STATISTICAL PROJECTION

S. N. Afriat

Econometric Research Program
Research Memorandum No. 16
10 June 1960

Princeton University
Econometric Research Program
92-A Nassau Street
Princeton, N. J.

STATISTICAL PROJECTION

An Investigation of the Role of Orthogonal Projectors in Regression Theory

By

S. N. Afriat

PREFACE

The operation of linear regression in multivariate analysis has, from its first principles, an involvement with Euclidean concepts. Some of these are immediately conspicuous, for instance in the way in which sums of squares and of cross-products feature, or in the fundamental role of orthogonality; and others are arrived at through further construction.

This exposition attempts to exhibit the Euclidean framework in as complete and explicit a form as possible, by reference to a Euclidean space and its transformations, sets of its points and volumes they determine, and the subspaces they span and their mutual relations.

It is a familiar scheme to show regression as resolving vectors of measurements into orthogonal components, one, the regressional part, lying in a space spanned by vectors of measurements, and the other, the residual part, lying in the orthogonal complement of that space. But the step from recognition of this scheme, to consideration of the pair of symmetric idempotent linear transformations which derive these components, defining a complementary pair of orthogonal projectors, and then to systematic formulation of analysis in terms of these projectors, is one that does not appear to have been explored, though basic elements of geometric method are already well established in the subject. The advantage such a step yields, apart from any possibility of finding natural approaches to useful new concepts, are that, joined with a fitting symbolic notation, it leads to a compact and penetrating formalism for handling the familiar processes; and also it gives the terms for some simple proofs and concise formulae. At the same time it makes a valuable enlargement of didactic method, by directly displaying the subject in its intuitively comprehensible algebraical-geometrical aspects, which is illuminating for the description and explanation of what is being done.

Whatever the essential addition that might be made in concepts of statistical analysis, which can be treated as a separate matter, it is a formalism which is to be elaborated, and which is closer to the writer's knowledge. Proofs of the main propositions involved have already been given in a separate algebraical paper (Afriat 1957; and also 1956), of which this is the promised statistical sequel. Now there mainly has to be shown the definitional and propositional scheme in which symmetric idempotent matrices, defining orthogonal projectors, have the fundamental role. Only a most naive, but also the most fundamental, view is taken of multivariate regression analysis; not the more developed one founded on the normal distribution, which justifies the naive view, and in this gets a critical part of its own justification. When reference is made to a variate, variance, covariance or correlation, which terms are properly for concepts relating to a distribution, all that is meant is a factor measured on objects, and certain functions of measurements of factors, which, when a factor is interpreted as a variate, and a trial as a population sample, can also be interpreted as sample measures which, in the case of normality, correspond in the usual way to those things in the population. Reference is made to Halmos (1948) for theory of linear spaces, and about projectors; and to Anderson (1958), and the bibliography there contained, for the general subject of multivariate analysis.

Hotelling (1935) has given an analysis of the relation between two sets of variates, in his canonical correlation theory. It can be formulated as an analysis of the relative position of two subspaces of a Euclidean space, spanned by vectors of measurements, as he has pointed out. This analysis is made complete by a set of orthogonal invariants, which characterize the figure formed by the spaces, and which are given by a certain set of angles determined between them. Such a characterization of a pair of spaces by angles was demonstrated by Jordan (1875), using synthetic methods; and a further such account has been given by Somerville (1929). Algebraical method for determination of the angles has been investigated by Schoute (1905), and more recently by Flanders (1948). Hotelling's theory implies another method, equivalent to consideration of directions whose variation in the spaces leaves the angle between them stationary. A further method is developed in Afriat (1956 and 1957), and has an improved description here, which turns on considera-tion of the characteristic values and latent vectors of the pair of products of the orthogonal projectors on the spaces. Still a further method is stated in a note following this exposition, which consists in the simultaneous transformation to diagonal form of a certain pair of symmetric matrices one of which is positive definite. The generalized problem, for the relation between subspaces of a Hilbert space, has been treated by Dixmier (1948). Another treatment of this problem is suggested here, in an appended note.

The account of canonical analysis, in the form in which it is stated here, founds it directly in its geometrical interpretation, and amplifies the algebra accordingly. Combinations of one set which are uncorrelated with every combination of the other must exist in any case where the variates are different in number, and generally when there is a subspace of the space spanned by the observation vectors of one set which is orthogonal to the same space for the other. This case is the same as that studied by Anderson (1951), in which a matrix of regression coefficients is of defective rank; so the regression components obtained admit between themselves a system of linear relations. This dissociation of one set with another, resulting in dependent sets of regression components, exists automatically when the dimensions are unequal; and it is a mutual relation in the event that the total multiplicity of positive canonical correlations is defective. An entirely different process of calculation applies to these residual components, which have zero correlation with all the rest. Another matter that has wanted a certain elaboration is the treatment of multiple canonical roots, and the complete system of orthogonality relations which hold between canonical variates in this case, irrespective of whether or not they belong to roots which are distinct. It is to be noted that, between the algebraical processes which arise here, and in Mann (1960), there are some striking resemblances of form, though a difference of interpretation.

While a complete set of canonical angles, or their cosines, which define the canonical correlations, give a complete specification of the relation between the spaces, or the sets of variates, it is desirable also to have coefficients which express aspects of this relation. Wilks (1932) and Hotelling (1935) have defined coefficients of correlation, and alienation, which have this character, and which are algebraically equivalent to the here-defined coefficients of inclusion and separation of one space with another. Presented in the geometrical way, and determined in terms of the orthogonal projectors on the spaces, their significance and properties are immediately evident. Some further coefficients are now defined, which have special properties, both in regard to a pair of spaces, and when they are given between the three pairs taken from a set of three spaces, in which case certain inequalities are obtained, limiting the association between two spaces in terms of their association with a third. Coefficients are defined in an analogous way for sets of three or more spaces, which take critical values according to the character of the configuration the spaces form together. There may well be problems for which these various coefficients provide appropriate statistics. One of these coefficients, for a pair of spaces, gives a direct generalization, for sets of arbitrary numbers of variates, of the multiple correlation coefficient, defines between a single variate and a set of several, this itself being a direct generalization of the correlation coefficient of Pearson defined between a simple pair. This coefficient is non-negative, and at most the dimension of each of the spaces, and also of the common dimension of the orthogonal projections of the spaces in each other, this defining their dimension of inclination; it is zero if and only if the spaces are orthogonal; and equal to the dimension of inclination if and only if the spaces cut at right angles, which is to say they are orthogonally incident, with the orthogonal complements in each of their intersections mutually orthogonal; and it is equal to the dimension of one of the spaces if and only if that space is included in the other. The properties of this coefficient have been derived in Afriat (1957); and subsequently it has had consideration, in a different but equivalent form, by Hooper (1957), who has applied it to the simultaneous equation method in econometrics. Another, related coefficient, defined for a pair of spaces, is bounded between zero and unity, being zero just when the spaces are orthogonal, and unity just when they are identical; and it has special properties when taken between the pairs from any set of spaces, such as are considered in an additional note.

A multivariate generalization of the concept of variance has been studied by Wilks (1932) and Anderson (1958). Here a formula is shown which gives the variance of a set of variates, in terms of the separate variances of complementary subsets, together with the coefficient of separation between them. Geometrically, it is a generalization of the familiar formula for the area of a parallelogram, in terms of the lengths of a pair of edges and the angle between them.

If a multiple regression is partitioned, into a sum of complementary partial regressions, then the associated orthogonal projector, represented by a symmetric idempotent, is correspondingly split into a sum of mutually annihilating oblique projectors, that is non-symmetric idempotents whose product together in either order is null. A formula is given for these oblique projectors, which obtain partial regressions, just as orthogonal projectors obtain total regressions; and then a formula is immediately deduced for the matrices of partial regression coefficients. This formula, derived in Afriat (1957), also appears in Scheffé (1959, p. 203). It becomes directly evident that a partial regression matrix, obtained in this way, involving three sets of variates, is identical with the total regression matrix between the residuals in the regressions of two of the sets on the third; which may generalize a proposition at least familiar for when the sets consist of single variates.

The basic formula for the orthogonal projector on the space spanned by a given set of vectors, shown in Afriat (1957), is derived again. This formula is also fundamental for the computations of the gradient-projection method of linear programming, account of which has been given by Rosen (1960).

Further observations are contained in a number of additional notes.

ACKNOWLEDGEMENTS

To Dr. Clive Granger, Mr. Francis Sand, Professor J. Tukey and Professor S. S. Wilks I am indebted for discussions and useful comments. I have to express my thanks to Professor Oskar Morgenstern, who has encouraged and facilitated this work. Also, my thanks are due to Mrs. Ouida Taylor for very skillful assistance in dealing with the manuscript.

Princeton, N. J. S. N. Afriat

TABLE OF CONTENTS

CONTENTS

1. Factors and measurement.

There is to be considered a universe $\mathcal{U}$ of objects. The objects are characterized by factors forming a variety $\mathcal{L}$ of different elements, each object having each simple factor-element to a numerical extent decided by measurement. The result of measuring a simple factor $x \in \mathcal{L}$ on a single object $a \in \mathcal{U}$ is a point in the number scale $\mathcal{F}$, which is denoted by M_x^a, thus,

$$M_x^a \in \mathcal{F} \quad (a \in \mathcal{U} \; ; \; x \in \mathcal{L}) .$$

A combination of simple factors, each measured by a single number, provides a multiple factor, measured by a vector. The measurement of a multiple factor on any object is equivalent to the measurement of each of its elements; so a factor $x \in \mathcal{L}$ of dimension p, or a p-factor, measured on any object $a \in \mathcal{U}$, is represented by a point in a p-dimensional vector space $\mathcal{F}_p$:

$$M_x^a = (M_{x_1}^a, \ldots, M_{x_p}^a) \in \mathcal{F}_p ,$$

where $x = (x_1, \ldots, x_p)$, and $a \in \mathcal{U}$.

2. Experiment.

A trial is constituted by any multiplicity of objects; and an experiment, with some factor, is a trial measurement with that factor, that is the measurement of the factor on all the objects in some trial. There is obtained a measurement matrix, each row of which shows the vector measuring all the elements of the factor on one object of the trial; and correspondingly, each column gives the measurement, on all the objects of the trial, of one element of the factor.

Thus, let an experiment be made with a factor x of dimension p and a trial $\mathcal{T}$ of N objects. Then there is obtained a matrix $M_x^{\mathcal{T}}$ of order $N \times p$, which will be denoted just by M_x when, as is usual, a certain trial $\mathcal{T}$ is taken as implicitly understood. The i^{th} row is a vector $M_x^i \in \mathcal{F}_p$, represented in a vector space $\mathcal{F}_p$ of dimension p, which gives the factor x measured on the i^{th} object of the trial ($i = 1, \ldots, N$). The columns are to be represented as points in a <u>measurement space</u>, which is a Euclidean space $\mathcal{E} = \mathcal{E}_N$ of dimension equal to the multiplicity N of objects in the trial. Thus M_x, first formed as a set of N measurement-vectors of order p, is also considered as a set of p vectors $M_{x_j} \in \mathcal{E}$ ($j = 1, \ldots, p$) of order N, each of which gives the N trial measurements of one element x_j of x.

An experiment may be performed in which various multiple factors $x, y, \ldots$ are measured in conjunction on each of the multiplicity of objects in some trial, its purpose being for the examination of experimentally defined relations between the factors, founded on analysis of their measurement matrices $M_x, M_y, \ldots$ These matrices can be considered as components of order $N \times p, N \times q, \ldots$ of a complete measurement matrix $M_{x,y,\ldots} = (M_x, M_y, \ldots)$ of order $N \times (p + q + \ldots)$, where N is the trial-multiplicity and $p, q, \ldots$ are the different component factor dimensions. They determine sets of points, which are p, q, ... in number, respectively, and in general independent, in a Euclidean space $\mathcal{E}$ of N dimensions.

