal variates with zero mean to the square root of an independent estimate of their common variance based on $n'$ degrees of freedom, then the required simultaneous confidence intervals for the parametric functions (15.2.19) are

\[(15.2.26)\quad y_h - y_{h'} - \frac{\sqrt{(n,n')}}{1-\rho} \leq \varphi_h - \varphi_{h'} \leq y_h - y_{h'} + \frac{\sqrt{(n,n')}}{1-\rho}.

In particular $y_1, y_2, \ldots, y_n$ may be the means of $n$ random samples of equal size drawn from normal populations with a common (unknown) variance, or may be the estimated treatment effects in a randomized block or a balanced incomplete block experiment.

(vi) In factorial experiments we are usually interested in estimating linear functions of treatment effects, whose estimates are independently and normally distributed with a common variance, which can be independently estimated by an appropriate multiple of the error mean square in the analysis of variance. The distribution needed for simultaneous estimation in this case is slightly different from that occurring in (v).

Suppose, for example, that we have observations for a $2 \times 2 \times 2 \times 2$ factorial experiment with factors $A, B, C, D$, and that we are interested in simultaneously estimating the main effects and two factor interactions only. We shall suppose that the experiment is so laid out that none of these is confounded in any replication.

Let $t_{11}, t_{22}, t_{33}, t_{44}$ denote the true main effects and $t_{12}, t_{13}, t_{14}, t_{23}, t_{24}, t_{34}$ the true two factor interactions. The order of the subscripts in $t_{ij}$ is immaterial, that is, $t_{ij} = t_{ji}$. We can then write in the usual notation
(15.2.27) \[ t_{11} = \frac{1}{\sqrt{8}}(a-1)(b+1)(c+1)(d+1) \]

(15.2.28) \[ t_{12} = \frac{1}{\sqrt{8}}(a-1)(b-1)(c+1)(d+1) \]

with similar expressions for other main effects and interactions. Let \( y_{ij} \) be the estimate of \( t_{ij} \). Then reasoning as before we get the following simultaneous confidence intervals for \( t_{ij} \):

(15.2.29) \[ y_{ij} - s_a(n, n') \leq t_{ij} \leq y_{ij} + s_a(n, n') \]

where \( s^2 \) is an estimate of \( V(y_{ij}) \), based on \( n' \) degrees of freedom available for the estimate of error, and where \( n \), which is 10 in this particular example, is the number of linear functions to be estimated.

The meaning of \( x_a(n, n') \) is as follows. Let \( x_1, x_2, \ldots, x_n \) be independent normal variates with zero mean and variance \( \sigma^2 \). Let \( |x| \) be the maximum of \( |x_1| \), \( |x_2| \), \ldots, \( |x_n| \) and let \( s^2 \) be an independent estimate of \( \sigma^2 \) based on \( n' \) degrees of freedom. Then \( x_a(n, n') \) is the upper \( \alpha \)-point of the distribution of \( |x|/s \).

In a factorial experiment in which each factor is at more than two levels, the above will still apply if the \( n \) linear functions to be simultaneously estimated (or tested for vanishing) are so chosen that their estimates are independently distributed with a common variance.

15.3. Applications to variances of one or two normal populations.

(i) Given a random sample of size \( n+1 \) (mean: \( \bar{x} \) and S.D.: \( s \)) from an \( N(\mu, \sigma^2) \), we take over from (7.3.1) the following statement with probability \( 1-\alpha \):

(15.3.1) \[ \chi^2_{1\alpha}(n) \leq ns^2/\sigma^2 \leq \chi^2_{2\alpha}(n), \]

where \( \chi^2_{1\alpha}(n) \) and \( \chi^2_{2\alpha}(n) \) are the upper \( \alpha \) and lower \( \alpha \) point of \( \chi^2 \)-distribution with D.F. \( n \) and \( \alpha \) is partitioned with \( \alpha_1 \) and \( \alpha_2 \) such that (a) \( \alpha_1 + \alpha_2 = \alpha \) and (b) the complement of (15.3.1), i.e., the critical region is locally unbiased (in the neighborhood of \( \sigma \)) and has also been shown to have the monotonicity property. We now rewrite (15.3.1) as

(15.3.2) \[ ns^2/\chi^2_{2\alpha}(n) \leq \sigma^2 \leq ns^2/\chi^2_{1\alpha}(n), \]
which gives confidence bounds on $\sigma^2$ with a confidence coefficient $1-\alpha$ and having properties in terms of shortness similar to those possessed by (15.3.1) of the second kind of error, already discussed.

(ii) Given two random samples of sizes $n_1+1$ (mean: $\overline{x}_h$ and S.D.: $s_h$) ($h = 1, 2$) from two $N(\xi_h, \sigma_h^2)$, we take over from (7.3.2) the following statement with probability $1-\alpha$:

\begin{equation}
F_{1a}(n_1, n_2) \leq \frac{s_1^2}{s_2^2} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq F_{2a}(n_1, n_2),
\end{equation}

where $F_{1a}(n_1, n_2)$ and $F_{2a}(n_1, n_2)$ are the upper $a_1$ and lower $a_2$ points of $F$-distribution with D.F. $n_1$ and $n_2$ and $a$ is partitioned with $a_1$ and $a_2$ such that (a) $a_1 + a_2 = a$ and (b) the compliment of (15.3.1), i.e., the critical region is locally unbiased (in the neighborhood of $\sigma_1/\sigma_2$) and has also been shown to have the monotonicity property. We now rewrite (15.3.3) as

\begin{equation}
\frac{s_1^2}{s_2^2} / F_{2a}(n_1, n_2) \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} / F_{1a}(n_1, n_2),
\end{equation}

which gives confidence bounds on $\sigma_1^2/\sigma_2^2$ with a confidence coefficient $1-\alpha$ and having properties in terms of shortness similar to those possessed by (15.3.3) in terms of the second kind of error, already discussed.

15.4. Application to the coefficient of regression for a bivariate normal population. Let $x_1$ and $x_2$ be distributed as a bivariate normal with variances $\sigma_1^2$ and $\sigma_2^2$ and correlation coefficient $\rho$, and let the sample variances (on a sample of size $n+2$) be denoted by $s_1^2$ and $s_2^2$, and the sample correlation coefficient by $r$. Also let $b_1 = s_1^2 r/s_2$ and $\theta_1 = \sigma_1^2 \rho / \sigma_2$. It is easy to check that the variates $(x_1 - \theta_1 x_2)$ and $x_2$ are uncorrelated, so that when the population parameters are $\sigma_1$, $\sigma_2$, and $\rho$, $n^{1/2} r^*/(1-r^*)^{1/2}$ has the t-distribution with $n$ D.F. Here $r^*$ stands for the sample correlation between $(x_1 - \theta_1 x_2)$ and $x_2$, that is,
\[ (15.4.1) \quad r^* = \frac{s_1s_2r - \beta_1s_2^2}{(s_1^2 - 2\beta_1s_1s_2r + \beta_1^2s_2^2)^{1/2}s_2} \]

\[ = \frac{(s_1 - \beta_1s_2)^2}{(s_1 - \beta_1s_2)^2 + (1-r^2)s_2^2} \frac{1}{(s_1 - \beta_1)^2 + (1-r^2)s_2^2} \frac{1}{s_2} \]

and, therefore,

\[ (15.4.2) \quad \frac{r^*}{1-r^*} = \frac{s_2}{s_1} \frac{b_1 - \beta_1}{(1-r^2)^{1/2}}. \]

Now consider the statement

\[ (15.4.3) \quad -t_a(n) \leq n^{1/2}r^*/(1-r^*2)^{1/2} \leq t_a(n), \]

where \( t_a(n) \) gives the upper \( a/2 \)-point of the \( t \)-distribution with \( n \) D.F. This is easily seen to reduce to the following confidence statement on \( \beta_1 \) (with a confidence coefficient \( 1-a \)):

\[ (15.4.4) \quad b_1 - \frac{t_a(n)}{\sqrt{n}} (1-r^2)^{1/2} \frac{1}{s_2} s_1 \leq \beta_1 \leq b_1 + \frac{t_a(n)}{\sqrt{n}} (1-r^2)^{1/2} \frac{1}{s_2} s_1. \]

By inversion of (15.4.4) the test that we obtain for the associated hypothesis

\( H_0: \beta_1 = 0, \) that is \( \rho = 0, \) is easily checked to be the customary test based on "\( r\)" and hence just the \( t \)-test. Similar procedures would go through for "partial regressions" or "multiple regressions". The interesting point here is that it would be far more difficult to give corresponding confidence bounds to \( \rho, \) because this would have to be done by inverting the distribution of the noncentral \( r, \) which is quite complicated.

16. Multivariate confidence bounds \( \text{[22,24].} \)

16.1 A convenient notation. From now on we shall make use of a rather convenient notation. A random sample of size \( n \) from a \( p \)-variate normal population, i.e., an \( X(p \times n) \) having the p.d.f.

\[ (2\pi)^{-pn/2} |\Sigma|^{-n/2} \exp \left[-\frac{1}{2} \text{tr} \Sigma^{-1}(X-\xi) (X' - \xi') \right] \]
where \( \xi(p \times n) \) stands for a \( p \times n \) matrix each column of which is the same \( p \times 1 \) vector \( \xi \) (with components \( \xi_1, \ldots, \xi_p \)) will be referred to as \( X(p \times n): N(\xi, \Sigma) \). A matrix \( Y(p \times n) \) having the p.d.f.

\[
(2\pi)^{-pn/2} |\Sigma|^{-n/2} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} YY' \right\},
\]

will be referred to as \( Y(p \times n): N(0, \Sigma) \). We recall from (5.14) - (5.14) that starting an \( X(p \times (n + 1)) \): \( N(\xi, \Sigma) \) and transforming and integrating we can always have an \( Y(p \times n): N(0, \Sigma) \), such that

\[
(16.1.1) \quad nS(p \times p) = Y(p \times n) Y'(n \times p) = X(p \times (n+1)) X'((n+1) \times p)
\]

\[= (n+1) \overline{X}(p \times 1) \overline{X}'(1 \times p),\]

where \( \overline{X}(p \times 1) = \frac{1}{n} X(p \times n) 1(n \times 1) \), \( 1(n \times 1) \) being an \( n \times 1 \) column vector with components \( 1, 1, \ldots, 1 \).

16.7. Confidence bounds relating to \( \xi \) from an \( N(\xi, \Sigma) \). Given an \( X(p \times (n+1)) \): \( N(\xi, \Sigma) \), suppose we try to obtain simultaneous confidence bounds on arbitrary linear compounds of the population mean vector \( \xi \). Consider the statement that

\[
\frac{1}{n+1} |a' (\overline{X} - \xi)| / (a' Sa)^{1/2} \leq c,
\]

or

\[
(16.2.1) \quad (n+1) a' (\overline{X} - \xi) (\overline{X}' - \xi') a/a' Sa \leq c^2,
\]

where \( \overline{X} \) is the sample mean vector and \( S \) is the sample covariance matrix, already defined in Section 11.1, and \( a(p \times 1) \) is an arbitrary nonnull nonstochastic column vector and \( c \) is a given positive constant. The statement (16.2.1) stems from the customary Student's t-test and the associated confidence interval (both having well known optimum properties) relating to the parameter \( a' \xi \). Now, for a given (positive) \( c \) and given \( \overline{X}, \xi, S \) and of course \( n \), the set of all statements (16.2.1) for all possible nonnull vectors \( a \) is exactly equivalent to the statement that
(16.2.2) \[ \sup_a (n+1) a^t (\bar{x} - \bar{x}^*) (\bar{x}^* - 1) a / a^t S a \leq c^2. \]

It is well known that this "sup" comes out as \( \text{tr}(n+1) S^{-1}(\bar{x} - \bar{x}^*) (\bar{x}^* - 1) \) or as \( \text{tr}(n+1) (\bar{x}^* - 1) S^{-1}(\bar{x} - \bar{x}^*) \) (since \( \text{tr} AB = \text{tr} BA \)), or simply as \( (n+1)(\bar{x}^* - 1) S^{-1}(\bar{x} - \bar{x}^*) \) (since \( \text{tr} \) scalar = scalar). It is also well known that under the null hypothesis, this is distributed as the central Hotelling's \( T^2 \) with D.F. \( p \) and \( n+1-p \) and that where in this statistic \( \bar{x} \) is replaced by \( \bar{x}^* \) (\( \neq \bar{x} \)), the resulting statistic is distributed as the noncentral Hotelling's \( T^2 \) with the same D.F. and with the non-centrality parameter \( \gamma^2 = (\bar{x}^* - \bar{x}^*) S^{-1}(\bar{x}^* - \bar{x}). \) Going back to (16.2.1) it is thus easy to see that if, for all \( \bar{x} \) and all nonnull \( a \),

\[
(16.2.3) \quad \frac{(n+1) a^t (\bar{x} - \bar{x}^*) (\bar{x}^* - 1) a}{a^t S a} \leq c^2 \quad \text{if} \quad \bar{x}^* = \bar{x} \quad \text{and} \quad \gamma = 1 - \alpha,
\]

then \( c^2 = T^2_\alpha \) is the upper \( \alpha \)-point of the central Hotelling's \( T^2 \)-distribution with D.F. \( p \) and \( n+1-p \) and can be conveniently written as \( T^2_\alpha (p, n+1-p) \). From (16.2.3) we have thus, with a confidence coefficient \( 1-\alpha \), the set of simultaneous or multiple confidence bounds (for all \( \bar{x} \) and all nonnull \( a \)):

\[
(16.2.4) \quad a^t \bar{x} - \sqrt{T^2_\alpha (a^t S a)/(n+1)} \|
\frac{1}{2} \leq a^t \bar{x} \leq a^t \bar{x} + \sqrt{T^2_\alpha (a^t S a)/(n+1)} \|
\frac{1}{2}.
\]

It should be noted that (16.2.4) gives the simultaneous confidence bounds on all arbitrary linear compounds of the \( p \) components of the population mean vector \( \bar{x} \). The shortness (in the sense of probability) of this set of confidence bounds, that is, the probability of these bounds covering \( \bar{x}^* \) when, in fact, \( \bar{x}^* \neq \bar{x} \), is obviously

\[ 1 - P \left[ T^2_{\text{noncentral}} > T^2_\alpha \right] \gamma^2 \quad \gamma = 1 - \alpha. \]

From the well known fact that the power function of Hotelling's \( T^2 \)-test is a monotonically increasing function of the nonnegative \( \gamma \), it follows, therefore, that the shortness of the confidence bound (16.2.4) tends to zero as \( \gamma \to \infty \).

16.3 Confidence bounds relating to mean differences from \( N(\bar{x}_h, \Sigma) \) (\( h = 1, 2, \ldots, k \)).

Given \( X_h (p \times (n_h + 1)) : N(\bar{x}_h, \Sigma) \), (\( h = 1, 2, \ldots, k \)) let us try to obtain a set of simultaneous confidence bounds on all arbitrary double linear compounds of the \( p \)-components
of the \( k \) population mean vectors measured from the weighted grand mean vector. Consider now the statement

\[
(16.3.1) \quad \left| \sum_{h=1}^{k} b_h a_i^j (n_h + 1)^{1/2} (\bar{x}_{h} - \bar{x}_+ + \xi) \right| \leq (k-1) g^2 a_i^j S_a^{-1/2}
\]

where \( \bar{x}_h \) is the mean vector for the \( h \)th sample,

\[
\bar{x} = \frac{1}{k} \sum_{h=1}^{k} n_h \bar{x}_h / \sum_{h=1}^{k} n_h + 1, \quad \xi = \frac{1}{k} \sum_{h=1}^{k} \xi_h / \sum_{h=1}^{k} n_h + 1,
\]

where \( S \) is the pooled "within" covariance matrix of the \( k \)-samples, given by

\[
(k \sum_{h=1}^{k} n_h) S = \sum_{h=1}^{k} n_h \bar{x}_h \bar{x}_h' - (n_h + 1) \bar{x}_h \bar{x}_h' / n_h + 1,
\]

and \( g \) is a given positive constant, \( a(p \times 1) \) is an arbitrary nonnull stochastic column vector and the \( b_h \)'s are arbitrary coefficients subject to \( \sum_{h=1}^{k} b_h^2 = 1. \)

If we now use the result that

\[
\left| \sum_{h=1}^{k} b_h y_h \right| \leq \sqrt{d^2} \iff \sum_{h=1}^{k} y_h^2 \leq d^2,
\]

then it directly follows that, given all the other quantities including \( a \), and under all possible variations of \( b_h \)'s subject to \( \sum_{h=1}^{k} b_h^2 = 1 \), the statement (16.3.1) is precisely equivalent to the statement that

\[
\sum_{h=1}^{k} (n_h + 1)^{1/2} (\bar{x}_h - \bar{x}_+ + \xi) / (k-1) a_i^j S_a^{-1/2} \leq g^2,
\]

or

\[
(16.3.2) \quad \sum_{h=1}^{k} (n_h + 1)(\bar{x}_h - \bar{x}_+ + \xi)(\bar{x}_h' - \bar{x}_+ + \xi') a / (k-1) a_i^j S_a^{-1/2} \leq g^2.
\]

Letting now \( a \) vary and putting

\[
(16.3.3) \quad (k-1) S^* = \sum_{h=1}^{k} (n_h + 1)(\bar{x}_h - \bar{x}_+ + \xi)(\bar{x}_h' - \bar{x}_+ + \xi')
\]

the statement (16.3.2), for all possible values of the nonnull \( a \), is precisely equivalent to:

\[
(16.3.4) \quad \sup_{a} \int a_i^j S_a^* a / a_i^j S_a^{-1/2} \leq g^2.
\]

As observed after (7.4.7) \( S \) is, a.e., p.d. and \( S^* \) is, a.e., p.s.d. of rank -
\[ q = \min(p, k-1) \] (p.s.d. if \( p > k-1 \) and p.d. if \( p \leq k-1 \)) and \( \sup_{a} S_{a}^* / a' S_{a} a \) is just the largest root \( c_{q} \) of the \( p \)th degree determinantal equation in \( c : \left| S^* - cS \right| = 0 \). Of this equation all roots are nonnegative, \( p-q \) of them always zero and \( q \) are, a.c., positive. Thus (16.3.4) and hence (16.3.2) under all permissible variations of \( a \) and the \( b_{h} \)'s, turns out to be equivalent to:

\[ (16.3.5) \quad c_{q} \leq g^2. \]

The distribution of \( c_{q} \) on the null hypothesis is known and relatively easy and involves as parameters \( p, k-1, \Sigma n_{h} \). Computation of the 5 per cent and 1 per cent points is in progress. Thus if

\[ (16.3.6) \quad P_{c_{q} \leq c_{a} \mid \text{null hypothesis}} = 1-a, \]

we can write \( c_{a} = c_{a}(p, k-1, \Sigma n_{h}) \), and now combining (16.3.1)-(16.3.6) we have, with a confidence coefficient \( 1-a \), the following set of multiple confidence statements

(for all \( \xi_{h} \)'s, all nonnull \( \xi_{h} \)'s and all \( b_{h} \)'s subject to \( \Sigma b_{h}^2 = 1, \)

\[ (16.3.7) \quad \sum_{h=1}^{k} b_{h} a_{h} (n_{h}+1)^{1/2} (x_{h} - \bar{x}) - \sum_{h=1}^{k} (k-1)c_{a} a_{h} S_{a} a \leq \sum_{h=1}^{k} b_{h} a_{h} (n_{h}+1)^{1/2} (\xi_{h} - \bar{\xi}) \leq \sum_{h=1}^{k} b_{h} a_{h} (n_{h}+1)^{1/2} (x_{h} - \bar{x}) + \sum_{h=1}^{k} (k-1)c_{a} a_{h} S_{a} a \leq \sum_{h=1}^{k} (\xi_{h} - \bar{\xi}) \]

where \( c_{a} = c_{a}(p, k-1, \Sigma n_{h}) \). This gives simultaneous confidence bounds on all arbitrary double linear compounds of the \( p \) components of the difference between the \( k \) population mean vectors \( \xi_{h} \)'s and the weighted grand mean of those which is \( \bar{\xi} \). To discuss the shortness of (16.3.7) consider the non-central distribution of \( c_{q} \), where \( c_{q} \) is defined after (16.3.4), i.e., \( c_{q} \) is the largest root of the equation in \( c \):

\[ (16.3.8) \quad \left| S^* - cS \right| = 0, \text{ where } S^* \text{ is given by (16.3.3)}. \]

It is easy to see that the distribution of the non-central \( c_{q} \) is really the distribution of \( c_{q} \), where \( c_{q} \) is the largest root of the equation in \( c \) obtained by (i) replacing in (16.3.2), \( \xi_{h} \) and \( \bar{\xi} \) by \( \xi_{h}^* (\neq \bar{\xi}_{h}) \) and \( \bar{\xi}^* (\neq \bar{\xi}) \) and (ii) substituting the
resulting value of $S^*$ in (16.3.8) and (iii) assuming that the true population parameters are $\xi_h$ and $\xi$. The distribution is extremely difficult but is well known (see (8.6)) to involve as parameters the positive roots $\gamma_1, \ldots, \gamma_s$ ($s \leq \min(p,k-1)$) of the determinantal equation in $\gamma$: $\det (S - \gamma \Sigma) = 0$, where $\Sigma$ is the common covariance matrix of the $k$ populations and $\Sigma = (k-1)^{-1} \sum_{h=1}^{k} (n_h+1)(\xi^*_h - \xi_h + \xi^* - \xi)^T (\xi^*_h - \xi_h + \xi^* - \xi)$. This $\Sigma$ is necessarily at least p.s.d. of rank $\leq \min(p,k-1) = s$ (say), so that out of the $p$ roots of the equation in $\gamma$, $p-s$ are zero and $s$ positive. Using (11.3) and (13.3) we observe that there is a good upper bound to the shortness of (16.3.7) and that the shortness is a monotonically decreasing function of the deviation parameters and tends to zero as those tend to infinity. With two populations (and samples), we have $q = \min(p,1) = 1$, and thus only one positive sample root, say $c$, and at the most one positive population root, say $\gamma$. It is easy to check that in this case

$$(16.3.9) \quad c = \frac{(n_1+1)(n_2+1)}{n_1 + n_2 + 2} \text{tr} \ S^{-1}(\bar{x}_1 - \bar{x}_2)^T \Sigma w_1 \bar{x}_2, \quad \gamma = \frac{(n_1+1)(n_2+1)}{n_1 + n_2 + 2} \text{tr} \ S^{-1}(\xi^*_1 - \xi_1 + \xi^* - \xi_1)^T \Sigma (\xi^*_1 - \xi_1 + \xi^* - \xi_1)$$

and it is well known that, on the null hypothesis, $c$ is distributed as central Hotelling $T^2$ with D.F. $p$ and $n_1+n_2+1-p$, and on the alternatives as non-central Hotelling $T^2$ with the same D.F. and with a deviation parameter $\gamma$. It is also easy to check that in this case the confidence statement (16.3.7) reduces to

$$(16.3.10) \quad a'(\bar{x}_1 - \bar{x}_2) + \frac{n_1+n_2+2}{(n_1+1)(n_2+1)} T^2 a S_a \rightarrow^{1/2} a'(\xi^*_1 - \xi_1)$$

$$\leq a'(\bar{x}_1 - \bar{x}_2) + \frac{n_1+n_2+2}{(n_1+1)(n_2+1)} T^2 a S_a \rightarrow^{1/2},$$

where $T^2 a = T^2_a(p,n_1+n_2+1-p)$ is the upper $a$-point of Hotelling's $T^2$. The shortness of (16.3.10) is exactly known and of course tends to zero as $\gamma \rightarrow \infty$.

16.4. An important subset of the set of bounds (16.3.7). Suppose now that, instead
of all contrasts of the type: \( \sum b_h a_i(n_h + 1)^{1/2}(x_i - \bar{x}) \) (with given restrictions on \( a \) and the \( b \)'s), we are interested in contrasts of the type: \( a'(x_i - \bar{x}) \), for all nonnull \( a \) and all \( h \neq \ell = 1, 2, \ldots, k \). It is easy to offer a multiple set of confidence bounds for contrasts of this type, which can be regarded as one kind of multivariate (under unequal sample sizes) analogue of a somewhat similar set given by Tukey \( \ell \sum \) for the corresponding univariate situations, and discussed in Section (15.2). The proposed set is built up as follows. With the same notation as before, and with

\[
T_{h\ell}^2 = n_h(x_i - x_\ell) S_a^{-1}(x_i - x_\ell) + T_{a}^2 \sum a_i(x_i - x_\ell) \frac{S_a}{n_h} \frac{T_a^2 a_i S_a}{n_h} \frac{1}{\sqrt{2}}.
\]

Thus, for a given pair \((h, \ell)\), the statement that \( T_{h\ell}^2 \leq T_a^2 \) is exactly equivalent to the statement that, for all nonnull \( a \)'s,

\[
a'(x_i - x_\ell) - \sqrt{T_a^2 a_i S_a} \frac{1}{n_h} \frac{1}{2} \leq a'(x_i - x_\ell) \leq a'(x_i - x_\ell) + \sqrt{T_a^2 a_i S_a} \frac{1}{n_h} \frac{1}{2}.
\]

We observe that when the true population means are \( \bar{x}_h \)'s, \( T_{h\ell}^2 \) is distributed as Hotelling's \( T^2 \) with \( D.F. p \) and \( \sum_{h=1}^k n_h + 1 - p \).

Now, considering all pairs \((h, \ell)\) out of \( k \) samples (and \( k \) populations), it is easy to see that the statement: all \( T_{h\ell}^2 \)'s \( \leq T_a^2 \), is precisely equivalent to the statement that the largest \( T_{h\ell}^2 \) out of all pairs is \( \leq T_a^2 \), which again is equivalent to the statement that, for all nonnull \( a \)'s and all pairs \((h, \ell)\) out of \( k \),

\[
(16.4.1) \quad a'(x_i - x_\ell) - \sqrt{T_a^2 a_i S_a} \frac{1}{n_h} \frac{1}{2} \leq a'(x_i - x_\ell) \leq a'(x_i - x_\ell) + \sqrt{T_a^2 a_i S_a} \frac{1}{n_h} \frac{1}{2}.
\]

If the confidence coefficient of \( (16.4.1) \) is to be \( 1 - \alpha \), then \( T_a = T_a(p, n_1, n_2, \ldots, n_k) \) will be given by

\[
(16.4.2) \quad P \left[ \text{Largest } T_{h\ell}^2 \text{ out of } \binom{k}{2} \text{ pairs } \geq T_a^2 \left| \text{null hypothesis } H = \alpha. \right. \right.
\]

It is obvious that the distribution of the largest \( T_{h\ell}^2 \) involves as parameters just \( p \) and \( n_1, n_2, \ldots, n_k \). It is easy to see that the distribution is manageable only when the number of parameters is small. In particular, the case that \( n_1 = n_2 = \ldots = n_k \) and
p=1, is identical with the one considered in Section 15.2. It may also be noted that when k=2, the largest $T^2_{h,\ell}$ will of course be Hotelling's $T^2$ distributed with D.F. p and $n_1+n_2+1-p$. Also the shortness of the confidence bounds (16.4.1) can be formally written as

$$P \left( \begin{array}{c} k \\ 2 \end{array} \right) \text{Largest } T^2_{h,\ell} \text{ out of } \left( \begin{array}{c} k \\ 2 \end{array} \right) \text{ pairs } \leq T^2_a (p, n_1, n_2, \ldots, n_k) \right| \text{ alternative } \beta.$$

It is important to observe that while each $T^2_{h,\ell}$ is individually distributed (on the null hypothesis) as a central Hotelling's $T^2$ with D.F. p and $\sum_{h=1}^{k} n_h + 1 - p$, the $\left( \begin{array}{c} k \\ 2 \end{array} \right)$'s are not independent, nor do we know what the distribution of the largest central $T^2_{h,\ell}$ is, to say nothing of the noncentral case, so that the confidence statement (16.4.1) has not been reduced to concrete terms as was done for the other cases discussed in this section. The distribution problem arising in this situation is now under investigation.

For the associated problem of testing $H_0: \xi_1 = \cdots = \xi_k$, we set up as before the rule that if, for all nonnull $\xi$ and all pairs $(h, \ell)$, the bounds (16.4.1) include zero, we accept $H_0$ and reject it otherwise. The properties (including power) of this test are tied up in an obvious manner with those of the multiple confidence interval statement (16.4.1).

Notice that so far, in testing of hypotheses by inversion of confidence statements, we have considered two-decision problems. Suppose, at this point, for purposes of illustration, we offer a multi-decision procedure, namely that, for a given pair $(h, \ell)$, we accept or reject $H(\bar{\xi}_h = \bar{\xi}_\ell)$ according as all those bounds (16.4.1) which involve $\bar{x}_h$ and $\bar{x}_\ell$ only include or exclude zero. It is obvious that in all other situations considered so far we could set up similar multi-decision procedures.

16.5. Further observations. In many situations it might be of greater physical interest to be able to make, instead of (16.3.7) or (16.4.1), a set of just $p \times \left( \begin{array}{c} k \\ 2 \end{array} \right)$ confidence interval statements, each relating to just one variate and difference
between one of \( \binom{k}{2} \) pairs. In other words, if \( \xi_h = (\xi_{1h}, \xi_{2h}, \ldots, \xi_{ph}) \) \( (h=1,2,\ldots,k) \) denote the \( p \) means for the \( h^{th} \) population, then we would like to make a statement of the form

\[
(f_{jhh'}(X_1, X_2, \ldots, X_k) \leq \xi_{jhh'} \leq F_{jhh'}(X_1, X_2, \ldots, X_k))
\]

(with obvious applications to subsection 15.2), for all \( h \neq h' = 1,2,\ldots,k \) and all \( j = 1,2,\ldots,p \), where \( f_{jhh'} \) and \( F_{jhh'} \) are supposed to be two different functions of the whole set of \( p \times \prod_{h=1}^{k} (n_h + 1) \) raw observations. It is clear that (16.5.1) is a subset of (16.4.1) which again is a subset of (16.3.7). Whether it is possible to make a statement like (16.5.1) in an elegant and useful way (i.e., with manageable functions \( f_{jhh'} \) and \( F_{jhh'} \)) and with a given joint confidence coefficient \( 1-\alpha \), that is, free of the nuisance parameters \( \Sigma \), is still an open question. It may well be that a range (not too wide) for the confidence coefficient itself is called for. Furthermore, whatever set of confidence intervals like (16.5.1) we propose, be it under a fixed confidence coefficient or under a confidence coefficient lying in a short range, the "goodness" of such a set would pose further questions. It is believed that in this situation a more promising approach might be one involving a suitable two-stage procedure.

16.6. Confidence bounds connected with a general linear hypothesis. In place of the set-up of section (16.3) consider the more general set-up of (iii c) of section 6, which is the following. We have an \( X(p \times n) \) whose column vectors are independently distributed, the \( r^{th} \) vector \( x_r \) being \( N(E(x_r), \Sigma) \) \( (r=1,2,\ldots,n) \). It is also given that \( E(X'(n \times p)) = A(n \times m)\xi(m \times p) \) where \( \xi \) is a set of unknown parameters and \( A(n \times m) \) is given by the experimental situation such that it is of rank \( r \leq m < n \). Also setting

\[
(A'_{1}(m \times n) = \begin{pmatrix} A_1^T \\ A_2^T \end{pmatrix}^{r \times (m-r)}
\]

in
let $A_1$ be a basis of $A'$, i.e., of $A$. Next consider a matrix $C(q \times m)$ of rank $s \leq \min(q, r) \leq m < n$ with a structure given by

$$
C(q \times m) = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}^{s}_{q-s} \begin{pmatrix} r \\ m-r \end{pmatrix}
$$

such that $(C_{11} \ C_{12})$ forms a basis and also that $C$ satisfies (15.2.12). Then combining the results (15.2.9)-(15.2.17) with the results (16.3.1)-(16.3.7) and (16.3.8) it is easy to check that the following is a set of simultaneous confidence bounds (with a confidence coefficient $1-\alpha$)

$$
(16.6.3) \ a'(1 \times p)X(p \times n)A_1(n \times r)(A_1'A_1)^{-1}(r \times r)C_{11}(r \times s)U(s \times s)b(s \times 1)
$$

$$
-(a'Sa)^{1/2} \int_{S_a}^{(p, s, n-r)} \frac{1}{(a'Sa)^{1/2}} \leq a' \eta'(p \times s)B'(s \times n)A_1(A_1'A_1)^{-1}C_{11}U_b
$$

$$
\leq a' \ A_1(A_1'A_1)^{-1}C_{11}U_b + (a'Sa)^{1/2} \int_{S_a}^{(p, s, n-r)} \frac{1}{(a'Sa)^{1/2}}
$$

for all non-null $a'(1 \times p)$ and all $b(s \times 1)$ subject to $b'b = 1$, where $B$ is given by (15.2.15), $U$ by (15.2.16), $S$ by

$$
(16.6.4) \ S(p \times p) = \int XX'X_1A_1'(A_1'A_1)^{-1}A_1'X_1/(n-r),
$$

$\eta(s \times p)$ is given by

$$
(16.6.5) \ s(C_{11} \ C_{12})\xi(m \times p) = \eta(s \times p),
$$

and $c_a(p, s, n-r)$ is the upper a point of the distribution of the largest root of (7.4.7) with D.F. $(p, s, n-r)$. The confidence bounds (16.6.3) are thus seen to be really on arbitrary double linear compounds of $(C_{11} \ C_{12})\xi$.

16.7. **Confidence bounds on departures from a particular kind of multicollinearity of means.** For $k(p+q)$-variate $N(\xi_1, \Sigma)$ (with $k > p+q$), where $\Sigma((p+q) \times (p+q))$ is symmetric p.d. with submatrices $\Sigma_{11}(p \times p)$, $\Sigma_{22}(q \times q)$ and $\Sigma_{12}(p \times q)$, and $\xi_1((p+q) \times 1)$ has column subvectors $\xi_{11}(p \times 1)$ and $\xi_{21}(q \times 1)$, let us consider the hypothesis $H_0: \xi_{11} - \Sigma_{12} 22^\Sigma_{21} = 0$ ($i = 1, 2, \ldots, k$). Notice (from (16.11.12)) that $\Sigma_{12} \Sigma_{22}^{-1}$ can be appropriately regarded as the matrix of regression of the set of p
variates on the set of q variates. The hypothesis \( H_0 \) can thus be stated otherwise as the hypothesis that the matrix of means of the first p variates, viz,
\[
(\xi_{11} \quad \xi_{12} \quad \cdots \quad \xi_{1k})
\]
is equal to the matrix of means of the remaining q variates, premultiplied by the regression matrix of the p variates on the q variates. We are now interested in setting confidence bounds on
\[
(16.7.1) \quad \xi_{1i} - \Sigma_{12}^{-1} \Sigma_{22}^{-1} \xi_{2i} = 0 \quad (i = 1, 2, \ldots, k),
\]
which, naturally, are departures from \( H_0 \). More properly speaking, we shall be interested in setting simultaneous confidence bounds on arbitrary bilinear compounds
\[
a'(1 \times p)b(p \times k)b'(k \times 1),
\]
where \( b \) is a \((p \times k)\) matrix with column vectors given by
\[
(16.7.1). \quad \text{Now taking the 'residuals' of the first p variates with respect to the remaining q variates after the manner of (A.3.17) it is easy to check that for the i^\text{th} \text{ population the residuals will be distributed as a p-variate normal with a covariance matrix } \Sigma_{11}^{-1} \Sigma_{12}^{-1} \Sigma_{22}^{-1} \Sigma_{12}^{-1} \text{ and about the mean vector } \xi_{1i} - \Sigma_{12}^{-1} \Sigma_{22}^{-1} \xi_{2i} \text{ (with } i = 1, 2, \ldots, p). \text{ Also the 'within' covariance matrix of the 'residuals', pooled from k samples of size, say n each, will be given by}
\]
\[
(16.7.2) \quad S_{1.2} = S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}^{-1},
\]
where \( S_{11} \) \((p \times p)\), \( S_{22} \) \((q \times q)\) and \( S_{12} \) \((p \times q)\) stand for the submatrices of the 'within' covariance matrix (of the p+q variates) pooled from the k samples. The mean vector for the i^\text{th} sample will be given by
\[
(16.7.3) \quad \xi_{1i} - S_{12}^{-1} S_{22}^{-1} \xi_{2i} = 0 \quad (with \ i = 1, 2, \ldots, k).
\]
Let \( B(p \times k) \) stand for the \((p \times k)\) matrix with the k column vectors given by
\[
(16.7.3).
\]
Thus exactly as in section (16.6) we have with a confidence coefficient, say \( 1 - \alpha \), the following set of simultaneous confidence bounds (for all arbitrary non-null \( a' \) \((1 \times p)\) and unit length \( b(k \times 1)\):
\[
(16.7.4) \quad a'Bb - \sqrt{k(a'S_1 .2a)c_a(p,k,nk-k)} \gamma^1/2 \leq \alpha^1b \leq a'Bb + \sqrt{k(a'S_1 .2a)c_a(p,k,nk-k)} \gamma^2/2,
\]
where \( c_a(p,k,nk-k) \) is the upper \( \alpha \)-point of the distribution of the (central) largest
characteristic root based on $p$, $k$ and $n-k$ degrees of freedom. The test for the associated hypothesis $H_0$ is also easily obtained, the critical region being given by

$$c_p \geq c_{a}(p,k,nk-k),$$

where $c_p$ is the largest root of $\frac{1}{k}(BB')S_{1.2}^{-1}$. Notice that $\beta$ and $B$ are each a $(p \times k)$ matrix with $k$ column vectors given respectively by (16.7.1) and (16.7.3).

16.8. **Confidence bounds on departures from another kind of multicollinearity of means.** It seems that when the population covariance matrix $\Sigma$ is not supposed to be known there are two kinds of multicollinearity of means (and departures from it) which can be properly handled, namely, (i) that the matrix of means of the first $p$ variates is a constant matrix times the matrix of means of the remaining $q$ variates, the constant matrix factor being equal to the regression matrix of the $p$-set on the $q$-set, whatever this regression matrix might be and (ii) that the matrix of means of the first $p$ variates is a constant (and given) matrix times the matrix of means of the remaining $q$ variates. Case (i) is the one discussed in section 16.7 while case (ii) belongs to linear hypothesis in multivariate analysis of variance of means and has already been discussed in section 16.6.

