AD-A082 849 PRINCETON UNIV NJ DEPT OF STATISTICS F/G 12/1
THE INTERPRETATIONS AND APPLICATIONS OF THE INDEX OF REDUNDANCY--ETC(U)
JAN 80 D E TYLER N00014-67-A-0151-0017
UNCLASSIFIED TR-161-SER-2 NL

1 OF 1
AD A
082849

END
DATE
FILMED
5-80
DTIC

ADA082849

LEVEL II

# THE INTERPRETATIONS AND APPLICATIONS OF THE INDEX OF REDUNDANCY AND THE REDUNDANCY TRANSFORMATIONS

DAVID E. TYLER

TECHNICAL REPORT #161, SERIES 2
DEPARTMENT OF STATISITCS
PRINCETON UNIVERSITY
JANUARY, 1980

Research supported in part by ONR Grants N0014-67A0151-0017 and N00014-75-C-0453, ARO Grant DAHC-04-74-G0178 and ERDA Grant E-111-2310. The author wishes to thank Lawrence S. Mayer for initially suggesting the topic.

ABSTRACT

The index of redundancy has been receiving increasing attention in disciplines which employ applied multivariate techniques, particularly psychology and education. This index purports to measure the degree to which one random vector can predict another random vector. In this paper attention is focused on the present applications and interpretations of the index of redundancy and to the relationship between the index and other multivariate techniques. Also, simultaneous transformations of the two random vectors, which differ from the standard canonical transformstions, are derived and motivated. These simultaneous transformations are shown to be naturally related to the index of redundancy.

KEY WORDS: Canonical correlation and variate analysis; Index of Redundancy; Total variance.

## 1. Introduction

Steward and Love (1968) proposed an index to measure the degree to which one random vector can predict another random vector, or equivalently, how redundant one random vector is relative to another random vector. Their index is commonly referred to as the "index of redundancy", and is becoming popular in some disciplines which employ applied multivariate techniques, particularly psychology and education. The index of redundancy is included in the applied multivariate analysis books of Cooley and Lohnes (1971, 1976), Timm (1975) and Cohen and Cohen (1975), which are popular in these fields. It is also included in articles which review multivariate techniques in these areas, such as Tatsuoka (1973) and Darlington, Weinberg and Walberg (1975). More recently, the index of redundancy has been introduced to the disciplines of geography [Briggs and Leonard (1977a)], business [Yoram (1978)], and public health [Laessig and Duckett (1979)].

Since its introduction, a number of papers have been written on the interpretation, applications and properties of the index of redundancy in the applied literature. These papers are usually written in the jargon of the respective fields, and this has led to some misconceptions and unresolved debates concerning the applications and interpretations of the index of redundancy. (See Wood (1972), Nicewander and Wood (1974, 1975), Miller (1975a), Gleason (1977), Cohen and Cohen (1977), and Cramer and Nicewander (1979).)

More statisticians are likely to eventually encounter the index of redundancy, and so this paper is intended to be partially expository. In this paper, attention is focused on the present applications of the index of redundancy, and to the relationship between the index and other multivariate techniques. This treatment will hopefully help clarify some of the issues debated in the applied literature. Also, in the appendix of this paper, a commonly cited "property" of the index of redundancy is shown to be incorrect by means of a counterexample.

In addition, simultaneous transformations of the two random vectors, which differ from the standard canonical transformations, are derived and motivated in Sections 4 and 5. These simultaneous transformations, labeled the "redundancy transformations", are shown to be naturally related to the index of redundancy. The redundancy transformations are suggested for use when analyzing the relationship between two random vectors in studies where the index of redundancy is considered a valid summary index.

## 2. Preliminaries

Let $\underset{\sim}{Y}$ be a p-dimensional random vector and let $\underset{\sim}{X}$ be a q-dimensional random vector, which without loss of generality are both assumed to have zero mean. Denote the joint variance-covariance matrix of $\underset{\sim}{Y}$ and $\underset{\sim}{X}$ by

$$\Sigma_{Y,X} = \begin{pmatrix} \Sigma_Y & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_X \end{pmatrix} . \tag{2.1}$$

Unless otherwise stated, $\Sigma_{Y,X}$ is assumed to be nonsingular. In this paper, $\underset{\sim}{Y}$ is regarded as the dependent vector and $\underset{\sim}{X}$ is regarded as the independent vector.

I. TOTAL VARIANCE. A commonly used measure of the overall dispersion of the random vector $\underset{\sim}{Y}$ is the total variance of $\underset{\sim}{Y}$, which by definition is $\text{Trace}(\Sigma_Y)$. For B, an arbitrary $(p \times k)$ matrix with rank(B) = k, the variance explained or the variance extracted by the set of linear combinations $B'\underset{\sim}{Y}$ is defined to be the difference between the total variance of $\underset{\sim}{Y}$ and the total variance of the residual vector $\underset{\sim}{Y} - \underset{\sim}{Y}_B$, where

$$(2.2) \qquad \underset{\sim}{Y}_B = \Sigma_Y B(B'\Sigma_Y B)^{-1}B'\underset{\sim}{Y}$$

is the linear regression of $\underset{\sim}{Y}$ on $B'\underset{\sim}{Y}$. The amount of the total variance of $\underset{\sim}{Y}$ which can be explained by the set of linear combinations is therefore

$$(2.3) \qquad V_e(\underset{\sim}{Y}:B'\underset{\sim}{Y}) \equiv \text{Trace}[\Sigma_Y B(B'\Sigma_Y B)^{-1}B'\Sigma_Y].$$

The variance extracted by uncorrelated linear combinations of $\underset{\sim}{Y}$ are additive. That is, if $B = [B_1 \vdots B_2]$ with $B_1'\Sigma_Y B_2 = 0$, then

$$(2.4) \qquad V_e(\underset{\sim}{Y}:B'\underset{\sim}{Y}) = V_e(\underset{\sim}{Y}:B_1'\underset{\sim}{Y}) + V_e(\underset{\sim}{Y}:B_2'\underset{\sim}{Y}).$$

In particular, if $\underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_p$ are a set of non-null vectors such that $\underset{\sim}{b}_i' \Sigma_Y \underset{\sim}{b}_j = 0$ for $i \neq j$, then

$$\text{(2.5)} \qquad \text{Trace}(\Sigma_Y) = \Sigma_{i=1}^{p} V_e(\underset{\sim}{Y}:\underset{\sim}{b}'\underset{\sim}{Y}).$$

When the concept of explained variance is used in practice, the random vector $\underset{\sim}{Y}$ is often scaled so that each component of $\underset{\sim}{Y}$ has equal variance. The more general case, though, is to be used in this paper. That is, the components of $\underset{\sim}{Y}$ are not necessarily assumed to have equal variances.

