Generalized Canonical Variables

Ashis Sengupta

Indian Statistical Institute, Applied Statistics Unit, Kolkata, India, ashis@isical.ac.in

Abstract. Often the random vector variable, X, being encountered, for example, in atmospheric, biological, economic and environmental and other research, is of a large dimension but admits of meaningful grouping(s) into two or more mutually exclusive and exhaustive subvectors. Dimension reduction techniques are then sought to obtain "representative" new variables for each group to be formed by taking suitable compounds of the components within that group. This results in the substantially reduced dimensional vector variable Y to represent X. This greatly facilitates the associated computational and inferential statistical analyses. Hotelling's construction of canonical variables for the case of two groups of quantitative variables has been generalized in various directions to yield Generalized Canonical Variables that encompass more than two groups, both quantitative and qualitative components in X, order constraints imposed by the user on the compounding coefficients, and so on. The availability of •packages makes the implementation of such techniques for real-life problems quite feasible and attractive. On the inferential side, a variety of new types of hypotheses is posed: determination of optimal number of groups, best grouping for the same number of groups, and so on. While the corresponding distribution theory is quite involved, some interesting results are, nevertheless, emerging also. This topic is a fertile area and is in need of further theoretical and applied research, which should be very useful to a wide variety of practi-

•Q1

tioners.

Keywords and Phrases. canonical variables, dimension reduction, isotonic regression, OVERALS subroutine in SPSS package, permutation limit distribution, standardized generalized variance.

Blind Entry.

AMS Subject Classification.

Introduction

In the statistical analysis of $k \geqslant 2$ sets of variables, one is primarily concerned with two related aspects: study of relationships between the sets and reduction of dimensionality of the variables per set. The latter is achieved by forming k new variables, one per set, at each stage and over several stages, based on the relationships between the sets. In the study of relationships between the sets, it is natural to seek appropriate similarity or distance measures. For numeric or quantitative variables, geometrically a reasonable measure of closeness is the angle between the two variables. This algebraically is equivalent to the correlation between the two variables and hence the extent of closeness translates to the extent of linear dependence in this case. Hotelling's enhancement of canonical correlation (CC) as a "generalized correlation" to measure the relationship between two sets of numeric random variables has thus had a common appeal. In addition to examining the similarity/dependence between two sets, CC analysis yields a method for dimension reduction also through the resulting canonical variables (CV). Each CV may be looked upon as an optimal one-dimensional summary statistic representative of the corresponding set composed of many variables.

The problems considered above are faced with nonnumeric or qualitative/categorical data too. Thus, substantial research—many research papers and at least five Ph.D. theses: References 8, 10, 14, 24, and 36 have been devoted to generalizations of CC analysis to more than two sets.

As in CV analysis with two sets, in extending it to the case of $k > 2$ sets, the aims are to obtain *(i)* a measure of similarity/dependence between the k sets usually based on some function of the correlation matrix and *(ii)* representative variables, one for each set, which optimize this measure. Such extensions have yielded, what we will term, generalized canonical correlations (GCCs) and generalized canonical variables (GCVs). For definitions and an earlier review of the work done on GCV analysis, the reader is referred to Reference 37 and references therein. Other types of generalizations exist, for example, see References 51, 3, 43, and 16. Here we emphasize more on the generalizations under the set up of *(i)* and *(ii)*.

Further, these generalizations of CC analysis, while expanding the scope of its applications, also expose both constructional and inferential problems specific to the several sets case.

The constructional generalizations suggested may broadly be classified into three approaches:

(a) conditional analysis with possibly two restricted linear compounds of the variables corresponding to two chosen sets while partializing out the other sets;

(b) general (unconditional) analysis incorporating all the sets simultaneously that are motivated by a *(i)* scalar-valued and *(ii)* matrix-valued GCC as a measure of dependence between the k sets; and

(c) nonlinear analysis to encompass "nonlinear," (not to be confused with nonlinear compounds) that is, nonnumeric or categorical, both nominal and ordinal, variables.

The inferential generalizations may be listed as:

(i) estimation and tests of hypotheses pertaining to GCC and GCV;

(ii) determination of the optimal grouping of the variables subject to retaining the number of sets to be the same;

(iii) determination of the optimal number of sets and the variables therein; and

(iv) independence of some and of all of the sets.

