REPEATABILITY OF SPATIAL CONFIGURATIONS

DEFINED BY INTERACTION MATRICES

by

ADA RAY CORRAL

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

DEPARTMENT OF STATISTICS

RALEIGH

1 9 8 0

APPROVED BY:

John Bishir

Thomas M. Gerig

Charles Proctor

Larry A. Nelson
Co-Chairman of Advisory Committee

John O. Rawlings
Co-Chairman of Advisory Committee

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

1. INTRODUCTION

Interaction effects, in general, have been studied to get a better understanding of the model representing the data. For example, Mandel (1971), in order to simplify the model, partitions the interaction into a significant and a non-significant part, according to the significance or non-significance of the eigenvalues of the matrix $D'D$ where d_{ij} are the interaction effects, and ($'$) denotes the transpose. In particular, genotype by environment interactions have been studied because of their importance in developing improved varieties and in variety testing programs. Knowledge of the magnitude and pattern of the genotype by environment interaction has become essential in helping the plant breeder reach many decisions concerning his breeding programs. For example, Eberhart and Russell (1966) reduced genotype by environment interaction by selecting stable genotypes, those that interact less with the environment, and proposed a model to provide the criteria necessary to rank cultivars for stability. Abou-El-Fittouh, et al. (1969) defined cultivar testing regions for cotton by using as a criterion the reduction in the within region genotype by location (environment) interaction. By reducing the within region interaction, the precision for comparing cultivars was increased and relative difference in performance of varieties will be more stable over the environment within the region.

While this particular application provided the motivation for this research and the language is in terms of "cultivars" and "locations", the formulation of the problem and the results obtained are generally applicable to any interaction analysis.

The genotype by environment interaction can be controlled by classifying locations according to the similarity of their interactions with a fixed set of cultivars (genotypes). Such a procedure would result in dividing a set of locations into not necessarily adjacent regions, such that the interaction of cultivars and locations within each region would be small. Any classification of locations based on cultivar by location interaction will depend to some extent on the set of cultivars used. Therefore, to do a classification of locations based on the cultivar by location interaction which would be useful, the set of cultivars chosen must be a representative set of the cultivars to be used in the future; that is, the cultivar effects either must be considered to be random effects or the classification must be considered to be conditional on the set of cultivars used.

Different methods have been used to measure the similarity of locations. Let each location be represented by a vector $(v \times 1)$ of estimated cultivar by location interaction effects, $\underset{\sim}{\ell}_i' = (\widehat{\alpha\beta}_{1i}, \widehat{\alpha\beta}_{2i}, \ldots, \widehat{\alpha\beta}_{vi})$, $i = 1, \ldots, \ell$, where ℓ is the number of locations, and v is the number of cultivars, and let each "true" location be represented by a vector of the parametric cultivar by location interaction effects, $\underset{\sim}{\lambda}_i' = (\alpha\beta_{1i}, \alpha\beta_{2i}, \ldots, \alpha\beta_{vi})$, $i=1,2,\ldots,\ell$. Conditional on the set of cultivars being used, the $\alpha\beta_{ij}$ are fixed interaction effects for the ith location. One measure of similarity is the set of Euclidean distances between all pairs of locations (ℓ_i). Another measure of similarity used is the product moment correlation coefficient between ℓ_i and ℓ_j. The classification is then done by applying a clustering algorithm on the similarity measure.

Although classification of locations and of cultivars has been done in this manner, little is known about the reliability of the classification. That is, little is known about the relation of this classification to the "best" classification, where the "best" classification is defined to be that which is obtained when the same clustering algorithm is applied to the true interaction effects. The problem of how close the classification of locations, based on the estimated location vectors, is to the "best" classification, is equivalent to the problem of determining how similar the configuration of the estimated location points, $\underset{\sim}{\ell}_i$, is to the configuration of the true location points, $\underset{\sim}{\lambda}_i$. The two problems are equivalent because any consistent clustering procedure depends only on the relative spatial arrangement of both sets of points, and not on the absolute distances between the points of each set. Therefore, the impact of random error in any clustering procedure can be studied by analyzing the agreement between the configurations of the two sets of points.

The purpose of this thesis is to measure the repeatability of this spatial configuration, where the location points are defined by the vectors of location by cultivar interaction effects, for various levels of random error, sizes of interaction matrices, and true configurations of location points. To do this, a Monte Carlo experiment was carried out using 2000 independent sets of random errors to generate the estimated interaction effects for each parametric combination.

A closely related problem that will be discussed is the relative gain or loss of precision in the classification when the dimensionality of the space of estimated location points is reduced to include only the more important dimensions as reflected in the eigenvalues of $D'D$ where

$d_{ij} = \widehat{\alpha\beta}_{ij}$. A reduction in the dimensionality of the space of estimated location points according to the importance of the eigenvalues, by eliminating the smallest eigenvalues, corresponds to a partitioning of the interaction effects (Mandel (1971)). If, in fact, certain components of the interaction effects which are essentially random error can be eliminated then a gain in precision might be expected by this reduction in dimensionality of the space.

Since the eigenvalues were used as the basis for reducing the dimensionality of the space of the estimated location points, the reliability of the estimation of the eigenvalues is basic to the reliability of the classification using less than full dimension. Because of this, a small section was dedicated to the discussion of some distributional properties of the eigenvalues of the singular matrix $D'D$ for different sizes of the matrices, different true interaction effects and different choices of random errors.

2. LITERATURE REVIEW

Interactions of the cultivars with the environments in which they are grown are of major importance to the plant breeder. Because of this, different methods have been proposed to evaluate the cultivar performance in several locations.

Plaisted and Peterson (1959) suggested that the portion of the cultivar by location variance component contributed by a single cultivar could be used as a measure of cultivar dependability. Finlay and Wilkinson (1963) computed a linear regression of yield on the mean yield of all cultivars for each location and season, for each cultivar. Eberhart and Russell (1966), based on Finlay and Wilkinson (1963), used the regression of each cultivar on an environmental index, and a function of the squared deviations from this regression to provide estimates of cultivar stability. The environmental index was defined as the mean of all cultivars at the jth environment minus the grand mean. Okuno, et al. (1971) proposed two other methods for measuring the performance level of the environment. The first is a version of Finlay and Wilkinson's method, which uses one or more independent measures of the environments in place of the average performance of a large number of cultivars for each location and season. The second method is to extract one or more hypothetical factors for evaluating the responses of cultivars to different environments by applying principal components analysis to the elements of genotype by environment interaction.

The classification of environments (locations) into more or less homogeneous groups to control the within group genotype by environment interaction is another procedure of interest used in cultivar testing.

Abou-El-Fittouh, et al. (1969) used cluster analysis as a tool for classifying locations in order to minimize the within cluster genotype by environment interaction. Data for lint yield per hectare in upland cotton was used. Each location was represented by a vector (v x 1) of estimated interaction effects, $\underline{\ell}_i' = (\widehat{\alpha\beta}_{1i}, \widehat{\alpha\beta}_{2i}, \ldots, \widehat{\alpha\beta}_{vi})$, where v is the number of cultivars. Two similarity measures were computed, the distance coefficient, $d_{ij} = [\underline{\ell}_i'\underline{\ell}_i/v]^{\frac{1}{2}}$, which was found to be the more efficient for the purpose of defining regions when the criterion is the minimization of within region genotype by environment interaction, and the product-moment correlation coefficient, $r_{ij} = \text{corr}(\underline{\ell}_i, \underline{\ell}_j)$. These two measures of similarity were computed for all pairs of locations, $\underline{\ell}_i$. Using both sets of similarity measures they proceeded to do the classification using a clustering algorithm. The results suggested some modifications in the currently recognized zones of adaptation for cotton.

Most methods of classification operate on a matrix of the similarities between all pairs of objects. Then a criterion is used to fit a dendogram (Clifford and Stephenson, 1975). Fitting a dendogram to a given set of distances might be regarded as optimizing some goodness of fit criterion, $f(d_{ij}, d_{ij}^{+})$, where d_{ij} is any Euclidean similarity measure and d_{ij}^{+} is apparent Euclidean distance read from the resulting dendogram (Gower and Banfield, 1975). Several possible forms of this function have been suggested. Gower (1971) discussed some preliminary work on one of these criteria, a statistic R^2 which compares two different sets of distances without using the dependent pairwise measures of distance. He discussed the problem of fitting a configuration Y of n points, to another configuration X, also of n points. The vertices of X, P_i, are fixed, while those of Y, Q_i, are

allowed to translate, rotate, and dilate to cope with possible changes of scale between the two configurations.

The sum of squares, R^2, of the distances between corresponding vertices is minimized after the centroids of both sets of points are made to coincide. Assuming P and Q have been translated to have coincident centroids, $R^2 = \mathrm{tr}(P-\delta QH)(P-\delta QH)'$, where δ is a dilation factor and H is a rotation matrix. Minimization with respect to δ and H, gives

$$R^2_{min} = \mathrm{tr}\ PP' - \delta\,\mathrm{tr}\ P'QH\ ,$$

where $H = ZW'$, Z is the $\ell \times \ell$ matrix of the orthonormalized eigenvectors of $Q'PP'Q$, $Z'Z = ZZ' = I$, and W is the $\ell \times \ell$ matrix of the orthonormalized eigenvectors of $P'QQ'P$, $W'W = WW' = I$, $H'H = HH' = I$ and $\delta = \mathrm{tr}\ P'QH/\mathrm{tr}\ QQ'$. After substitution for δ,

$$R^2_{min} = \mathrm{tr}\ PP' - (\mathrm{tr}\ P'QH)^2/\mathrm{tr}\ QQ'\ .$$

Gower and Banfield (1975) studied the behavior of this criterion, with six other criteria, when dendograms are computed using the single linkage algorithm (Gower and Ross, 1969). They examined the distributions of the seven criteria when fitting single linkage clusters to n samples drawn from a spherical multinormal distribution in a varying number of dimensions v. The R^2_{min} criterion used by Gower and Banfield (1975) differed from the R^2_{min} discussed by Gower (1971) in that only rotation, not dilation, was permitted. Therefore, in Gower and Banfield (1975), $R^2_{min} = \mathrm{tr}\ PP' - 2\mathrm{tr}\ P'QH + \mathrm{tr}\ QQ'$ where H is defined as in Gower (1971). For this statistic a regression through the origin, weighted by the inverse of the statistic's variance, gave

mean(R^2min) = 1.022 nv, which was consistent with the simple relation mean(R^2min) = nv.

The conclusions found by Gower and Banfield (1975) were that although R^2min is not easily computed, its mean and variance are easily interpolated from a table of means and variances of R^2min for different n and v. Also R^2min is not a function of dependent measures as are the other criteria. All criteria were found to approach normality as n and v increased. The statistic R^2min seems best if metric dependencies are to be avoided, and it also has good distributional properties.

Mandel (1970) developed some theorems on the eigenvalues of the interaction covariance matrix, one of which states that if $(d_{ij}) = D$ is an m x n interaction matrix, where d_{ij} is normally distributed with mean zero, and cov(vec D) = $B \otimes A\sigma^2$ where vec D is a mn x 1 vector formed by stacking the columns of D one under the other, A is an m x m matrix of rank r, and B is an n x n matrix of rank s, then the non-zero eigenvalues of DD' are the same as those of the matrix TT' where $T = (t_{ij})$ is a r x s matrix of normally and independently distributed variates, t_{ij}, having zero mean and common variance σ^2. So he proved that the non-zero eigenvalues of DD' are the same as those of TT' where TT' is distributed as a Wishart with s degrees of freedom, $TT' \sim W_r(s, I\sigma^2)$.

Later, Mandel (1971) did a partitioning of the interaction effects, $\alpha\beta_{ij}$, i = 1,...,v, j = 1,...,λ, into the sum of multiplicative functions of i and j, $\alpha\beta_{ij} = \sum_{k=1}^{s} \vartheta_k u_{ik} v_{jk}$ where ϑ_k^2, k = 1,...,s, are the non-zero eigenvalues of $M'M$ (M is the matrix of true interaction effects, $\alpha\beta_{ij}$), (u_{ik}) is the eigenvector of MM' corresponding to ϑ_k^2, and (v_{jk}) is the eigenvector of $M'M$ corresponding to ϑ_k^2 .

Only a few multiplicative terms of the form $\theta\, v_j u_i$ (generally one or two) are retained in the model, the remaining terms are pooled together and regarded as experimental error. The same partitioning is done on the estimated interaction effects $\widehat{\alpha\beta}_{ij} = \sum_{k=1}^{s} \hat{\theta}_k \hat{u}_{ik} \hat{v}_{jk}$. He had shown in a previous paper, Mandel (1969), that there is a partitioning of the sum of squares of interaction of the usual analysis of variance that parallels the partitioning of the interaction effects. He developed an empirical criterion by which a decision could be made as to where the partitioning should be ended. For the sum of all terms involving $\hat{\theta}_k$ a sum of squares and a "degrees of freedom" were obtained. The interaction sum of squares associated with the kth non-zero root, $\hat{\theta}_k$, was shown to be $\hat{\theta}_k^2$. The clue to an appropriate definition of degrees of freedom, ν_k, was that in the absence of true systematic interaction effects, the quotients $\hat{\theta}_k^2/\nu_k$, $k = 1,\ldots,s$, should be estimates of only experimental error. Therefore, if σ^2 is the experimental error variance, and if none of the terms $\theta_k u_{ki} v_{kj}$ occurs in the real model, then $E(\hat{\theta}_k^2/\nu_k) = \sigma^2$ for each k, providing a definition for the degrees of freedom, $\nu_k = E(\hat{\theta}_k^2/\sigma^2)$. To obtain estimates of ν_k, a Monte Carlo experiment was carried out using 625 independent sets of identically and independently distributed (iid) normal (0,1) deviates for each combination of v and ℓ and ν_k was calculated as $\text{mean}(\hat{\theta}_k^2)$. Using Mandel's empirical estimates of ν_k and observing the magnitude of $\hat{\theta}_k^2/\nu_k$ in the partitioned interaction relative to σ^2, the error variance, a judgement can be made on how far the partitioning process should be carried.

3. METHODS AND MATERIALS

3.1. Description of Interaction Effects -- Model and Assumptions

To provide the interaction effects needed to accomplish the objectives described in the introduction, a two-way linear model with interaction was adopted;

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ijk} ,$$

$$i = 1,\ldots, v,\ j = 1,\ldots, \ell,\ k = 1,\ldots, n ,$$

where α_i is the ith cultivar effect, β_j is the jth location effect, $\alpha\beta_{ij}$ is the interaction effect of the ith cultivar and the jth location, and e_{ijk} are independently and identically distributed (iid) normal with zero mean and variance σ_e^2. The $\alpha\beta_{ij}$, conditional on the set of cultivars being used, are regarded as fixed effects and $\sum_i \alpha\beta_{ij} = \sum_j \alpha\beta_{ij} = 0$. The estimated interaction effects, $\widehat{\alpha\beta}_{ij}$, were calculated as $\widehat{\alpha\beta}_{ij} = y_{ij.} - y_{i..} - y_{.j.} + y_{...} = \alpha\beta_{ij} + (e_{ij.} - e_{i..} - e_{.j.} + e_{...})$ where (.) represents the average taken over that subscript.

Let $A = (a_{ij})$ be the $v \times \ell$ matrix of true interaction effects, i.e., $a_{ij} = \alpha\beta_{ij}$ and therefore $A'\underline{1}_v = \underline{0}_\ell$, $A\underline{1}_\ell = \underline{0}_v$, where $\underline{1}_k$ is a $k \times 1$ vector of ones and $\underline{0}_k$ is a $k \times 1$ vector of zeroes. The columns, $\underline{a}_j$, of the matrix A are regarded as the coordinates of the true location points as defined by the $v \times 1$ vector of true interaction effects. Therefore, the spatial configuration of the ℓ true location points is contained in the A matrix. The distance between locations i and j is given by $(\underline{a}_i - \underline{a}_j)'(\underline{a}_i - \underline{a}_j)$ which is equal

to $\sum_{k=1}^{m} \delta_k^2(\tau_{ik} - \tau_{jk})^2$ where δ_k^2, $k = 1,\ldots,m$, are the non-zero eigenvalues of $A'A$ and $(\tau_{1k},\ldots,\tau_{\ell k})$, $k = 1,\ldots,m$, are the orthonormalized eigenvectors of $A'A$ corresponding to the non-zero roots. There are, of course, an infinite number of possible parametric spatial configurations as defined by the A matrix which might be considered. The specific configurations that were considered in this study are described in Section 3.4.

Let $\hat{A} = (\hat{a}_{ij})$ be the $v \times \ell$ matrix of estimated interaction effects, i.e., $\hat{a}_{ij} = \widehat{\alpha\beta}_{ij} = \alpha\beta_{ij} + e_{ij.} - e_{i..} - e_{.j.} + e_{...}$ and $\hat{A}'\underline{1}_v = \underline{0}_\ell$, $\hat{A}\,\underline{1}_\ell = \underline{0}_v$. Therefore, in matrix notation

$$\hat{A} = A + (L_v - \tfrac{1}{v} J_v)E(I_\ell - \tfrac{1}{\ell} J_\ell)$$

where I_k is a $k \times k$ identity matrix, J_k is a $k \times k$ matrix of ones, and $E = (e_{ij.})$, $e_{ij.}$ is iid $N(0,\sigma_e^2/n = \sigma^2)$.

The interaction sum of squares, $SS(\alpha\beta)$, which is equal to $\sum_{i=1}^{v} \sum_{j=1}^{\ell} \hat{a}_{ij}^2$ can also be expressed in matrix notation as $\operatorname{tr} \hat{A}\hat{A}'$. A singular value decomposition of $\hat{A}$ gives $\hat{A} = UDV'$, where D^2 is a $k \times k$ diagonal matrix containing the non-zero eigenvalues, d_i^2, of $\hat{A}'\hat{A}$, U is a $v \times k$ matrix of the orthonormalized eigenvectors of $\hat{A}\hat{A}'$ corresponding to the non-zero eigenvalues, $U'U = I$, and V is a $\ell \times k$ matrix of the eigenvectors of $\hat{A}'\hat{A}$ corresponding to the non-zero eigenvalues, $V'V = I$. Then $ss(\alpha\beta) = \operatorname{tr}(\hat{A}'\hat{A}) = \operatorname{tr}(VDU'UDV') = \operatorname{tr} D^2 = \sum_{i=1}^{k} d_i^2$.