16.9. **Confidence bounds connected with $\Sigma$ of an $N(\xi, \Sigma)$.** Let us start from a $Y(p \times n): N(\Omega, \Sigma)$, where $\Sigma(p \times p)$ is supposed to be p.d. (so that its characteristic roots are all positive). For simplicity we also assume that $p \leq n$, so that, a.e., $YY'$, that is, $nS$ is p.d., and hence all its characteristic roots are positive. We now recall the well known result (A.3.3) that there exists an orthogonal $\Gamma(p \times p)$ such that $\Sigma(p \times p) = \Gamma(p \times p)D_{\gamma}(p \times p)\Gamma'(p \times p)$ where the $\gamma$'s are the characteristic roots of $\Sigma$. If the roots are distinct then by a convention, say by taking all the elements of the first row of $\Gamma$ to be positive, the transformation could be made one-to-one. However, we do not need this for our present purpose. Note that the number of independent elements on both sides is the same. Except for the factor $(-\frac{1}{2})$ the argument under the exponential in the probability density of $Y$ can now be written, if
if we put \( \Delta = \gamma^{-1/2} \), as

\[
\text{tr}(\gamma^{\Gamma'})^{-1}YY' = \text{tr} \gamma \gamma_{\Delta} \gamma_{\Delta}'YY' = \text{tr}(\gamma \gamma_{\Delta} \gamma_{\Delta}'YY').
\]

If we put \( Z = \gamma \gamma_{\Delta} \gamma_{\Delta}'Y \), it is easy to check that the probability density of \( Z \) is

\[
(16.9.1) \quad \sqrt{2\pi}^{-p/2} \exp - \frac{1}{2} \text{tr} Z Z'.
\]

For all nonnull nonstochastic \( \gamma(p \times 1) \) consider now the simultaneous statement that

\[
(16.9.2) \quad g_1^2 \leq a'Z Z' \gamma / a' \gamma \leq g_2^2 \quad \text{or} \quad g_1^2 \leq a'(\gamma \gamma_{\Delta} \gamma_{\Delta}'YY' \gamma_{\Delta}) \gamma / a' \gamma \leq g_2^2.
\]

This statement, for a given \( Z \) and \( g_1^2 \) and \( g_2^2 \) is precisely equivalent to the statement that

\[
g_1^2 \leq \inf_{\gamma} \frac{a'Z Z' \gamma}{a' \gamma} \leq \sup_{\gamma} \frac{a'Z Z' \gamma}{a' \gamma} \leq g_2^2,
\]

or that

\[
(16.9.3) \quad g_1^2 \leq c_1 \leq c_p \leq g_2^2,
\]

where \( c_1 \) and \( c_p \) are the smallest and largest characteristic roots of the matrix \( ZZ' \), both, a.e., positive. The relevant distribution on the null hypothesis, i.e., when the true population matrix is \( \Sigma \), is known from Chapter 9 and we now put

\[
(16.9.4) \quad g_1^2 = c_{1d}(p, n) \text{ and } g_2^2 = c_{2d}(p, n),
\]

where \( c_{1d}(p, n) \) and \( c_{2d}(p, n) \) are constants taken over from (7.4.3). If we now tie up (16.9.2), (16.9.3) and (16.9.4) we have, with a confidence coefficient \( 1-\alpha \), the set of multiple or simultaneous confidence interval statements for all nonnull \( \gamma \) and all permissible values of the unknown parameters \( \Gamma \) and \( \Delta \):

\[
(16.9.5) \quad a' \gamma c_{1d}(p, n) \leq a'(\gamma \gamma_{\Delta} \gamma_{\Delta}'YY' \gamma_{\Delta}) \gamma \leq a' \gamma c_{2d}(p, n),
\]

or, remembering that \( n\Sigma = YY' \),

\[
a' \gamma c_{1d}(p, n) \leq a'(\gamma \gamma_{\Delta} \gamma_{\Delta}'n\Sigma \gamma_{\Delta}) \gamma \leq a' \gamma c_{2d}(p, n).
\]

The shortness of the confidence bounds (16.9.5) is tied up with the power of the test (7.4.3), which has been already discussed in section 8.
Far more meaningful confidence bounds than (16.9.5) can be obtained in the following way, starting from (16.9.5). As before denoting the characteristic roots of a (square) matrix $M$ by $c(M)$ and the largest and smallest roots of $M$ (if $M$ is at least p.s.d.) by $c_{\max}(M)$ and $c_{\min}(M)$, and remembering that $\Delta = \gamma^{-1/2}$ and finally using (A.2.5) we can rewrite (16.9.5) as

\begin{equation}
\frac{1}{n}c_{1a}(p,n) \leq c(D_{1/\gamma}^{1/\gamma} (\text{S(D)}_{1/\gamma}^{1/\gamma})^*) \leq \frac{1}{n}c_{2a}(p,n).
\end{equation}

Now using (A.1.18) we note that

\begin{equation}
c(D_{1/\gamma}^{1/\gamma} (\text{S(D)}_{1/\gamma}^{1/\gamma})^*) = c(\text{S(D)}_{1/\gamma}^{1/\gamma}) = c(\Sigma^{-1}),
\end{equation}

and obtain with a confidence coefficient $1-\alpha$, the confidence bounds

\begin{equation}
\frac{1}{n} c_{1a}(p,n) \leq c(\Sigma^{-1}_{\alpha}) \leq \frac{1}{n} c_{2a}(p,n), \text{ or } n c_{1a}(p,n) \geq c(\Sigma^{-1}_{\alpha}) \geq n c_{2a}(p,n).
\end{equation}

We now recall (A.1.22) and deduce from it that

\begin{equation}
c_{\min}(B^{-1}) c_{\min}(AB) \leq c(A) \leq c_{\max}(S^{-1}) c_{\max}(AB).
\end{equation}

By using (16.9.9) it is easy to see that the statement (16.9.8) \(\implies\) the following

\begin{equation}
nc_{1a}(p,n)c_{\max}(S) \geq c(\Sigma) \geq nc_{2a}(p,n)c_{\min}(S).
\end{equation}

We now use the following result of set-theoretic logic, namely that

\begin{equation}
\text{"If } E_1, \text{ then } E_2" \iff \text{"E}_2 \text{ is a necessary condition for } E_1" \iff E_1 \subseteq E_2 \implies P(E_1) \leq P(E_2),
\end{equation}

to observe that if the probability of (16.9.8) is $1-\alpha$, the probability of (16.9.10) is $1-\alpha$. Thus (16.9.10) is a set of simultaneous confidence bounds with probability $1-\alpha$. Also using (A.1.21) we observe that (16.9.8) \(\implies\) the following

\begin{equation}
\sum nc_{1a}(p,n) \sum t \geq c(\Sigma) \geq \sum nc_{2a}(p,n) \sum t \geq c(\Sigma), \text{ (t = 1, 2, ..., p)},
\end{equation}

which, by using (16.9.11), is thus another set of simultaneous confidence bounds on $c(\Sigma)$'s.

16.10. Confidence bounds connected with $\Sigma_1$ and $\Sigma_2$ from $N(\Sigma_1, \Sigma_1)$ and $N(\Sigma_2, \Sigma_2)$. Let us start from $Y_n(p \times n_1) \sim N(0, \Sigma_1)$ (h = 1, 2), where we assume that $p \leq n_1, n_2$, and $\Sigma_1$
and \( \Sigma_2 \) are both p.d. so that the characteristic roots of \( \Sigma_2^{-1} \) are all positive and those of \( Y_1^{-1}(Y_2^{-1})^{-1} \), i.e., of \( (n_1/n_2)S_1S_2^{-1} \) are, a.e., all positive. We recall that there exists a nonsingular \( \mu(p \times p) \) such that \( \Sigma_1 = \mu \Sigma_1 \mu' \) and \( \Sigma_2 = \mu \Sigma_2 \mu' \), where the \( \gamma \)'s are the characteristic roots of \( \Sigma_2^{-1} \). If these roots are distinct, then by a convention, say taking all the elements of the first row of \( \mu \) to be positive, the transformation could be made one-to-one. Noting that the number of independent elements on both sides is the same we shall work in terms of \( \mu \) and the \( \gamma \)'s instead of \( \Sigma_1 \) and \( \Sigma_2 \). (As in Section (16.9) we put \( \Delta = \gamma^{-1/2} \).) Except for the factor \((-1/2)\) the argument under the exponential in the probability density of \( Y_1 \) and \( Y_2 \) can be written as

\[
(16.10.1) \quad \text{tr} \left[ (\mu \Sigma_1^{-1} \gamma)^{-1} Y_1^{-1} (\mu \Sigma_1^{-1} \gamma)^{-1} Y_2^{-1} \right] = \text{tr} \left[ (\mu \Sigma_1^{-1} \gamma)^{-1} Y_1^{-1} (\mu \Sigma_1^{-1} \gamma)^{-1} Y_2^{-1} \right].
\]

If we now put \( Z_1 = \mu^{-1} Y_1 \) and \( Z_2 = \mu^{-1} Y_2 \), it is easy to check that the probability density of \( Z_1 \) and \( Z_2 \) is

\[
(2\pi)^{-\frac{(p(n_1+n_2)/2)}{2}} \exp \left[ -\frac{1}{2} \text{tr} \left( Z_1Z_1^* + Z_2Z_2^* \right) \right].
\]

For all nonnull nonstochastic \( a(p \times 1) \) consider the set of statements

\[
g_1^2 \leq a'Z_1Z_1a / a'Z_2Z_2a \leq g_2^2 \quad \text{or}
\]

\[
(16.10.2) \quad g_1^2 \leq a'(D_{\Delta}^{-1} Y_1)(D_{\Delta}^{-1} Y_1)' a / a'(-Y_2)(-Y_2)' a \leq g_2^2 \quad \text{or}
\]

\[
\frac{n_2}{n_1} g_1^2 \leq a'(D_{\Delta}^{-1} S_1^{-1} S_2^{-1} D_{\Delta})a / a'(-Y_2)(-Y_2)' a \leq \frac{n_2}{n_1} g_2^2.
\]

For given \( Z_1, Z_2, g_1^2 \) and \( g_2^2 \), this statement is precisely equivalent to the statement that

\[
g_1^2 = \inf_{a} \frac{a'Z_1Z_1a}{a'Z_2Z_2a} \leq \sup_{a} \frac{a'Z_1Z_1a}{a'Z_2Z_2a} = g_2^2 \quad \text{or}
\]

\[
(16.10.3) \quad g_1^2 \leq c_1 \leq c_2 \leq g_2^2
\]

where \( c_1 \) and \( c_2 \) are the smallest and largest characteristic roots of the matrix

\((Z_1Z_1')(Z_2Z_2')^{-1} \), both, a.e., positive. The relevant distribution on the null hypothesis, i.e., when \( \Sigma_1 = \Sigma_2 \), is known from Chapter 9 and we now put
\[ c_1^2 = c_{1a}(p, n_1, n_2) \text{ and } c_2^2 = c_{2a}(p, n_1, n_2), \]

where \( c_{1a}(p, n_1, n_2) \) and \( c_{2a}(p, n_1, n_2) \) are constants taken over from (7.4.6).

If we now tie up (16.10.2) and (16.10.3) and put \( s_1^{-1} = b' \), we have (with a confidence coefficient \( 1-\alpha \)), the set of simultaneous confidence interval statements for all nonnull \( b \) and all permissible values of the unknown parameters \( \mu \) and \( \gamma \):

\[ \frac{n_2}{n_1} c_{1a}(p, n_1, n_2) b' S_0 \leq b' (\mu D_0)^{-1} S_1^{-1} D_0 (\mu' D_0 b') \leq \frac{n_2}{n_1} c_{2a}(p, n_1, n_2) b' S_2 b. \]

The shortness of the confidence bounds (16.10.4) is tied up with the power of the test (7.4.5), which has been already discussed in section 10.

As in the previous case far more meaningful bounds than (16.10.4) can be obtained from (16.10.4) in the following way. Rewrite (16.10.4) as

\[ \frac{n_2}{n_1} c_{1a}(p, n_1, n_2) b' S_0 \leq b' (\mu D_0)^{-1} S_1^{-1} D_0 (\mu' D_0 b') \leq \frac{n_2}{n_1} c_{2a}(p, n_1, n_2) b' S_2 b, \]

which again, by using (A.2.2) and (A.1.18), rewrite as

\[ \frac{n_2}{n_1} c_{1a}(p, n_1, n_2) \geq b' (\mu D_0)^{-1} S_1^{-1} D_0 (\mu' D_0 b') \geq \frac{n_2}{n_1} c_{2a}(p, n_1, n_2). \]

Using (A.5.6) and (16.9.9) we have

\[ c_{\text{max}} S_2 (\mu')^{-1} D_0 \mu S_1^{-1} D_0 (\mu' D_0)^{-1} \geq c_{\text{max}} S_2^{-1} \]

i.e., all \( c(S_1^{-1} S_2) \geq c_{\text{min}} S_2^{-1} \)

where

\[ c_{\text{min}} = \min \left( \frac{c(S_1^{-1} S_2^{-1} D_0 \mu S_1^{-1} D_0 \mu')}{c(S_2^{-1})} \right) \]

In the same way we have

\[ c_{\text{max}} S_1^{-1} \geq c_{\text{min}} S_1^{-1} \]

Furthermore noting that

\[ c(\mu D_0 \mu^{-1}) = c(D_0 \mu^{-1}) S_1^{-1} \]

and using (A.1.22), we have
(16.10.11) \( \gamma_0(\mathbf{S}) \geq \gamma_0^{(1)}(\mathbf{S}^\dagger), \) i.e., all \( \gamma_0^{(1)}(\mathbf{S}) \), i.e., all \( \gamma_0^{(1)}(\mathbf{S}) \geq \gamma_0^{(1)}(\mathbf{S}) \).

Combining (16.10.7), (16.10.9) and (16.10.11) we have

\[
\begin{align*}
(16.10.12) \quad c_{\max}(S_2^{(1)}(\mu)^{-1}) &
\geq c_{\min}(S_2^{(1)}(\mu)^{-1})\quad (16.10.13) \quad c_{\max}(S_2^{(1)}(\mu)^{-1})
\geq c_{\min}(S_2^{(1)}(\mu)^{-1})
\end{align*}
\]

From this it is easy to check that (16.10.6) \( \rightarrow \) the following:

\[
(16.10.13) \quad \frac{n_1}{n_2} c^{-1}(p,n_1,n_2)c_{\max}(S_2^{(1)})c_{\max}(S_1) \geq c_{\max}(S_2^{(1)})
\]

which is thus by using (16.9.11) a set of simultaneous confidence bounds with a confidence coefficient \( > 1 - \alpha \). Notice that

\[
c_{\max}(S_2^{(1)}) = 1/c_{\min}(S_2^{(1)}) \quad \text{and} \quad c_{\min}(S_2^{(1)}) = 1/c_{\max}(S_2^{(1)}).
\]

Confidence bounds in terms of \( t \) could also be given as in (16.9.12), but in this case the bounds would be more complicated and do not appear to be so worthwhile as in the previous case.

16.11. Confidence bounds on regression like parameters.

(i) Some preliminary observations. We now start with a random sample of size \( n \) \( (> p+q; \ p \leq q) \) from a \((p+q)\)-variate normal population, and next reduce for the means and set

\[
\begin{pmatrix}
S_1 & S_{12} \\
S_{12} & S_{22}
\end{pmatrix}
\begin{pmatrix}
p \\
q
\end{pmatrix}
= \begin{pmatrix}
Y_1 \\
Y_2
\end{pmatrix}
\begin{pmatrix}
Y_1 \\
Y_2
\end{pmatrix}^{-1}
\]

where \( S_1, S_{22} \) and \( S_{12} \) stand respectively for the sample dispersion sub-matrices of the p-set, the q-set and that between the p-set and the q-set and where \( Y_1 \) and \( Y_2 \) have the p.d.f.

\[
(16.11.1) \quad \text{const. exp. } \left[ -\frac{1}{2} \text{tr} \left( \begin{pmatrix}
S_{11} & S_{12} \\
S_{12} & S_{22}
\end{pmatrix}^{-1}
\begin{pmatrix}
Y_1 \\
Y_2
\end{pmatrix}
\begin{pmatrix}
Y_1 \\
Y_2
\end{pmatrix}
\right] \right].
\]
We next recall that there exist non-singular \( \mu_1(p \times p) \) and \( \mu_2(q \times q) \) such that

\[
\Sigma_{11}(p \times p) = \mu_1(p \times p) \Sigma_{11}(p \times p) \mu_1(p \times p), \quad \Sigma_{22}(q \times q) = \mu_2(q \times q) \Sigma_{22}(q \times q) \mu_2(q \times q) \quad \text{and}
\]

\[
\Sigma_{12}(p \times q) = \mu_1(p \times p) D_{\gamma} \mu_2(q \times q),
\]

where \( D_{\gamma} \) stands for a diagonal matrix whose diagonal elements are the positive square roots the characteristic roots (all non-negative) of the matrix \( \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12} \) (i.e., the squares of the population canonical correlations between the p-set and the q-set). Now, denoting by \( I(m) \) an \( m \times m \) identity matrix, we have

\[
(16.11.3) \quad \left( \begin{array}{cc}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{12} & \Sigma_{22}
\end{array} \right)^{-1} = \left( \begin{array}{cc}
\mu_1 & 0 \\
0 & \mu_2
\end{array} \right) \left( \begin{array}{cc}
I(p) & (D_{\gamma} 0) \\
(D_{\gamma}) & I(q)
\end{array} \right)^{-1} \left( \begin{array}{cc}
\mu_1 & 0 \\
0 & \mu_2
\end{array} \right)^{-1}
\]

\[
= \left( \begin{array}{cc}
\mu_1 & 0 \\
0 & \mu_2
\end{array} \right) \left( \begin{array}{cc}
D/1-\gamma & (D_{\gamma} 0) \\
0 & I(q)
\end{array} \right)^{-1} \left( \begin{array}{cc}
\mu_1 & 0 \\
0 & \mu_2
\end{array} \right)
\]

\[
= \left( \begin{array}{cc}
\mu_1^{-1} & 0 \\
0 & \mu_2^{-1}
\end{array} \right) \left( \begin{array}{cc}
D/1-\gamma & 0 \\
-(D/\gamma/1-\gamma) & I(q)
\end{array} \right)^{-1} \left( \begin{array}{cc}
\mu_1^{-1} & 0 \\
0 & \mu_2^{-1}
\end{array} \right).
\]

Going back to (16.11.1) and using (A.1.5) we have now

\[
(16.11.4) \quad \text{tr} \left( \begin{array}{cc}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{12} & \Sigma_{22}
\end{array} \right)^{-1} \left( \begin{array}{cc}
Y_1 \\
Y_2
\end{array} \right) \left( \begin{array}{cc}
Y_1 \\
Y_2
\end{array} \right)
\]

\[
= \text{tr} \left( \begin{array}{cc}
D/1-\gamma & -(D/\gamma/1-\gamma) \\
0 & I(q)
\end{array} \right) \left( \begin{array}{cc}
\mu_1^{-1} & 0 \\
0 & \mu_2^{-1}
\end{array} \right) \left( \begin{array}{cc}
Y_1 \\
Y_2
\end{array} \right) \left( \begin{array}{cc}
Y_1 \\
Y_2
\end{array} \right)
\]

\[
x \left( \begin{array}{cc}
\mu_1^{-1} & 0 \\
0 & \mu_2^{-1}
\end{array} \right) \left( \begin{array}{cc}
D/1-\gamma & 0 \\
-(D/\gamma/1-\gamma) & I(q)
\end{array} \right) = \text{tr} \left( \begin{array}{cc}
Z_1 \\
Z_2
\end{array} \right) \left( \begin{array}{cc}
Z_1 \\
Z_2
\end{array} \right),
\]
where
\[(16.11.5) \quad z_1 = D^{1/1-\gamma} \mu_1^{-1} y_1 - (D^{1/1-\gamma} : 0) \mu_2^{-1} y_2, \quad z_2 = \mu_2^{-1} y_2.\]

Thus it is easy to check from (16.11.1), (16.11.4) and (16.11.5) that \((z_1, z_2)\) have the p.d.f.

\[(16.11.6) \quad \text{const. exp.} \left\{ -\frac{1}{2} \text{tr} \left( \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \begin{pmatrix} z_1 & z_2 \end{pmatrix} \right) \right\}.\]

Consider now, for any two arbitrary non-null vectors \(a_1(p \times 1)\) and \(a_2(q \times 1)\) and for a fixed positive \(g^2\), the statement

\[(16.11.7) \quad \frac{(a_1^T z_1 a_1)(a_2^T z_2 a_2)}{(a_1^T z_1 a_1)(a_2^T z_2 a_2)} \leq g^2,\]

\[(16.11.8) \quad \frac{\sum_{a_1} a_1^T \left( D^{1/1-\gamma} \mu_1^{-1} y_1 - (D^{1/1-\gamma} : 0) \mu_2^{-1} y_2 \right) a_1 - \frac{1}{2} a_2 J}{(a_2^T \mu_2^{-1} y_2 a_2)(a_1^T \mu_1^{-1} y_1 a_1)} \leq g^2,\]

where

\[(16.11.9) \quad \xi = D^{1/1-\gamma} \mu_1^{-1} y_1 - (D^{1/1-\gamma} : 0) \mu_2^{-1} y_2.\]

Now putting

\[(16.11.10) \quad b_1(1 \times p) = a_1^{T} D^{1/1-\gamma} \mu_1^{-1} \quad \text{and} \quad b_2(1 \times q) = a_2^{T} \mu_2^{-1},\]

we check that (16,11,8) reduces to

\[\frac{\sum_{b_1} b_1(1) y_1^T y_1 b_1 - 3 y_2^T y_2 b_2}{(b_2^{T} y_2 b_2) \sum_{b_1} b_1(1) y_1^T y_1 b_1} \leq g^2\]

or

\[(16.11.11) \quad \frac{\sum_{b_1} b_1(1) y_1^T y_1 b_1 - 3 y_2^T y_2 b_2}{(b_2^{T} y_2 b_2) \sum_{b_1} b_1(1) y_1^T y_1 b_1} \leq g^2,\]

where

\[(16.11.12) \quad \beta(p \times q) = \mu_1^{T} D^{1/1-\gamma} : 0) \mu_2^{-1} = z_1 z_2^{-1}.\]

\(\beta\) defined by (16,11,12) can be appropriately called the matrix of population regression of the p-set on the q-set and it is the only set of population parameters that occurs in the statement (16,11,11).
(ii) Confidence bounds on the regression matrix $\beta$. It is well known that
the statement (16.11.11), for all arbitrary non-null $b_1$ and $b_2$, is exactly equivalent to

$$
(16.11.13) \quad \text{all } c_i' \text{ s} \leq \varepsilon^2 \text{ or } c_p \leq \varepsilon^2,
$$

where $c_i$'s $(i = 1, 2, \ldots, p; 0 \leq c_1 \leq \ldots \leq c_p \leq 1)$ are the roots of the determinantal equation in $c$:

$$
(16.11.14) \quad \left| c(S_{11} - S_{12} \beta' - S_{12} \beta S_{22} \beta'') - (S_{12} - S_{22})S_{22}^{-1}(S_{12} - S_{22} \beta') \right| = 0.
$$

Now put $e = c/l-c$, so that we have from (16.11.14), the determinantal equation in $e$

$$
(16.11.15) \quad \left| e(S_{11} - S_{12} S_{22}^{-1} S_{12}) - (S_{12} S_{22}^{-1} S_{22} S_{22}^{-1} S_{12} - \beta) \right| = 0.
$$

Notice that the statement (16.11.12) can now be replaced by the statement that the
largest characteristic root $e_p \leq \varepsilon^2/(1-\varepsilon)^2$, i.e.,

$$
(16.11.16) \quad \text{all } c_i \left| (S_{11} - S_{12} S_{22}^{-1} S_{12})^{-1}(S_{12} S_{22} S_{22}^{-1} S_{12} - \beta) \right| \leq \varepsilon^2/(1-\varepsilon)^2,
$$

where

$$
(16.9.17) \quad B(p \times q) = S_{12} S_{22}^{-1},
$$

which may be appropriately called the matrix of sample regression of the $p$-set on the $q$-set.

We note that (16.11.16) $\iff$ (16.11.13) $\iff$ (16.11.7), so that $c_p$ is
the largest characteristic root of the matrix $(Z_1 Z_1')^{-1}(Z_1 Z_1')(Z_2 Z_2')^{-1}(Z_2 Z_2')$, where
$(Z_1, Z_2)$ have the p.d.f. (16.11.6). The joint distribution of these central $c_i$'s,
and also of the largest root $c_p$ being known, all that we have to do to make
(16.11.16), i.e., (16.11.13), i.e., (16.11.7), a simultaneous confidence statement
with a joint coefficient $1-\alpha$ is to choose $\varepsilon^2 = c_{\alpha}(p, q, n-1)$ where the quantity on the
right side is defined by

$$
(16.11.18) \quad \text{P(central } c_p \geq c_{\alpha}(p, q, n-1) = \alpha.
$$

Substituting now $c_{\alpha}(p, q, n-1)$ (to be sometimes denoted more simply by $c_{\alpha}$) for $\varepsilon^2$ in (16.11.16), we have a simultaneous confidence statement with a joint confidence
coefficient $1-\alpha$. 
Now applying (A.1.15), (A.1.22) and (16.11.11) (in the same manner as in the previous sections), we have from (16.11.16), now with a joint confidence coefficient \( \geq 1-\alpha \), the following simultaneous confidence statement

\[
(16.11.19) \quad \text{all } c \sqrt{\frac{(B-B')(B'-B')}{-\gamma}} \leq \frac{c}{1-c}\frac{c_{\text{max}}(S_{11}^{-1} S_{12}^2 S_{12}^{-1} S_{12})}{c_{\text{max}}(S_{22}^{-1})}.
\]

Now note that \( c_{\text{max}}(S_{22}^{-1}) = \frac{1}{c_{\text{min}}(S_{22})} \), \( c_{\text{max}}(S_{11}^{-1} S_{12}^2 S_{12}^{-1} S_{12}) \)

\[
= c_{\text{max}}(S_{11})(1 - S_{11}^{-1} S_{12}^{-1} S_{12}^{-1} S_{12}) \quad \text{and} \quad c_{\text{max}}(1 - S_{11}^{-1} S_{12}^{-1} S_{12}^{-1} S_{12}) = 1 - c_{\text{min}}(S_{11}^{-1} S_{12}^{-1} S_{12}).
\]

Using these, we check that (16.11.19) can be replaced by the following (with a confidence coefficient \( \geq 1-\alpha \)):

\[
(16.11.19a) \quad \text{all } c \sqrt{\frac{(B-B')(B'-B')}{-\gamma}} \leq \frac{c}{1-c}\frac{c_{\text{max}}(S_{11}^{-1} S_{12}^2 S_{12}^{-1} S_{12})}{c_{\text{max}}(S_{11})/c_{\text{min}}(S_{22})}.
\]

We next recall the following two well-known results (A.2.5) and (A.2.7) which we remember for convenience as

\[
(16.11.20) \quad \text{all } c(M) \leq \gamma \quad (\text{for a } p \times p \text{ real matrix } M \text{ with real roots}) \quad \iff \quad d_1(1 \times p)^H(p \times p)d_1(p \times 1) \leq \gamma \quad (\text{for all arbitrary unit vectors } d_1(p \times 1) \text{ and } d_2) \quad \text{ and } \quad (16.11.21) \quad x'(1 \times q)x(q \times 1) \leq \gamma(>0) \quad \iff \quad x'(1 \times q)d_2(q \times 1) \leq \sqrt{\gamma} \quad (\text{for all arbitrary unit vectors } d_2(q \times 1)).
\]

Applying (16.11.20) and (16.11.21) to (16.11.19a) we have (with a joint confidence coefficient \( \geq 1-\alpha \)) the following simultaneous confidence statement (for all arbitrary unit vectors \( d_1(p \times 1) \) and \( d_2(q \times 1) \)),

\[
(16.11.22) \quad d_1'(E - B)d_2 \leq \sqrt{\text{right side of } (16.11.19a)} \sqrt{\gamma/2},
\]

or ultimately

\[
(16.11.23) \quad d_1'Bd_2 - \sqrt{\gamma} \leq d_1'Bd_2 \leq d_1'Bd_2 + \sqrt{\gamma},
\]

where

\[
(16.11.24) \quad \gamma = \frac{c}{1-c}\frac{c_{\text{max}}(S_{11}^{-1} S_{12}^2 S_{12}^{-1} S_{12})}{c_{\text{max}}(S_{11})/c_{\text{min}}(S_{22})}. \gamma.
\]

A set of simultaneous confidence bounds on just the elements \( B_{ij} \) of the \( B \)-matrix would be a subset of the bounds on the total set \( d_1'Bd_2 \). It is worthwhile to check
that if $n = q = 1$, (16.11.23) reduces, as it should, to (15.4.4). Also if $p = 1$, we should have another special case of (16.11.23) giving a set of simultaneous confidence bounds on all linear functions of the partial regressions of one variate on several others. Thus, in several ways, (16.11.23) seems to be an appropriate generalization of (15.4.4).
A.1. Some preliminary results in matrix theory.

(A.1.1) Given four matrices $A(p \times p)$, $B(p \times q)$, $C(q \times p)$ and $D(q \times q)$, if $D$ is non-singular, then

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D| \begin{vmatrix} A - BD^{-1}C \end{vmatrix}.$$

Proof.\[
\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} I(p) & O(p \times q) \\ D^{-1}C & I(q) \end{vmatrix} = \begin{vmatrix} A - BD^{-1}C & B \\ O & D \end{vmatrix} = |D| \begin{vmatrix} A - BD^{-1}C \end{vmatrix}.
\]

(A 1.2) $\sqrt[\gamma]{A(p \times q)B(q \times s)} \leq \min \sqrt[\gamma]{r(A), r(B)}$, where $\min(x, y)$ denotes the lesser of two real numbers $x$ and $y$.

(A.1.3) $\sqrt[\gamma]{A(p \times q)} = \sqrt[\gamma]{B(p \times p)A(p \times q)} = \sqrt[\gamma]{A(p \times q)C(q \times q)}$, if $B$ and $C$ are non-singular.

(A.1.4) $\sqrt[\gamma]{A(p \times q)} = \sqrt[\gamma]{A'(q \times q)} = \sqrt[\gamma]{A(p \times q)A'(q \times q)}$.

(A.1.5) $\text{tr} \sqrt[\gamma]{A(p \times q)B(q \times p)} = \text{tr} \sqrt[\gamma]{B(q \times p)A(p \times q)}$.

Proof. If $A = (a_{ij})$ and $B = (b_{ij})$, then by the definition of trace we have

$$\text{tr}(AB) = \sum_{i=1}^{p} \sum_{j=1}^{q} a_{ij}b_{ji} = \sum_{j=1}^{q} \sum_{i=1}^{p} b_{ji}a_{ij} = \text{tr}(BA).$$

(A.1.6) $\sqrt[\gamma]{A(p \times q)} = \sqrt[\gamma]{A(p \times q)B(q \times t)} = \sqrt[\gamma]{C(s \times p)A(p \times q)}$, if $q \leq t$, $p \leq s$ and $B$ and $C$ are respectively of ranks $q$ and $p$.

Proof. Using (A.1.2) - (A.1.3) we have

$$\sqrt[\gamma]{A(p \times q)} = \sqrt[\gamma]{A(p \times q)B(q \times t)B'(t \times q)} \leq \min \sqrt[\gamma]{r(AB), r(B')} \leq r(AB),$$

i.e., $\leq r(AB)$. But $r(AB) \leq r(A)$, whence $r(A) = r(AB)$. Likewise starting with $A'C'$ and noting that $r(CA) = r(A'C')$, we should have, in an exactly similar manner, $r(CA) = r(A)$, which completes the proof of (A.1.6).

(A.1.7) If $L_1(p \times n)$ $(p < n)$ is subject to $L_1L_1' = I(p)$, there exists an $L_2(n \times p)$ such that $\begin{pmatrix} L_1 \\ L_2 \end{pmatrix}$ is $\bot$. $L_2$ will be called an arbitrary completion
of $I_n$.

(A.1.8) If $M(p \times p)$ is symmetric and at least p.s.d. of rank $r(\leq p)$, then, out of
the $p$ $c(M)'s$ (i.e., roots of the determinantal equation in $c$: $|M - cI| = 0$), $r$ are
positive and the rest, $p - r$ in number, are zero. If $r = p$, the number of non-zero
roots will of course be $p$.

(A.1.9) If $M_1(p \times p)$ is symmetric and at least p.s.d. of rank $r(\leq p)$ and
$M_2(p \times p)$ is symmetric and p.d., there are exactly $r$ positive roots of the following equation
in $c$: $|M_1 - cM_2| = 0$, the rest, $p - r$ in number, being 0. If $r = p$, the number of
positive roots will of course be $p$.

(A.1.10) $X(p \times n)X'(n \times p)$ will be symmetric and at least p.s.d. of the same rank
as $X$ or $X'$, the common rank $r$ being $\leq \min(p, n)$, where the symbol (which will be
frequently used later) denotes the lesser of $p$ and $n$. It is easy to see that if
$p \leq n$ and $X$ is of rank $p$, then $XX'$ is p.d.

(A.1.11) If $A(q \times q)$ is symmetric p.d., $B(p \times q)A(q \times q)B'(q \times p)$ is symmetric
and at least p.s.d. of the same rank as $B$.

Proof. Since $A$ is symmetric p.d., there exists, by (A.3.9), a non-singular $\tilde{T}(q \times q)$
such that $A = \tilde{T}\tilde{T}'$. Hence $BAB' = (B\tilde{T})(B\tilde{T})'$, which, by (A.1.10), is symmetric and at
least p.s.d. of the same rank as $B\tilde{T}$. But $B\tilde{T}$ is of the same rank as $B$, since $\tilde{T}$ is
non-singular, whence the theorem follows.

(A.1.12) If $A(p \times p)$ is symmetric and at least p.s.d. of rank $r \leq p$ and $B(p \times p)$
is non-singular, $BAB'$ is symmetric and at least p.s.d. of rank $r$.

Proof. If $A$ is symmetric and at least p.s.d. of rank $r$, then by (A.3.9) there exist
a non-singular $\tilde{T}_1(r \times r)$ and a $\tilde{T}_2(p-r \times r)$ such that $A = \begin{pmatrix} \tilde{T}_1 \\ \tilde{T}_2 \end{pmatrix}$.

Therefore, $BAB' = \int B \begin{pmatrix} \tilde{T}_1 \\ \tilde{T}_2 \end{pmatrix} \int B \begin{pmatrix} \tilde{T}_1 \\ \tilde{T}_2 \end{pmatrix} \int'$, which, by (A.1.10), is symmetric and at
least p.s.d. of the same rank as $B \begin{pmatrix} \tilde{T}_1 \\ \tilde{T}_2 \end{pmatrix}$. But, since $B$ is non-singular and
\[
\begin{pmatrix}
\hat{\mu}_1 \\
\hat{\mu}_2 \\
\end{pmatrix}
\] is obviously of rank \( r \), therefore, \( B \begin{pmatrix}
\hat{\mu}_1 \\
\hat{\mu}_2 \\
\end{pmatrix} \) is of rank \( r \) and thus \( BAB' \) is of rank \( r \).

(A.1.13) If \( M_1(p \times p) \) is symmetric and at least p.s.d. of rank \( r \leq p \) and \( M_2(p \times p) \) is symmetric p.d., then (i) all the roots of the equation in \( c: |M_1 - cM_2| = 0 \) are zero if and only if \( M_1 = 0 \), and (ii) all the roots are unity if and only if \( M_1 = M_2 \).

**Proof.** Part (i) of (A.1.13) is a direct consequence of (A.1.9). To prove part (ii), put \( c = 1 - e \). We have then the equation in \( e: |(M_1 - M_2) - eM_2| = 0 \) whence it follows that all roots of the equation are zero, i.e., all roots of the equation in \( c \) are unity, if and only if \( M_1 - M_2 = 0 \), i.e., \( M_1 = M_2 \), which proves part (ii) of (A.1.13).

(A.1.14) If \( M \) is a \((p + q) \times (p + q)\) symmetric matrix given by, say,
\[
M = \begin{pmatrix}
M_{11} & M_{12} \\
M_{12}^\prime & M_{22} \\
\end{pmatrix},
\]
and if \( M_{22} \) is non-singular, then (i) \( M_{11} - M_{12}M_{22}^{-1}M_{12}^\prime \) is symmetric, and (ii) is also of rank \( r-q \), where \( q \) is (evidently) the rank of \( M_{22} \) and \( r \) denotes the rank of \( M \) (evidently satisfying: \( q \leq r \leq p + q \)).

**Proof.** Part (i) is obvious if we remember that \( M_{11}, M_{22} \) (and thus \( M_{22}^{-1} \)) are symmetric and so also \( M_{12}M_{22}^{-1}M_{12}^\prime \). For part (ii) we first observe that the rank of \( M \) would be unaltered if it were pre-multiplied and/or post-multiplied by two conformable non-singular matrices. Post-multiply \( M \) by the conformable non-singular matrix (of rank \( p + q \)):
\[
\begin{pmatrix}
I & 0 \\
-M_{12}M_{22}^{-1}M_{12}^\prime & I \\
\end{pmatrix}
\]
and premultiply by the transpose of this matrix. Then we have
\[
\text{rank of } M = \text{rank of } \begin{pmatrix}
I(p) & -M_{12}M_{12}^{-1} \\
-M_{12}M_{12}^{-1} & I(q)
\end{pmatrix} \begin{pmatrix}
M_{11} & M_{12} \\
M_{12} & M_{22}
\end{pmatrix} \begin{pmatrix}
I(p) & 0 \\
-M_{12}M_{12}^{-1} & I(q)
\end{pmatrix}, \text{ i.e.,}
\]
\[
\text{rank of } \begin{pmatrix}
M_{11} - M_{12}M_{12}^{-1}M_{12} & 0 \\
0 & M_{22}
\end{pmatrix}. \text{ But the rank of this last matrix is evidently the same as that of } M_{22} \text{ (which is } q) \text{ plus that of } (M_{11} - M_{12}M_{12}^{-1}M_{12}). \text{ This proves part (ii) of (A.1.14).}
\]

(A.1.15) If \( M \) has the same structure as in (A.1.14) and is in addition at least p.s.d. of rank \( r \) (\( q \leq r \leq p+q \)), then \( M_{11} - M_{12}M_{12}^{-1}M_{12} \) is also at least p.s.d. of rank \( r-q \).

Proof. Since \( M \) is symmetric and at least p.s.d. of rank \( r \) (\( q \leq r \leq p+q \)), pre-multiplying and post-multiplying it by the same conformable non-singular matrices as in the proof of (A.1.14) and using next (A.1.12), we observe that
\[
\begin{pmatrix}
M_{11} - M_{12}M_{12}^{-1}M_{12} & 0 \\
0 & M_{22}
\end{pmatrix}^p \text{ is at least p.s.d. of rank } r. \text{ Hence } M_{11} - M_{12}M_{12}^{-1}M_{12} \text{ is evidently at least p.s.d. and since (A.1.14) shows that it is of rank } r-q, \text{ the theorem (A.1.15) follows.}
\]

(A.1.16) If \( M = \begin{pmatrix}
M_{11} & M_{12} \\
M_{12} & M_{22}
\end{pmatrix} \) \( p \text{ is symmetric and at least p.s.d. of rank } r \) (\( q \leq r \leq p+q \)), and if \( p \leq q \) and \( M_{11} \text{ and } M_{22} \text{ are both non-singular (i.e., in this situation both p.d. of ranks } p \text{ and } q \text{ respectively) and if } s \text{ denotes the rank of } M_{12}^p(p \times q) \text{ (evidently } s \leq p \leq q) \text{ then the } p \text{ roots of the } p^{\text{th}} \text{ degree equation in } c: \left| cM_{11} - M_{12}M_{22}^{-1}M_{12} \right| = 0 \text{ has the following properties: (i) } 0 \leq \text{ all } c\text{'s } \leq 1, \text{ (ii) out of the } p \text{ c's, } r-q \text{ are } \neq 1 \text{ and the rest, } p - (r-q) \text{ (} = p + q - r < p \text{) in number, are } 1; \text{ (iii) also out of the } p \text{ c's, } s \leq p \text{ are } \neq 0 \text{ and the rest, } p-s \leq p \text{) \}}
are = 0.