Constructions

Assume that a meaningful subgrouping of a vector variable $\mathbf{X} : p \times 1$, into k disjoint subvector variables $\mathbf{X}_i : p_i \times 1, \Sigma_{i=1}^{k} p_i = p$, is given. Consider first the case in which $\mathbf{X}$ is a quantitative or numeric variable with dispersion matrix of $\mathbf{X}$ denoted by $\Sigma_{\mathbf{X}}$. Let $\mathbf{Y} = (Y_1, \ldots, Y_k)', Y_i \equiv f_i(\mathbf{X}_i), i = 1, \ldots, k$, for some functions $f_i(\cdot)$. Usually, $f_i(\cdot)$s will be each zero mean, unit variance linear functions. In general, we will refer to $\mathbf{Y}$ as a GCV and its components as canonical variables (variates, when normalized). Denote by $\Sigma_{\mathbf{Y}}$ the correlation matrix of $\mathbf{Y}$. We discuss here mostly the first stage GCC and GCV, primarily for notational simplicity and also since the role of the higher stage ones are similar to that of the ones for CC.

In approach (a), we consider two preferred sets conditioned on or partializing out the remaining sets in a variety of ways. The canonical correlation analysis is then carried out on these two conditional sets at the first and higher stages in the usual way (see e.g., Ref. 1). The resulting CCs and CVs are the GCCs and GCVs respectively. However, for each of these conditionings, the corresponding GCVs at the various stages are obtained as the normalized eigenvectors associated with the eigenvalues, or equivalently GCCs, corresponding to some determinantal equation. This equation may be presented in a unified form as

$$|\Sigma_{od}^* - (k-1)\rho\Sigma_d^*| = 0, \Sigma_{od}^* + \Sigma_d^* = \Sigma_{\mathbf{X}}^*.$$

Σ_d^* and Σ_{od}^* are the diagonal and off-diagonal super matrices associated with the k sets and $\Sigma_{\mathbf{X}}^*$ is the modified covariance matrix of $\mathbf{X}$, modified by the specific generalization adopted. This approach has yielded computationally quite convenient - though of limited generality - part, partial, bipart, bipartial, [34, p.26; 35; 42], and g_1- and g_2-bipartial [26, 27] GCCs and their corresponding GCVs.

Certain real-life applications may demand that the compounding coefficients for •Q2 some or all of the sets be restricted; say, obey some ordering, or have •linear constraints (see e.g., Ref. 52 for such CC analysis), or have a mixture of equality and inequality constraints [10]. Das and Sen [11] demonstrate that this problem can be reduced to that of a CV analysis with nonnegative constraints. For two sets, this leads to restricted canonical correlation (RCC) and corresponding restricted canonical variable (RCV). Formally, Das and Sen [12] define

$$RCC = \max[\alpha'\Sigma_{12}\beta : \alpha'\Sigma_{11}\alpha = \beta'\Sigma_{12}\beta = 1, \alpha \in R^+_{p_1}, \beta \in R^+_{p_2}]$$

where $R^+_{p_i}$ denotes the nonnegative orthant of $R_{p_i}, i = 1, 2$. They then solve the problem by an appeal to the Kuhn-Tucker Lagrangian method. This method of construction easily generalizes to yield restricted conditional generalized canonical correlations (RCGCCs) and restricted conditional generalized canonical variables (RCGCVs) corresponding to the conditional GCCs and GCVs in (a) discussed above, for example, restricted g_1- and g_2-bipartial GCCs and their corresponding GCVs. Omladič and Omladič [33] give an alternative approach of obtaining RCC and RCV by translating the problem into that of determining generalized eigenvalues of a real symmetric matrix in the metric of a positive definite matrix. GCC analysis under linear constraints has also been considered by Yanai [53].

Consider now approach b(i). It is convenient for interpretation, comparison, and inference purposes to have a scalar measure of dependence of k sets like the CC coefficient for two sets. McKeon [29] obtained a (first stage) GCC as a generalization of a modified intraclass correlation coefficient. SenGupta [36] constructed equicorrelated GCV (at each stage) and this equicorrelation coefficient may be taken as a scalar GCC. Following the sample space formulation of Carroll, Coelho [8] also presents a scalar GCC λ, which we define rigorously in the next section because of its connection with some of the approaches discussed there.