When $\{\alpha\beta\}$ are fixed effects and $\widehat{\alpha\beta}_{ij}$ is normally distributed, as assumed herein, $ss(\alpha\beta)/\sigma^2$ is distributed as a non-central Chi-square with $(\ell-1)(v-1)$ degrees of freedom and non-centrality parameter $\lambda = \tfrac{1}{2} \sum_{i,j} \alpha\beta_{ij}^2/\sigma^2$; i.e., $ss(\alpha\beta)/\sigma^2 = \sum d_i^2/\sigma^2$ is distributed as

$\chi'^2_{(\ell-1)(v-1)}(\lambda)$ (Searle (1971)). Since $E(\chi'^2_\nu(\lambda)) = \nu + 2\lambda$ and variance $(\chi'^2_\nu(\lambda)) = 2\nu + 8\lambda$ (see Searle (1971)),

$$E\left(\sum_{i=1}^{k} d_i^2/\sigma^2\right) = (\ell-1)(v-1) + \sum_{i=1}^{v}\sum_{j=1}^{\ell} \alpha\beta_{ij}^2/\sigma^2$$

$$= (\ell-1)(v-1) + \sum_{s=1}^{m} \delta_s^2/\sigma^2 = t$$

and

$$\mathrm{Var}\left(\sum_{i=1}^{k} d_i^2/\sigma^2\right) = 2(\ell-1)(v-1) + 4\sum_{i=1}^{v}\sum_{j=1}^{\ell} \alpha\beta_{ij}^2/\sigma^2$$

$$= 2(\ell-1)(v-1) + 4\sum_{s=1}^{m} \delta_s^2/\sigma^2$$

Scaling the d_i^2 by $1/t\sigma^2$ gives $s_i^2 = d_i^2/t\sigma^2$ with

3.1 $$E\left(\sum_{i=1}^{k} s_i^2\right) = 1$$

and

3.2 $$\mathrm{Var}\left(\sum_{i=1}^{k} s_i^2\right) = \frac{2(\ell-1)(v-1)\sigma^4 + 4\sum_{i=1}^{m} \delta_i^2\,\sigma^2}{\left[(\ell-1)(v-1)\sigma^2 + \sum_{i=1}^{m} \delta_i^2\right]^2}$$

$$= \frac{2(1-r^2)}{(\ell-1)(v-1)}$$

where $r = \sum_{i=1}^{m} \delta_i^2 \,/\, \left(\sum_{i=1}^{m} \delta_i^2 + (\ell-1)(v-1)\sigma^2\right)$.

In all studies reported herein, the magnitude of error, $\sigma^2 = \sigma_e^2/n$, relative to true interaction is controlled by choice of r, $0 \leq r \leq 1$. The relative spatial configuration is controlled by choice of the relative magnitudes of the non-zero roots, δ_i.

3.2 Measuring the Agreement Between the True and the Estimated Location Points.

Since the classification of locations depends only on the relative spatial configuration of the estimated location points, a standardized version of the R^2_{min} statistic (Gower, 1971) was used to measure the degree to which the estimated location points, $\underline{\hat{\lambda}}_i' = (\widehat{\alpha\beta}_{1i}, \widehat{\alpha\beta}_{2i}, \cdots, \widehat{\alpha\beta}_{vi})$, reproduced the spatial configuration of true location points, $\underline{\lambda}_i' = (\alpha\beta_{1i}, \alpha\beta_{2i}, \cdots, \alpha\beta_{vi})$. R^2_{min} is the minimum possible sum of squared distances between corresponding vertices of two configurations X and Y after allowing translation, rotation and dilation of one set of points to "match" as closely as possible the other. Minimization of

$$R^2 = \mathrm{tr}(A - \delta\hat{A}H)(A - \delta\hat{A}H)'.$$

with respect to δ and H, a dilation scaler and rotation matrix, respectively, gives

$$R^2_{min} = \mathrm{tr}\, AA' - \delta\,\mathrm{tr}(\hat{A}\, HA')$$

$$= \mathrm{tr}\, AA' - [\mathrm{tr}\,(\hat{A}\, HA')]^2/\mathrm{tr}\, \hat{A}\hat{A}'$$

where $A'\hat{A}$ was decomposed as WCZ' where C is a $(\ell \times \ell)$ diagonal matrix containing the positive square roots of the eigenvalues of $(A'\hat{A})'(A'\hat{A})$, W is the $(\ell x \ell)$ matrix of the orthonormalized eigenvectors of $(A'\hat{A})(A'\hat{A})'$, $W'W = WW' = I$, and Z is the $(\ell x \ell)$ matrix of the

orthonormalized eigenvectors of $(A'\hat{A})'(A'\hat{A})$, $ZZ' = Z'Z = I$, and $H = ZW'$, $H'H = HH' = I$, $\delta = \text{tr } \hat{A} H A'/\text{tr } \hat{A} \hat{A}'$ (see Gower, 1971).

For all results reported herein, R^2_{min} was standardized by dividing by $\text{tr } AA'$ to facilitate comparisons over different A matrices. Thus,

$$m^2 = R^2_{min}/\text{tr } AA'$$

$$= 1 - (\text{tr } \hat{A} H A')^2/(\text{tr } \hat{A}\hat{A}' \text{ tr } AA').$$

Note that $1 - m^2$ is the square of the correlation between the elements of A and the elements of $\hat{A} H$, since, in both cases, the centroids are zero. Also note that $\text{tr } AA'$ is the sum of squared distances of the true location points defined by A from their centroid so that this standardization puts all A matrices on a common basis in terms of the dispersion of the true location points, or equivalently in terms of the sums of squares of true interaction effects. Thus, m^2 is the standardized minimum sum of squares of the distances between corresponding vertices, $\underline{a}_i$ and $\underline{\hat{a}}_i$, that can be obtained by allowing rotation and dilation of $\hat{A}$ or $1 - m^2$ is the maximum correlation that can be obtained between the elements of A and the elements of $\hat{A}H$.

Clearly, $0 \leq 1 - m^2 \leq 1$, because it is the square of a correlation coefficient, or $0 \leq m^2 \leq 1$. The lower bound of m^2, zero, is obtained when $\hat{A}$ equals A; i.e., when $\sigma^2 = 0$. The other limit of one is reached if and only if $(\text{tr } \hat{A}HA')^2/\text{tr } AA' = 0$; that is, if and only if $\text{tr } \hat{A}HA' = 0$, assuming $\text{tr } AA'$ finite. $\text{Tr } \hat{A}HA' = 0$ implies $\text{tr } WCZ'ZW' = \text{tr } C = \Sigma c_i = 0$ (using the definition of H and the decomposition of $A'\hat{A}$) or that $c_i = 0$ for all i since $c_i \geq 0$ by definition. Therefore, $m^2 = 1$ if and only if $A'\hat{A} = 0$, which means that

every column of A must be orthogonal to every column of $\hat{A}$. But

$$\text{prob}(m^2=1) = \text{prob}(A'\hat{A} = 0)$$

$$= \text{prob}\{A'[A+(I-\tfrac{1}{v}J)E(I-\tfrac{1}{\ell}J)] = 0\}$$

$$= \text{prob}\{A'(I-\tfrac{1}{v}J)E(I-\tfrac{1}{\ell}J) = -A'A\} = 0.$$

since e_{ij} is an absolutely continuous random variable and a_{ij} are fixed numbers. Therefore, m^2 is less than one with probability one. A better upper bound for $E(m^2)$ was sought, and it was proven (see appendix) that $E(m^2) \leq 1 - r$, where $r = \sum_{i=1}^{m}\delta_i^2/(\sum_{i=1}^{m}\delta_i^2 + (\ell-1)(v-1)\sigma^2)$.

From this proof, the bias in the statistic m^2, if any, was determined to be negative, $E(m^2) \leq 1 - \rho^2_{A,\hat{A}H}$ (see appendix).

3.3 Monte Carlo Study of Repeatability of Spatial Configuration

3.3.1 Definition of Parametric A Matrices. To assess the repeatability with which $\hat{A}$, the matrix of estimated interaction effects, reproduces the spatial configuration defined by the A matrix of true interaction effects, a Monte Carlo study was run using several parametric definitions of A.

The A matrices were defined in terms of the components of the singular value decomposition of A, $v^A{}_\ell = v^T{}_m \Delta_m T'_\ell$, where $T'T = I$ and $T'T = I$. The rows of $T(\ell \times m)$ may be regarded as coordinates of the ℓ locations in R^m, in a standardized coordinate system; standardized in the sense that the centroid is zero, the axes are orthogonal, and the sum of squares of the values of the i^{th} coordinate over all the location points equals one for all i, $T'T = I$, m is the number of non-zero eigenvalues of AA'. Choice of Δ, the $m \times m$ diagonal matrix

containing the positive square roots of the non-zero eigenvalues of AA' determines the final scaling of each axis. Thus, $\Delta T'$ completely defines the spatial configuration, including all pairwise distances, of the ℓ location points in $\mathbb{R}^m$. The matrix $\Gamma(\ell \times m)$ serves as a one-to-one transformation carrying the ℓ location points in $\mathbb{R}^m$ to $\mathbb{R}^v$, $m < v$. The spatial configuration, including all pairwise distances, is unaffected since $\Gamma'\Gamma = I$; i.e., the location points in $\mathbb{R}^v$ lie on an m-dimensional hyperplane in $\mathbb{R}^v$.

The following A matrices, $A = \Gamma\Delta T'$, were used:

1. $\ell = 8$, $v = 8$, $m \leq 3$

$$\Gamma' = \tfrac{1}{2}\begin{bmatrix} 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & -1 & -1 \end{bmatrix}$$

$$T' = \frac{1}{2\sqrt{2}}\begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

$\Delta = \text{diag}(\delta_1, \delta_2, \delta_3)$ for all choices of δ_i^2 shown in Table 3.1. See Figure 3.1 for the spatial relation of the eight locations.

Table 3.1. Values of eigenvalues (δ_i^2), and standardized eigenvalues, $\delta_i^+ = \delta_i^2 / \sum_{i=1}^{m} \delta_i^2$, used for Δ when $\ell = v = 8$.

Case	δ_1^2	δ_2^2	δ_3^2	δ_1^+	δ_2^+	δ_3^+
(i)	128	128	128	.333	.333	.333
(ii)	.02	.02	.02	.333	.333	.333
(iii)	128	128	8	.485	.485	.030
(iv)	128	8	8	.889	.056	.056
(v)	1.28	.72	0	.640	.360	0
(vi)	2	0	0	1.0	0	0

2. $\ell = 4$, $v = 8$, $m \leq 3$.

Γ' is as shown in 1, above.

$$\Gamma' = \tfrac{1}{2}\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$\Delta = \text{diag}(\delta_1, \delta_2, \delta_3)$ for all choices of δ_i^2 shown in Table 3.2.

See Figure 3.2 for the spatial relation of the four locations.

Table 3.2. Values of eigenvalues (δ_i^2), and standardized eigenvalues, $\delta_i^+ = \delta_i^2 / \sum_{i=1}^{m} \delta_i^2$, used for Δ when $\ell = 4$, $v = 8$.

Case	δ_1^2	δ_2^2	δ_3^2	δ_1^+	δ_2^+	δ_3^+
(i)	64	64	64	.333	.333	.333
(ii)	64	64	4	.485	.485	.030
(iii)	64	4	4	.889	.056	.056

Figure 3.1. Configuration of the location points given by the columns of the $v \times \ell$ interaction matrix A, when $\ell = 8$, $v = 8$ and $A'A$ has at most three non-zero eigenvalues, δ_1, δ_2, and δ_3. $\delta_i^* = \delta_i / 2\sqrt{2}$

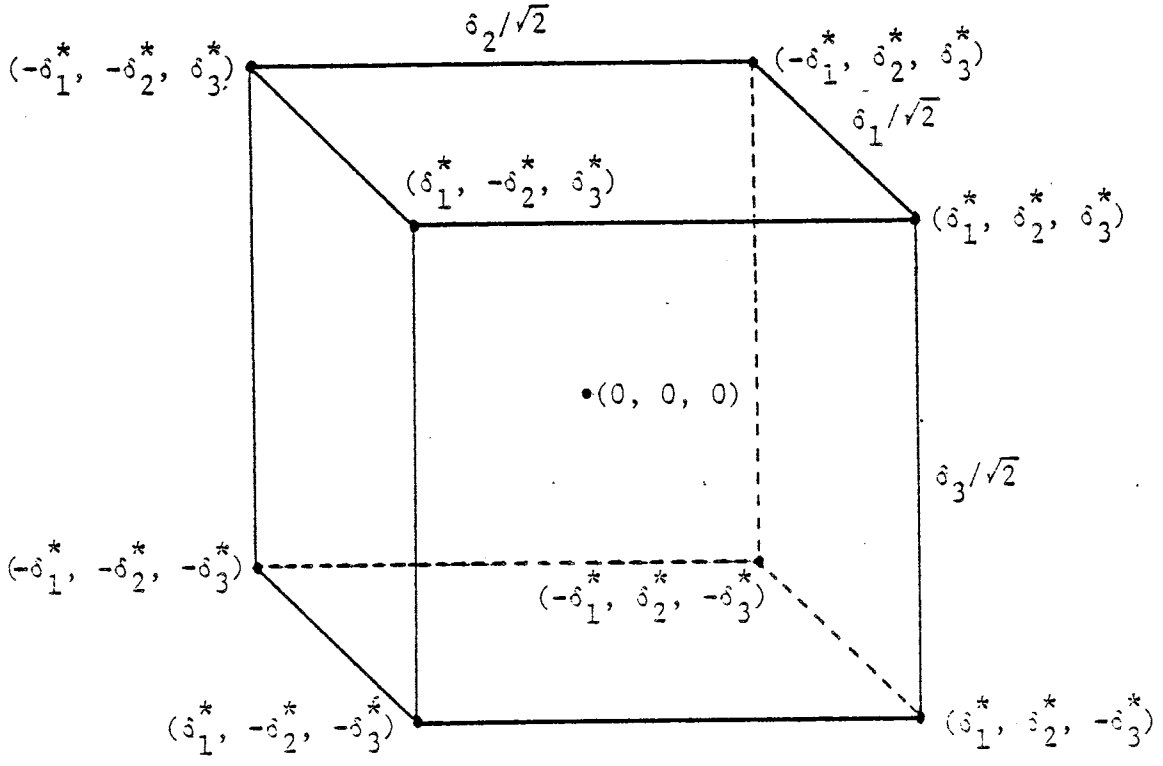
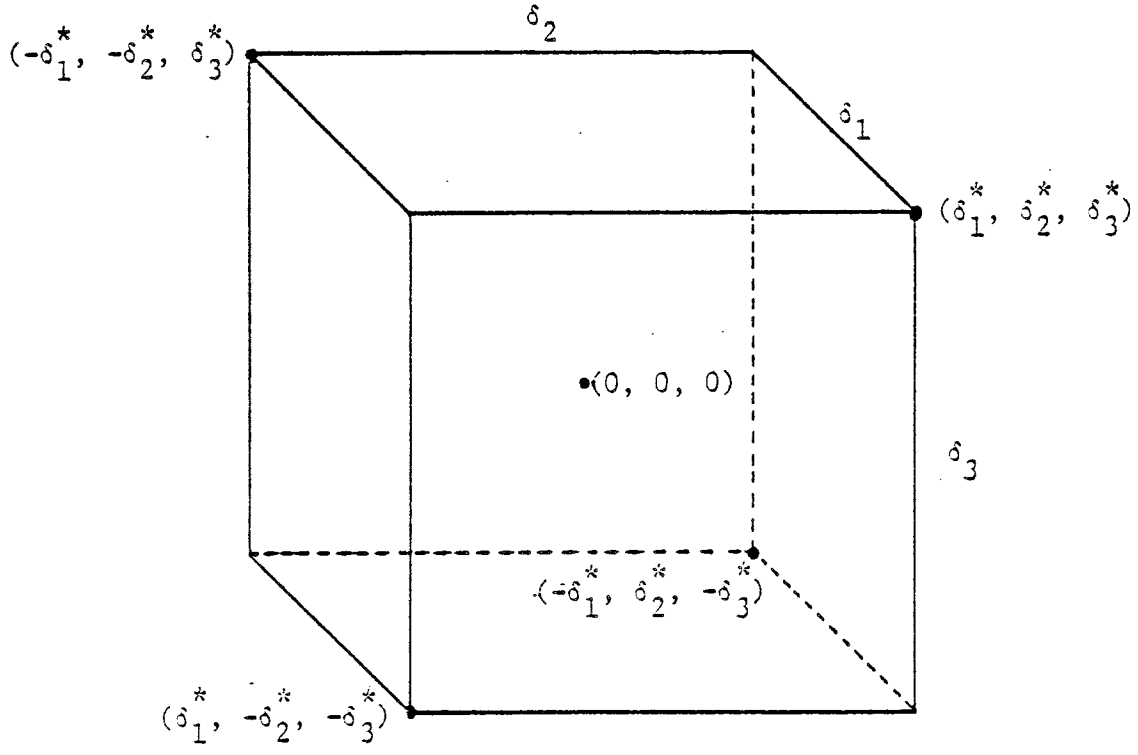

$\delta_2/\sqrt{2}$
$(-\delta_1^*, -\delta_2^*, \delta_3^*)$
$(-\delta_1^*, \delta_2^*, \delta_3^*)$
$\delta_1/\sqrt{2}$
$(\delta_1^*, -\delta_2^*, \delta_3^*)$
$(\delta_1^*, \delta_2^*, \delta_3^*)$
$(0, 0, 0)$
$\delta_3/\sqrt{2}$
$(-\delta_1^*, -\delta_2^*, -\delta_3^*)$
$(-\delta_1^*, \delta_2^*, -\delta_3^*)$
$(\delta_1^*, -\delta_2^*, -\delta_3^*)$
$(\delta_1^*, \delta_2^*, -\delta_3^*)$

Figure 3.2. Configuration of the four location points given by the columns of the $v \times \ell$ interaction matrix A, when $\ell = 4$, $v = 8$ and $A'A$ has at most three non-zero eigenvalues, δ_1, δ_2, and δ_3. $\delta_i^* = \delta_i/2$.


δ_2
$(-\delta_1^*, -\delta_2^*, \delta_3^*)$
δ_1
$(\delta_1^*, \delta_2^*, \delta_3^*)$
$(0, 0, 0)$
δ_3
$(-\delta_1^*, \delta_2^*, -\delta_3^*)$
$(\delta_1^*, -\delta_2^*, -\delta_3^*)$

3. $A = 0$. All possible combinations of ℓ and v for $\ell, v = 3,4,5,6,7,$ and 8 were used.

4. $A = (I - \frac{1}{8} J_8)$ for $\ell = 8$, $v = 8$. There are seven equal non-zero eigenvalues for a matrix of this form.

3.3.2. Generation of $\hat{A}$, the matrix of estimated interaction effects. For each A matrix, the $(v \times \ell)$ matrix $E = (e_{ij.})$ was generated using a random number generator called Super Duper, developed at McGill University (Dickey, 1978), so the $e_{ij.}$, $i = 1,\ldots,v$, $j = 1,\ldots,\ell$ were iid normal random variables with mean zero and variance $\sigma_e^2/n = \sigma^2$.

The variance of the random error, σ^2, was chosen so that the ratio

$$r = \frac{\sum_{i=1}^{m} \delta_i^2}{\sum_{i=1}^{m} \delta_i^2 + (\ell-1)(v-1)\sigma^2}$$

was equal to .001, 1., .25, .5, .75 in case 1, and .1, .25, .5, .75 in case 2. r is always zero in case 3, and r is always .001 in case 4. The values of σ^2 are given in Table 3.3.

Table 3.3. Values of the error variance, σ^2, used to generate the random error $e_{ij.}$, for each choice of parameters used.

case ratio	1-i	1-ii	1-iii	1-vi	1-v,vi	2-i	2-ii	2-iii
.001	7828.90	1.22	5382.37	2935.84	40.78			
.1	70.53	.011	48.49	26.45	.37	82.29	56.57	30.86
.25	23.51	.0037	16.16	8.82	.12	27.43	18.86	10.29
.50	7.84	.0012	5.39	2.94	.041	9.14	6.29	3.43
.75	2.61	.00041	1.80	.98	.014	3.05	2.10	1.14

The matrix of estimated effects, $\hat{A}$, was then calculated as

$$\hat{A} = A + (I_v - \frac{1}{v} J_v)E(I_\ell - \frac{1}{\ell} J_\ell)$$

where I_k is the $k \times k$ identity matrix and J_k is a $k \times k$ matrix of ones.