II. CANONICAL ANALYSIS. The most developed procedure for analyzing the linear relationship between two random vectors is canonical correlation and variable analysis. The largest <u>canonical correlation</u> between the random vectors $\underset{\sim}{X}$ and $\underset{\sim}{Y}$, denoted by $\rho_{(1)}$, is the maximum absolute correlation between a linear combination of $\underset{\sim}{X}$, say $\underset{\sim}{a}_{(1)}'\underset{\sim}{X}$, and a linear combination of $\underset{\sim}{Y}$, say $\underset{\sim}{b}_{(1)}'\underset{\sim}{Y}$. The variables $\underset{\sim}{a}_{(1)}'\underset{\sim}{X}$ and $\underset{\sim}{b}_{(1)}'\underset{\sim}{Y}$ are called the first <u>canonical variables</u> for the $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ vectors respectively. The second canonical correlation $\rho_{(2)}$ is the maximum absolute correlation between a linear combination of $\underset{\sim}{X}$ uncorrelated with $\underset{\sim}{a}_{(1)}'\underset{\sim}{X}$ and a linear combination of $\underset{\sim}{Y}$ uncorrelated with $\underset{\sim}{b}_{(1)}'\underset{\sim}{Y}$, say $\underset{\sim}{a}_{(2)}'\underset{\sim}{X}$ and $\underset{\sim}{b}_{(2)}'\underset{\sim}{Y}$ respectively. The canonical correlations and variables $\rho_{(i)}$, $\underset{\sim}{a}_{(i)}'\underset{\sim}{X}$, and $\underset{\sim}{b}_{(i)}'\underset{\sim}{Y}$, $i = 3, 4, \ldots, \min(p,q)$ are defined analogously.

If $q > p$, define $\underset{\sim}{a}_{(i)}$, $i = p+1, p+2, \ldots, q$, to be any vectors such that $\underset{\sim}{a}'_{(i)}\underset{\sim}{X}$ is uncorrelated with $\underset{\sim}{a}'_{(j)}\underset{\sim}{X}$ for $i \neq j$, $j = 1, 2, \ldots, q$, and $i = p+1, p+2, \ldots, q$. If $p > q$, define $\underset{\sim}{b}_{(i)}$, $i = q+1, q+2, \ldots, p$, in an analogous manner. For completeness, $\underset{\sim}{a}'_{(i)}\underset{\sim}{X}$, $i = p+1, p+2, \ldots, q$, or $\underset{\sim}{b}'_{(i)}\underset{\sim}{X}$, $i = q+1, q+2, \ldots, p$, are considered canonical variables associated with the canonical correlation $\rho_{(i)} = 0$, $p+1 \le i \le q$ or $q+1 \le i \le p$, whichever the case.

To further specify the canonical variables, it is conventional to choose $\underset{\sim}{a}_{(i)}$ and $\underset{\sim}{b}_{(j)}$ so that $\underset{\sim}{a}'_{(i)}\Sigma_X\underset{\sim}{a}_{(i)} = 1$, $i = 1, 2, \ldots, q$, and $\underset{\sim}{b}'_{(j)}\Sigma_Y\underset{\sim}{b}_{(j)} = 1$, $j = 1, 2, \ldots, p$.

Let $A_*$ be a $(q \times q)$ matrix with columns $\underset{\sim}{a}_{(1)}, \underset{\sim}{a}_{(2)}, \ldots, \underset{\sim}{a}_{(q)}$ and let $B_*$ be a $(p \times p)$ matrix with columns $\underset{\sim}{b}_{(1)}, \underset{\sim}{b}_{(2)}, \ldots, \underset{\sim}{b}_{(p)}$. When the transformations $A_*'$ and $B_*'$ are applied simultaneously to the random vectors $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ respectively, they are to be referred to as the <u>canonical</u> <u>transformations</u>. The transformed vectors $B_*'\underset{\sim}{Y}$ and $A_*'\underset{\sim}{X}$ have the much simplified joint variance-covariance matrix,

$$\begin{pmatrix} B_*'\Sigma_Y B_* & B_*'\Sigma_{YX}A_* \\ A_*'\Sigma_{XY}B_* & A_*'\Sigma_X A_* \end{pmatrix} = \begin{pmatrix} I & C \\ C' & I \end{pmatrix} , \tag{2.6}$$

where

$$C = \begin{cases} [\Lambda \,\vdots\, 0], & \text{if } p \le q \\ \begin{bmatrix} \Lambda \\ \hline 0 \end{bmatrix}, & \text{if } p > q \end{cases}$$

with $\Lambda$ being a diagonal matrix of order $\min(p,q)$ and having $\rho_{(i)}$

as the ith diagonal element.

As defined, the canonical correlations and vectors satisfy the following identities

$$
(2.7)\qquad
\begin{aligned}
\Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX}\underset{\sim}{a}_{(i)} &= \rho_{(i)}^2\Sigma_X\underset{\sim}{a}_{(i)},\\
\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}\underset{\sim}{b}_{(i)} &= \rho_{(i)}^2\Sigma_Y\underset{\sim}{b}_{(i)}, \text{ and}\\
\rho_{(i)}\underset{\sim}{a}_{(i)} &= \Sigma_X^{-1}\Sigma_{XY}\underset{\sim}{b}_{(i)}.
\end{aligned}
$$

An important property of canonical correlations and variables is that they are co-ordinate free concepts. That is, they are invariant under nonsingular linear transformations of $\underset{\sim}{Y}$ and non-singular linear transformations of $\underset{\sim}{X}$.

## 3. The Index of Redundancy

Many scalar-valued indices which are strictly functions of the canonical correlations have been proposed to measure the relationship between two random vectors. In a recent paper, Cramer and Nicewander (1979) discuss many such indices.

However, the co-ordinate free property of the canonical correlations is not always a desirable property. For example, the concept of total variance is not co-ordinate free. The ability of $\underset{\sim}{X}$ to predict a linear combination of $\underset{\sim}{Y}$ which accounts for a large proportion of the total variance of $\underset{\sim}{Y}$ may be of more interest than the ability of $\underset{\sim}{X}$ to predict a linear combination of $\underset{\sim}{Y}$ which accounts for a small proportion of the total variance of $\underset{\sim}{Y}$. This distinction cannot be considered in an index which is strictly a function of the canonical correlations.

In consideration of this argument, Stewart and Love (1968) proposed a measure which they called an "index of redundancy." This index is a weighted average of the squared canonical correlations, where the weights are the proportions of total variance of $\underset{\sim}{Y}$ which is explained by each of the canonical variates $\underset{\sim}{b}'_{(i)}\underset{\sim}{Y}$.