For approach b(ii), various criteria for obtaining GCV on the basis of the correlation matrix $\Sigma_{\mathbf{Y}}$ have been proposed, for example, SUMCOR, MAXVAR, SSQCOR, MINVAR, GENVAR, MAXBET, MAXRAT, and MAXNEAR (see e.g., Refs. 18, 37, 41, 44, and 45). Though these are general methods, yet unlike the methods in (a), in general these methods do not lead to eigenvalue and eigenvector solutions and require iterative techniques. This fact leads to difficulties in developing inference procedures. However, the modification of GENVAR proposed by SenGupta [36], to be termed restricted GENVAR (RGENVAR), yields GCC and GCV in terms of eigenvalues and eigenvectors. RGENVAR uses the criterion of minimizing the generalized variance of $\Sigma_{\mathbf{Y}}$ subject to its off-diagonal elements, that is, the pairwise correlations, being all equal. RGENVAR analysis is particularly amenable to statistical inference. We discuss briefly the MAXVAR method because of its connections with several other methods and with later developments. For further details on the others, the reader is referred to Reference 37 and the references therein. The MAXVAR method was suggested by Horst [23, 24]. The compounding coefficients in $\mathbf{Y}$ are determined such that $\Sigma_{\mathbf{Y}}$ achieves the best least squares or Euclidean norm approximation to a rank one matrix. Methods proposed by McKeon [29], McDonald [30], and Carroll [5] from different viewpoints coincide with the MAXVAR method. Carroll's approach can be given several interpretations. In his original approach, Carroll obtained sample GCV, $\mathbf{Y}$, and an auxiliary sample variate U such that they are most closely related in the sense that the sum of the squared correlations of U with each Y_j is maximum. Further, U such obtained can be related to the first principal component line for $\mathbf{Y}$. By providing a random variable development of Carroll's approach, Kettenring [70] bridges it with MAXVAR approach. Proceeding with the sample space framework of Carroll, Coelho [8] presents a geometric-algebraic description of those GCV. To be more specific, now let $\mathbf{X} : n \times p$ be the data matrix partitioned as $\mathbf{X} = [\mathbf{X}_1, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_k], \mathbf{X}_i : n \times p_i, i = 1, \ldots, k$, commensurable with the given partitioning of the p variables into the k sets. Denote by $c_i : n \times 1, i = 1, \ldots, k$, the ith canonical variate. Then, Coelho defines a GCV $c : n \times 1$ as *a linear combination of all the p variables.* He shows that $c = \sum_{i=1}^{k} c_i$ and it comprises of the variables that lie closest, "on an average", to all the m groups of variables, that is, closest on an average to all the k subspaces $\mathcal{E}_i$ spanned by the columns of each X_i. By "closest on the average," geometrically it is meant that c is the vector variable that maximizes $\sum_{i=1}^{k} \cos^2 \theta(c, c_i)$, where $\theta(.,.)$ denotes the angle between the two arguments. Algebraically, this translates to: c maximizes $\lambda =$ the sum of the squares of their multiple correlation coefficients with the variables in each of the k sets. The scalar λ is then taken as the corresponding GCC.

Some other interesting generalizations are due to Dauxois and Pousse [13] and Van de Geer [44].

We now return to Carroll's approach in order to trace the developments leading to methods for (iii). Let now the auxiliary *vector variate $U : n \times 1$* have n scores. For each set i, there will be m weights denoted by $a_i : m \times 1$, yielding one weighted sum. Thus, the data for the set i is represented by $\mathbf{X}_i : n \times m$. Carroll's method seeks $\max_{(U,a_j)} \sum_{j=1}^{k} \text{cor}(U, \mathbf{X}_j a_j)^2$. van der Burg [46] expresses this GCV problem as one of minimization of a sum of squared (SSQ) difference loss function, specifically, $\min_{(U,a_j)} \sum_{j=1}^{k} SSQ(U - \mathbf{X}_j a_j)$, for unit normalized, that is, zero mean and unit variance U and $\mathbf{X}_j$ (per column). Proceeding to higher stages, van der Burg, de Leeuw, ●Q3 and Dijksterhuis [49] (henceforth VDD●), seek, as is usual,

$$\min_{(\mathbf{U},\mathbf{A}_j)} \sum_{j=1}^{k} SSQ(\mathbf{U} - \mathbf{X}_j \mathbf{A}_j), \tag{1}$$

for unit-normalized $\mathbf{U}$ and $\mathbf{X}_j$, $\mathbf{U}$ uncorrelated across columns. This yields their version of a GCV analysis performed by their general technique, OVERALS, for a three-way table corresponding to the special case of only quantitative variables that are measured in interval scales.