3. Components.

A factor y is defined to be a component of a factor x if y is measured on any object whenever x is measured on that object, with result determined in the form

$$M_y^a = M_x^a \alpha \quad ,$$

where α defines the coefficient matrix of y as a component of x, with row and column orders equal to the dimensions of x and y. Symbolically, any component y of a factor x may be indicated by

$$y = x\alpha \quad ;$$

and there is the equation

$$M_{x\alpha}^a = M_x^a \alpha \quad .$$

The relation which two factors x, y have when y is a component of x will be denoted by

$$y \lhd x \quad .$$

It is a reflexive and transitive relation, that is

$$x \lhd x \; , \quad x \lhd y \wedge y \lhd z \,.\Longrightarrow.\, x \lhd z \; .$$

If $y \lhd x$, let α_{xy} denote the coefficient matrix of y as a component of x. Then the coefficient matrices, implicitly involved in the reflexivity and transitivity conditions, have the properties

$$\alpha_{xx} = 1 \ , \quad \alpha_{xy}\alpha_{yz} = \alpha_{xz} \ .$$

Two factors x, y will be defined to be equivalent if they are components of each other, the relation thus defined between them, which will be indicated by

$$x \diamond y \ ,$$

being reflexive, symmetric and transitive. It is necessary for equivalence that

$$\alpha_{xy}\alpha_{yx} = 1_p \ , \quad \alpha_{yx}\alpha_{xy} = 1_q \ ,$$

where 1_p, 1_q denote the unit matrices of order p, q the dimensions of x, y. But this is only possible if $p = q$; and then α_{xy}, α_{yx} must be regular, mutually inverse matrices.

The linear spaces, with the naturally defined operations of scalar multiplication and addition, formed by the simple components of a multiple factor x, define its range $\mathcal{R}_x$:

$$y \in \mathcal{R}_x \equiv y \lhd x \ ;$$

and then

$$y \lhd x \iff \mathcal{R}_y \subset \mathcal{R}_x \ .$$

4. Span.

The measurement of a factor x of dimension p in a trial of multiplicity N obtains a set of p measurement vectors $M_i \in \mathcal{E}$ $(i = 1, \ldots, p)$ of order N, these forming the columns of the measurement matrix M_x of order $N \times p$, and represented as points in a Euclidean space $\mathcal{E} = \mathcal{E}_N$ of dimension N . These p vectors in $\mathcal{E}$, in general independent when $N > p$, span a subspace $\mathcal{E}_x \subset \mathcal{E}$, in general of dimension p , which is to define the <u>span</u> of the p-factor x , in the experiment with that factor and the given trial. Thus, having

$$[M_x] = \{M_x\alpha \; ; \; \alpha \in \mathcal{E}_p\}$$

as the range of the matrix M_x , which is the space spanned by the vectors forming its columns, there is made the definition

$$\mathcal{E}_x = [M_x] \; ,$$

for the experimental span of a factor.

The criterion that $\mathcal{E}_x$ attains the dimension p , or equivalently that the p vectors forming the columns of M_x be independent, is that the $p \times p$ matrix $M_x'M_x$ be regular.

Every vector $M \in \mathcal{E}_x$, in the experimental span $\mathcal{E}_x$ of x , is of the form $M = M_x\alpha$, where $\alpha \in \mathcal{E}_p$ is a unique vector of order p given by

$$\alpha = (M_x'M_x)^{-1}M_x'M \quad .$$

Accordingly, any vector $M \in \mathcal{E}$ is the measurement vector $M_y = M_x\alpha$, in the experiment, of a uniquely determined simple component $y = x\alpha$ of x . Conversely, the measurement vector $M_{x\alpha} = M_x\alpha$ of any component $x\alpha$ of x belongs to the span $\mathcal{E}_x$. In this way there is an isomorphism between the range $\mathcal{R}_x$ of x , that is the linear space formed by its components, and the span $\mathcal{E}_x$ of x obtained in any experiment:

$$y \in \mathcal{R}_x \Leftrightarrow M \in \mathcal{E}_x \quad (M = M_y \; ; \; y = x\alpha) \; .$$

The residual span of x, in the experiment, is now defined as the orthogonal complement $\bar{\mathcal{E}}_x$ in $\mathcal{E}$ of the span $\mathcal{E}_x$, thus:

$$\bar{\mathcal{E}}_x = \mathcal{E} \ominus \mathcal{E}_x \; .$$

A variety of experimental relations between multiple factors may be founded on the various possible relations between their experimental spans, which are to be instrumental for deciding expectations about one factor from knowledge of the other.

5. Projectors.

A p-factor x in an N-trial experiment has span $\mathcal{E}_x$, which is a subspace of dimension p in a space $\mathcal{E}$ of dimension N, and residual span $\bar{\mathcal{E}}_x$ of dimension $N-p$, so that

$$\mathcal{E}_x \perp \bar{\mathcal{E}}_x \; , \quad \mathcal{E}_x \oplus \bar{\mathcal{E}}_x = \mathcal{E} \; .$$

Every vector $Z \in \mathcal{E}$ has a unique resolution into a sum of components Z_x, $\bar{Z}_x$ in the span and the residual span of x, which is determined by linear transformations e_x, $\bar{e}_x$ of Z:

$$Z = Z_x + \bar{Z}_x \; ,$$

where

$$Z_x = e_x Z \in \mathcal{E}_x \; , \quad \bar{Z}_x = \bar{e}_x Z \in \bar{\mathcal{E}}_x \; .$$

The thus-defined linear transformation e_x is called the projector on the space $\mathcal{E}_x$, it being more explicitly the projector on $\mathcal{E}_x$ parallel to its orthogonal complement $\bar{\mathcal{E}}_x$, that is to say the orthogonal projector on $\mathcal{E}_x$; and the component Z_x defines the <u>projection on</u> $\mathcal{E}_x$ of Z. The similarly defined linear transformation $\bar{e}_x$ is the complementary projector of e_x, it being the projector on the orthogonal complement $\bar{\mathcal{E}}_x$ of $\mathcal{E}_x$.

It is directly evident that

$$e_x + \bar{e}_x = 1 \ , \quad e_x{}'e_x = 0 \ ,$$

from which it follows that

$$e_x{}' = e_x \ , \quad e_x^2 = e_x \ ;$$

and similarly for $\bar{e}_x$. Thus the complementary orthogonal projectors are complementary symmetric idempotent linear transformations, that is with sum equal to the identity, with the range of each the same as the null-space of the other, and the ranges together forming the complementary orthogonal subspaces $\mathcal{E}_x$ and $\bar{\mathcal{E}}_x$.

Thus the projector e_x on $\mathcal{E}_x$ is a symmetric idempotent with $\mathcal{E}_x$ as its range: $\mathcal{E}_x = e_x\mathcal{E}$. Conversely, any symmetric idempotent with $\mathcal{E}_x$ as range must be identical with e_x.

The trace of any idempotent is equal to its rank, which is the same as the dimension of its range. Therefore

$$\text{trace } e_x = \text{rank } e_x = p \ .$$

Consider the matrices

$$E_x = M_x(M_x{}'M_x)^{-1}M_x \ , \quad \bar{E}_x = 1 - E_x \ ,$$

defined in the case of $\mathcal{E}_x = [M_x]$ being of full dimension p ;

in which case $|M_x'M_x| \neq 0$. They are complementary symmetric idempotents, with

$$[E_x] \subset [M_x] = \mathcal{E}_x ,$$

where the inclusion will be an identity if the spaces have the same dimension. But

$$\text{trace } E_x = \text{trace } M_x(M_x'M_x)^{-1}M_x'$$

$$= \text{trace } M_x'M_x(M_x'M_x)^{-1}$$

$$= \text{trace } 1_p = p ;$$

so the dimensions are the same. Therefore

$$[E_x] = \mathcal{E}_x ;$$

and hence the orthogonal projector e_x, on $\mathcal{E}_x = [M_x]$, is calculated by the formula

$$e_x = E_x ;$$

whence also

$$\bar{e}_x = \bar{E}_x .$$

From the measurement matrices M_x, M_y of factors x, y in some joint experiment, there can be computed the projectors e_x, e_y on their spans $\mathcal{E}_x$, $\mathcal{E}_y$; and then criteria for the various relations between the spans can be applied to these. For, since $\mathcal{E}_x$, $\mathcal{E}_y$ can be derived as the ranges of the matrices e_x, e_y which are defined as the unique symmetric idempotents having these spaces for ranges, any condition on the spaces can be expressed as a condition on the projectors.

6. Orthogonality.

The condition that $\mathcal{E}_x = e_x\mathcal{E}$, $\mathcal{E}_y = e_y\mathcal{E}$ be orthogonal is that $e_x'e_y = 0$. But e_x, e_y are symmetric; so the condition is $e_xe_y = 0$, and equivalently $e_ye_x = 0$, thus showing the condition for the orthogonality of the spaces in terms of the projectors on them.

7. Incidence.

Two spaces are said to be incident if they have a non-null intersection, and otherwise to be separate. The criterion for the incidence of the spaces $\mathcal{E}_x$, $\mathcal{E}_y$ is expressible in terms of the orthogonal projectors e_x, e_y on them by the condition

$$|1 - e_xe_y| = 0 .$$

An algorithm for obtaining the intersection of the spaces is by determining it as the null-space of the matrix $1 - e_xe_y$, or equivalently, of the matrix $1 - e_ye_x$; for

$$Z \in \mathcal{E}_x \cap \mathcal{E}_y \iff (1 - e_xe_y)Z = 0 .$$

8. Orthogonal incidence.

Spaces are said to be orthogonally incident if the orthogonal complements in each of their intersections are mutually orthogonal

$$\mathcal{E}_x \ominus \{\mathcal{E}_x \cap \mathcal{E}_y\} \perp \mathcal{E}_y \ominus \{\mathcal{E}_x \cap \mathcal{E}_y\} ;$$

and otherwise they are obliquely incident. The criterion for orthogonal incidence, in terms of the orthogonal projectors on the spaces, is given by the condition

$$e_y e_x = e_x e_y \quad .$$

In this, and only this case, $e_x e_y$ is a symmetric idempotent, with the intersection of the spaces for its range; and it is therefore the orthogonal projector on this intersection.

9. Inclination.

If two spaces are such that there is no subspace of one which is orthogonal to the other, then the one space is said to be inclined to that other, and otherwise to be disinclined. Thus, for $\mathcal{E}_y$ to be inclined to $\mathcal{E}_x$, which relation between the spaces will be indicated by

$$\mathcal{E}_y \ \text{◖} \ \mathcal{E}_x \ ,$$

the condition is that $\mathcal{E}_y$ has no intersection with the orthogonal complement $\bar{\mathcal{E}}_x$ of $\mathcal{E}_x$; and the criterion is therefore

$$|1 - \bar{e}_x e_y| \neq 0 \quad .$$

For the inclination of one space to the other, it is necessary that the dimension of the one space be at most the dimension of the other:

$$\mathcal{E}_y \ \text{◖} \ \mathcal{E}_x \Rightarrow q \leq p \quad .$$

Therefore if the spaces are completely inclined, each being inclined to the other, which relation is indicated by

$$\mathcal{E}_y \ \text{●} \ \mathcal{E}_x \ ,$$

they must be of the same dimension: $q = p$; otherwise the spaces will be incompletely inclined. If

$$r = \text{rank } e_x e_y \leq \text{rank } e_y = q$$

defines the dimension of inclination r of the spaces, then

$$\mathcal{E}_y \,◖\, \mathcal{E}_x \Longleftrightarrow r = q \,.$$

10. Inclusion.

Consider the condition $e_x e_y = e_y$, which, by the symmetry of e_x, e_y is equivalent to the condition $e_y e_x = e_y$. It is that $e_x(e_y Z) = (e_y Z)$ for all Z; that is, $e_x Y = Y$ for all $Y = e_y Z \in \mathcal{E}_y$. But $e_x Y = Y$ if and only if $Y \in \mathcal{E}_x$. Thus the considered condition is that $\mathcal{E}_y \subset \mathcal{E}_x$.

Thus if spaces are given, by sets of base vectors, their inclusions may be thus decided from the orthogonal projectors on them, computed from the base vectors by the formula which has been given.

11. Imitation and dependence.

If two factors x, y have equal measurements for all objects in a trial $\mathcal{J}$, thus

$$M_y^a = M_x^a \ (a \in \mathcal{J}) \,, \quad \text{or} \quad M_y = M_x \,,$$

they are said to imitate each other in the trial, which relation they have may be also indicated by

$$y \underset{\mathcal{J}}{=} x \,.$$

More generally, one factor y may be said to be experimentally dependent on another factor x, in regard to some trial $\mathcal{J}$, if it is imitated by some component $x\alpha$ of x , thus:

$$y \underset{\mathcal{J}}{=} x\alpha \ ;$$

or equivalently,

$$M_y = M_x\alpha \ .$$

The experimental dependence relation thus defined between y and x may be indicated by

$$y \prec x \ ,$$

with the trial $\mathcal{J}$ understood. It is evident that if y is a component of x , then it is dependent on x in every experiment:

$$y \lhd x \Rightarrow y \prec x \ .$$

The relation of experimental dependence between factors is equivalent to the relation of inclusion between their spans:

$$y \prec x \Leftrightarrow \mathcal{E}_y \subset \mathcal{E}_x \ .$$

There is thus the criterion

$$y \prec x \Leftrightarrow e_y = e_x e_y$$

for experimental dependence in terms of the orthogonal projectors on the spans, from which it follows that

$$y \prec x \Leftrightarrow M_y = M_x r_{xy} \ ,$$

where

$$r_{xy} = (M_x{}'M_x)^{-1}M_x{}'M_y \ .$$

But

$$M_x r_{xy} = M_{xr_{xy}} \ .$$

Accordingly, if y is experimentally dependent on x , as decided by the condition $e_y = e_x e_y$, then it is experimentally imitated by the component xr_{xy} of x, which may be denoted by $\hat{y}(x)$:

$$y \underset{\mathcal{J}}{\equiv} \hat{y}(x) , \quad \text{equivalently} \quad M_y = M_{\hat{y}(x)} ,$$

where $\hat{y}(x) = xr_{xy}$.