Proof. We note that since $M_{22}$ and hence $M_{22}^{-1}$ is p.d. of rank $q$ and $M_{12}$ is of rank $s$, therefore, from (A.1.11), $M_{12}M_{22}^{-1}M_{12}$ is symmetric and at least p.s.d. of rank $s$ ($\leq p \leq q$). Also $M_{11}(p \times p)$ is supposed to be p.d. Hence by (A.3.3), out of the $p$ roots of the equation in $c$: 
\[
|cM_{11} - M_{12}M_{22}^{-1}M_{12}| = 0, \ s \ are \ > \ 0 \ and \ the \ rest, \ p-s \ in \ number, \ are = 0. \ \text{Next, putting} \ c = 1 - e, \ \text{we have the equation in} \ e:
\]
\[
|cM_{11} - (M_{11} - M_{12}M_{22}^{-1}M_{12})| = 0. \ \text{But} \ M_{11} \ \text{is symmetric p.d. and, by (A.1.14) and (A.1.15),} \ M_{11} - M_{12}M_{22}^{-1}M_{12} \ \text{is symmetric and at least p.s.d. of rank} \ r-q (\leq p). \ \text{Hence, out of the} \ p \ \text{roots of the equation in} \ e, \ r-q \ \text{are} > 0 \ \text{and the rest,} \ p - (r-q) (= p+q-r \leq p) \ \text{in number, are} = 0. \ \text{Since} \ c = 1 - e, \ \text{this means that, out of the} \ p \ \text{roots of the equations in} \ c, \ r - q \ \text{are} < 1 \ \text{and the rest,} \ p + q - r \ \text{in number, are} = 1. \ \text{This completes the proof of (A.1.16).}
\]

(A.1.17) With the same set-up as in (A.1.16), the roots of the equation in $c$:
\[
|cM_{11} - M_{12}M_{22}^{-1}M_{12}| = 0 \ \text{are all zero if and only if the rank of} \ M_{12} \ \text{is zero, i.e.,} \ M_{12} \ \text{is the null matrix. This is a direct consequence of (A.1.16). With regard to theorems (A.1.16) and (A.1.17) we observe that in statistical applications we shall always be considering the special case,} \ r = p + q, \ \text{that is, the case where} \ M \ \text{is symmetric p.d. In this situation we state and prove two theorems on transformations, (A.3.16) and (A.3.17).}
\]

(A.1.18) Every non-zero characteristic root of $A(p \times q)B(q \times p)$ is a (non-zero) characteristic root of $B(q \times p)A(p \times q)$ and vice versa.

Proof. If $c$ is any (non-zero) characteristic root of $AB$, we have by definition,
\[
|AB - cI| = 0 \ \text{or, by using A.1.1,}
\]
\[
\begin{bmatrix}
cI & A & p \\
B & I & q
\end{bmatrix}
\]
\[
p \ q
\]

Since $c$ is non-zero we can obviously rewrite this as
\[(A.1.18)\]
\[
\begin{vmatrix}
  cI & B \\
  A & I \\
  p & q
\end{vmatrix} = 0,
\]

or by using \((A.1.1)\) again,

\[(A.1.18.3)\] \(|BA - cI| = 0\), which proves \((A.1.18)\).

\[(A.1.19)\]
(i) If \(B(p \times p)\) is non-singular, the roots of the equation in

\[
c: \quad |A(p \times p) - cB(p \times p)| = 0
\]

are the same as the characteristic roots of \(B^{-1}\) or

of \(B^{-1}A\) and (ii) in \((A.1.16)\) the roots of the equation in

\[
c: \quad |cM_{11} - M_{12}M_{22}^{-1}M_{12}| = 0
\]
are the same as the characteristic roots of $M_{11}^{-1} M_{12} M_{22}^{-1} M_{12}$ or of $M_{12}^{-1} M_{22} M_{12}^{-1}$ (with the exception of zero roots in the case where $p < q$). The proof is obvious.

\[(A.1.20) \quad \text{tr}_t^A (p \times p) = \sum_{i_1 \neq i_2 \neq \ldots \neq i_t} c_{i_1} (A) c_{i_2} (A) \ldots c_{i_t} (A), \text{ where } \text{tr}_t^A \]

stands for the sum of all $t \times t$ minors (found by the intersection of any $t$ rows of $A$ with $t$ columns bearing the same number), and, in particular,

\[\text{tr}_1^A = \sum_{i=1}^p c_{i_1} = \sum_{i=1}^p a_{i_1} \quad \text{and} \quad \text{tr}_p^A = \prod_{i=1}^p c_{i_1} = |A| .\]

\[(A.1.18) \text{ coupled with } (A.1.20) \text{ supplies another proof of the relation:} \]
\[\text{tr}(AB) = \text{tr}(BA) \quad (\text{see } (A.1.5)).\]

\[(A.1.21) \quad \text{If } (a) \quad d_1 \leq \text{all } c \quad (AB^{-1}) \leq d_2 \quad (d_2 > 0), \text{ then } (b) \quad (d_1)^t \text{tr}_t (B) \leq \text{tr}_t (A) \leq (d_2)^t \text{tr}_t (B) \quad (t = 1, 2, \ldots, p), \text{ where } A \text{ and } B \text{ are two } p \times p \text{ p.d. matrices. Notice that } (b) \text{ is a necessary (though not a sufficient) condition for } (a). \]

**Proof.** It is easy to check that "$d_1 < \text{all } c \quad (AB^{-1})$" $\iff (A - d_1 B)$ is p.d.

\[\iff (A_t - d_1 B_t) \quad (t = 1, \ldots, p) \text{ is p.d. (where } A_t - d_1 B_t \text{ is a submatrix formed by the intersection of any } t \text{ rows of } (A - d_1 B) \text{ with } t \text{ columns bearing the same numbers}) \iff d_1 < \text{all } c \quad (A_t B_t^{-1}) \quad (t = 1, \ldots, p). \]

Now, if all $c \quad (A_t B_t^{-1}) > d_1$, one consequence is that

\[(A.1.21.1) \quad \prod_{i=1}^t c_{i_1} (A_t B_t^{-1}) > (d_1)^t, \text{ i.e., } |a_t| / |B_t| > (d_1)^t.\]

For a given $t$, summing over different possible submatrices, we have

\[(A.1.21.2) \quad \text{tr}_t^A > (d_1)^t \text{tr}_t B.\]

Using the same kind of argument for the other half of the inequality and remembering that $t = 1, 2, \ldots, p$, and combining, we have the result that

\[(A.1.21.3) \quad \text{if } d_1 < \text{all } c(AB^{-1}) < d_2, \text{ then } (d_1)^t \text{tr}_t (B) \leq \text{tr}_t (A) \leq (d_2)^t \text{tr}_t (B) \quad (t = 1, 2, \ldots, p).\]
By a slight rephrasing (which is obviously permissible here) we have the result
(A.1.21).

(A.1.22) If \( M(p \times p) \) is symmetric p.d. and \( B(p \times p) \) is symmetric and at least
p.s.d., then (i) all \( c(AB) \) are non-negative and (ii) \( c(AB) = \min \min \max \max \) where \( c(M) \) and \( c(M) \) stand respectively for the largest and smallest roots (both non-
negative) of any \( M \) which is symmetric and at least p.s.d.

Proof. By (A.3.3) there are \( \perp \) matrices \( L_A(p \times p) \) and \( L_B(p \times p) \) such that
\[
A = L_A^Dc(A)L_A^! \quad \text{and} \quad B = L_B^Dc(B)L_B^!, \quad \text{and thus} \quad AB = L_A^Dc(A)L_A^!L_B^Dc(B)L_B^!.
\]
Now using (A.1.18) (and noting that here \( p = q \), so that all characteristic roots are the same in both
products), we have the two-way relation
\[
c(A^M) = c(D_{c(A)}^MC_{c(B)}^M)^! = c(D_{c(A)}^MC_{c(B)}^!M'), \quad \text{where} \quad M \text{ stands for} \quad L_A^!L_B!.
\]
Notice that \( MM' = L_A^!L_B^!L_A^L_B^!L_A^!L_B^! = I(p) \) (since \( L_A \) and \( L_B \) are each \( \perp \)), so that \( M \) itself is \( \perp \). Also note that
\( D_{c(A)}^M \) is non-singular since \( M \), being \( \perp \), is non-singular and \( D_{c(A)}^M \) is non-singular,
because all the \( c(A) \)'s are positive. Using (A.1.18) again we find that
\[
c(AB) = c(D_{c(A)}^MD_{c(B)}^!M'/c(A))^!, \quad \text{and since} \quad D_{c(B)}^M \text{ is obviously symmetric p.d. by vir-
tue of} \quad B \text{ being p.d., we notice by using (A.1.11) that} \quad D_{c(A)}^MD_{c(B)}^!M'/c(A) \text{ is sym-
matic and at least p.s.d., and thus all} \quad c(AB) \text{ are non-negative. This proves part}
\]
(i). For part (ii), let us go back to \( D_{c(A)}^MD_{c(B)}^M \). denote by \( \lambda_i \) and \( \mu_j \) the
characteristic roots of \( A \) and \( B \), observe that here all \( \lambda_i > 0 \) and all \( \mu_j > 0 \) and
next observe that, if \( c \) is to be a characteristic root of \( AB \) (here all roots are
non-negative), there exists a set of (real) numbers \( x_1, x_2, \ldots, x_p \), not all of which
are zero, such that the following set of equations are satisfied.

\[
(A.1.22.1) \quad \sum_{j,k=1}^{p} \lambda_i m_{ij} \mu_j^{m_{kj}} x_k = c x_i \quad (i = 1, 2, \ldots, p) \quad (\text{notice that} \quad (M')_{jk} = (M)_{kj}).
\]

Remembering that here \( \lambda_i > 0 \) and \( \mu_j > 0 \) (both sets being real), dividing
by \( \lambda_i \), and squaring any member of \( (A.1.22.1) \) and summing over \( i = 1, 2, \ldots, p \), we
have
\[(A.1 \textit{??}) \quad \sum_{i,j,j',k,k'} m_{ij} m_{i'j'} \mu_{j} \mu_{j'} k_{j} k_{j'} x_{k} x_{k'} = c^2 \sum_{i=1}^{\text{P}} x_{i}^2 \lambda_{i}. \] 

Now, since \( M \) is \( \perp \), we have \( \sum_{i} m_{ij} m_{i'j'} = \delta_{jj'} \) (where \( \delta \) is the Kronecker symbol), so that \((A.1.22.2)\) reduces to

\[(A.1.22.3) \quad c^2 \sum_{i} x_{i}^2 / \lambda_{i} = \sum_{j,k,k'} \mu_{j} m_{jk} k_{j} k_{k'} x_{k} x_{k}. \]

It is easy to check that the coefficients of \( \lambda_{i} \) on the left hand side and those of \( \mu_{j} \) on the right hand side are each at least non-negative. Hence, if we replace all \( \mu_{j} \)'s by \( \mu_{\text{max}} \) and all \( \lambda_{i} \)'s by \( \lambda_{\text{max}} \) the right hand side is increased (or at least not diminished) and the left hand side is diminished (or at least not increased). We have thus

\[(A.1.22.4) \quad (c^2 / \lambda_{i}) \sum_{i \in \text{max}} x_{i}^2 \leq \mu_{\text{max}} \sum_{j,k,k'} m_{jk} k_{j} k_{k'} x_{k} x_{k} \quad \text{i.e.,} \quad \leq \mu_{\text{max}} \sum_{j} \delta_{kk'} x_{k} x_{k} \quad \text{(since } M \text{ is } \perp \text{)}, \quad \text{i.e.,} \quad \leq \mu_{\text{max}} \sum_{i} x_{i}^2. \]

Since \( \sum_{i} x_{i}^2 \) is positive, it follows that \( c^2 \leq \lambda_{\text{max}} \mu_{\text{max}} \), i.e., \( c \leq \lambda_{\text{max}} \mu_{\text{max}} \) (taking the positive square root on both sides).

Thus we have

\[(A.1.22.5) \quad \text{all } c(AB) \leq c_{\text{max}}(A) c_{\text{max}}(B). \]

Likewise, in \((A.1.22.3)\), replacing all \( \lambda_{i} \)'s by \( \lambda_{\text{min}} \) and all \( \mu_{j} \)'s by \( \mu_{\text{min}} \) and arguing in a similar manner, we have

\[(A.1.22.6) \quad c_{\text{min}}(A) c_{\text{min}}(B) \leq \text{all } c(AB). \]

Combining \((A.1.22.5)\) and \((A.1.22.6)\) we have part \((ii)\) of \((A.1.22)\).

Replacing \( A \) by a complex non-singular \( A \), \( B \) by any complex \( B \), remembering that \( AA^* \) is hermitian p.d. and \( BB^* \) is hermitian and at least p.s.d. we have the following more general theorem, proved elsewhere:

\[(A.1.23) \quad c_{\text{min}}((AA^*)c_{\text{min}}((BB^*)) \leq \text{all } c(AB) \leq c_{\text{max}}((AA^*)c_{\text{max}}((BB^*)). \]

However, this result will not be needed in the present report.

\section{A.2. Some results in quadratic forms.}

\((A.1)\) If \( A(p \times p) \) is symmetric and at least p.s.d. of rank \( r \) \((\leq p)\), then

\((i)\) \( a^t(1 \times p) A(p \times p) a(p \times 1) \) is a p.s.d. quadratic form in \( a_{i} \)'s \((i = 1, \ldots, p)\),
(ii) the stationary values of $a'\bar{A}a/a'_a$ (under variation of $a$ over all non-null $a$'s) are the characteristic roots of $\bar{A}$ (all non-negative) and (iii) in particular, the largest and smallest values of $a'\bar{A}a/a'_a$ (under variation of $a$) are the largest and smallest characteristic roots of $\bar{A}$.

Proof. Part (i) is given in all text books and need not be proved. For part (ii), putting $a'\bar{A}a/a'_a = c$ and differentiating $c$ with respect to the elements of $a$, we have the vector equation giving the stationary values of $c$: $\bar{A}_a - c_2 = 0$, whence by eliminating $a$ we have, for the stationary values of $c$, the $p^{\text{th}}$ degree determinantal equation in $c$: $|\bar{A} - cI| = 0$. The roots of this are the so-called characteristic roots of $\bar{A}$, which proves part (ii). In this case the proof of part (iii) is obvious and will not be separately discussed.

(A.2.2) If $B(p \times p)$ is symmetric p.d. and $A(p \times p)$ is symmetric and at least p.s.d. of rank $r$ (\textless; $p$), then for all non-null $a$'s (i) $a'(1 \times p)A(p \times p)a(p \times 1)/a'(1 \times p)B(p \times p)a(p \times 1)$ is non-negative, (ii) the stationary values of $a'\bar{A}a/a'\bar{B}a$ (under variation of $a$) are the roots of the determinantal equation in $c$: $|\bar{A} - cB| = 0$ and (iii) in particular, the largest and smallest values of $a'\bar{A}a/a'\bar{B}a$ are the largest and smallest roots of the determinantal equation.

Proof. Part (i) is obvious. For part (ii), putting $a'\bar{A}a/a'\bar{B}a = c$ and differentiating $c$ with respect to the elements of $a$, we have the vector equation giving the stationary values of $c$: $\bar{A}_a - cB_a = 0$, whence by eliminating $a$ we have, for the stationary values of $c$, the $p^{\text{th}}$ degree determinantal equation in $c$: $|\bar{A} - cB| = 0$,

which proves part (ii). The proof of part (iii) is now obvious.

(A.2.3) If $M = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^t & M_{22} \end{pmatrix}$ $(p \leq q)$ is symmetric p.d. (from which it follows easily that $M_{11}$ and $M_{22}$ are each symmetric p.d.), then, for all non-null $a_1(p \times 1)$ and $a_2(q \times 1)$, (i) $\sum_{1}^{q} (M_{11} a_1 a_2^t + (M_{11})_{11} a_1 a_1^t a_2 a_2) / (\sum_{1}^{q} M_{11} a_1 a_1^t)$ is non-negative,

(ii) the stationary values of this expression are just the roots of the equation in
c: \[ c^{M_{11}} - M_{12}^{-1}M_{12} \] = 0 and (iii) in particular, the largest and smallest values of the expression are the largest and smallest roots of the determinantal equation.

\textbf{Proof.} Part (i) is obvious. For part (ii), putting \( a_{11}^{M_{12}a_{2}} = a_{12} \), \( a_{11}^{M_{12}a_{1}} = a_{11} \) and \( a_{22}^{M_{12}a_{2}} = a_{22} \), and \((a_{12})^2/a_{11}a_{22} = c\) (say), and differentiating c with respect to the elements of \( a_{1} \) and \( a_{2} \), we have the vector equation giving the stationary values of c:

\[ M_{12}a_{2} - (a_{12}/a_{11})M_{11}a_{1} = 0 \text{ and } a_{1}^{M_{12}} - (a_{12}/a_{22})a_{2}^{M_{22}} = 0 \text{ or } (a_{12}/a_{22})M_{22}a_{2} - M_{12}a_{1} = 0. \]

Eliminating \( a_{2} \) and \( a_{1} \) between the two vector equations, we have, for the stationary values of c, the \( p^{th} \) degree determinantal equation in c:

\[
\begin{vmatrix}
M_{12} & (a_{12}/a_{11})M_{11} \\
(a_{12}/a_{22})M_{22} & M_{12}
\end{vmatrix} = 0 \text{ or, by using (A.1.1) and remembering that } c = a_{12}^{2}/a_{11}a_{22},
\]

(A.2.3.2) \[ c^{M_{11}} - M_{12}^{-1}M_{12} \] = 0, which proves part (ii). The proof of part (iii) is now obvious.

(A.2.4) If \( M = \begin{pmatrix}
M_{11} & M_{12} & M_{13} \\
M_{12} & M_{22} & M_{23} \\
M_{13} & M_{23} & M_{33}
\end{pmatrix} \) is symmetric p.d., then, for all \( p \leq q \) is symmetric p.d., then, for all \( p \leq q \),

\[
\begin{pmatrix}
M_{11} & M_{12} & M_{13} \\
M_{12} & M_{22} & M_{23} \\
M_{13} & M_{23} & M_{33}
\end{pmatrix}
\]

\( p \quad q \quad r \)

non-null \( a_{1} ((p+r) x 1) \) and \( a_{2} ((q+r) x 1) \), (i) \[ \sum a_{1} \begin{pmatrix} M_{12} & M_{13} \\ M_{13} & M_{33} \end{pmatrix} \begin{pmatrix} a_{2}^{-2}/a_{1} \\ M_{11} & M_{13} \end{pmatrix} \]

\[ a_{1} \cdot \sum a_{2}^{2} \begin{pmatrix} M_{22} & M_{23} \\ M_{23} & M_{33} \end{pmatrix} \begin{pmatrix} a_{2}^{-2}/a_{1} \\ M_{11} & M_{13} \end{pmatrix} \]

\[ a_{2}^{-2} \begin{pmatrix} a_{2}^{-2}/a_{1} \\ M_{11} & M_{13} \end{pmatrix} \]

\[ a_{2}^{-2} \begin{pmatrix} a_{2}^{-2}/a_{1} \\ M_{11} & M_{13} \end{pmatrix} \]

\[ a_{2}^{-2} \]

is non-negative, (ii) the stationary values of this expression are just the roots of the equation in c:

\[
\begin{vmatrix}
c(M_{11} - M_{12}^{-1}M_{13}) - (M_{12}^{-1}M_{13}^{-1}M_{12}^{-1}) \\
(M_{12}^{-1}M_{13}^{-1}M_{12}^{-1})^{-1}(M_{12}^{-1}M_{13}^{-1}M_{12}^{-1})
\end{vmatrix} = 0 \text{ and (iii) in particular, the largest and}
\]

the smallest values of the expression are the largest and smallest roots of the determinantal equation.
Proof. As before, part (i) is obvious. For part (ii), putting the expression under (i) = c (say) and proceeding in exactly the same manner as in (A.2.3) we have for the stationary values of c, the determinantal equation in c:

\[
\begin{vmatrix}
  c & (M_{11} - M_{13}M_{33}^{-1}M_{33}^{-1}) \\
  (M_{13} - M_{33}) & (M_{23} - M_{23}M_{33}^{-1}M_{33}^{-1}) \\
  (M_{33}^{'}) & (M_{33}^{'}) \\
  (M_{33}^{'}) & (M_{33}^{'}) \\
\end{vmatrix} = 0.
\]

As in (A.1.14) - (A.1.15), premultiply the left-hand side of (A.2.4.1) by the determinant of the non-singular matrix F (and postmultiply by its transpose), where F is given by

\[
F = \begin{pmatrix}
  I & -M_{13}M_{33}^{-1} \\
  0 & I \\
  0 & I \\
  0 & I \\
\end{pmatrix}
\]

The equation now reduces to

\[
\begin{vmatrix}
  (M_{11} - M_{13}M_{33}^{-1}M_{33}^{-1}) & (M_{12} - M_{13}M_{33}^{-1}M_{33}^{-1}) \\
  0 & (M_{22} - M_{23}M_{33}^{-1}M_{33}^{-1}) \\
  (M_{33}^{'}) & (M_{33}^{'}) \\
\end{vmatrix} = 0.
\]
or
\[ (A.2.4.2) \quad \left| c(M_{11}^{M_{13}^{M_{33}^{M_{13}}} - M_{12}^{M_{13}^{M_{33}^{M_{23}}} (M_{22}^{M_{23}^{M_{33}^{M_{23}})}})}}\right| = 0. \]

Arguing as in (A.1.14) - (A.1.16) it is easy to see that (i) the roots of this \( p \)th degree equation in \( c \) all lie between 0 and 1, (ii) if \( M \) of (A.2.4) is p.d., then all roots are < 1 and (iii) if \( M_{12}^{M_{13}^{M_{33}^{M_{23}}}} \) is of rank \( r \) (\( r \) \leq p), then \( r \) of these roots are > 0 and the rest, i.e., \( p - r \) are = 0. All the roots are zero if and only if
\( M_{12}^{M_{13}^{M_{33}^{M_{23}}}} = 0. \)

(A.2.5) If \( M(p \times p) \) is symmetric and at least p.s.d., the statement:
"\( g_1 \leq \lambda'(1 \times p)M(p \times p)\lambda(p \times 1)/\lambda'\lambda \leq g_2 \) for all non-null \( \lambda \)" is exactly equivalent to "\( g_1 \leq c_1 \leq c_p \leq g_2 \)," where \( c_1 \) and \( c_p \) stand for the smallest and largest characteristic roots (both non-negative) of \( M \). Notice that the last statement gives also the lowest permissible value of \( g_2 \) and the highest permissible value of \( g_1 \), both in terms of the roots of \( M \). The proof is obvious from (A.2.1).

(A.2.6) If \( M_2(p \times p) \) is symmetric p.d. and \( M_1(p \times p) \) is symmetric and at least p.s.d., the statement: "\( g_1 \leq \lambda'(1 \times p)M_1(p \times p)\lambda(p \times 1)/\lambda'\lambda \leq g_2 \) for all non-null \( \lambda \)" is exactly equivalent to "\( g_1 \leq c_1 \leq c_p \leq g_2 \)," where \( c_1 \) and \( c_p \) stand for the smallest and largest roots of the equation in \( c \) (all positive):
\[ \left| M_1 - cM_2 \right| = 0. \] Notice that the last statement gives also, in terms of the roots of \( M_1M_2^{-1} \), the lowest permissible value of \( g_2 \) and the highest permissible value of \( g_1 \). The theorem is a direct consequence of (A.2.2).

(A.2.7) The statement: "\( x'(1 \times p)x(p \times 1) \leq g \) (\( g > 0 \))" is exactly equivalent to "\( -\sqrt{g} \leq x'(1 \times p)x(p \times 1) \leq +\sqrt{g} \) (for all \( x \) subject to \( x'x = 1 \))." The proof follows easily from Cauchy's inequality in algebra.

1.3. Some results in transformations.

(A.3.1) If \( x(n \times 1) = A(n \times n)y(n \times 1) \), where \( A \) is \( n \times n \), then \( x'x = y'\Lambda y \) where \( \Lambda = y'y \).
If \(x(n \times 1) = A(n \times n)y(n \times 1)\) and \(y(n \times 1) = A(n \times n)v(n \times 1)\), where \(A\) is \(\perp\), then \(x'u = y'A'Av = y'v\).

If \(M(p \times p)\) is symmetric and at least p.s.d. of rank \(r(\leq p)\), then denoting by \(c\) the roots \(c(M)\) of (A.1.8), there exists an orthogonal matrix \(A(p \times p)\) (not necessarily unique) such that \(M = AD_cA'\).

Under the conditions of (A.1.9), there exists a non-singular \(A(p \times p)\) (not necessarily unique) such that \(M_1 = AD_cA'\) and \(M_2 = AA'\).

The matrix \(A\) of (A.3.3) will be unique, except for a post-factor \(D_k\), if \(M\) is p.d. and all \(c(M)\)'s are distinct.

Proof. Suppose there are two orthogonal \(A\)'s, say \(A_1\) and \(A_2\), satisfying the condition of (A.3.3). Then we have \(A_1D_cA_1' = A_2D_cA_2'\) or \(A_2^{-1}A_1D_c = D_cA_2'(A_1')^{-1}\) or \(A_2^{-1}A_1D_c = D_cA_2A_1\) (since, for an orthogonal \(A\), \(A^{-1} = A'\)). If we now denote \(A_2A_1\) by \(B\) with elements \(b_{ij}\), then the above equation gives

\[b_{ij}c_j = c_ib_{ij}\] or \(b_{ij}(c_i-c_j) = 0\). Thus, if \(i \neq j\) and \(c_i \neq c_j\), \(b_{ij} = 0\), which means that \(B\) is a diagonal matrix \(D_b\) with elements, say \(b_1, \ldots, b_p\). Since \(D_b = B = A_2A_1\), we have

\[D_bD_b' = D_{b^2} = A_1'\lambda A_1' = I(p)\), so that \(b^2 = +1\) and so \(b = \pm 1\).

Thus \(D_b = D_k\) and hence \(A_2A_1 = D_k\) or \(A_1 = A_2D_k\). This proves (A.3.5).

We note that \(A\) can thus be made unique by adopting the convention, say, that its first row is to be positive. It is easy to check that the transformation is now one-to-one.

If \(X(p \times n)\) \((p \leq n)\) is of rank \(p\) (in which case, by (A.1.10), \(XX'\) is symmetric p.d.), then there exists a transformation \(X(p \times n) = A(p \times p)D_c(p \times p)L\) \((p \times n)\), where \(A\) is \(\perp\), \(LL' = I(p)\) and where \(c\)'s are the characteristic roots (all positive) of the matrix \(XX'\). If all \(c\)'s are distinct this transformation is unique except for a post factor \(D_k\) to go with \(A\).
Proof. By (A.3.3) and (A.3.5) there exists an orthogonal \( A(p \times p) \), which may not be unique, such that \( XX' = AD_cA' \). We now define a \( L(p \times n) \) by \( X = AD_cL \) and note that, given \( X \) and hence \( c \)'s and \( A \) (which we can find but which may not be unique), this is a linear equation in \( L \) uniquely solvable in terms of the above elements. Also \( LL' = D_{1/c}^{-1} A^{-1} A'^{-1} D_{1/c} = D_{1/c}^{-1} A^{-1} A'^{-1} A^{-1} A'^{-1} D_{1/c} = I(p) \).

We have thus the transformation \( X = AD_c^{-1}L \), where \( A \) is \( \perp \) and \( LL' = I(p) \). Notice that, if the \( c \)'s are distinct, \( A \) is unique except for a post factor \( D_k \) and that \( L \) will go with \( A \), being defined by \( L = D_{1/c}A^{-1}X \). This proves (A.3.6). It is easy to check for distinct roots the transformation can be made one to one by adopting the convention, say, that the first row of \( A \) is to be positive.

(A.3.7) The matrix \( A \) of (A.3.4) will be unique, except for a factor \( D_k \), if \( M_1 \) is p.d. and all the roots are distinct. \[ ^{\text{10.7}} \]

Proof. Suppose there are two non-singular \( A \)'s, say \( A_1 \) and \( A_2 \), satisfying the conditions of (A.3.4). Then we have

\[
A_1 D_c M_1 = A_2 D_c M_2 \quad \text{and} \quad A_1 A_1' = A_2 A_2'.
\]

These lead, after a little manipulation, to

\[
A_2^{-1}A_1 D_c = D_c A_2^{-1} A_1 \quad \text{or} \quad BD_c = D_c B, \quad \text{where} \quad A_2^{-1} A_1 = B. \]

If now \( B = (\delta_{ij}) \), say, then (A.3.7.2) leads to

\[
\lambda_{ij} c_j = c_i \delta_{ij} \quad \text{or} \quad \delta_{ij} (c_i - c_j) = 0 \quad \text{or} \quad \delta_{ij} = 0 \quad \text{if} \quad i \neq j \quad \text{and} \quad c_i \neq c_j. \]

Thus \( B = D_b \) (say) and so we have

\[
D_b D_b' = D_{b_2} = B B' = A_2^{-1} A_1 (''I'')' = A_2^{-1} A_1 (''I'')^{-1} = I(p), \quad \text{so that} \quad b = \perp 1,
\]

i.e., \( D_b = D_k \).

(A.3.7.5) Thus \( A_2^{-1} A_1 = D_k \) or \( A_1 = A_2 D_k \), which proves (A.3.7). As before, we note that \( A \) can be made unique by adopting the convention, say, that its first row is to be positive. Check that the transformation in this case is one-to-one.

(A.3.8) If \( X_1(p \times n_1), X_2(p \times n_2), (p \leq n_1, n_2) \) are each of rank \( p \) (in which case, by (A.1.10), \( X_1'X_1 \) and \( X_2'X_2 \) are both symmetric p.d.), then there exists a
transformation \( X_1(p \times n_1) = \alpha(p \times p)D/\sigma(p \times p)L_1(p \times n_1) \), and \( X_2(p \times n_2) = \lambda(p \times p)L_2(p \times n_2) \) where \( \lambda \) is non-singular, \( \alpha \)'s are the roots (all positive) of the equation \[ \left| X_1 X_1^* - c X_2 X_2^* \right| = 0 \], and \( L_1 L_1^* = L_2 L_2^* = I(p) \). If all \( \alpha \)'s are distinct then this transformation is unique except for a post factor \( D_k \) to go with \( \lambda \).

**Proof.** By (A.1.9) and (A.3.7) there exists a non-singular \( \lambda \), which may not be unique, such that \( X_1 X_1^* = \lambda D_1 D_1^* \) and \( X_2 X_2^* = \lambda D_2 D_2^* \). We now define \( L_1(p \times n_1) \) and \( L_2(p \times n_2) \) by \( X_1 = \lambda D_1 L_1 \) and \( X_2 = \lambda D_2 L_2 \) and note that, given \( X_1 \), \( X_2 \) and \( \alpha \)'s and \( \lambda \) (which may not be unique), \( L_1 \) and \( L_2 \) are uniquely solvable in terms of these. Also \( L_1 L_1^* = D_1 D_1^* \lambda^{-1} X_1 X_1^* \lambda^{-1} D_1 D_1^* = I(p) \) and \( L_2 L_2^* = \lambda^{-1} X_2 X_2^* \lambda^{-1} = I(p) \). This proves the existence of the transformation (A.3.8). Notice that if all \( \alpha \)'s are distinct, then by (A.3.7) \( \lambda \) is unique except for a post factor \( D_k \) and that \( L_1 \) and \( L_2 \) will go with \( \lambda \) being defined by \( L_1 = D_1 D_1^* \lambda^{-1} X_1 \) and \( L_2 = \lambda^{-1} X_2 \). Check that the transformation in this case is one-to-one if we adopt the convention, say, that the first row of \( \lambda \) is to be positive.

(A.3.9) If \( M(p \times p) \) is symmetric and p.d., then there exists a non-singular \( \tilde{\lambda}(p \times p) \) such that \( \tilde{\lambda} = \tilde{\lambda} \tilde{\lambda}^* \), and this \( \tilde{\lambda} \) is unique except for a post factor \( D_k \), and so \( \tilde{\lambda} \) will be called near unique.

**Proof of the near uniqueness.** Suppose there are two \( \tilde{\lambda} \)'s, say \( \tilde{\lambda}_1 \) and \( \tilde{\lambda}_2 \), satisfying the condition. Notice from (A.1.10) that since \( M \) is p.d. \( \tilde{\lambda} \) must necessarily be non-singular. Thus we have

(A.3.9.1) \[ \tilde{\lambda}_1 \tilde{\lambda}_1^* = \tilde{\lambda}_2 \tilde{\lambda}_2^* \] or \[ \tilde{\lambda}_2^{-1} \tilde{\lambda}_1 = \tilde{\lambda}_1 \tilde{\lambda}_2^{-1} \tilde{\lambda}_2 \].

Now making use of the remarks made after (1.1) we note that \( \tilde{\lambda}_2^{-1} \tilde{\lambda}_1 \) is a triangular matrix with the same configuration as \( \tilde{\lambda}_1 \) and \( \tilde{\lambda}_2^{-1} \tilde{\lambda}_1^{-1} \) of opposite configuration.

Thus it is obvious that

(A.3.9.2) \[ \tilde{\lambda}_2^{-1} \tilde{\lambda}_1 = D_2 (say) \], whence \( D_2 D_2^* = D_2 D_2^* \tilde{\lambda}_2 \tilde{\lambda}_2^{-1} \tilde{\lambda}_2^{-1} = \tilde{\lambda}_2^{-1} \tilde{\lambda}_1 \tilde{\lambda}_2^{-1} \tilde{\lambda}_2^{-1} \tilde{\lambda}_2^{-1} \tilde{\lambda}_2^{-1} \tilde{\lambda}_2^{-1} = I(p) \), so that \( a = \pm 1 \), i.e., \( D_2 = D_k \). Thus

(A.3.9.3) \[ \tilde{\lambda}_2^{-1} \tilde{\lambda}_1 = D_k \] or \( \tilde{\lambda}_1 = \tilde{\lambda}_2 D_k \), which proves the near uniqueness. It is easy
to check that \( \tilde{T} \) can be made unique by adopting the convention, say, that the diagonal elements of \( \tilde{T} \) are to be all positive. The transformation in this case is one-to-one.

(A.3.10) If \( M = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^t & M_{22} \end{pmatrix} \), \( p \) is symmetric and p.s.d. of rank \( p \) and if the first \( p \) rows can be taken as a basis, then there exists a non-singular \( \tilde{T}_1(p \times p) \) and a \( T_2(q \times p) \) such that

\[
\begin{pmatrix}
M_{11} & M_{12} \\
M_{12}^t & M_{22}
\end{pmatrix} = \begin{pmatrix} \tilde{T}_1 \\ T_2 \end{pmatrix} \begin{pmatrix} \tilde{T}_1^t \\ T_2^t \end{pmatrix},
\]

and furthermore \( \begin{pmatrix} \tilde{T}_1 \\ T_2 \end{pmatrix} \) is unique except for a post factor \( D_k \).

Proof. First notice that if the first \( p \) rows of \( M \) can be taken as the basis then there exists a non-null \( A(q \times p) \) such that \( M_{12} = AM_{11} \) and \( M_{22} = AM_{12} \). Combining the two we have \( M_{22} = M_{12}A^{-1}M_{11} \) (note that \( M_{11} \) is non-singular and thus we can take the inverse). We next observe that in this set-up \( M_{11} \) is obviously p.d. Therefore, by (A.3.9), we can find a non-singular \( \tilde{T}_1(p \times p) \), unique except for a post-factor \( D_k \), such that \( M_{11} = \tilde{T}_1 \tilde{T}_1^t \). Now find a \( T_2 \) defined by \( T_2 = \tilde{T}_1^{-1}M_{12} \) and check that

\[
M_{12} = (\tilde{T}_1^{-1}T_2)^t = T_2^{-1} \tilde{T}_1 \text{ and } M_{22} = M_{12}A^{-1}M_{11}^{-1}M_{12} = T_2^{-1}T_2^t,
\]

which proves (A.3.10). We observe that, as in (A.3.9), \( \begin{pmatrix} \tilde{T}_1 \\ T_2 \end{pmatrix} \) can be made unique by adopting the convention that the diagonal elements of \( \tilde{T}_1 \) are to be positive. Check that the transformation is now one-to-one.

(A.3.11) If \( X(p \times n) \) (\( p \leq n \)) is of rank \( r \) (\( \leq p \)) such that, say, the first \( r \) rows of \( X \) can be taken as a basis, then there exists a transformation

\[
\begin{pmatrix}
X_1 \\
X_2
\end{pmatrix} \begin{pmatrix} r \\ r-p \end{pmatrix} = \begin{pmatrix} \tilde{T}_1 \\ T_2 \end{pmatrix} \begin{pmatrix} L(r \times n) \end{pmatrix},
\]

where \( \tilde{T}_1 \) is unique except for a post factor \( D_k \).
where \( L_L' = I(r) \) and \( \Omega_1 \) is non-singular and unique except for a post factor \( D_k \).

**Proof.** By (A.1.10) \( X_1X_1' \) is symmetric p.d. of rank \( r \) and by (A.3.9) there exists a (non-singular) \( \Omega_1 \) (unique except for a post factor \( D_k \)) such that \( X_1X_1' = \Omega_1\Omega_1' \). We now define an \( L \) by \( L(r \times n) = \Omega_1^{-1}(r \times r)X_1(r \times n) \) and note that given \( X_1 \) and hence \( \Omega_1 \) (which is unique except for a post factor \( D_k \)), \( L \) is uniquely solvable in terms of these. Also \( LL' = \Omega_1^{-1}X_1'(\Omega_1^{-1})' = \Omega_1^{-1}L\Omega_1'(\Omega_1^{-1})^{-1} = I(r) \). Next define a \( T_2 \) by \( T_2L = X_2 \) or \( T_2LL' = X_2L' \) or \( T_2 = X_2L' \) and note that, given \( X_1, X_2 \) and hence \( L \) (which is near unique), \( T_2 \) is also uniquely solvable in terms of these. We note further that now \( X_2 = T_2L = T_2\Omega_1^{-1}X_1 = B(p-r \times r)X_1 \) (say), where \( B = T_2\Omega_1^{-1} \). This is obviously the condition that \( X \) is of rank \( r \) and \( X_1 \) is a basis. Hence the transformation is proved to exist with the near uniqueness already stated. By adopting a convention, say that of (A.3.10), the transformation can be checked to be one-to-one.

(A.3.12) If \( X_1(p \times n_1), X_2(p \times n_2) \) (\( p \leq n_1, n_2 \)) are each of rank \( p \) (see (A.3.8)), then there exists a transformation: \( X_1(p \times n_1) = \tilde{T}(p \times p)L(p \times p)D_c(p \times p)L_1(p \times n_1) \) and \( X_2(p \times n_2) = \tilde{T}(p \times p)L_3(p \times n_2) \), where \( \tilde{T} \) is non-singular, \( L \) is \( L_1 \), \( L_1L_1' = L_3L_3' \) = \( I(p) \) and the \( c \)'s are the (all positive) roots of the equation in \( c: \left| X_1X_1' - cX_2X_2' \right| = 0 \). If the \( c \)'s are distinct, the transformation can be made one-to-one by letting \( \tilde{T} \) have a positive diagonal.