For (iii), we now turn to canonical variable analysis for k sets of qualitative variables, which is sometimes termed as nonlinear CC analysis, see, for example, Reference 50 and VDD. Coelho [8] introduces indicator variables to represent categorical (when nominal) variables as in a multinomial setup. He also performs GCV analysis in terms of the Burt matrix when categories (say, ordinal) can be assigned numerical values. VDD attempts to present a unified treatment for all these three different types of variables by introducing optimal scaling for the variables represented by the columns of $\mathbf{X}_j$. Let C_t denote the class of transformations that satisfy the measurement constraints for a vector variable $x_t, t = 1, \ldots, km$, obtained from $\mathbf{X}$. Then, C_t contains all linear transformations of x_t for numerical variables, all monotone ascending transformations for ordinal variables and all isomorphic transformations of x_t for nominal variables. Then, with optimal scaling, Equation 1 can be given the general representation

$$\min_{(\mathbf{U},\mathbf{A}_j,\mathbf{Q}_j)} \sum_{j=1}^{k} SSQ(\mathbf{U} - \mathbf{X}_j \mathbf{A}_j), \tag{2}$$

for unit-normalized $\mathbf{U}$ and $\mathbf{Q}_j$, $\mathbf{U}$ uncorrelated across columns with the vector $q_t, t = 1, \ldots, km$, a column of $Q_1, \ldots, Q_k$, constrained to satisfy the given measurement restrictions. OVERALS of VDD is a slightly extended version of model (2) in that it incorporates sets differing possibly both in number and type. For related models and discussions see References 4 and 30.

GCV analysis has been extended to encompass complex variables [31].

With large-dimensional random variables, the problem of having singular covariance matrices may arise. Determination of the maximum number of stages that yield nonzero GCC and other related problems then need to be addressed. SenGupta [42] gives a unified approach using generalized inverses - see also References 24, 8, and 54.

Inference

Parametric statistical inference procedures for GCV analysis have mostly been based on the assumption of multivariate normal distribution of **X**. For (a), since the problem basically reduces to that for two sets, the usual approaches adopted for CC analysis ($k = 2$) may be invoked (see e.g., Refs. 12, p. 53; and 19, p.4). Thus, one may, with obvious modifications, adopt the parametric inference procedures on CC for normal populations from References 1 and 20, and for nonnormal populations from Reference 32, and so on, to study relevant MLEs, Likelihood Ratio, and Union-Intersection tests for (a). When normality is suspected, resampling methods may be invoked. Observing that the associated statistics in (a) are smooth functions of the sample covariance matrix, it follows from the result of Beran and Srivastava [2] that the sampling distribution of such statistics may be estimated by the corresponding bootstrapped distributions - see, for example, Reference 12.

Consider now the general methods under (b). Under (asymptotic) multivariate normality, tests for a specified value of the first stage GCC and for optimal grouping as well as for optimal dimensionality of GCV, that is, problems (i)–(iii), have been derived for RGENVAR in Reference 38 using isotonic regression. For GENVAR, such problems under general alternatives may also be dealt with using the tests for Standardized Generalized Variances given by SenGupta [39]. It may be worthwhile to cast these problems in the arena of Ranking and Selection, where certain results on Generalized Variances do exist.

•Q4 For nonnormal data, de Leeuw •and van der Burg [56] discuss the permutation limit distribution of certain GCCs.

For the nonlinear methods in (c), an underlying multinomial distribution is often a reasonable choice. Constructing the covariance matrix accordingly, counterparts of the multivariate normal approaches mentioned above may be explored. Coelho [8] considers three types of nested hypotheses:

H_0: The k sets are all independent,

$H_0^{(m)}$: The mth set is independent of all the sets from $m+1$ to $k(m = 1, \ldots, k-1)$, and

$H_0^{i(m)}$: The mth and ith sets are independent ($i > m$), conditionally on all the sets $m+1$ through k (but not the ith one), ($m = 1, \ldots, k-1; i = m+1, \ldots, k$).