Experimental equivalence is defined by experimental mutual dependence. The criterion is the identity of spans, for which the condition is $e_y = e_x$. Experimentally equivalent factors must be of the same dimension, with r_{xy} and r_{yx} regular and mutually inverse:

$$r_{xy}r_{yx} = r_{yx}r_{xy} = 1 .$$

Factors which are incident, in that their spans in an experiment are incident, are such that they have common factors, given by components of each which imitate each other in the experiment. The common factors are represented by parts of their ranges which are experimentally identified. Thus, if $\mathcal{E}_x \cap \mathcal{E}_y \neq 0$, then there exist components xr, ys of x, y which are measured equal on all the objects in the trial.

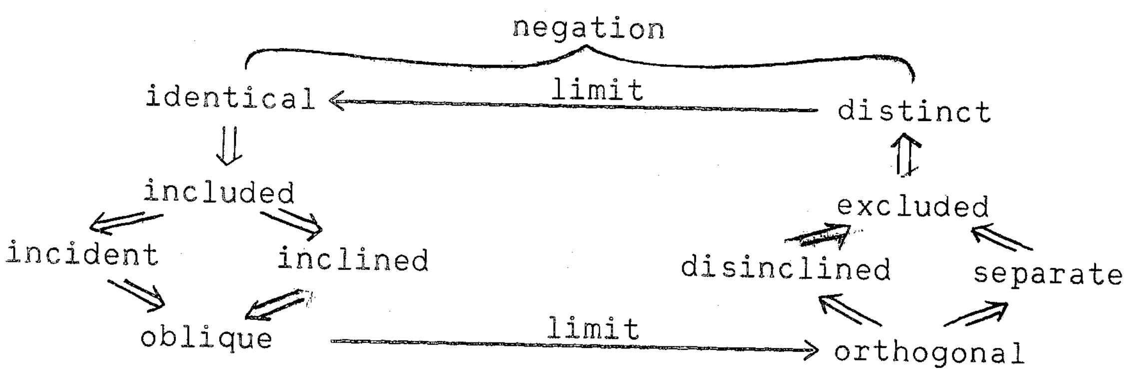
$$M_{xr} = M_{ys} .$$

They can be considered as determining a common factor z of x, y, whose span lies in the intersection of their spans.

12. Configuration coefficients.

Two linear subspaces of a Euclidean space may be identical; one may be included in the other, or each may contain a part excluded from the other; they may be separate or incident; oblique or orthogonal; one may be inclined to the other, with no elements in the one orthogonal to all elements of the other, or disinclined; they may be orthogonally incident, with the orthogonal complements in each of their intersections mutually orthogonal, or obliquely incident. With all these various relations, which are invariant under orthogonal transformations, and therefore characteristic of the configuration formed by the subspaces in the Euclidean space, some are opposites, being negations of each other, and some form sequences of increasing restriction. There are to be considered some different coefficients, that is non-negative real-valued functions defined over the pairs of spaces, which have certain critical values when the spaces exactly attain to certain relations; but more generally, they each will put the configuration formed by a pair of spaces in a scale, measuring the extent to which there is attainment to or departure from some relation, in a way which is invariant under orthogonal transformations.

These considered relations form a scheme in which certain pairs are linked by negation or by implication or by one being a limiting form of the other:



Now various coefficients are to be introduced in the following theorems, each of which, by its properties there stated, fulfills a role as a configuration coefficient, having certain critical values when the spaces satisfy certain relations.

THEOREM: _If_ $R^2_{xy} = \text{trace } e_x e_y$, $d_x = \text{trace } e_x = \text{rank } e_x$ _and_ $d_{xy} = \text{rank } e_x e_y$ _then_

$$0 \leq R^2_{xy} \leq d_{xy} \leq dx \quad ,$$

$$R^2_{xy} = 0 \iff \mathcal{E}_x \text{ orthogonal to } \mathcal{E}_y \; (e_x e_y = 0)$$

$$= d_{xy} \iff \mathcal{E}_x \text{ orthogonally incident with } \mathcal{E}_y \; (e_x e_y = e_y e_x)$$

$$= d_x \iff \mathcal{E}_x \text{ included in } \mathcal{E}_y \; (e_x e_y = e_x)$$

The critical values of R^2_{xy} are thus 0 , the dimensions d_x, d_y of the spaces $\mathcal{E}_x$, $\mathcal{E}_y$ and their dimension of inclination d_{xy} , which is the dimension of the orthogonal projection of either space on the other; and its general values can be taken as measure of <u>correlation</u> between the spaces.

COROLLARY: _If_ $Q^2_{xy} = \text{trace } e_x \bar{e}_y$ _then_ $Q^2_{xy} + R^2_{xy} = d_x$, $0 \leq Q^2_{xy} \leq d_x$

$$Q^2_{xy} = 0 \iff \mathcal{E}_x \text{ included in } \mathcal{E}_y$$

$$= d_x \iff \mathcal{E}_x \text{ orthogonal to } \mathcal{E}_y$$

Here it appears that the dimension d_x of $\mathcal{E}_x$ is partitioned into two parts Q^2_{xy} and R^2_{xy} which measure the extents to which this space is associated with the other, and with the orthogonal complement, respectively. While R_{xy} symmetrically measures the correlation between the spaces, Q_{xy} is to measure the residual correlation of $\mathcal{E}_x$ to $\mathcal{E}_y$.

COROLLARY: If $C_{xy}^2 = (\text{trace } e_x e_y)^2/(\text{trace } e_x)(\text{trace } e_y)$ then

$$0 \leq C_{xy} \leq 1$$

$$C_{xy} = 0 \iff \mathcal{E}_x, \mathcal{E}_y \text{ orthogonal } (e_x e_y = 0)$$

$$= 1 \iff \mathcal{E}_x, \mathcal{E}_y \text{ identical } (e_x = e_y) .$$

Thus C_{xy}, which is to define the coefficient of association between the spaces, appears with properties like those of the cosine of the acute angle between a pair of directions; and it sets a pair of spaces in a scale between extremes of orthogonality and identity. Similarly, the non-negative real number S_{xy} satisfying

$$S_{xy}^2 = 1 - C_{xy}^2 ,$$

which is to define the coefficient of dissociation, has properties analogous to the sine.

THEOREM: If $s_{xy} = |1 - e_x e_y|^{\frac{1}{2}}$ then

$$0 \leq s_{xy} \leq 1$$

$$s_{xy} = 0 \iff \mathcal{E}_x, \mathcal{E}_y \text{ incident}$$

$$= 1 \iff \mathcal{E}_x, \mathcal{E}_y \text{ orthogonal} .$$

The number s_{xy} has properties again analogous to the sine. It measures the separation of the spaces, being a minimum when they are incident, and a maximum when they are orthogonal.

COROLLARY: If $c_{xy} = |1 - e_x \bar{e}_y|^{\frac{1}{2}}$ then

$$0 \leq c_{xy} \leq 1$$

$$c_{xy} = 0 \iff \mathcal{E}_x \text{ disinclined to } \mathcal{E}_y$$

$$= 1 \iff \mathcal{E}_x \text{ included in } \mathcal{E}_y$$

Thus c_{xy}, analogous to a cosine, measures the extent of inclusion of one space in relation to another. It is the application of the coefficient s to one space and the orthogonal complement of the other, and is a maximum when the relation of inclusion is attained, and a minimum when there is a part of one space which is orthogonal to the other. Like the Q-coefficient, but unlike the others, it is not a symmetric function of the two spaces.

13. Parallelipiped volumes.

Any set of vectors determines a parallelipiped, with the origin for a vertex, and with fundamental edges, at that vertex, given by the vectors. The space spanned by the fundamental edges contain the parallelipiped, and define its supporting space; and the dimension of the supporting space gives the dimension of the parallipiped, which is called regular if the dimension and the number of the fundamental edges are equal.

A matrix M_x of order $N \times p$, considered as making by its columns a set of p vectors of order N, spanning a subspace $\mathcal{E}_x = [M_x]$ in the Euclidean space $\mathcal{E}$ of dimension N, determines a parallelipiped π_x whose volume V_x is the non-negative real number determined by the formula

$$V_x = |M_x' M_x|^{\frac{1}{2}} ,$$

and is positive just when the parallelipiped is regular.

Two parallelipipeds together determine a third, their resultant, whose fundamental edges are their fundamental edges taken together. If the parallelipipeds are regular, and their supporting subspaces separate, then and only then their resultant is regular, of dimension equal to the sum of their dimensions, and with supporting space the union of their supporting spaces; and the volume is the product of their volumes with the coefficient of separation between the supporting spaces. Thus, with $M_{x,y} = (M_x, M_y)$, there is the identity

$$|M_{x,y}' M_{x,y}| = |M_x' M_x| |M_y' M_y| |1 - e_x e_y| ,$$

from which it follows that

$$V_{x,y} = V_x V_y s_{xy} ,$$

generalizing the familiar formula for the area of a parallelogram in terms of the lengths of a pair of edges and the sine of the angle between them.

Now the parallelipiped π_x may be orthogonally projected onto the orthogonal complement $\overline{\mathcal{E}}_y$ of the space $\mathcal{E}_y$, to obtain a parallelipiped which may be denoted by $\bar{e}_y \pi_x$, whose fundamental edges are determined by the matrix $\bar{e}_y M_x$; and the square of its volume is, by the symmetry and the idempotence of $\bar{e}_y$, given by

$$|(\bar{e}_y M_x)'\bar{e}_y M_x| = |M_x'\bar{e}_y'\bar{e}_y M_y|$$

$$= |M_x'\bar{e}_y M_x| \quad .$$

But there is the identity

$$|M_x'\bar{e}_y M_x| = |M_x'M_x||1 - e_x e_y| \quad .$$

So it follows that the volume of the projected parallelipiped is the product of the original one with the coefficient of separation between its supporting space and the orthogonal complement of the space onto which it is projected.

14. Multivariance.

The multivariance of a multiple factor in an experiment is defined by the volume of the parallelipiped which has for a set of edges at a vertex the experimental measurement vectors of its elements.

Thus the multivariance of a factor x of dimension p is the volume V_x of the p-dimensional parallelipiped which has edges determined by the vectors forming the columns of the measurement matrix M_x . The multivariance of a simple factor, or simply its variance, is given by the length of its measurement vector.

Two multiple factors x, y of dimension p, q taken together constitute a further multiple factor (x, y) of dimension $p + q$. The multivariance of two factors thus taken in conjunction is determined by their separate multivariances together with the coefficient of separation between them, according to the formula

$$V_{x,y} = V_x V_y \, |1 - e_x e_y|^{\frac{1}{2}} \; ,$$

$$= V_x V_y s_{xy} \; .$$

15. Projection and regression.

Consider an experiment in which a pair of multiple factors are measured together on a multiplicity of objects. The experience thus gained concerning the association of the factors has to be analyzed from the data of measurements, in order to have, on any new object, an expectation of one factor, from an inspection of the other.

The factors x, y in the experiment have measurement matrices M_x, M_y and spans $\mathcal{E}_x, \mathcal{E}_y$ on which the orthogonal projectors e_x, e_y are computed.