DEFINITION 3.1. The index of redundancy is a measure of how "redundant $\underset{\sim}{Y}$ is given $\underset{\sim}{X}$," and is defined to be

$$R^2(\underset{\sim}{Y} : \underset{\sim}{X}) \equiv \Sigma_{i=1}^{p} \rho^2_{(i)} V_e(\underset{\sim}{Y} : \underset{\sim}{b}'_{(i)}\underset{\sim}{Y})/\text{Trace}(\Sigma_Y).$$

The index of redundancy is an asymmetric index. That is, in general $R^2(\underset{\sim}{Y} : \underset{\sim}{X}) \neq R^2(\underset{\sim}{X} : \underset{\sim}{Y})$. The index $R^2(\underset{\sim}{Y} : \underset{\sim}{X})$ distinguishes between a dependent vector ($\underset{\sim}{Y}$) and an independent vector ($\underset{\sim}{X}$). If the dependent vector is univariate, then the index of redundancy is equivalent to the square of the multiple correlation coefficient.

An important representation of the index of redundancy is given in the following lemma. This representation is discussed by Stewart and Love (1968) without justification. A proof can be found in Gleason (1976).

LEMMA 3.2 $R^2(\underset{\sim}{Y} : \underset{\sim}{X}) = \text{Trace}(\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY})/\text{Trace}(\Sigma_Y)$.

This lemma states that the index of redundancy is the percent reduction from the total variance of $\underset{\sim}{Y}$ to the total variance of the residual vector $\underset{\sim}{Y} - \hat{\underset{\sim}{Y}}$, where

(3.1) $$\hat{\underset{\sim}{Y}} \equiv \Sigma_{YX}\Sigma_X^{-1}\underset{\sim}{X}$$

is the linear regression of $\underset{\sim}{Y}$ on $\underset{\sim}{X}$. In other words, the index is the proportion of the total variance of $\underset{\sim}{Y}$ which can be explained by the linear regression of $\underset{\sim}{Y}$ on $\underset{\sim}{X}$. It should be noted that Rao (1964) informally used this concept to measure what he called the "predictive efficiency" of $\underset{\sim}{X}$ for $\underset{\sim}{Y}$. Consequently, by statement (8.6) in Rao (1964), we note that the index of redundancy can be decomposed over any complete set of uncorrelated linear combinations of the independent vector. That is, if $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_q$ is any set of non-zero vectors such that $\underset{\sim}{a}_i'\Sigma_X\underset{\sim}{a}_j = 0$ for $i \neq j$, then

(3.2) $$R^2(\underset{\sim}{Y} : \underset{\sim}{X}) = \Sigma_{i=1}^{q} R^2(\underset{\sim}{Y} : \underset{\sim}{a}_i'\underset{\sim}{X}).$$

It is interesting to observe that statement (3.2) is a generalized version of the summation given in the definition of the index of redundancy, since

(3.3) $$R^2(\underset{\sim}{Y} : \underset{\sim}{a}_{(i)}'\underset{\sim}{X}) = \rho_{(i)}^2 V_e(\underset{\sim}{Y} : \underset{\sim}{b}_i'\underset{\sim}{Y})/\mathrm{Trace}(\Sigma_Y).$$

By using the representation for the index of redundancy given in Lemma 3.2, we see that the index has the following important property.

Lemm 3.3. If P is orthogonal and A nonsingular, then
$R^2(\underset{\sim}{Y} : \underset{\sim}{X}) = R^2(P\underset{\sim}{Y} : A\underset{\sim}{X})$.

The index of redundancy, however, is not invariant under arbitrary nonsingular transformations of the dependent vector.

REMARK. In the definition of the index of redundancy, it is assumed that the joint variance-covariance matrix of the vectors is nonsingular. This is a consequence of the use of canonical correlations and variables in the definition. In view of the representation of the index given by Lemma 3.2, Nicewander and Wood (1975) note that the index of redundancy can be logically extended in the following manner. If $\text{rank}(\Sigma_Y) \geq 1$ and if $\Sigma_X$ is nonsingular, then define

$$R^2(\underset{\sim}{Y} : \underset{\sim}{X}) \equiv \text{Trace}(\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY})/\text{Trace}(\Sigma_Y). \tag{3.4}$$

If $\text{rank}(\Sigma_Y) \geq 1$ and $\text{rank}(\Sigma_X) = r \leq q$, then define

$$R^2(\underset{\sim}{Y} : \underset{\sim}{X}) \equiv R^2(\underset{\sim}{Y} : B'\underset{\sim}{X}), \tag{3.5}$$

where B is a $(t \times r)$ matrix such that $B'\Sigma_X B$ is nonsingular. This definition does not depend upon the choice of B. These extensions of the index of redundancy also represent the proportion of the total variance of $\underset{\sim}{Y}$ which can be explained by the linear regression of $\underset{\sim}{Y}$ on $\underset{\sim}{X}$.

## 4. The Redundancy Transformations

In practice, the index of redundancy is usually used as a summary index in conjunction with canonical correlation and variable analysis. (For example, see Stewart (1967), Briggs and Leonard (1977a, 1977b), Oostendorp and Berlyne (1978) or Cohen, Gaughran and Cohen (1979).) This practice is suggested not only by Stewart and Love, but also in the review paper by Tatsuota (1973) and in the applied multivariate analysis books by Cooley and Lohnes (1971) and Timm (1975). In addition, the index of redundancy is included in a recent canonical analysis computer program by Thompson and Frankiewicz (1979).

It is argued, though, by Nicewander and Wood (1974, 1975) and by Cramer and Nicewander (1979) that the association of the index of redundancy with canonical correlation and variable analysis is somewhat aritficial. Canonical correlation and variable analysis does not distinguish between a depenuent and an independent vector, whereas the index of redundancy does. In addition, the index of redundancy is only invariant under orthogonal transformations of the dependent vector, whereas the canonical correlations and variables are invariant under all nonsingular transformations of the dependent vector.

In view of this argument, simultaneous transformations of the two random vectors are introducted in the next theorem which would be more appropriate to use in conjunction with the index of redundancy than the standard canonical transformations. Being more general, these transformations should prove to be useful in studies where a distinction is made between the dependent vector and the independent vector, and where only invariance under orthogonal transformations of the dependent vector is desirable.