H_0 is of obvious relevance. It may also be a judicious choice to have some (see Ref. 25) but not all of the sets to be independent. Observe that our approach of viewing GCC to measure the closeness of the sets, naturally demands that the nullity of the GCC over all the stages should be equivalent to H_0 above, under multivariate normality. This is true for the methods in (a) and for several of the methods in (b) including GENVAR, RGENVAR, SUMCOR, MAXVAR, SSQCOR, and MINVAR. Anderson [1] discusses Wilk's statistic for testing independence of k sets, and Coelho [9] provides "near-exact" distribution of this statistic based on the generalized near-integer Gamma distribution. Coelho [8] also proposes Likelihood Ratio type tests for the above three nested hypotheses, while van der Burg and de Leeuw [48] show how to use bootstrap and jackknife methods under the multinomial model for categorical data. There is a serious need for the development of optimal statistical inference procedures for the various generalizations of CC and CV discussed above.

Softwares

For (a), simple computations are needed for obtaining the associated eigen values and eigenvectors. Timm and Carlson [42] provide a program for part and bipartial CC analysis. Carroll and Chang [6] and CANON of Chen and Kettenring [7] provide programs for several of the generalizations mentioned in (b). CORALS, the program of Young, de Leeuw, and Takane [55], and CANALS of van der Burg and de Leeuw [47] and Gifi [18] provide programs for CC analysis of nonmetric data. CORALS has also been extended by Gifi [17] to cover a variety of measurement scales for variables as well as data grouped in $k \geqslant 2$ sets. van der Burg, de Leeuw, and Dijksterhuis [49] have prepared the program OVERALS, which performs GCC analysis of two or more sets with the three measurement levels of data : numerical, ordinal, and nominal. This program is included in the SPSS package. However, programs for conducting statistical inference procedures on GCC and GCV even with linear data are yet to be developed.

Applications

Broadly formulated, GCV analysis subsumes the aspects, and is intimately related to and provides generalizations of CVs analysis, Multivariate Multiple Regression analysis, Multivariate Analysis of Variance and Covariance, Discriminant analysis, Fisher's Optimal Scoring technique, Principal Component analysis (one variable per set), Multiple Correspondence analysis (one variable per set and multiple nominal transformations–Gower, [21]), Multiple Factorial (not factor) analysis [8], and the Gifi system [17] of exploratory multivariate data analysis.

There exist several examples in the literature of applications of GCV analysis to real-life data sets—see, for example, References 22, 23, 25, 37, and 38, for the numerical psychometric data of Thurstone and Thurstone; and References 21 and 8, for the nonnumerical medical data of Taylor.

References

[1] Anderson, T.W. (1984). *An Introduction to Multivariate Analysis Techniques*, 2nd ed. John Wiley & Sons, New York.

[2] Beran, R. and Srivastava, M.S. (1985). Bootstrap tests and covariance regions for functions of a covariance matrix. *Ann. Stat.*, **13**, 95–115.

[3] Brillinger, D.R. (1969). "The Canonical Analysis of Stationary Time Series". In *Multivariate Analysis II*, P.R. Krishnaiah, ed. Academic Press, New York, pp. 331–350.

[4] Buja, A. (1990). Remarks on functional canonical variates, alternating least squares methods and ACE. *Ann. Stat.*, **18**, 1032–1069.

[5] Carroll, J.D. (1968). Generalization of canonical correlation analysis to three or more sets of variables. *Proc. 76th Convent. Am. Psych. Assoc.*, **3**, 227–228.

[6] Carroll, J.D. and Chang, J.J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition. *Psychometrika*, **35**, 283–319.

[7] Chen, X. and Kettenring, J.R. (1972). *CANON: A Computer Program Package for the Multi-Set Canonical Correlation Analysis*. Bell Laboratories Memorandum, Bell Laboratories, Murray Hill, N.J.

[8] Coelho, C.A. (1992). *Generalized Canonical Analysis*. Ph.D. Dissertation, Department of Statistics, The University of Michigan, Ann Arbor, Mich.

[9] Coelho, C.A. (2002). The generalized near-integer Gamma distribution. *J. Multivariate. Anal.*; to appear•.

•Q5

[10] Das, S. (1993). *Restricted Canonical Correlations*. Ph.D. Dissertation, Mimeo Series 2095, The Institute of Statistics, University of North Carolina, Chapel Hill, N.C.