All parametric combinations having the same ℓ and v were compared using the same set of 2000 random error matrices. This was done to increase the precision of the comparison of results from different A matrices.

3.3.3. Analysis of the $\hat{A}$ matrices. The reproducibility of the A matrix by the $\hat{A}_i$ matrix was measured by computing m^2, as defined previously, for each $\hat{A}_i$ matrix, $i = 1,2,\ldots,2000$, where 2000 is the number of independent sets of random errors generated. In addition, to determine the loss or gain of precision in reproducing the true spatial configuration one might realize by considering only the "dominant" dimensions of $\hat{A}_i$, the $\hat{A}_i$ matrix was decomposed according to the singular value decomposition (SVD), using the SVD decomposition of IMSL (1975), i.e., $\hat{A}_i = U_i D_i V_i'$ where U_i is the $v \times k$ matrix of the orthonormalized eigenvectors of $\hat{A}_i\hat{A}_i'$ corresponding to the non-zero eigenvalues, $U_i'U_i = I$, V_i is the $(\ell \times k)$ matrix of the orthonormal eigenvectors of $\hat{A}_i'\hat{A}_i$ corresponding to the non-zero eigenvalues, $V_i'V_i = I$, and D_i^2 is a $(k \times k)$ diagonal matrix containing the non-zero eigenvales of $\hat{A}_i\hat{A}_i'$ $(d_{1i}^2 > d_{2i}^2 > \ldots > d_{ki}^2)$.

Then the matrices $\hat{A}_i(p)$ were defined as $\hat{A}_i(p) = U_i(p) \cdot D_i(p) \cdot V_i(p)'$, $p = 1,2,3$, and k, where $D_i(p)$ is the diagonal matrix of the p largest roots of $\hat{A}_i\hat{A}_i'$, and $U_i(p)$ and

$V_i(p)$ are the matrices of the eigenvectors of $\hat{A}_i\hat{A}_i'$ and $\hat{A}_i'\hat{A}_i$, respectively, corresponding to the p largest roots; note $\hat{A}_i(k) = \hat{A}_i$. Thus, for example, $A_i(3) = (\underline{u}_{1i},\underline{u}_{2i},\underline{u}_{3i})\text{diag}(d_{1i},d_{2i},d_{3i})(\underline{v}_{1i}\underline{v}_{2i}\underline{v}_{3i})' = \sum_{t=1}^{3} d_{ti}\underline{u}_{ti}\underline{v}_{ti}'$.

For each $\hat{A}_i(\bar{B})$, $R_i(p) = A'\hat{A}_i(p)$ $(\ell \times \ell)$, $p = 1,2,3$, and k, was calculated, and the diagonal $(\ell \times \ell)$ matrix, $C_i(p)$ of the positive square root of the eigenvalues of $R_i(p)'R_i(p)$ were found, using the SVD subroutine of IMSL (1975).

Then the statistic m^2 was calculated as

$$m_i^2(p) = 1 - (\text{tr } C_i(p))^2/(\text{tr } AA \text{ tr } \hat{A}_i(p)\ \hat{A}_i(p)'),\ p=1,2,3, \text{ and } k\ (m_i^2(k) = m_i^2).$$

m_i^2 is calculated for each of the $\hat{A}_i(p)$ matrices so they can be compared to determine the gain or loss of precision in the fitting of the true spatial configuration of location points when the dimensionality of the space of the estimated location points has been reduced.

The process described above was repeated 2000 times (2000 was taken to be a sufficiently large number to give accepted estimates of the mean and variance of the statistics), and the first four moments around the origin were obtained for $m_i^2(p)$, $p=1,2,3$, and k, and as well as for each of the estimated eigenvalues, d_{ji}^2, $j=1,2,\ldots,\ell - 1$.

Also, the histograms of $m_i^2(p)$, $p=1,2,3$, and k, and of d_{ji}, $j=1,2,\ldots, \ell - 1$, were obtained. The histograms and the moments were calculated using a set of FORTRAN subroutines for generating and summarizing large Monte Carlo studies (Dickey, 1978).

4. REPRODUCIBILITY OF SPATIAL CONFIGURATION - DISTRIBUTIONAL PROPERTIES OF m^2.

4.1 Results.

The statistic m^2 expresses the sum of squared distances of the observed location points, represented by the columns of the $v \times \ell$ interaction matrix, $\hat{A}$, after rotation and dilation, from the corresponding true location points, defined by the columns of A, as a proportion of the sum of squared distances of the true location points from their centroid. Alternately, $1 - m^2$ is the square of the product moment correlation between the elements of A and $\hat{A}H$, where H is the the rotation matrix. In these senses, m^2 reflects the degree to which an observed configuration of location points reproduces the original or "true" configuration. An assessment of the reproducibility of a spatial configuration and, consequently, the degree to which a classification of locations according to cultivar by location interaction might be successful is obtained by studying the distributional properties of m^2 for different parametric situations.

In Table 4.1 is presented the observed means of m^2, $\bar{m}$, based on 2000 samples, for different A matrices (spatial configurations) and values of r, $r = \sum_{i=1}^{m} \delta_i^2 / (\sum_{i=1}^{m} \delta_i^2 + (\ell-1)(v-1)\sigma^2)$, where δ_i^2, $i=1,\ldots,m$, are the non-zero eigenvalues of AA', $0 \leq r \leq 1$. Clearly, at the upper limit of r, $\hat{A} = A$ (σ^2 goes to zero) and $m^2 = 0$. It is clear from Table 4.1 that $\bar{m}$ is a monotonically decreasing function toward zero from some upper limit, $b(A) < 1$, which depends on the parametric situation defined by A (Table 4.2). Also, from section 3.2 m^2 is strictly less than one with probability one, and from the appendix $E(m^2)$

Table 4.1 Mean values of the statistic m^2 for different interaction matrices A (v x ℓ), and different ratios r.

v,ℓ	Relative sizes of the non-zero roots of A'A	ratio (r) .001	.1	.25	.5	.75	1.0[e]
8,8	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.597[b]	.544[b]	.457[b]	.300[b]	.145[a]	0.0
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.664[b]	.596[b]	.490[b]	.315[b]	.150[a]	0.0
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.721[b]	.634[b]	.511[b]	.322[b]	.151[a]	0.0
	$9\delta_1^2 = 16\delta_2^2 > 0$	.723[b]	.652[b]	.541[b]	.354[b]	.172[a]	0.0
	$\delta_1^2 > 0$	.848[b]	.761[b]	.630[b]	.413[b]	.202[a]	0.0
8,4	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$		.652[c]	.579[c]	.413[c]	.209[b]	0.0
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2$		.675[c]	.582[c]	.401[c]	.204[b]	0.0
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$		.694[c]	.581[c]	.393[c]	.203[b]	0.0
	$\delta_1^2 > 0$	.856[b]				.219[b]	
4,8	$\delta_1^2 > 0$	.668[c]				.171[a]	
4,4	$\delta_1^2 > 0$	.664[d]				.176[c]	

[a] Standard error of mean = 0.01.

[b] Standard error of mean = 0.02.

[c] Standard error of mean = 0.03.

[d] Standard error of mean = 0.05.

[e] Results for r = 1.0 are exact.

Table 4.2. Estimates of the upper bound of m^2, b(A), obtained from the linear regression of m^2 on (1-r), where possible.

Relative sizes of the non-zero roots of AA′	ℓ=8, v=8	ℓ=4, v=8	ℓ=8, v=4	ℓ=4, v=4
$\delta_1^2 = \delta_2^2 = \ldots = \delta_7^2 > 0$	.249[a]			
$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.602	.759	.286[a]	
$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.657	.769		
$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.698	.776		
$9\delta_1^2 = 16\delta_2^2 > 0$	.721			
$\delta_1^2 > 0$	.843	.856[a]	.668[a]	.664[a]

[a]Values for r = 0.001.

is less than or equal to $1 - r$. In Figures 4.1 and 4.2 are shown a nearly linear relationship between $\bar{m}$ and $1 - r$, for each of the A matrices. In all cases, a linear regression through the origin of $\bar{m}$ on $(1-r)$ gave a good representation of the change in $\bar{m}$ for changing r, $\bar{m} = b(A)\cdot(1-r)$. For a given A matrix, the regression coefficient $b(A)$ is the ordinary least squares estimate of the upper bound of $\bar{m}$ for that A, (Table 4.2). The smallest coefficient of determination obtained was $R^2 = 0.997$ and the greatest difference between the $b(A)$ calculated from the regression and the values of $\bar{m}$ for $r = 0.001$ was 0.023.

From Table 4.1, the mean value of m^2 is functionally dependent on the relative sizes of the non-zero roots of the AA' matrix. (The results obtained for m^2 were not changed when the values of the roots of AA' were changed, while keeping their relative sizes constant.) For a given value of ℓ and v, the mean of m^2 is largest when AA' has only one non-zero eigenvalue and smallest when AA' has $\ell - 1$ equal eigenvalues. The mean of m^2 for fixed ℓ and v is not only dependent on the relative sizes of the non-zero roots of AA', but it also depends on the number of non-zero roots. As the number of non-zero roots of AA' increased, the value of the mean of m^2 decreased.

Changes in v appear to have greater impact on $\bar{m}$ than do changes in ℓ. When AA' has only one non-zero eigenvalue and ℓ was changed from $\ell = 8$ to $\ell = 4$, keeping $v = 8$ constant, $\bar{m}$ changed very little, from 0.848 to 0.856 ($r = 0.001$), but when v was decreased from $v = 8$ to $v = 4$, keeping $\ell = 8$ constant, $\bar{m}$ changed appreciably from 0.848 to 0.668 ($r = 0.001$). The same effect was seen when ℓ was kept constant at $\ell = 4$ and v was changed from 8 to 4 ($\bar{m}$ for

Figure 4.1. The statistic $\bar{m}$ against $1 - r$, for each of the 8×8 interaction matrices A.

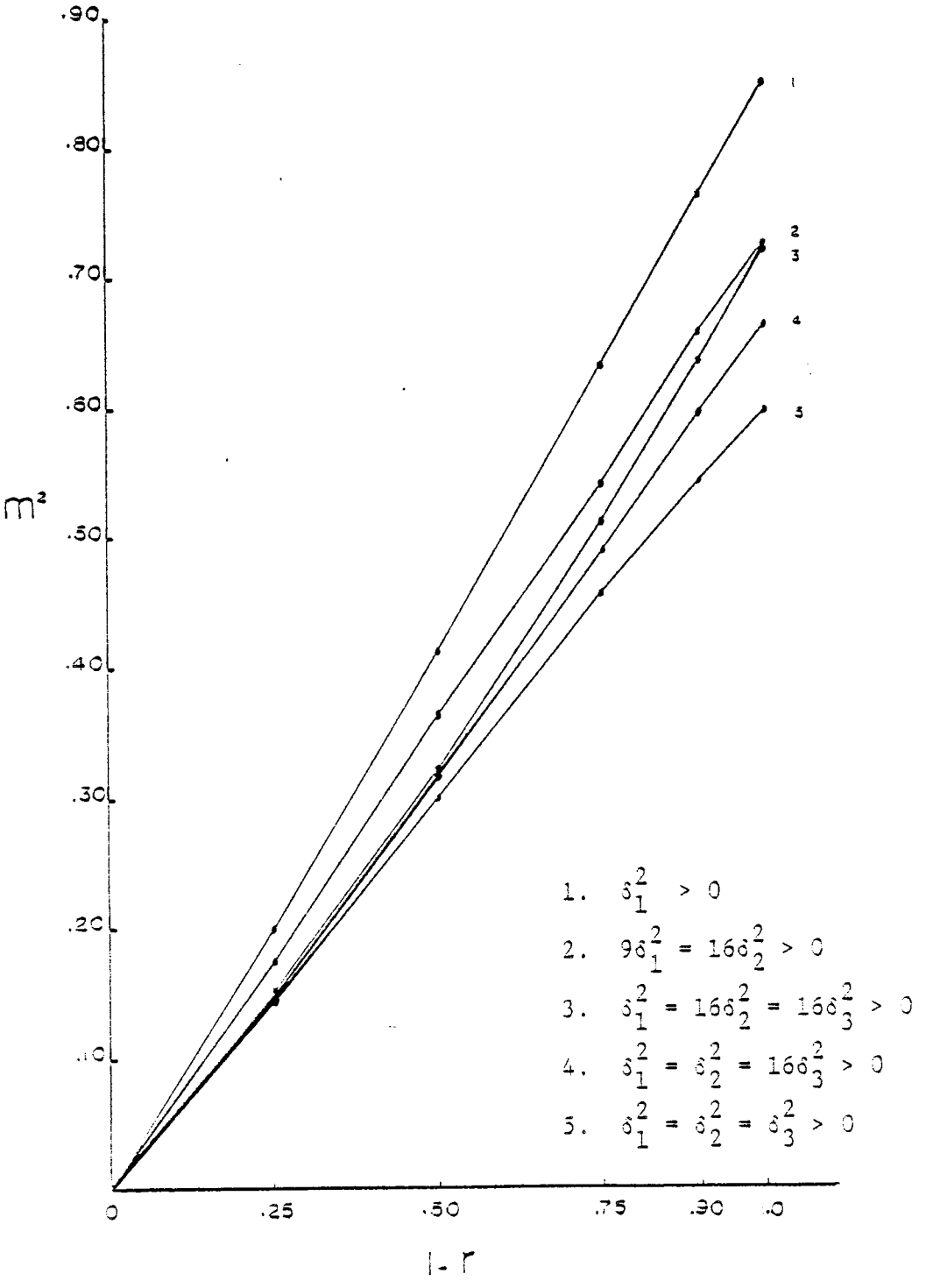
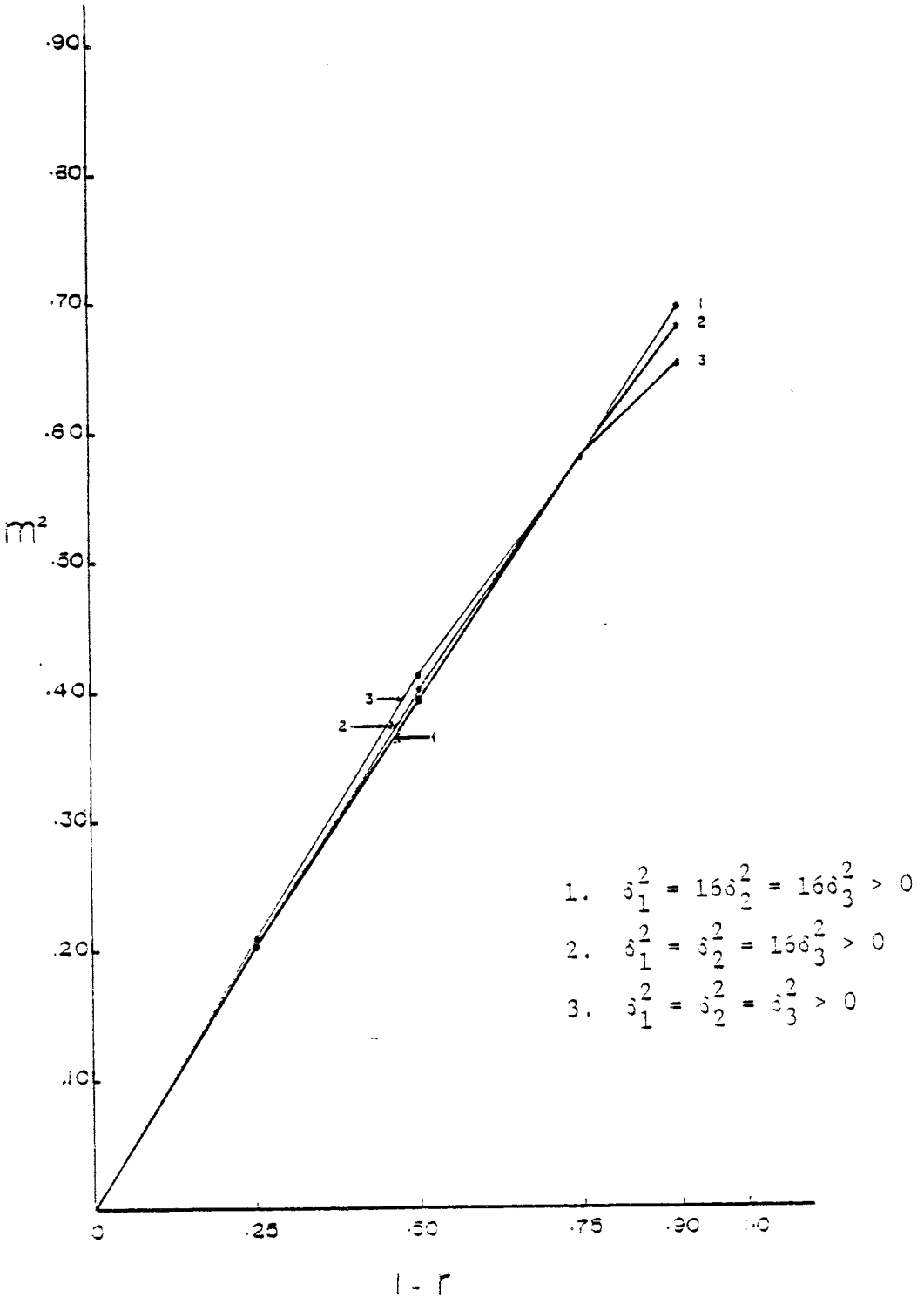

.90
.80
.70
.60
.50
.40
.30
.20
.10
m^2
0
.25
.50
.75
.90
1.0
$1-r^2$
1
2
3
4
5
1. $\delta_1^2 > 0$
2. $9\delta_1^2 = 16\delta_2^2 > 0$
3. $\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$
4. $\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$
5. $\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$

Figure 4.2. The statistic $\bar{m}$ against $1 - r$, for each of the 8×4 interaction matrices A.



1. $\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$
2. $\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$
3. $\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$

$r = 0.001$ changed from .856 to .664), compared to when ℓ was changed from 8 to 4 keeping $v = 4$ constant ($\bar{m}$ for $r = 0.001$ changed from .668 to .664).

After trying, unsuccessfully, to obtain

$$E(m^2) = 1 - E\,\frac{(\text{tr } A'AH)^2}{\text{tr } AA' \text{tr } \hat{A}\hat{A}'},$$

it was decided to approximate $E(m^2)$ by

$$1 - \frac{E(\text{tr } A'AH)^2}{\text{tr } AA'\, E(\text{tr } \hat{A}\hat{A}')} .$$

This was accomplished for the case where AA' has only one non-zero eigenvalue, because in this case

$$(\text{tr } A'\hat{A}H)^2 = (\text{tr } C)^2 = \text{tr } C^2 = \text{tr } A'\hat{A}\hat{A}'A \qquad (A'\hat{A} = WCZ',\ H = ZW').$$

Therefore, when there is only one non-zero root,

$$E(m^2) \doteq 1 - \frac{E(\text{tr } A'\hat{A}\hat{A}'A)}{\text{tr } AA'\, E(\text{tr } \hat{A}\hat{A}')} = 1 - \frac{\text{tr}(AA')^2 + (\ell-1)\sigma^2\ \text{tr } AA'}{\text{tr } AA'(\text{tr } AA' + (\ell-1)(v-1)\sigma^2)}$$

$$= 1 - \frac{\delta_1^4 + \delta_1^2\,(\ell-1)\,\sigma^2}{\delta_1^2(\delta_1^2 + (\ell-1)(v-1)\sigma^2)} = \frac{v-2}{v-1}\,(1-r) .$$

This approximation for the expected value of m^2 was very close to the observed values of $\bar{m}$; compare $\bar{m} = 0.843$, 0.856, and 0.664 (Table 4.1, $r = 0.001$) with the approximation, $E(m^2) \doteq 0.856$, 0.856, and 0.666 for the cases $(\ell,v) = (8,8)$, $(4,8)$, and $(4,4)$, respectively. It also can be proved that this approximation for the $E(m^2)$ when

AA' has only one non-zero root is an upper bound to $E(m^2)$ for the other cases of A matrices.

