DISCRIMINANT

No analysis is done for any subfile group for which the number of non-empty groups is less than 2 or the number of cases or sum of weights fails to exceed the number of non-empty groups. An analysis may be stopped if no variables are selected during variable selection or the eigenanalysis fails.

Notation

The following notation is used throughout this chapter unless otherwise stated:

g	Number of groups
p	Number of variables
q	Number of variables selected
X_{ijk}	Value of variable i for case k in group j
f_{jk}	Case weights for case k in group j
m_j	Number of cases in group j
n_j	Sum of case weights in group j
n	Total sum of weights

Basic Statistics

Mean

$$\overline{X}_{ij} = \left(\sum_{k=1}^{m_j} f_{jk} X_{ijk} \right) \Big/ n_j \qquad (\text{variable } i \text{ in group } j)$$

$$\overline{X}_{i\bullet} = \left(\sum_{j=1}^{g} \sum_{k=1}^{m_j} f_{jk} X_{ijk} \right) \Big/ n \qquad (\text{variable } i)$$

Variances

$$S_{ij}^2 = \frac{\left(\sum_{k=1}^{m_j} f_{jk} X_{ijk}^2 - n_j \overline{X}_{ij}^2 \right)}{\left(n_j - 1 \right)} \qquad \left(\text{variable } i \text{ in group } j \right)$$

$$S_{i\bullet}^2 = \frac{\left(\sum_{j=1}^{g} \sum_{k=1}^{m_j} f_{jk} X_{ijk}^2 - n \overline{X}_i^2 \right)}{(n-1)} \qquad (\text{variable } i)$$

Within-groups Sums of Squares and Cross-product Matrix (*W*)

$$w_{il} = \sum_{j=1}^{g} \sum_{k=1}^{m_j} f_{jk} X_{ijk} X_{ljk} - \sum_{j=1}^{g} \left(\sum_{k=1}^{m_j} f_{jk} X_{ijk} \right) \left(\sum_{k=1}^{m_j} f_{jk} X_{ljk} \right) \Big/ n_j \qquad i, l = 1, \ldots, p$$

Total Sums of Squares and Cross-product Matrix (*T*)

$$t_{il} = \sum_{j=1}^{g} \sum_{k=1}^{m_j} f_{jk} X_{ijk} X_{ljk} - \left(\sum_{j=1}^{g} \sum_{k=1}^{m_j} f_{jk} X_{ijk} \right) \left(\sum_{j=1}^{g} \sum_{k=1}^{m_j} f_{jk} X_{ljk} \right) \Big/ n$$

Within-groups Covariance Matrix

$$\mathbf{C} = \frac{\mathbf{W}}{(n-g)} \qquad n > g$$

Individual Group Covariance Matrices $\left(\mathbf{C}^{(j)}\right)$

$$c_{il}^{(j)} = \frac{\left(\sum_{k=1}^{m_j} f_{jk} X_{ijk} X_{ljk} - \overline{X}_{ij}\overline{X}_{lj} n_j\right)}{\left(n_j - 1\right)}$$

Within-groups Correlation Matrix (*R*)

$$r_{il} = \begin{cases} \dfrac{w_{il}}{\sqrt{w_{ii}w_{ll}}} & \text{if } w_{ii}w_{ll} > 0 \\ \text{SYSMIS} & \text{otherwise} \end{cases}$$

Total Covariance Matrix

$$\mathbf{T}' = \frac{\mathbf{T}}{n-1}$$

Univariate *F* and Λ for Variable *I*

$$F_i = \frac{(t_{ii} - w_{ii})(n-g)}{w_{ii}(g-1)}$$

with $g-1$ and $n-g$ degrees of freedom

$$\Lambda_i = \frac{w_{ii}}{t_{ii}}$$

with 1, $g-1$ and $n-g$ degrees of freedom.

Rules of Variable Selection

Both direct and stepwise variable entry are possible. Multiple inclusion levels may also be specified.

Method = Direct

For direct variable selection, variables are considered for inclusion in the order in which they are written on the ANALYSIS = list. A variable is included in the analysis if, when it is included, no variable in the analysis will have a tolerance less than the specified tolerance limit (default = 0.001).