**Proof.** Start from the transformation (A.3.8) and, using (A.3.11), put \( \Lambda(p \times p) = \tilde{T}(p \times p)L(p \times p) \) where \( L \) is \( L_1 \). Next put \( LL_2 = L_3 \) and note that \( L_3L_3' = LL_2L_2' \) = \( I(p) \). The proof of near uniqueness in the case of distinct roots follows along the same lines as in (A.3.8) and (A.3.11). This completes the proof of (A.3.12).

(A.3.13) If \( M_2(p \times p) \) is symmetric p.d. and \( M_1(p \times p) \) is symmetric p.s.d. of rank \( r < p \) such that the first row vectors can be taken as the basis, then there exists a transformation

\[
(A.3.13.1) \quad M_1(p \times p) = p-r \begin{pmatrix} \Lambda_1 & D_c^* (r \times r) (\Lambda_1')^{-1} \\ \Lambda_2 & r-p \end{pmatrix}, \quad \text{and}
\]

\[
\Lambda_1' = \begin{pmatrix} \Lambda_1 & r \\ \Lambda_2 & r \end{pmatrix}
\]
(A.3.13.2) \( M_2(p \times p) = p-r \begin{pmatrix}
\Lambda_1 & \Lambda_3 \\
\Lambda_2 & \Lambda_4
\end{pmatrix} \begin{pmatrix}
\Lambda_1 & \Lambda_2 \\
\Lambda_3 & \Lambda_4
\end{pmatrix} r \),

\[ r \begin{pmatrix}
\Lambda_1 & \Lambda_4 \\
\Lambda_3 & \Lambda_4
\end{pmatrix} \begin{pmatrix}
\Lambda_1 & \Lambda_2 \\
\Lambda_3 & \Lambda_4
\end{pmatrix} p-r \]

where the \( c \)'s of \( D_c^* \) stand for the \( r \) non-zero roots of the equation in \( c : \left| M_1-c\tilde{M}_2 \right| = 0 \)

and where \( \Lambda = \begin{pmatrix}
\Lambda_1 & \Lambda_3 \\
\Lambda_2 & \Lambda_4
\end{pmatrix} \) is non-singular. If the non-zero roots are distinct, the matrix \( \Lambda \) is unique except for a post factor \( D_k(p) \).

**Proof.** It is obvious from (A.1.9) and (A.3.4) that in this case we can find a transformation

(A.3.13.3) \( M_1 = p-r \begin{pmatrix}
\Lambda_1 & \Lambda_4 \\
\Lambda_2 & \Lambda_4
\end{pmatrix} D_c^* (r \times r) \begin{pmatrix}
\Lambda_1 & \Lambda_2 \\
\Lambda_3 & \Lambda_4
\end{pmatrix} r \),

\[ r \begin{pmatrix}
\Lambda_1 & \Lambda_4 \\
\Lambda_2 & \Lambda_4
\end{pmatrix} \begin{pmatrix}
\Lambda_1 & \Lambda_2 \\
\Lambda_3 & \Lambda_4
\end{pmatrix} p-r \]

(A.3.13.4) \( M_2 = p-r \begin{pmatrix}
\Lambda_1 & B_3 \\
\Lambda_2 & B_4
\end{pmatrix} \begin{pmatrix}
\Lambda_1 & \Lambda_2 \\
\Lambda_3 & \Lambda_4
\end{pmatrix} r \), where \( \begin{pmatrix}
\Lambda_1 & B_3 \\
\Lambda_2 & B_4
\end{pmatrix} \) is to be

\[ r \begin{pmatrix}
\Lambda_1 & B_3 \\
\Lambda_2 & B_4
\end{pmatrix} \begin{pmatrix}
\Lambda_1 & \Lambda_2 \\
\Lambda_3 & \Lambda_4
\end{pmatrix} p-r \begin{pmatrix}
\Lambda_1 & \Lambda_4 \\
\Lambda_2 & \Lambda_4
\end{pmatrix} \begin{pmatrix}
\Lambda_1 & \Lambda_2 \\
\Lambda_3 & \Lambda_4
\end{pmatrix} p-r \]

non-singular but is obviously not necessarily unique. Assuming now that \( B_3 \) is non-singular (since \( M_2 \) is p.d. we can obviously find a non-singular \( B_3 \) satisfying the conditions of the problem), we can, by (A.3.11), find a transformation \( B_3(p-r \times p-r) = \tilde{A}_3(p-r \times p-r)L(p-r \times p-r) \) where \( L \) is \( I \). Now put \( B_4(r \times p-r) = \Lambda_4(r \times p-r)L(p-r \times p-r) \) (which defines \( \Lambda_4 \) in a unique way in terms of \( B_4 \) and \( L \)). Thus we have

\[ \begin{pmatrix}
\Lambda_1 & B_3 \\
\Lambda_2 & B_4
\end{pmatrix} = \begin{pmatrix}
\Lambda_1 & \tilde{A}_3 \\
\Lambda_2 & \Lambda_4
\end{pmatrix} \begin{pmatrix}
I(r) & 0 \\
0 & L
\end{pmatrix} p-r \]

and thus (A.3.13.4) is replaced by

(A.3.13.5) \( M_2 = \begin{pmatrix}
\Lambda_1 & \tilde{A}_3 \\
\Lambda_2 & \Lambda_4
\end{pmatrix} \begin{pmatrix}
\Lambda_1 & \Lambda_2 \\
\tilde{A}_3 & \Lambda_4
\end{pmatrix} \).

(A.3.13.3) and (A.3.13.5) taken together give us (A.3.13.1) and (A.3.13.2). Now for
the near uniqueness in the case of distinct roots under $D_c$. Remembering that
\[ D_c = r \begin{pmatrix} D^* & 0 \\ 0 & 0 \end{pmatrix}, \]
and putting
\[ U = \begin{pmatrix} A_1 & A_2 & A_3 \\ & A_4 \end{pmatrix}, \]
let us write $M_1 = UD_cU'$ and $M_2 = UU'$. If now there is another matrix $V$ satisfying the same conditions, then arguing in the same manner as (A.3.7) we have $V^{-1}UD_c = D_cV^{-1}U$ or $BD_c = D_cB$, where $B = V^{-1}U = (b_{ij})$, say. This, as in (A.3.7), leads to the equation $b_{ij}c_j = c_i b_{ij}$, whence $b_{ij} = 0$ if $i \neq j$ and $c_i \neq c_j$. Note that here $c_i = 0$ ($i = r+1, \ldots, p$). This shows that the $B$ matrix is of the form
\[ \begin{pmatrix} D_a & 0 \\ 0 & \text{solid} \end{pmatrix} r = \begin{pmatrix} D_a & 0 \\ 0 & S \end{pmatrix} \]
(say).

Remembering that $B = V^{-1}U$, we have
\[ BB' = \begin{pmatrix} D_a^2 & 0 \\ 0 & SS' \end{pmatrix} = V^{-1}UU'(V^{-1})' = V^{-1}VV'(V')^{-1} = I(p), \]
whence $D_a = D_k$ and $S$ is soon to be an orthogonal matrix. Using the structure of $U$ and $V$ and the relation $V^{-1}U = B$ we have
\[ \begin{pmatrix} U_1 & \tilde{U}_3 \\ U_2 & U_4 \end{pmatrix} = \begin{pmatrix} V_1 & \tilde{V}_3 \\ V_2 & V_4 \end{pmatrix} \begin{pmatrix} D_k & 0 \\ 0 & S \end{pmatrix} = \begin{pmatrix} V_1D_k & \tilde{V}_3S \\ V_2D_k & V_4S \end{pmatrix}, \]
where $S$ is an $\perp$ matrix. But, since $\tilde{U}_3 = \tilde{V}_3S$, therefore, $S$ must also be a triangular matrix. Hence $\tilde{S}$ is necessarily of the form $D_k(v-r)$. Thus $B$ is of the form $D_k(p)$, which proves the near uniqueness in the case of distinct non-zero roots. This completes the proof of (A.3.13). In this case the transformation is easily checked to be one-to-one if we adopt the convention that the first row of $\tilde{A}_3$ and the diagonal elements of $\tilde{A}_3$ are to be positive.

(A.3.14) If $X_1(p \times n_1)$ ($p > n_1$) be of rank $n_1$ such that the last $n_1$ rows which form a square matrix is non-singular and $X_2(p \times n_2)$ ($p \leq n_2$) be of rank $p$, then there exists a transformation
\[ X_1(p \times n_1) = \frac{p-n_1}{n_1} U_1 \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right) D/c (n_1 \times n_1) L_1(n_1 \times n_1), \]
\[ X_2(p \times n_2) = \frac{p-n_1}{n_1} U_1 \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right) U_3 \left( \begin{array}{c} U_3 \\ U_4 \end{array} \right) L_2(p \times n_2), \]

such that \( L_1 \) is \( \perp \) and \( L_2 L_2^T = I(p) \) where \( c \)'s stand for the non-zero roots of the equation in \( c \): \[ |X_1X_1^T - cX_2X_2^T| = 0, \] and \( U = \left( \begin{array}{cc} U_1 & U_3 \\ U_2 & U_4 \end{array} \right) \) is non-singular. Also, if all the non-zero roots \( c \) are distinct, \( U \) is unique except for a post factor \( D_k \).

Notice that, by \((A.1.2)\), all the \( c \)'s are anyway non-negative and \( n_1 \) of them are positive, the rest being zero.

**Proof.** By \((A.1.12)\) there exists an \( \left( \begin{array}{cc} U_1 & \tilde{U}_3 \\ U_2 & \tilde{U}_4 \end{array} \right) \) \( p-n_1 \) not necessarily unique such that \( X_1X_1^T = \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right) D/c (n_1 \times n_1) \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right) \) and \( X_2X_2^T = \left( \begin{array}{cc} U_1 & U_3 \\ U_2 & U_4 \end{array} \right) \left( \begin{array}{cc} U_3 & U_4 \end{array} \right) \). Now define an \( L_1(n_1 \times n_1) \) and \( L_2(p \times n_2) \) by

\[ X_1(p \times n_1) = \frac{p-n_1}{n_1} Y_1 \left( \begin{array}{c} Y_1 \\ Y_2 \end{array} \right) = \frac{p-n_1}{n_1} Y_1 \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right) D/c (n_1 \times n_1) L_1(n_1 \times n_1), \]
\[ X_2(p \times n_2) = \frac{p-n_1}{n_1} Y_3 \left( \begin{array}{c} Y_3 \\ Y_4 \end{array} \right) L_2(p \times n_2), \]

and notice that \( L_2 \) is given uniquely by \( L_2 = U^{-1}X_2 \), in terms of \( X_2 \) and \( U \) which itself may not be unique and similarly \( L_1 \) is given uniquely (in the same sense) by
\[ L_1 = D_1 U_2^{-1} Y_2 \]. Next we check that \( L_2 L_2' = U^{-1} X_2 X_2'(U^{-1})' = U^{-1} U U' (U')^{-1} = I(p) \) and 
\[ L_1 L_1' = D_1 U_2^{-1} Y_2 U_2' (U_-1)' D_1 / c = D_1 / c U_2 U_2' D_1 / c = I(n_1) \]. We observe also if the non-zero roots are unique, then, by (A.3.13), \( U \) is unique except for a post-factor \( D_k(p) \) and thus \( L_1 \) and \( L_2 \) which hang on \( U \) are also indeterminate to the same extent. This completes the proof of (A.3.14). As in the case of (A.3.13) so also here, for distinct roots the transformation can be made one-to-one by adopting the same convention as at the end of (A.3.13).

(A.3.15) If \( X_1(p \times n_1), X_2(p \times n_2) \) \((n_1 < p \leq n_2)\) are of ranks \( n_1 \) and \( p \) respectively, then there exists a transformation: \( X_1'(n_1 \times p) = L(n_1 \times n_1) D / c (n_1 \times n_1) L (n_1 \times p) \), \( X_2'(p \times p) = \tilde{T}(p \times p) L_2(p \times n_2) \), where \( L \) is \( \perp \), \( L_1 L_1' = I(n_1) \), \( L_2 L_2' = I(p) \) and \( c \)'s are the \( n_1 \) characteristic roots (all positive) of \( X_1'(X_2'X_2')^{-1}X_1 \).

For distinct roots the transformation can be made one-to-one by letting \( \tilde{T} \) have a positive diagonal.

**Proof.** Using (A.3.11), put \( X_2(p \times n_2) = \tilde{T}(p \times p) L_2(p \times n_2) \), subject to \( L_2 L_2' = I(p) \) and now using (A.3.7) put \( X_1(n_1 \times p) = (\tilde{T}^{-1}(p \times p) = L(n_1 \times n_1) D / c \), \( \perp \), \( L_1 L_1' = I(n_1) \) and \( c \)'s are the roots of \( X_1'(X_2'X_2')^{-1}X_1 \), i.e., of \( X_1'(\tilde{T}^{-1})^{-1}X_1 \), i.e., of \( X_1'(X_2'X_2')^{-1}X_1 \). Post-multiplying both sides by \( \tilde{T} \) we have: 
\[ X_1' = L D / c \tilde{T} \], and for \( X_2 \) we already have \( X_2 = \tilde{T} L_2 \). Near uniqueness, in the case of distinct roots, follows along the same lines as in (A.3.11) and (A.3.14). Check, by using (A.1.18), that these \( c \)'s of (A.3.15) are the same as the non-zero roots of the equation in \( c \) (considered in (A.3.14): \[ X_1 X_1' - c X_2 X_2' = 0 \].

(A.3.16) If \( M = \begin{pmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{pmatrix} \) \( p \leq q \) is symmetric p.d. (note that, in this situation, \( M_{11} \) and \( M_{22} \) are both necessarily symmetric p.d.) and if \( M_{12} \) is of rank \( s \), \( s \leq q \), and if \( D_c(s \times s) \) is the diagonal matrix based on the \( s \) non-zero roots of the \( p \) th degree equation in \( c \): 
\[ c D_c s - M_{12} D_c s M_1 \] = 0, then there exist non-singular
\( A(p \times p) \) and \( B(q \times q) \) of the structure \( A = \begin{pmatrix} \hat{\lambda}_1 & \hat{\lambda}_3^* \\ \hat{\lambda}_2 & \hat{\lambda}_4^* \end{pmatrix} \) \( p-s \) and \( B = \begin{pmatrix} B_1 & B_3^* \\ B_2 & B_4^* \end{pmatrix} \) \( q-s \)

\( \hat{\lambda}_3 \) and \( \hat{\lambda}_3^* \) being non-singular such that \( M_{11}(p \times p) = A(p \times p)A'(p \times p) \), \( M_{22}(q \times q) = B(q \times q)B'(q \times q) \) and \( \tilde{M}_{12}(p \times q) = A(p \times p) \begin{pmatrix} D_c(s \times s) & 0(s \times q-s) \\ 0(p-s \times s) & 0(p-s \times q-s) \end{pmatrix} B'(q \times q) \)

\( = \begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix} D_c(B_1 \quad B_2^*) \); also, if the c's are distinct, \( A \) is unique except for a post factor \( D_k(p) \) and, for a given choice of \( A \), \( B \) is unique except for a post factor \( D_k(q-s) \) to go with \( \tilde{B}_3^* \).

**Proof.** Since \( M_{11} \) is symmetric p.d. and \( M_{12}^{-1}M_{12}^{-1} \) is symmetric and at least p.s.d. of rank \( s \leq p \), there exists, by (A.3.13), a transformation \( M_{11} = A' \) and

\[
M_{12}^{-1}M_{12}^{-1} = \begin{pmatrix} \hat{\lambda}_1 & \hat{\lambda}_3 \\ \hat{\lambda}_2 & \hat{\lambda}_4 \end{pmatrix} \begin{pmatrix} D_c & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\lambda}_1 & \hat{\lambda}_3' \\ \hat{\lambda}_2 & \hat{\lambda}_4' \end{pmatrix} \]

where \( \begin{pmatrix} \hat{\lambda}_1 & \hat{\lambda}_3' \\ \hat{\lambda}_2 & \hat{\lambda}_4' \end{pmatrix} \) is non-singular, \( \hat{\lambda}_3' \) is non-singular and c's stand for the \( s \) non-zero roots of the equation in \( c \), the rest, \( p-s \) in number, being zero. Next, since \( M_{22}(q \times q) \) is unique, it follows from (A.3.3) that there is an orthogonal \( E(q \times q) \) such that \( M_{22} = ED \) \( E \) where \( c = c_1, \ldots, c_q \) denotes the \( q \) characteristic roots (all positive) of the p.d. matrix \( M_{22} \). Substituting this in \( M_{12}^{-1}M_{12}^{-1} \) we have

\[
(A.3.16.1) \quad M_{12}(ED \times E)^{-1}M_{12} = \begin{pmatrix} D_c & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\lambda}_1 & \hat{\lambda}_3 \end{pmatrix} \begin{pmatrix} \hat{\lambda}_1' & \hat{\lambda}_3' \end{pmatrix} \quad \text{cr (since } E \text{ and } A \text{ non-singular)}
\]

\[
(A.3.16.2) \quad \hat{\lambda}_1^{-1}M_{12}ED_{1/0}E'M_{12}(\hat{\lambda}_1^{-1})' = \begin{pmatrix} D_c & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\lambda}_1 & \hat{\lambda}_3 \end{pmatrix} \begin{pmatrix} \hat{\lambda}_1' & \hat{\lambda}_3' \end{pmatrix} \begin{pmatrix} D_c & 0 \\ 0 & 0 \end{pmatrix}
\]

We now define a \( G_1(s \times q) \) by

\[
(A.3.16.3) \quad D_{/c}(s \times s)G_1(s \times q) \text{ the submatrix formed by the first } s \text{ rows of } (\hat{\lambda}_1^{-1}M_{12}ED_{1/0}s) \quad p \text{. It is easy to check that, given the other elements, (A.3.16.3)}
\]

\[ q \]
defines $G_1$ uniquely and also that $G_1(s \times q)G_1(q \times s) = I(s)$. It is well known that if $G_1(s \times q)$ $(s < q)$ satisfies $G_1G_1^t = I(s)$, then we can adjoin an $G_2(q-s \times q)$ to $G_1$ such that \( \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} \) is an $\perp$ matrix. With this adjunction we can now write

\[(\lambda.3.16.4) \quad (\lambda^{-1}\lambda_{12}^{\text{ED}}D^{-1/2})p = s \begin{pmatrix} D/c & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} q-s \quad \text{or} \quad s q-s q \]

\[(\lambda.3.16.5) \quad (\lambda^{-1}\lambda_{12}^{\text{ED}}D^{-1/2})(q_1 q_2^t) = \begin{pmatrix} D/c & 0 \\ 0 & 0 \end{pmatrix} s \quad \text{Next put} \quad (\lambda\lambda_{12}^{\text{ED}})(G_1 G_2^t) = F^{-1} (\text{say}, so that} \]

\[(\lambda.3.16.6) \quad F'(q \times q) = s \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} D/c E' \quad \text{(remembering that $E$ is $\perp$). Notice that,}

given the submatrices of the $M$ matrix, we can find a non-singular $\lambda$, an $\perp E$ and an $\perp \begin{pmatrix} G_1 \\ G_2 \end{pmatrix}$ (note being necessarily unique) and thus a (non-singular but not necessarily unique) $F$ given by $(\lambda.3.16.6)$. Using now $(\lambda.3.16.5)$, $(\lambda.3.16.6)$ and the definition of $\lambda$ (in the beginning of the proof), we check that we have non-singular $\lambda$ and $F$ satisfying

\[(\lambda.3.16.7) \quad M_{11}(p \times p) = \lambda(p \times p) \lambda'(p \times p), M_{12}(p \times q) = \lambda(p \times p) \begin{pmatrix} D/c & 0 \\ 0 & 0 \end{pmatrix} s \quad q-s \]

$F'(q \times q)$ and $M_{22}(q \times q) = F(q \times q)F'(q \times q)$. We next partition $F$ into

\[ \begin{pmatrix} F_1 & F_3 \\ F_2 & F_4 \end{pmatrix} q-s \quad \text{, assume that $F_3$ is non-singular (as we obviously can), note that} \]

$F_3$ and $F_4$ do not occur in the factorization of $M_{12}$ and put $F_1 = B_1, F_2 = B_2$,
\[ F_3(q-s \times q-s) = \tilde{E}_3(q-s \times q-s)L(q-s \times q-s) \quad \text{(where } L \text{ is } \perp) \quad \text{and } F_4(s \times q-s) = B_4(s \times q-s)L(q-s \times q-s). \]

As in (A.3.13), remembering the structure of \( A \), we now rewrite (A.3.16.7) as

\[
M_{11}(p \times p) = \begin{pmatrix} \Lambda_1 & \Lambda_3 \\ \Lambda_2 & \Lambda_4 \end{pmatrix} \begin{pmatrix} \Lambda_1 & \Lambda_3 \\ \Lambda_2 & \Lambda_4 \end{pmatrix} \begin{pmatrix} p-s \\ p-s \end{pmatrix},
\]

\[
M_{12}(p \times q) = \begin{pmatrix} \Lambda_1 \\ \Lambda_2 \end{pmatrix} D \begin{pmatrix} s \\ s \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \begin{pmatrix} s \\ q-s \end{pmatrix}
\]

\[
M_{22}(q \times q) = \begin{pmatrix} B_1 & \tilde{B}_3 \\ B_2 & \tilde{B}_4 \end{pmatrix} \begin{pmatrix} B_1 & \tilde{B}_3 \\ B_2 & \tilde{B}_4 \end{pmatrix} \begin{pmatrix} q-s \\ q-s \end{pmatrix}
\]

which establishes the existence of the transformation (A.3.16).

Now for the near uniqueness of \( A \) and \( B \) where the \( c \)'s are distinct. In this case we first recall the definition of \( A \) and observe, as in the proof of (A.3.13), that \( A \) is unique except for a post factor \( D_k(p) \). The second equation of (A.3.16.8) shows that at this stage \( B_1 \) and \( B_2 \) are unique except for the post factor that goes with \( A \). Now consider the third equation of (A.3.16.8) and partition \( M_{22} \) into four submatrices and rewrite the equation as

\[
(M_{22}^{(1)} \quad M_{22}^{(3)})_{q-s} = \begin{pmatrix} B_1 B_1 + \tilde{B}_3 B_3 \\ B_2 B_2 + \tilde{B}_4 B_4 \end{pmatrix},
\]

\[
(M_{22}^{(2)} \quad M_{22}^{(4)})_{q-s} = \begin{pmatrix} B_2 B_1 + \tilde{B}_4 B_3 \\ B_2 B_2 + \tilde{B}_4 B_4 \end{pmatrix}
\]

whence, from the relation: \( B_1 B_1 + \tilde{B}_3 B_3 = M_{22}^{(1)} \), remembering that \( B_1 \) is already known and using (A.3.10), we see that \( \tilde{B}_3 \) is uniquely determined except for a post factor \( D_k(q-s) \). The equation \( B_1 B_1 + \tilde{B}_3 B_3 = M_{22}^{(3)} \) now uniquely defines \( B_4 \) except for the post factors that go with the other \( B_1 \)'s. This completes the proof of the near uniqueness in the case of distinct \( c \)'s. If \( s = p \), i.e., if \( M_{12} \) is of rank \( p \), then
all roots become positive, i.e., $D_c$ becomes $p \times p$, $L$ becomes a solid matrix while $B$ retains its own structure with $q$-s being replaced by $q$-p. If $q = p$, then $B$ itself becomes a solid matrix. As before, for the case of distinct roots, the transformation is checked to be one-to-one by adopting the convention, say, that the first row of $\lambda_1$ and the diagonal elements of $\lambda_2$ and $B_3$ are to be positive.

\[(A.3.17) \quad \text{If } \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}^p_q \quad (p \leq q, p + q \leq n) \text{ is of rank } p + q \]

and $X_1X_2^t$ is also of rank $p$, then there exists a transformation: $X_2(q \times n) = \hat{T}(q \times q)L_2(q \times n)$ and $X_1(p \times n) = U(p \times p)M_1(p \times n-q)$ : $D_{/c}(p \times p)M_2(p \times q)$ : $x \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}^n_q$, where $\hat{T}$ and $U$ are non-singular, $M_1^1 = M_2^1 = I(p), L_2L_2^t = I(q)$, and $L_1$ is a completion of $L_2$ (see (A.1.6)) such that $\begin{pmatrix} L_1 \\ L_2 \end{pmatrix}$ is $\bot$ and $c_i = (1-c_i)/c_i$ or $c_i = 1/(1+c_i)$ \((i = 1, \ldots, p)\) and $c_i$'s are the roots of the equation in $c$:

\[c(X_1^2) - (X_1^2)(X_2^2)^{-1}(X_2^2) = 0.\]

Proof. Using (A.3.11) put $X_2(q \times n) = \hat{T}(q \times q)L_2(q \times n)$ where $\hat{T}$ is non-singular and $L_2L_2^t = I(q)$. Complete $L_2$ by an $L_1$ such that $\begin{pmatrix} L_1 \\ L_2 \end{pmatrix}$ is $\bot$. Now using (A.3.8), put

\[X_1(p \times n)M_1(n \times n-q) L_2(n \times q) = U(p \times p)M_1(p \times n-q) : D_{/c}(p \times p)M_2(p \times q)
\]

where $U$ is non-singular, $M_1^1 = M_2^1 = I(p)$ and $c$'s are the roots of the equation in $c$:

\[c(X_1^2L_1^2X_1^1) - (X_1^2L_2^2X_1^1) = 0.\]

Multiplying both sides of the $X_1$-equation by $\begin{pmatrix} L_1 \\ L_2 \end{pmatrix}$ and taking into account the $X_2$-equation we have the transformation $(A.3.17)$, except for the required interpretation of $c$, which is as follows.
\( L_2 = (\overline{T}^\top)^{-1}X_2, \) so that \( L_2^*L_2 = X_2'(\overline{T}^\top)^{-1}X_2. \) Also \( \overline{L}_2\overline{L}_2 = I - L_2^*L_2 = I - X_2'(\overline{T}^\top)^{-1}X_2. \)

Hence the equation in \( Q \) becomes:

\[
\begin{vmatrix}
X_2 \sqrt{I} - X_2'(\overline{T}^\top)^{-1}X_2\overline{X}_1' \\
\frac{1}{X_2'} X_1X_1' - X_1X_2'(X_2X_2')^{-1}X_2X_1'
\end{vmatrix} = 0
\]

or

\[
\begin{vmatrix}
\frac{1}{X_2'} X_1X_1' - X_1X_2'(X_2X_2')^{-1}X_2X_1'
\end{vmatrix} = 0
\]

(since \( X_2X_2' = \overline{T}^\top \)), which completes the proof of (A.3.17).

\[(A.3.18)\] If \( X = \begin{pmatrix} \lambda_1 \\ 1 \\ X_2 \end{pmatrix} \) \( (p \leq q, p + q \leq n; \text{rank} = p + q) \) is such that \( X_1X_2' \) is of rank \( s \leq p \) (in which case it is easy to check that \( X_1X_2' \) and \( X_2X_2' \) are each symmetric p.d. and \( X_1X_2'(X_2X_2')^{-1}X_2X_1' \) is symmetric and at least p.s.d. of rank \( s \), so that \( s \) roots of the \( p \)th degree equation in \( c: \left| c(X_1X_2') - (X_1X_2')(X_2X_2')^{-1}(X_2X_1') \right| = 0 \) are positive, the rest being zero), then there exists a transformation

\[(A.3.18.1)\]

\[
X_1(p \times n) = \begin{pmatrix} L_1 \\ \vdots \\ L_s \end{pmatrix}
\]

\[
+ \begin{pmatrix} L_{s+1} \\ \vdots \\ L_p \end{pmatrix}
\]

where the \( L \) matrices are subject to

\[(A.3.18.3)\]

\[
X_1X_2' = \begin{pmatrix} A_1 \\ \vdots \\ A_s \end{pmatrix} \begin{pmatrix} A_1' \\ \vdots \\ A_s' \end{pmatrix}
\]

and where the \( L \) matrices are subject to
\[
\begin{pmatrix}
    L_1' & L_2' & L_3' & L_4'
\end{pmatrix} n = I(p + q) \cdot
\]

Proof.

(\(A.3.18.4\)) Put \(X_1(p \times n) = \begin{pmatrix}
    A_1 & A_3 \\
    A_2 & A_4
\end{pmatrix} M \begin{pmatrix}
    L_1 & L_2 \\
    L_3 & L_4
\end{pmatrix}
\) and

\[
\begin{pmatrix}
    B_1 & B_3 \\
    B_2 & B_4
\end{pmatrix} \begin{pmatrix}
    L_1' & L_2' \\
    L_3' & L_4'
\end{pmatrix}
\]

(\(A.3.18.5\)) \(X_2(q \times n) = \begin{pmatrix}
    B_1 & B_3 \\
    B_2 & B_4
\end{pmatrix} \begin{pmatrix}
    L_1 & L_2 \\
    L_3 & L_4
\end{pmatrix}
\)

Now check that, since \(A\) and \(B\) are non-singular, the above equation defines \(M\), \(L_1\), \(L_2\), \(L_3\), \(L_4\) uniquely except for the indeterminacy in \(A\) and \(B\). Now, using the first two equations of \((A.3.18.3)\), it is easy to check that

\[
\begin{pmatrix}
    M \\
    L_2
\end{pmatrix} \begin{pmatrix}
    M' & L_2'
\end{pmatrix} = I(p)\quad\text{and}\quad\begin{pmatrix}
    L_3 \\
    L_4
\end{pmatrix} \begin{pmatrix}
    L_3' & L_4'
\end{pmatrix} = I(q).
\]

Substituting for \(X_1\) and \(X_2\) (in terms of the \(A\), \(B\) and \(M\), \(L_2\), \(L_3\) and \(L_4\)) in the third equation of \((A.3.18.3)\) we have

\[
\begin{pmatrix}
    A_1 & A_3 \\
    A_2 & A_4
\end{pmatrix} \begin{pmatrix}
    D' & 0 \\
    0 & 0
\end{pmatrix} \begin{pmatrix}
    B_1 & B_2 \\
    B_3 & B_4
\end{pmatrix} = \begin{pmatrix}
    A_1 & A_3 \\
    A_2 & A_4
\end{pmatrix} \begin{pmatrix}
    M \\
    M'
\end{pmatrix} \begin{pmatrix}
    L_3 & L_4 \\
    L_3' & L_4'
\end{pmatrix} \begin{pmatrix}
    B_1 & B_2 \\
    B_3 & B_4
\end{pmatrix},
\]

whence it follows that

\[
\begin{pmatrix}
    M \\
    L_2
\end{pmatrix} \begin{pmatrix}
    L_3 & L_4
\end{pmatrix} = \begin{pmatrix}
    D' & 0 \\
    0 & 0
\end{pmatrix} \begin{pmatrix}
    B_1 & B_2 \\
    B_3 & B_4
\end{pmatrix}
\]

Let us now put

\[
A.3.18.6 & (A.3.18.9)\]

\[
\begin{pmatrix}
    M(x \times n) = D' & 0 \\
    0 & 0
\end{pmatrix} \begin{pmatrix}
    L_3 & L_4
\end{pmatrix}
\]

\[
A.3.18.7 & (A.3.18.8)
\]

\[
\begin{pmatrix}
    D' & 0 \\
    0 & 0
\end{pmatrix} \begin{pmatrix}
    B_1 & B_2 \\
    B_3 & B_4
\end{pmatrix}
\]

\[
\begin{pmatrix}
    M(x \times n) + M_1(x \times n)
\end{pmatrix}
\]

which uniquely defines \(M_1\) in terms of \(M\), \(L_3\) and \(c's\). Now substituting in the equations \((A.3.18.6)\) and \((A.3.18.8)\) for \(M\) the right hand side of \((A.3.18.9)\), we have
\[ M^1(L_2', L_3', L_4') = (0, C, 0) \] and

\[ I(s) = M^{11} = D^{1/c} L_2 L_3 D^{1/c} + M^{11} = D_t + M^{11}. \]

It follows from (A.3.18.11) that

\[ M^{11} = I(s) - D_t = D_{1-c}, \]

so that if we put

\[ M_1(s \times n) = D^{1/c}(s \times s)L_1(s \times n), \]

we shall have, from (A.3.18.12) and (A.3.18.10)

\[ L_1 L_1' = I(s) \text{ and } L_1(L_2', L_3', L_4') = (0, 0, 0). \]

Substituting from (A.3.18.13) for \( M_1 \) in (A.3.18.9) we have

\[ M(s \times n) = D^{1/c}(s \times s)L_3(s \times n) + D^{1/c}(s \times s)L_1(s \times n), \]

where \( L_1 \) satisfies (A.3.18.14). Now substituting for \( M \) from (A.3.18.15) in (A.3.18.14) and using (A.3.18.6), (A.3.18.8) and (A.3.18.14) we have

\[ X_1 = \begin{pmatrix} \Lambda_1 & \Lambda_3' \\ \Lambda_2 & \Lambda_4' \end{pmatrix} \begin{pmatrix} D^{1/c}L_1 + D^{1/c}L_3 \\ L_2 \end{pmatrix} = \begin{pmatrix} \Lambda_1 & \Lambda_3' \\ \Lambda_2 & \Lambda_4' \end{pmatrix} \begin{pmatrix} D^{1/c} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} \]

\[ + \begin{pmatrix} \Lambda_1' \\ \Lambda_2' \end{pmatrix} D^{1/c}L_3 \]

and

\[ X_2 = \begin{pmatrix} B_1 & \widetilde{B}_3 \\ B_2 & B_4 \end{pmatrix} \begin{pmatrix} L_3 \\ L_4 \end{pmatrix}, \text{ where the } L's \text{ satisfy} \]

\[ \begin{pmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \end{pmatrix} = I(p + q). \text{ This proves (A.3.18). If} \]

\[ s = p \text{ (which is the case that will be actually considered in this report), } L_2 \text{ will be absent and } q-s = q-p \text{ and we shall have} \]

\[ X_1(s \times n) = A(p \times n) \begin{pmatrix} D^{1/c} & D^{1/c} \\ p & p \end{pmatrix} \begin{pmatrix} L_1 \\ L_3 \end{pmatrix} \]

and

\[ X_2 = \begin{pmatrix} B_1 & \widetilde{B}_3 \\ B_2 & B_4 \end{pmatrix} \begin{pmatrix} L_3 \\ L_4 \end{pmatrix}. \]
\[ x_2(q \times n) = p \begin{pmatrix} B_1 & B_3 \\ B_2 & B_4 \end{pmatrix} \begin{pmatrix} L_3 \\ L_4 \end{pmatrix}^{q-p} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}^{p-q-p} n \]

where the \( L \)'s satisfy

\[ p \begin{pmatrix} L_1 \\ L_3 \end{pmatrix} \begin{pmatrix} L_1^* & L_3^* & L_4^* \end{pmatrix} n = I(p + q) \]

As to the indeterminacy on the right hand side of (A.3.18.1) and (A.3.18.2) (for the case \( s < p \) and of (A.3.18.19) (for the case \( s = p \)), it is easy to check that in either case, if the non-zero roots are all distinct, there is no unique in the sense of (A.3.16), the only indeterminacy arising out of a post factor \( D_k(p) \) going with the total \( \Lambda \)-matrix and a post factor \( D_k(q-s) \) going with \( \tilde{B}_3^* \). In this case the transformation can be made one-to-one by adopting the same convention as, say, at the end of (A.3.16).

\[ (A.3.19) \quad \text{For } \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} \mathbb{1} \\ q \\ r \end{pmatrix} \text{ (} p \leq q, p + q + r \leq n, \text{ rank } p + q + r \text{)}, \]

(i) there exists a transformation: \( X_3(r \times n) = \tilde{T}(r \times r)L_3(r \times n) \) subject to \( L_2L_3^* = I(r) \) and

\[ x_1(p \times n) = p(z_{11} z_{12})^{n-r} (L_3)_r n \\ x_2(q \times n) = q(z_{21} z_{22})^{n-r} (L_3)_r n \\ \]

where \( L \) is just a completion of \( L_3 \) so that \( \begin{pmatrix} L \\ L_3 \end{pmatrix} \) is \( \perp \). (ii) Putting \( M = XX^t \) (observe that, by (A.1.10), \( M \) will be symmetric p.d.), the roots of the equation in \( e \), namely (A.2.4.1) or (A.2.4.2) are the same as the characteristic roots of

\[ (z_{11} z_{11})^{-1}(z_{11} z_{21})(z_{21} z_{21})^{-1}(z_{21} z_{11}) \]
Proof. The proof of (i) is obvious from the preceding sections. For (ii) we observe
that $L_3 = (\mathbf{T})^{-T}X_3$ so that $L_3^T = X_3(\mathbf{T})^{-1}$ whence $L_3^T L_3 = X_3^T(X_3^{-1})^T X_3 = X_3^T X_3^{-1} X_3$.

Therefore $L_3^T L_3 = I(n) - L_3^T L_3 = I(n) - X_3^T X_3^{-1} X_3$ and thus

$$Z_{11} Z_{11} = X_1 L_1^T L_1 X_1 = X_1^T X_1 - X_1^T X_3^{-1} X_3 X_1 = M_{11} - M_{13} M_{33}^{-1} M_{13},$$

$$Z_{12} Z_{12} = X_1 L_1^T X_2 = X_1^T X_2 - X_1^T X_3^{-1} X_3 X_2 = M_{12} - M_{13} M_{33}^{-1} M_{23},$$

and

$$Z_{21} Z_{21} = X_2 L_1^T L_1 X_2 = X_2^T X_2 - X_2^T X_3^{-1} X_3 X_2 = M_{22} - M_{23} M_{33}^{-1} M_{23}. $$

This completes the proof of (ii).

(1.3.20) For an $M$ of the structure (1.2.4) there exists the transformation

$$p \begin{pmatrix} 1 \quad 0 \quad \lambda_1 \\
0 \quad 0 \quad \lambda_2 \\
0 \quad 0 \quad \lambda_3 \\
r \end{pmatrix} q = \begin{pmatrix} 1 \quad 0 \quad \lambda_1 \\
0 \quad 0 \quad \lambda_2 \\
0 \quad 0 \quad \lambda_3 \\
r \end{pmatrix} q 
\begin{pmatrix} D/\mathcal{C} & 0 \\
0 & I \\
0 & I \\
0 & I \end{pmatrix} \begin{pmatrix} 1 \quad 0 \quad \lambda_1 \\
0 \quad 0 \quad \lambda_2 \\
0 \quad 0 \quad \lambda_3 \\
r \end{pmatrix} q^{-1},
$$

where

$$\begin{pmatrix} 1 \quad 0 \quad \lambda_1 \\
0 \quad 0 \quad \lambda_2 \\
0 \quad 0 \quad \lambda_3 \\
r \end{pmatrix}$$

is a non-singular matrix and $c$'s are the roots of the equation

$$(\lambda_1 \lambda_2 \lambda_3)^{1/2}$$
in $c$, (1.2.4.1) or (1.2.4.2).