The variances of m^2 are presented in Table 4.3 for the different ratios r, and for the various A matrices and Figure 4.3 represents the plot of the variance of m^2, $V(m^2)$, against r. The variances of m^2 reach a maximum value when the ratio is in the vicinity of 0.1 and 0.25, depending on the A matrix. This interval for which the largest variance is obtained also is the interval where the skewness of the distribution of m^2 changes from negative to positive (skewness coefficients are not shown). As the number of locations ℓ and the number of cultivars v were decreased, the variances of m^2 increased. The increase in $V(m^2)$ is more drastic when v is decreased than when ℓ is decreased; compare $V(m^2) = 0.538$, 1.136, 1.937, and 4.168 for the cases $(v,\ell) = (8,8)$, $(8,4)$, $(4,8)$, and $(4,4)$, respectively.

In Figure 4.4 a-h are given box plots of the empirical distributions of m^2. The rectangular box extends from the first to the third quartile with the median shown by the middle horizontal line. The two horizontal lines above the box represent the 95th, and 97.5th percentiles, respectively; the two horizontal lines below the box represent the 5th and 2.5th percentiles, respectively. The $(\cdot)$ represents the extreme observations from the sample of 2000. The box plots have been spaced according to the ratio r. From these, it is clear that the distribution of m^2 changes from negative skewness when r is small, to positive skewness as r increases which is reflecting the impact of the finite boundaries on m^2. Also, as r increases toward one, and ℓ and v are fixed, the empirical distributions of m^2 for those A matrices for which $A'A$ have an equal number of non-zero roots become very similar.

Table 4.3. Variance of the statistic m^2 (multiplied by 100) for the different interaction matrices A (v x ℓ), and for different ratios r.

v,ℓ	Relative sizes of the non-zero roots of A'A	r .001	.1	.25	.5	.75	1.0[a]
8,8	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.954	1.011	.950	.550	.140	0.0
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.899	1.045	1.011	.581	.151	0.0
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.836	1.113	1.083	.600	.152	0.0
	$9\delta_1^2 = 16\delta_2^2$	.807	1.001	1.010	.607	.166	0.0
	$\delta_1^2 > 0$	.538	.891	1.029	.670	.190	0.0
8,4	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$		1.437	1.612	1.429	.507	0.0
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$		1.713	1.954	1.442	.467	0.0
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$		1.962	2.154	1.441	.446	0.0
	$\delta_1^2 > 0$	1.136				.507	
4,8	$\delta_1^2 > 0$	1.937				.401	
4,4	$\delta_1^2 > 0$	4.163				.959	

[a]Results for r = 1.0 are exact.

Figure 4.3. The variance of m^2 (x 100) against the ratio (r), for the 8 x 4 and 8 x 8 interaction matrices A.

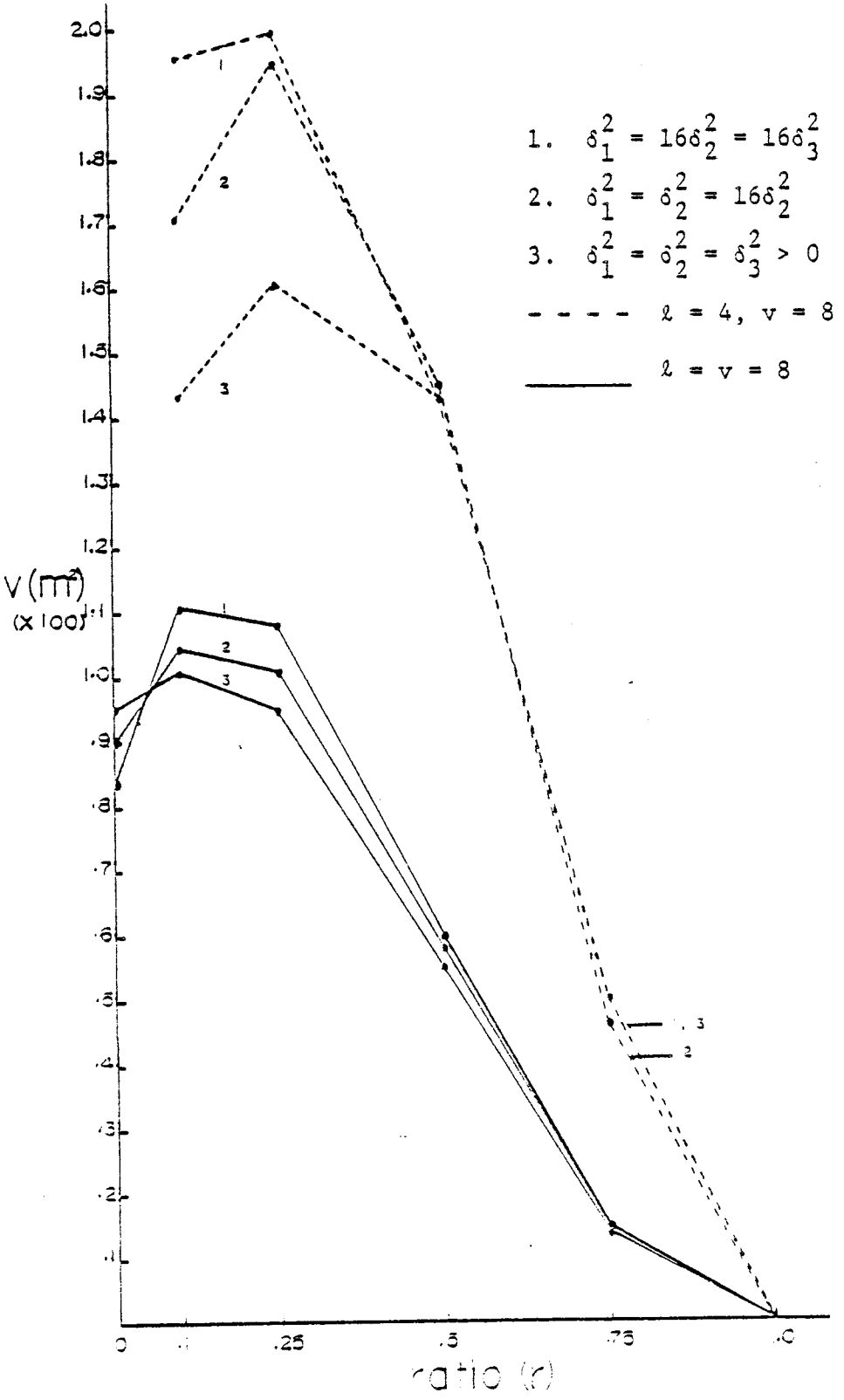

1. $\delta_1^2 = 16\delta_2^2 = 16\delta_3^2$
2. $\delta_1^2 = \delta_2^2 = 16\delta_2^2$
3. $\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$
- - - - $\ell = 4, v = 8$
——— $\ell = v = 8$
V(m²)
(x 100)
2.0
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.25
.5
.75
1.0
ratio (r)

Figure 4.4 a-h. The empirical conditional distributions of the statistic m^2, shown as box plots, conditional on the value of $r = \sum_{i=1}^{m} \delta_i^2 / (\sum_{i=1}^{m} \delta_i^2 + (\lambda-1)(\nu-1)\sigma^2)$. Each rectangular box extends from the first to the third quartiles of the distribution of m^2 for that value of r. The middle horizontal line is the median; the pair of horizontal lines above the box represent the 95th and 97.5th percentiles and the pair of horizontal lines below the box represent the 5th and 2.5th percentiles. The (·) represents the extreme values in the sample of 2000.

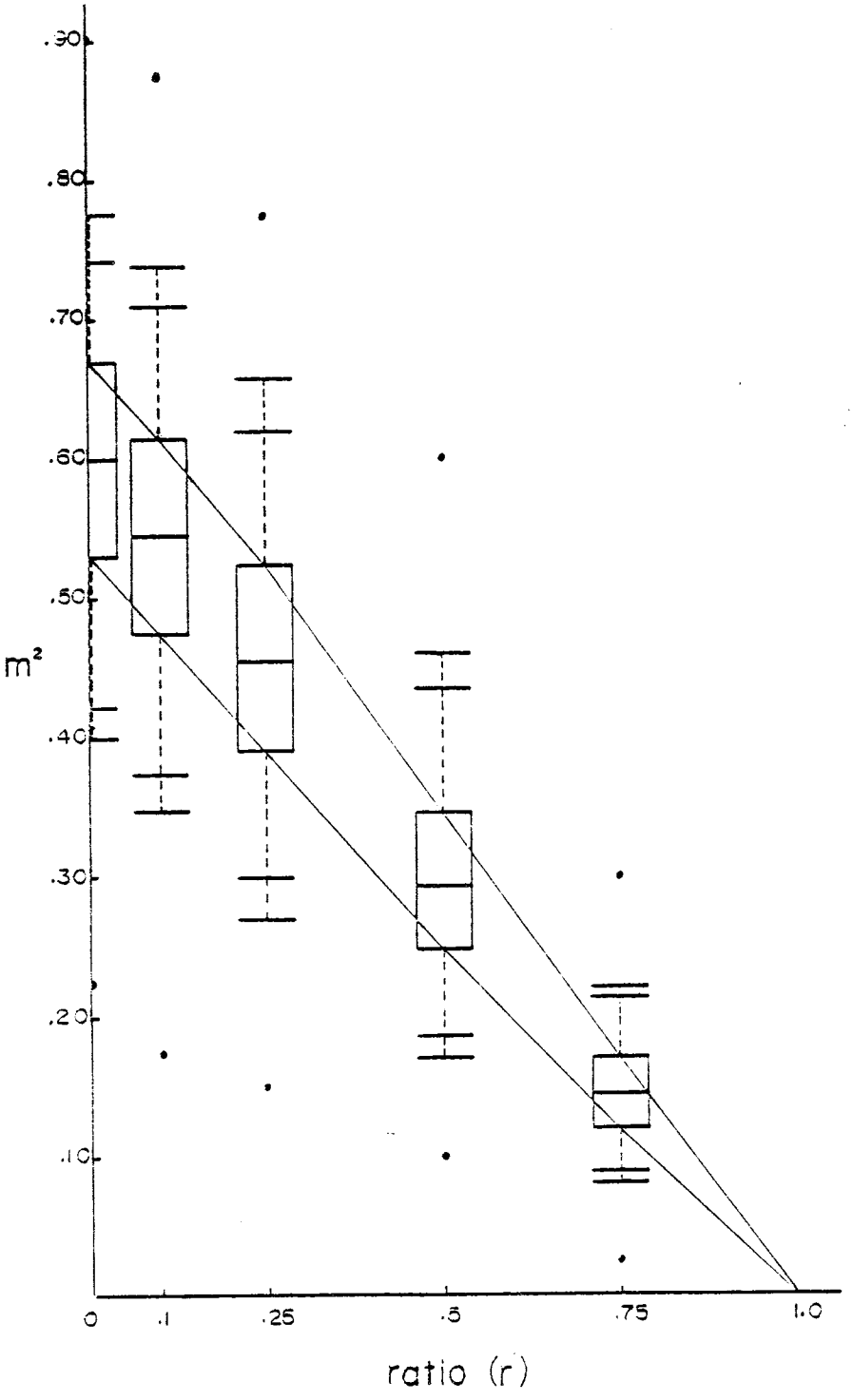


Figure 4.4a. $v = 8$, $\ell = 8$ and AA' has three equal non-zero roots, $\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$.

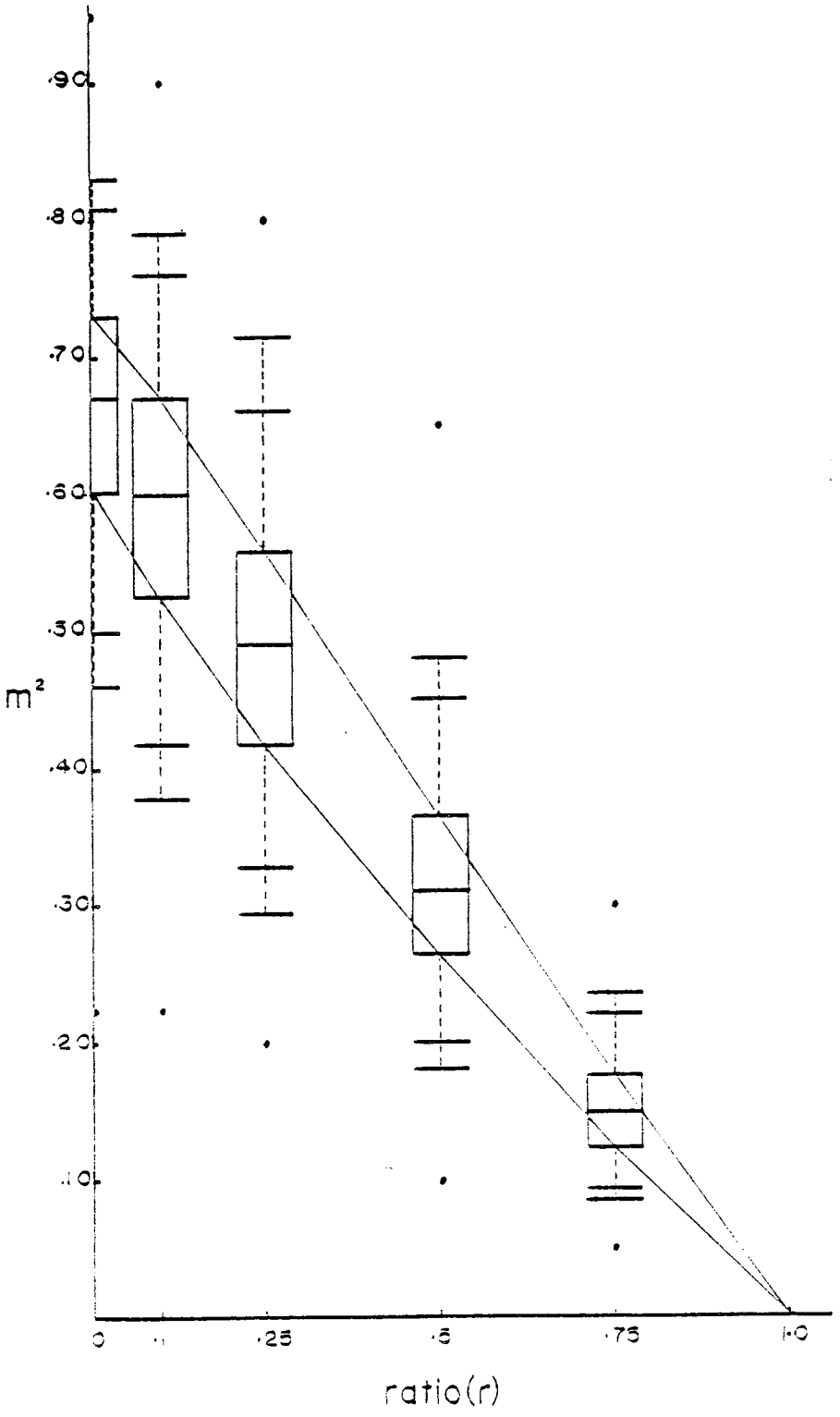


Figure 4.4b. $v = 8$, $\ell = 3$ and AA' has two large equal roots and one small root, $\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$.

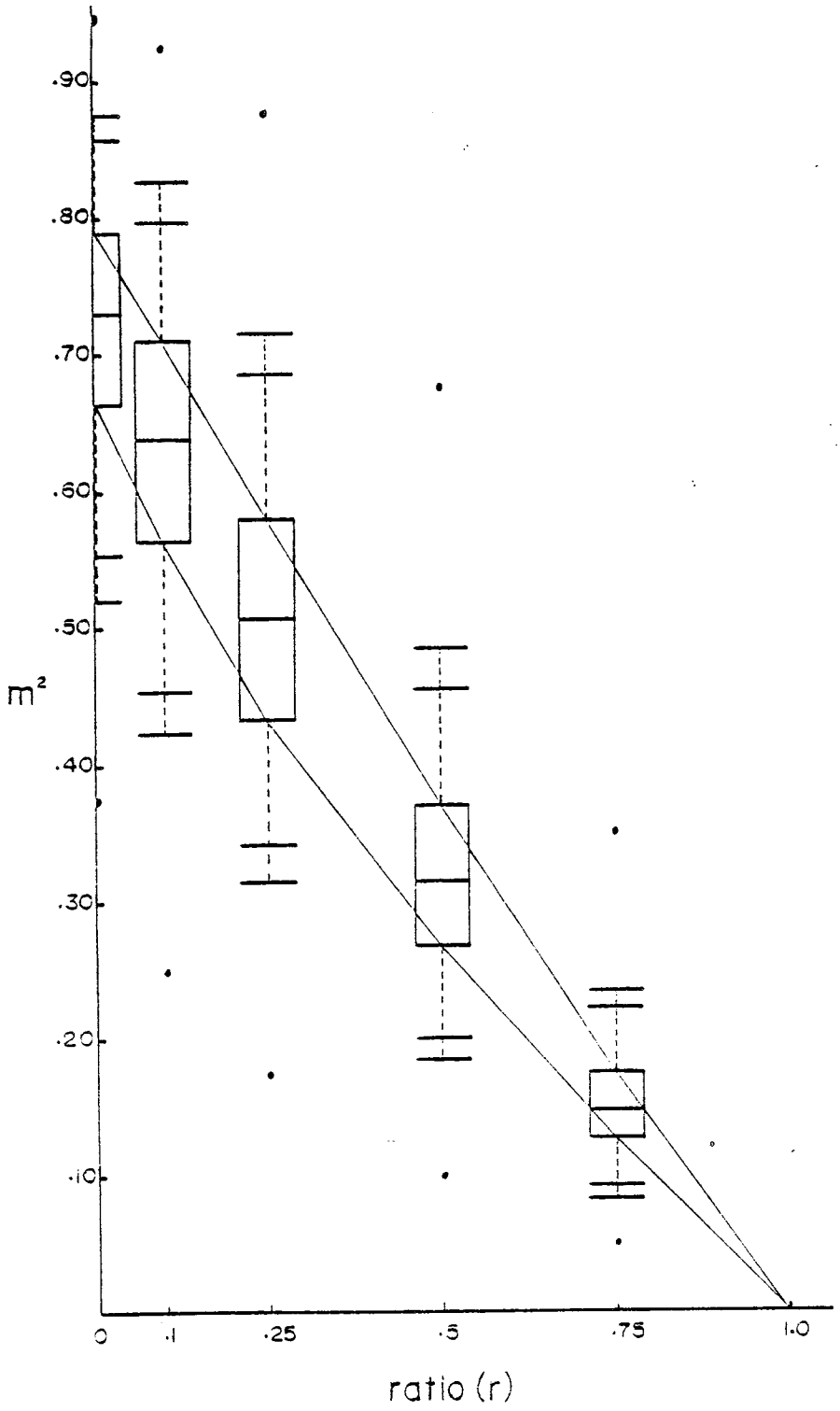


Figure 4.4c. $v = 8$, $\ell = 8$ and AA' has one large and two small roots, $\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$.