THEOREM 4.1 (The Redundancy Transformations)

Let $\Sigma_{Y,X} = \begin{bmatrix} \Sigma_Y & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_X \end{bmatrix}$ be a positive definite matrix of order $(p+q)$. There exists an orthogonal matrix $\bar{\underline{Y}}$ and a non-singular matrix $\bar{\underline{X}}$ such that

$$\begin{bmatrix} \bar{\underline{Y}}' & 0 \\ 0 & \bar{\underline{X}}' \end{bmatrix} \Sigma_{Y,X} \begin{bmatrix} \bar{\underline{Y}} & 0 \\ 0 & \bar{\underline{X}} \end{bmatrix} = \begin{bmatrix} \bar{\underline{Y}}'\Sigma_Y\bar{\underline{Y}} & D \\ D' & I \end{bmatrix},$$

where

$$D = \begin{cases} [\Delta \,\vdots\, 0], & \text{if } p \le q \\ \begin{bmatrix} \Delta \\ \hline 0 \end{bmatrix}, & \text{if } p > q \end{cases}$$

and $\Delta$ is a diagonal matrix of order min(p,q) with diagonal entries

$$\lambda^{\frac{1}{2}}_{(1)} \ge \lambda^{\frac{1}{2}}_{(2)} \ge \ldots \ge \lambda^{\frac{1}{2}}_{(\min[p,q])} \ge 0.$$

PROOF. The matrix $\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}$ is positive semi-definite of order p. Let $\lambda_{(1)} \geq \lambda_{(2)} \geq \ldots \geq \lambda_{(p)} \geq 0$ be its eigenvalues, and choose $\underline{Y}$ such that its ith column, denoted by $\underset{\sim}{y}_i$, is an eigenvector of $\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}$ associated with the eigenvalue $\lambda_{(i)}$ and chosen such that $\underset{\sim}{y}_i'\underset{\sim}{y}_j = \delta_{ij}$, where $\delta_{ij}$ represents the Kronecker delta. By construction, $\overline{\underline{Y}}$ is an orthogonal matrix. Choose $\overline{\underline{X}}$ such that its jth column is $\underset{\sim}{x}_j = \lambda_{(j)}^{-\frac{1}{2}}\Sigma_X^{-1}\Sigma_{XY}\underset{\sim}{y}_j$ for $j = 1, 2, \ldots, r$ where $r = \text{rank}(\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY})$. It is easy to verify that $\underset{\sim}{x}_j$, $j = 1, 2, \ldots, r$ satsified the equation $\Sigma_X^{-1}\Sigma_{XY}\Sigma_{YX}\underset{\sim}{x}_j = \lambda_{(i)}\underset{\sim}{x}_j$. Thus, we have $\underset{\sim}{x}_j'\Sigma_X\underset{\sim}{x}_j = \delta_{ij}$ for $i,j = 1, 2, \ldots, r$. To complete the definition of $\overline{\underline{X}}$, if $q > r$, let its remaining columns, denoted by $\underset{\sim}{x}_{r+1}$, $\underset{\sim}{x}_{r+2}$, ..., $\underset{\sim}{x}_q$, be the solutions to the equation $\Sigma_X^{-1}\Sigma_{XY}\Sigma_{YX}\underset{\sim}{x} = 0$ such that $\underset{\sim}{x}_i'\Sigma_X\underset{\sim}{x}_j = \delta_{ij}$ for $i,j = r+1, r+2, \ldots, q$. So, by construction we have $\underset{\sim}{x}_i'\Sigma_X\underset{\sim}{x}_j = \delta_{ij}$ for $i,j = 1, 2, \ldots, q$. Finally, for $i \leq r$, $\underset{\sim}{y}_j'\Sigma_{YX}\underset{\sim}{x}_i = \lambda_{(i)}^{-\frac{1}{2}}\underset{\sim}{y}_j'\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}\underset{\sim}{y}_i = \lambda_{(i)}^{\frac{1}{2}}\delta_{ij}$, and for $i > r$, $\underset{\sim}{y}_j'\Sigma_{YX}\underset{\sim}{x}_i = 0$. Thus, the proof is complete.

The $\overline{\underline{X}}$ and $\overline{\underline{Y}}$ of Theorem 4.1 are not unique. The next theorem, however, shows that any $\overline{\underline{X}}$ and $\overline{\underline{Y}}$ which satisfies Theorem 4.1 must be of the form constructed in the proof.

THEOREM 4.2. If $\overline{\underline{Y}}$ and $\overline{\underline{X}}$ are matrices which satisfy Theorem 4.1 with $\underset{\sim}{y}_i$ and $\underset{\sim}{x}_i$ the ith columns of $\overline{\underline{Y}}$ and $\overline{\underline{X}}$ respectively, then

$$\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}\underset{\sim}{y}_i = \lambda_{(i)}\underset{\sim}{y}_i,$$

$$\Sigma_X^{-1}\Sigma_{XY}\Sigma_{YX}\underset{\sim}{x}_i = \lambda_{(i)}\underset{\sim}{x}_i, \text{ and}$$

$$\lambda_{(i)}^{\frac{1}{2}}\underset{\sim}{x}_i = \Sigma_X^{-1}\Sigma_{XY}\underset{\sim}{y}_i,$$

with $\lambda_{(1)} \geq \lambda_{(2)} \geq \ldots \geq \lambda_{(r)} > 0$ being the non-zero roots of $\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}$, and $\lambda_{(r+1)} = \lambda_{(r+2)} = \ldots = \lambda_{(\max[p,q])} = 0$.

PROOF. Let $\underline{\bar{Y}}$ and $\underline{\bar{X}}$ be matrices satisfying Theorem 4.1, then $\underline{\bar{Y}}'\Sigma_{YX}\underline{\bar{X}} = D$ and $\underline{\bar{X}}'\Sigma_X\underline{\bar{X}} = I$ implies

$$\Sigma_{YX} = \underline{\bar{Y}}D\underline{\bar{X}}^{-1} \text{ and } \Sigma_X^{-1} = \underline{\bar{X}\bar{X}}'. \tag{4.1}$$

Thus, $\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY} = \underline{\bar{Y}}D\underline{\bar{X}}^{-1}\underline{\bar{X}\bar{X}}'(\underline{\bar{X}}')^{-1}D'\underline{\bar{Y}}' = \underline{\bar{Y}}DD'\underline{\bar{Y}}' = \underline{\bar{Y}}\Lambda_1\underline{\bar{Y}}'$ where $\Lambda_1 = \text{diagonal}(\lambda_{(1)}, \lambda_{(2)}, \ldots, \lambda_{(p)})$. This implies that $\underset{\sim}{y}_i$ and $\lambda_{(i)}$, $i = 1, 2, \ldots, p$ satisfy $\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}\underset{\sim}{y}_i = \lambda_{(i)}\underset{\sim}{y}_i$. Likewise, $\Sigma_X^{-1}\Sigma_{XY}\Sigma_{YX} = \underline{\bar{X}\bar{X}}'(\underline{\bar{X}}')^{-1}D'\underline{\bar{Y}}'\underline{\bar{Y}}D\underline{\bar{X}}^{-1} = \underline{\bar{X}}\Lambda_2\underline{\bar{X}}^{-1}$, where $\Lambda_2 = \text{diagonal}(\lambda_{(1)}, \lambda_{(2)}, \ldots, \lambda_{(q)})$, implies that $\underset{\sim}{x}_i$ and $\lambda_{(i)}$, $i = 1, 2, \ldots, p$ satisfy $\Sigma_X^{-1}\Sigma_{XY}\Sigma_{YX}\underset{\sim}{x}_i = \lambda_{(i)}\underset{\sim}{x}_i$. Finally, $\Sigma_X^{-1}\Sigma_{XY}\underset{\sim}{y}_i = \underline{\bar{X}\bar{X}}'(\underline{\bar{X}}')^{-1}D'\underline{\bar{Y}}\underset{\sim}{y}_i = \lambda_{(i)}^{1/2}\underset{\sim}{x}_i$. Thus, the proof is complete.