Stepwise Variable Selection

At each step, the following rules control variable selection:

- Eligible variables with higher inclusion levels are entered before eligible variables with lower inclusion levels.
- The order of entry of eligible variables with the same even inclusion level is determined by their order on the ANALYSIS = specification.
- The order of entry of eligible variables with the same odd level of inclusion is determined by their value on the entry criterion. The variable with the “best” value for the criterion statistic is entered first.
- When level-one processing is reached, prior to inclusion of any eligible variables, all already-entered variables which have level one inclusion numbers are examined for removal. A variable is considered eligible for removal if its *F*-to-remove is less than the *F* value for variable removal, or, if probability criteria are used, the significance of its *F*-to-remove exceeds the specified probability level. If more than one variable is eligible for removal, that variable is removed that leaves the “best” value for the criterion statistic for the remaining variables. Variable removal continues until no more variables are eligible for removal. Sequential entry of variables then proceeds as described previously, except that after each step, variables with inclusion numbers of one are also considered for exclusion as described before.
- A variable with a zero inclusion level is never entered, although some statistics for it are printed.

Ineligibility for Inclusion

A variable with an odd inclusion number is considered ineligible for inclusion if:

- The tolerance of any variable in the analysis (including its own) drops below the specified tolerance limit if it is entered, or
- Its *F*-to-enter is less than the *F*-value for a variable to enter value, or
- If probability criteria are used, the significance level associated with its *F*-to-enter exceeds the probability to enter.

A variable with an even inclusion number is ineligible for inclusion if the first condition above is met.

Computations During Variable Selection

During variable selection, the matrix **W** is replaced at each step by a new matrix $\mathbf{W}^*$ using the symmetric sweep operator described by Dempster (1969). If the first q variables have been included in the analysis, **W** may be partitioned as:

$$\begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{bmatrix}$$

where $\mathbf{W}_{11}$ is $q \times q$. At this stage, the matrix $\mathbf{W}^*$ is defined by

$$\mathbf{W}^* = \begin{bmatrix} -\mathbf{W}_{11}^{-1} & \mathbf{W}_{11}^{-1}\mathbf{W}_{12} \\ \mathbf{W}_{21}\mathbf{W}_{11}^{-1} & \mathbf{W}_{22} - \mathbf{W}_{21}\mathbf{W}_{11}^{-1}\mathbf{W}_{12} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{11}^* & \mathbf{W}_{12}^* \\ \mathbf{W}_{21}^* & \mathbf{W}_{22}^* \end{bmatrix}$$

In addition, when stepwise variable selection is used, T is replaced by the matrix T^*, defined similarly.

The following statistics are computed:

Tolerance

$$\text{TOL}_i = \begin{cases} 0 & \text{if } w_{ii} = 0 \\ w_{ii}^* / w_{ii} & \text{if variable } i \text{ is not in the analysis and } w_{ii} \neq 0 \\ -1 / \left(w_{ii}^* w_{ii} \right) & \text{if variable } i \text{ is in the analysis and } w_{ii} \neq 0. \end{cases}$$

If a variable's tolerance is less than or equal to the specified tolerance limit, or its inclusion in the analysis would reduce the tolerance of another variable in the equation to or below the limit, the following statistics are not computed for it or any set including it.

***F*-to-Remove**

$$F_i = \frac{\left(w_{ii}^* - t_{ii}^* \right)(n - q - g + 1)}{t_{ii}^* (g - 1)}$$

with degrees of freedom $g-1$ and $n-q-g+1$.

***F*-to-Enter**

$$F_i = \frac{\left(t_{ii}^* - w_{ii}^* \right)(n - q - g)}{w_{ii}^* (g - 1)}$$

with degrees of freedom $g-1$ and $n-q-g$.

Wilks' Lambda for Testing the Equality of Group Means

$$\Lambda = \left| \mathbf{W}_{11} \right| / \left| \mathbf{T}_{11} \right|$$

with degrees of freedom q, $g-1$, and $n-g$.

The Approximate *F* Test for Lambda (the "overall *F*"), also known as Rao's *R* (Tatsuoka, 1971)

$$F = \frac{\left(1-\Lambda^{s}\right)\left(r/s+1-qh/2\right)}{\Lambda^{s} qh}$$

where

$$s = \begin{cases} \sqrt{\dfrac{q^2+h^2-5}{q^2h^2-4}} & \text{if } q^2+h^2 \neq 5 \\ 1 & \text{otherwise} \end{cases}$$

$$r = n-1-(q+g)/2$$

$$h = g-1$$

with degrees of freedom qh and $r/s+1-qh/2$. The approximation is exact if q or h is 1 or 2.