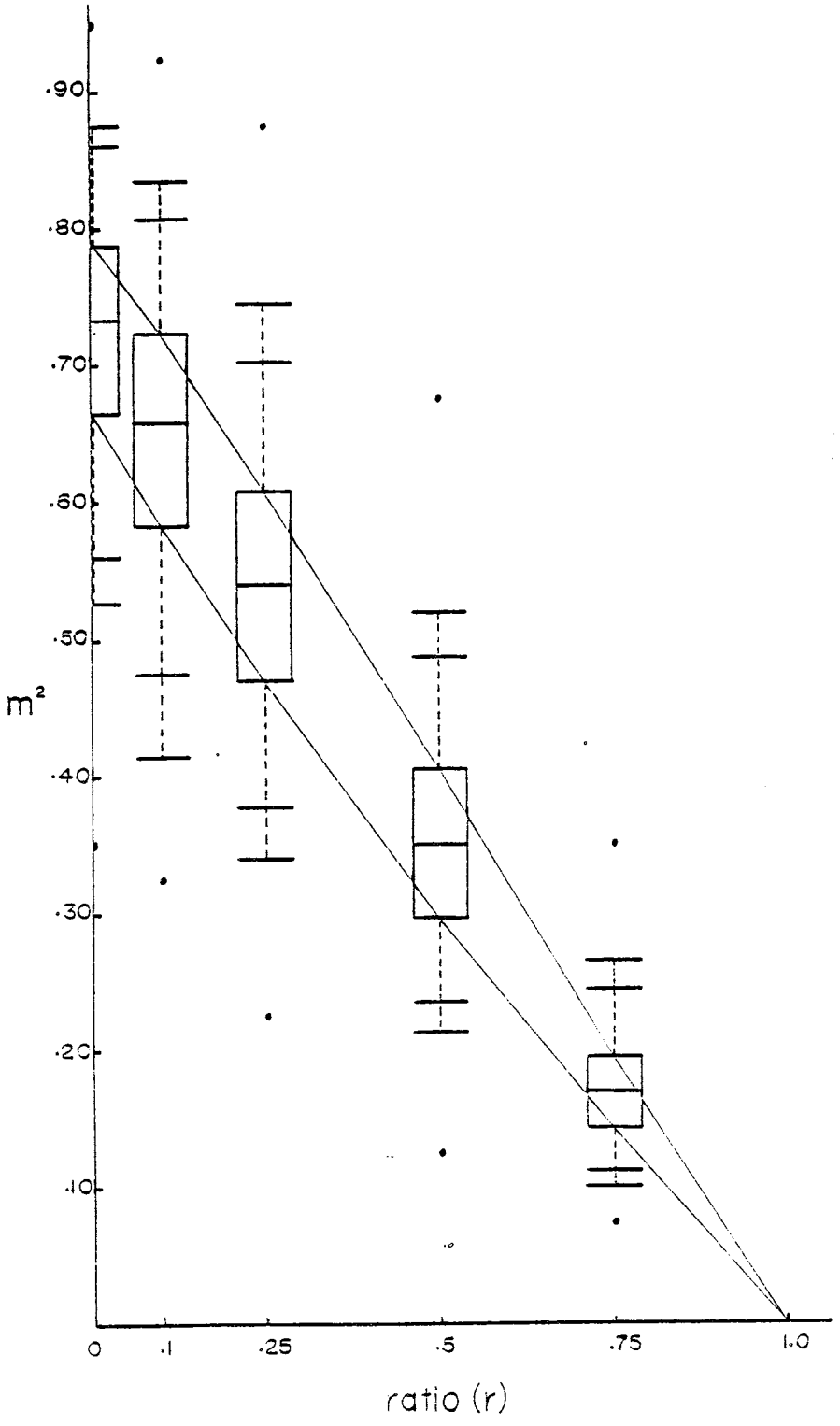


Figure 4.4d. $v = 8$, $\ell = 8$ and AA' has two almost equal non-zero roots, $9\delta_1^2 = 16\delta_2^2 > 0$.

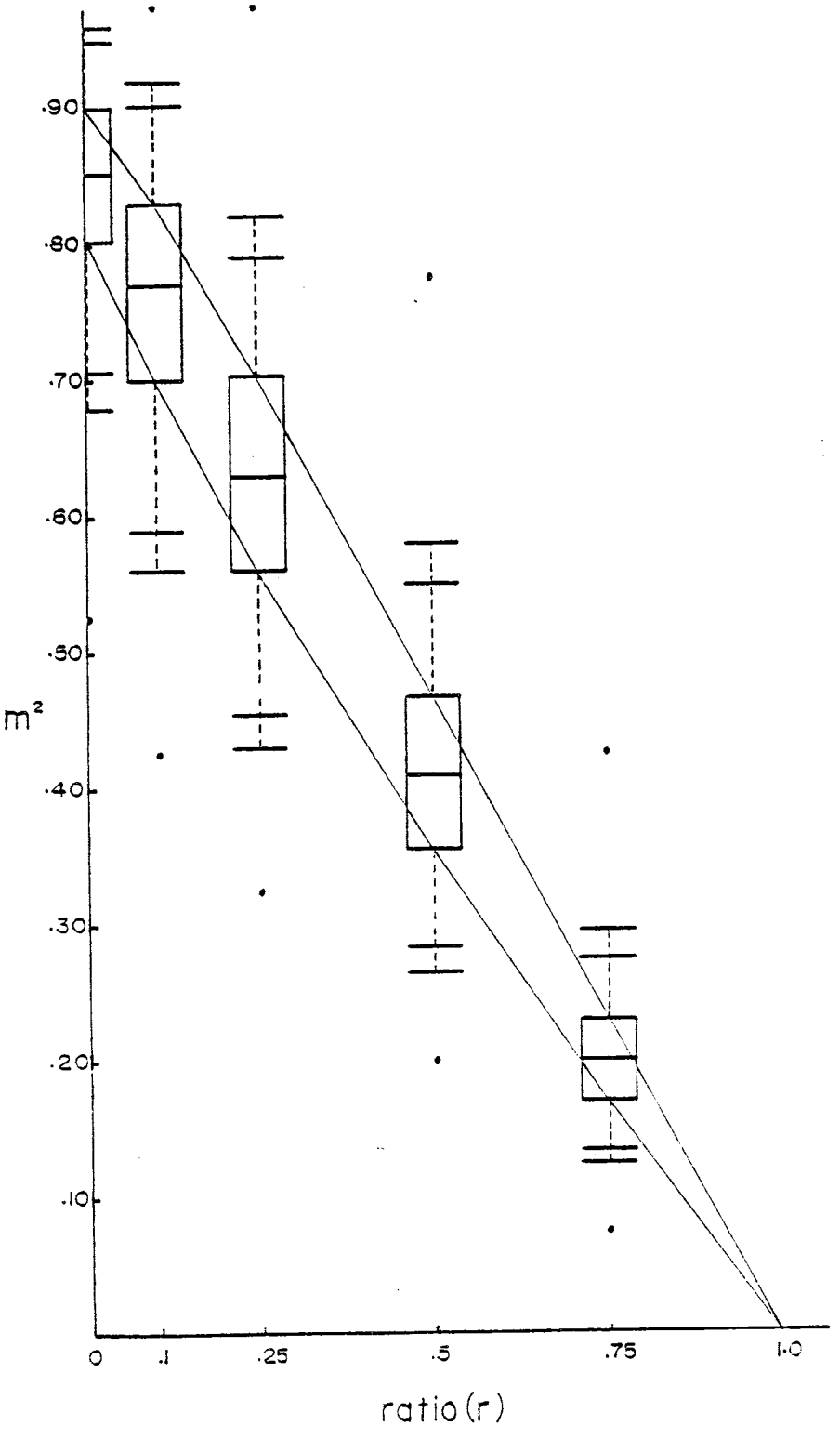


Figure 4.4e. $v = 8$, $\ell = 8$ and AA' has only one non-zero root, $\delta_1^2 > 0$.

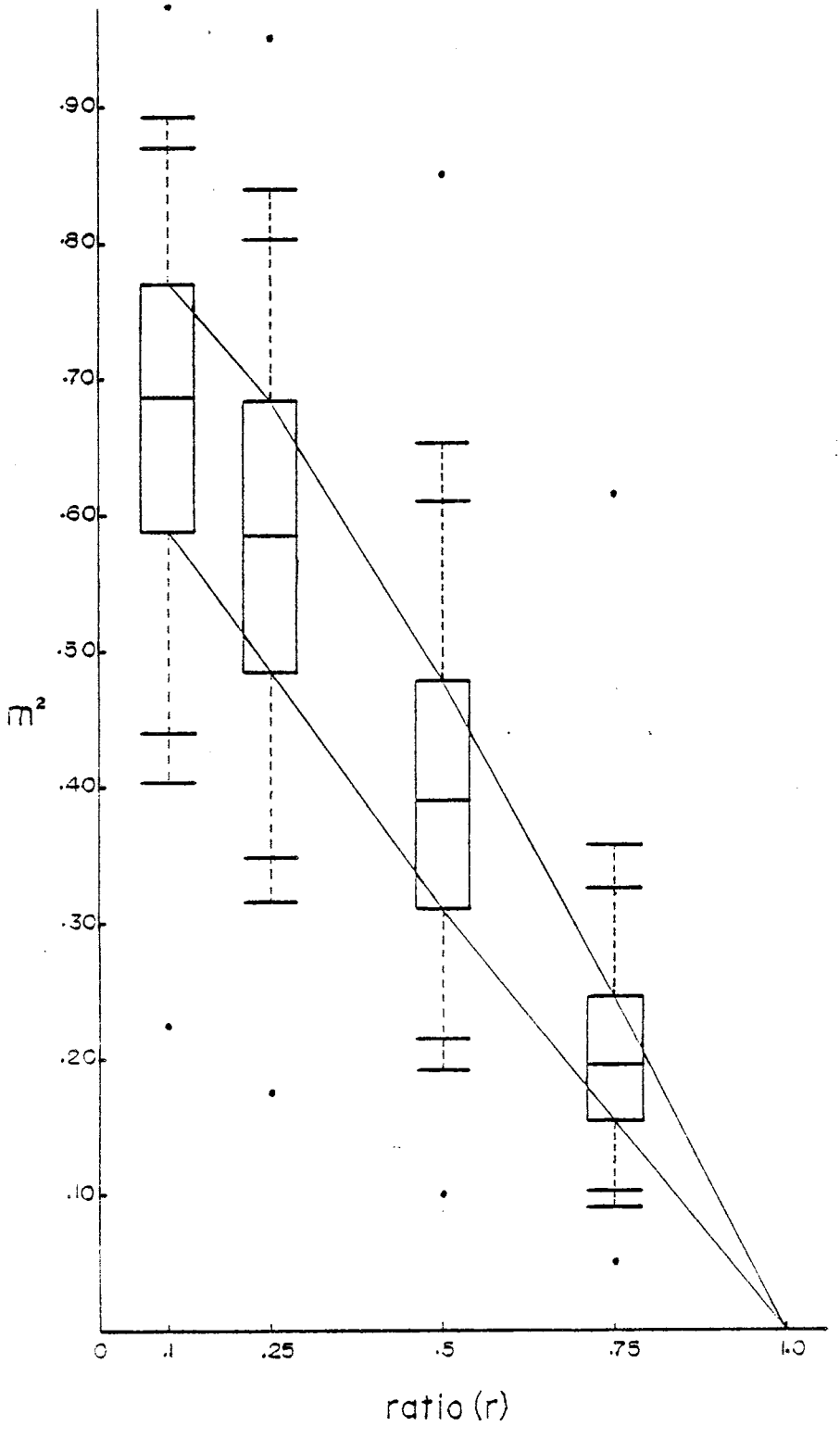


Figure 4.4f. $v = 8$, $\ell = 4$ and AA' has three equal nonzero roots, $\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$.

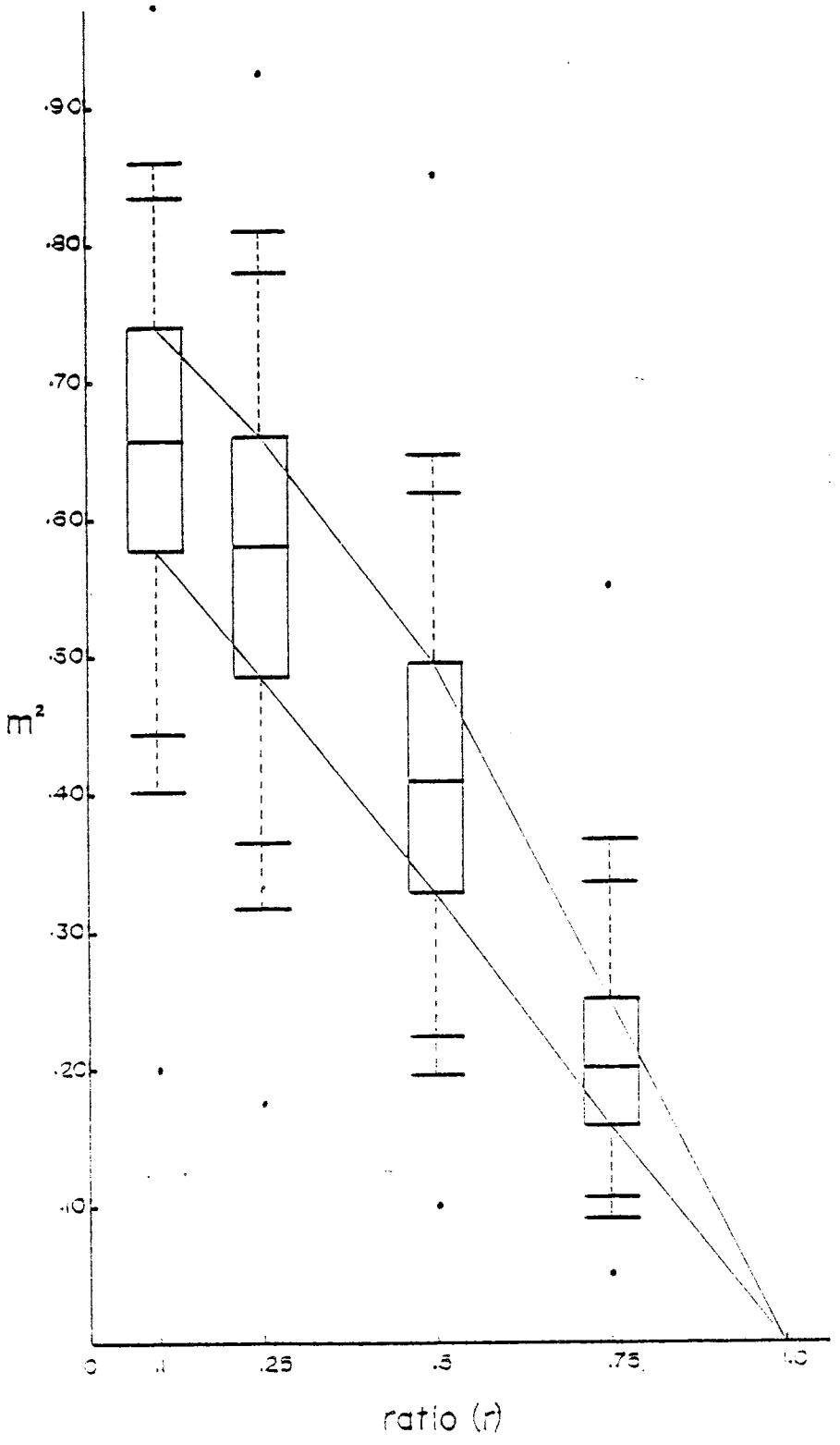


Figure 4.4g. $v = 8$, $\ell = 4$ and AA'_2 has two large equal roots and one small, $\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$.

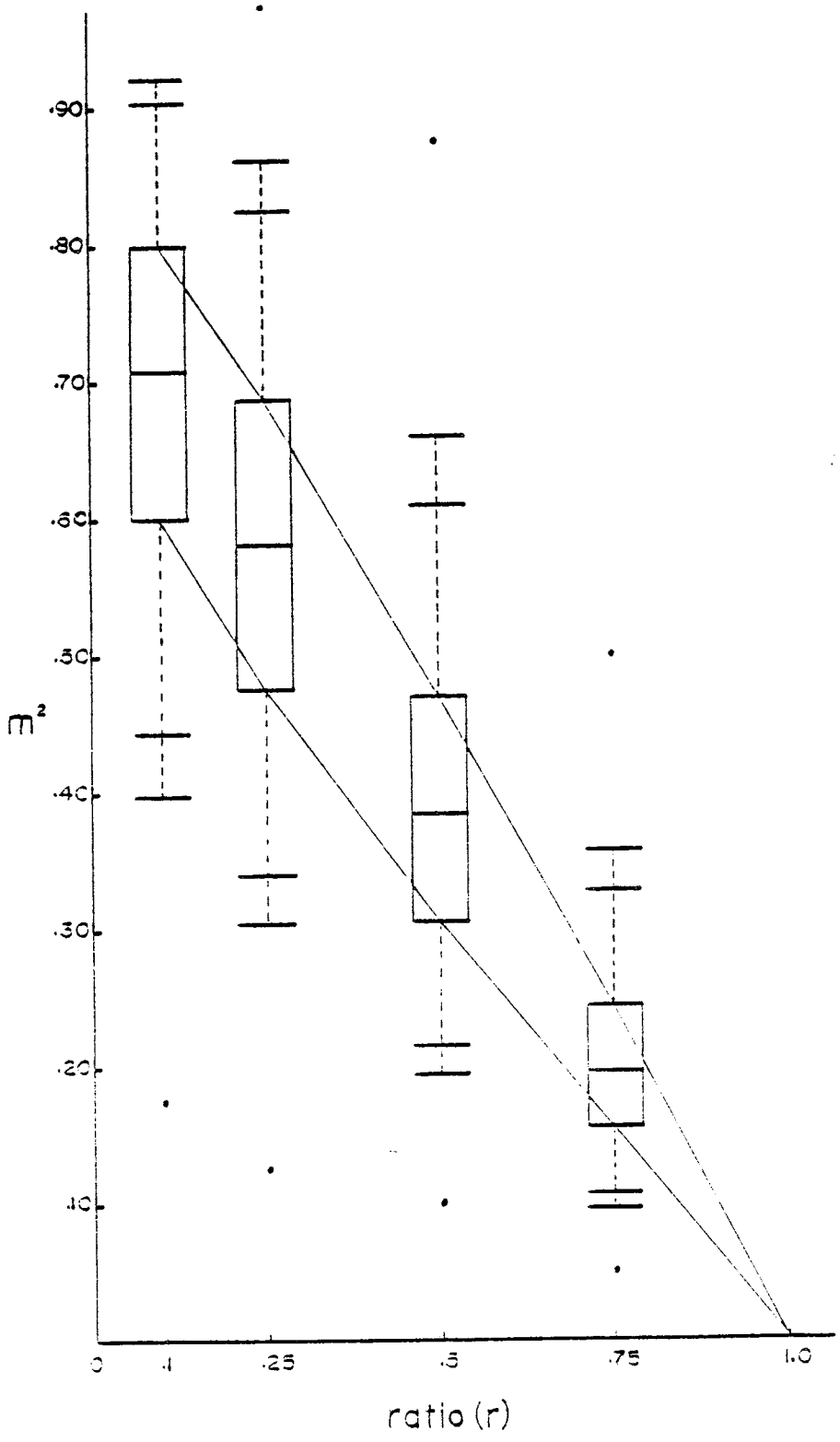


Figure 4.4h. $v = 8$, $\ell = 4$ and AA' has one large and two small roots, $\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$.