The transformations $\underline{\bar{Y}}'$ and $\underline{\bar{X}}'$ when simultaneously applied to $\underset{\sim}{Y}$ and $\underset{\sim}{X}$ respectively, are to be referred to as the <u>redundancy transformations</u>. When $\underset{\sim}{Y}$ and $\underset{\sim}{X}$ are thus transformed, Theorem 4.1 gives the resulting joint variance-covariance matrix. In addition, $\underset{\sim}{y}_i'\underset{\sim}{Y}$ and $\underset{\sim}{x}_i'\underset{\sim}{X}$ are to be referred to as the <u>ith</u> <u>redundancy</u> <u>variables</u>, $\underset{\sim}{y}_i$ and $\underset{\sim}{x}_i$ as the <u>ith</u> <u>redundancy</u> <u>vectors</u> and $\lambda_{(i)}$ as the <u>ith</u> <u>redundancy</u> <u>root</u>. It easily follows that the redundancy roots and variables are invariant under orthogonal transformations of $\underset{\sim}{Y}$ and non-singular transformations of $\underset{\sim}{X}$. The redundancy vectors are equivalent under these transformations.

As an exploratory technique, the redundancy transformations do not simplify the joint variance-covariance matrix to the extent in which the canonical transformations do. This is to be expected since less information on the joint variance-covariance matrix is sacrificed when only considering orthogonal transformations of $\underset{\sim}{Y}$. In particular, the index of redundancy is preserved. That is, by applying Lemma 3.3 with $P = \underline{\overline{Y}}'$ and $A = \underline{\overline{X}}'$, we obtain

$$(4.2) \qquad R^2(\underset{\sim}{Y} : \underset{\sim}{X}) = R^2(\underline{\overline{Y}}'\underset{\sim}{Y} : \underline{\overline{X}}'\underset{\sim}{X}) = \Sigma_{i=1}^{\min(p,q)} \lambda_{(i)}/\mathrm{Trace}(\Sigma_Y).$$

## 5. The Optimality of the Redundancy Transformations

When the index of redundancy is used in conjunction with canonical correlation and variable analysis, the value of (3.3) is usually used to help determine which canonical variables deserve intepretation and further attention rather than simply using the canonical correlations themself. This approach, for example, is applied in the previously mentioned studies of Briggs Leonard (1977a, 1977b), Oostendorp and Berlyne (1978), Cohen, Goughran and Cohen (1979), and Laessig and Duckett (1979). Likewise, this approach is recommended by Stewart and Love (1968), and also in the review paper by Tatsuoka (1973) and in the books by Cooley and Lohnes (1971) and Timm (1975).

This practice of using the value of (3.3) for each of the canonical variables to reduce, in essence, the dimensionality of the two sets of multivariate responses is not an optimal procedure.

The canonical variables are extracted because they best explain the intercorrelations between the sets of responses. They are not necessarily the best linear combinations to consider when attempting to account for the overall size of the index of redundancy. It is shown in this section that the redundancy variables are best suited for this purpose. Before doing so, it is first necessary to extent the concept of the contribution made by a canonical variable to the overall size of the index of redundancy, which is given by (3.3), to the contribution made by any set of linear combinations of $\underset{\sim}{X}$ or by any set of linear combinations of $\underset{\sim}{Y}$ to the overall size of the index.

A natural extension for an arbitrary set of linear combinations of the dependent vector is the proportion of the total variance of the dependent vector which can be explained by its linear regression on these linear combinations only. That is, the value of $R^2(\underset{\sim}{Y} : A'\underset{\sim}{X})$ can be considered as the contribution made by the set of linear combinations $A'\underset{\sim}{X}$ to the overall size of the index of redundancy. Thus defined, the contributions to the index made by uncorrelated linear combinations of the independent vector are additive. If $A = [A_1 : A_2]$ with $A_1'\Sigma_X A_2 = 0$, then

$$(5.1) \qquad R^2(\underset{\sim}{Y} : A'\underset{\sim}{X}) = R^2(\underset{\sim}{Y} : A_1'\underset{\sim}{X}) + R^2(\underset{\sim}{Y} : A_2'\underset{\sim}{X}).$$

In particular, if $A_o$ is a $(q \times k)$ matrix with $\text{rank}(A_o) = k$ and whose columns are a subset of the canonical vectors, say $\{\underset{\sim}{a}_{(i)}, i \in I\}$, then we have the desired result

$$(5.2) \qquad R^2(\underset{\sim}{Y} : A_o'\underset{\sim}{X}) = \Sigma_{i\in I}\rho^2_{(i)}V_e(\underset{\sim}{Y} : \underset{\sim}{b}'_{(i)}\underset{\sim}{Y})/\text{Trace}(\Sigma_Y).$$

For an arbitrary set of linear combinations of the dependent vector, a suitable extension is not obvious. One extension proposed by Miller (1969) and Miller and Farr (1971) for any linear combination $\underset{\sim}{b}'\underset{\sim}{Y}$ is the product $R^2(\underset{\sim}{b}'\underset{\sim}{Y} : \underset{\sim}{X})R^2(\underset{\sim}{Y} : \underset{\sim}{b}'\underset{\sim}{Y})$. Their work is discussed in more detail in the appendix of this paper. In particular, it is shown in the appendix that it is possible for this product to be greater than the index of redundancy itself. Thus, an alternative generalization is needed.

To motivate an alternative generalization, note that

$$(5.3) \qquad R^2(\underset{\sim}{Y} : B_o'\hat{\underset{\sim}{Y}}) = \Sigma_{i\in I}\rho^2_{(i)}V_e(\underset{\sim}{Y} : \underset{\sim}{b}'_{(i)}\underset{\sim}{Y})/\text{Trace}(\Sigma_Y),$$

where $B_o$ is a $(p \times k)$ matrix with $\text{rank}(B_o) = k$ and whose columns are the canonical vectors $\{\underset{\sim}{b}_{(i)}, i \in I\}$. So, in general, it is proposed that $R^2(\underset{\sim}{Y} : B'\hat{\underset{\sim}{Y}})$ be considered as the contribution made by the set of linear combinations $B'\underset{\sim}{Y}$ to the overall size of the index of redundancy. This quantity represents the proportion of the total variance of $\underset{\sim}{Y}$ which can be accounted for by the linear regression of $B'\underset{\sim}{Y}$ on $\underset{\sim}{X}$. As defined, the contributions to the index of redundancy made by uncorrelated linear combinations of the dependent vector are not necessarily additive. The contributions made by the linear combinations of $\underset{\sim}{Y}$ whose linear regression on $\underset{\sim}{X}$ are uncorrelated, though, are additive. That is, if $B = [B_1 : B_2]$ with $B_1' \Sigma_{YX} \Sigma_X^{-1} \Sigma_{XY} B_2 = 0$, then

$$R^2(\underset{\sim}{Y} : B'\hat{\underset{\sim}{Y}}) = R^2(\underset{\sim}{Y} : B_1'\hat{\underset{\sim}{Y}}) + R^2(\underset{\sim}{Y} : B_2'\hat{\underset{\sim}{Y}}). \tag{5.4}$$