Proof. We can write

$$\begin{pmatrix} M_{11} & M_{12} & M_{13} \\
M_{12} & M_{22} & M_{23} \\
M_{13} & M_{23} & M_{33} \end{pmatrix} = \begin{pmatrix} M_{11} & M_{13}^{-1} M_{13} & M_{12} & M_{13}^{-1} M_{13} & M_{13}^{-1} M_{13} & M_{13}^{-1} M_{13} \\
M_{12} & M_{22} & M_{23} & M_{23} & M_{23} & M_{23} \\
M_{13} & M_{23} & M_{33} & M_{33} & M_{33} & M_{33} \end{pmatrix} = \begin{pmatrix} M_{11} & M_{13}^{-1} M_{13} & M_{12} & M_{13}^{-1} M_{13} & M_{13}^{-1} M_{13} & M_{13}^{-1} M_{13} \\
M_{12} & M_{22} & M_{23} & M_{23} & M_{23} & M_{23} \\
M_{13} & M_{23} & M_{33} & M_{33} & M_{33} & M_{33} \end{pmatrix}$$

Using (1.3.9) and (1.3.16) we can now put $M_{33} = \frac{\lambda_3 \lambda_4 \lambda_5}{2}$, $M_{11} - M_{13} M_{33}^{-1} M_{13} = \Lambda_1 \Lambda_1'$, $M_{22} - M_{23} M_{33}^{-1} M_{23} = \Lambda_2 \Lambda_2'$, and $M_{12} - M_{13} M_{33}^{-1} M_{13} = \Lambda_1 (p \times q)(D/(C) 0) (r)$ $\Lambda_2' (q \times q)$. If we next put $M_{13} (p \times q) = \frac{\Lambda_3 (p \times q) \Lambda_1 (q \times r)}{2}$ and $M_{23} (q \times r) = \Lambda_4 (q \times r) \Lambda_2' (q \times r)$, we observe that $\Lambda_3$ and $\Lambda_4$ are determinate. We check furthermore that now $M_{13} M_{33}^{-1} M_{13} = \Lambda_3 \Lambda_3'$, $M_{23} M_{33}^{-1} M_{23} = \Lambda_4 \Lambda_4'$, and $M_{13} M_{33}^{-1} M_{13} = \Lambda_3 \Lambda_3'$, so that altogether we have
\[ M = \begin{pmatrix}
\lambda_1^{A_1^1 + A_3^1} & \lambda_1^{D/C} & 0 & \lambda_2^{A_2^1 + A_3^2} & \lambda_3^5
\\
\lambda_2^{D/C} & \lambda_1^{A_1^1 + A_3^2} & \lambda_2^{A_2^1 + A_4^1} & \lambda_4^5
\\
\lambda_3^5 & \lambda_5^1 & \lambda_5^1 & \lambda_5^1
\\
\lambda_5^1 & \lambda_5^1 & \lambda_5^1 & \lambda_5^1
\\
\end{pmatrix} \]

which proves \((A.3.20)\).

\((A.3.21)\) The passage from \(L\) matrices to \(L_T\) variables. Consider the transformations \((A.3.6), (A.3.8), (A.3.11), (A.3.14), (A.3.15), (A.3.17)\) and \((A.3.18, 19)\) and notice that everywhere we have, on the right hand side, a post factor of the form \(L(p \times n) \ (p \leq n)\) subject to the constraint \(LL' = I(p)\). Check that the actual number of independent constraints is just \(p(p + 1)/2\). Suppose now that instead of transforming to \(L\) subject to \(LL' = I(p)\), we take a slightly different set of variates in the following way. Putting

\[ L(p \times n) = \begin{pmatrix}
\lambda_{11} & \cdots & \lambda_{1n} \\
\vdots & \ddots & \vdots \\
\lambda_{pl} & \cdots & \lambda_{pn}
\end{pmatrix} = \begin{pmatrix}
\lambda_{11}' \\
\vdots \\
\lambda_{pl}'
\end{pmatrix} \]  

(say),

we notice that \(LL' = I(p) \iff \lambda_{ij}' \lambda_{ij} = \delta_{ij} \ (i, j = 1, 2, \ldots, p)\), so that, by virtue of the \(p(p + 1)/2\) constraints, \(L\) really consists of \(pn - p(p + 1)/2\) independent elements, although the \((p \times n)\) matrix itself is naturally one of \(pn\) elements. From \(L\) let us choose an independent set, say, \((\lambda_{11}', \lambda_{12}', \ldots, \lambda_{1,n-1}', \lambda_{21}', \lambda_{22}', \ldots, \lambda_{2,n-2}', \ldots, \lambda_{pl}', \lambda_{p2}', \ldots, \lambda_{p,n-p}'\) and let us call this set \(L_T\). Throughout this report \(L_T\) will stand uniformly for this set of variates.

\((A.3.22)\) It will now be shown that if no elements of \(L = 0\), then, the correspondence between \(L_T\) and \(L\) is one-to-one.
Proof. Having regard to the constraint \( LL' = I(p) \), under our set-up, we are going to treat \( \ell_{ij} \) \((i = 1, 2, \ldots, p; j = 1, 2, \ldots, n-1)\) (= \( \ell_{1} \) say) as the (so-called) independent variates and \( \ell_{ij} \) \((i = 1, 2, \ldots, p; j = n-i+1, \ldots, n)\) (= \( \ell_{0} \) say) as the (so-called) dependent variates. This notation will be uniformly followed. We have now the following equations in the dependent variates (in terms of the independent):

For the first row of the \( L \) matrix

\[(A.3.22.1) \quad \ell_{1n}^2 = 1 - \sum_{j=1}^{n-1} \ell_{1j}^2.\]

For the \( 2^{nd} \) row of the \( L \) matrix

\[(A.3.22.2) \quad \ell_{2,n-1} \ell_{1,n-1} + \ell_{2n} \ell_{1n} = -\sum_{j=1}^{n-2} \ell_{1j} \ell_{2j}; \quad \ell_{2,n-1}^2 + \ell_{2n}^2 = 1 - \sum_{j=1}^{n-2} \ell_{2j}^2.\]

And in general for the \( i^{th} \) row of the \( L \) matrix (with \( i = 1, 2, \ldots, p \))

\[(A.3.22.3) \quad \sum_{j=n-i+1}^{n-i} \ell_{ij} \ell_{i'j} = -\sum_{j=1}^{n-i} \ell_{ij} \ell_{i'j}; \quad \sum_{j=n-i+1}^{n-i} \ell_{ij}^2 = 1 - \sum_{j=1}^{n-i} \ell_{ij}^2,\]

for \( i' = 1, 2, \ldots, i-1 \).

It is easy to see that, for the first row of \( L \), the equation \((A.3.22.1)\) gives (in this case) two real and distinct values of \( \ell_{1n} \) in terms of \((\ell_{11}, \ldots, \ell_{1,n-1})\).

Next, for the second row of \( L \), the equations \((A.3.22.2)\) give (in this case) two real and distinct pairs of values for \((\ell_{2,n-1}, \ell_{2n})\) in terms of the first row (now supposed to be given), and so on. In general, for the \( i^{th} \) row of \( L \), the equations \((A.3.22.3)\) give (in this case) two real and distinct sets of values for \((\ell_{i,n-i+1}, \ldots, \ell_{in})\) in terms of the \((i-1)\) previous rows (now supposed to be given). This proves \((A.3.22)\).

A.4. Invariance of the characteristic roots under certain linear transformations.

\[(A.4.1) \quad \text{If } X(p \times n) \text{ (} p \leq n \text{) is of rank } p \text{ (in which case, by } (A.1.10), XX' \text{ is symmetric c.d.), then the characteristic roots of } XX' \text{ are invariant under the transformation: } X(p \times n) = \Lambda(p \times p)X(p \times n)B(n \times n), \text{ where } \Lambda \text{ and } B \text{ are any two } \times \text{ matrices.}\]
**Proof.** \( c(XY') = c(X'B'B'Y'Y') = c(Y'Y') \) (since \( B \) is \( \perp \)) = \( c(YY'Y') \) (using (A.1.18)) = \( c(YY') \) (since \( A \) is \( \perp \)), which completes the proof of (A.4.1).

(A.4.2) If \( X_1(p \times n_1) \), \( X_2(p \times n_2) \) (\( p \leq n_1, n_2 \)) are each of rank \( p \) (in which case, by (A.1.10), \( X_1X_1 \) and \( X_2X_2 \) are both symmetric p.d.), then the characteristic roots of \((X_1X_1'(X_2X_2)^{-1}) \) are invariant under the transformation: \( X_1(p \times n_1) = \lambda(p \times p)X_1(p \times n_1)B_1(n_1 \times n_1) \) and \( X_2(p \times n_2) = \lambda(p \times p)X_2(p \times n_2)B_2(n_2 \times n_2) \), where \( \lambda \) is any non-singular matrix and \( B_1 \) and \( B_2 \) any two \( \perp \) matrices.

**Proof.** \( c((X_1X_1')(X_2X_2)^{-1}) = c(\{A_{X_1}B_{X_1}Y_{X_1}Y_{X_1}\}) = c(\{A_{X_2}B_{X_2}Y_{X_2}Y_{X_2}\}) \) (since \( B_1 \) and \( B_2 \) are \( \perp \))

\[ = c(\{X_1X_1'(X_2X_2)^{-1}\}) \] (using (A.1.18)), which completes the proof of (A.4.2).

(A.4.3) If \( X_1(p \times n_1) \) be of rank \( n_1(p \perp p) \) and \( X_2(p \times n_2) \) of rank \( p \), then the characteristic roots of \((X_1X_1'(X_2X_2)^{-1}) \) are invariant under the transformation: \( X_1(p \times n_1) = \lambda(p \times p)X_1(p \times n_1)B_1(n_1 \times n_1) \) and \( X_2(p \times n_2) = \lambda(p \times p)X_2(p \times n_2)B_2(n_2 \times n_2) \), where \( \lambda \) is any non-singular matrix and \( B_1 \) and \( B_2 \) two arbitrary \( \perp \) matrices.

The proof is on the lines of that of (A.4.2) and is thus obvious.

(A.4.4) For \( X = \begin{pmatrix} X_1 & \end{pmatrix} \) \( (p \leq q, p + q \leq n, \text{rank } = p + q) \) the characteristic roots of \((X_1X_1'^{-1}(X_2X_2'^{-1}) \) are invariant under the transformation: \( X_1(p \times n) = \lambda_1(p \times p)X_1(p \times n)B(n \times n) \) and \( X_2(q \times n) = \lambda_2(q \times q)X_2(q \times n)B(n \times n) \), where \( \lambda_1 \) and \( \lambda_2 \) are any two non-singular matrices and \( B \) is any \( \perp \) matrix.

**Proof.** \( c((X_1X_1'^{-1}(X_2X_2'^{-1}) \) (since \( B \) is \( \perp \))

\[ = c((X_1X_1'^{-1}(X_2X_2'^{-1}) \) \] (A.1.18)
= c \left((Y_1 Y_1')^{-1}(Y_1 Y_2)(Y_2 Y_2')^{-1}(Y_2 Y_1')^{-1} A_1^{-1} \right)

= c \left((Y_1 Y_1')^{-1}(Y_1 Y_2)(Y_2 Y_2')^{-1}(Y_2 Y_1')^{-1} A_1^{-1} \right) (\text{using } (A.4.18))

= c \left((Y_1 Y_1')^{-1}(Y_1 Y_2)(Y_2 Y_2')^{-1}(Y_2 Y_1')^{-1} A_1^{-1} \right), \text{ which proves } (A.4.4).

\begin{align*}
(A.4.5) \quad & \text{For } X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \end{pmatrix}^p \\
& \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \end{pmatrix} \begin{pmatrix} q \\ q \\ q \\ \vdots \end{pmatrix}, \quad (p \leq q, p + q + r \leq n, \text{ rank } p + q + r), \\
& \text{the roots of the equation in } c \text{ of the form } (A.2.4.1), \text{ i.e., of }
\left| c \begin{pmatrix} X_1 X_1 & X_1 X_3 \\ X_3 X_1 & X_3 X_3 \end{pmatrix} & \begin{pmatrix} X_1 X_2 \\ X_3 X_2 \end{pmatrix} \\ \begin{pmatrix} X_2 X_1 & X_2 X_3 \\ X_3 X_1 & X_3 X_3 \end{pmatrix} & \begin{pmatrix} X_2 X_2 \\ X_3 X_2 \end{pmatrix} \right| = 0,
\end{align*}

\begin{align*}
\text{i.e., of }
\begin{align*}
(A.4.5.1) \quad & \left| c \begin{pmatrix} X_1 \\ X_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & \lambda_3 \\ \lambda_1 & \lambda_3 \end{pmatrix} \\ \begin{pmatrix} X_2 \\ X_3 \end{pmatrix} \begin{pmatrix} \lambda_2 & \lambda_3 \\ \lambda_2 & \lambda_3 \end{pmatrix} \right| = 0 \\
& \begin{pmatrix} X_1 \\ X_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & \lambda_3 \\ \lambda_1 & \lambda_3 \end{pmatrix} \\ \begin{pmatrix} X_2 \\ X_3 \end{pmatrix} \begin{pmatrix} \lambda_2 & \lambda_3 \\ \lambda_2 & \lambda_3 \end{pmatrix}
\end{align*}
\end{align*}

are invariant under the transformation
\begin{align*}
\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \end{pmatrix}^p = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = B(n \times n) \left( \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} \right),
\end{align*}

where \( B \) is \( \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \) is any non-singular matrix.

**Proof.** The proof by noting that \( B \) will pass out of the picture and the equation...
(4.4.5.1) can be written in terms of $Y$'s and $\Lambda$'s as

\[
\begin{pmatrix}
\Lambda_1 & \Lambda_3 \\
0 & \Lambda_5
\end{pmatrix}
\begin{pmatrix}
0 \\
0
\end{pmatrix}
\begin{pmatrix}
Y_1 \\
Y_3
\end{pmatrix}
= \begin{pmatrix}
Y_1 \\
Y_3
\end{pmatrix}
\begin{pmatrix}
Y_1' \\
Y_3'
\end{pmatrix}
\begin{pmatrix}
Y_2 \\
Y_3'
\end{pmatrix}
= 0,
\]

and also observing furthermore that, since $\Lambda$ is non-singular, \begin{pmatrix}
\Lambda_1 & 0 \\
0 & \Lambda_5
\end{pmatrix}
and
\begin{pmatrix}
\Lambda_2 & 0 \\
0 & \Lambda_5
\end{pmatrix}
are both easily checked to be non-singular.

A.5. Some general theorems in Jacobians.

(A.5.1) If $x(n \times l) = A(n \times n)y(n \times l)$, where $A$ is non-singular, then $J(x : y) = \det A$.

(A.5.2) If $X(m \times n) = A(m \times m)Y(m \times n)$, where $A$ is non-singular, then $J(X : Y) = \det A^n$.

(A.5.3) If $X(m \times n) = A(m \times m)Y(m \times n)B(n \times n)$, where $A$ and $B$ are non-singular, then $J(X : Y) = \det A^n \det B^m$.

(A.5.4) If $A$ and $B$ are each $l$, then $\det A = \det B = 1$ and (A.5.1) and (A.5.2) - (A.5.3) will reduce respectively to $J(x : y) = 1$ and $J(X : Y) = 1$.

(A.5.5) If $y_i = f_i(x_1, ..., x_m, x_{m+1}, ..., x_{m+n})$ $i = 1, ..., m$ where $x_j$'s $j = 1, 2, ..., m+n$ are subject to $n$ constraints

\[f_i(x_1, ..., x_m, x_{m+1}, ..., x_{m+n}) = 0 \quad (i = m+1, ..., m+n),\]

then (under the usual conditions for the existence of the Jacobian, including the non-vanishing of the numerator and the denominator in the following) we have, $J_{20}$,
\[ J(y_1, \ldots, y_m : x_1, \ldots, x_n) = \frac{\partial f_1, \ldots, f_m, f_{m+1}, \ldots, f_{m+n}}{\partial (x_1, \ldots, x_m, x_{m+1}, \ldots, x_{m+n})}. \]

**Proof.** Let us denote by \( \frac{\partial y_i}{\partial x_j} \), i, j = 1, \ldots, m, the partial differential coefficient of \( y_j \) with regard to \( x_j \) after having expressed \( y_i \) (i = 1, \ldots, m) in terms of \( (x_1, \ldots, x_m) \), that is, after eliminating \( (x_{m+1}, \ldots, x_{m+n}) \) with the help of the constraints. Next denote by

\[ \left| \frac{\partial y_i}{\partial x_j} \right|, \ i, j = 1, 2, \ldots, m, \]

the absolute value of the determinant of the \( m \times m \) (square) matrix

\[ \left( \frac{\partial y_i}{\partial x_j} \right), \ i, j = 1, 2, \ldots, m. \]

Then we have

\[ (\lambda.5.5.1) \quad J(y_1, \ldots, y_m : x_1, \ldots, x_n) = \left| \frac{\partial y_i}{\partial x_j} \right|, \ i, j = 1, 2, \ldots, m, \]

\[ = \left| \frac{\partial f_i}{\partial x_j} + \sum_{k=m+1}^{m+n} \frac{\partial f_i}{\partial x_k} \frac{\partial x_k}{\partial x_j} \right|, \ i, j = 1, 2, \ldots, m. \]

Notice that in \( \frac{\partial f_i}{\partial x_j} \) or \( \frac{\partial f_i}{\partial x_k} \), \( f_i \) is supposed to be expressed in terms of all the \( (m+n) \) \( x \)'s and the partial differentiation is supposed to be with regard to \( x_j \) or \( x_k \) assuming all the other \( (m+n-1) \) independent variates to be kept fixed, while in \( \frac{\partial y_i}{\partial x_j} \) or \( \frac{\partial x_k}{\partial x_j} \) it is supposed that \( y_i \) (i = 1, 2, \ldots, m) or \( x_k \) (k = m+1, \ldots, m+n) has first been expressed in terms of \( x_j \)'s (j = 1, 2, \ldots, m) and then the partial differentiation is made with respect to a particular \( x_j \), assuming the other \( (m-1) \) 'independent' variates to be kept fixed. Now from the set of n constraints on \( x_j \)'s (j = 1, 2, \ldots, m+n) given by the conditions of (\( \lambda.5.5 \)) we have

\[ (\lambda.5.5.2) \quad \frac{\partial f_i}{\partial x_j} + \sum_{k=m+1}^{m+n} \frac{\partial f_i}{\partial x_k} \frac{\partial x_k}{\partial x_j} = 0 \quad (i = m+1, \ldots, m+n, \text{ and } j = 1, \ldots, m), \]

or matrix wise,
\[ - \left( \frac{\partial f_i}{\partial x_j} \right) = \left( \frac{\partial f_i}{\partial x_k} \right) \left( \frac{\partial x_k}{\partial x_j} \right) \quad (i, k = m+1, \ldots, m+n; j = 1, \ldots, m), \] or

\[ \left( \frac{\partial x_k}{\partial x_j} \right) = -\left( \frac{\partial f_i}{\partial x_k} \right)^{-1} \left( \frac{\partial f_i}{\partial x_j} \right) \] (note that, by the conditions of \( \Lambda.5.5 \), \( \frac{\partial f_i}{\partial x_k} \) can be assumed to be non-singular). Substituting from \((\Lambda.5.5.3)\) in \((\Lambda.5.5.1)\) we have

\[ \begin{vmatrix} \frac{\partial f_i}{\partial x_k} & \frac{\partial f_i}{\partial x_j} & \cdots & \frac{\partial f_i}{\partial x_k} \end{vmatrix}_{k=m+1, \ldots, m+n}^{j=1, \ldots, m} \]

\[ = \begin{vmatrix} \frac{\partial f_i}{\partial x_k} & \frac{\partial f_i}{\partial x_j} & \cdots & \frac{\partial f_i}{\partial x_k} \\ i, j = 1, \ldots, m & k = m+1, \ldots, m+n \end{vmatrix} \begin{vmatrix} \frac{\partial f_i}{\partial x_k} & \frac{\partial f_i}{\partial x_j} \end{vmatrix}_{k=m+1, \ldots, m+n}^{j=1, \ldots, m} \]

\[ = \begin{vmatrix} \frac{\partial f_i}{\partial x_k} \\ k, j = m+1, \ldots, m+n \end{vmatrix} \begin{vmatrix} \frac{\partial f_i}{\partial x_j} \\ i, j = m+1, \ldots, m+n \end{vmatrix} \]

(by using \((\Lambda.1.1)\))

which proves \((\Lambda.5.5)\).

The real use of this theorem (as also the next one) is in those situations where it would be difficult to express \(y_i\)'s in terms of \(x_j\)'s (\(j = 1, \ldots, m\)) (after elimination of \(x_{m+1}, \ldots, x_{m+n}\) with the help of the constraints), but where it is much easier to express the right hand side of \((\Lambda.5.5.4)\) in terms of \((x_1, \ldots, x_m)\), or where even this explicit expression is not directly needed.

\[ \text{(\Lambda.5.6) If } F_i(y_1, \ldots, y_m, x_1, \ldots, x_m, x_{m+1}, \ldots, x_{m+n}) = 0 \quad (i = 1, 2, \ldots, m+n) \]
are a set of equations solvable in the real domain in the sense that corresponding to real \((x_1, \ldots, x_m)\) we can find real \((y_1, \ldots, y_n)\) and \((x_{m+1}, \ldots, x_{m+n})\), then, under the other usual conditions for the existence of the Jacobian (including the non-vanishing of the numerator and the denominator in the following), we have, \[ J(y_1, \ldots, y_n : x_1, \ldots, x_m) = \frac{\partial(F_1, \ldots, F_{m+n})}{\partial(x_1, \ldots, x_{m+n})} \div \frac{\partial(F_1, \ldots, F_{m+n})}{\partial(y_1, \ldots, y_n, x_{m+1}, \ldots, x_{m+n})} \]

Proof. As before we have

\[ J(y_1, \ldots, y_n : x_1, \ldots, x_m) = \left| \frac{\partial y_i}{\partial x_j} \right|_{i,j = 1, \ldots, m} \]

But from the basic conditions of (1.5.6) we have

\[ \sum_{i=1}^{m} \frac{\partial F_k}{\partial y_i} \frac{\partial y_i}{\partial x_j} + \sum_{j=m+1}^{m+n} \frac{\partial F_k}{\partial x_j} \frac{\partial x_j}{\partial x_j} + \frac{\partial F_k}{\partial x_j} = 0 \quad (k = 1, \ldots, m+n; j = 1, 2, \ldots, m) \]

Notice that in \( \frac{\partial F_k}{\partial y_i}, \frac{\partial F_k}{\partial x_j}, \frac{\partial F_k}{\partial x_j} \), \( F_k \) is supposed to be expressed in terms of all the \((2m+n)\) variates \((y_1, \ldots, y_n, x_1, \ldots, x_m, x_{m+1}, \ldots, x_{m+n})\), and the partial differentiation is with respect to \( y_i \) or \( x_j \) or \( x_{m+1} \), keeping all the other \((2m+n) - 1\) variates fixed.

Also notice that, in \( \frac{\partial y_i}{\partial x_j} \) or \( \frac{\partial x_j}{\partial x_j} \), \( y_i \) (or \( x_j \)) \((i = 1, 2, \ldots, m; j = m+1, \ldots, m+n)\) is supposed to be expressed in terms of \( x_j \)'s \((j = 1, 2, \ldots, m)\) and then the partial differentiation is with respect to a particular \( x_j \), keeping all the other \((m-1)\) of the \( x_j \)'s fixed.

Matrixwise, (1.5.6.1) could be written as

\[ \begin{pmatrix} \frac{\partial F_k}{\partial y_i} \\ \frac{\partial y_i}{\partial x_j} \end{pmatrix} + \begin{pmatrix} \frac{\partial F_k}{\partial x_j} \\ \frac{\partial x_j}{\partial x_j} \end{pmatrix} = 0, \]

where each side is a \((m+n) \times m\) matrix.

Taking, say, the first \( m \) rows of this matrix equation (1.5.6.2) we shall have the square \((m \times m)\) matrix equation:
\[ (i. 5.6.3) \quad \left( \frac{\partial F_{k_1}}{\partial y_1} \right) \left( \frac{\partial y_1}{\partial x_j} \right) + \left( \frac{\partial F_{k_1}}{\partial x_{\ell}} \right) \left( \frac{\partial x_{\ell}}{\partial x_j} \right) + \left( \frac{\partial F_{k_1}}{\partial x_j} \right) = 0 , \]

where now \( i, j, k_1 = 1, 2, \ldots, m \), and \( \ell = m+1, \ldots, m+n \), and \( \left( \frac{\partial F_{k_1}}{\partial y_1} \right) \) is square \((m \times m)\).

Again taking the last \( n \) rows of the matrix equation \((i. 5.6.2)\) we have

\[ (i. 5.6.4) \quad \left( \frac{\partial F_{k_2}}{\partial y_1} \right) \left( \frac{\partial y_1}{\partial x_j} \right) + \left( \frac{\partial F_{k_2}}{\partial x_{\ell}} \right) \left( \frac{\partial x_{\ell}}{\partial x_j} \right) + \left( \frac{\partial F_{k_2}}{\partial x_j} \right) = 0 , \]

where now \( i, j = 1, 2, \ldots, m \) and \( k_2, \ell = m+1, \ldots, m+n \), so that \( \left( \frac{\partial F_{k_2}}{\partial x_{\ell}} \right) \) is now square \((n \times n)\).

Treating \((i. 5.6.3)\) and \((i. 5.6.4)\) as a pair of simultaneous equations in \( \left( \frac{\partial y_1}{\partial x_j} \right) \) \((i, j = 1, \ldots, m)\) and \( \left( \frac{\partial x_{\ell}}{\partial x_j} \right) \) \((\ell = m+1, \ldots, m+n \text{ and } j = 1, \ldots, m)\), and solving for them, we have for \( \left( \frac{\partial y_1}{\partial x_j} \right) \) the following:

\[ (i. 5.6.5) \quad \left( \frac{\partial y_1}{\partial x_j} \right) = - \left[ \left( \frac{\partial F_{k_1}}{\partial y_1} \right) - \left( \frac{\partial F_{k_1}}{\partial x_{\ell}} \right) \left( \frac{\partial x_{\ell}}{\partial x_j} \right) \right]^{-1} \left( \frac{\partial F_{k_1}}{\partial x_j} \right) \]

Hence

\[ \left| \frac{\partial y_1}{\partial x_i} \right| = \left| \left( \frac{\partial F_{k_1}}{\partial x_j} \right) - \left( \frac{\partial F_{k_1}}{\partial x_{\ell}} \right) \left( \frac{\partial x_{\ell}}{\partial x_j} \right) \right| \div \left| \left( \frac{\partial F_{k_1}}{\partial y_1} \right) - \left( \frac{\partial F_{k_1}}{\partial x_{\ell}} \right) \left( \frac{\partial x_{\ell}}{\partial y_1} \right) \right| \times \left( \frac{\partial F_{k_2}}{\partial y_1} \right) \]

But we have by \((i. 1, 1)\),
\[
\begin{align*}
&\left(\begin{array}{c}
m \\
n
\end{array}\right) \left(\begin{array}{c}
\frac{\partial F_{k_1}}{\partial \varphi_j} \\
\frac{\partial F_{k_1}}{\partial \varphi'_l}
\end{array}\right) = \left(\begin{array}{c}
m \\
n
\end{array}\right) \left(\begin{array}{c}
\frac{\partial F_{k_2}}{\partial \varphi_j} \\
\frac{\partial F_{k_2}}{\partial \varphi'_l}
\end{array}\right) - \left(\begin{array}{c}
\frac{\partial F_{k_1}}{\partial \varphi_j} \\
\frac{\partial F_{k_2}}{\partial \varphi'_l}
\end{array}\right)^{-1} \\
&\times \left(\begin{array}{c}
\frac{\partial F_{k_1}}{\partial \varphi_j} \\
\frac{\partial F_{k_2}}{\partial \varphi'_l}
\end{array}\right)
\end{align*}
\]

and

\[
\begin{align*}
&\left(\begin{array}{c}
m \\
n
\end{array}\right) \left(\begin{array}{c}
\frac{\partial F_{k_1}}{\partial \psi_1} \\
\frac{\partial F_{k_1}}{\partial \psi'_l}
\end{array}\right) = \left(\begin{array}{c}
m \\
n
\end{array}\right) \left(\begin{array}{c}
\frac{\partial F_{k_2}}{\partial \psi_1} \\
\frac{\partial F_{k_2}}{\partial \psi'_l}
\end{array}\right) - \left(\begin{array}{c}
\frac{\partial F_{k_1}}{\partial \psi_1} \\
\frac{\partial F_{k_2}}{\partial \psi'_l}
\end{array}\right)^{-1} \\
&\times \left(\begin{array}{c}
\frac{\partial F_{k_2}}{\partial \psi_1} \\
\frac{\partial F_{k_2}}{\partial \psi'_l}
\end{array}\right)
\end{align*}
\]

Substituting from (1.5.6,6) in (1.5.6,5) we have

\[
(1.5.6.7) \quad J(y_1, \ldots , y_m : x_1, \ldots , x_n) = \left| \frac{\partial \psi_1}{\partial \varphi_j} \right|
\]

\[
\begin{align*}
&\left(\begin{array}{c}
m \\
n
\end{array}\right) \left(\begin{array}{c}
\frac{\partial F_{k_1}}{\partial \psi_j} \\
\frac{\partial F_{k_1}}{\partial \psi'_l}
\end{array}\right) = \frac{\partial F_{k_2}}{\partial \psi_1} \left(\begin{array}{c}
m \\
n
\end{array}\right) \left(\begin{array}{c}
\frac{\partial F_{k_2}}{\partial \psi_1} \\
\frac{\partial F_{k_2}}{\partial \psi'_l}
\end{array}\right) \\
&\quad + \frac{\partial F_{k_1}}{\partial \psi_1} \left(\begin{array}{c}
m \\
n
\end{array}\right) \left(\begin{array}{c}
\frac{\partial F_{k_1}}{\partial \psi_1} \\
\frac{\partial F_{k_1}}{\partial \psi'_l}
\end{array}\right)
\end{align*}
\]

\[
\begin{align*}
&= \frac{\partial (F_1, \ldots , F_{m+n})}{\partial (x_1, \ldots , x_m, x_{m+1}, \ldots , x_{m+n})} + \frac{\partial (F_1, \ldots , F_{m+n})}{\partial (y_1, \ldots , y_m, x_{m+1}, \ldots , x_{m+n})}
\end{align*}
\]

which proves (1.5.6).

(1.5.5) is really a special case of (1.5.6), which can be shown by putting in

\[(1.5.5), \quad F_i = y_i - f_i(x_1, \ldots , x_{m+n}) \quad (i = 1, 2, \ldots , m) \quad \text{and next} \quad F_i = f_i(x_1, \ldots , x_{m+n}).\]
\(x_{m+n}, \quad \ldots, m+n)\), that is, by assuming that the last \(n\) equations are free from the \(y_i\)'s. Substituting in the right hand side of (A.5.6.7), we easily check that it goes over into the right hand side of (A.5.5.4).

It seems that (A.5.6) is a very general theorem in Jacobians and yields as special cases practically all the usual well-known Jacobian theorems.

### 3.6. Jacobians of certain specific transformations.

We shall consider the transformations (A.3.6), (A.3.8), (A.3.11) with rank = \(p\), (A.3.14), (A.3.15), (A.3.17) and (A.3.18.19) and, in each case, pass on to \(L_1\) from the postfactor and prefactor of the form \(L \times M\) (subject to \(LL' = I\)) and discuss, for the different cases, the respective Jacobians 

(i) \(J(X; M_1, c's, L_1)\), 
(ii) \(J(X_1, X_2; A, c's, L_{11}, L_{21})\), 
(iii) \(J(X; \tilde{T}, L_1)\), 
(iv) \(J(X_1, X_2; U_1, U_2, \tilde{U}_3, U_4, c's, L_{11}, L_{21})\), 
(v) \(J(X_1, X_2; \tilde{T}, c's, L_1, L_{11}, L_{21})\), 
(vi) \(J(X_1, X_2; \tilde{T}, U, c's, \tilde{M}_{11}, \tilde{M}_{21}, L_{21})\) and 
(vii) \(J(X_1, X_2; A, B_1, B_2, B_3, B_4, c's, L_1)\), where, in (vii), \(L_1\)'s are respectively the (so-called) independent elements formed, as in section (A.3.21), out of the matrices

\[
\begin{pmatrix}
L_1 \\
L_3 \\
L_4
\end{pmatrix}
\begin{pmatrix}
p \\
p \\
q-p
\end{pmatrix}
\begin{pmatrix}
\begin{bmatrix}
L_1 \\
L_3 \\
L_4
\end{bmatrix}
\end{pmatrix}
\begin{pmatrix}
p \\
p \\
q-p
\end{pmatrix}
\end{array}
\]

We shall first obtain the following two Jacobians which will be basic to the derivations of all the other ones.

(A.6.1) \(\text{Jacobian of the transformation (A.3.11) (with rank = } \pi\text{) i.e., } J(X; \tilde{T}, L_1)\)

where \(\tilde{T}(p \times p)\) is non-singular with a positive diagonal. To obtain the Jacobian from \(X\) to \(\tilde{T}\) and \(L_1\) we use (A.5.5), remembering that now \(X = \tilde{T}L\) takes the place of \(y_i = f_i\) and \(LL' - I(p) = 0\) takes the place of \(f_1 = 0\). We also note that \(d(LL' - I(p)) = d(LL')\). We have now using (A.5.5)

(A.6.1.1) \(J(X; \tilde{T}, L_1) = \left| \frac{\partial (X, LL')}{\partial (\tilde{T}, L)} \right| \left| \frac{\partial (LL')}{\partial (L, L')} \right| = \left| \frac{\partial (X, LL')}{\partial (\tilde{T}, L)} \right| \left| \frac{\partial (LL')}{\partial (L)} \right| \left| \frac{\partial (LL')}{\partial (L')} \right| \left| \frac{\partial (LL')}{\partial (L')} \right| L_1\)
where on the extreme right, for practical usability, everything is expressed in terms of \( T \) and \( L \). To calculate the numerator of (A.6.1.1) we proceed as follows.

\[
\begin{align*}
X &= \begin{pmatrix}
  x_{11} & \cdots & x_{1n} \\
  \vdots & \ddots & \vdots \\
  x_{p1} & \cdots & x_{pn}
\end{pmatrix} = \begin{pmatrix}
  x_{1}^{l} \\
  \vdots \\
  x_{p}^{l}
\end{pmatrix} \quad \text{(say)};
L &= \begin{pmatrix}
  \ell_{11} & \cdots & \ell_{1n} \\
  \vdots & \ddots & \vdots \\
  \ell_{p1} & \cdots & \ell_{pn}
\end{pmatrix} = \begin{pmatrix}
  \ell_{1} \\
  \vdots \\
  \ell_{p}
\end{pmatrix} \quad \text{(say)}.
\end{align*}
\]

Also put \( LL' = K \) with elements \( k_{ij} \) \((i, j = 1, 2, \ldots, p; k_{ij} = k_{ji})\). Then (A.3.11) can be written as \( x_{i}^{l} = (t_{11} \ldots t_{ii} \ldots 0 \ldots 0) \times L \) \((i = 1, 2, \ldots, p)\) or

\[
\begin{pmatrix}
  t_{11} \\
  \vdots \\
  t_{ii} \\
  0 \\
  \vdots \\
  0
\end{pmatrix} = \begin{pmatrix}
  \ell_{11} & \ell_{12} & \cdots & \ell_{1p} \\
  \ell_{21} & \ell_{22} & \cdots & \ell_{2p} \\
  \vdots & \vdots & \ddots & \vdots \\
  \ell_{p1} & \ell_{p2} & \cdots & \ell_{pp}
\end{pmatrix}
\]

\((i = 1, 2, \ldots, p)\).