4.2 Discussion

With the characterization of the distribution of m^2 given in section 4.1, it is now possible to determine the effect of the various parameters, within the limits of the parameter values used, on the degree to which the configuration of the true points is reproduced by the configuration of the observed points.

When the number of locations was decreased the mean of m^2 is changed very little (Table 4.1), but its variance increased (Table 4.3). When the number of cultivars was decreased, $\bar{m}$ decreased, but the increase in its variance was twice as big as the increase in $V(m^2)$ when ℓ was changed by the same amount. Therefore, the number of cultivars used affects more the degree to which the true location points are reproduced by the observed location points, than does the number of locations.

From the box plots presented in Figure 4.4 a-h and from Tables 4.1 and 4.3, the distribution of m^2 is seen to depend on the A matrix, that is, the repeatability of the location points is seen to depend on the configuration of the true points. As the dimension of the space in which the true location points lie is increased (ℓ and v fixed), the degree of agreement between both sets of points increases.

Although it is clear from Figure 4.4 a-h that the distribution of m^2 depends on the configuration of the true location points, it is also clear that as r approaches one, the distributions become very similar. Also, as r approaches one, both the mean and variance of m^2 decrease; hence, the reproducibility of the true location points by the estimated location points is improved. Therefore, knowledge on the magnitude of

r in a particular trial, used in conjunction with the conditional distributions of m^2, would provide a basis for forming a judgment on how "successful" any classification would be. Two approaches might be taken. Recalling that $\sigma^2 = \sigma_\epsilon^2/n$, prior information, if available, on the magnitude of σ_ϵ^2 and the magnitude of the interaction effects can be used in the designing of the experiment to make the ratio, r, as close as desired to one (by choice of n). In the absence of prior information, or the lack of opportunity to choose n, the experimental results themselves provide information on the magnitude of r.

Information on the magnitude of r in any one experiment can be obtained from the f-ratio,

$$f = \frac{SS(\alpha\beta)/(\ell-1)(v-1)}{SS(ERROR)/\eta}$$

from the analysis of variance where η is the degrees of freedom for error. This f is distributed as a non-central F with $(\ell-1)(v-1)$ and η degrees of freedom and non-centrality parameter $\lambda = \frac{1}{2}\sum_{i=1}^{m} \delta_i^2/\sigma^2$. From Searle (1971),

$$\frac{(v-1)(\ell-1)f}{(\ell-1)(v-1) + 2\lambda}$$

is approximately distributed as a central F with $S = [(\ell-1)(v-1) + 2\lambda]^2/[(\ell-1)(v-1) + 4\lambda]$ and η degrees of freedom. The non-centrality parameter can be expressed as a function of r as $\lambda = \frac{1}{2}r(\ell-1)(v-1)/(1-r)$ so that $f(1-r)$ is approximately distributed as a central F with $S = (\ell-1)(v-1)/(1-r^2)$ and η degrees of freedom. Using this approximation, the null hypothesis H_o: $r = r_o$ against

H_1: $r > r_o$ can be tested or alternatively, an approximate confidence interval estimator for r would be

$$P\{F_{\alpha_1} < (1-r)f < F_{1-\alpha_2}\} = P\left\{1 - \frac{F_{1-\alpha_2}}{f} < r < 1 - \frac{F_{\alpha_1}}{f}\right\} = 1 - \alpha_1 - \alpha_2$$

where α_1 and α_2 are the lower and upper tail confidence coefficients, respectively. Now, this confidence interval gives information on r, which in turn gives information on $E(m^2)$; that is, on the degree to which the estimated points reproduce the true points. Thus, a decision rule as to whether or not one should attempt classification of locations according to the cultivar by location interaction could be based on the magnitude of the f-ratio of the cultivar by location interaction mean square to the error mean square once a decision is made as to what constitutes an acceptable value of m^2.

In the absence of direct information on the probability of "correct" classification conditional on m^2, which in turn would be dependent on the clustering algorithm used and on a somewhat arbitrary definition of what constitutes a "correct" classification, one must resort to geometrical considerations in forming a judgment on the desired magnitude of m^2. The ℓ true location points are distributed in a $m \leq (v-1)$ dimensional space with the average squared distance to their centroid being $\sum_{i=1}^{m} \delta_i^2/\ell$. The estimated location points are distributed about the same centroid with the average squared distance of a given estimated point to its corresponding true point being $m^2 \sum_{i=1}^{m} \delta_i^2/\ell$. Thus, $m^2 = .25$, say, implies that the average squared distance of an observed point to its true point is one-fourth the average squared distance of the

true points to their centroid. Letting the term "standard distance" mean the square root of the average squared distance, and scaling so that standard distance to the centroid is the unit of measure, the standard distance between an observed and its true point is $m = \sqrt{m^2}$ and the maximum diameter of the configuration of true points is 2. Thus, $m^2 = .25$, or $m = .5$, would be sufficiently small to clearly differentiate the more dominant axes in the location space with a high degree of success. Classification applied to situations where m^2 is of the order 0.25 or less would be expected to be successful and repeatable for the dominant dimensions.

Joint consideration of the mean and variance of m^2 for the Monte Carlo studies when $v = \ell = 8$ shows $m^2 \leq .5$ with approximate probability .95 if $r = .5$ and $m^2 \leq .25$ with approximate probability .95 if $r = .75$. When $v = 8$, $\ell = 4$ the corresponding results were $m^2 \leq .61$ with approximate probability 0.95 if $r = 0.5$ and $m^2 \leq .33$ with approximate probability .95 if $r = 0.75$. Thus, it would appear that requiring $r \geq .75$ for $\ell = 8$ and $r \geq .8$ (linear interpolation from the box plots) for $\ell = 4$ would suffice.

The minimum value of r needed so that the classification can be expected to be successful ($m^2 \leq .25$) for the cases of $\ell = 8$ and $\ell = 4$ were .75 and .80, respectively. When the number of locations is between 4 and 8, linear interpolation can be used to determine r, or the value of r for $\ell = 4$ can be used to be conservative. When ℓ is greater than 8, it is clear that the minimum value of r needed will be smaller than .75, but from this study it is not possible to determine how much smaller.

The one-sided confidence interval of r is given by $P(r > 1 - \frac{1}{f} F_\alpha) = 1 - \alpha$ where F_α is the critical value of the F-distribution with $s = (\ell-1)(v-1)/(1-r^2)$ and n degrees of freedom. Since s depends on r, a point estimate of r must be used to be able to determine F_α. A biased point estimate of r, $\hat{r}$, is given by $1 - \frac{1}{f}$. It should be noted that the effect of the adjustment for r in in s is to increase the apparent degrees of freedom for the numerator. Consequently, conservative results are obtained by using $(\ell-1)(v-1)$ for s.

5. EFFICACY OF USING REDUCED DIMENSIONALITY TO CHARACTERIZE SPATIAL CONFIGURATIONS

5.1 Results

The m^2 statistic as defined in the previous sections used the full dimensionality (k) of the estimation space with no consideration given to the relative contributions of the different dimensions to the estimated interaction sum of squares. It is conceivable that reproducibility of the true location configuration might be improved if some of the minor dimensions of the estimated space which, presumably represent only random error, are ignored. This is equivalent to considering, for classification purposes, only the significant components of the interaction matrix as detected by, for example, Mandel's (1971) interaction analysis.

To this end, $m^2(p)$ was defined as the standardized sum of squared distances between the estimated location points in the reduced space (of p dimensions, p = 1, 2, 3, and k), and the true location points. Any improvement in reproducing the configuration of true location points will be reflected in a decrease in $\bar{m}(p)$, compared to $\bar{m}$, with a concomitant decrease, or at least not a drastic increase, in the variance of $m^2(p)$. Such a result would suggest that greater precision in classification of locations on the basis of interactions would be realized by using in the classification only the "dominant" dimensions of the space of location points, which are represented by the columns of the interaction matrix. If there is shown to be some advantage in using reduced dimensionality, a second consideration is to define a

reasonably consistent method of determining the "appropriate" dimensionality in any one case.

The mean values of $m^2(p)$ (Table 5.1), showed a clear tendency for the minimum to occur when p corresponds to the number of dominant eigenvalues of AA'. Thus, for example, when AA' had three equal non-zero roots, the minimum $\bar{m}(p)$ (for those values of p studied) occurs at $p = 3$ for both $r = 0.5$ and 0.75. In a few cases where this pattern does not seem to hold it could be argued that the error variance is small enough that the secondary, but non-zero, roots carry some information; for example, when $r = 0.75$ and AA' had one large and two small roots, and when AA' had two almost equal roots, the minimum $\bar{m}(p)$ occurs at p equal to the total number of non-zero roots. The potential gain, in terms of reduced m^2, ranges from trivial, when $\ell = v = 8$, $r = 0.5$ and AA' has three equal non-zero roots, to approximately a 75% decrease in $\bar{m}(p)$ when AA' has only one non-zero eigenvalue. It should also be noted that the potential loss in reducing dimensionality too much is much greater than any potential gain; compare $\bar{m}(1)$ with $\bar{m}(k)$.

The variances of m^2 did not appear to have any distinct pattern of change, when the dimensionality was changed (Table 5.2). A fairly definite pattern did emerge, however, when the mean and the variances were combined by using the value of the 95^{th} percentile (Table 5.3). For the three cases studied with $\ell = 4$, there was no advantage to be gained in reducing the dimensionality below three. When $\ell = 8$, however, there were distinct advantages of reduced dimensionality in several cases. These results suggest that knowledge of the relative

Table 5.1. Mean values of m^2 calculated using the first p eigenvalues of the estimated interaction covariance matrix, $\hat{A}\hat{A}'$, for the different interaction matrices A, and different ratios r.

v,ℓ	Relative sizes of the non-zero roots of AA′	r	p = 1	p = 2	p = 3	p = k
8,8	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.5	.722	.476	.293	.300
		.75	.688	.388	.107	.145
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.5	.602	.274	.301	.315
		.75	.547	.112	.131	.150
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.5	.247	.304	.317	.322
		.75	.155	.151	.141	.151
	$9\delta_1^2 = 16\delta_2^2 > 0$	.5	.521	.244	.320	.354
		.75	.424	.080	.144	.172
	$\delta_1^2 > 0$	.5	.127	.308	.380	.413
		.75	.042	.134	.179	.202
8,4	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.5	.752	.545	.413	
		.75	.708	.437	.209	
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.5	.658	.414	.401	
		.75	.579	.199	.204	
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.5	.377	.397	.393	
		.75	.255	.215	.203	

Table 5.2. Variances of m^2 calculated using the first p eigenvalues of the estimated interaction covariance matrix, $\hat{A}\hat{A}'$, for the different interaction matrices A and different ratios r.

v,ℓ	Relative sizes of the non-zero roots of AA′	r	p = 1	p = 2	p = 3	p = k
8,8	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.5	.00146	.00470	.01293	.00550
		.75	.00029	.00061	.00147	.00140
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.5	.00299	.01352	.00795	.00581
		.75	.00039	.00117	.00173	.00151
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.5	.00707	.00674	.00737	.00600
		.75	.00063	.00122	.00214	.00152
	$9\delta_1^2 = 16\delta_2^2$	.5	.00898	.01321	.00737	.00607
		.75	.00257	.00110	.00144	.00166
	$\delta_1^2 > 0$	.5	.00591	.00639	.00675	.00670
		.75	.00057	.00121	.00168	.00190
8,4	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.5	.00292	.00774	.01429	
		.75	.00080	.00209	.00507	
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.5	.00647	.02196	.01442	
		.75	.00146	.00532	.00467	
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.5	.02786	.01547	.01441	
		.75	.00328	.00406	.00446	

Table 5.3. Values of the 95th percentile of the distribution of $m^2(p)$ $p = 1,2,3$, and k, for the different A matrices and different ratios r.

v, ℓ	Relative sizes of the non-zero roots of AA'	r	p = 1	p = 2	p = 3	p = k
8,8	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.5	.798	.604	.514	.435
		.75	.717	.433	.176	.214
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.5	.705	.507	.463	.450
		.75	.585	.175	.207	.220
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.5	.402	.447	.466	.454
		.75	.203	.215	.225	.222
	$9\delta_1^2 = 16\delta_2^2$	.5	.699	.477	.474	.488
		.75	.530	.142	.212	.244
	$\delta_1^2 > 0$	.5	.273	.447	.523	.550
		.75	.087	.195	.253	.279
8,4	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.5	.859	.708	.620	
		.75	.764	.518	.336	
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.5	.812	.681	.611	
		.75	.652	.330	.325	
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.5	.709	.624	.608	
		.75	.310	.331	.327	

magnitudes of the roots of AA' would provide a basis for determining the appropriate dimensionality.

The dimensionalities suggested by the above were compared to those obtained by applying Mandel's (1971) partitioning of the interaction sum of squares, $SS(\alpha\beta)$, for the different cases of A matrices. The results obtained applying Mandel's (1971) partitioning to the average results are given in Table 5.4. The numbers in the table are what Mandel calls the "mean squares"for the partitioned interaction effects divided by the error variance, namely $d_i^2/\nu_i\sigma^2$, where ν_i, Mandel's estimate of "degrees of freedom", is the ith eigenvalue of $\hat{A}\hat{A}'$ when $A = 0$ and $\sigma^2 = 1$ (these values of ν_i can be obtained from a table in Mandel (1971)). Thus, the values should be approximately 1.0 if that root is reflecting only random error. The vertical lines in Table 5.4 mark the partitioning which was done; only the eigenvalues to the left of the lines will be used. To determine the cut-off point, the critical values of the F-distribution ($\alpha = 0.05$) with ν_i and η degrees of freedom was used. Since the error degrees of freedom were not known η was taken to be $2(v-1)$, corresponding to $n = 2$ in a randomized complete block design. Although $d_i^2/\nu_i\sigma^2$ is not distributed as an F, the partition obtained was acceptable. Only the results for the cases in which r is greater than or equal to .5 were presented, because, as was reported in section 4.2, there would be little interest in doing a classification of environments unless interactions were relatively important.

Table 5.4. Values of $d_i^2/(\sigma^2\nu_i)$ for the different interaction matrices A and different ratios r.

v,ℓ	Relative sizes of the non-zero roots of AA′	r	i						
			1	2	3	4	5	6	7
8,8	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.5	2.031	2.102	2.056	1.769	1.597	1.614	0.135
		.75	3.968	4.591	5.023	2.140	1.835	1.808	1.673
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.5	2.230	2.177	1.628	1.496	1.425	1.434	1.335
		.75	4.763	5.334	2.011	1.743	1.610	1.614	1.507
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.5	2.764	1.510	1.396	1.356	1.300	1.343	1.269
		.75	6.896	2.003	1.806	1.651	1.533	1.580	1.425
	$9\delta_1^2 = 16\delta_2^2 > 0$	.5	2.340	2.091	1.532	1.414			
		.75	5.708	5.172	1.803	1.626			
	$\delta_1^2 > 0$	.5	3.001	1.312					
		.75	2.553	0.443					

Table 5.4. Continued.

v, ℓ	Relative sizes of the non-zero roots of AA′	r	i						
			1	2	3	4	5	6	7
8,4	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.5	1.868	2.081	2.343				
		.75	3.360	4.365	6.158				
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.5	1.981	2.071	1.824				
		.75	3.850	4.709	2.677				
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2$	.5	2.261	1.604	1.528				
		.75	5.139	2.212	2.170				

5.2 Discussion

The mean, together with the variance of $m^2(p)$, and the value of the 95th percentile of the distribution of $m^2(p)$ were used to determine the "best" p, and the gain in precision to be realized by reducing to p the number of eigenvalues of $\hat{A}\hat{A}'$ used. When $\ell = 8$ and AA' had only one non-zero root, it was advantageous to use only the first root of $\hat{A}\hat{A}'$ because both the mean and the variance of $m^2(1)$ were the smallest. By going to one dimension in such cases, $\bar{m}(1)$ was reduced to about 1/3 to 1/5, depending on r, of the original $\bar{m}$. The variance was reduced only slightly when $r = 0.5$, but by 2/3 when $r = 0.75$. Therefore, precision is gained by using only the first root of $\hat{A}\hat{A}'$, in this case. When AA' had two almost equal eigenvalues and $r = 0.75$, there is a similar result; both the mean and variance of $m^2(2)$ are the smallest, $2\bar{m}(2) \doteq \bar{m}$.

Comparing the values of the 95th percentile of the distribution of $m^2(p)$, $p = 1, 2, 3$, and k, the "best" p coincided, in general, with the number of dominant roots of $\hat{A}\hat{A}'$. Therefore, in general, the reproduction of the spatial configuration is improved by reducing the dimension of the estimated location space to the number of "dominant" dimensions of the true space when $\ell = 8$. When $\ell = 4$ there seems to be no advantage in reducing the dimension of the space.

Mandel's (1971) approach gave a partitioning of the interaction effects which was similar to the above. While Mandel's (1971) partitioning was done on the average of 2000 runs, not on the individual trials, the results suggest that this criterion for partitioning the interaction sum of squares into a significant part, would be useful for gaining precision in the classification of locations by reducing the dimensionality

of the estimated location space. How reliable the procedure is when Mandel (1971) is applied to a single trial is a problem that will need further investigation.

6. MEANS AND VARIANCES OF EIGENVALUES ESTIMATED FROM $v \times \ell$ INTERACTION COVARIANCE MATRICES

6.1 Results

Since the eigenvalues were used as the basis for reducing the dimensionality of the estimated location space, the reliability in the estimation of the eigenvalues is basic to the decision as to how much the dimensionality should be reduced. Because of this, the means and variances of the eigenvalues of $\hat{A}\hat{A}'$ for different A matrices were studied.

The means of the eigenvalues of $\hat{A}\hat{A}'$ standardized by $E(\text{tr } \hat{A}\hat{A}')$, S_i^2, estimated from the Monte Carlo study are presented in Table 6.1. The i^{th} eigenvalue (normalized) of all $\hat{A}\hat{A}'$ converges, as the ratio r goes to zero, to a point determined only by the dimensions of the matrix A, v and ℓ. These limiting values of d_i^2 give the sample values of the distribution of the eigenvalues of a $(v-1) \times (v-1)$ Wishart matrix with $(\ell-1)$ degrees of freedom and associated covariance matrix $I\sigma^2$ (Mandel (1970)). Therefore, d_1^2/σ^2 (when $r = 0$) were compared to tables for the largest eigenvalues of the Wishart matrix (Clemm, et al. (1971), Schuurmann, et al. (1973)), and they were found to agree up to two decimals.