In particular, if $\underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_p$ are any set of vectors such that $\underset{\sim}{b}_i' \Sigma_{YX} \Sigma_X^{-1} \Sigma_{XY} \underset{\sim}{b}_j = 0$ for $i \neq j$, then

$$R^2(\underset{\sim}{Y} : \underset{\sim}{X}) = \Sigma_{i=1}^{p} R^2(\underset{\sim}{Y} : \underset{\sim}{b}'\hat{\underset{\sim}{Y}}). \tag{5.5}$$

In view of these extensions of (3.3), the main optimality property of the redundancy transformations is given in the next theorem. This theorem states that of all sets of k pairs of linear combinations of $\underset{\sim}{X}$ and $\underset{\sim}{Y}$, the redundancy variables associated with the k largest redundancy roots best account for the overall size of the index of redundancy.

THEOREM 5.1. Let $\underset{\sim}{x}_i$ and $\underset{\sim}{y}_i$ be defined as in Theorem 4.2, let $\underline{\bar{X}}_k$ be a $(q \times k)$ matrix with columns $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \cdots \underset{\sim}{x}_k$, and let $\underline{\bar{Y}}_k$ be a $(p \times k)$ matrix with columns $\underset{\sim}{y}_1, \underset{\sim}{y}_2, \ldots, \underset{\sim}{y}_k$.

(i) For any $(q \times s)$ matrix A with $\mathrm{rank}(A) \le k$,
$R^2(\underset{\sim}{Y} : A'\underset{\sim}{X}) \le R^2(\underset{\sim}{Y} : \underline{\bar{X}}_k'\underset{\sim}{X})$.

(ii) For any $(p \times s)$ matrix B with $\mathrm{rank}(B) \le k$,
$R^2(\underset{\sim}{Y} : B'\hat{\underset{\sim}{Y}}) \le R^2(\underset{\sim}{Y} : \underline{\bar{Y}}_k'\underset{\sim}{Y})$.

Before proving Theorem 5.1, it is interesting to note that

$$(5.6) \qquad R^2(\underset{\sim}{Y} : \underline{\bar{X}}_k' \underset{\sim}{X}) = R^2(\underset{\sim}{Y} : \underline{\bar{Y}}_k' \underset{\sim}{Y}) = \Sigma_{i=1}^{k} \lambda_{(i)} / \text{Trace}(\Sigma_Y).$$

PROOF OF THEOREM 5.1. Part (i) follows from the results of Rao (1964) section 8. In that paper, it is shown that the quantity $\text{Trace}[\Sigma_Y - \Sigma_{YX}A'(A\Sigma_X A')^{-1}A'\Sigma_{XY}]$ is minimized over all A of order $(k \times q)$ with $\text{rank}(A) = k$ by choosing $A' = \underline{\bar{X}}_k$. For all such A, the inequality in part (i) holds since $r(\underset{\sim}{Y} : AX) = \text{Trace}[\Sigma_{YX}A'(A\Sigma_X A')^{-1}A'\Sigma_{XY}]/\text{Trace}(\Sigma_Y)$. The inequality easily extends to any A of order $(s \times q)$ with $\text{rank}(A) \le k$, (see the remark at the end of section 3.)

To prove part (ii), we note that for all B of rank less than or equal to k, $r(\underset{\sim}{Y} : B'\hat{\underset{\sim}{Y}})$ is maximized by choosing B such that $B'\Sigma_{YX}\Sigma_X^{-1} = M\underline{\bar{X}}_k'$, where M has full rank. This follows from part (i). If $k \le r$, where $r = \text{rank}(\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY})$, then by using the representation for $\Sigma_{XY}$ and $\Sigma_X^{-1}$ given in (4.1), we have $\underline{\bar{Y}}_k\Sigma_{YX}\Sigma_X^{-1} = \underline{\bar{Y}}_k'\underline{\bar{Y}}\underline{D}\underline{\bar{X}}^{-1}\underline{\bar{X}\bar{X}}' = D_k\underline{\bar{X}}_k'$ where $D_k = \text{diagonal}(\lambda_{(1)}^{1/2} \vdots \lambda_{(2)}^{1/2} \vdots \ldots \lambda_{(k)}^{1/2})$. If $k > r$, part (ii) is immediate, since $R^2(\underset{\sim}{Y} : \underline{\bar{Y}}_k'\underset{\sim}{Y}) = R^2(\underset{\sim}{Y} : \underset{\sim}{X})$.

After reducing a multivariate response to a smaller set of linear combinations of the response, it is customary in practice to consider linear transformations of the reduced set of linear combinations. These linear transformations are usually made to facilitate the interpretation of the reduced set. So, it is important to note that the optimality property for $\underline{X}_k'\underset{\sim}{X}$ and $\underline{Y}_k'\underset{\sim}{Y}$ given in Theorem 5.1 still holds if either is transformed by a nonsingular linear transformation. However, in view of the discussion in section 4, only orthogonal transformations of $\underline{Y}_k'\underset{\sim}{Y}$ would be appropriate.

## 6. Concluding Remarks

The redundancy transformation for the independent vector X was first introduced by Rao (1964). He referred to this transformation as the principal components transformation for the instrumental variable $\underset{\sim}{X}$ with respect to the variable $\underset{\sim}{Y}$. This transformation also arises in reduced rank regression problems, (see Brillinger (1975) Theorem 10.21, or Izenman (1976).) The redundancy transformation for the dependent vector $\underset{\sim}{Y}$ is the principal components transformation for $\hat{\underset{\sim}{Y}}$.

In this paper, these two transformations are viewed as being naturally related to each other and to the index of redundancy. In exploring the relationship between two multivariate responses, it should prove desirable to have a transformation for one of the responses which is accompanied by a suitable transformation for the other response.

It must be acknowledged that Van den Wallenberg (1977) also relates the index of redundancy with the redundancy transformation for the vector $\underset{\sim}{X}$. He does not refer to Rao's paper and derives it independently. In Van den Wallenberg's paper, it is suggested that $\underset{\sim}{Y}$ be transformed in a manner similar to the transformation for $\underset{\sim}{X}$, that is, to use the eigenvectors of

$\Sigma_Y^{-1}\Sigma_{YX}\Sigma_{XY}$. In this approach, the transformation for $\underset{\sim}{Y}$ is not related to the transformation for $\underset{\sim}{X}$.