To calculate \( \frac{\partial (X, LL')}{\partial (T, L)} = \frac{\partial (X, K)}{\partial (T, L)} \), we display below the partial differential coefficients of \( X \) and \( K \) \((= LL')\) with respect to the elements of \( T \) and \( L \) (all elements of \( L \) being temporarily regarded as independent for purposes of the present differentiation):

\[
\begin{pmatrix}
  \frac{\partial x_{11}}{\partial t_{11}} & \frac{\partial x_{11}}{\partial t_{12}} & \cdots & \frac{\partial x_{11}}{\partial t_{ii}} & \frac{\partial x_{11}}{\partial t_{ii}} & \cdots & \frac{\partial x_{11}}{\partial t_{i1}} & \frac{\partial x_{11}}{\partial t_{i2}} & \cdots & \frac{\partial x_{11}}{\partial t_{ip}} \\
  \frac{\partial x_{12}}{\partial t_{11}} & \frac{\partial x_{12}}{\partial t_{12}} & \cdots & \frac{\partial x_{12}}{\partial t_{ii}} & \frac{\partial x_{12}}{\partial t_{ii}} & \cdots & \frac{\partial x_{12}}{\partial t_{i1}} & \frac{\partial x_{12}}{\partial t_{i2}} & \cdots & \frac{\partial x_{12}}{\partial t_{ip}} \\
  \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
  \frac{\partial x_{1n}}{\partial t_{11}} & \frac{\partial x_{1n}}{\partial t_{12}} & \cdots & \frac{\partial x_{1n}}{\partial t_{ii}} & \frac{\partial x_{1n}}{\partial t_{ii}} & \cdots & \frac{\partial x_{1n}}{\partial t_{i1}} & \frac{\partial x_{1n}}{\partial t_{i2}} & \cdots & \frac{\partial x_{1n}}{\partial t_{ip}} \\
  \frac{\partial x_{21}}{\partial t_{11}} & \frac{\partial x_{21}}{\partial t_{12}} & \cdots & \frac{\partial x_{21}}{\partial t_{ii}} & \frac{\partial x_{21}}{\partial t_{ii}} & \cdots & \frac{\partial x_{21}}{\partial t_{i1}} & \frac{\partial x_{21}}{\partial t_{i2}} & \cdots & \frac{\partial x_{21}}{\partial t_{ip}} \\
  \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
  \frac{\partial x_{2n}}{\partial t_{11}} & \frac{\partial x_{2n}}{\partial t_{12}} & \cdots & \frac{\partial x_{2n}}{\partial t_{ii}} & \frac{\partial x_{2n}}{\partial t_{ii}} & \cdots & \frac{\partial x_{2n}}{\partial t_{i1}} & \frac{\partial x_{2n}}{\partial t_{i2}} & \cdots & \frac{\partial x_{2n}}{\partial t_{ip}} \\
  \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
  \frac{\partial x_{pn}}{\partial t_{11}} & \frac{\partial x_{pn}}{\partial t_{12}} & \cdots & \frac{\partial x_{pn}}{\partial t_{ii}} & \frac{\partial x_{pn}}{\partial t_{ii}} & \cdots & \frac{\partial x_{pn}}{\partial t_{i1}} & \frac{\partial x_{pn}}{\partial t_{i2}} & \cdots & \frac{\partial x_{pn}}{\partial t_{ip}}
\end{pmatrix}
\]

\[
\begin{pmatrix}
  \frac{\partial \ell_{11}}{\partial t_{11}} & \frac{\partial \ell_{11}}{\partial t_{12}} & \cdots & \frac{\partial \ell_{11}}{\partial t_{ii}} & \frac{\partial \ell_{11}}{\partial t_{ii}} & \cdots & \frac{\partial \ell_{11}}{\partial t_{i1}} & \frac{\partial \ell_{11}}{\partial t_{i2}} & \cdots & \frac{\partial \ell_{11}}{\partial t_{ip}} \\
  \frac{\partial \ell_{12}}{\partial t_{11}} & \frac{\partial \ell_{12}}{\partial t_{12}} & \cdots & \frac{\partial \ell_{12}}{\partial t_{ii}} & \frac{\partial \ell_{12}}{\partial t_{ii}} & \cdots & \frac{\partial \ell_{12}}{\partial t_{i1}} & \frac{\partial \ell_{12}}{\partial t_{i2}} & \cdots & \frac{\partial \ell_{12}}{\partial t_{ip}} \\
  \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
  \frac{\partial \ell_{1n}}{\partial t_{11}} & \frac{\partial \ell_{1n}}{\partial t_{12}} & \cdots & \frac{\partial \ell_{1n}}{\partial t_{ii}} & \frac{\partial \ell_{1n}}{\partial t_{ii}} & \cdots & \frac{\partial \ell_{1n}}{\partial t_{i1}} & \frac{\partial \ell_{1n}}{\partial t_{i2}} & \cdots & \frac{\partial \ell_{1n}}{\partial t_{ip}} \\
  \frac{\partial \ell_{21}}{\partial t_{11}} & \frac{\partial \ell_{21}}{\partial t_{12}} & \cdots & \frac{\partial \ell_{21}}{\partial t_{ii}} & \frac{\partial \ell_{21}}{\partial t_{ii}} & \cdots & \frac{\partial \ell_{21}}{\partial t_{i1}} & \frac{\partial \ell_{21}}{\partial t_{i2}} & \cdots & \frac{\partial \ell_{21}}{\partial t_{ip}} \\
  \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
  \frac{\partial \ell_{2n}}{\partial t_{11}} & \frac{\partial \ell_{2n}}{\partial t_{12}} & \cdots & \frac{\partial \ell_{2n}}{\partial t_{ii}} & \frac{\partial \ell_{2n}}{\partial t_{ii}} & \cdots & \frac{\partial \ell_{2n}}{\partial t_{i1}} & \frac{\partial \ell_{2n}}{\partial t_{i2}} & \cdots & \frac{\partial \ell_{2n}}{\partial t_{ip}} \\
  \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
  \frac{\partial \ell_{pn}}{\partial t_{11}} & \frac{\partial \ell_{pn}}{\partial t_{12}} & \cdots & \frac{\partial \ell_{pn}}{\partial t_{ii}} & \frac{\partial \ell_{pn}}{\partial t_{ii}} & \cdots & \frac{\partial \ell_{pn}}{\partial t_{i1}} & \frac{\partial \ell_{pn}}{\partial t_{i2}} & \cdots & \frac{\partial \ell_{pn}}{\partial t_{ip}}
\end{pmatrix}
\]
\[
\begin{array}{cc|cccc}
& t_{11} & t_{21} & \cdots & t_{p1} & t_{22} & \cdots & t_{p2} & \cdots & t_{pp} \\
\hline
x_1 & \xi_1 & 0 & 0 & 0 & 0 & \cdots & 0 & D_{t11} & 0 & \cdots & 0 \\
x_2 & 0 & \xi_1 & 0 & \xi_2 & 0 & \cdots & 0 & D_{t21} & D_{t22} & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
x_p & 0 & \cdots & \xi_1 & 0 & \xi_2 & \cdots & \xi_p & D_{tp1} & D_{tp2} & \cdots & D_{tp,p-1} & D_{tp,p} \\
\hline
k_{11} & & & & & & & & & & & & \\
\vdots & & & & & & & & & & & & \\
k_{1p} & & & & & & & & & & & & \\
k_{12} & 0 & & & & & & & & & & & \\
\vdots & & & & & & & & & & & & \\
k_{1p} & & & & & & & & & & & & \\
\vdots & & & & & & & & & & & & \\
k_{p-1,p} & & & & & & & & & & & & \\
\end{array}
\]

where \( D_{a} \) will stand for a diagonal matrix with diagonal elements all equal to \( a \). Recall that \( x_i' \) is \( 1 \times n \), \( \xi_i' \) is also \( 1 \times n \) \((i = 1, \ldots, p)\) and \( K(p \times p) \) has \( p(p+1)/2 \) independent elements so that the above is really a \((np+p(p+1)/2) \times (np+p(p+1)/2)\) matrix. Now put

\[(4.6.1.3) \quad M_{11}(p(p+1)/2 \times p(p+1)/2) = 0; \quad M_{12}(p(p+1)/2 \times np) =
\begin{pmatrix}
\xi_1' & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \xi_p' & \xi_p' \\
\xi_1' & \xi_1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \xi_p' & \xi_{p-1}' \\
\end{pmatrix};
\]
\[
M_{21}(np \times p(p+1)/2) = \begin{pmatrix}
\xi_1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \xi_1 & 0 & \xi_2 & 0 & 0 & \cdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \xi_1 & 0 & \xi_2 & \xi_p
\end{pmatrix}
\]

and

\[
M_{22}(np \times np) = \begin{pmatrix}
D_{t_11} & 0 & 0 & 0 \\
D_{t_{12}} & D_{t_{22}} & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots \\
D_{t_{p1}} & D_{t_{p2}} & \cdots & D_{t_{pp}}
\end{pmatrix}
\]

(notice that each \(D \) is \( n \times n \).

By (A.1.1) we shall now have

\[
\frac{\partial (X,K)}{\partial (T,L)} = 2^p \begin{vmatrix}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{vmatrix} np = \frac{p(p+1)}{2} np
\]

Recalling the structure of \( \tilde{T} \), we have

\[
\tilde{T}^{-1} = \begin{pmatrix}
t_{11} & 0 & 0 \\
t_{12} & t_{22} & 0 \\
\cdots & \cdots & \cdots \\
t_{1p} & t_{2p} & t_{pp}
\end{pmatrix}
\]

so that \( M_{22} = |\tilde{T}|^{-n} \). We have furthermore

\[
M_{22}^{-1}M_{21} = \begin{pmatrix}
\xi_1 t_{11} & 0 & 0 & 0 & 0 & 0 \\
\xi_1 t_{12} & \xi_1 t_{22} & 0 & \xi_1 t'_{22} & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\xi_1 t_{1p} & \xi_1 t_{2p} & \xi_1 t'_{p2} & \xi_1 t'_{pp} & \xi_1 t'_{pp} & \xi_1 t'_{pp}
\end{pmatrix}
\]
\[ \begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
t_{11} & 0 \\
\vdots & \vdots \\
t_{1p} & t_{pp} \\
0 & \cdots & 0 \\
0 & \cdots & 0 \\
\end{pmatrix}
\begin{pmatrix}
p \\
p-1 \\
0 \\
0
\end{pmatrix}
= M_{21} N_{22} \quad \text{(say)} \text{ where } N_{22} \text{ stands for the right matrix factor. We note that } N_{22} \text{ is } p(n+1)/2 \times p(p+1)/2 \text{ and is non-singular if } T' \text{ is non-singular. We note also that } M_{12}^{-1} (p(p+1)/2 \times p) M_{21} (p(n) \times p(p+1)/2) \text{ is } p(p+1)/2 \times p(p+1)/2 \text{ and non-singular, so that}
\]
\[ (A.6.1.6) \quad \left| M_{12} M_{21}^{-1} M_{21} \right| = \left| M_{12} M_{21}^{-1} \right| \left| N_{22} \right|. \]

It is easy to check that
\[ (A.6.1.7) \quad \left| N_{22} \right| = \prod_{i=1}^{p} t_{ii}^{-1} / \left| T \right|^n \quad \text{and} \quad \left| T \right| = \prod_{i=1}^{p} t_{ii}. \]

It is also easy to verify, by using the condition \( LL' = I(p) \), that
\[ (A.6.1.8) \quad \left| M_{12} M_{21} \right| LL' = I(p) = 1. \]

Hence \( (A.6.1) \) will now reduce to \( (10,11,7) \),
\[ (A.6.1.9) \quad J(X; T', L') = \left| \frac{\partial (X, LL')}{\partial (T', L)} \right|_{T', L'} = \left| \frac{\partial (LL')}{\partial (L')} \right|_{L'} = 2^p \prod_{i=1}^{p} t_{ii}^{n-i} / \left| \frac{\partial (LL')}{\partial (L')} \right|_{L'}, \]
so that we have
\[ (A.6.1.10) \quad dX \rightarrow J(X; T', L') \, dT' \, dL' \] when \( J \) is given by \( (A.6.1.9) \).

It is easy to check that, with \( nS = T'T' \), we have
\[ (A.6.1.11) \quad J(S; T') = 2^p \prod_{i=1}^{p} t_{ii}^{n-i+1} / n^{(p+1)/2}, \]
so that

\[(A.6.1.12) \quad d\mathbf{T} \rightarrow \sum_{i=1}^{n} \left( p+1 \right) / 2 \prod_{j=1}^{p} t_{i}^{-i+1} ds. \]

\[(A.6.2) \quad \text{Jacobian of the transformation (A.3.6), i.e., } J(X_1, X_2 : \lambda, c's, L_{11}, L_{21}), \text{ where } X_1(\lambda \times n_1), X_2(\lambda \times n_2) (p \leq n_1, n_2) \text{ are each of rank } p, \text{ c's are distinct, and } \lambda \text{ is solid } \lambda \times p \text{ non-singular with a positive first row. Putting } \]

c_{i}^{1/2} = t_{i} (i = 1, 2, \ldots, p) \text{ and using (A.1.1) we have}

\[(A.6.2.1) \quad J(X_1, X_2; \lambda, t's, L_{11}, L_{21}) = \left| \frac{\partial (X_1, X_2, L_{11}, L_{21})}{\partial (\lambda, t's, L_{11}, L_{21})} \right|_{L_{11}, L_{21}} \]

\[\frac{1}{2} \left| \frac{\partial (L_{11})}{\partial (L_{11})} \right|_{L_{11}} \times \left| \frac{\partial (L_{21})}{\partial (L_{21})} \right|_{L_{21}}. \]

To evaluate the numerator we proceed as follows. Denote, as before, the row vectors of \(L_1, L_2, X_1, X_2\) by \(L_{11}, L_{21}, X_{11}, X_{21}\) \((i = 1, 2, \ldots, p)\) and \(L_{11}'\) by \((k_{1i1})\) and \(L_{21}'\) by \((k_{2i1})\). Then the transformation can be written as

\[(A.6.2.2) \quad x_{1i} = (L_{1i} \cdots L_{1p}) \begin{pmatrix} a_{1i} & t_{1} \\ \vdots \\ a_{ip} & t_{p} \end{pmatrix}; \quad x_{2i} = (L_{2i} \cdots L_{2p}) \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ip} \end{pmatrix}\]

\((i = 1, 2, \ldots, p), \text{ or in full}\)
The scheme of partial differentiation is given below.

<table>
<thead>
<tr>
<th></th>
<th>(a^t)</th>
<th>(t^t)</th>
<th>(\mathcal{L}_1^t)</th>
<th>(\mathcal{L}_2^t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_1)</td>
<td>((\mathcal{L}_3^t, t))</td>
<td>((a_1^t, a_3))</td>
<td>((a_1^t))</td>
<td>0</td>
</tr>
<tr>
<td>(x_2)</td>
<td>((\mathcal{L}_4^t))</td>
<td>0</td>
<td>0</td>
<td>((a_2))</td>
</tr>
<tr>
<td>(x_{1}^t)</td>
<td>0</td>
<td>0</td>
<td>((\mathcal{L}_5^t))</td>
<td>0</td>
</tr>
<tr>
<td>(x_{2}^t)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>((\mathcal{L}_6^t))</td>
</tr>
</tbody>
</table>

where \(a^t = (a_{11} \cdot a_{12} \cdot a_{1p} \cdot a_{pp})\), \(t^t = (t_1 \cdot t_p)\), \(\mathcal{L}_1^t = (\mathcal{L}_{11}^t \cdot \mathcal{L}_{12}^t \cdot \mathcal{L}_{1p}^t \cdot \mathcal{L}_{pp}^t)\), 
\(\mathcal{L}_2^t = (\mathcal{L}_{21}^t \cdot \mathcal{L}_{22}^t \cdot \mathcal{L}_{2p}^t \cdot \mathcal{L}_{pp}^t)\), 
\(\mathcal{L}_3^t = \begin{pmatrix} x_{11}^t \\ x_{1p}^t \end{pmatrix}\), 
\(\mathcal{L}_4^t = \begin{pmatrix} x_{21}^t \\ x_{2p}^t \end{pmatrix}\).

\[ k_1 = \begin{pmatrix} k_{111} \\ k_{1pp} \\ k_{112} \\ k_{11p} \\ k_{1p-1,p} \end{pmatrix}, \quad k_2 = \begin{pmatrix} k_{211} \\ k_{2pp} \\ k_{212} \\ k_{21p} \\ k_{2p-1,p} \end{pmatrix}, \quad (\mathcal{L}_3^t, t) = \begin{pmatrix} \mathcal{L}_{11}^{t1} \cdot \mathcal{L}_{1p}^{tp} \cdot 0 \cdot 0 \\ \cdots \cdots \cdots \cdots \cdots \\ 0 \cdots \mathcal{L}_{11}^{t1} \cdot \mathcal{L}_{1p}^{tp} \end{pmatrix}, \]

\[ (a_1^t, \mathcal{L}_3^t) = \begin{pmatrix} a_{11} \mathcal{L}_{11}^t \cdot a_{1p} \mathcal{L}_{1p}^t \\ \cdots \cdots \cdots \cdots \cdots \\ a_{1p} \mathcal{L}_{11}^t \cdot a_{pp} \mathcal{L}_{pp}^t \end{pmatrix}, \quad (a_1^t, t) = \begin{pmatrix} D_{a_{11}}^{t1} (n_1) \cdots D_{a_{1p}}^{tp} (n_1) \\ \cdots \cdots \cdots \cdots \cdots \\ D_{a_{1p}}^{t1} (n_1) \cdots D_{a_{pp}}^{tp} (n_1) \end{pmatrix}, \]

\[ (\mathcal{L}_4^t) = \begin{pmatrix} \mathcal{L}_{21} \cdot \mathcal{L}_{2p} \cdot 0 \cdot 0 \\ \cdots \cdots \cdots \cdots \cdots \\ 0 \cdots \mathcal{L}_{21} \cdot \mathcal{L}_{2p} \end{pmatrix}, \quad (a_2^t) = \begin{pmatrix} D_{a_{11}} (n_2) \cdots D_{a_{1p}} (n_2) \\ \cdots \cdots \cdots \cdots \cdots \\ D_{a_{1p}} (n_2) \cdots D_{a_{pp}} (n_2) \end{pmatrix}, \]

\[ (\mathcal{L}_5^t) = \begin{pmatrix} \mathcal{L}_{11}^t \cdot 0 \cdot 0 \\ \cdots \cdots \cdots \cdots \cdots \\ 0 \cdots \mathcal{L}_{1p} \cdot \mathcal{L}_{1p}^t \end{pmatrix}, \quad (\mathcal{L}_6^t) = \begin{pmatrix} \mathcal{L}_{21}^t \cdot 0 \cdot 0 \\ \cdots \cdots \cdots \cdots \cdots \\ 0 \cdots \mathcal{L}_{2p} \cdot \mathcal{L}_{2p}^t \end{pmatrix}. \]
we are interested in the absolute value of the determinant of the above matrix 
(which is really the numerator in the Jacobian) and which is:
\(\{ p^2 + p + (n_1 + n_2)p \} \times \{ p^2 + p \} \). After some obvious manipulations we can take out a factor

\( p^2 \prod_{i=1}^{n_1-p} t_i \)

so that we have the whole determinant reducing to

\[
(\lambda_6, 2, 3) \quad 2^p \prod_{i=1}^{p} t_i \quad \begin{vmatrix}
M_{11} & M_{12} & M_{13} & 0 \\
M_{21} & 0 & 0 & M_{24} \\
0 & 0 & M_{33} & 0 \\
0 & 0 & 0 & M_{44}
\end{vmatrix}
\begin{pmatrix}
p n_1 \\
p n_2 \\
p (n_1 + 1)/2 \\
p (n_1 + 1)/2
\end{pmatrix},
\text{where}

\[
M_{11}(p n_1 \times p^2) = \begin{pmatrix}
\mathcal{L}_{11} t_1 \cdot \mathcal{L}_{1p} t_p & 0 & 0 \\
\vdots & \ddots & \vdots \\
0 & \mathcal{L}_{11} t_1 & \mathcal{L}_{1p} t_p
\end{pmatrix};
M_{21}(p n_2 \times p^2) = \begin{pmatrix}
\mathcal{L}_{21} \cdot \mathcal{L}_{2p} & 0 & 0 \\
\vdots & \ddots & \vdots \\
0 & \mathcal{L}_{21} & \mathcal{L}_{2p}
\end{pmatrix}
\]

\[
M_{1p}(p n_1 \times p^2) = \begin{pmatrix}
\mathcal{L}_{11} & a_{1p} \mathcal{L}_{1p} \\
\vdots & \ddots \\
\mathcal{L}_{11} & a_{1p} \mathcal{L}_{1p}
\end{pmatrix};
M_{13}(p n_1 \times p n_1) = \begin{pmatrix}
a_{11} (n_1) \cdot D_{a_{1p}} (n_1) \\
\vdots & \ddots \\
a_{11} \cdot D_{a_{1p}}
\end{pmatrix};
\]

\[
M_{24}(p n_2 \times p n_2) = \begin{pmatrix}
a_{11} (n_2) \cdot D_{a_{1p}} (n_2) \\
\vdots & \ddots \\
a_{11} \cdot D_{a_{1p}}
\end{pmatrix};
\]

\[
M_{33}(p (n_1 + 1)/2 \times p n_1) = \begin{pmatrix}
\mathcal{L}_{11} t_1 & 0 & 0 \\
0 & \mathcal{L}_{11} t_1 & \mathcal{L}_{1p} t_p \\
0 & \mathcal{L}_{11} t_1 & \mathcal{L}_{1p} t_p
\end{pmatrix}
\]
\[
\hat{\mathcal{H}}_{\mu_4} \left( \frac{x(x+1)}{2} \right) \times \text{pm}_2 = \begin{pmatrix}
\mathcal{L}^1_{21} & 0 & 0 & 0 \\
\mathcal{L}^1_{22} & \mathcal{L}^1_{21} & \cdots & 0 \\
\mathcal{L}^1_{23} & \mathcal{L}^1_{22} & \mathcal{L}^1_{21} & \cdots & 0 \\
0 & \cdots & \mathcal{L}^1_{24} & \mathcal{L}^1_{23} & \cdots & \mathcal{L}^1_{22} & \mathcal{L}^1_{21}
\end{pmatrix}
\]

Hence we should have

\[
\begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} & \mathcal{M}_{13} & 0 \\
\mathcal{M}_{21} & 0 & 0 & \mathcal{M}_{24} \\
0 & \mathcal{M}_{33} & 0 & 0 \\
0 & 0 & \mathcal{M}_{44} & 0 \end{pmatrix} = \begin{pmatrix} \mathcal{M}_{13} & 0 \\
0 & \mathcal{M}_{24} \\
0 & \mathcal{M}_{34} & 0 & \mathcal{M}_{44} \end{pmatrix} = \begin{pmatrix} \mathcal{M}_{13} & 0 \\
0 & \mathcal{M}_{24} \\
0 & \mathcal{M}_{34} & \mathcal{M}_{44} \end{pmatrix}
\]

It is now easy to check that

\[
\begin{pmatrix} \mathcal{M}_{13} \\
\mathcal{M}_{24} \\
\mathcal{M}_{34} & \mathcal{M}_{44} \end{pmatrix} = \mathcal{L}^1_{12} \mathcal{N}_{12} \mathcal{N}_{24} \mathcal{L}^1_{23} \mathcal{L}^1_{22} \mathcal{L}^1_{21}, \quad \begin{pmatrix} \mathcal{M}_{13} \\
\mathcal{M}_{24} \\
\mathcal{M}_{34} & \mathcal{M}_{44} \end{pmatrix} = \mathcal{L}^1_{12} \mathcal{N}_{12} \mathcal{N}_{24} \mathcal{L}^1_{23} \mathcal{L}^1_{22} \mathcal{L}^1_{21}
\]

and \( \mathcal{M}_{24} \) is exactly of this form each \( a \) being of \( n_2 \) dimensions. Hence we shall have

\[
\begin{pmatrix} \mathcal{M}_{13}^{-1} \mathcal{M}_{11} \\
\mathcal{M}_{24}^{-1} \mathcal{M}_{21} \\
\mathcal{M}_{34}^{-1} \mathcal{M}_{32} & \mathcal{M}_{44}^{-1} \mathcal{M}_{43} \end{pmatrix} = \begin{pmatrix} \mathcal{D}_{11}(n_1) & \mathcal{D}_{12}(n_1) \\
\mathcal{D}_{21}(n_1) & \mathcal{D}_{22}(n_1) \\
\mathcal{D}_{31}(n_1) & \mathcal{D}_{32}(n_1) \end{pmatrix}
\]

\[
= \begin{pmatrix} \hat{\mathcal{M}}_{11} & \mathcal{D}_{12}(n_1) \\
\mathcal{D}_{21}(n_1) & \mathcal{D}_{22}(n_1) \\
\mathcal{D}_{31}(n_1) & \mathcal{D}_{32}(n_1) \end{pmatrix}
\]

where we denote the right hand matrix factor by \( \mathcal{D} \). In an exactly similar manner we have \( \mathcal{M}_{24}^{-1} \mathcal{M}_{21} = \mathcal{M}_{21} \mathcal{D} \). Next we have
(A.6.2.7) \[ M^{-1}_{13}M_{12} = \begin{pmatrix} k_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & k_{1p} \end{pmatrix} = M_{1p} \text{(say)} \] (using \(WW^{-1} = I(p)\)). Thus we have

\[
M^{-1}_{13}(M_{11}^t;M_{12}) = (M_{11}D;N_{12}) = p^n_1(M_{11}^t;M_{12}) \begin{pmatrix} D & 0 \\ 0 & I \end{pmatrix} \frac{p^2}{p^2} \frac{p^2}{p}
\]

and \(M^{-1}_{24}(M_{21};0) = p^n_2(M_{21};0) \begin{pmatrix} D & 0 \\ 0 & I \end{pmatrix} \frac{p^2}{p} \frac{p^2}{p}\), so that (A.6.2.4) now reduces to

\[
\begin{vmatrix} A \end{vmatrix} = \begin{vmatrix} \frac{1}{A} \end{vmatrix} = \begin{vmatrix} \frac{M_{11} \times \frac{M_{11}}{M_{11}}; \frac{M_{11}}{M_{11}} \times \frac{N_{12}}{N_{12}}}{} & D & 0 \\ \frac{M_{11} \times \frac{M_{11}}{M_{11}}; \frac{M_{11}}{M_{11}} \times \frac{N_{12}}{N_{12}}}{} & 0 & I(p) \end{vmatrix} \]

\[
= \begin{vmatrix} A \end{vmatrix} \begin{vmatrix} \frac{M_{11} \times \frac{M_{11}}{M_{11}}; \frac{M_{11}}{M_{11}} \times \frac{N_{12}}{N_{12}}}{} & D & 0 \\ \frac{M_{11} \times \frac{M_{11}}{M_{11}}; \frac{M_{11}}{M_{11}} \times \frac{N_{12}}{N_{12}}}{} & 0 & I(p) \end{vmatrix} \] (Remembering that \(|D| = |A|^{-p}\) and \(|I(p)| = 1\). Using \(L_1L_1^t = L_2L_2^t = I(p)\) the structure and reduction of the 2nd factor (which is a determinant) can now be displaced and visualized by considering \(p=3\) which will make immediately obvious the corresponding structure and mechanism of reduction for the general case. Below is given the case of \(p=3\).

\[
\begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & t_2^2 & 0 & t_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & t_2^2 & 0 & 0 & 0 & t_1^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & t_2^3 & 0 & t_1^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{vmatrix} = \text{mod}(t_1^2-t_2^2)(t_1^2-t_3^2)(t_2^2-t_3^2). \]
In the general case this is easily checked to be replaceable by mod \( \prod_{i<j=1}^{p^2} (t_i-t_j) \)
so that substituting in (A.6.2.3) and noting that \( t_i = c_i \) we have

\[
(A.6.2.10) \quad \left| \frac{\partial(X_1, X_2, L_1 L_1 L_2 L_2)}{\partial(\Lambda, \tau, L_1 L_2)} \right|_{\Lambda, \tau, L_1 L_2} = 2^p \prod_{i=1}^{p-1} \frac{n_1-p}{t_i} \frac{n_1+n_2-p}{\mod \prod_{i<j=1}^{p^2} (t_i-t_j)} ,
\]

so that

\[
(A.6.2.11) \quad J(\tau, X_2; \Lambda, c, L_1 L_2) = \left| \frac{\partial(X_1, X_2, L_1 L_1 L_2 L_2)}{\partial(\Lambda, c, L_1 L_2)} \right| \div \left| \frac{\partial(L_1 L_2)}{\partial(L_{1D})} \right| = 2^p \prod_{i=1}^{n_1+n_2-p} \frac{n_1-p}{\prod_{i=1}^{n_2-p} (c_i - c_j)} \div \left| \frac{\partial(L_1 L_2)}{\partial(L_{1D})} \right| \left| \frac{\partial(L_{1L})}{\partial(L_{2D})} \right|_{L_1 L_2} \left| \frac{\partial(L_{1L})}{\partial(L_{2D})} \right|_{L_2} \left| \frac{\partial(L_{1L})}{\partial(L_{2D})} \right|_{L_2 L_1} .
\]

It may be noticed now that (A.6.2.11) is the Jacobian (ii) mentioned in the beginning of (A.6), i.e., \( J(X_1, X_2; \Lambda, c, L_1 L_2) \) and (A.6.1.9) is the Jacobian (iii) mentioned there, i.e., \( J(X_1, L_1); \Lambda \) 10, 11 7.

(A.6.3) Jacobian of the transformation (A.3.6), i.e., \( J(X; L, c, s, L_1) \), where \( X(p \times n) (p < n) \) is of rank \( p \), \( c, s \)'s are distinct and \( M \) has a positive first row.

By straightforward methods of exactly the kind used in the preceding subsection (A.6.1) which is rather lengthy it can be shown that

\[
(A.6.3.1) \quad J(X; L, c, s, L_1) = 2^p \prod_{i=1}^{n-p-1} \frac{n-p}{c_i} \frac{p}{\prod_{i<j=1}^{n-p} (c_i - c_j)} \div \left| \frac{\partial(M L_1)}{\partial(L_1)} \right| \left| \frac{\partial(L_1)}{\partial(L_2)} \right|_{L_1} \left| \frac{\partial(L_1)}{\partial(L_2)} \right|_{L_2} .
\]

But a shorter proof of this result can be given by combining (A.6.1.9) and (A.6.2.11) in the following way. By (A.6.2.11) we have

\[
(A.6.3.2) \quad J(X_1, X_2; \Lambda, c, s, L_1 L_2) = 2^p \prod_{i=1}^{n_1+n_2-p} \frac{n_1-p}{t_i} \frac{n_1+n_2-p}{\mod \prod_{i<j=1}^{p^2} (t_i-t_j)} \div \left| \frac{\partial(L_1 L_1)}{\partial(L_{1D})} \right| \left| \frac{\partial(L_2 L_1)}{\partial(L_{2D})} \right|_{L_1} \left| \frac{\partial(L_2 L_1)}{\partial(L_{2D})} \right|_{L_2} \left| \frac{\partial(L_2 L_1)}{\partial(L_{2D})} \right|_{L_1} ,
\]

where \( X_1(p \times n) = \Lambda D L_1 \) and \( X_2(p \times n_2) = \Lambda L_2 \). Also, using (A.3.11), we put
\((\text{A.6.3.3)}\) \(A(p \times n) = \tilde{r}(p \times n)M(p \times p)\) where \(M\) is \(\perp\), and using \((\text{A.6.1.9})\), we have
\[(\text{A.6.3.4)}\) \(J(\Lambda; \tilde{T}, M_1) = 2^p \prod_{i=1}^{p} t_{ii}^{-1} / \left| \frac{\partial (M_1)}{\partial (M_{\tilde{T}})} \right|_{M_1} \). Next put
\[(\text{A.6.3.5)}\) \(x_1(p \times n) = \tilde{r}(p \times p)X(p \times n), M(p \times p)L_2(p \times n_2) = M_2(p \times n_2)\) (say) so that \(X = MD_{L_2}X_2 = \tilde{M}_2\) and note from orthogonality of \(M\) that
\[(\text{A.6.3.6)}\) \(|\Lambda| = |\tilde{T}| = \prod_{i=1}^{p} t_{ii} \) and \(M_2M_1 = ML_2L_2^\perp = I(p)\). We thus have
\[(\text{A.6.3.7)}\) \(J(x_1, x_2; \Lambda, c's, L_1, L_2) = J(x_1, x_2; \tilde{T}, M_1, c's, L_1, L_2) / J(\Lambda; \tilde{T}, M_1) \)
\(= J(x_1, x_2) J(x_1, x_2; \tilde{T}, M_1, c's, L_1, L_2) / J(\Lambda; \tilde{T}, M_1) \)
\(= J(x_1, x_2) J(x_1; M_1, c's, L_1) J(x_2; \tilde{T}, M_1) / J(\Lambda; \tilde{T}, M_1) \).
Now notice that
\[(\text{A.6.3.8)}\) \(J(x_1; \tilde{T}) = \prod_{i=1}^{p} t_{ii}^{-n}, J(x_2; \tilde{T}, M_1) = 2^p \prod_{i=1}^{p} t_{ii}^{n_2-1} / \left| \frac{\partial (M_1)}{\partial (M_2)} \right|_{M_2} \),
and \(J(\Lambda; \tilde{T}, M_1) = 2^p \prod_{i=1}^{p} t_{ii}^{-1} / \left| \frac{\partial (M_1)}{\partial (M_{\tilde{T}})} \right|_{M_1} \).
Now to evaluate \(J(M_2; L_2)\) we temporarily regard \(\Lambda\) as a constant but \(\perp\) matrix, notice that \(L_2L_2^\perp = I(p)\) is equivalent to \(M_2M_2^\perp = I(p)\) and now using \((\text{A.5.5)}\) we find
\[(\text{A.6.3.9)}\) \(J(M_2; L_2) = \left| \frac{\partial (M_2 - ML_2, M_2^\perp)}{\partial (M_2, L_2)} \right| / \left| \frac{\partial (M_2 - ML_2, M_2^\perp)}{\partial (M_2, L_2)} \right| \)
\(= \left| \frac{\partial (M_2 - ML_2, M_2^\perp)}{\partial (M_2, L_2)} \right| / \left| \frac{\partial (M_2 - ML_2, M_2^\perp)}{\partial (M_2, L_2)} \right| \).
Now substituting in the left hand side of \((\text{A.6.3.7)}\) from \((\text{A.6.3.2)}\) and \((\text{A.6.3.6)}\) and in the right hand side from \((\text{A.6.3.8)}\) and \((\text{A.6.3.9)}\) and putting \(L_1 = L\) (say), we have the Jacobian \((\text{A.6.3.1)}\).
\[(\text{A.6.4)}\) <br>\text{Jacobian of the transformation} \((\text{A.3.15})\), i.e., \(J(x_1, x_2; \tilde{T}, c's, L_1, L_1, L_2)\) where \(\tilde{T}\) is non-singular with a positive diagonal and \(c's\) are distinct. Using
\[(\text{A.6.1.9)}\) we have \(J(x_2; \tilde{T}, L_2) = 2^p \prod_{i=1}^{p} t_{ii}^{n_2-1} / \left| \frac{\partial (L_2L_2^\perp)}{\partial (L_2L_2^\perp)} \right| \). Next, using \((\text{A.6.3.1)}\) we
\[ J(x_1^{p-1}, L_1, c_i s, L_{II}) = 2^n \prod_{i=1}^{n-1} \frac{p-n_i-1}{c_i} \mod \prod_{i<j=1}^{n-1} (c_i - c_j) \cdot \frac{\frac{\partial (L_{II})}{\partial \langle L \rangle_{II}}}{\frac{\partial (L_1)}{\partial \langle L \rangle_1}} . \]

From these it is easy to check that

\[ (\lambda.6.4.1) \ J(x_1 x_2, c_i s, L_1, L_{II}, L_2) = 2^{n_1+n_2-1} \prod_{i=1}^{n_1} t_{ii} \prod_{i=1}^{n_1} \frac{p-n_i-1}{c_i} \]

\[ \times \mod \frac{n_i-1}{\prod_{i<j=1}^{n_i} (c_i - c_j)} \frac{\frac{\partial (L_{II})}{\partial \langle L \rangle_{II}}}{\frac{\partial (L_1)}{\partial \langle L \rangle_1}} \frac{\frac{\partial (L_2)}{\partial \langle L \rangle_2}}{\frac{\partial (L_2)}{\partial \langle L \rangle_2}} . \]

\[ (\lambda.6.5) \ Jacobian \ of \ the \ transformation \ (\lambda.3.17), \ i.e., \ J(x_1 x_2; \bar{T}, U, c_i s, M_{II}, M_{I}, L_{II}, L_2), \] when the c's are distinct, \( \bar{T} \) is non-singular with a positive diagonal and \( U \) is non-singular solid with a positive first row. Using (\lambda.6.19) we have \( J(x_2; T, L_2) \)

\[ = 2^q \prod_{i=1}^{n-1} t_{ii}^{n-1} / \frac{\partial (L_{II})}{\partial \langle L \rangle_{II}} \]

Next, notice that \( J(x_1; L_1, L_2) = 1 \) (since \( L_1, L_2 \) is \( L \)). Next, using (\lambda.6.2.12), we have \( J(x_1 x_2; L_1 L_2; U, c_i s, M_{II}, M_{I}) \)

\[ = 2^n \prod_{i=1}^{n-p} t_{ii}^{n-p} \prod_{i=1}^{n-p} (e_i - e_j) / \frac{\partial (M_{II})}{\partial \langle M \rangle_{II}} \frac{\partial (M_{I})}{\partial \langle M \rangle_{I}} . \]

It is easy to check, by combining the three Jacobians, that

\[ (\lambda.6.5.1) \ J(x_1 x_2; U, e_i s, M_{II}, M_{I}, L_{II}, L_2) = 2^{p+q} \prod_{i=1}^{n-1} t_{ii}^{n-1} \prod_{i=1}^{n-p} e_i^{n-p} \prod_{i=1}^{n-q-p-1} \frac{\partial (M_{II})}{\partial \langle M \rangle_{II}} \frac{\partial (M_{I})}{\partial \langle M \rangle_{I}} \frac{\partial (L_{II})}{\partial \langle L \rangle_{II}} \frac{\partial (L_2)}{\partial \langle L \rangle_2} . \]

We recall from (\lambda.3.17) that if \( e_i = (c_i - 1)/c_i \) \( (i=1, \ldots, p) \) where \( c_i \)'s are the roots of the equation in \( c: |c(x_1 x_1)(x_2 x_2) - 1(x_2 x_2)| = 0 \). In terms of the c's therefore, we should have the Jacobian given by
(6.5.2) \( J(x_1, x_2, U, c', s, M_{11}, M_{21}, L_{21}) = 2^{p+q} \frac{d}{\prod_{i=1}^{q} t_{1i}^{n-i}} |U|^{n-p} \times \prod_{i=1}^{n-p-q} \frac{n-p-q-1}{c_i} \mod \prod_{i<j=1}^{n-q+2} (c_i - c_j) \div \frac{\partial (M_{11})}{\partial (L_{11})} \left| M_{11} \right| \frac{\partial (M_{21})}{\partial (L_{21})} \left| M_{21} \right| \frac{\partial (L_{21})}{\partial (L_{21})} \left| L_{21} \right| \). 

(\text{A.6.6}) \text{ Jacobian of the transformation (A.3.15), i.e., } J(x_1, x_2, x_3; z_{11}, z_{12}, z_{21}, z_{22})

\[ Z_{22}, T, L_{31} \]. \text{ Using (A.6.1.9) we have } J(x_3; T, L_{31}) = 2^{r} \left( \frac{r}{n-i} \right) \left| \frac{\partial (L_{33})}{\partial (L_{D})} \right| \left| L_{31} \right| \).

Next we notice that \( J(x_1, x_2, z_{11}, z_{12}, z_{21}, z_{22}) = J \left( \begin{pmatrix} x_1 \\ x_2 \\ z_{11} \\ z_{12} \\ L \end{pmatrix}, \begin{pmatrix} z_{21} \\ z_{22} \\ L_3 \end{pmatrix} \right) = 1, \) since \( \left( \begin{pmatrix} L \\ L_3 \end{pmatrix} \right) \text{ is } \perp. \) Therefore it is easily checked that the total Jacobian

\[ = 2^{r} \frac{r}{n-i} \left| \frac{\partial (L_{33})}{\partial (L_{D})} \right| \left| L_{31} \right| \).

(\text{A.6.7}) \text{ Jacobian of the transformation (A.3.11), i.e., } J(x_1, x_2, U, U_2, U_3, U_4, c', s, L_{11}, L_{21}), \text{ where } x_1(p \times n_1) (p > n_1) \text{ is of rank } n_1, x_2(p \times n_2) (p < n_2) \text{ is of rank } p, \text{ the } c' \text{ s are distinct and } U_1 \text{ has a positive first row and } \tilde{U}_3 \text{ a positive diagonal.}

\text{We start with the transformation (A.3.15) and use the Jacobian result (A.6.4.1) and rename the symbols. The transformation is } x_1(p \times n_1) = \tilde{T}(p \times p) L(p \times n_1) \frac{t_1}{c_1(n_1 \times n_1)} x_1(n_1 \times n_1), x_2(p \times n_2) = \tilde{T}((p \times p) M_2(p \times n_2) \text{ subject to } L_1 \text{ being } \perp, K_{21} = I(p), L' = I(n_1) \text{ and } \tilde{T} \text{ being non-singular, and the Jacobian being given by}

(\text{A.6.7.1}) \ J(x_1, x_2, T, c', s, L_1, L_{11}, L_{21}) = 2^{p+n_1} \frac{n_1}{\prod_{i=1}^{n_1} t_{1i}^{n_1-n_2+i}} \prod_{i=1}^{n_1-n_2+i} \left( c_i - c_j \right) \div \frac{\partial (L_{11})}{\partial (L_D)} \left| L_{11} \right| \frac{\partial (L_{21})}{\partial (L_{21})} \left| L_{21} \right| \frac{\partial (M_{21})}{\partial (M_{21})} \left| M_{21} \right| \). \text{ Let us write}

\[ = 2^{p+n_1} \frac{n_1}{\prod_{i=1}^{n_1} t_{1i}^{n_1-n_2+i}} \prod_{i=1}^{n_1-n_2+i} \left( c_i - c_j \right) \div \frac{\partial (L_{11})}{\partial (L_D)} \left| L_{11} \right| \frac{\partial (L_{21})}{\partial (L_{21})} \left| L_{21} \right| \frac{\partial (M_{21})}{\partial (M_{21})} \left| M_{21} \right| \). \]
\[ L(p \times n_1) = \begin{pmatrix} K_1 \\ K_2 \\ \vdots \\ K_p \end{pmatrix}_{n_1}^{p-n_1}. \]

To this \( L \) now, if we adjoin, as we could, a matrix \( \begin{pmatrix} K_3 \\ K_4 \end{pmatrix}_{n_1}^{p-n_1} \) such that

\[ \kappa(p \times p) = \begin{pmatrix} K_1 & K_3 \\ K_2 & K_4 \end{pmatrix}_{n_1}^{p-n_1} \]

is orthogonal (note that this could be done since \( L' L = I(n_1) \)), it will be seen that the number of independent elements in \( K \) is the same as in \( L \). This is verified as follows.