The values of S_1^2 for the case of no interaction, $r = 0$, are given in Table 6.2 for several choices of ℓ and v up to $\ell = v = 8$. These values, multiplied by $(\ell-1)(v-1)$, $(\ell-1)(v-1)\, S_1^2 = d_1^2/\sigma^2$, were plotted against $(\ell-2)$ (Figure 6.1). A regression through the origin of $(\ell-1)(v-1)\, S_1^2 - (|v-\ell| + 1)$ on $|v-\ell|(r(\hat{A}) - 1)$, $r(\hat{A}) - 1$, and $(r(\hat{A}) - 1)^2$ gave an approximation of the change in $(\ell-1)(v-1)\, S_1^2$ as

Table 6.1. Mean values of s_i^2 (for those values calculated) for the different interaction matrices A, and different ratios r.

v,ℓ	Relative sizes of the non-zero roots of AA′	r	s_1^2	s_2^2	s_3^2	s_4^2	s_5^2	s_6^2	s_7^2	Σs_1^2
8,8	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.001	.4273	.2615	.1614	.0913				
		.1	.4287	.2619	.1615	.0911				
		.25	.4311	.2633	.1611	.0897	.04273	.01466	.00209	1.0048
		.5	.4337	.2748	.1660	.0807	.03558	.01211	.00165	1.0045
		.75	.4237	.3001	.2027	.0488	.02045	.00678	.00093	1.0034
		1[a]	.3333	.3333	.3333	0	0	0	0	1.0000
	$\delta_1^2 = \delta_2^2 = 1.6\delta_3^2 > 0$	.001	.4272	.2616	.1614	.09123				
		.1	.4303	.2616	.1608	.09071				
		.25	.4425	.2626	.1561	.08613	.04110	.01397	.00196	1.0044
		.5	.4718	.2846	.1310	.06823	.03175	.01076	.00148	1.0040
		.75	.5085	.3487	.0812	.03975	.01794	.00606	.00083	1.0030
		1[a]	.4848	.4848	.0303	0	0	0	0	1.0000
	$\delta_1^2 = 1.6\delta_2^2 = 1.6\delta_3^2 > 0$	.001	.4271	.2616	.1614	.09125				
		.1	.4336	.2607	.1592	.08972				
		.25	.4711	.2491	.1470	.08186	.03885	.01338	.00191	1.0033
		.5	.5902	.1975	.1127	.06184	.02898	.01008	.00140	1.0027
		.75	.7363	.1310	.0729	.03766	.01708	.00593	.00079	1.0015
		1[a]	.8889	.05556	.0556	0	0	0	0	1.0000

Table 6.1. Continued.

v,ℓ	Relative sizes of the non-zero roots of AA′	r	s_1^2	s_2^2	s_3^2	s_4^2	s_5^2	s_6^2	s_7^2	Σs_1^2
8,8	$9\delta_1^2 = 16\delta_2^2 > 0$	.001	.4271	.2616	.1614	.0912				
		.1	.4308	.2617	.1603	.0906				
		.25	.4478	.2612	.1537	.0850				
		.5	.4997	.2734	.1237	.0645				
		.75	.5649	.3138	.0674	.0344				
		1[a]	.64	.36	0	0	0	0	0	1.0000
	$\delta_1^2 > 0$	.001	.4270	.2616						
		.1	.4360	.2601						
		.25	.4850							
		.5	.6410	.1716						
		.75	.8179	.0871						
		1[a]	1	0						
8,4	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.1	.6055	.2903	.1108					1.0066
		.25	.5981	.2921	.1137					1.0039
		.5	.5694	.3023	.1290					1.0007
		.75	.5121	.3171	.1695					.9987
		1[a]	.3333	.3333	.3333					1.0000
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.1	.6090	.2903	.1100					1.0092
		.25	.6082	.2906	.1089					1.0077
		.5	.6038	.3009	.1004					1.0047
		.75	.5867	.3421	.0737					1.0025
		1[a]	.4848	.4848	.0303					1.0000

Table 6.1. Continued.

v,ℓ	Relative sizes of the non-zero roots of AA'	r	s_1^2	s_2^2	s_3^2	s_4^2	s_5^2	s_6^2	s_7^2	Σs_i^2
8,4	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.1	.6119	.2889	.1092					1.0101
		.25	.6289	.2767	.1033					1.0089
		.50	.6893	.2330	.0841					1.0065
		.75	.7832	.1607	.0597					1.0032
		1[a]	.8889	.0556	.0556					1.0000

[a]Exact results.

Figure 6.1. The largest eigenvalue of ${}_{v}\hat{A}_{\ell}\,\hat{A}_{v}'$ (standardized by σ^2), $d_1^2/\sigma^2 = (\ell-1)(v-1)S_1^2$ against $\ell - 2$ for the case where there is no interaction $A = 0$, and for different choices of v.

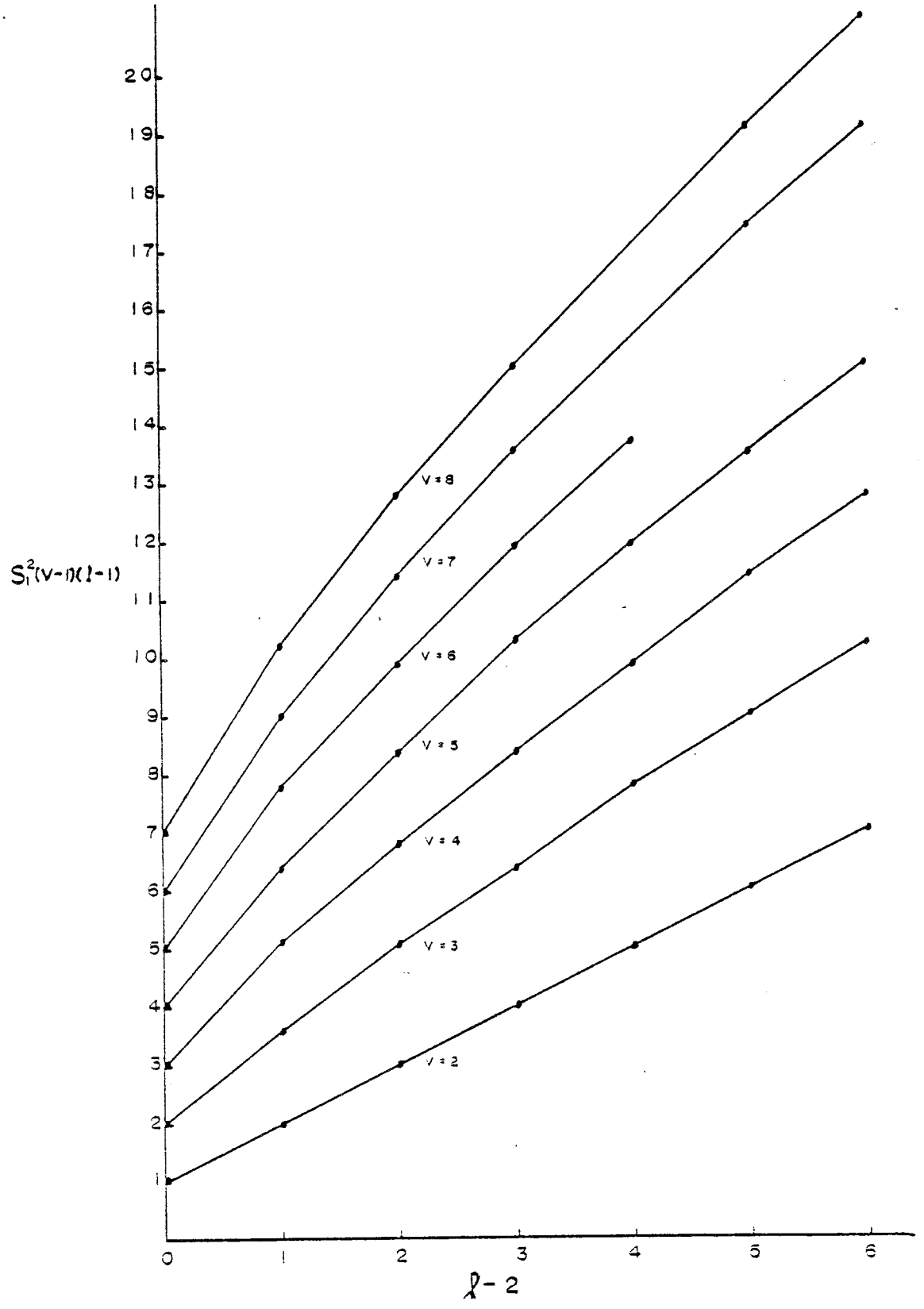

$S_1^2(v-1)(l-1)$
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
V = 8
V = 7
V = 6
V = 5
V = 4
V = 3
V = 2
0
1
2
3
4
5
6
$l-2$

Figure 6.2. The largest eigenvalue of $\hat{A}\hat{A}'$ (standardized), s_1^2, against the ratio, r, for each of the 8 x 8 and 8 x 4 interaction matrices A.

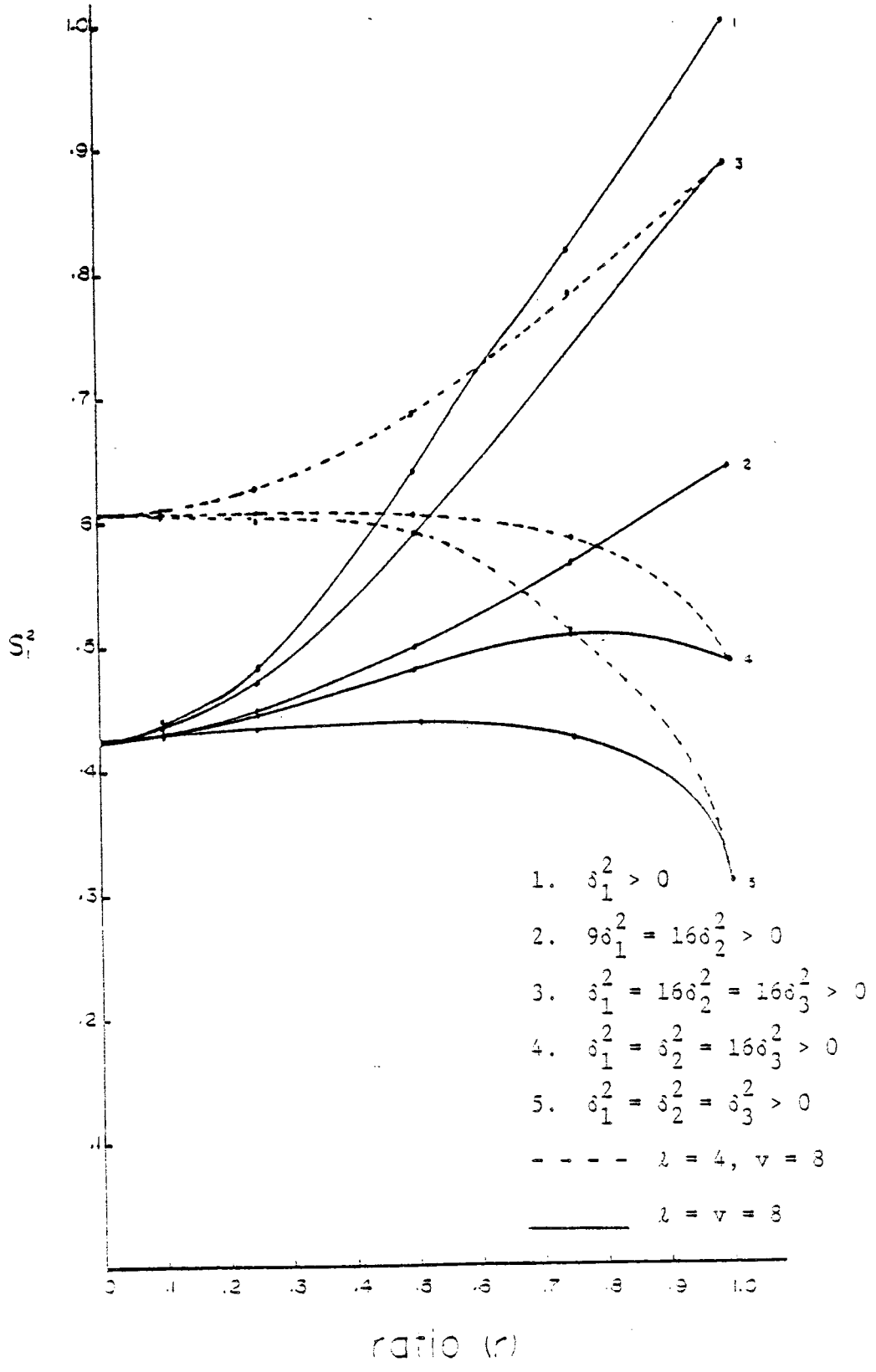

1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
S_1^2
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
ratio (r)
1
2
3
4
5
1. $\delta_1^2 > 0$
2. $9\delta_1^2 = 16\delta_2^2 > 0$
3. $\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$
4. $\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$
5. $\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$
- - - - $\ell = 4$, $v = 8$
——— $\ell = v = 8$

Table 6.4. Variances of $S_{\bar{i}}^2$ (for those values calculated) for the different interaction matrices A, and for different ratios r.

	Relative sizes of the non-zero roots of AA′	r	s_1^2	s_2^2	s_3^2	s_4^2	s_5^2	s_6^2	s_7^2
8,8	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.001	.01190	.004328	.002206	.001004			
		.1	.01194	.004438	.002187	.001048			
		.25	.01165	.004423	.002163	.001001	.000278	.000106	.0000095
		.5	.00947	.004421	.002253	.000819	.000279	.000074	.0000058
		.75	.00512	.002757	.002104	.000359	.000095	.000023	.0000019
		1[a]	0	0	0	0	0	0	0
	$\delta_1^2 = \delta_2^2 = 16\delta_3^2 > 0$	.001	.01192	.004329	.002205	.001004			
		.1	.01228	.004416	.002154	.001030			
		.25	.01296	.004511	.002018	.000928	.000354	.000101	.0000081
		.5	.01188	.005263	.001525	.000620	.000225	.000006	.0000048
		.75	.00712	.004342	.0006854	.000224	.000075	.000019	.0000015
		1[a]	0	0	0	0	0	0	0
	$\delta_1^2 = 16\delta_2^2 = 16\delta_3^2 > 0$	.001	.01192	.004331	.002213	.001005			
		.1	.01318	.004451	.002145	.000993			
		.25	.01742	.004345	.001800	.000834	.000314	.0000859	.0000088
		.5	.02077	.003047	.001118	.000486	.000178	.0000492	.0000048
		.75	.01441	.001320	.000484	.000181	.000624	.0000176	.0000014
		1[a]	0	0	0	0	0	0	0
	$9\delta_1^2 = 16\delta_2^2 > 0$	.001	.01192	.004328	.002206	.001004			
		.1	.01251	.004452	.002159	.001020			
		.25	.01388	.004600	.001975	.000903			
		.5	.01415	.005457	.001413	.000537			
		.75	.01000	.004886	.000477	.000154			
		1[a]	0	0	0	0			

Table 6.4. Continued.

ℓ	Relative sizes of the non-zero roots of AA'	r	s_1^2	s_2^2	s_3^2	s_4^2	s_5^2	s_6^2	s_7^2
8,8		.001	.01193						
		.1	.1366						
		.25							
	$\delta_1^2 > 0$	.5	.02318						
		.75	.01611						
		1[a]	0						
		.1	.04244	.01415	.004273				
		.25	.03741	.01361	.004415				
	$\delta_1^2 = \delta_2^2 = \delta_3^2 > 0$	.5	.02645	.01133	.004740				
		.75	.01300	.00677	.004234				
		1[a]	0	0	0				
		.1	.04435	.01392	.004272				
		.25	.04144	.01354	.004176				
8,4	$\delta_1^2 = \delta_2^2 = 16\delta_3^2$	.5	.03268	.01276	.003506				
		.75	.01798	.01015	.001830				
		1[a]	0	0	0				
		.1	.04490	.01394	.004275				
		.25	.04707	.01266	.003797				
	$\delta_1^2 = 16\delta_2^2 = .16\delta_3^2 > 0$	.5	.04932	.00905	.002497				
		.75	.03452	.00404	.001172				
		1[a]	0	0	0				

Table 6.2. Mean values of S_1^2 in the case where there is no interaction, $A = 0$, for different choices of ℓ and v.

ℓ \ v	1	3	4	5	6	7	8
2	1	1	1	1	1	1	1
3	1	.9098	.8551	.8029	.7802	.7501	.7310
4	1	.8551	.7563	.6981	.6596	.6345	.6096
5	1	.8029	.6981	.6434	.5960	.5649	.5388
6	1	.7802	.6596	.5960	.5474		
7	1	.7501	.6345	.5649		.4822	.4553
8	1	.7310	.6096	.5388		.4553	.4271

ℓ and v changed, $R^2 = .9998$. The regression equation obtained was

$$(\ell-1)(v-1)S_1^2 = d_1^2/\sigma^2 = 2.83(r(\hat{A})-1)\{.081|v-\ell| + .029(r(\hat{A})-1) + 1\}$$

$$+ (|v-\ell| + 1)$$

where $r(\hat{A})$ is the rank of $\hat{A}$.

The limiting results as r approaches one are the standardized values of the parametric eigenvalues of the AA' matrix. This is illustrated in Figure 6.2, where the mean values of S_1^2, for the different A matrices are plotted against the ratio r. Clearly, the largest values of S_1^2 are obtained when $A'A$ has only one non-zero root, and the smallest values are obtained when AA' has $(\ell-1)$ equal non-zero roots, as is expected.

Table 6.3 presents the values of the variances of S_1^2 when $A = 0$ $(r = 0)$. As ℓ and v increased, the variances of S_1^2 were monotonically decreasing, $\ell, v \geq 3$, as did the mean. From a regression

through the origin of $(\ell-1)^2(v-1)^2 v(S_1^2)$ on $[(\ell-2)(v-2)]^{\frac{1}{2}}$ it was obtained that $(\ell-1)^2(v-1)^2 v(S_1^2) = v(d_1^2/\sigma^2) \doteq 5.26[(\ell-2)(v-2)]^{\frac{1}{2}}$, $R^2 = .988$. The variance of S_1^2, for the different cases of A matrices and values of r are presented in Table 6.4. The same effects as in Table 6.3 are seen. As ℓ decreased from 8 to 4 the variances of S_1^2 increased. For the cases where AA' had three equal roots, two large and one small, or one large and two small, the variances of S_1^2 and S_2^2 for $\ell = 4$ varied from 2.4 to 3.6 times as large as for $\ell = 8$. When ℓ and v are fixed, the variances of S_i^2 appear to reach a maximum in the region of r = .1 to r = .25 and then monotonically decrease to zero as r increases to 1.0. It is also clear that S_1^2 always has the largest variance.

Table 6.3. Values of the variance of S_1^2 in the null case, for different choices of ℓ and v.