APPENDIX. A counterexample to a result by Miller and Farr.

In defining the index of redundancy, Stewart and Love refer to the value of (3.3) as "the proportion of variance of the $\underset{\sim}{Y}$ set explained by the correlation between $\underset{\sim}{a}'_{(i)}\underset{\sim}{X}$ and $\underset{\sim}{b}'_{(i)}\underset{\sim}{Y}$." Stewart and Love observe that this quantity is the proportion of the variance of the $\underset{\sim}{Y}$ set "extracted" by the canonical variate $\underset{\sim}{b}'_{(i)}\underset{\sim}{Y}$ times the proportion of the variance of $\underset{\sim}{b}'_{(i)}\underset{\sim}{Y}$ which is "predictable" from $\underset{\sim}{X}$.

Recognizing that linear combinations other than the canonical vectors are often of interest, Miller (1969) and Miller and Farr (1971) proposed a generalization of the above concept. To quote them using the notation established in this paper, they suggest that for any linear combination of $\underset{\sim}{Y}$, the product

$$R^2(\underset{\sim}{Y} : \underset{\sim}{b}'\underset{\sim}{Y})R^2(\underset{\sim}{b}'\underset{\sim}{Y} : \underset{\sim}{X}) \tag{A.1}$$

can be considered as "the proportion of the total variance in $\underset{\sim}{Y}$ explained by $\underset{\sim}{X}$ with respect to the component $\underset{\sim}{b}'\underset{\sim}{Y}$." This quantity is the proportion of the total variance of $\underset{\sim}{Y}$ which can be "extracted" by the variable $\underset{\sim}{b}'\underset{\sim}{Y}$ times the proportion of the variance of $\underset{\sim}{b}'\underset{\sim}{Y}$ which can be "explained" by $\underset{\sim}{X}$. Miller and Farr call this concept a "multiplication law."

In addition, they claim that if $\underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_p$ are a set of vectors such that $\underset{\sim}{b}_i' \Sigma_X \underset{\sim}{b}_j = 0$ for $i \neq j$, then "because of the orthogonality of the components $\underset{\sim}{b}_i' \underset{\sim}{Y}$," the proportions of the total variance in $\underset{\sim}{Y}$ explained by $\underset{\sim}{X}$ with respect to each of the components can be added to obtain $R^2(\underset{\sim}{Y} : \underset{\sim}{X})$. That is,

$$\text{(A.2)} \qquad R^2(\underset{\sim}{Y} : \underset{\sim}{X}) = \Sigma_{i=1}^{p} R^2(\underset{\sim}{Y} : \underset{\sim}{b}_i' \underset{\sim}{Y}) R^2(\underset{\sim}{b}_i' \underset{\sim}{Y} : \underset{\sim}{X}).$$

This summation is equivalent to the summation defining the index of redundancy if the set of vectors $\{b_i\}$ are chosen to be the canonical vectors, since $R^2(\underset{\sim}{b}_{(i)}' \underset{\sim}{Y} : \underset{\sim}{X}) = \rho_{(i)}^2$ and $R^2(\underset{\sim}{Y} : \underset{\sim}{b}_{(i)}' \underset{\sim}{Y}) = V_e(\underset{\sim}{Y} : \underset{\sim}{b}_{(i)}' \underset{\sim}{Y}) / \mathrm{Trace}(\Sigma_Y)$.

In both Miller (1969) and Miller and Farr (1971), statement (A.2) is justified by informal arguments. They accompany their arguments by an example using the principal component vectors. Statement (A.2), however, is incorrect. Also, to view $R^2(\underset{\sim}{Y} : \underset{\sim}{b}' \underset{\sim}{Y}) R^2(\underset{\sim}{b}' \underset{\sim}{Y} : \underset{\sim}{X})$ as the proportion of the total variance in $\underset{\sim}{Y}$ explained by $\underset{\sim}{X}$ with respect to the components $\underset{\sim}{b}' \underset{\sim}{Y}$ is misleading. The following counterexample to (A.2) shows that it is possible to have $R^2(\underset{\sim}{Y} : \underset{\sim}{b}' \underset{\sim}{Y}) R^2(\underset{\sim}{b}' \underset{\sim}{Y} : \underset{\sim}{X}) > R^2(\underset{\sim}{Y} : \underset{\sim}{X})$.

COUNTEREXAMPLE. Let the joint variance-covariance matrix of $\underset{\sim}{Y}_{2\times 1}$ and $\underset{\sim}{X}_{1\times 1}$ be

$$\Sigma_{Y,X} = \begin{pmatrix} 1.0 & 0.1 & 0.1 \\ 0.1 & 1.0 & 0.0 \\ 0.1 & 0.0 & 1.0 \end{pmatrix}.$$

This gives $R^2(\underset{\sim}{Y} : \underset{\sim}{X}) = .005$. Also, if $\underset{\sim}{b}' = (1.0\ 0.0)$, then $R^2(\underset{\sim}{Y} : \underset{\sim}{b}'\underset{\sim}{Y}) = .505$ and $R^2(\underset{\sim}{b}'\underset{\sim}{Y} : \underset{\sim}{X}) = .01$. Thus $R^2(\underset{\sim}{Y} : \underset{\sim}{b}'\underset{\sim}{Y})R^2(\underset{\sim}{b}'\underset{\sim}{Y} : \underset{\sim}{X}) = .00505$. Statement (A.2) is then countradicted by choosing $\underset{\sim}{b}_1' = (1.0\ 0.0)$ and $\underset{\sim}{b}_2' = (-0.1\ 1.0)$.

It is interesting to note that if the vectors $\{\underset{\sim}{b}_i\}$ in statement (A.2) are chosen to be the principal component vectors, then statement (A.2) is valid. To show this, observe that if B is an orthogonal matrix, then $R^2(\underset{\sim}{Y} : \underset{\sim}{X}) = \text{Trace}(B'\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}B)/\text{Trace}(\Sigma_Y)$, or equivalently

$$\text{(A.3)} \qquad R^2(\underset{\sim}{Y} : \underset{\sim}{X}) = \Sigma_{i=1}^{p} R^2(\underset{\sim}{b}_i'\underset{\sim}{Y} : \underset{\sim}{X})\,\text{var}(\underset{\sim}{b}_i'\underset{\sim}{Y})/\text{Trace}(\Sigma_Y),$$

where $\underset{\sim}{b}_i$ represents the ith column of B and "var" designates variance. If $\underset{\sim}{b}_i$ is a principal component vector, then it is well known that $\text{var}(\underset{\sim}{b}_i'\underset{\sim}{Y}) = V_e(\underset{\sim}{Y} : \underset{\sim}{b}_i'\underset{\sim}{Y})$, and so $\text{var}(\underset{\sim}{b}_i'\underset{\sim}{Y})/\text{Trace}(\Sigma_Y) = R^2(\underset{\sim}{Y} : \underset{\sim}{b}_i'\underset{\sim}{Y})$. Therefore, statement (A.2) and (A.3) are equivalent when the set of vectors $\{b_i\}$ is taken to be the principal component vectors. This discussion explains why the example given by Miller and Farr using the principal component variables works correctly. The summation over the principal component variables and the summation over the canonical variables however are not two special cases of the more general statement (A.2).