[11] Das, S. and Sen, P.K. (1994). Restricted canonical correlations. *Lin. Alg. Appl.*, **210**, 29–47.

[12] Das, S. and Sen, P.K. (1996). Asymptotic distribution of restricted canonical correlations and relevant resampling methods. *J. Multivariate Anal.*, **56**, 1–19.

[13] Dauxois, J. and Pousse, A. (1976). Essai de synthese sur quelques problemes mathematiques de l'analyse des donnees. *Stat. Anal. Donn.*, **1**, 47–67.

[14] de Leeuw, J. (1973). *Canonical Analysis of Categorical Data*. Ph.D. Dissertation, University of Leiden, Leiden, The Netherlands.

[15] de Leeuw, J. and van der Burg, E. (1986). "The Permutation Limit Distribution of Generalized Canonical Correlations". In *Data Analysis and Informatics*. Vol. IV, E. Diday et al.•, ed. North Holland, Amsterdam, pp. 509–521.

•Q6

[16] Escofier, B. and Pagés, J.P. (1982). *Comparaison de Groupes de Variables Definis sur le meme Ensemble d'individus*. Rapport IRISA No. 166, Rennes, France.

[17] Gifi, A. (1981). *Nonlinear Multivariate Analysis*. Department of Data Theory, University of Leiden, Leiden, Netherlands.

[18] Gifi, A. (1990). *Nonlinear Multivariate Analysis*. John Wiley & Sons, New York.

[19] Gittins, R. (1985). *Canonical Analysis: A Review with Applications in Ecology*. Springer-Verlag, New York.

[20] Glynn, W.J. and Muirhead, R.J. (1978). Inference in canonical correlation analysis. *J. Multivariate Anal.*, **8**, 468–478.

[21] Gower, J.C. (1990). Fisher's optimal scores and multiple correspondence analysis. *Biometrics*, **46**, 947–961.

[22] Horst, P. (1961). Generalized canonical correlations and their applications to experimental data. *J. Clin. Psychol.*, Monograph Supplement, **14**, 331–347.

[23] Horst, P. (1965). *Factor Analysis of Data Matrices*. Holt, Rinehart and Winston, New York.

[24] Kettenring, J.R. (1969). *Canonical Analysis of Several Sets of Variables*. Ph.D. Dissertation, Department of Statistics, University of North Carolina, Chapel Hill, N.C.

[25] Kettenring, J.R. (1971). Canonical analysis of several sets of variables. *Biometrika*, **58**, 433–451.

[26] Lee, S.-Y. (1978). Generalizations of the partial, part and bipartial canonical correlation analysis. *Psychometrika*, **43**, 427–431.

[27] Lee, S.-Y. (1979). Errata to 'Generalizations of the partial, part and bipartial canonical correlation analysis'. *Psychometrika*, **44**, 131.

[28] McDonald, R.P. (1968). A unified treatment of the weighting problem. *Psychometrika*, **33**, 351–381.

[29] McKeon, J.J. (1966). Canonical analysis: some relations between canonical correlation, factor analysis, discriminant analysis and scaling theory. *Psychom. Monogr.*, **13**●.
●Q7

[30] Michailidis, G. and de Leeuw, J. (1998). The Gifi system of descriptive multivariate analysis. *Stat. Sci.*, **13**, 307–336.

[31] Miyata, M. (1970). Complex generalization of canonical correlation and its application to a sea-level study. *J. Marine Res.*, **28**, 202–214.

[32] Muirhead, R.J. and Waternaux, C.M. (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for non-normal populations. *Biometrika*, **67**, 31–43.

[33] Omladič, M. and Omladič, V. (2000). More on restricted canonical correlations. *Lin. Alg. Appl.*, **321**, 285–293.

[34] Raja Rao, B. (1969). Partial canonical correlations. *Trabajo Estadistica Investig. Oper.*, **20**, 211–219.

[35] Roy, S.N. (1957). *Some Aspects of Multivariate Analysis.* John Wiley & Sons, New York.

[36] SenGupta, A. (1979). *On the Problems of Construction and Statistical Inference Associated with a Generalization of Canonical Variables.* Ph.D. Dissertation, Department of Statistics, The Ohio State University, Ohio, Columbus.