ℓ \ v	2	3	4	5	6	7	8
2	0	0	0	0	0	0	0
3	0	.4170	.2606	.1692	.1242	.1013	.0859
4	0	.2606	.1441	.0940	.0710	.0576	.0466
5	0	.1692	.0940	.0636	.0452	.0363	.02841
6	0	.1242	.0710	.0452	.0312		
7	0	.1013	.0576	.0363		.0180	.0151
8	0	.0859	.0466	.0284		.0152	.0119

The last column of Table 6.1 shows the sum of the means of the S_i^2, that is, the mean of $\sum_{i=1}^{k} S_i^2$ over the 2000 runs. From formula 3.1 it is known that the expectation of this quantity is one.

$$E\left(\sum_{i=1}^{k} S_i^2\right) = 1 \ .$$

In the cases presented in Table 6.1, the observed means are slightly greater than one, but not significantly so. The fact that almost all are greater than 1.0 reflects the fact that the same random error samples were used for all cases having the same values of ℓ and v.

6.2 Discussion.

It is clear that the means of the eigenvalues of $\hat{A}\hat{A}'$ depend on the relative sizes of the non-zero eigenvalues of AA', where A is the true interaction matrix, and within each case of A matrix, it depends on the error variance σ^2(or r). In the case when r approaches zero, an approximate relation to characterize the mean and variance of d_1^2/σ^2 was obtained. The importance of these limiting values of d_i^2/σ^2 is seen in papers like Mandel (1971), who uses these values to test the hypothesis that $E(d_i^2)/\nu_i = \sigma^2$ where $\nu_i = d_i^2/\sigma^2$ in the null case, and then partitions the interaction according to the result of the test. Only those d_i^2 for which the hypothesis is rejected are kept. When the values of ν_i obtained from the Monte Carlo study were compared to the values of ν_i presented in Mandel (1971), they agreed to one or two decimal digits depending on ℓ and v. As explained in section 5.1, Mandel's partitioning can be used to reduce the dimensionality of the space where the estimated location points lie.

The eigenvalues are used to determine the partitioning of the interaction effects, that is, to reduce the dimensionality of the estimated location space, hence the precision with which they are calculated is basic to the precision in the decision made in reducing the dimensionality. Therefore, since the eigenvalues are estimated with greater precision when the ratio r is close to one, or equivalently when the ratio of the interaction mean square to the error mean square is large, the closer r is to one, the greater the precision in the decision made in reducing the dimensionality of the estimated space. (The confidence interval estimate of r is given in section 4.2.)

7. SUMMARY AND CONCLUSIONS

Analysis of two-way interaction matrices is of interest for many reasons. In this particular study the motivating objective was to assess the degree to which the true spatial configuration of locations, as defined by the cultivar by location interaction effects, is reproduced by the estimated interaction matrix. The ultimate purpose was to define conditons under which it would be reasonably consistent and successful to classify locations so as to minimize within cluster cultivar by location interactions.

A Monte Carlo experiment was carried out to study the repeatability of the configuration of the ℓ location points each represented by a $(v \times 1)$ vector of estimated interaction effects. This repeatability was measured by the statistic m^2, a standardized version of R^2_{min}, (Gower (1971)), which measures the degree to which the estimated interaction effects in any trial reproduces the true, or parametric, spatial configuration. This statistic, m^2, is the standardized sum of squared distances of the observed location points, after rotation and dilation, from the corresponding true points. The rationale for allowing rotation and dilation in the assessment of repeatability was that any consistent classification procedure would be dependent only on relative distances between locations, not on the absolute distances, nor on the orientation of the configuration.

The two aspects of the problem with potential impact on the behavior of m^2 were the nature of the true spatial configuration of the location points and the relative importance of cultivar by location interaction to the experimental error in the estimation of the interaction effects. The true spatial configurations were defined by the matrix of true

interaction effects, $A(v \times \ell)$. The three dimensional figures used for the configurations of the true location points were rectangular parallelepipeds. The relative sizes of the sides of the rectangular parallelepiped were controlled by choice of the eigenvalues of AA', δ_1^2, δ_2^2, and δ_3^2. Such a definition allows for spatial configurations ranging from all locations being "equally" spaced in three dimensions, if $\delta_1^2 = \delta_2^2 = \delta_3^2$, to increasing degrees of similarity of pairs or quadruplets of points as δ_2^2 and δ_3^2 decrease in size relative to δ_1^2. The relative importance of the interaction effects to the error variance was controlled by specification of the ratio

$$r = \sum_{i=1}^{m} \delta_i^2 \Big/ \left[\sum_{i=1}^{m} \delta_i^2 + (\ell-1)(v-1)\sigma^2 \right]$$

where $\sigma^2 = \sigma_e^2/n$ is the variance of the estimates of the cultivar-location means $(y_{ij.})$.

In any specific classification problem, ℓ and v are known and completely under the control of the researcher. The ratio r is not known a priori but is subject to some control by the researcher in the choice of n and may be estimated from the data. The true spatial configuration, however, is completely beyond the control of the experimenter and all information derives from the data.

The statistic m^2 measured the degree to which the data reproduced the true spatial configuration in our Monte Carlo studies but, in any set of real data, m^2 cannot be computed. Therefore, since m^2 was not uniformly small enough over the cases studied to provide a uniformly high degree of success of classification, criteria must be developed for judging whether or not classification would be reliable.

The study of m^2 showed that $\bar{m}$, the mean of m^2, was a monotonically decreasing function of r toward zero from an upper limit, $b(A) < 1$, dependent on A. In all cases of A matrices studied, a linear regression through the origin of $\bar{m}$ on $(1-r)$ gave a good fit, $\bar{m} = b(A)\cdot(1-r)$. The distribution of m^2 depended on the number and relative sizes of the non-zero roots of AA'. However, as r increased toward one these differences in the distributions decreased to the point that all cases studied, for a given ℓ and v, gave satisfactorily low values of m^2 for nearly the same values of r. Thus, specification of an upper bound on m^2 can be translated, for practical purposes, into specification of a lower bound on r, conditional on the values of ℓ and v.

A confidence interval estimate of r can be calculated by using the approximation of a non-central F distribution by a central F given by Searle (1971). The ratio $f = MS(\text{interaction})/MS(\text{error})$ is distributed as a non-central F with $(\ell-1)(v-1)$ and n degrees of freedom (n are the error degrees of freedom) and non-centrality parameter λ, and $(\ell-1)(v-1)f/[(\ell-1)(v-1) + 2\lambda]$ is approximately distributed as a central F with $[(\ell-1)(v-1) + 2\lambda]^2/[(\ell-1)(v-1) + 4\lambda]$ and n degrees of freedom. This can be expressed in terms of r as $f(1-r)$ is approximately distributed as a central F with $(\ell-1)(v-1)/(1-r^2)$ and n degrees of freedom. A point estimate of r, $\hat{r}$, can be calculated as $\hat{r} = 1 - \frac{1}{f}$. The approximate confidence interval estimate of $\hat{r}$ is given by

$$P\left\{1 - \frac{F_{1-\alpha_2}}{f} < r < 1 - \frac{F_{\alpha_1}}{f}\right\} = 1 - \alpha_1 - \alpha_2 .$$

To determine what constitutes a desirable value for m^2 a geometric argument was given in section 4.2, and it was decided that an m^2 $m^2 \leq .25$ was "sufficiently" small.

The distribution of m^2 was seen to change as ℓ and v were changed. The effect of varying ℓ was greater than that of varying v when the values of the 95th percentiles of the distribution of m^2 for the ratios of interest were compared. When v was changed keeping ℓ constant, and $r \geq .75$, the change in the value of the 95th percentile of m^2 was almost negligible. For the cases of ℓ studied ($\ell=8$ and $\ell=4$) the minimum value of r necessary so that $m^2 \leq .25$ with approximate probability .95 was .75 for $\ell=8$ and .80 for $\ell=4$. For values of ℓ between 8 and 4 a linear interpolation can be used to determine r. For ℓ greater than 8 it is clear that a smaller value of r would be permissible, although, from the cases studied, it is not possible to make any inference on how much smaller the value of r will be for ℓ greater than 8.

Another problem discussed was the relative gain or loss of precision in the classification when the dimension of the space of estimated location points is reduced to include only the more important dimensions as reflected in the eigenvalues of the interaction covariance matrix.

The Monte Carlo results clearly indicate that there are situations when repeatability of the spatial configuration is improved by imposing a reduction in dimensionality on the estimation space. It was determined, in general for the cases studied, that when the dimension of the space of estimated points was reduced so that it coincided with the number of dominant dimensions of the space of true location points, there was an improvement in the reproducibility of the true points by

the observed points. But it was also found that it is better to keep more dimensions than optimal than to keep fewer. To reduce the dimensionality in a specific problem requires information on the eigenvalues of the interaction covariance matrix, and it appears that comparison of the estimated eigenvalues to their averages in the null case (no interaction effects), provides a basis for judging what dimensionality should be used. This approach is explained in Mandel (1971) where he also provides a table of expected values of the first three largest eigenvalues of the interaction covariance matrix for the null case. To make the decision as to whether or not the dimensionality should be reduced the hypothesis $E(d_i^2/\nu_i) = \sigma^2$ is tested where d_i^2 is the ith eigenvalue of the interaction covariance matrix and ν_i is the expected value of the ith eigenvalue in the null case, σ^2 is the error variance. If the hypothesis is not rejected, then that component of the interaction sum of squares $\left(\sum_{i=1}^{k} d_i^2\right)$ is assumed to be random error. Therefore, now $SS(\alpha\beta) = \sum_{i=1}^{k'} d_i^2$ for some $k' < k$, and since the number of non-zero eigenvalues of the covariance matrix reflects the dimensionality of the location space, the dimension of the estimated space has been reduced. Mandel (1971) uses as a cut-off point to determine the partitioning that point in which there is a 'jump' in d_i^2/ν_i, and the d_i^2/ν_i below this point are declared as random error. In section 5.1, the table of the F-distribution was used to test these hypotheses, and the conclusions obtained using this F table agreed with the conclusions obtained by using the statistic $m^2(p)$, $p = 1, 2, 3,$ and k.

Recognizing that only a small part of the parametric space has been studied, the results suggest that, when presented with a real problem the steps to follow to determine whether or not a classification of

locations according to the cultivar by location interaction and a reduction in dimensionality should be attempted are:

1. Perform the analysis of variance.
2. Using MS(AB) from 1 calculate $f = MS(AB)/MSE$ and test $H_o: r = r_o$ against $H_1: r > r_o$ by using the approximation $f(1-r) \doteq F_{s,\eta}$ where $s = (\ell-1)(v-1)/(1-r^2)$ or compute an approximate confidence interval estimate for r

$$P\left[1 - \frac{F_{1-\alpha_2}}{f} < r < 1 - \frac{F_{\alpha_1}}{f}\right] = 1 - \alpha_1 - \alpha_2 .$$

3. If r is sufficiently large, say $r \geq .75$ with probability $1 - \alpha$ and $\ell \geq 8$ or $r > .75$ and $4 \leq \ell < 8$, then proceed to step 4. Otherwise stop.
4. Calculate the interaction covariance matrix and decompose it according to the singular value decomposition.
5. Test the "significance" of the eigenvalues of the interaction covariance matrix by using Mandel (1971) or any similar procedure.
6. Recalculate the interaction covariance matrix by using only the "significant" eigenvalues, and proceed to do the classification using this matrix.

8. LIST OF REFERENCES

Abou-El-Fittouh, H. A., J. O. Rawlings, and P. A. Miller. 1969. Classification of environments to control genotype by environment interaction with an application to cotton. Crop Sci. 9:135-140.

Clemm, D. C., P. R. Krishnaiah, and V. B. Waikar. 1971. Tables for the extreme roots of the Wishart Matrix. Aerospace Research Laboratories, ARL 71-0264, December 1971.

Clifford, H. T. and W. Stephenson. 1975. An Introduction to Numerical Classification. Academic Press, New York.

Dickey, D. A. 1978. A program for generating and analyzing large Monte Carlo studies. Unpublished manuscript, N. C. State University at Raleigh.

Eberhart, S. A. and W. A. Russell. 1966. Stability parameters for comparing varieties. Crop Sci. 6:36-40.

Finlay, K. W. and G. N. Wilkinson. 1963. The analysis of adaptation in a plant-breeding programme. Aust. J. Agric. Res. 14:742-754.

Gower, J. C. 1971. Statistical methods of comparing different multivariate analyses of the same data. Mathematics in the Archaeological and Historical Sciences. pp. 138-149.

Gower, J. C. and C. F. Banfield. 1975. Goodness-of-fit criteria for hierarchical classification and their empirical distributions. Proc. 8th Int. Biometric Conf.

Gower, J. C. and G. J. S. Ross. 1969. Minimum spanning trees and single linkage cluster analysis. Appl. Statist. 18:54-64.

Hoel, P. G., S. C. Port, and C. J. Stone. 1971. Introduction to Probability Theory. Houghton Mifflin Company, Boston.

IMSL, Computer Subroutine Libraries in Mathematics and Statistics. 1975. International Mathematical and Statistical Libraries, Inc., Houston, Texas.

Mandel, J. 1970. The distribution of eigenvalues of covariance matrices of residuals in analysis of variance. Journal of Research of the National Bureau of Standards - B. Mathematical Sciences. Vol. 74B, No. 3.

Mandel, J. 1971. A new analysis of variance model for non-additive data. Technometrics 13(1):1-18.

Okuno, T., F. Kikuchi, K. Kumagai, C. Okuno, M. Shiyomi, and H. Tabuchi. 1971. Evaluation of varietal performance in several environments. The Bulletin of the National Institute of Agricultural Sciences, Series A18, 93-147.

Plaisted, R. L. and L. C. Peterson. 1959. A technique for evaluating the ability of selection to yield consistently in different locations or seasons. Am. Potato J. 36:381-385.

Schuurmann, F. J., P. R. Krishnaiah, and A. K. Chattopadhay. 1973. Tables for the distributions of the ratios of the extreme roots to the trace of Wishart Matrix. Aerospace Research Laboratories, ARL 73-0010, February 1973.

Searle, S. R. 1971. Linear Models. John Wiley and Sons, Inc., New York.

9. APPENDIX

AN UPPER BOUND FOR THE EXPECTATION OF m^2

In the following it is proven that $1 - r$ is an upper bound for the expected value of the statistic m^2, and moreover that the bias in m^2, if any, is negative.

Given that $m^2 = 1 - (\text{tr } A'\hat{A}H)^2/(\text{tr } AA' \text{tr } \hat{A}\hat{A}')$ where $E(\hat{A}) = A(v \times \ell)$, and $E(\text{tr } \hat{A}\hat{A}') > 0$, $A'\hat{A} = WCZ'$ where W is the $(\ell \times \ell)$ matrix of the orthonormalized eigenvectors of $(A'\hat{A})(A'\hat{A})'$, $W'W = WW' = I$, Z is the $(\ell \times \ell)$ matrix of the orthonormalized eigenvectors of $(A'\hat{A})'(A'\hat{A})$, $Z'Z = ZZ' = I$, and C^2 is a $(\ell \times \ell)$ diagonal matrix containing the eigenvalues of $(A'\hat{A})'(A'\hat{A})$, and $H = ZW'$; $r = \text{tr } AA'/E(\text{tr } \hat{A}\hat{A}')$, then $E(m^2) \leq 1 - r$

Proof:

From Schwarz's inequality (see Hoel, et al. (1971)), $E^2(XY) \leq E(X^2) \cdot E(Y^2)$ and equality holds if and only if $P(X = aY) = 1$ for some constant a.

Let

$$X = \frac{\text{tr } A'\hat{A}H}{\sqrt{\text{tr } \hat{A}\hat{A}'}} \quad \text{and} \quad Y = \sqrt{\text{tr } \hat{A}\hat{A}'}$$

Then,

$$E^2(\text{tr } A'\hat{A}H) \leq E\left\{\frac{\text{tr } A'\hat{A}H}{\sqrt{\text{tr } \hat{A}\hat{A}'}}\right\}^2 \cdot \{E(\sqrt{\text{tr } \hat{A}\hat{A}'})^2\}$$

and equality holds if and only if

$$P\left\{\frac{\text{tr } A'\hat{A}H}{\sqrt{\text{tr } \hat{A}\hat{A}'}} = a\sqrt{\text{tr } \hat{A}\hat{A}'}\right\} = 1$$

for some a. Therefore,

$$\frac{E^2(\text{tr } A'\hat{A}H)}{\text{tr } AA'} \leq E\left[\frac{(\text{tr } A'\hat{A}H)^2}{\text{tr } AA' \cdot \text{tr } \hat{A}\hat{A}'}\right] \cdot E(\text{tr } \hat{A}\hat{A}')$$

$$E\left[\frac{(\text{tr } A'\hat{A}H)^2}{\text{tr } AA' \cdot \text{tr } \hat{A}\hat{A}'}\right] \geq \frac{E^2(\text{tr } A'\hat{A}H)}{E(\text{tr } \hat{A}\hat{A}') \cdot \text{tr } AA'} = \rho^2_{A,\hat{A}H}$$

where $\rho_{A,\hat{A}H}$ is the correlation coefficient betwenn the elements of A and $\hat{A}H$. But now the left hand side term of this inequality is equal to $1 - E(m^2)$. Therefore,

$$1 - E(m^2) \geq \rho^2_{A,\hat{A}H} \quad \text{and equality holds}$$

if and only if $P\{\text{tr } A'\hat{A}H = a \text{ tr } \hat{A}\hat{A}'\} = 1$ or equivalently if and only if $P(1 - m^2 = b \text{ tr } \hat{A}\hat{A}') = 1$ for some constant b.

Now, to prove that $E(m^2) \leq 1 - r$ we need to prove that $\rho^2_{A,\hat{A}H} \geq r$. The proof of this is as follows:

Proof.

It was given that $A'\hat{A} = WCZ'$, and $WW' = W'W = I = Z'Z = ZZ'$, $H = ZW'$.

Then $\text{tr}(A'\hat{A}H) = \text{tr } C = \sum_{i=1}^{\ell} c_i$ and $\text{tr}(A'\hat{A}) = \text{tr}(WCZ') = \text{tr } CZ'W$.

Let $T = Z'W$. Then $T'T = I$, and therefore $\sum_{k=1}^{\ell} t_{ik}^2 = 1$ for all i, which implies that $|t_{ik}| \leq 1$ for all i and k.

So we have

$$\text{tr } A'\hat{A} = \text{tr } CZ'W = \text{tr } CT = \sum_{i=1}^{\ell} c_i t_{ii} \leq \sum_{i=1}^{\ell} c_i |t_{ii}| \leq \sum_{i=1}^{\ell} c_i .$$

Therefore, $tr(A'\hat{A}) \leq tr\ C = tr\ A'\hat{A}H$ which implies

$0 \leq E(tr\ A'\hat{A}) = tr\ A'A \leq E(tr\ A'\hat{A}H)$.

All terms are positive, so both sides can be squared,

$$(tr\ (A'A))^2 \leq E^2(tr\ A'\hat{A}H)$$

$$\frac{E^2(tr\ A'\hat{A}H)}{(tr\ A'A)^2} \geq 1$$

$$\frac{E^2(tr\ A'\hat{A}H)}{(tr\ A'A)^2} \cdot r = \frac{E^2(tr\ A'\hat{A}H)}{tr\ A'A \cdot E(tr\ \hat{A}\hat{A}')} = \rho^2_{A,\hat{A}H} \geq r$$

So

$$\rho^2_{A,\hat{A}H} \geq r \ .$$

Therefore

$$E(m^2) \leq 1 - \rho^2_{A,\hat{A}H} \leq 1 - r \ .$$

Also, the bias of $1 - m^2$, $E(1-m^2) - \rho^2_{A,\hat{A}H}$, can be seen to be greater than or equal to zero.