Rao's *V* (Lawley-Hotelling trace) (Rao, 1952; Morrison, 1976)

$$V = -(n-g)\sum_{i=1}^{q}\sum_{l=1}^{q} w_{il}^{*}\left(t_{il}-w_{il}\right)$$

When $n-g$ is large, *V*, under the null hypothesis, is approximately distributed as χ^2 with $q(g-1)$ degrees of freedom. When an additional variable is entered, the change in *V*, if positive, has approximately a χ^2 distribution with $g-1$ degrees of freedom.

The Squared Mahalanobis Distance (Morrison, 1976) between groups *a* and *b*

$$D_{ab}^{2} = -(n-g)\sum_{i=1}^{q}\sum_{l=1}^{q} w_{il}^{*}\left(\overline{X}_{ia}-\overline{X}_{ib}\right)\left(\overline{X}_{la}-\overline{X}_{lb}\right)$$

The *F* value for Testing the Equality of Means of Groups *a* and *b*

$$F_{ab} = \frac{(n-q-g+1)n_a n_b}{q(n-g)(n_a+n_b)} D_{ab}^2$$

The Sum of Unexplained Variations (Dixon, 1973)

$$R = \sum_{a=1}^{g-1} \sum_{b=a+1}^{g} 4 \Big/ \left(4 + D_{ab}^2\right)$$

Classification Functions

Once a set of q variables has been selected, the classification functions (also known as Fisher's linear discriminant functions) can be computed using

$$b_{ij} = (n-g) \sum_{l=1}^{q} w_{il}^{*} \overline{X}_{lj} \qquad i = 1, 2, \ldots, q, j = 1, 2, \ldots, g$$

for the coefficients, and

$$a_j = \log p_j - \frac{1}{2} \sum_{i=1}^{q} b_{ij} \overline{X}_{ij} \qquad j = 1, 2, \ldots, q$$

for the constant, where p_j is the prior probability of group j.

Canonical Discriminant Functions

The canonical discriminant function coefficients are determined by solving the general eigenvalue problem

$$(\mathbf{T} - \mathbf{W})\mathbf{V} = \lambda \mathbf{W}\mathbf{V}$$

where **V** is the unscaled matrix of discriminant function coefficients and `l` is a diagonal matrix of eigenvalues. The eigensystem is solved as follows:

The Cholesky decomposition

$$\mathbf{W} = \mathbf{LU}$$

is formed, where **L** is a lower triangular matrix, and $\mathbf{U} = \mathbf{L}'$.

The symmetric matrix $\mathbf{L}^{-1}\mathbf{B}\mathbf{U}^{-1}$ is formed and the system

$$\left(\mathbf{L}^{-1}(\mathbf{T}-\mathbf{W})\mathbf{U}^{-1} - \lambda\mathbf{I}\right)(\mathbf{UV}) = 0$$

is solved using tridiagonalization and the QL method. The result is *m* eigenvalues, where $m = \min(q, g-1)$ and corresponding orthonormal eigenvectors, **UV**. The eigenvectors of the original system are obtained as

$$\mathbf{V} = \mathbf{U}^{-1}(\mathbf{UV})$$

For each of the eigenvalues, which are ordered in descending magnitude, the following statistics are calculated:

Percentage of Between-Groups Variance Accounted for

$$\frac{100\lambda_k}{\sum_{k=1}^{m}\lambda_k}$$

Canonical Correlation

$$\sqrt{\lambda_k/(1+\lambda_k)}$$

Wilks' Lambda

Testing the significance of all the discriminating functions after the first *k*:

$$\Lambda_k = \prod_{i=k+1}^{m} 1/(1+\lambda_i) \qquad k = 0, 1, \ldots, m-1$$

The significance level is based on

$$\chi^2 = -\left(n - (q+g)/2 - 1\right)\ln \Lambda_k,$$

which is distributed as a χ^2 with $(q-k)(g-k-1)$ degrees of freedom.