In \( L \) (by virtue of \( L' L = I(n_1) \)), the number of independent elements are

\[ p n_1 - n_1(n_1+1)/2. \]

In \( K \) the total number of elements are \( p^2 - (p-n_1)(p-n_1-1)/2 \) and by virtue of \( K K' = I(p) \), the number of constraints are \( p(p+1)/2 \), so that the number of independent elements are \( p^2 - (p-n_1)(p-n_1-1)/2 - p(p+1)/2 = p n_1 - n_1(n_1+1)/2 \). If we now put

\[ U(p \times p) = \begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix}_{n_1}^{p-n_1} \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}_{n_1}^{p-n_1} \begin{pmatrix} K_1 & K_3 \\ K_2 & K_4 \end{pmatrix}_{n_1}^{p-n_1} \]

(by examining the right hand side we note that the left hand side is really of the structure indicated), we observe that the number of independent elements in \( U \) which is \( p^2 - (p-n_1)(p-n_1-1)/2 \), is the same as in \((T, K)\), i.e., as in \((\tilde{T}, K_1, K)\), which is \( p(p+1)/2 + p n_1 - n_1(n_1+1)/2 \). It will be shown in the next article (and we assume the result here) that

\[ J(U; \tilde{T}, K_1) = J(U; T, L_I) = 2 \sum_{i=1}^{n_1} \left( \begin{array}{c} p-n_1 \\ 3 \\ \vdots \\ p-1 \\ i \\ 1 \\ 1 \end{array} \right) \frac{\partial (L^{(l_1, L, d)})}{\partial (L_I)} \bigg|_{L_I} \]

so that, by taking the inverse, we should have
(6.6.7.4) \( J(T, L_1; U) = J(T, L_1; U) = \prod_{i=1}^{n_1} \left( u_{3ii} \right)^{-n_1-1} \left| \frac{\partial (L_1L)}{\partial (T)} \right|_{L_1} \times 2 \prod_{i=1}^{n_1} \frac{n}{t_{ii}} \) 

(6.6.7.5) Also if we put \( L_2(p \times n_2) = K(p \times p)M_2(p \times n_2) \) (where by virtue of \( KK^T = K^TK = M_2M_2^T = I(p) \) we have \( L_2L_2^T = I(p) \)), then exactly as in (6.6.3) (treating \( K \) as a constant \( \perp \) matrix) we have

\[ J(M_2^T; L_2^T) = \left| \frac{\partial (M_2^T)}{\partial (M_2)} \right|_{M_2^T} \times \left| \frac{\partial (L_2^T)}{\partial (L_2)} \right|_{L_2^T} \] 

Thus we have

\[ J(T, L_1; M_2^T; U, L_2^T) = \left| \frac{\partial (M_2^T)}{\partial (M_2)} \right|_{M_2^T} \times \left| \frac{\partial (L_1L)}{\partial (T)} \right|_{L_1} \prod_{i=1}^{n_1} \left( u_{3ii} \right)^{-n_1-1} \]

\[ \times \frac{2^p \prod_{i=1}^{n} \frac{n}{t_{ii}} \left| \frac{\partial (L_2^T)}{\partial (L_2)} \right|_{L_2^T}}{ U } \]

Using these and remembering that \( |U| = |T| = \prod_{i=1}^{n} t_{ii} \), we have

\[ J(X_1, X_2; c, U, L_1L, L_2L) = J(X_1, X_2; c, T, L_1L, L_2L, M_2^T; U, L_2L) \]

\[ = 2^p \left| \frac{\partial (L_1L)}{\partial (T)} \right|_{L_1L} \times \frac{\prod_{i=1}^{n_1+n_2-p} \frac{n_1}{c_i}}{2^{p-n_1-1}} \times \prod_{i<j=1}^{n_1} \frac{n_1}{(c_i - c_j)} \frac{n_1}{(u_{3ii})^{p-n_1-1}} \times \frac{\partial (L_1L)}{\partial (T)} \] 

Now for the proof of (6.6.7.3) with a transformation of the form (4.3.14) we proceed as follows. We start from (6.6.7.5), post multiply both sides by the \( p \times p \) matrix

\[ \left( \begin{array}{ccc} 1 & 0 \\ 0 & M \end{array} \right)^{n_1} \]

where \( M \) is \( \perp \) and then write

\[ U = \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right) \left( \begin{array}{ccc} 1 & 0 \\ 0 & M \end{array} \right)^{n_1} = \left( \begin{array}{c} U_1 \\ U_2 \end{array} \right) \left( \begin{array}{cc} V_1 & V_3 \\ V_2 & V_4 \end{array} \right) \] (say) = \( V \) (say)
\[ \begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix} = T_N \text{ (say). Then using (A.6.1.9) and (A.6.5.5) and (A.6.6.5) we have} \]

\[ J(V:U:K) = 2^n \sum_{i=1}^{P} t_{ii} \left| \frac{\partial (NN')} {\partial (N)} \right|_N \times \left| \frac{\partial (MM')} {\partial (M)} \right|_M, \]

and also

\[ J(V:U:K) = 2^n \sum_{i=1}^{n} (u_{3ii})^{n-1} \left| \frac{\partial (LL')} {\partial (L)} \right|_L. \]

Taking account of the remarks after (A.6.6.1) it is easy to check that

\[ \left| \frac{\partial (KK')} {\partial (K)} \right|_K = \left| \frac{\partial (LL')} {\partial (L)} \right|_L. \]

Now combining (A.6.7.9), (A.6.7.10), (A.6.7.11) and (A.6.7.12), we have

\[ dV \rightarrow 2^n \sum_{i=1}^{n} (u_{3ii})^{n-1} dU dM / \left| \frac{\partial (MM')} {\partial (M)} \right|_M, \]

which proves (A.6.7.3).

(A.6.8) Jacobian of the transformation (A.3.18.19), i.e.,

\[ J(X_1, X_2, \Lambda, B_1, B_2, B_3, B_4, c's, L), \]

where the c's are distinct, \( \Lambda \) is non-singular with a positive first row.

\( B_3 \) has a positive first row and \( B = \begin{pmatrix} \Lambda & 0 \\ B_2 & B_4 \end{pmatrix} \) is non-singular. This Jacobian can be derived in the same manner as in sub-section (A.6.2). We shall not need it in this report and so will not derive it. We merely state without proof that

\[ J(X_1, X_2, \Lambda, B, c's, L) = 2^n \left| \begin{array}{c} n-p \\ B \end{array} \right|^q \left( \frac{\partial B}{\partial \Lambda} \right)^{q-p} \times \left( \frac{\partial B}{\partial \Lambda} \right)^{q-p} \mod \left( c_i - c_j \right) / \left| \frac{\partial (LL')} {\partial (L)} \right|_L, \]

where \( L = \begin{pmatrix} L_1 \\ L_3 \\ L_4 \end{pmatrix} \) and is subject to \( LL' = I(p+q), \] 10.7.
A.7. Canonical reduction of certain distribution problems.

(A.7.1) If \( X(p \times n) \) (\( p \leq n \)) has the probability law (5.13): \( \left\{ \frac{p^n}{(2\pi)^{n/2} \Gamma(n/2)} \right\} \)
\( x \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} XX' \right\} \) \( dX \), then the distribution of the characteristic roots of \( XX' \) (to be called \( c \)'s) could not involve as parameters anything except the characteristic roots of \( \Sigma \) (to be called \( \gamma \)'s).

Proof. Note that, a.e., \( XX' \) is p.d. so that, a.e., all roots \( c(XX') \) are positive. Notice also that, a.e., they are also distinct. It is of course assumed that \( \Sigma \) is symmetric p.d., so that all \( c(\Sigma)'s \), i.e., \( \gamma \)'s are positive. Using (A.3.3), set \( \Sigma = \mu D_\mu' \), where \( \mu \) is \( \perp \). We have now \( \text{tr} \Sigma^{-1} XX' = \text{tr}(\mu D_\mu')^{-1} XX' \)
\( = \text{tr} D_{1/\gamma} \mu' XX' \mu D_{1/\gamma} \) (using (A.1.5) and the orthogonality of \( \mu \)). Now put \( \mu'X = Y \) or \( X(p \times n) = (p \times p)Y(p \times n) \) and observe that, by (A.4.1), \( c(XX') = c(YY') \).

Also by (A.5.2), \( J(X; Y) = |\mu|^n = 1 \). Remembering further that \( |\Sigma| = |\mu|^2 \sum_{i=1}^{p} \gamma_i \), it is easy to check that \( Y \) has the probability law:

\[
(A.7.1.1) \quad \left[ \frac{1}{1/(2\pi)^{p/2} \Gamma(p/2)} \right] \exp \left\{ -\frac{1}{2} \text{tr} D_{1/\gamma} YY' \right\} \) \( dY \)

which, in view of the fact that \( c(XX') = c(YY') \), proves (A.7.1). For the distribution of \( c(XX') \), therefore, we can, without any loss of generality, start directly from the above form of probability law which is accordingly a canonical law for this purpose.

(A.7.2) If \( X_1(p \times n_1), X_2(p \times n_2) \) (\( p \leq n_1, n_2 \)) have the joint probability law:
\[
\left[ \frac{1/(2\pi)^{(n_1+n_2)/2}} {\Gamma((n_1+n_2)/2)} \right] \exp \left\{ -\frac{1}{2} \text{tr} (\Sigma_1^{-1} X_1 X_1' + \Sigma_2^{-1} X_2 X_2') \right\} dX_1 dX_2 ,
\]
\( \Sigma_1 \) and \( \Sigma_2 \) being each symmetric p.d., then the distribution of \( c((X_1 X_1')(X_2 X_2'))^{-1} \) (to be called \( c \)'s) could not involve as parameters anything except the \( c(\Sigma_1^{-1} \Sigma_2^{-1})'s \) (to be called \( \gamma \)'s).
Proof. Notice that, a.e., $c(X_1^{-1}X_2^{-1})$ are positive and distinct. Since $\Sigma_1$ and $\Sigma_2$ are each p.d., use (5.3.4) to set $\Sigma_1 = \mu \Sigma_1^{-1}$ and $\Sigma_2 = \mu \Sigma_2^{-1}$, where $\mu$ is non-singular and all $\gamma$'s are positive. We have now, using (5.1.5),

$$\text{tr } \Sigma_1^{-1}X_1X_1^{-1} = \text{tr } D_{1/\gamma}^{-\frac{1}{2}}X_1X_1^{-1}D_{1/\gamma}^{-\frac{1}{2}}$$

and $\text{tr } \Sigma_2^{-1}X_2X_2^{-1} = \text{tr } \mu^{-\frac{1}{2}}X_2X_2^{-1}\mu^{-\frac{1}{2}}$. Now put $\mu^{-\frac{1}{2}}X_1 = Y_1$ and $\mu^{-\frac{1}{2}}X_2 = Y_2$, i.e., $X_1(p \times n_1) = \mu(p \times p)Y_1(p \times n_1)$ and $X_2(p \times n_2) = \mu(p \times p)Y_2(p \times n_2)$ and observe that, by (5.1.2), $c(X_1^{-1}X_2^{-1}) = c(Y_1^{-1}Y_2^{-1})$. Also, by

$$\frac{n_1n_2}{n_1+n_2}$$

(5.5.2),

$$J(X_1, X_2 ; Y_1, Y_2) = |\mu|^{-\frac{n}{n_1+n_2}}.$$  Remememring further that $|\Sigma_1| = |\mu|^\frac{n}{n_1+n_2}$ and $|\Sigma_2| = |\mu|^\frac{n}{n_1+n_2}$, we check that $Y_1$ and $Y_2$ have the joint probability law:

$$(\text{5.7.1}) \left[ \frac{1}{(2\pi)^{\frac{n}{2}}} \prod_{i=1}^{n} \frac{1}{\gamma_i^{\nu_i}} \right] \times \exp \left\{ -\frac{1}{2} \text{tr} \left( D_{1/\gamma}^{-\frac{1}{2}}Y_1Y_1^{-1} + Y_2Y_2^{-1} \right) \right\} dY_1 dY_2,$$

which, in view of the fact that $c(X_1^{-1}X_2^{-1}) = c(Y_1^{-1}Y_2^{-1})$, proves (5.7.2).

If we are interested in the distribution of these roots, i.e., of the $c$'s, we can, without any loss of generality, start right away from the above form which, for the purpose of this problem, will thus be called a canonical distribution law.

(5.7.3) If $X = \left( \begin{array}{c} X_1 \\ X_2 \end{array} \right)$ has the probability law (5.15):

$$\frac{1}{(2\pi)^{\frac{n}{2}}} \left| \Sigma \right|^{-\frac{n}{2}} |\mu|^{-\frac{n}{n_1+n_2}} \exp \left\{ -\frac{1}{2} \text{tr} \left( D_{1/\gamma}^{-\frac{1}{2}}Y_1Y_1^{-1} + Y_2Y_2^{-1} \right) \right\} dX_1 dX_2,$$

where $\Sigma = \left( \begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{array} \right)$ is supposed to be symmetric p.d., then the distribution of $c\sqrt{(X_1^{-1}X_2^{-1})(X_1^{-1}X_2^{-1})}^{-1} (X_1^{-1}X_2^{-1})^{-1}$, (to be called $c$'s) could not involve as parameters anything except $c(\Sigma^{-1}X_1^{-1}X_2^{-1}X_1^{-1})$ (to be called $\gamma$'s).

Proof. Notice that, a.e., the $p$ $c$'s are positive and distinct and also that $\gamma$'s are all non-negative. Use (5.3.16) to set $\Sigma_1(p \times p) = \mu_1(p \times p)$, $\mu_1(p \times p)$,
\[ \Sigma'_{\omega}(q \times q) = \mu_2(q \times q) \Sigma_{12}(q \times q) \] and \[ \Sigma_{12}(q \times q) = \mu_1(q \times p)(D_{\gamma})^{0}(p)\mu_2(q \times q), \]

when \( \mu_1 \) and \( \mu_2 \) are non-singular. We have now

\[ (A.7.3.1) \quad \Sigma^{-1} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \mu_1^{-1} & 0 \\ 0 & \mu_2^{-1} \end{pmatrix} \begin{pmatrix} I(p) & (D_{\gamma})^{0}(p) \\ (D_{\gamma})^{0}(p) & I(q) \end{pmatrix}^{-1} \begin{pmatrix} \mu_1^{-1} & 0 \\ 0 & \mu_2^{-1} \end{pmatrix}. \]

Also

\[ (A.7.3.2) \quad \begin{pmatrix} I(p) & (D_{\gamma})^{0}(p) \\ (D_{\gamma})^{0}(p) & I(q) \end{pmatrix} = \begin{pmatrix} I(p) & (D_{\gamma})^{0}(p) \\ (D_{\gamma})^{0}(p) & I(q-p) \end{pmatrix} = (D_{\gamma}/(D_{\gamma}/I-\gamma)^{0}) \begin{pmatrix} I(p) & (D_{\gamma})^{0}(p) \\ (D_{\gamma})^{0}(p) & I(q-p) \end{pmatrix}, \]

where we notice that, on the right hand side, one matrix factor is the transpose of the other matrix factor. Taking the inverse on both sides of (A.7.3.2) we have

\[ (A.7.3.3) \quad \begin{pmatrix} I(p) & (D_{\gamma})^{0}(p) \\ (D_{\gamma})^{0}(p) & I(q-p) \end{pmatrix}^{-1} = \begin{pmatrix} I(p) & -(D_{\gamma}/I-\gamma)^{0} \\ (D_{\gamma}/I-\gamma)^{0} & I(q-p) \end{pmatrix}. \]

Taking into account (A.7.3.1), (A.7.3.2) and (A.7.3.3) and using (A.1.5) we have

\[ (A.7.3.4) \quad \text{tr} \Sigma^{-1} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = M_{\gamma}(\gamma) \begin{pmatrix} \mu_1^{-1} & 0 \\ 0 & \mu_2^{-1} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \begin{pmatrix} \mu_1^{-1} & 0 \\ 0 & \mu_2^{-1} \end{pmatrix} \]

\[ = M(\gamma) M_{\gamma}(\gamma) \text{ (say)}. \]

Now put \( \mu_1^{-1}X_1 = Y_1 \) and \( \mu_2^{-1}X_2 = Y_2 \), i.e., \( X_1(p \times n) = \mu_1(p \times p)X_1(q \times n) \) and \( X_2(q \times n) = \mu_2(q \times q)Y_2(q \times n) \) and observe that, by (A.4.4), \( c(x_1x_1)^{-1}(x_1x_2)(x_2x_2)^{-1}(x_2x_1) \)

\[ = c(y_1y_1)^{-1}(y_1y_2)(y_2y_2)^{-1}(y_2y_1). \]

Also, by (A.5.2), \( J(x_1x_2; y_1, y_2) = |\mu_1|^{n} |\mu_2|^{n} \).
Next check that $|\Sigma|^2 = |\mu_1| = |\mu_2|^n \operatorname{tr} \left( \begin{bmatrix} I(p) & (D \nu) (0) \\ (D \nu) & I(q) \end{bmatrix} \right) = |\mu_1| |\mu_2|^n \prod_{i=1}^p \lambda_i^{1-\gamma_i^2} 
}

and finally check that $(Y_1, Y_2)$ have the probability law:

$$(A.7.3.5) \quad \frac{1}{(2\pi)^{n/2}} \prod_{i=1}^n \lambda_i^{1/2} \exp \left[ -\frac{1}{2} \operatorname{tr} N'(\gamma) \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} M(\gamma) \right] dX_1 dX_2.
$$

In view of the fact that $c(X_1 X_1)^{-1}(X_1 X_2)(X_2 X_2)^{-1}(X_2 X_1)^{-1} = c(Y_1 Y_1)^{-1}(Y_1 Y_2)(Y_2 Y_2)^{-1}$

$$(Y_1 Y_1)$$

the probability law $(A.7.3.5)$ proves $(A.7.3)$. Thus, as in the two previous subsections, if we are interested in the distribution of these roots, i.e., the $c$'s, we can, without any loss of generality, start from the probability law $(A.7.3.5)$, which is thus a canonical form for this purpose.

$$(A.7.4) \quad \text{If } X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \text{ (\text{r1} \leq q, \text{r2}+q+r \leq n) has the probability law (5.15):}

\begin{align*}
\frac{1}{(2\pi)^{(n+q+r)/2}} &\prod_{i=1}^n \lambda_i^{n/2} 
\cdot \exp \left[ -\frac{1}{2} \Sigma^2 X X^T \right] dX \text{ where } \Sigma = \\
&= \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{12} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{13} & \Sigma_{23} & \Sigma_{33} \end{pmatrix}
\end{align*}

is supposed to be symmetric p.d., then the distribution of the $c$'s could not involve as parameters anything except $\gamma$'s where $c$'s and $\gamma$'s are respectively the characteristic roots of $\begin{pmatrix} X_1 X_1 & -X_1 X_2 & -X_1 X_3 \\ X_2 X_1 & X_2 X_2 & -X_2 X_3 \\ X_3 X_1 & X_3 X_2 & X_3 X_3 \end{pmatrix}$ and $\Gamma_{11} - \Gamma_{13} \Gamma_{33}^{-1} \Gamma_{13} \Gamma_{22} - \Gamma_{13} \Gamma_{33}^{-1} \Gamma_{13} \Gamma_{22} + \Gamma_{11}$

- $\Gamma_{12} = \Sigma_{12}$

- $\Gamma_{13} = \Sigma_{13}$

- $\Gamma_{22} = \Sigma_{22}$

- $\Gamma_{23} = \Sigma_{23}$

- $\Gamma_{33} = \Sigma_{33}$

- $\Gamma_{11} = \Sigma_{11}$

- $\Gamma_{12} = \Sigma_{12}$

- $\Gamma_{13} = \Sigma_{13}$

- $\Gamma_{22} = \Sigma_{22}$

- $\Gamma_{23} = \Sigma_{23}$

- $\Gamma_{33} = \Sigma_{33}$

Proof. Notice as in the previous section that, a.o., the $c$'s are positive and distinct and that the $\gamma$'s are all non-negative. Use $(A.3.20)$ to set
(4.7.4.1) \[
\begin{pmatrix}
\Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\
\Sigma_{12} & \Sigma_{22} & \Sigma_{23} \\
\Sigma_{13} & \Sigma_{23} & \Sigma_{33}
\end{pmatrix}
= \begin{pmatrix}
\mu_1 & 0 & \mu_3 \\
0 & \mu_2 & \mu_4 \\
0 & 0 & \tilde{\mu}_5
\end{pmatrix}
\begin{pmatrix}
I & (D \Gamma) & 0 \\
(D \Gamma) & I & 0 \\
0 & 0 & I
\end{pmatrix}
\begin{pmatrix}
\mu_1 & 0 & c \\
0 & \mu_2 & 0 \\
\mu_3 & \mu_4 & \mu_5
\end{pmatrix}
\]
and check that

(4.7.4.2) \[
\begin{pmatrix}
\mu_1 & 0 & \mu_3 \\
0 & \mu_2 & \mu_4 \\
0 & \tilde{\mu}_5
\end{pmatrix}
^{-1}
\]
is of the form

\[
\begin{pmatrix}
\nu_1 & 0 & \nu_3 \\
0 & \nu_2 & \nu_4 \\
0 & \tilde{\nu}_5
\end{pmatrix}
\]
Proceeding as in the previous section we have now

(4.7.4.3) \[
\Sigma^{-1} = \begin{pmatrix}
\mu_1 & 0 & 0 \\
0 & \mu_2 & 0 \\
\mu_3 & \mu_4 & \mu_5
\end{pmatrix}
^{-1}
\begin{pmatrix}
I & -(D \Gamma / I - \gamma) & 0 \\
(D \Gamma / I - \gamma) & (D \Gamma / I - \gamma) & 0 \\
0 & 0 & I
\end{pmatrix}
\begin{pmatrix}
\mu_1 & 0 & \mu_3 \\
0 & \mu_2 & \mu_4 \\
0 & \tilde{\mu}_5
\end{pmatrix}
\]
\[
= \mu_1^{-1} M(\gamma) M^*(\gamma) \mu_1^{-1} \text{ (say)}
\]
and thus, as in the previous section,

(4.7.4.4) \[
\text{tr} \Sigma^{-1} \begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix}
\begin{pmatrix}
X_1 \\
X_2 \\
X_3
\end{pmatrix}
= \text{tr} \begin{pmatrix}
X_1 \\
X_2 \\
X_3
\end{pmatrix}
\mu_1^{-1} M(\gamma) M^*(\gamma) \mu_1^{-1} \begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix}
\]

Now set \[
\begin{pmatrix}
X_1 \\
X_2 \\
X_3
\end{pmatrix}
= \begin{pmatrix}
\mu_1 & 0 & \mu_3 \\
0 & \mu_2 & \mu_4 \\
0 & 0 & \tilde{\mu}_5
\end{pmatrix}
\begin{pmatrix}
Y_1 \\
Y_2 \\
Y_3
\end{pmatrix}
\]
and note that the c's are invariant under this transformation. Also \[
J(X;Y) = |\mu|^n \text{ and } |Z| = |\mu|^n \prod_{i=1}^{n} (1-\gamma_i)^{n_i}. \text{ Thus, finally Y has the probability law}
\]

Y has the probability law
(4.7.4.5) \[
\left[ \frac{1}{2\pi^{n/2}} \prod_{i=1}^{n} (1-\gamma_i^2)^{-1} \right] \exp \left[ -\frac{1}{2} \text{tr} \cdot \Sigma^{-1} \cdot \left( \begin{array}{c} Y_1 \\ Y_2 \\ Y_3 \end{array} \right) \cdot \left( \begin{array}{c} Y_1 \\ Y_2 \\ Y_3 \end{array} \right)^\top \cdot \Sigma^{-1} \right] dY_1 \, dY_2 \, dY_3
\]
which proves (4.7.4) which we take to be a canonical form.

(4.7.5) If \( X_1(p \times n_1) \) and \( X_2(p \times n_2) \) (\( p \leq n_2 \) but might be \( \leq p \) or \( > n_1 \)) have the joint probability law (5.21):
\[
\frac{1}{2\pi^{(n_1+n_2)/2}} \frac{1}{2^{1/2}} \sqrt{(n_1+n_2)} \exp \left[ -\frac{1}{2} \text{tr} \cdot \Sigma^{-1} \cdot \left( \begin{array}{c} X_1 \\ X_2 \end{array} \right) \cdot \left( \begin{array}{c} X_1 \\ X_2 \end{array} \right)^\top \cdot \Sigma^{-1} \right] dX_1 \, dX_2,
\]
where \( \Sigma(p \times p) \) is symmetric p.d. and \( \xi \) is \( p \times n_1 \), then the distribution of \( cX_1X_1^\top(X_2X_2^\top)^{-1} \), to be called \( c' \)'s could not involve as parameters anything except \( c(\xi_2^\top \Sigma^{-1} \xi_2) \), to be called \( \gamma ' \)'s.

Proof. Notice that, a.c., out of the \( c' \)'s, \( r \) are positive and \( r-r \) are zero, where \( r = \min(p,n_1) \) and also that \( \gamma ' \)'s are all non-negative and out of them \( s \) are positive and \( n-s \) are zero where \( s \leq \min(r,n_1) \) is the rank of \( \Sigma \), i.e., of \( \xi_2^\top \Sigma^{-1} \xi_2 \). Assuming, as we can without any loss of generality that the last \( s \) rows of \( \xi_2^\top \Sigma^{-1} \xi_2 \) are all zero, the last \( s \) rows of \( \Sigma \) can be taken as the basis, use (4.3.13) to set

\[
(4.7.5.1) \quad (\xi_2^\top \Sigma^{-1} \xi_2)(p \times p) = \begin{pmatrix} \mu_1^{1/2} \\ \mu_2^{1/2} \end{pmatrix} s \quad \text{and} \quad D_\Sigma(s \times s) \left( \begin{array}{cc} \mu_1^{1/2} & 0 \\ 0 & \mu_2^{1/2} \end{array} \right) s \quad \text{and}
\]

\[
\Sigma(p \times p) = \begin{pmatrix} \mu_1^{1/2} & \mu_3^{1/2} \\ \mu_3^{1/2} & \mu_4^{1/2} \end{pmatrix} s \quad \text{and}
\]

where \( \mu = \begin{pmatrix} \mu_1 & \mu_3 \\ \mu_2 & \mu_4 \end{pmatrix} \) and \( \mu_3^{1/2} \) are non-singular and \( D_\Sigma^{1/2} \) stands for the diagonal matrix

with the \( s \) (non-zero and hence positive) roots. If we now set

\[
(4.7.5.2) \quad \xi(p \times n_1) = \begin{pmatrix} \xi_1^\top \\ \xi_2^\top \end{pmatrix} s = \begin{pmatrix} \mu_1^{1/2} \\ \mu_2^{1/2} \end{pmatrix} s \quad \text{and} \quad D_\Sigma^{1/2} \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix} s \quad \text{and}
\]

it is easy to check that \( \nu \) is determined by \( \nu = D_\Sigma^{1/2} \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix} \begin{pmatrix} \xi_2^\top \\ \xi_2^\top \end{pmatrix} \) and that \( \nu \nu^\top = I(s) \). Let

\[
D_\Sigma(p \times p) = \begin{pmatrix} D_\Sigma^{1/2} & 0 \\ 0 & \mu_p \end{pmatrix} s \quad \text{and}
\]
Recall that \( s \leq \min(n_1, n_2) \). Recalling now that \( \text{tr} X_1 \xi' = \text{tr} \xi' X_1 \) and using (\( \cdot.1.4 \)), (\( \cdot.7.5.1 \)) and (\( \cdot.7.5.2 \)) we have

\[
(\cdot.7.5.3) \quad \text{tr} \left\{ \frac{X_2 X_1' + (X_1 \xi')(X_1 \xi')'}{\xi'} \right\} = \mu^{-1} \left\{ \frac{X_2 X_1' + X_1 \xi' X_1' - 2X_1 \Sigma Y Y'}{\Sigma} \chi_1 \mu^2 \right\} + \mu D \mu' \mu^{-1}.
\]

Now using (\( \cdot.1.7 \)) complete \( \nu'(n_1 \times s) \) into an \( \Sigma \delta'(n_1 \times n_1) \) and rewrite the right hand side of (\( \cdot.7.5.3 \)) as \( \mu^{-1} \left\{ \frac{X_2 X_1' + X_1 \xi' X_1' - 2X_1 (p \times n_1) \delta'(n_1 \times n_1)}{\Sigma} \right\} \mu' \cdot \mu'. \mu^{-1} \mu \cdot \mu'. \mu^{-1} \). But now \( \mu^{-1} X_2 = Y_2 \) and \( \mu^{-1} X_1 \delta' = Y_1 \), i.e., \( X_1 (p \times n_1) = \mu(p \times p) Y_1 (p \times n_1) \delta(n_1 \times n_1) \) and \( X_2 (p \times n_2) = \mu(p \times p) Y_2 (p \times n_2) \) and observe that, by (\( \cdot.4.3 \)), \( c(\chi_1 \chi_1' (X_2 Y_2')^{-1}) = c(Y_1 Y_1' (Y_2 Y_2')^{-1}) \) since \( \mu \) is non-

singular and \( \delta' \) is \( \Sigma \). Also, by (\( \cdot.5.2 \)), \( J(X_1 \chi_2', Y_1, Y_2) = \mu^2 \). Finally check that \( (Y_1 Y_2) \) has the probability law

\[
(\cdot.7.5.4) \left[ \frac{1}{2\pi} \right] \left[ \frac{1}{2\pi} \right] \exp \left[ -\frac{1}{2} \left\{ \text{tr} \left\{ Y_1 Y_2 + D \Sigma + Y_1 Y_1' - 2Y_1 (p \times n_1) \right\} \right\} \right] \left[ \frac{1}{2\pi} \right] \left[ \frac{1}{2\pi} \right] \exp \left[ -\frac{1}{2} \left\{ \text{tr} \left\{ D \Sigma \right\} \right\} \right] \right] dY_1 \ dY_2,
\]

which, in view of the fact that \( c(\chi_1' Y_1' (X_2 Y_2')^{-1}) = c(Y_1 Y_1' (Y_2 Y_2')^{-1}) \), completes the proof of (\( \cdot.7.5 \)). We also note as before that for the purpose of any discussion of the distribution of \( c \)'s the probability law (\( \cdot.7.5.4 \)) can be taken as a canonical form. In (\( \cdot.7.5.4 \)) notice that

\[
(\cdot.7.5.5) \quad \text{tr} \left( Y_1 (p \times n_1) \left[ \frac{D}{\Sigma} \left[ \frac{0}{0} \right] \right] \right) = \sum_{i=1}^{s} \left( Y_1 \gamma_i \right) \left( Y_1 \gamma_i \right)' \left[ \frac{1}{2\pi} \right] \left[ \frac{1}{2\pi} \right] \exp \left[ -\frac{1}{2} \left\{ \text{tr} \left\{ Y_1 Y_1' + Y_2 Y_2' \right\} + \sum_{i=1}^{s} \left( Y_1 \gamma_i \right)' \left( Y_1 \gamma_i \right) - 2\left( Y_1 \gamma_i \right)' \left( Y_2 \gamma_i \right) \right\} \right] dY_1 \ dY_2.
\]

Using (\( \cdot.7.5.5 \)) the canonical form (\( \cdot.7.5.4 \)) can thus be reduced to the more convenient form

\[
(\cdot.7.5.6) \quad \frac{p(n_1 + n_2)}{2\pi} \exp \left[ -\frac{1}{2} \left\{ \text{tr} \left( Y_1 Y_1' + Y_2 Y_2' \right) + \sum_{i=1}^{s} \left( Y_1 \gamma_i \right)' \left( Y_1 \gamma_i \right) - 2\left( Y_1 \gamma_i \right)' \left( Y_2 \gamma_i \right) \right\} \right] dY_1 \ dY_2.
\]
A.8. Some results in integration.

\[(A.8.1)\] \[\int \frac{n}{\prod_{i=1}^{n} x_i^{p_i-1} dx_i} = \frac{n}{\prod_{i=1}^{n} \Gamma\left(\frac{p_i}{q_i}\right) a_i^{p_i}} / \Gamma\left(\sum_{i=1}^{n} \frac{p_i}{q_i} + 1\right) \prod_{i=1}^{n} q_i \], \[\sum(x_i/a_i) \leq 1\]

where \(x_i \geq 0\) and \(p_i, q_i, a_i > 0\), \(i = 1, 2, \ldots, n\). An important special case is where \(a_i = r, p_i = 1\) and \(q_i = 2\), in which case we have

\[(A.8.2)\] \[\int_{\sum x_i^2 < r^2(x_i > 0)} \frac{n}{\prod_{i=1}^{n} dx_i} = \sqrt{\Gamma\left(\frac{1}{2}\right)} \Gamma^n r^{n/2} \Gamma\left(\frac{n}{2} + 1\right).\]

If, however, we integrate over \(x_i\)'s in the domain \(\sum_{i=1}^{n} x_i^2 < r^2\), after dropping the restriction that \(x_i \geq 0\), i.e., if \(x_i\)'s could take both +ve and -ve values, subject to \(\sum_{i=1}^{n} x_i^2 < r^2\), then from considerations of symmetry we shall have

\[(A.8.3)\] \[\int_{\sum x_i^2 < r^2} \frac{n}{\prod_{i=1}^{n} dx_i} = \sqrt{\Gamma\left(\frac{1}{2}\right)} \Gamma^n r^{n} / \Gamma\left(\frac{n}{2} + 1\right).\]

Differentiating the above on both sides w.r.t. \(r\) we have

\[(A.8.4)\] \[\int_{r \leq (\sum_{i=1}^{n} x_i^2)^{1/2} \leq r+dr (r \geq 0)} \frac{n}{\prod_{i=1}^{n} dx_i} = n\sqrt{\Gamma\left(\frac{1}{2}\right)} \Gamma^n r^{n-1} dr / \Gamma\left(\frac{n}{2} + 1\right).\]

\[(A.8.5)\] \[\int_{r \leq (\sum_{i=1}^{n} x_i^2)^{1/2} \leq r+dr (r \geq 0)} \frac{n}{\prod_{i=1}^{n} dx_i} = \frac{(n-1)\sqrt{\Gamma\left(\frac{1}{2}\right)} \Gamma^n r^{n-1}}{\Gamma\left(\frac{n-1}{2} + 1\right)} dr (\sin \theta)^{n-2} d\theta .\]

\[r \leq (\sum_{i=1}^{n} x_i^2)^{1/2} \leq r+dr (r \geq 0)\]

\[\theta \leq \cos^{-1}\left(\sum_{i=1}^{n} x_i a_i / \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \left(\sum_{i=1}^{n} a_i^2\right)^{1/2}\right) \leq \theta+\delta\theta\]
Proof. Make a transformation $y_1 = \sum_{i=1}^{n} x_i^1a_i/(\sum_{i=1}^{n} a_i^2)^{1/2} = r^* \cos \theta^*$, say, and

$$y_i = \sum_{j=1}^{n} u_{ij} x_j \quad (i = 2, 3, \ldots, n)$$

such that

$$\begin{pmatrix}
a_1/a & a_2/a & \cdots & a_n/a \\
\mu_{21} & \mu_{22} & \cdots & \mu_{2n} \\
\mu_{31} & \mu_{32} & \cdots & \mu_{3n} \\
\vdots & \vdots & \ddots & \vdots \\
\mu_{n1} & \mu_{n2} & \cdots & \mu_{nn}
\end{pmatrix}
\begin{pmatrix}
a_1^2 \\
a_2^2 \\
a_3^2 \\
\vdots \\
a_n^2
\end{pmatrix}
= \begin{pmatrix}
\sum_{i=1}^{n} a_i^2 \\
\sum_{i=1}^{n} \mu_{ij}^2 a_i^2
\end{pmatrix}
$$

is an orthogonal matrix. Then using (4.5.4) and remembering that $J(x:y) = 1/J(y:x)$, we shall have $\int \prod_{i=1}^{n} dx_i \rightarrow \int \prod_{i=1}^{n} dy_1$, i.e., $\int \prod_{i=1}^{n} dy_i$. We have furthermore $\sum_{i=1}^{n} y_i^2 = y_1^2 + \sum_{i=2}^{n} y_i^2 = \sum_{i=1}^{n} x_i^2 = r^* x^2$, so that $\sum_{i=1}^{n} y_i^2 = r^* y_1^2 + \sum_{i=2}^{n} y_i^2 = r^* \sin^2 \theta^*$, whence $(\sum_{i=1}^{n} y_i^2)^{1/2} = r^* \sin \theta^* = u^*$ (say). It is easy to see that the domain: $r \leq r^* \leq r + dr$ and $\theta \leq \theta^* \leq \theta + d\theta$, is exactly equivalent to $u \leq u^* \leq u + du$ and $v \leq y_1 \leq v + dv$, so that

$$\int \prod_{i=1}^{n} \frac{dx_i}{r} = \int \prod_{i=1}^{n} \frac{dy_1}{v} = dv \int \prod_{i=2}^{n} \frac{dy_i}{u^*}$$

$$r \leq r^* \leq r + dr \quad v \leq y_1 \leq v + dv \quad u \leq (\sum_{i=1}^{n} y_i^2)^{1/2} \leq u + du$$

$$\theta \leq \theta^* \leq \theta + d\theta \quad u \leq u^* \leq u + du$$

$$= dv \, (n-1) \sqrt{T(\frac{1}{2})} \, \prod_{i=1}^{n-1} u^{-2} du / (\frac{n-1}{2} + 1) \quad (\text{using (4.8.4)}), = (n-1) \sqrt{T(\frac{1}{2})} \, \prod_{i=1}^{n-1} r^{-1} r^{-1} dr$$

$$x (\sin \theta)^{n-2} d\theta / (\frac{n-1}{2} + 1), \text{which proves (4.8.5).} \text{ Notice that } y_1 = r^* \cos \theta^* \text{ and } u^* = r^* \sin \theta^*, \text{ whence } J'(y_1, u^*, \theta^*) = r^*, \text{ so that } dy_1 du^* \rightarrow r^* dr^* d\theta^*.$$ 