If y is experimentally dependent on x, for which the criterion is

$$e_x e_y = e_y \; ,$$

then

$$M_y = M_{x r_{xy}} = M_{\hat{y}(x)}$$

so that y is exactly imitated by the component $\hat{y}(x) = x r_{xy}$ of x throughout the experiment. The expectation then is that the imitation will persist on a new object a; so that the expected measurement $\hat{M}_y^a$ of y on a, when only x is observed, with result M_x^a, is defined by

$$\hat{M}_y^a = M_{\hat{y}(x)}^a = M_{xr_{xy}}^a = M_x^a r_{xy} \quad .$$

However, y not being given as a component of x, it will not generally be found experimentally to be exactly dependent on x. There will be a discrepancy D_r between the measurement matrices M_y of y and $M_{xr} = M_x r$ of any component xr of x, thus:

$$M_y = M_{xr} + D_r \quad (D_r \neq 0) \quad .$$

This discrepancy matrix D_r can be considered the measurement matrix of a further factor, to be denoted by $y - xr$:

$$D_r = M_{y-xr} \quad .$$

It appears that the value of the coefficient matrix r which obtains the sum trace $D_r'D_r$ of the squares of the elements of D_r an absolute minimum is given by

$$r = r_{xy} \quad ,$$

where again

$$r_{xy} = (M_x'M_x)^{-1}M_x'M_y \quad .$$

Accordingly, corresponding to this value of r, there is the resolution of the measurement matrix M_y of y into two parts, associated, and dissociated with x, respectively, each with the form

$$M_{\hat{y}(x)} = M_{xr_{xy}} = e_x M_y \ , \quad M_{y-\hat{y}(x)} = D_{r_{xy}} = \bar{e}_x M_y \ ,$$

thus:

$$M_y = e_x M_y + \bar{e}_x M_y = M_{\hat{y}(x)} + M_{y-\hat{y}(x)}$$

The function $\hat{y}(x)$ of x obtained in this way defines the value of y to be expected when x is given. Then $y - \hat{y}(x)$ is the deviation of y from its expected value. Evidently, since $e_x \bar{e}_x = 0$,

$$\widehat{(y - \hat{y}(x))}(x) = 0 \quad ,$$

or the expected deviation of y from its expected value is zero, on all the objects in the experiment.

The component $\hat{y}(x) = xr_{xy}$ of x thus determined experimentally in regard to y is to be called the regressional image of y in x , in the experiment; and r_{xy} defines the regression coefficient matrix of y on x . The part $e_x M_y$ in the corresponding resolution of M_y constitutes the regressional part, it being the measurement matrix of the regressional image $\hat{y}(x) = xr_{xy}$; and $\bar{e}_x M_y = M_{y-\hat{y}(x)}$ is the residual part, being the measurement matrix of $y - \hat{y}(x)$. There is said to be a null regression when the regressional part is null, and a perfect regression when the residual part is null.

The regressional image of y in x in an experiment can be defined again as that uniquely determined component xr of x whose measurement vectors, forming $M_{xr} = M_x r$, are the orthogonal projections $e_x M_y$ in the span $\mathcal{E}_x$ of x of the measurement vectors M_y of y , thus:

$$M_x r = e_x M_y \quad ;$$

for, from this relation, it follows that

$$M_x{}'M_x r = M_x{}'e_x M_y = M_x{}'M_y ,$$

and then that

$$r = r_{xy} .$$

This form of the derivation of the component xr_{xy} of x in relation to y in an experiment may be called statistical projection; and the operations of least squares regression and statistical projection have here appeared equivalent.

If the x-image of y is not null, y is said to be correlated with x, in the experiment; and the condition is that the span of y be oblique to that of x. But obliquity is a symmetrical relation between the spans, whence so is correlation between the factors. Now complete correlation of y with x is defined by the condition that none of the elements of its x-image xr_{xy} be null; and this is equivalent to there being no part of the span $\mathcal{E}_y$ of y which is orthogonal to the span $\mathcal{E}_x$ of x, which is to say that $\mathcal{E}_y$ is inclined to $\mathcal{E}_x$, the non-symmetry of which relation carries with it the non-symmetry of the relation of complete correlation.

From a regressional decomposition

$$M_y = e_x M_y + \bar{e}_x M_y = M_{\hat{y}} + M_{y-\hat{y}}$$

of M_y there follows the decomposition

$$M_y{}'M_y = M_y{}'e_x M_y + M_y{}'\bar{e}_x M_y = M_{\hat{y}}{}'M_{\hat{y}} + M_{y-\hat{y}}{}'M_{y-\hat{y}}$$

of $M_y{}'M_y$ into two parts with a similar form, using the symmetry and idempotence of the projectors; and then, taking traces, the decomposition

$$\text{trace } M_y{}'M_y = \text{trace } M_{\hat{y}}{}'M_{\hat{y}} + \text{trace } M_{y-\hat{y}}{}'M_{y-\hat{y}}$$

of the sum of the squares of the elements of M_y into sums of squares of elements of $M_{\hat{y}}$ and $M_{y-\hat{y}}$. The second part gives the absolute minimum value of the sum of the squares of the discrepancies $D_r = M_{y-xr}$ in the dependence of y on x. Between the extremes of being perfect and null, a regression in general has a certain intermediate extent, which may be measured by a comparison of the parts trace $M_y{}'e_xM_y$ and trace $M_y{}'\bar{e}_xM_y$ in the partition of trace $M_y{}'M_y$. Alternatively, multiply by $(M_y{}'M_y)^{-1}$ first, and then take traces; and by a cyclical permutation of factors, which leaves traces unchanged, there follows the decomposition

$$q = R^2_{xy} + Q^2_{xy}$$

of the dimension of y into the coefficients of correlation and residual correlation of y with x.

Now the variances of y, and of the x-image and x-residual of y, are given by

$$V^2_y = |M_y{}'M_y| , \quad V^2_{\hat{y}(x)} = |M_y{}'e_xM_y| , \quad V^2_{y-\hat{y}(x)} = |M_y{}'\bar{e}_xM_y| ,$$

and then their comparison is shown in the identities

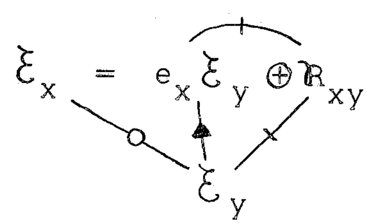
$$V_{\hat{y}(x)}/V_y = c_{yx} , \quad V_{y-\hat{y}(x)}/V_y = s_{xy} ,$$

where c_{yx}, the coefficient of inclusion of y in x, is a maximum just when the regression of y on x is perfect, and a minimum just when it is null, and where s_{xy}, the coefficient of separation, is a maximum when the regression is null, and a minimum when the spaces are incident, in which case there are components of one factor which regress perfectly on the other.

16. <u>Principal reduction</u>.

Any two spaces which are oblique may be reduced to a pair of components which are inclined to each other, and a pair which are orthogonal to these and also to each other. The principal pair, which are mutually inclined, are obtained as the orthogonal projections of the spaces in each other; and the other, the residual pair, are the orthogonal complements in each of the spaces of either space in the principal pair. The principal pair also have the property of being reciprocals, in that they are the orthogonal projections of each other in the spaces.

The fundamental proposition from which this reduction is obtained is expressed in the scheme:



where the relations indicated are that $\mathcal{E}_x$ and $\mathcal{E}_y$ are oblique, that is $e_x \mathcal{E}_y$ is inclined to $\mathcal{E}_y$, and that $\mathcal{R}_{xy}$, which appears as orthogonal to both $e_x \mathcal{E}_y$ and $\mathcal{E}_y$, can be defined in three equivalent ways:

$$\mathcal{R}_{xy} = \begin{cases} \mathcal{E}_x \ominus e_x\mathcal{E}_y \\ \mathcal{E}_x \ominus \mathcal{E}_y \\ \mathcal{E}_x \ominus e_y\mathcal{E}_x \end{cases},$$

that is, as the orthogonal complement in $\mathcal{E}_x$ of the orthogonal projection in $\mathcal{E}_x$ of $\mathcal{E}_y$, or of simply $\mathcal{E}_y$, or of the orthogonal projection of $\mathcal{E}_x$ in $\mathcal{E}_y$. Applying this scheme to one space with the other, and then again with the spaces interchanged, there is obtained the further scheme

$$\begin{aligned} \mathcal{E}_x &= e_x\mathcal{E}_y \oplus \mathcal{R}_{xy} \\ \mathcal{E}_y &= e_y\mathcal{E}_x \oplus \mathcal{R}_{yx} \end{aligned}$$

which defines the principal reduction of the pair.

The reciprocal property of the principal components is stated by the relation

$$e_x(e_y\mathcal{E}_x) = e_x\mathcal{E}_y$$

and the same relation with the spaces interchanged.

An algorithm for obtaining the reduction is by computing $e_x\mathcal{E}_y$ as the range of e_xe_y, and $\mathcal{R}_{xy}$ as the null-space of $1 - e_x\bar{e}_y$, the orthogonal projectors having been computed, from any bases for the spaces, by the formula which has been given.

17. Reciprocal and reflexive directions.

In a pair of spaces, reciprocal directions are those which are the orthogonal projections of each other in the spaces. Thus

$$e_x \mathcal{V} = \mathcal{U} \,, \quad e_y \mathcal{U} = \mathcal{V} \,,$$

for reciprocal directions $\mathcal{U}$, $\mathcal{V}$ in spaces $\mathcal{E}_x$, $\mathcal{E}_y$ on which the orthogonal projectors are e_x, e_y. The products

$$e_x e_y \,, \quad e_y e_x$$

define the dual pair of biprojectors on the spaces. Since the orthogonal projectors are symmetric, the dual biprojectors are the transposes of each other.

A reflexive direction in one space relative to another is one which is identical with the orthogonal projection in that space of its orthogonal projection in the other. Accordingly,

$$\mathcal{U} = e_x e_y \mathcal{U}$$

for a reflexive direction $\mathcal{U}$ in $\mathcal{E}_x$ in regard to $\mathcal{E}_y$. Evidently a reciprocal pair of directions in the spaces are each a reflexive direction in one space in regard to the other. Conversely, any reflexive direction in one space in regard to another forms with its orthogonal projection in the other a reciprocal pair of directions in the spaces; for

if $\mathcal{U} = e_x e_y \mathcal{U}$, and $\mathcal{V} = e_y \mathcal{U}$, then $\mathcal{U} = e_x \mathcal{V}$.

Thus the determination of the reciprocal directions between a pair of spaces is equivalent to the determination of the reflexive directions in either one in regard to the other; and these are the invariant directions of either one of the biprojectors.

18. Proper angles.

A proper angle between the spaces is defined as an angle made by a pair of reciprocal directions. Thus, if $\mathcal{U}, \mathcal{V}$ are a pair of reciprocal directions, the acute angle $\widehat{\mathcal{U},\mathcal{V}}$ which they make between them determines a proper angle between the spaces. If U, V is any pair of vectors spanning the directions $\mathcal{U}, \mathcal{V}$ then

$$\cos^2 \widehat{\mathcal{U},\mathcal{V}} = (U'V)^2/(U'U)(V'V) ,$$

where

$$0 < \cos \widehat{\mathcal{U},\mathcal{V}} \leq 1 , \quad \text{with} \quad 0 \leq \widehat{\mathcal{U},\mathcal{V}} < \pi/2 .$$

While any reciprocal pair of directions determines a unique proper angle, a given proper angle may be made by many distinct reciprocal pairs.

19. Reciprocal vectors.

Now reciprocal vectors are defined as those which span reciprocal directions. Thus, if U, V are reciprocal vectors in $\mathcal{E}_x, \mathcal{E}_y$ then

$$e_x V = U\rho , \quad e_y U = V\sigma ,$$

for some multipliers $\rho, \sigma \neq 0$, which, it appears, are given by

$$\rho = (U'U)^{-1}U'V , \quad \sigma = (V'V)^{-1}V'U ;$$

from which it follows that

$$\rho\sigma = \cos^2 \widehat{U,V} .$$

Moreover,

$$e_x e_y U = U\lambda \ , \quad e_y e_x V = V\lambda \ ,$$

where $\lambda = \rho\sigma$; so U, V are latent vectors of the dual biprojectors $e_x e_y$, $e_y e_x$ with common characteristic value

$$\lambda = \rho\sigma = \mu^2 \ ,$$

which is the product of the multipliers ρ, σ and also is the square of the cosine

$$\mu = \cos \widehat{U, V}$$

of the angle between them, which is also a proper angle between the spaces.

If, further, U, V are unit vectors, then

$$\rho = U'V = \sigma \ .$$

In this case, $\sigma = \mu\varepsilon$ where $\varepsilon = \pm 1$; and then

$$e_y U = (V\varepsilon)\mu \ , \quad e_x(V\varepsilon) = U\mu \ ,$$

showing U, $V\varepsilon$ to be a reciprocal pair of unit vectors with multipliers which are equal to each other, and positive; with which properties they are to define a <u>normal</u> reciprocal pair.