The Standardized Canonical Discriminant Coefficient Matrix D

The standard canonical discriminant coefficient matrix **D** is computed as

$$\mathbf{D} = \mathbf{S}_{11}^{-1}\mathbf{V}$$

where

$\mathbf{S}$	$= \text{diag}\left(\sqrt{w_{11}}, \sqrt{w_{22}}, \ldots, \sqrt{w_{pp}}\right)$
$\mathbf{S}_{11}$	= partition containing the first q rows and columns of S
$\mathbf{V}$	= matrix of eigenvectors such that
$\mathbf{V}'\mathbf{W}_{11}\mathbf{V}$	$= I$

The Correlations Between the Canonical Discriminant Functions and the Discriminating Variables

The correlations between the canonical discriminant functions and the discriminating variables are given by

$$\mathbf{R} = \mathbf{S}_{11}^{-11}\mathbf{W}_{11}\mathbf{V}$$

If some variables were not selected for inclusion in the analysis $(q < p)$, the eigenvectors are implicitly extended with zeroes to include the nonselected variables in the correlation matrix. Variables for which $W_{ii} = 0$ are excluded from *S* and *W* for this calculation; *p* then represents the number of variables with non-zero within-groups variance.

The Unstandardized Coefficients

The unstandardized coefficients are calculated from the standardized ones using

$$\mathbf{B} = \sqrt{(n-g)}\mathbf{S}_{11}^{-1}\mathbf{D}$$

The associated constants are:

$$a_k = -\sum_{i=1}^{q} b_{ik}\overline{X}_{i\bullet}$$

The group centroids are the canonical discriminant functions evaluated at the group means:

$$\bar{f}_{kj} = a_k + \sum_{i=1}^{q} b_{ik}\overline{X}_{ij}$$

Tests For Equality Of Variance

Box's *M* is used to test for equality of the group covariance matrices.

$$M = (n-g)\log|\mathbf{C}'| - \sum_{j=1}^{g}(n_j - 1)\log\left|\mathbf{C}^{(j)}\right|$$

where

$\mathbf{C}'$ = pooled within-groups covariance matrix excluding groups with singular covariance matrices

$\mathbf{C}^{(j)}$ = covariance matrix for group j.

Determinants of $\mathbf{C}'$ and $\mathbf{C}^{(j)}$ are obtained from the Cholesky decomposition. If any diagonal element of the decomposition is less than 10^{-11}, the matrix is considered singular and excluded from the analysis.

$$\log\left|\mathbf{C}^{(j)}\right| = 2\sum_{i=1}^{p}\log l_{ii} - p\log\left(n_j - 1\right)$$

where l_{ii} is the ith diagonal entry of L such that $\left(n_j - 1\right)\mathbf{C}^{(j)} = \mathbf{L}'\mathbf{L}$.

Similarly,

$$\log\left|\mathbf{C}'\right| = 2\sum_{i=1}^{p}\log l_{ii} - p\log(n' - g)$$

where

$(n' - g)\mathbf{C}' = \mathbf{L}'\mathbf{L}$

n' = sum of weights of cases in all groups with nonsingular covariance matrices

The significance level is obtained from the F distribution with t_1 and t_2 degrees of freedom using (Cooley and Lohnes, 1971):

$$F = \begin{cases} M/b & \text{if } e_2 > e_1^2 \\ \dfrac{t_2 M}{t_1(b - M)} & \text{if } e_2 < e_1^2 \end{cases}$$

where

$$e_1 = \left(\sum_{j=1}^{g} \frac{1}{n_j - 1} - \frac{1}{n-g} \right) \frac{2p^2 + 3p - 1}{6(g-1)(p+1)}$$

$$e_2 = \left(\sum_{j=1}^{g} \frac{1}{(n_j - 1)^2} - \frac{1}{(n-g)^2} \right) \frac{(p-1)(p+2)}{6(g-1)}$$

$$t_1 = (g-1)p(p+1)/2$$

$$t_2 = \left(t_1 + 2\right) \Big/ \left| e_2 - e_1^2 \right|$$

$$b = \begin{cases} \dfrac{t_1}{1 - e_1 - t_1/t_2} & \text{if } e_2 > e_1^2 \\ \dfrac{t_2}{1 - e_1 - 2/t_2} & \text{if } e_2 < e_1^2 \end{cases}$$

If $e_1^2 - e_2$ is zero, or much smaller than e_2, t_2 cannot be computed or cannot be computed accurately. If

$$e_2 = e_2 + 0.0001\left(e_2 - e_1^2\right)$$

the program uses Bartlett's χ^2 statistic rather than the *F* statistic:

$$\chi^2 = M(1 - e_1)$$

with t_1 degrees of freedom.

For testing the group covariance matrix of the canonical discriminant functions, the procedure is similar. The covariance matrices $\mathbf{C}^{(j)}$ and $\mathbf{C}'$ are replaced by $\mathbf{D}_j$ and $\mathbf{D}'$, where

$$\mathbf{D}_j = \mathbf{B}'\mathbf{C}^{(j)}\mathbf{B}$$

is the group covariance matrix of the discriminant functions.