[37] SenGupta, A. (1983). "Generalized Canonical Variables". In *Encyclopedia of Statistical Sciences*, Vol. 3, N. Johnson and S. Kotz, eds. John Wiley & Sons, New York, pp. 326–330.

[38] SenGupta, A. (1986). "On Tests Under Order Restrictions in Reduction of Dimensionality". In *Advances in Order Restricted Statistical Inference*, Vol. 37, R. Dykstra et al., ed. Lecture notes in Statistics, Springer-Verlag, New York, pp. 249–256.

[39] SenGupta, A. (1987). Tests for standardized generalized variances of multivariate normal populations of possibly different dimensions. *J. Multivariate Anal.*, **23**, 209–219.

[40] SenGupta, A. (1991). Generalized correlations in the singular case. *J. Stat. Plann. Inference*, **28**, 241–245.

[41] Ten Berge, J.M.F. (1988). Generalized approaches to the Maxbet problem and the Maxdiff problem with applications to canonical correlations. *Psychometrika*, **53**, 487–494.

[42] Timm, N.H. and Carlson, J.E. (1976). Part and bipartial canonical correlation analysis. *Psychometrika*, **41**, 159–176.

[43] Van de Geer, J.P. (1971). *Introduction to Multivariate Analysis for the Social Sciences*. Freeman, San Francisco, Calif.

[44] Van de Geer, J.P. (1984). Linear relations among k sets of variables. *Psychometrika*, **49**, 79–94.

[45] Van de Geer, J.P. (1986). *Introduction to Linear Multivariate Analysis*, Vol. 2. DSWO Press, Leiden, Netherlands.

[46] van der Burg, E. (1988). *Nonlinear Canonical Correlation and Some Related Techniques*. DSWO Press, Leiden, Netherlands.

[47] van der Burg, E. and de Leeuw, J. (1983). Non-linear canonical correlation. *Br. J. Math. Stat. Psych.*, **38**, 54–80.

[48] van der Burg, E. and de Leeuw, J. (1988). Use of the multinomial jackknife and bootstrap in generalized canonical correlation analysis. *Appl. Stochastic Model. Data Anal.*, **4**, 154–172.

[49] van der Burg, E., de Leeuw, J., and Dijksterhuis, G.B. (1994). OVERALS Nonlinear canonical correlation with k sets of variables. *Comput. Stat. Data Anal.*, **18**, 141–163.

[50] van der Burg, E. and Dijksterhuis, G.B. (1989). "Nonlinear Canonical Correlation of Multiway Data". In *Multiway Data Analysis*, R. Coppi and S. Bolasco, eds. Elsevier Science, Amsterdam.

[51] Vinograde, B. (1950). Canonical positive definite matrices under internal linear transformations. *Proc. Am. Math. Soc.*, **1**, 159–161.

[52] Yanai, H. (1992). Canonical correlation analysis with linear constraints. *Lin. Alg. Appl.*, **176**, 75–89.

[53] Yanai, H. (1998). "Generalized Canonical Correlation Analysis with Linear Constraint". In *Studies in Classification, Data Analysis and Knowledge Organization: Data Science, Classification and Related Methods*, C. Hayashi et al., eds. Springer-Verlag, Tokyo, Japan.

[54] Yanai, H. and Puntanen, S. (1993). "Partial Canonical Correlations Associated with the Inverse and Some Generalized Inverse of a Partitioned Dispersion Matrix". In *Statistical Sciences and Data Analysis*, K. Matusita, M.L. Puri, and
●Q8 T. Hayakawa, eds.●

[55] Young, F.W., de Leeuw, J., and Takane, Y. (1976). Regression with qualitative and quantitative variables: an alternating least squares method with optimal scaling features. *Psychometrika*, **41**, 505–529.

Queries in Article ess6054

Q1. As per the style of this encyclopedia, the words 'recent' and 'recently' are not allowed in the articles. We have therefore deleted 'recently developed' from 'recently developed packages'. Please confirm if this change is acceptable.

Q2. We have modified this sentence. Please clarify if this is fine.

Q3. Please spell out this abbreviation at this instance.

Q4. This reference has not been provided in the reference list. Please provide the complete details.

Q5. Please clarify if this article has since been published. If so, please provide the complete details for this reference.

Q6. Please list out all the authors' names for References 15, 38 and 53, and please provide the place of citation for this reference.

Q7. Please provide the page range for this reference.

Q8. Please provide the complete details for this reference.