Miller and Farr's results are often cited in papers pertaining to the index of redundancy, such as Tatsuoka (1973), Darlington, Weinberg and Walberg (1975), Dawson (1976), Briggs and Leonard (1977a, 1977b) and Cramer and Nicewander (1979). They are also cited in the multivariate analysis book by Cooley and Lohnes (1971), and by one of the authors, Miller (1975a, 1975b). However, no application of statement (A.2) has appeared in practice for which the set of vectors $\{\underset{\sim}{b}_i\}$ are not the canonical vectors or the principal component vectors, even though Miller and Farr recommend its use in general. Apparently, this accounts for the previously undetected error in statement (A.2).

REFERENCES


Briggs, R. and Leonard, W. A. (1977a). Advances in canonical theory. Canadian Geographer 21, 133-147.

Briggs, R. and Leonard, W. A. (1977b). Mortality and ecological structure: a canonical approach. Science and Medicine 11, 757-762.

Brillinger, D. R. (1975). Time Series: Data Analysis and Theory. Holt, Rinehart and Winston, New York.

Cohen, J. and Cohen, P. (1975). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Wiley, New York.

Cohen, P. and Cohen, J. (1977). Inefficient redundancy. Multivariate Behavior Research. 12, 167-169.

Cohen, P. Caughran, E. and Cohen, J. (1979). Age patterns of childbearing: a canonical analysis. Multivariate Behavior Research 14, 75-89.

Cooley, W. W. and Lohnes, P. R. (1971). Multivariate Data Analysis. Wiley, New York.

Cooley, W. W. and Lohnes, P. R. (1976). Evaluation Research in Education. Wiley, New York.

Cramer, E. M. and Nicewander, W. A. (1979). Some symmetric, invariant measures of multivariate association. Psychometrika 44, 43-54.

Darlington, R. B., Weinberg, S. L. and Walberg, H. J. (1975). Canonical variate analysis and related techniques. Introductory Multivariate Analysis. (Amick, D. J. and Walberg, H. J., Ed.), pp. 90-112, McCutchan Pub. Co., Berkeley, Calif.

Dawson, B. K. (1977). Predicting the population value of the canonical redundancy statistic. _ASA Proceedings of the Social Statistics Section_, Part II, 897-902.

Gleason, T. C. (1976). On redundancy in canonical analysis. _Psychological Bulletin_ 83, 1004-1006.

Izenman, A. J. (1976). Reduced-rank regression for the multivariate model. _J. Multivariate Anal_. 5, 248-264.

Laessig, R. E. and Duckett, E. J. (1979). Canonical correlation analysis: potential for environmental health planning. _Americal Journal of Public Health_ 69, 353-359.

Miller, J. K. (1969). The development and application of bi-multivariate correlation. Ph.d. dissertation, State University of New York at Buffalo.

Miller J. K. (1975a). In defense of the general canonical correlation index: reply to Nicewander and Wood. _Psychological Bulletin_ 82, 207-209.

Miller, J. K. (1975b). The sampling distribution of the bi-multivariate redundancy statistic: a Monte Carlo study. _Multivariate Behavior Research_ 10, 233-244.

Miller, J. K. and Farr, S. D. (1971). Bi-multivariate redundancy: a comprehensive measure of interbattery relationship. _Multivariate Behavior Research_ 6, 313-324.

Nicewander, W. A. and Wood, D. A. (1974). Comments on "A general canonical correlation index." _Psychological Bulletin_ 81, 92-94.

Nicewander, W. A. and Wood, D. A. (1975). On the mathematical bases of the general canonical correlation index: rejoinder to Miller. Psychological Bulletin 82, 210-212.

Oostendorp, A. and Berlyne, D. E. (197 ). Dimensions in the perception of architecture. Scand. J. Psychol. 19, 73-82.

Rao, C. R. (1964). The use and interpretation of principal components analysis in applied research. Sankhya, J. Psychol. 19, 73-82.

Stewart, D. K. (1967). Support for dissent: a study of trends in the United States. Ph.D. dissertation, University of Pittsburgh.

Stewart, D. K. and Love, W. (1968). A general canonical correlation index. Psychological Bulletin 70, 160-163.

Tatsuoka, M. M. (1973). Multivariate analysis in educational research. Review of Research in Education. (F. N. Kerlinger, Ed.), pp. 273-319. Peacock, Itaska, Ill.

Thompson, B. and Frankiewicz, R. G. (1979). CANON: A computer program which produces canonical structure and index coefficients. Educational and Psychological Measurement 39, 219-221.

Timm, N. H. (1975). Multivariate Analysis with Applications in Education and Psychology. Brooks, Cole Pub. Co., Monterey, Calif.

Van den Wallenberg, A. L. (1977). Redundancy analysis: An alternative to canonical correlation analysis. Psychometrika 42, 202-220.

Wood, D. A. (1972). Toward the interpretation of canonical dimensions. Multivariate Behavioral Research 7, 477-482.

Yoram, W. (1978). Issues and advances in segmentation research. J. Marketing Research 15, 317-337.

David E. Tyler
Department of Mathematical Sciences
Old Dominion University
Norfolk, Virginia 23508

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

# REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: T.R.#161, Series 2

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): THE INTERPRETATIONS AND APPLICATIONS OF THE INDEX OF REDUNDANCY AND THE REDUNDANCY TRANSFORMATIONS

5. TYPE OF REPORT & PERIOD COVERED: Technical Report

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): David E. Tyler

8. CONTRACT OR GRANT NUMBER(s): N0014-67-0151-0017 and N00014-75-C-0453

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Department of Statistics
Princeton University
Princeton, N. J. 08540

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS:
Office of Naval Research (Code 436)
Arlington, Virginia 22217

12. REPORT DATE: January 1980

13. NUMBER OF PAGES: 27

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

SPECIAL NOTE: All requests for this report should be directed to the author at his present location, i.e., Department of Math. Sci.
Old Dominion University
Norfolk, Virginia 23508

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Canonical correlation and variate analysis; index of redundancy; total variance.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The index of redundancy has been receiving increasing attention in disciplines which employ applied multivariate techniques, particularly psychology and education. This index purports to measure the degree to which one random vector can predict another random vector. In this paper attention is focused on present applications and interpretations of the index of redundancy, and relationship between the index and other multivariate techniques. Also, simultaneous transformations of the two random vectors, which differ from the standard canonical transformations, are derived and motivated. These simultaneous transformations are shown to be naturally related to the index of redundancy

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)