The pooled covariance matrix in this case is an identity, so that

$$\mathbf{D}' = (n-g)\mathbf{I}_m - \sum_j (n_j - 1)\mathbf{D}_j$$

where the summation is only over groups with singular $\mathbf{D}_j$.

Classification

The basic procedure for classifying a case is as follows:

- If $\mathbf{X}$ is the $1 \times q$ vector of discriminating variables for the case, the $1 \times m$ vector of canonical discriminant function values is

$$\mathbf{f} = \mathbf{XB} + \mathbf{a}$$

- A chi-square distance from each centroid is computed

$$\chi_j^2 = \left(\mathbf{f} - \bar{\mathbf{f}}_j\right)\mathbf{D}_j^{-1}\left(\mathbf{f} - \bar{\mathbf{f}}_j\right)'$$

where $\mathbf{D}_j$ is the covariance matrix of canonical discriminant functions for group *j* and $\bar{\mathbf{f}}_j$ is the group centroid vector. If the case is a member of group *j*, χ_j^2 has a χ^2 distribution with *m* degrees of freedom. $P(\mathbf{X}|\mathbf{G})$ is the significance level of such a χ_j^2.

- The classification, or posterior probability $P(\mathbf{G}_j|\mathbf{X})$, is

$$P(\mathbf{G}_j|\mathbf{X}) = \frac{P_j\left|\mathbf{D}_j\right|^{-1/2} e^{-\chi_j^2/2}}{\sum_{j=1}^{g} P_j\left|\mathbf{D}_j\right|^{-1/2} e^{-\chi_j^2/2}}$$

where p_j is the prior probability for group *j*. A case is classified into the group for which $P(\mathbf{G}_j|\mathbf{X})$ is highest.

The actual calculation of $P(\mathbf{G}_j|\mathbf{X})$ is

$$g_j = \log P_j - \frac{1}{2}\left(\log\left|\mathbf{D}_j\right| + \chi_j^2\right)$$

$$P(\mathbf{G}_j|\mathbf{X}) = \begin{cases} \dfrac{\exp\left(g_j - \max\limits_j g_j\right)}{\sum\limits_{j=1}^{g} \exp\left(g_j - \max\limits_j g_j\right)} & \text{if } g_j - \max\limits_j g_j > -46 \\ 0 & \text{otherwise} \end{cases}$$

If individual group covariances are not used in classification, the pooled within-groups covariance matrix of the discriminant functions (an identity matrix) is substituted for $\mathbf{D}_j$ in the above calculation, resulting in considerable simplification.

If any $\mathbf{D}_j$ is singular, a pseudo-inverse of the form

$$\begin{bmatrix} \mathbf{D}_{j11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

replaces $\mathbf{D}_j^{-1}$ and $\left|\mathbf{D}_{j11}\right|$ replaces $\left|\mathbf{D}_j\right|$. $\mathbf{D}_{j11}$ is a submatrix of $\mathbf{D}_j$ whose rows and columns correspond to functions not dependent on preceding functions. That is, function 1 will be excluded only if the rank of $\mathbf{D}_j = 0$, function 2 will be excluded only if it is dependent on function 1, and so on. This choice of the pseudo-inverse is not optimal for the numerical stability of $\mathbf{D}_{j11}^{-1}$, but maximizes the discrimination power of the remaining functions.

Cross-Validation

The following notation is used in this section:

$\underset{\sim}{X}_{jk}$ $(X_{1jk},\ldots,X_{qjk})^T$

$\underset{\sim}{M}_j$ Sample mean of *jth* group

$$\underset{\sim}{M}_j = \frac{1}{n_j}\sum_{k=1}^{m_j} f_{jk}\,\underset{\sim}{X}_{jk}$$

$\underset{\sim}{M}_{jk}$ Sample mean of *j*th group excluding point $\underset{\sim}{X}_{jk}$

$$\underset{\sim}{M}_{jk} = \frac{1}{n_j - f_{jk}}\sum_{\substack{l=1\\ l\neq k}}^{m_j} f_{jl}\,\underset{\sim}{X}_{jl}$$