The characteristic values of the dual biprojectors on the spaces are the same, and non-negative; and they are all zero only if the spaces are orthogonal, which is when the biprojectors are null. Therefore, if the spaces are oblique, let μ^2 denote any non-zero characteristic value, where it is taken that $\mu > 0$; and let U be any corresponding unit latent vector of $e_x e_y$, so that

$$e_x e_y U = U\mu^2 \ , \quad U'U = 1 \ ;$$

and define V by

$$e_y U = V\mu \quad .$$

Then

$$e_x V = U\mu \ , \ \text{and} \quad V'V = 1 \ ;$$

when U, V are a normal reciprocal pair of vectors in the spaces $\mathcal{E}_x, \mathcal{E}_y$.

Accordingly, to any non-zero characteristic value μ^2 of the biprojectors on the spaces, there corresponds a normal reciprocal pair of vectors making a proper angle between the spaces with cosine μ. It has already appeared that the square of the cosine of every proper angle between the spaces is a characteristic value of the biprojectors. Thus it is seen that the characteristic values of the biprojectors on the spaces are precisely the squares of the cosines of the proper angles between them.

20. Rank and multiplicity.

Let $\mathcal{E}_{x,\alpha}$ $(0 \leqslant \alpha \leqslant \pi/2)$ denote the null-space of the matrix $\lambda 1 - e_x e_y$, where $\lambda = \cos^2\alpha$. Then evidently

$$\mathcal{E}_{x,\alpha} \subset e_x \mathcal{E}_y \quad ;$$

and $\mathcal{E}_{x,\alpha} \neq \mathcal{O}$ just when α is a proper angle between the spaces, which is just when λ is a biprojector characteristic value, in which case $\mathcal{E}_{x,\alpha}$ is composed of the latent vectors of $e_x e_y$ for the characteristic value λ, which span reflexive directions associated with a proper angle α.

The space $\mathcal{E}_{x,\alpha}$ and the corresponding space $\mathcal{E}_{y,\alpha}$

have the same dimension r_α, which will be taken to define the rank of α as a proper angle between the given spaces.

Since latent vectors of a matrix corresponding to different characteristic values are independent, the spaces $\mathcal{E}_{x,\alpha}$ are independent, and therefore the dimension of their union is the sum of their dimensions. Thus

$$\bigoplus_\alpha \mathcal{E}_{x,\alpha} \subset e_x \mathcal{E}_y ,$$

and correspondingly

$$\sum_\alpha r_\alpha \leq r ,$$

where r is the dimension of inclination of the spaces, being the dimension of $e_x \mathcal{E}_y$, which is the rank of $e_x e_y$. Here the relation between spaces must be an identity if the relation between dimensions is, as will appear in fact to be the case, an equality.

Let the multiplicity m_α of α as a proper angle be defined as the multiplicity of $\lambda = \cos^2\alpha$ as a biprojector characteristic value. But there is the general proposition that, with e_x, e_y symmetric and idempotent, the nullity of the matrix $\lambda 1 - e_x e_y$ is equal to the multiplicity of λ as a root of the equation $|\lambda 1 - e_x e_y| = 0$. It follows that

$$r_\alpha = m_\alpha ,$$

and also that the rank r of $e_x e_y$ is equal to the multiplicity of its non-zero characteristic values,

$$r = \sum_\alpha m_\alpha ;$$

whence

$$\sum_{\alpha} r_{\alpha} = r \quad .$$

Accordingly,

$$e_x \xi_y = \bigoplus_{\alpha} \xi_{x,\alpha} \quad ,$$

which shows the reflexive directions, associated with the different proper angles, spanning the orthogonal projections of the given spaces in each other.

21. Orthogonality relations.

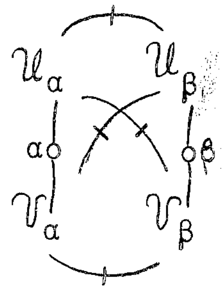
Let $\mathcal{U}_{\theta}, \mathcal{V}_{\theta}$ $(\theta = \alpha, \beta)$ be two pairs of directions, making angles α, β. Then a direction in one pair can be orthogonal to a direction in the other pair in four possible ways, defining four possible, in general independent, orthogonality relations between the pairs. However, should the pairs be formed of reciprocal directions in a pair of spaces, there is then in this case the proposition that the four orthogonality relations are equivalent, any one of them implying the other three. Thus, in this case, there are the equivalences

$$\mathcal{U}_{\alpha} \perp \mathcal{U}_{\beta} \iff \mathcal{U}_{\alpha} \perp \mathcal{V}_{\beta} \quad ,$$

and the others deduced from symmetry, by interchanges in the symbols $\mathcal{U}, \mathcal{V}$ and α, β. If the four orthogonality relations between the members of the pairs are taken together to define orthogonality as between the pairs, then there is the further proposition that reciprocal pairs making distinct proper angles are orthogonal:

$$\alpha \neq \beta \Longrightarrow (\mathcal{U}_{\alpha}, \mathcal{V}_{\alpha}) \perp (\mathcal{U}_{\beta}, \mathcal{V}_{\beta}) \quad .$$

Thus, in the following scheme, any one of the orthogonalities indicated implies all four; and all four are implied by the distinctness of α, β.



22. Isogonality.

Consider a pair of spaces which are such that any direction in one space is a reflexive direction of that space with regard to the other. An equivalent condition is that the spaces be of the same dimension, and have the proper angles between them all the same. Such a pair of spaces will be called isogonal. A necessary and sufficient condition that spaces $\mathcal{E}_x, \mathcal{E}_y$ of dimension p be isogonal, with proper angle α, is that

$$(\cos^2 \alpha 1 - e_x e_y)^p = 0 .$$

In an isogonal pair of spaces, any orthogonal pair of directions in one space projects orthogonally onto the other space into an orthogonal pair of directions; then, pairing the directions with their orthogonal projections, two reciprocal pairs of directions between the spaces are obtained which are mutually orthogonal.

Accordingly, by taking an orthogonal set of directions spanning one space, and then taking the orthogonal projections of these directions in the other, there is obtained an orthogonal set of directions spanning the other. These two sets of directions are arranged in a set of reciprocal pairs. Each direction, in either one of the two sets, is oblique to its reciprocal, in the other set, making with it the unique proper angle between the spaces; but it is orthogonal to every other direction, both in its own and in the other set.

The construction of such reciprocal orthogonal bases of directions in the isogonal spaces constitutes the form for their total decomposition, which is now to be extended to a general pair of spaces.

23. Total reduction.

A pair of spaces is to be considered totally reduced in relation to each other when together they are resolved into canonical directions with the property that any canonical direction in one space is orthogonal to every other in that space, and also to every one in the other space with the possible exception of at most one. The angles made by the oblique canonical directions define the canonical angles between the spaces, in the decomposition.

It appears immediately that oblique canonical directions, in such a reduction of the spaces, must be the orthogonal projections of each other in the spaces, and thus form reciprocal pairs of directions. The canonical angles are then identified with the proper angles, uniquely defined; so in every total decomposition, the canonical angles obtained are the same, and given by the proper angles, the squares of the cosines of which are determined algebraically as the biprojector characteristic values.

A total reduction of any oblique pair of spaces may now be obtained by the following algorithm. First take the principal reduction of the spaces, into their orthogonal projections in each other, and the orthogonal residuals:

$$\mathcal{E}_x = e_x \mathcal{E}_y \oplus \mathcal{R}_{xy} , \quad \mathcal{E}_y = e_y \mathcal{E}_x \oplus \mathcal{R}_{yx} .$$

Then there is the further scheme of decomposition:

$$e_x \mathcal{E}_y = \bigoplus_{\alpha} \mathcal{E}_{x,\alpha} , \quad e_y \mathcal{E}_x = \bigoplus_{\alpha} \mathcal{E}_{y,\alpha} ,$$

where the components

$$\mathcal{E}_{x,\alpha} , \quad \mathcal{E}_{y,\alpha}$$

are isogonal; in particular $\mathcal{E}_{x,o}$ $\mathcal{E}_{y,o}$ are identical, being the intersection

$$\mathcal{E}_{x,o} = \mathcal{E}_x \cap \mathcal{E}_y = \mathcal{E}_{y,o} ;$$

but all other components are orthogonal. Now take the total decomposition of the isogonal pairs, into mutually orthogonal reciprocal pairs of directions all making the same angle. Finally, take any orthogonal bases of directions in the residual spaces $\mathcal{R}_{xy}$, $\mathcal{R}_{yx}$. All the directions thus constructed have together the property that they constitute a total reduction of the spaces.

Matrices

$$U = M_x r , \quad V = M_y s$$

of order $N \times p$, $N \times q$ where r, s are any regular square matrices of order p, q are any bases for the spaces $\mathcal{E}_x = [M_x]$, $\mathcal{E}_y = [M_y]$. They are orthonormal bases, composed of orthogonal sets of unit vectors, if

$$U'U = 1 \,, \quad V'V = 1 \,,$$

in which case the orthogonal projectors on the spaces have the simpler form

$$e_x = UU' \,, \quad e_y = VV' \,.$$

If, moreover,

$$U'V = \begin{pmatrix} \mu_1 & & & & & & 0 \\ & \cdot & & & & & \\ & & \cdot & & & & \\ & & & \mu_r & & & \\ & & & & 0 & & \\ & & & & & \cdot & \\ 0 & & & & & & 0 \end{pmatrix} = \begin{pmatrix} \mu & 0 \\ 0 & 0 \end{pmatrix}$$

where $1 \geq \mu_1 \geq \cdots \geq \mu_r > 0$, they will be called a <u>canonical pair of bases</u> for the spaces. In this case (U_i, V_i) $(i = 1, \ldots, r)$ will be normal reciprocal pairs of vectors on the spaces, with

$$e_x \mathcal{E}_y = [U_1, \ldots, U_r], \quad e_y \mathcal{E}_x = [V_1, \ldots, V_r]$$

and

$$\mathcal{R}_{xy} = [U_{r+1}, \ldots, U_p], \quad \mathcal{R}_{yx} = [V_{r+1}, \ldots, V_q] \,.$$

The problem for constituting a total reduction of a pair of spaces is, to give it a pure matrix formulation, that of transforming the bases, with which the spaces happen to be given, into an equivalent canonical pair; alternatively, to simultaneously rotating orthonormal bases into a canonical pair. The analysis which has been given shows the possibility of this, together with an algebraic algorithm for the realization.

24. Analysis of configuration.

Two pairs of subspaces of a Euclidean space may be considered equivalent, or the spaces in the pairs to present the same configuration in their relation to each other, if they can be rigidly rotated into coincidence; that is to say one pair is the image of the other pair under an orthogonal transformation. A linear transformation which preserves orthogonality also preserves angles. It follows that two equivalent pairs of spaces have total reductions with the same canonical angles. For, an orthogonal transformation which sends one pair of spaces into the other sends canonical directions in one pair into canonical directions in the other, with angles unchanged. It has already been settled that different total reductions of the same pair of spaces always obtain the same canonical angles, given by the proper angles; and now the question arises as to whether different spaces with the same canonical angles, thus uniquely defined for each, are equivalent; so that the angles will constitute a complete analysis of the configuration formed by the spaces. It is only necessary to present a canonical form, into which pairs of spaces with given canonical angles can be rotated. It will then be established that two pairs of spaces are equivalent if and only if the proper angles between them are the same; so the proper angles between spaces gives a complete characterization of their configuration, invariant under orthogonal transformations.

25. Normal form.

An orthogonal transformation in a Euclidean space is determined by a correspondence between one complete orthogonal set of unit vectors and another. Let the coordinate vectors of the space be taken as one such set; and let the other be defined as follows, in relation to a given pair of subspaces. Take the vectors in a canonical pair of bases; let the duplicates among these, belonging to the intersection, be removed; and augment the remainder by an orthonormal base for the orthogonal complement of the union of the spaces; and let every reciprocal pair U, V making an angle α, determined by $\cos \alpha = U'V$, be replaced by an orthogonal pair

$$U \cos \alpha/2 + V \sin \alpha/2 , \quad U \cos \alpha/2 - V \sin \alpha/2 .$$

The set of vectors thus obtained constitute the second orthonormal base. The orthogonal transformation thus defined makes the given pair of spaces the image of a pair of spaces with canonical bases in the following normal form: reciprocal pairs of base vectors not in the intersection have the form

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \cos \alpha/2 \\ \sin \alpha/2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \cos \alpha/2 \\ -\sin \alpha/2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

while all others have the form

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Every pair of subspaces is the orthogonal image of a pair with bases in such a form.