(4.8.6) The integral $\int_{L_I} \frac{\partial (LL^I)}{\partial (L)} \bigg|_{L_I} = F(p,n)$ (say), where $L$ is $p \times n$

where $p < n$. This can be evaluated directly but we shall use an artifice to derive this. Consider the integral

$$\int \frac{p \pi}{2} \exp \left\{ -\frac{1}{2} tr YY' \right\} \, dY$$

where the elements of $Y(p \times n) \ (p < n)$ vary from $-\infty$ to $\infty$. It is of course known
that this integral is equal to 1. Using now the transformation (A.3.11) we have
\( X(p \times n) = \tilde{T}(p \times p)L(p \times n) \) under \( LL' = I(p) \). Notice that almost everywhere \( YY' \) and so \( \tilde{T} \) will be non-singular. The \( t_{ii}'s \) vary from 0 to \( \infty \) and \( t_{ij}'s \) (\( i \neq j \)) vary from \(-\infty \) to \( \infty \). Observe that \( YY' = \tilde{T}\tilde{T}' \) and \( \text{tr } YY' = \sum_{i,j=1}^{P} t_{ij}'^2 \). Using now the result (A.6.1.9) we have

\[
(l.8.6.2) \quad 1 = \int_{Y} \left[ \frac{2n}{(2\pi)^{\frac{Pn}{2}}} \right] \exp \left[ -\frac{1}{2} \text{tr } YY' \right] dY \\
= \int_{LL' = I(p)} \left\{ \frac{dL_L}{|\frac{\partial (LL')}{\partial (L'I)}|} \right\}^{\frac{Pn}{2}} \int_{l/2}^{l/2} \frac{dY}{(2\pi)^{\frac{Pn}{2}}} \\
x \int_{0}^{\infty} \frac{dL_L}{L_{I}} \left| \frac{\partial (LL')}{\partial (L'I)} \right|^{\frac{Pn}{2}} \int_{0}^{\infty} \exp \left[ -\frac{1}{2} \sum_{i,j=1}^{P} t_{ij}'^2 \right] \prod_{i=1}^{P} \frac{t_{ii}'}{2} t_{i}' \prod_{i,j=1}^{P} \frac{1}{2} \Gamma \left( \frac{n-i+1}{2} \right) dt_{ij} .
\]

But the last integral on the right hand side of (A.8.6.2) is easily evaluated to be
\[ 2^{-\frac{n+P}{2}} \pi^{\frac{P}{2}(p-1)} \prod_{i=1}^{P} \Gamma \left( \frac{n-i+1}{2} \right). \]
Hence we have the following result (to be repeatedly used)

\[
(l.8.6.3) \quad F(p,n) = \int_{LL' = I(p)} \frac{dL_L}{L_{I}} \left| \frac{\partial (LL')}{\partial (L'I)} \right|^{\frac{Pn}{2}} \int_{\frac{1}{2}}^{\frac{1}{2}} \frac{dY}{(2\pi)^{\frac{Pn}{2}}} \\
\prod_{i=1}^{P} \Gamma \left( \frac{n-i+1}{2} \right).
\]

(A.8.7) The integral \( \int_{U} \exp \left[ -\frac{1}{2} \text{tr } UU' \right] U(U')^q dU \), where \( U(p \times p) \) has its elements varying from \(-\infty \) to \( \infty \). Using the transformation (A.3.11) we have \( U(p \times p) = \tilde{T}(p \times p) \) with an orthogonal \( L \). Notice that \( |U| = |\tilde{T}| = \prod_{i=1}^{P} t_{ii}' \) and that almost everywhere \( \tilde{T} \) is non-singular. Also as before \( \text{tr } UU' = \sum_{i,j=1}^{P} t_{ij}'^2 \). Hence we
have, by using (A.6.1.9) and (A.8.6.3), $\int_{10,11} L_7$

\[(A.8.7.1) \quad \int \exp \left[ -\frac{1}{2} \text{tr} \ U U' I \right] |U| q \, dU = 2^p \int \text{d}L_I \left| \frac{\partial (L_I')}{\partial M_I} \right| \]

\[\times \quad \int \exp \left[ -\frac{1}{2} \sum_{i=j=1}^{p} t_{ij}^2 \frac{1}{i!} \sum_{i=j=1}^{p} t_{ij}^{p+q-1} \frac{1}{i!} \right] \, dt_{ij} \]

\[0 \leq t_{ii} < \infty, \quad -\infty < t_{ij} < \infty \quad (i > j) \]

\[= 2^p (p+q-2)/2 \frac{p^2}{4} \sum_{i=1}^{p} \frac{1}{i!} \frac{\Gamma(q+q+1)}{\Gamma(q+q+1)} \frac{1}{i!} \frac{\Gamma(p+q+1)}{\Gamma(p+q+1)} \]

\[= 2^p (p+q-2)/2 \frac{p^2}{4} \sum_{i=1}^{p} \frac{1}{i!} \frac{\Gamma(q+q+1)}{\Gamma(q+q+1)} \frac{1}{i!} \frac{\Gamma(p+q+1)}{\Gamma(p+q+1)} \]

\[(A.8.8) \quad \text{The integral } \int \exp \left[ -\frac{1}{2} \text{tr} \ U U' I \right] |U| q \sum_{i=1}^{p} (\tilde{U}_3')_{i=1}^{p-n} | \sum_{i=1}^{p-n} \text{d}U , \]

where $U(p \times p) = \begin{pmatrix} U_1 & \tilde{U}_3 \\ U_2 & U_4 \end{pmatrix} \quad \text{p-n}$

the diagonal elements of $\tilde{U}_3$ vary from 0 to $\infty$ and

the rest from $-\infty$ to $\infty$. Let $V = \begin{pmatrix} V_1 & V_3 \\ V_2 & V_4 \end{pmatrix} \quad \text{p-n}$

$\begin{pmatrix} U_1 & \tilde{U}_3 \\ U_2 & U_4 \end{pmatrix} \quad \text{p-n}$

$\begin{pmatrix} V_1 & V_3 \\ V_2 & V_4 \end{pmatrix} \quad \text{p-n}$

Then we have $VV' = UU'$ and

\[J(V; U) = J(V_3; V_3; M_1; U_1) = J(V_3; \tilde{U}_3; M_1; U_1) = 2^{p-n} \sum_{i=1}^{p-n} (\tilde{U}_3')_{i=1}^{p-n} |M_1| \]

\[\frac{\partial (\tilde{U}_3')_{i=1}^{p-n}}{\partial M_1} \]

\[= 2^{p-n} \sum_{i=1}^{p-n} (\tilde{U}_3')_{i=1}^{p-n} |M_1| \]

Thus we have

\[(A.8.8.1) \quad \int \exp \left[ -\frac{1}{2} \text{tr} \ U U' I \right] |U| q \, dU = \int \exp \left[ -\frac{1}{2} \text{tr} \ U U' I \right] |U| q \sum_{i=1}^{p-n} (\tilde{U}_3')_{i=1}^{p-n} q \, dU \]
\[ x \int \frac{dW_i}{\partial_i} \mid \frac{d(x_i)}{dx_i} \mid. \] Now substituting from (A.8.7.1) in the left hand side of (A.8.7.1) and from (A.8.8.3) in the right hand side of (A.8.8.1) we have, \[ \mathcal{V}_{107}, \]

(A.8.8.2) \[ \int \exp \left\{ -\frac{1}{2} \text{tr} \, \mathbf{U} \mathbf{U}^T, \right\} \, \mathcal{V}_{107} \]

\[ = 2^{\frac{p(p+4)/2}{4}} \cdot \frac{\Gamma\left(\frac{p+4-i+1}{2}\right)}{\Gamma\left(\frac{p-i+1}{2}\right)} \cdot \prod_{i=1}^{p} \frac{\Gamma\left(\frac{p-n-i+1}{2}\right)}{\Gamma\left(\frac{p-n-i+1}{2}\right)}. \]

A.9. Some results in integration connected with the distribution of the largest and the smallest roots, \[ \mathcal{V}_{12,13,17,7}. \]

(A.9.1) **Evaluation of the integral**

\[ \int x^s x^{s-1} \cdots x_1^s \, dx_1 \]

\[ = \beta \cdot x; m_s, n_s; s-1, n_{s-1}; \ldots; m_1, n_1; \mathcal{V} \text{(say)} = \beta \cdot x; m_s, n_s; s-1, n_{s-1}; \ldots; m_1, n_1; \]

The last expression is in the form of a pseudo-determinant whose meaning is made clear by considering, for illustration, the case of \( s = 3 \), for which

(A.9.1.1) \[ \beta \cdot \left[ \begin{array}{ccc}
    m_3, n_3 & m_2, n_2 & m_1, n_1 \\
    m_3, n_3 & m_2, n_2 & m_1, n_1 \\
    m_3, n_3 & m_2, n_2 & m_1, n_1 \\
  \end{array} \right] 
\]

\[ = \int x_3^m (1-x_3)^{n_3} \, dx_3 \int x_2^m (1-x_2)^{n_2} \, dx_2 \int x_1^m (1-x_1)^{n_1} \, dx_1 \]
\[
\begin{align*}
&\left[ x^3 \int_{x_2(1-x_2)^{n_2}dx_2}^{x_1(1-x_1)^{n_2}dx_1} - \int_{x_2(1-x_2)^{n_2}dx_2}^{x_1(1-x_1)^{n_2}dx_1} x^2 \int_{x_2(1-x_2)^{n_2}dx_2}^{x_1(1-x_1)^{n_2}dx_1} x^3 \int_{x_2(1-x_2)^{n_2}dx_2}^{x_1(1-x_1)^{n_2}dx_1} 
\end{align*}
\]

In opening out the pseudo-determinant it is very important to stick to the order of the factors, indicated in the expansion on the right side of \((.9.1.1)\) and to keep in mind that the factors are non-commutative. It is also clear that the whole expression will be zero if any two columns become equal in

\[
\begin{pmatrix}
m_s, n_s & m_{s-1}, n_{s-1} & \cdots & m_1, n_1 \\
m_s', n_s & m_{s-1}', n_{s-1} & \cdots & m_1', n_1 \\
\vdots & \vdots & \ddots & \vdots \\
m_s, n_s & m_{s-1}, n_{s-1} & \cdots & m_1, n_1
\end{pmatrix}
\]

Next we use the notation

\[(.9.1.?) \quad \Theta(x; m_s, n_s; m_{s-1}, n_{s-1}; \ldots; m_1, n_1) = \int_{x_s}^{x_{s-1}} x_s^{n_s(1-x_s)} dx_s \int_{x_{s-1}}^{x_{s-1}} x_{s-1}^{m_{s-1}(1-x_{s-1})} dx_{s-1}\]
\[
\int_0^x x (1-x)^n \, dx = \int_0^{x_1} x_1^{n_1} \, dx_1 + \int_0^{x_2} x_2^{n_2} \, dx_2 + \int_0^{x_3} x_3^{n_3} \, dx_3,
\]
so that \( \theta(x; a, n) = \int_0^x x_1^{n_1} \, dx_1 = \) the incomplete \( \theta \)-function. Also let \( x^m(1-x)^n = \theta_0(x; m, n) \). In terms of (4.9.1.2), the expression (4.9.1.1) can be rewritten as

(4.9.1.3) \( \theta(x; m_2, n_2; m_1, n_1) - \theta(x; m_3, n_3; m_1, n_1) - \theta(x; m_2, n_2; m_3, n_3) \)

\[+ \theta(x; m_2, n_2; m_3, n_3) + \theta(x; m_1, n_3; m_3, n_3) - \theta(x; m_1, n_3; m_2, n_2) \]

and (4.9.1) can be rewritten as

(4.9.1.4) \( \Sigma + \theta(x; s, n; s-1, n-1; n, n) \), where \( (m^1, n^1), (m^2, n^2), \ldots, (m^i, n^i), \) is any permutation of \( (m_s, n_s), (m_{s-1}, n_{s-1}), \ldots, (m_1, n_1) \), the summation is taken over all such permutations, the positive or negative sign is taken exactly as in the usual expansion of a determinant, care being taken to preserve the order of factorization from \( x_s \) through \( x_{s-1}, x_{s-2} \) down to \( x_1 \).

(4.9.2) \textit{Lemma.}

\[
\int_0^x x^m(1-x)^n f(x) \, dx = \frac{1}{m+n+1} \int_{-x_0^m(1-x_0)^{n+1}} f(x_0) + \int_0^x (1-x)^{n+1} f'(x) + x^n f(x) \, dx.
\]

where \( m, n > -1 \), \( x_0 \leq 1 \) and \( f(x) \) is such that \( f'(x) \) and the two integrals on two sides of (4.9.2) all exist. The proof is obvious and need not be given here.

By handling the second term on the right side of (4.9.2) it is easy to check that

(4.9.2.1) \( \int_0^x x^m(1-x)^n f(x) \, dx = \frac{1}{m+n+1} \int_{-x_0^m(1-x_0)^{n+1}} f(x_0) \)

\[+ \int_0^x x^n(1-x)^{n+1} f'(x) \, dx + m \int_0^x x^{n-1}(1-x)^n f(x) \, dx.\]
Theorem. \( \Sigma \theta(x;m_1^{i_1},n_1^{i_1};m_{s-1}^{i_{s-1}},n_{s-1}^{i_{s-1}};\ldots;m_1^{i_1},n_1^{i_1}) = \prod_{i=1}^{s} \beta(x;m_i^{i_1},n_i^{i_1}), \) where on the left hand side \((m_1^{i_1},n_1^{i_1}),\ldots,(m_1^{i_1},n_1^{i_1})\) is any permutation of \((m_1^{i_1},n_1^{i_1}),\ldots,(m_1^{i_1},n_1^{i_1})\), the summation is taken over all such permutations and where the factors on the right hand side have been already defined.

Proof. The nature of the proof will be evident by considering, for simplicity of algebra, the case of \(s = 2\). We have

\[
\frac{x}{x_2^2(1-x_2)^2} \frac{x}{x_1^2(1-x_1)^2} \frac{1}{x_2} \frac{1}{x_1} \frac{x}{x_2} \frac{x}{x_1} = \int_0^x \int_0^{m_2(1-x_2)^2} \int_0^{m_1(1-x_1)^2} \frac{dx_1}{dx_2} \int_0^x \int_0^{m_2(1-x_2)^2} \int_0^{m_1(1-x_1)^2} \frac{dx_2}{dx_1}.
\]

(which is obtained by interchanging, in the second term on the left side of (A.9.3.1), the variables \(x_2\) and \(x_1\) and rewriting the domain of integration in the appropriate manner)

\[
\frac{x}{x_2^2(1-x_2)^2} \frac{x}{x_1^2(1-x_1)^2} = \theta(x; m_2, n_2) \theta(x; m_1, n_1).
\]

Lemma. \( \Sigma \theta_r(x;m_{s-1}^{r},n_{s-1}^{r};\ldots;m_r^{r},n_r^{r};\ldots;m_1^{i_1},n_1^{i_1}) \)

\(= \theta(x; m_r^{r},n_r^{r}) \theta(x;m_{s-1}^{r},n_{s-1}^{r};\ldots;m_1^{i_1},n_1^{i_1}), \) where \(\theta_r\) is the result of putting \((m,n)\) in the \(r\)th place and filling up the other positions with \((m_{s-1}^{r},n_{s-1}^{r}),(m_{s-2}^{r},n_{s-2}^{r}),\ldots,(m_1^{i_1},n_1^{i_1}), \) \(r\) running from 1 to \(s\). Notice that each \(\theta_r\) is an \(s\)-fold integral, while \(\theta(x;m_{s-1}^{r},n_{s-1}^{r};\ldots;m_1^{i_1},n_1^{i_1})\) is an \((s-1)\)-fold integral.

Proof. The mechanism of the proof is brought out by considering, in particular, the case \(s = 3\), where we have

\[
\theta_1(x;m_2^{i_2},n_2^{i_2};m_1^{i_1},n_1^{i_1}) + \theta_2(x;m_2^{i_2},n_2^{i_2};m_1^{i_1},n_1^{i_1}) + \theta_3(x;m_1^{i_2},n_1^{i_2};m_2^{i_1},n_2^{i_1})
\]
\[
  = \int_{0}^{x} x_{2}^{m(1-x_{1})n_{1} dx_{1}} + \int_{0}^{x} x_{3}^{m(1-x_{2})n_{2} dx_{2}} + \int_{0}^{x} x_{3}^{m(1-x_{3})n_{3} dx_{3}}
\]

(by interchanging the variables and suitably adjusting the domain of integration)

\[
  = \Theta(x; m, n) \Theta(x; m_2, n_2; m_1, n_1).
\]

(\ref{9.5}) \textbf{Lemma},

\[
  \sum (-1)^{r-1} \int_{r} \begin{pmatrix}
    m_s, n_s & \cdots & m_1, n_1 \\
    m_s, n_s & \cdots & m_1, n_1 \\
    \vdots & \cdots & \vdots
  \end{pmatrix}
  \begin{pmatrix}
    m_s, n_s & \cdots & m_1, n_1 \\
    m_s, n_s & \cdots & m_1, n_1 \\
    \vdots & \cdots & \vdots
  \end{pmatrix}
\]

\[
  = \sum (-1)^{r-1} \int_{r} \Theta(x; m_{s-r+1}, n_{s-r+1}) \Theta_{rr}
  \begin{pmatrix}
    m_s, n_s & \cdots & m_1, n_1 \\
    m_s, n_s & \cdots & m_1, n_1 \\
    \vdots & \cdots & \vdots
  \end{pmatrix}
  \begin{pmatrix}
    m_s, n_s & \cdots & m_1, n_1 \\
    m_s, n_s & \cdots & m_1, n_1 \\
    \vdots & \cdots & \vdots
  \end{pmatrix}
\]

where \( \Theta \) on the left side is the result of replacing the \( r \)-th row of \( \Theta \) by \((m_1', n_1'), \ldots, (m_1', n_1')\) and \( \Theta \) on the right side is the result of suppressing the \( r \)-th row and \( r \)-th column of \( \Theta \). Notice that \( \Theta \) is an \( s \times s \) and \( \Theta \) an \((s-1) \times (s-1)\) pseudo-determinant.

\textbf{Proof.} The mechanism of the proof will be made clear by considering for simplicity the case of \( s = 3 \) and picking out from the expansion of each pseudo-determinant on the left side of (\ref{9.5}) (for the case \( s = 3 \)) the term involving the index, say, \((m_3', n_3')\) and putting together all such terms (with index \((m_3', n_3')\)). We have thus the following contribution from such terms
\[
\beta(x;m_1^n, n_1^{m_1}, m_2^n, n_2^{m_2}, m_3^n, n_3^{m_3}) - \beta(x;m_1^n, n_1^{m_1}, m_2^n, n_2^{m_2}, m_2^n, n_2^{m_2}) + \beta(x;m_1^n, m_2^n, n_2^{m_2}, m_3^n, n_3^{m_3}) - \beta(x;m_1^n, m_2^n, n_2^{m_2}, m_3^n, n_3^{m_3})
\]

\[
= \beta(x;m_1^n, n_1^{m_1}) \delta \left[ x; \begin{pmatrix} m_1 & n_1 \\ m_2 & n_2 \end{pmatrix} \right] = \beta(x;m_1^n, n_1^{m_1}) \delta \left[ x; \begin{pmatrix} m_3 & n_3 \\ m_2 & n_2 \\ m_2 & n_2 \end{pmatrix} \right]
\]

(\text{using the notation introduced in the beginning of lemma (A.9.5))}. \text{ This immediately shows that if, in the general case, from the expansion of each pseudo-determinant (with the proper sign) on the left side of (A.9.5) we pick out the term with the index } (m_1^n, n_1^{m_1}) \text{ and add together such terms (with the same index } (m_1^n, n_1^{m_1}) \text{) we shall have the following contribution}

\[
\beta(x;m_1^n, n_1^{m_1}) \delta \left[ x; \begin{pmatrix} m_1 & n_1 \\ m_2 & n_2 \\ \ldots & \ldots \end{pmatrix} \right],
\]

(\text{whence the proof becomes obvious by combining different expressions like (A.9.5.2) involving the different indices } (m_r^n, n_r^{m_r}) \text{ (} r = 1, 2, \ldots, s).\]

(A.9.6) \text{ Reduction and evaluation of the integral}

\[
\beta \left[ x; \begin{pmatrix} m_1 & n_1 \\ m_2 & n_2 \\ \ldots & \ldots \end{pmatrix} \right],
\]

\text{where } m_s > m_{s-1} > \ldots > m_1 > -1 \text{ and } n > -1 \text{ and the } m's \text{ differ by integers. We have already seen from (A.9.1.4) that the pseudo-determinant can be expanded into}

\[\Sigma \beta(x;m_1^n, \ldots, m_1^n) \text{ where } (m_1^n, \ldots, m_1^n) \text{ is any permutation of } (m_s, \ldots, m_1)\],

\text{the summation is over all such permutations, } s! \text{ in number and the positive or negative sign is to be taken according as it is an even or an odd permutation. Recalling from (A.9.1) that } \beta \text{ will be zero if any two columns of the pseudo-determinant are equal, let us try to reduce } m_s \text{ to } m_{s-1} \text{ by successive integration by parts. Toward this end consider the typical term in the expansion and in that term let } m_r^n \text{ be the largest}
exponent = m_s (of course). To reduce this exponent by 1 we proceed as follows. By definition

\( a(x; m^1_s, n; \ldots; m^1_{r+1}, n; m^1_s, n; m^1_{r-1}, n; \ldots; m^1_1, n) = \int \frac{x^{m^1_r}}{x^s (1-x)^n} dx_s \)

\[ x \int_{0}^{x^{m^1_r+1}} x^{r+1} \int_{0}^{x^{r+1}} x^{m^1_s (1-x)^n} dx_r \int_{0}^{x^{m^1_{r-1} (1-x)^n}} x^{r-1} dx_{r-1} \ldots \int_{0}^{x^{m^1_1 (1-x)^n}} x_1^{m^1_1} (1-x_1)^n dx_1. \]

Now using (A.9.2) we have

\( b(x; m^1_s, n; \ldots; m^1_{r+1}, n; m^1_s, n; m^1_{r-1}, n; \ldots; m^1_1, n) \)

\[ = \int \frac{x^{r+1}}{x^s (1-x)^n} dx_r \beta(x; m^1_{r-1}, n; \ldots; m^1_1, n) \]

\[ = \frac{1}{m_s + n+1} \int_{-x^{m_s+1}}^{m_s} (1-x^{r+1} (1-x^{r+1})^{n+1}) \beta(x; m^1_{r-1+1}, n; \ldots; m^1_1, n) \]

\[ + \int \frac{x^{r+1}}{x^s (1-x)^n} \beta(x; m^1_{r-1}, n; \ldots; m^1_1, n) dx_r \]

\[ + m_s \int_{0}^{x^{r+1}} x^{m_s} (1-x)^n \beta(x; m^1_{r-1}, n; \ldots; m^1_1, n) dx_r \]

\[ = \frac{1}{m_s + n+1} \int_{-x^{m_s+1}}^{m_s} (1-x^{r+1})^{n+1} \beta(x; m^1_{r-1+1}, n; \ldots; m^1_1, n) \]

\[ + \beta(x; m^1_{r-1+1}, m^1_s, 2^{n+1}; m^1_{r-2}, \ldots; m^1_1, n) \]

\[ + m_s \beta(x; m^1_{r-1+1}, m^1_s, m^1_{r-1}, \ldots; m^1_1, n) \]

(notice that \( \beta(x; m^1_{r-1}, n; \ldots; m^1_1, n) = x_{r}^{m^1_{r-1}} (1-x_{r})^{n} \beta(x; m^1_{r-2}, n; \ldots; m^1_1, n) \) and also
that on the right hand side of (A.9.6.2), the first and second \( \beta \)'s are each an \((r-1)\)-fold integral while the third \( \beta \) is an \( r \)-fold integral. Now substituting the right hand side of (A.9.6.2) we have (A.9.6.1) reducing to

\[
\begin{align*}
(A.9.6.3) \quad &\frac{1}{m_s+n+1}\int \beta(x;m_s',n;\ldots;m_{r-1}'+m_s',2n+1;m_{r-1}',n;\ldots;m_{1}',n) \\
&+ \beta(x;m_s',n;\ldots;m_{r-1}',n;m_{r-1}'+m_s',2n+1;m_{r-2}','n;\ldots;m_{1}',n) \\
&+ m_s \beta(x;m_s',n;\ldots;m_{r-1}',n;m_{s-1}',n;m_{r-1}',n;\ldots;m_{1}',n) \tag{7},
\end{align*}
\]

where the first and second \( \beta \)'s are each an \((s-1)\)-fold integral while the third \( \beta \) is an \( s \)-fold integral with the index \( m_s \) reduced to \( m_s-1 \). It is easy to check through (A.9.6.1) to (A.9.6.3) that the reduction to (A.9.6.3) holds for \( r = s-1,s-2,\ldots,2 \).

If \( r = s \), it is easy to see that (A.9.6.3) will be replaced by

\[
\begin{align*}
(A.9.6.4) \quad &\frac{1}{m_s+n+1}\int -\beta(x;m_s',n+1)\beta(x;m_{s-1}',n;\ldots;m_{1}',n) + \beta(x;m_{s-1}',n+1,m_s',2n+1,m_{s-2}',n;\ldots;m_{1}',n) + m_s \beta(x;m_s',n;\ldots;m_{s-1}',n;m_{s-1}',n;\ldots;m_{1}',n),
\end{align*}
\]

and if \( r = 1 \), (A.12.6.3) will be replaced by

\[
\begin{align*}
(A.9.6.5) \quad &\frac{1}{m_s+n+1}\int -\beta(x;m_s',n;\ldots;m_{3}',n;m_{2}'+m_s',2n+1) + m_s \beta(x;m_s',n;\ldots;m_{2}',n;m_{s-1}',n) \tag{7}.
\end{align*}
\]

We can now use the rather convenient notation

\[
\begin{align*}
(A.9.6.6) \quad &\beta(x;m_s',n;\ldots;m_{r+1}',n+1,n+1;m_{r-1}',n;\ldots;m_{1}',n) = \beta(x;m_s',n;\ldots;m_{r+1}',n+1,m_{s-1}',n;\ldots;m_{1}',n) \tag{7},
\end{align*}
\]

where \((m_{s-1}', n+1)\) is supposed to be added to the \((m_{r+1}', n)\) on the left so as to reduce the integral by one dimension,

\[
\begin{align*}
(A.9.6.7) \quad &\beta_0(x;m_s',n+1)\beta(x;m_{s-1}',n;\ldots;m_{1}',n) = \beta(x;m_{s-1}',n+1,m_{s-1}',n;\ldots;m_{1}',n),
\end{align*}
\]

and

\[
\begin{align*}
(A.9.6.8) \quad &\beta(x;m_s',n;\ldots;m_{r-1}',n+1,m_{r-1}',n;\ldots;m_{1}',n) = \beta(x;m_s',n;\ldots;m_{r-1}',n+1,n+1,m_{r-2}',n;\ldots;m_{1}',n),
\end{align*}
\]

where \((m_{s-1}', n+1)\) is supposed to be added to the \((m_{r-1}', n)\) on the right so as to reduce the integral by one dimension. Using now (A.9.6.1) - (A.9.6.8) we have
\[
(..., 9.6.9) \quad \beta \left[ x; \begin{pmatrix}
\begin{array}{cccc}
m_s, n & m_{s-1}, n & \cdots & m_1, n \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\end{array}
\end{pmatrix}
\right]
= \frac{1}{m_{s+n+1}} \beta \left[ x; \begin{pmatrix}
\begin{array}{cccc}
\_{m_s, n+1} & m_{s-1}, n & \cdots & m_1, n \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\end{array}
\end{pmatrix}
\right]
+ \frac{1}{m_{s+n+1}} \beta \left[ x; \begin{pmatrix}
\begin{array}{cccc}
\rightarrow m_s, n+1 & m_{s-1}, n & \cdots & m_1, n \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\end{array}
\end{pmatrix}
\right]
+ \frac{m_s}{m_{s+n+1}} \beta \left[ x; \begin{pmatrix}
\begin{array}{cccc}
m_{s-1}, n & m_{s-1}, n & \cdots & m_1, n \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\end{array}
\end{pmatrix}
\right],
\]

where in the second pseudo-determinant □ indicates that the corresponding terms in the formal expansion are not to be considered at all, □ being introduced merely to write the pseudo-determinant in a complete form. Recalling the notation (A.9.6.6)-(A.9.6.8) and the lemma (A.9.5) it is easy to see that
\[
(A.9.6.10) \quad \beta \left[ x; \begin{pmatrix}
\begin{array}{cccc}
\_{m_s, n+1} & m_{s-1}, n & \cdots & m_1, n \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\end{array}
\end{pmatrix}
\right]
= s_0(x; m_{s+n+1}) \beta \left[ x; \begin{pmatrix}
\begin{array}{cccc}
m_{s-1}, n & \cdots & m_1, n \\
\cdot & \cdots & \cdot \\
\cdot & \cdots & \cdot \\
\cdot & \cdots & \cdot \\
\end{array}
\end{pmatrix}
\right]
+ \sum_{r=1}^{s-1} (-1)^r \beta \left[ x; \begin{pmatrix}
\begin{array}{cccc}
m_{s-1}, n & \cdots & m_1, n \\
\cdot & \cdots & \cdot \\
\cdot & \cdots & \cdot \\
\cdot & \cdots & \cdot \\
\end{array}
\end{pmatrix}
\right],
\]

where \( \beta \sqrt{J} \) is an \((s-1)\)-fold pseudo-determinant obtained by substituting
\((m_{s+m_{s-1}, n+1}), \ldots, (m_{s+m_1, n+1})\) for \((m_{s-1}, n), (m_{s-2}, n), \ldots, (m_1, n)\) in the \(r\)th row of
the \((s-1)\)-fold pseudo-determinant $\beta \left[ \begin{array}{c} x; \\
\begin{pmatrix}
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n}
\end{pmatrix}
\end{array} \right]$.

Thus \((A.9.6.10) = \beta_0(x; m_{s,n+1}) \beta \left[ \begin{array}{c} x; \\
\begin{pmatrix}
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n}
\end{pmatrix}
\end{array} \right] + \sum_{r=1}^{s-1} (-1)^{r-1} \beta(x; m_{s-r,n+1,2n+1}) \beta_{rr} \left[ \begin{array}{c} x; \\
\begin{pmatrix}
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n}
\end{pmatrix}
\end{array} \right]$; where $\beta_{rr}$ is the \((s-2)\)-fold integral obtained by suppressing the $r^{th}$ row and $r^{th}$ column of the \((s-1)\)-fold pseudo-determinant $\beta$ already referred to. We have likewise

\[(A.9.6.11) \beta \left[ \begin{array}{c} x; \\
\begin{pmatrix}
\scriptstyle m_{s,n+1} & \cdots & \scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s,n+1} & \cdots & \scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s,n+1} & \cdots & \scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n}
\end{pmatrix}
\end{array} \right] = \sum_{r=1}^{s-1} (-1)^{r-1} \beta \left[ \begin{array}{c} x; \\
\begin{pmatrix}
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n}
\end{pmatrix}
\end{array} \right] \beta \left[ \begin{array}{c} x; \\
\begin{pmatrix}
\scriptstyle m_{s,n+1} & \cdots & \scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s,n+1} & \cdots & \scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s,n+1} & \cdots & \scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n}
\end{pmatrix}
\end{array} \right] = \sum_{r=1}^{s-1} (-1)^{r-1} \beta(x; m_{s-r,n+1,2n+1}) \beta_{rr} \left[ \begin{array}{c} x; \\
\begin{pmatrix}
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n}
\end{pmatrix}
\end{array} \right]$.

Now substituting from \((A.9.6.10)\) and \((A.9.6.11)\) in the right side of \((A.9.6.9)\) we have

\[(A.9.6.12) \beta \left[ \begin{array}{c} x; \\
\begin{pmatrix}
\scriptstyle m_{s,n} & \scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s,n} & \scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n} \\
\scriptstyle m_{s,n} & \scriptstyle m_{s-1,n} & \cdots & \scriptstyle m_{1,n}
\end{pmatrix}
\end{array} \right] \]
\[
= - \frac{1}{n^{s+n+1}} \beta_0(x; n, n+1) \beta \left[ x; \begin{pmatrix} m_{s-1, n} & \cdots & m_{1, n} \\ \vdots & \ddots & \vdots \\ m_{s-1, n} & \cdots & m_{1, n} \end{pmatrix} \right] + \frac{2}{n^{s+n+1}} \sum_{r=1}^{s-1} (-1)^{r-1} \beta(x; m_s + m_{s-r}, n+1) \beta \left[ x; \begin{pmatrix} m_{s-1, n} & \cdots & m_{1, n} \\ \vdots & \ddots & \vdots \\ m_{s-1, n} & \cdots & m_{1, n} \end{pmatrix} \right] \\
+ \frac{m_s}{n^{s+n+1}} \beta \left[ x; \begin{pmatrix} m_{s-1, n} & m_{s-1, n} & \cdots & m_{1, n} \\ \vdots & \ddots & \vdots & \vdots \\ m_{s-1, n} & m_{s-1, n} & \cdots & m_{1, n} \end{pmatrix} \right].
\]

It may be noticed that the left hand side is an \( s \)th order pseudo-determinant while, on the right hand side, the first \( \beta \sum \beta \) is an \((s-1)\)th order pseudo-determinant, the second group of terms involves \( \beta \), each such \( \beta \) being an \((s-2)\)th order pseudo-determinant, and the last term has a \( \beta \) which is an \( s \)th order pseudo-determinant with the exponent \( m_s \) reduced to \( m_{s-1} \). It may be also noticed that \( \beta \) may also be conveniently written as

\[
\beta \left[ x; \begin{pmatrix} m_{s-1, n} & m_{s-1, n} & \cdots & m_{1, n} \\ \vdots & \ddots & \vdots & \vdots \\ m_{s-1, n} & m_{s-1, n} & \cdots & m_{1, n} \end{pmatrix} \right].
\]

(4.9.6.12) thus gives us a recurrence relation, whereby, proceeding along the chain and reducing \( m_s \) to \( m_{s-1} \) (in which case the pseudo-determinant will be zero) we have the following reduction of the integral by one dimension.

\[
(4.9.6.13) \quad \beta \left[ x; \begin{pmatrix} m_{s, n} & \cdots & m_{1, n} \\ \vdots & \ddots & \vdots \\ m_{s, n} & \cdots & m_{1, n} \end{pmatrix} \right] = -\beta \left[ x; \begin{pmatrix} m_{s-1, n} & \cdots & m_{1, n} \\ \vdots & \ddots & \vdots \\ m_{s, n} & \cdots & m_{1, n} \end{pmatrix} \right] + \sum_{r=1}^{s-1} \beta_0(x; m_s - r, n+1) \beta \left[ x; \begin{pmatrix} m_{s-1, n} & \cdots & m_{1, n} \\ \vdots & \ddots & \vdots \\ m_{s-1, n} & \cdots & m_{1, n} \end{pmatrix} \right] \sum_{r'=1}^{s-1} \beta_0(x; m_s - r' + 1, n+1) \beta \left[ x; \begin{pmatrix} m_{s-1, n} & \cdots & m_{1, n} \\ \vdots & \ddots & \vdots \\ m_{s-1, n} & \cdots & m_{1, n} \end{pmatrix} \right].
\]
where \((a)_{p}\) stands for \(m(m-1)...(m-p+1)\). The \(s^{th}\) order pseudo-determinant is thus thrown back on \((s-1)^{th}\) and \((s-2)^{th}\) order pseudo-determinants, and these again on \((s-2)^{th}\) and \((s-3)^{th}\) order ones and so on till we get to 1st order pseudo-determinants which are easily evaluated from the incomplete 3-function tables.

\[(A.9.7)\] **Evaluation of the integral**

\[
\int x^{s} x_{1}^{s} \int \ldots \int x_{s}^{s} x_{1}^{s} \ldots \int M \wedge dx_{i} = \beta \int x_{0} x_{i}^{s} a_{s}^{1}, n_{j} \ldots a_{1}, n_{j} \wedge (\text{say})
\]

\[
= \beta \left[ x_{0} x_{i}^{s} \left( \begin{array}{cccc}
  m_{s}, n & m_{s-1}, n & \ldots & m_{1}, n \\
  m_{s}, n & m_{s-1}, n & \ldots & m_{1}, n \\
  \vdots & \vdots & \ddots & \vdots \\
  m_{s}, n & m_{s-1}, n & \ldots & m_{1}, n \\
\end{array} \right) \right] (\text{say}),
\]

where \(M\) stands for the determinant under the integration sign in \((A.9.1)\). We shall also use the notation

\[(A.9.7.1)\] \(\beta(x_{0} x_{i}^{s} a_{s}^{1}, n_{j} ; a_{s-1}, n_{j} ; \ldots ; a_{1}, n_{j})\)

\[
= \beta \int_{x_{0}}^{x_{i}^{s}(1-x_{i}^{s})^{n-1}} \int_{x_{0}}^{x_{i}^{s}(1-x_{i}^{s})^{n-1}} \int_{x_{i}^{s}(1-x_{i}^{s})^{n-1}} \int_{x_{0}}^{x_{i}^{s}(1-x_{i}^{s})^{n-1}} \int_{x_{0}}^{x_{i}^{s}(1-x_{i}^{s})^{n-1}} dx_{i}.
\]

Proceeding now exactly as in sections \((A.9.1)\), \((A.9.2)\), \((A.9.3)\), \((A.9.4)\) and \((A.9.5)\) with obvious modifications at each stage we have in place of \((A.9.6.12)\) the following result

\[(A.9.7.2)\] \(\beta \left[ x_{0} x_{i}^{s} \left( \begin{array}{cccc}
  m_{s}, n & m_{s-1}, n & \ldots & m_{1}, n \\
  m_{s}, n & m_{s-1}, n & \ldots & m_{1}, n \\
  \vdots & \vdots & \ddots & \vdots \\
  m_{s}, n & m_{s-1}, n & \ldots & m_{1}, n \\
\end{array} \right) \right]
\]

\[
= -\beta \left[ x_{0} x_{i}^{s} \left( \begin{array}{cccc}
  m_{s-1}, n & \ldots & m_{1}, n \\
  \vdots & \ddots & \vdots \\
  m_{s-1}, n & \ldots & m_{1}, n \\
\end{array} \right) \right]
\]
\[
  x \sum_{r' = 1}^{m_s - m_s - 1} \left( \frac{(m_s)}{r' - 1} \right) \left( \frac{m_s + n + 1}{r'} \right) \sum_{r = 1}^{s - 1} \frac{m_s - m_s - 1}{(-1)^{r-1}} \beta_0(x; m_s - r' + 1, n + 1) \left( -1 \right)^{s-1} \beta_0(x_0; m_s - r' + 1, n + 1) \left( -1 \right)^{s-1} \\
  + 2 \sum_{r = 1}^{s - 1} \sum_{r' = 1}^{m_s - m_s - 1} (-1)^{r-1} \left[ x, x_0; \left( \begin{array}{cccc}
  m_s - 1, n & \cdots & m_{r+1}, n & m_{r-1}, n & \cdots & m_{l}, n \\
  m_{s-1}, n & \cdots & m_{r+1}, n & m_{r-1}, n & \cdots & m_{l}, n \\
  \vdots & \cdots & \vdots & \cdots & \cdots & \vdots \\
  m_{s-1}, n & \cdots & m_{r+1}, n & m_{r-1}, n & \cdots & m_{l}, n \\
 \end{array} \right) \right] \right] \\
  x \frac{(m_s)}{(m_s + r + 1)} \beta(x, x_0; m_s + m_s - r' + 1, n + 1),
\]

where \((m)_p = m(m-1) \cdots (m-p+1)\) and \(\beta_0(x; m, n)\) stands for \(x^m (1-x)^n\). The \(s\)th order pseudo-determinant is thus thrown back on the \((s-1)\)th and \((s-2)\)th order pseudo-determinants, and so on till we get to first order pseudo-determinants which can be easily evaluated from the incomplete beta function tables.
REFERENCES

The references given are by no means exhaustive, being merely such as would just suffice for an understanding of this report.


REFERENCES (Continued)


17. S. N. Roy, "The individual sampling distribution of the maximum, the minimum and any intermediate one of the p-statistics on the null hypothesis", Sankhyä, Vol. 7 (1945), pp. 133-158.


REFERENCES (Continued)