Σ Polled sample covariance matrix

Σ_j Sample covariance matrix of *j*th group

Σ_{jk} Polled sample covariance matrix without point $\underset{\sim}{X}_{jk}$

$$\Sigma_{jk}^{-1} = \frac{n-g-f_{jk}}{n-g}\left(\Sigma^{-1} + \frac{n_j\Sigma_j^{-1}(\underset{\sim}{X}_{jk}-\underset{\sim}{M}_j)(\underset{\sim}{X}_{jk}-\underset{\sim}{M}_j)^T\Sigma_j^{-1}}{(n_j-f_{jk})(n_j-g)-n_j(\underset{\sim}{X}_{jk}-\underset{\sim}{M}_j)^T\Sigma_j^{-1}(\underset{\sim}{X}_{jk}-\underset{\sim}{M}_j)}\right)$$

$$d_0^2(\underset{\sim}{a},\underset{\sim}{b}) = (\underset{\sim}{a}-\underset{\sim}{b})^T\Sigma_{jk}^{-1}(\underset{\sim}{a}-\underset{\sim}{b})^T$$

Cross-validation applies only to linear discriminant analysis (not quadratic). During cross-validation, SPSS loops over all cases in the data set. Each case, say $\underset{\sim}{X}_{jk}$, is extracted once and treated as test data. The remaining cases are treated as a new data set.

Here we compute $d_0^2(\underset{\sim}{X}_{jk}, \underset{\sim}{M}_{jk})$ and $d_0^2(\underset{\sim}{X}_{jk}, \underset{\sim}{M}_i)$ $(i = 1,\ldots,g.i \neq j)$. If there is an i $(i \neq j)$ that satisfies $(\log(P_i) - d_0^2(\underset{\sim}{X}_{jk}, \underset{\sim}{M}_i)/2 > \log(P_j) - d_0^2(\underset{\sim}{X}_{jk}, \underset{\sim}{M}_{jk})/2)$, then the extracted point $\underset{\sim}{X}_{jk}$ is misclassified. The estimate of prediction error rate is the ratio of the sum of misclassified case weights and the sum of all case weights.

To reduce computation time, the linear discriminant method is used instead of the canonical discriminant method. The theoretical solution is exactly the same for both methods.

Rotations

Varimax rotations may be performed on either the matrix of canonical discriminant function coefficients or on that of the correlation between the canonical discriminant functions and the discrimination variables (the structure matrix). The actual algorithm for the rotation is described in FACTOR.

For the Kaiser normalization

$$h_i^2 = \begin{cases} \underset{\text{(squared multiple correlation)}}{1 + 1/w_{ii}w_{ii}^*} & \text{if coefficients rotated} \\ \sum_{k=1}^{m} r_{ik}^2 & \text{if correlations rotated} \end{cases}$$

The unrotated structure matrix is

$$\mathbf{R} = \mathbf{S}_{11}^{-1}\mathbf{W}_{11}\mathbf{V}$$

If the rotation transformation matrix is represented by **K**, the rotated standardized coefficient matrix $\mathbf{D}_R$ is given by

$$\mathbf{D}_R = \mathbf{DK}$$

The rotated matrix of pooled within-groups correlations between the canonical discriminant functions and the discriminating variables $\mathbf{R}_R$ is

$$\mathbf{R}_R = \mathbf{RK}$$

The eigenvector matrix $\mathbf{V}$ satisfies

$$\mathbf{V}'(\mathbf{T} - \mathbf{W})\mathbf{V} = \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$$

where the λ_k are the eigenvalues.

The equivalent matrix for the rotated coefficient $\mathbf{V}_R$

$$(\mathbf{V}_R)'(\mathbf{T} - \mathbf{W})\mathbf{V}_R$$

is not diagonal, meaning the rotated functions, unlike the unrotated ones, are correlated for the original sample, although their within-groups covariance matrix is an identity. The diagonals of the above matrix may still be interpreted as the between-groups variances of the functions. They are the numerators for the proportions of variance printed with the transformation matrix. The denominator is their sum. After rotation, the columns of the transformation are exchanged, if necessary, so that the diagonals of the matrix above are in descending order.

References

Anderson, T. W. 1958. *Introduction to multivariate statistical analysis*. New York: John Wiley & Sons, Inc.

Cooley, W. W., and Lohnes, P. R. 1971. *Multivariate data analysis*. New York: John Wiley & Sons, Inc.

Dempster, A. P. 1969. *Elements of Continuous Multivariate Analysis*. Reading, Mass.: Addison-Wesley.

Dixon, W. J., ed. 1973. *BMD Biomedical computer programs*. Los Angeles: University of California Press.

Tatsuoka, M. M. 1971. *Multivariate analysis*. New York: John Wiley & Sons, Inc.