26. Stationary variation.

The canonical directions obtained in a total reduction of a pair of spaces have been characterized by the property that those which form oblique pairs are reciprocals, in that they are the orthogonal projections of each other in the spaces. Now there will be shown a different but equivalent characterization.

An angle α made by certain directions in a pair of spaces will be called a stationary angle between the spaces if it is stationary when the directions undergo constrained variation in the spaces. It appears that the stationary angles are precisely the proper angles between the spaces; and that directions characterized as reciprocals are characterized equivalently as directions which make stationary angles.

Thus let

$$U = M_x r \ , \quad V = M_y s$$

be any unit vectors in the spaces, making an angle with cosine μ, where r, s are now vectors of order p, q. Then

$$r'M_x'M_x r = 1 \ , \quad s'M_y'M_y s = 1$$

and

$$r'M_x'M_y s = \mu \ .$$

The condition for μ to be stationary under the constrained variation is

$$M_x'M_y s = \lambda_r M_x'M_x r \ , \quad M_y'M_x r = \lambda_s M_y'M_y s \ ,$$

where λ_r, λ_s are the Lagrangian multipliers corresponding to the constraints on r, s. It appears now that

$$\lambda_r = \mu = \lambda_s \ .$$

But, with $\lambda_r = \lambda_s$ one of the constraints becomes redundant; whence a necessary and sufficient condition for the stationarity under constraint is

$$\begin{pmatrix} -\mu M_x'M_x & M_x'M_y \\ M_y'M_x & -\mu M_y'M_y \end{pmatrix} \begin{pmatrix} r \\ s \end{pmatrix} = 0 \ , \quad r'M_x'M_x s = 1 \ .$$

Thus the stationary values of μ are those which satisfy the equation

$$\begin{vmatrix} -\mu M_x{}'M_x & M_x{}'M_y \\ M_y{}'M_x & -\mu M_y{}'M_y \end{vmatrix} = 0 \quad .$$

But, in view of the identity

$$(-\mu)^{2N}\begin{vmatrix} -\mu M_x{}'M_x & M_x{}'M_y \\ M_y{}'M_x & -\mu M_y{}'M_y \end{vmatrix} = (-\mu)^{p+q}|M_x{}'M_x|\,|M_y{}'M_y|\,|\mu^2 1 - e_x e_y|$$

this equation and the equation

$$|\mu^2 1 - e_x e_y| = 0$$

have the same positive roots, with the same multiplicities. The proper angles and the stationary angles are thus identical.

27. Canonical statistical analysis.

Consider two multiple factors x, y correlated in an experiment in which their spans $\mathcal{S}_x = [M_x]$, $\mathcal{S}_y = [M_y]$ have canonical bases $U = M_x r$, $V = M_y s$. They are equivalent to factors

$$u = xr \ , \quad v = ys$$

which have U, V for their measurement matrices:

$$M_u = M_{xr} = M_x r = U \ , \quad M_v = M_{ys} = M_y s = V \quad .$$

The elements of u, v define canonical components for x, y . They have the property that they are all of unit variance; the pairs u_i, v_i $(i = 1, \ldots, r)$ have positive correlation, given by numbers μ_i the squares of which are the characteristic values of $e_x e_y$, and which define the canonical correlation coefficients of x, y ; while all other pairs, taken from within and from between the two sets, are uncorrelated. The derivation of components with such a pattern of correlations constitutes the canonical analysis of the factors.

28. Oblique projections.

Any vector in a Euclidean space has a unique resolution into a sum of components in each of a supplementary set of subspaces, that is a set of subspaces the sum of those dimensions is the dimension of their union, and whose union is the whole space; and these components, obtained by linear transformations of the vector, define the projections of the vector relative to that supplementary set. Correspondingly, the identity is resolved into a sum of projectors, the idempotent linear transformations which obtain the projections, the range of each being one of the spaces, while its null-space is the union of the rest; so the product of any two of these projectors is null. Conversely, any set of mutually annihilating idempotents summing to the identity are the projectors on the supplementary set of subspaces formed by their ranges.

With any idempotent, its range and null-space are complementary, and it is identified with the projector on its range parallel to its null-space. It is an orthogonal projector if it is symmetric, in which case its range and null-space are orthogonal complements; and otherwise it is an oblique projector.

More generally, relative to any set of subspaces which are just independent, and not necessarily supplementary, the sum of whose dimensions is the dimension of their sum, not necessarily the whole space, there may be taken the orthogonal projection of any vector onto their union, and then, with the vector obtained, the oblique projections on the supplementary set which the spaces form relative to their union. In this way any vector is resolved into a pair of orthogonal components, one of which is resolved further into a set of oblique components in the given spaces, while the other belongs to the orthogonal complement of their union.

If the supplementary spaces are mutually orthogonal, then projection on each space relative to the set is the same as orthogonal projection on that space, that is the projection on that space parallel to its orthogonal complement.

29. Split orthogonal projectors.

The question now arises as to explicit formulae for the determination of projectors defined with respect to a pair of spaces, either on one parallel to the other, if they are complementary, or, more generally, on one parallel to the other relative to their union, in case they are just separate. A formula can be given directly in terms of the bases, or alternatively, in terms of the orthogonal projectors on the spaces.

If two spaces $\mathcal{E}_x = [M_x]$, $\mathcal{E}_y = [M_y]$ on which the orthogonal projectors are e_x, e_y are separate, for which condition the criteria are given by

$$|1 - e_x e_y| \neq 0 \text{ , and, equivalently, } |M'\bar{e}_y M_x| \neq 0 \text{ ,}$$

there is defined the projector, which may be denoted by $e_{x|y}$, on $\mathcal{E}_x$ parallel to $\mathcal{E}_y$ relative to their union $\mathcal{E}_{x,y}$, that is the projector on $\mathcal{E}_x$ parallel to the complementary space given by the union $\mathcal{E}_y \oplus \bar{\mathcal{E}}_{x,y}$ of $\mathcal{E}_y$ with the orthogonal complement $\bar{\mathcal{E}}_{x,y}$ of their union; so $e_{x|y}$ is the unique idempotent with range $\mathcal{E}_x$ and null-space $\mathcal{E}_y \oplus \bar{\mathcal{E}}_{x,y}$. Then, in terms of the orthogonal projectors on the spaces, there is the formula

$$e_{x|y} = (1 - e_x e_y)^{-1} e_x (1 - e_x e_y)$$

for the determination of $e_{x|y}$; or, more immediately in terms of bases, and in a form which directly generalizes the formula

$$e_x = M_x (M_x' M_x)^{-1} M_x'$$

for the orthogonal projector on a space in terms of a base, there is the formula

$$e_{x|y} = M_x (M_x' \bar{e}_y M_x)^{-1} M_x' \bar{e}_y \quad .$$

From here it is noted, incidentally, that

$$\bar{e}_y e_{x|y} = e_{\bar{e}_y \mathcal{E}_x} \quad .$$

The sum of the complementary oblique projectors on a pair of spaces relative to their union is the orthogonal projector on their union:

$$e_{x|y} + e_{y|x} = e_{x,y} \quad .$$

Now any vector Z in $\mathcal{E}$ has a unique resoltuion

$$Z = e_{x|y} Z + e_{y|x} Z + \bar{e}_{x,y} Z$$

into components in $\mathcal{E}_x$, $\mathcal{E}_y$ and $\bar{\mathcal{E}}_{x,y}$.

If $\mathcal{E}_x$, $\mathcal{E}_y$ are, moreover, complementary, so that

$$\mathcal{E}_{x,y} = \mathcal{E} \quad \text{and} \quad e_{x,y} = 1 \quad ,$$

for which the condition now is simply that

$$p + q = N \quad ,$$

then $e_{x|y}$, $e_{y|x}$ are complementary projectors, on and parallel to a complementary pair of spaces; and in this case

$$e_{x|y} + e_{y|x} = 1 \quad .$$

30. Multiply split projectors.

Now let $\mathcal{E}_x$, $\mathcal{E}_y$, $\mathcal{E}_z$, ... be any spaces of dimension $p, q, r, \ldots$ which are independent, and so form a supplementary set relative to their union $\mathcal{E}_{x,y,z,\ldots}$ of dimension $p+q+r+\ldots$. Then there is determined the projective resolution

$$Z = e_{x|y,z,\ldots}Z + e_{y|x,z,\ldots}Z + \ldots + \bar{e}_{x,y,z,\ldots}Z \quad ,$$

of any vector Z in $\mathcal{E}$ into components in $\mathcal{E}_x$, $\mathcal{E}_y$, ..., and $\bar{\mathcal{E}}_{x,y,z,\ldots}$. The orthogonal projector on the union has the decomposition

$$e_{x,y,z,\ldots} = e_{x|y,z,\ldots} + e_{y|x,z,\ldots} + \ldots$$

into the sum of mutually annihilating projectors, on each parallel to the union of the other with the orthogonal complement of their union all together. Thus, $e_{x|y,z,\ldots}$ is the projector on $\mathcal{E}_x$ parallel to the complementary space given by

$$\mathcal{E}_y \oplus \mathcal{E}_z \oplus \ldots \oplus \bar{\mathcal{E}}_{x,y,z,\ldots} \quad ;$$

and, moreover,

$$\text{rank } e_{x|y,z,\ldots} = \text{trace } e_{x|y,z,\ldots} = p \text{ , } \ldots$$

If $\overline{M}_{x,y,\ldots}$ is any base obtained for $\overline{\mathcal{E}}_{x,y,\ldots}$ then $(M_x \; M_y \; \ldots \; \overline{M}_{x,y,\ldots})$ is obtained as a regular square matrix, the components of the inverse of which, when transposed and conformably partitioned, define matrices N_x, N_y, ..., $\overline{N}_{x,y,\ldots}$, thus:

$$(M_x \; M_y \; \ldots \; \overline{M}_{x,y,\ldots})^{-1} = (N_x \; N_y \; \ldots \; \overline{N}_{x,y,\ldots})' \text{ .}$$

The projectors which have been considered can be computed thus:

$$e_{x|y,z,\ldots} = M_x N_x' \text{ , } \ldots$$

31. <u>Partial regression</u>.

Consider three factors x, y, z and the regression of z on the factor composed out of x, y together. It obtains the resolution

$$M_z = e_{x,y} M_z + \overline{e}_{x,y} M_z \text{ .}$$

For the regressional part here there is the separation

$$e_{x,y} M_z = M_{x,y} r_{x,y;z}$$

$$= M_x r_{x|y,z} + M_y r_{y|x,z}$$

into further parts corresponding to the partition

$$M_{x,y} = (M_x \; M_y) \text{ ;}$$

and, conformably,

$$r_{x,y;z} = \begin{pmatrix} r_{x|y,z} \\ r_{y|x,z} \end{pmatrix}$$

Now .

$$xr_{x|y,z} + yr_{y|x,z}$$

is the regression of z on x, y together, given as a sum of complementary partial regressions, defining the regression of z on x partially with respect to y, and on y partially with respect to x.

It is required to have a formula for the partial regression matrix $r_{x|y,z}$ belonging to the regression of z on x partially with respect to y. Splitting the orthogonal projector $e_{x,y}$ into oblique components, corresponding to the resolution of its range $\mathcal{E}_{x,y}$ into complements $\mathcal{E}_x$, $\mathcal{E}_y$:

$$e_{x,y}M_z = e_{x|y}M_z + e_{y|x}M_z \quad .$$

Substituting from the formula which has been given for the split projectors, it follows immediately that

$$r_{x|y,z} = (M_x'\bar{e}_yM_x)^{-1}M_x'\bar{e}_yM_z \ ,$$

directly generalizing the corresponding formula for a total regression.

From this formula it is seen that a partial regression matrix can be expressed as a total regression matrix. The partial regression matrix, of z on x partially with respect to y, appears the same as the total regression matrix of the residuals in the regression of z on y on the residuals in the regression of x on y. For these residuals form the matrices

$$M_x^* = \bar{e}_yM_x \ , \quad M_z^* = \bar{e}_yM_z \quad ,$$

which, considered as the measurement matrices of factors $z^* = z - \hat{z}(y)$, $x^* = x - \hat{x}(y)$, give

$$r_{x^*, z^*} = r_{x|y,z} \quad .$$

32. Inversion and partition.

Consider a factor $z = (x, y)$ composed of subfactors x, y. Its measurement matrix has the partitioned form $M_z = (M_x \; M_y)$; and correspondingly,

$$M_z{}'M_z = \begin{pmatrix} M_x{}'M_x & M_x{}'M_y \\ M_y{}'M_x & M_y{}'M_y \end{pmatrix} \quad .$$

The question now is to obtain the inverse in a similar form

$$(M_z{}'M_z)^{-1} = N_z{}'N_z \quad ,$$

where

$$N_z = (N_x \; N_y)$$

is partitioned conformably with M_z, and, moreover, is such that

$$N_z{}'M_z = 1 \quad .$$

It can be verified, though it is not immediately obvious, that this scheme is obtained by

$$N_x = \bar{e}_y M_x (M_x{}'\bar{e}_y M_x)^{-1} \quad , \quad N_y = \bar{e}_x M_y (M_y{}'\bar{e}_x M_y)^{-1} \quad .$$

It is noted that, for any matrix M with independent columns, there is a variety of matrices of the same order such that

$$(M'M)^{-1} = N'N \quad , \quad N'M = 1 \quad ;$$

for a special example,

$$N = M(M'M)^{-1} \quad ,$$

and there is a further example, as just indicated, associated with any partition in the columns of M .

Now

$$M_zN_z' = M_xN_x' + M_yN_y' ,$$

which obviously gives one projector as a sum of two mutually annihilating projectors. Then, in view of the formula for split projectors, there is made the identification

$$e_{x/y} = M_xM_x' , \quad e_{y/x} = M_yN_y' ,$$

and hence, though otherwise not obviously,

$$e_z = M_zN_z' .$$

33. Experimental uniformity and mean values.

In any experiment it is always possible to entertain a fictitious factor, to be denoted by I , which remains unchanged throughout: it defines the uniform factor of the experiment, taking the value 1 on any object. Its measurement matrix is thus

$$M_I = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} ,$$

the vector with all its elements equal to 1 .

Now consider the regression of any factor x on the uniform factor I , thus:

$$M_x = e_IM_x + \bar{e}_IM_x .$$

Since

$$M_I'M_I = N ,$$

it follows that

$$e_I = \frac{1}{N} M_I M_I' \quad ,$$

and therefore that

$$e_I M_x = M_I \bar{x} = M_{\hat{x}(I)}$$

where

$$\bar{x} = r_{x,I} = \frac{1}{N} M_I' M_x = \hat{x}(I)$$

defines the mean value $\bar{x}$ of x in the experiment, it being its vector of regression coefficients on the uniform factor; and it gives its expected value on any further object, subject to the fictitious uniformity. The residuals $\bar{e}_I M_x$ in this regression measure the deviation $x - \bar{x}$ of x from its mean, thus:

$$M_{x-\bar{x}} = \bar{e}_I M_x \quad .$$

34. Multiple regressions.

The regression of a factor w of a factor $(x,y,z,\ldots)$ of dimension $p+q+r+\ldots$ composed of subfactors $x,y,z,\ldots$ of dimension $p,q,r,\ldots$ has the form

$$(x,y,\ldots)r_{x,y,\ldots;w} = xr_{x|y,z,\ldots;w} + yr_{y|x,z,\ldots;w} + \ldots$$

where

$$r_{x,y,\ldots;w} = \begin{pmatrix} r_{x|y,z,\ldots;w} \\ r_{y|x,z,\ldots;w} \\ \vdots \end{pmatrix}$$

Just as the total has the determination

$$M_{(x,y,\ldots)}r_{x,y,\ldots;w} = e_{x,y,\ldots}M_w$$

by orthogonal projection, so the parts have the determinations

$$M_x r_{x|y,z,\ldots;w} = e_{x|y,z,\ldots} M_w \ , \ \ldots$$

by oblique projections, determined by the oblique projectors into which the orthogonal projector is split. But these have been obtained in the form

$$e_{x|y,z,\ldots} = M_x N_x' \ , \ \ldots$$

so that

$$M_x r_{x|y,z,\ldots;w} = M_x N_x' M_w \ , \ \ldots$$

whence the set of partial regression matrices are obtained in the form

$$r_{x|y,z,\ldots;w} = N_x' M_w \ , \ \ldots$$

35. Multiple configurations.

Just as various binary relations have been considered, applying to any pair of spaces, and therefore to any pair of factors through their spans; and coefficients have been defined which measure the extent in which any configuration of a pair of spaces attains to or departs from them, so the same might be done for the configurations of three or more spaces, by defining n-ary relations, holding between n spaces taken together, and which have criteria for attainment to or departure from them given by certain coefficients, defined as functions of the spaces. These relations and coefficients are considered functions of the configurations formed by the spaces, and remain invariant under orthogonal transformations.

The most obvious multiple relation to be considered for any number of spaces is mutual orthogonality. More special than this is the relation of mutual distinctness. More general than the negation of this relation, which is the occurrence of a coincidence, is a relation of linear dependence between the orthogonal projectors on the spaces, which may be termed <u>composability</u>. An equivalence of the condition of composability is the condition that there exist two subsets of the spaces, which have their unions identical, and within each of which the spaces are mutually orthogonal.

orthogonal	⟵ extremes ⟶	composable
⇓		⇑
distinct	opposites	coincident

A coefficient will be defined which sets the configuration formed by any three spaces in a scale between the extremes of orthogonality and composability. For three spaces, composability is that either a pair of them are identical, or a pair are orthogonal and the third is their union. A natural extension for the coefficient to any number of spaces is readily suggested, and partly established.

Let the number

$$K_{x,y,z} = \begin{vmatrix} 1 & C_{xy} & C_{xz} \\ C_{yx} & 1 & C_{yz} \\ C_{zx} & C_{zy} & 1 \end{vmatrix}$$

determined for any three spaces symmetrically in terms of the coefficients of association between their pairs, be taken to define their <u>coefficient of dissociation</u>. It can be shown to have the property of being non-negative, and equal to zero if and only if the three spaces are composable. It will now be seen to be at most one, and equal to one if and only if the three spaces are mutually orthogonal. For it is seen that

$$0 \leq K_{x,y,z} = (1 - C^2_{xy})(1 - C^2_{xz}) - (C_{yz} - C_{xy}C_{xz})^2$$

$$= S^2_{xy}S^2_{xz} - (C_{yz} - C_{xy}C_{xz})^2$$

$$\leq S^2_{xy}S^2_{xz} \leq 1 \quad ,$$

the bounds being obtained since the S and C coefficients are bounded between 0 and 1; and similarly with the spaces permuted. It follows that the K-coefficient between the three spaces equals 1 if and only if the S-coefficients between the pairs all equal 1, which is if and only if the spaces are mutually orthogonal.

Accordingly,

$$0 \leq K_{x,y,z} \leq 1 \quad ,$$

$$K_{x,y,z} = 0 \iff \text{composable}$$

$$= 1 \iff \text{orthogonal} \; .$$

In terms of orthogonal projectors, composability is one of the conditions

$$e_x = e_y \; , \quad \text{or} \quad e_xe_y = 0 \quad \text{and} \quad e_x + e_y = e_z \quad ,$$

or one of the conditions obtained by permuting x, y, z in these.

Now for any number of spaces, corresponding to $x, y, z, \ldots$, define

$$K_{x,y,z,\ldots} = \begin{vmatrix} 1 & C_{xy} & C_{xz} & \vdots \\ C_{yx} & 1 & C_{yz} & \vdots \\ C_{zx} & C_{zy} & 1 & \vdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix}$$

Then it can be shown that

$$K_{x,y,z,\ldots} \geq 0$$

$$= 0 \iff \text{composable}$$

It is obvious that

$$K_{x,y,z,\ldots} = 1 \Longleftarrow \text{orthogonal} \quad ;$$

and it is natural to conjecture that

$$K_{x,y,z,\ldots} \leq 1$$

$$= 1 \Longrightarrow \text{orthogonal} \; .$$

36. <u>Limits of association</u>.

There is now to be shown how the relation between a pair of spaces, or a pair of factors through their spans, is limited by intermediate relation to a third.

Already there has been shown the inequality

$$(C_{yz} - C_{xy}C_{xz})^2 \leq S^2_{xy}S^2_{xz} \quad ,$$

with equality if and only if the spaces are composable, the general discrepancy in the equality being given by $K_{x,y,z}$. It follows that C_{yz} is limited thus:

$$C_{xy}C_{xz} - S_{xy}S_{xz} \leq C_{yz} \leq C_{xy}C_{xz} + S_{xy}S_{xz} \quad ,$$

where equality to one or other of the limits is attained if and only if the spaces are composable.

In this way the coefficient of association between a pair is limited when their coefficients of association and dissociation with a third are given. There is to be noted an analogy with the formulae for the cosine of a sum and difference of angles.

37. Distribution characteristics.

With a factor z, through its measurement matrix M_z, there is associated a multinormal distribution with parameter matrix

$$A = (M_z{}'M_z)^{-1} \quad ;$$

and it is desirable to identify certain coefficients which have been defined as functions of the measurements with functions of the distribution parameters, so that they can appear as direct characteristics of the distribution.

These coefficients have been formed relative to a partition $z = (x,y)$ of z into subfactors x, y to analyze the statistical relation between x, y ; and the measurements have the corresponding partition

$$M_z = (M_x \; M_y) \quad ,$$

from which the coefficients are calculated. Correspondingly, the parameter matrix A, and its inverse

$$\Sigma = A^{-1} = M_z{}'M_z \quad ,$$

have partitions

$$A = \begin{pmatrix} A_{xx} & A_{xy} \\ A_{yx} & A_{yy} \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}$$

where the diagonal submatrices are square, or order equal to the dimensions p, q of x, y. Explicitly,

$$\Sigma_{xx} = M_x{}' M_x , \quad \Sigma_{xy} = M_x{}' M_y ,$$

$$A_{xx} = (M_x{}'\bar{e}_y M_x)^{-1} , \quad A_{xy} = (M_x{}'\bar{e}_y M_x)^{-1} M_x{}'\bar{e}_y \bar{e}_x M_y (M_y{}'\bar{e}_x M_y)^{-1} .$$

The formulae giving the coefficient of correlation R_{xy} and the coefficient of separation s_{xy} between x, y as functions of the distribution parameters are

$$R^2_{xy} = \text{trace } \Sigma_{yx}\Sigma^{-1}_{xx}\Sigma_{xy}\Sigma^{-1}_{yy} , \quad s^2_{xy} = \frac{1}{|A_{xx}|\,|\Sigma_{xx}|}$$

or the same formulae with x, y interchanged. Similarly, for the coefficients of residual correlation, and of inclusion,

$$Q^2_{xy} = p - R^2_{xy} , \quad c^2_{xy} = |1 - \Sigma_{xx} A_{xx}| .$$

The coefficients of association and dissociation are determined by

$$C^2_{xy} = R^4_{xy} / pq , \quad S^2_{xy} = 1 - C^2_{xy} .$$

The canonical correlation coefficients μ and their multiplicities are determined from the positive roots, together with their multiplicities, of the equation

$$|\nu^2 \Sigma_{xx} A_{xx} - 1| = 0$$

where $\nu^2 = 1 - \mu^2$; or the same equation with x, y interchanged.

Additional notes

i. Orthogonalization.

It is required to replace a set of vectors $M_1, \dots, M_p$ by an orthonormal set $U_1, \dots, U_p$ in such a way that, in the r^{th} stage, $U_r \in [M_1, \dots, M_r] = \mathcal{E}_r$ $(r = 1, \dots, p)$; in other words, to carry out the Gramm-Schmidt orthogonalization process.

Assume the $r-1^{th}$ stage complete, and the orthogonal projector e_{r-1} on $\mathcal{E}_{r-1}$ computed. The r^{th} stage is completed by taking

$$N_r = \bar{e}_{r-1} M_r \ , \quad U_r = (N_r' N_r)^{-\frac{1}{2}} N_r \ ;$$

and, moreover,

$$e_r = e_{r-1} + U_r U_r' \quad .$$

ii. Intersection.

The intersection of the subspaces $\mathcal{E}_x$, $\mathcal{E}_y$ of dimension p, q of a space $\mathcal{E}$ of dimension N has been determined as the null-space of the matrix $1 - e_x e_y$ of order N. It can also be determined from the null-space $\mathcal{A}$ of the matrix $M_x' \bar{e}_y M_x$ of smaller order p. For, if α is any vector, $M_x' \bar{e}_y M_x \alpha = 0$ is equivalent to $\bar{e}_y M_x \alpha = 0$, which is equivalent to $M_x \alpha \in \mathcal{E}_x \cap \mathcal{E}_y$. Accordingly,

$$\mathcal{A} = [\alpha] \iff \mathcal{E}_x \cap \mathcal{E}_y = [M_x \alpha] \quad .$$

An equivalent process is obtained when x, y are interchanged.

iii. Canonical pairs of bases.

(i) Since

$$M_x'M_x \; , \quad M_x'e_yM_x$$

are a pair of symmetric matrices, of which the first is positive definite, and the second non-negative definite, they can be simultaneously transformed into the unit matrix, and a non-negative diagonal matrix; thus, with some regular square matrix α ,

$$\alpha'M_x'M_x\alpha = 1 \; , \quad \alpha'M_x'e_yM_x\alpha = \begin{bmatrix} \mu^2 & 0 \\ 0 & 0 \end{bmatrix} ,$$

where μ is a real diagonal matrix with non-zero elements, of order the rank r of e_xe_y . Take

$$U = M_x\alpha = (U_0 \; U_1) \; , \quad V^* = e_yU = (V_0^* \quad V_1^*) \; ,$$

the partitions being at the r^{th} columns; so that

$$U'U = 1 \; , \quad U_0'V_0^* = V_0^{*\prime}V_0^* = \mu^2 \; ;$$

and take $V_0 = V_0^* \mu^{-1}$, so that

$$U_0'V_0 = \mu , \quad V_0'V_0 = 1 \; .$$

Now let β_1 be a base for the null-space of $M_ye_xM_y$; and let $V_1^{**} = M_y\beta_1$, this being a base for the intersection $\mathcal{E}_y \ominus \mathcal{E}_x$ of $\mathcal{E}_y$ with the orthogonal complement of $\mathcal{E}_x$, which has been seen to be orthogonal to $e_y\mathcal{E}_x$. Let V_1 be an orthonormal equivalent of V_1^{**} ; and let $V = (V_0 \; V_1)$. Then U, V are a canonical pair of bases, having the properties

$$U'U = 1 \; , \quad V'V = 1 \; , \; U'V = \begin{bmatrix} \mu & 0 \\ 0 & 0 \end{bmatrix} \; .$$

(ii) If some orthonormal base U^* of $\mathcal{E}_x$ has already been obtained, it can be rotated by an orthogonal transformation which will simultaneously transform $U^{*\prime} e_y U^*$ to diagonal form. Take $U = U^*\alpha$, $V^* = e_y U$, and proceed as in (i), where now, more simply, $e_x = UU'$.

iv. Inversion and partition.

Given a rectangular matrix M with $|M'M| \neq 0$, it has, according to the definition of Penrose (1955), a generalized inverse given by $M^{-1} = (M'M)^{-1}M'$. The matrix N defined by

$$M^{-1} = N' \quad ,$$

that is, which is the transpose of the generalized inverse, has the properties

$$(M'M)^{-1} = N'N \;, \quad N'M = 1 \quad .$$

It is now interesting to make further observation on the form of

$$(M_z{}'M_z)^{-1} = N_z{}'N_z \;, \quad N_z{}'M_z = 1 \;,$$

where

$$M_z = (M_x \; M_y) \;, \quad N_z = (N_x \; N_y) \quad .$$

Thus,

$$R_x = \bar{e}_y M_x \;, \quad R_y = \bar{e}_x M_y$$

are the residuals in the regressions of x,y on each other; and

$$R_x^{-1} = N_x{}' \;, \quad R_y^{-1} = N_y{}' \quad .$$

v. Quadratic decomposition.

The following is a reformulation of the algebraical proposition on which the theorem of Cochran (1934) on the distribution of quadratic forms depends, in which orthogonal projectors again have a role, and which shows a significance for the form in which a sum-of-squares decomposition has been derived from a regression.

If $a, b, \ldots$ are positive definite symmetric matrices of rank $p, q, \ldots$ and of order N, such that $a + b + \ldots = 1$, then

$$p + q + \ldots = N$$

if and only if there exists a decomposition of the unit matrix into mutually orthogonal projectors $e, f, \ldots$, that is

$$e + f + \ldots = 1 \ , \quad ef = 0 \ , \ \ldots \ ,$$

and an orthogonal transformation U, such that

$$a = U'eU \ , \quad b = U'fU \ , \ \ldots \ .$$

vi. <u>Separation and inclusion.</u>

While the squares of the coefficients of correlation and residual correlation of x with y have a fixed sum, prescribed by the dimension of y, thus

$$R^2_{xy} + Q^2_{xy} = q \ ,$$

which can be written

$$\text{trace } fef + \text{trace } f\bar{e}f = \text{trace } f$$

which shows a quadratic decomposition of the form just treated; and the same holds for the coefficients of association and dissociation, thus

$$C^2_{xy} + S^2_{xy} = 1 \ ,$$

there are no precise analogues for the coefficients of separation and inclusion, showing them as exactly partitioning a fixed integral quantity into sums of squares. However, they can be shown to satisfy the inequality

$$c_{xy}^2 + s_{xy}^2 \leq 1 ,$$

where the equality holds if and only if the first space is either included in or orthogonal to the second. Further, by an application of theorems of Holder and Minkowski, there is the stronger result

$$(c_{xy}^2)^{1/p} + (s_{xy}^2)^{1/p} \leq 1 ,$$

with equality again under the conditions just stated.

vii. Determinantal inequalities.

From the identity

$$\begin{vmatrix} M'M & M'N \\ N'M & N'N \end{vmatrix} = |M'M|\,|N'N|\,|1 - ef| ,$$

and the properties which have been established for the quantity $|1 - ef|$, it follows that

$$\begin{vmatrix} M'M & M'N \\ N'M & N'N \end{vmatrix} \leq |M'M|\,|N'N| \quad \text{(Fischer, 1908)} ,$$

with equality just if $M'N = 0$. It follows immediately by induction that if

$$A = \begin{bmatrix} A_{11} & A_{12} & \vdots \\ A_{21} & A_{22} & \vdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

is any symmetrically partitioned positive definite matrix. Then

$$|A| \leq |A_{11}| |A_{22}| \ldots$$

with equality just if all the non-diagonal component matrices A_{ij} are null.

In particular, with components all of order 1, Hadamard's inequality is obtained. An advantage of this approach is that the most general inequality is obtained directly, together with necessary and sufficient conditions for equality. Bellman (1960, p. 137) remarks that "Hadamard's inequality is one of the most proved results in analysis, with well over one hundred proofs in the literature."

viii. Dissociation.

An application of Hadamard's theory (1893) on the determinant of a positive definite matrix immediately establishes the conjecture that was made about the maximum of the coefficient $K_{x,y,\ldots}$ of dissociation between spaces $\mathcal{E}_x, \mathcal{E}_y, \ldots$ which is that its value is 1, and that it is attained just when the spaces are mutually orthogonal. Moreover, the more general form of this theorem, discovered by Fischer (1908), leads to more general results. Thus, for several sets of spaces $\mathcal{E}_x, \mathcal{E}_y, \ldots$ and $\mathcal{E}_a, \mathcal{E}_b, \ldots$ and so forth,

$$K_{x,y,\ldots,a,b,\ldots,\ldots} \leq K_{x,y,\ldots} K_{a,b,\ldots} \ldots$$

with equality if and only if spaces in different sets are orthogonal. It should be noted that the previously defined composability condition holds for a set if it holds for any subset. Now if there is made the definition

$$K_{x,y,\ldots;a,b,\ldots;\ldots} = K_{x,y,\ldots,a,b,\ldots,\ldots} / K_{x,y,\ldots} K_{a,b,\ldots} \ldots$$

then this coefficient also lies between 0 and 1; being zero just if all the spaces taken together are composable, and 1 just if orthogonality always holds between spaces taken from different sets. Still more general coefficients can be defined, interminably, when sets of spaces are combined into further sets, and so forth unrestrictedly, preserving an analogous scheme at every stage, showing an inexhaustible combinatorial, schematically reproductive property of the dissociation coefficient.

ix. <u>A generalization in Hilbert space</u>.

The analysis which has been given for pairs of subspaces of a finite dimensional Euclidean space generalizes with only the slightest modifications of method directly to a finite dimensional unitary space; and it can be put in a spectral form which can be interpreted in Hilbert space, and suggests that generalization. A different approach to the related question of the unitary invariants of a pair of subspaces has been made by Dixmier (1948).

Let $\mathcal{E}$, $\mathcal{F}$ be a pair of subspaces of a Hilbert space $\mathfrak{h}$, and let e, f be the orthogonal projectors on them; so e, f are the pair of Hermitian idempotent operators in $\mathfrak{h}$ which have $\mathcal{E}$, $\mathcal{F}$ for their ranges.

Now efe, fef are a pair of non-negative Hermitian operators which have the same spectrum I, excluding 0, which is a compact set on the real axis between 0 and 1. Then, following the formulation of the spectral theorem for Hermitian operators given in Halmos (1951), there exist unique spectral measures $E(\sigma)$, $F(\sigma)$ $(\sigma \in \mathscr{S})$, defined on subspaces $\mathscr{S}$ of I, such that

$$efe = \int \lambda dE(\lambda) , \quad fef = \int \lambda dF(\lambda) ,$$

where, in addition to the automatic orthogonalities

$$E(\sigma)E(\rho) = 0 , \quad F(\sigma)F(\rho) = 0 \quad (\sigma \cap \rho = 0) ,$$

carried by the general spectral theorem for Hermitian operators, there are the further orthogonalities

$$E(\sigma)F(\rho) = 0 \quad (\sigma \cap \rho = 0) ,$$

which arise from the peculiar form of construction of this pair of Hermitian operators out of a pair of Hermitian idempotents. Moreover, the spectral measures $E(I)$, $F(I)$ of the positive spectrum give the orthogonal projectors on the orthogonal projection $e\mathfrak{F}$, $f\mathcal{E}$. The matter is easily settled so far as the point spectrum is concerned; but it calls for future analysis in respect to the approximate part of the spectrum, using the fact that the operators considered can be approximated by operators with point spectra.

REFERENCES

Afriat, S. N. On the latent vectors and characteristic values of products of pairs of symmetric idempotents. Quart. J. Math. (2), 7 (1956), 76-78.

Afriat, S. N. Orthogonal and oblique projectors and the characteristics of pairs of vector spaces. Proc. Cambridge Phil. Soc. 53, 4 (1957), 800-816.

Anderson, T. W. Estimating linear restrictions on regression coefficients for multivariate normal distributions. Ann. Math. Stat. 22 (1951), 327-351.

Anderson, T. W. An Introduction to Multivariate Statistical Analysis (New York, 1958).

Bellman, Richard. Introduction to Matrix Analysis (New York, 1960).

Cochran, W. G. The distribution of quadratic forms in a normal system, with applications to the analysis of variance. Proc. Cambridge Phil. Soc. 30 (1934), 178-191.

Dixmier, Jacques. Position relative de deux variétés lineáres fermées dans us espace de Hilbert. Rev. Sci. Fasc. 7 (1948), 387-399.

Flanders, D. A. Angles between flat subspaces of a real n-dimensional Euclidean space. Studies and Essays Presented to R. Courant (New York, 1948).

Hadamard, Jacques. Resolution d'une question relative aux déterminants. Bull. Sci. Math. (2), 17 (1893), 240-246.

Halmos, P. R. Finite-dimensional Vector Spaces (Princeton, 1948).

Halmos, P. R. Introduction to Hilbert Space and the Theory of Spectral Multiplicity (New York, 1951).

Hooper, J. W. Simultaneous equations and canonical correlation theory. Econometrica 27 (1959), 245-256.

Hotelling, H. Relations between two sets of variates. Biometrica 28 (1935), 321-377.

Jordan, C. Essai sur la géométrie à n dimensions. Bull. Soc. Math. Paris 3 (1875), 103-174; 4 (1877), 92.

Julia, G. Sur les projections des systèmes orthonormaux de l'espace hilbertien. C. R. Ac. Sc. 218 (1944), 892.

Mann, H. B. The algebra of the linear hypothesis. Ann. Math. Stat. 31 (1960), 1-15.

Penrose, R. A generalized inverse for matrices. Proc. Cambridge Phil. Soc. 51 (1955), 406-413.

Rosen, J. B. The gradient projection method for nonlinear programming. Part I. Linear constraints. J. Soc. Indust. Appl. Math. 8 (1960), 181-217.

Roy, S. N. Some Aspects of Multivariate Analysis (New York, 1957).

Scheffé, Henry. The Analysis of Variance (New York, 1959).

Schoute, P. H. On the equations determining the angles of two polydimensional spaces. Proc. Sci. K. Akad. Wet. 7 (1905), 340-341.

Somerville, D. M. Y. An Introduction to the Geometry of N Dimensions (London, 1929).

Wilks, S. S. Certain generalizations in the analysis of variance. Biometrica 24 (1932), 471-494.

Wilks, S. S. Mathematical Statistics (Princeton, 1943).