Comment: Group Symmetry Covariance Models

Michael D. Perlman

Mark Schervish's review is a well-conceived and welcome Michelin Guide to Multivariate Analysis, classical and current. Simultaneously comprehensive and opinionated, it both surveys and evaluates the major features of this complex statistical landscape. Some of his opinions and emphasis may be arguable, but his views on the interplay between methodology and theory, the need for scientific relevance in statistical research, and the future directions of multivariate analysis should be considered carefully by every student of statistics.

MULTIVARIATE ANALYSIS IN COPENHAGEN

I will attempt to supplement Schervish's remarks with a brief description of some elegant and significant contributions by Danish statisticians to multivariate theory, particularly S. A. Andersson, H. Brøns and S. T. Jensen of the University of Copenhagen. They have developed an algebraic approach, not yet fully published nor in textbook form, which provides a unified mathematical framework for studying amenable problems of classical multivariate analysis. Their approach not only unifies much of the classical theory, but also extends its scope to include several new types of amenable models and testing problems. (Because they show that in some sense there are only finitely many types of amenable models and testing problems, their results also define the limits of classical multivariate theory.) Just as T. W. Anderson's classic book lighted the path of an entire generation of researchers, their approach could well become the standard for future work on multivariate statistical theory.

By "amenable problems of classical multivariate analysis," I loosely refer to statistical models and hypothesis testing problems consisting of families of multivariate normal distributions determined by linear constraints on the mean vector and/or covariance matrix and that, furthermore, allow explicit likelihood analysis, e.g. explicit (linear) maximum likelihood estimates, explicit likelihood ratio test statistics and explicit expressions for their distributions. Alternatively, one could loosely refer to "multivariate problems similar to those treated in T. W. Anderson's classic textbook."

These ideas are most easily illustrated by considering the three main hypothesis-testing problems for covariance matrices treated in Chapters 9 and 10 of Anderson:

(a) Testing independence of two or more sets of variates in a p-variate normal population $N_p(\mu, \Sigma)$, e.g., testing

$$H_0^{(a)}: \Sigma_{12} = 0 \quad \text{vs.} \quad H_1^{(a)}: \Sigma_{12} \neq 0,$$

where the covariance matrix Σ is partitioned as (Σ_{ij}) with $\Sigma_{ij}: p_i \times p_j$, $i, j = 1, 2$, $p_1 + p_2 = p$.

(b) Testing sphericity, i.e., testing

$$H_0^{(b)}: \Sigma = \sigma^2 I_p, \quad \sigma^2 > 0 \quad \text{vs.} \quad H_1^{(b)}: \Sigma \neq \sigma^2 I_p,$$

where I_p denotes the $p \times p$ identity matrix.

(c) Testing equality of two or more covariance matrices, e.g., testing

$$H_0^{(c)}: \Sigma_1 = \Sigma_2 \quad \text{vs.} \quad H_1^{(c)}: \Sigma_1 \neq \Sigma_2,$$

based on independent samples from the normal populations $N_{p_1}(\mu_1, \Sigma_1)$ and $N_{p_2}(\mu_2, \Sigma_2)$ with $p_1 = p_2$.

Anderson's treatments of each of these three "amenable" problems are both clear and complete. In each case, the maximum likelihood estimate (MLE), the likelihood ratio test (LRT) statistic and its null distribution are explicitly derived. The reader may notice, however, that the derivations and results have a common mathematical flavor, raising an important but unposed question: which other hypothesis testing problems for normal covariance matrices share this "common flavor"? More generally, which multivariate normal models and testing problems are amenable to explicit analysis? Can one characterize the "amenable problems in classical multivariate analysis," or, at least, describe general classes (as opposed to isolated examples) of models and testing problems that can be analyzed in an explicit and unified manner?

GROUP SYMMETRY COVARIANCE MODELS

These questions have been addressed in the work of Andersson, Brøns and Jensen (unpublished) in Copenhagen. By using standard mathematical tools, e.g., linear algebra, group representations and invariant

Michael D. Perlman is Professor, Department of Statistics, University of Washington, Seattle, Washington 98195.

measures, they have obtained answers that are both elegant and illuminating. Furthermore, the general classes of models they obtain and study have meaningful statistical interpretability. For example, as we shall show below, the statistical models in (a), (b), and (c) above are special cases of *group symmetry covariance models* (≡ *symmetry models*), and the three testing problems are special cases of the general problem of *testing one symmetry model against another*. Other examples of symmetry models include the intraclass correlation model, the circular symmetry model and the complex multivariate normal model. Each of these has been studied individually during the past twenty years without general recognition (outside of Denmark) that they are actually special cases of the group symmetry model. Thus, although Schervish rightly points to "an entire industry of ad hoc methods for dealing with multivariate data," he could also have noted the current ad hoc state of multivariate theory as well.

The general group symmetry covariance model was introduced by S. A. Andersson (1975). (See also Chapter 9 of Eaton, 1983.) Such a model describes symmetries present in the error structure of multivariate observations, e.g., on biological objects, on symmetrically arranged seismographs, etc. A symmetry model may be described formally as a family $\mathbf{S}_G^+$ of covariance matrices Σ that remain invariant under a finite group G of orthogonal transformations. More precisely, let $Y{:}p \times 1$ denote a multivariate observation (i.e., a random vector) with covariance matrix $\Sigma{:}p \times p$, a positive definite (pd) symmetric matrix, and let $G = \{g\}$ be a finite group of $p \times p$ orthogonal matrices. The symmetry model determined by G is the family of covariance matrices

$$(1)\quad \mathbf{S}_G^+ = \{\Sigma \mid \Sigma \text{ pd}, g\Sigma g' = \Sigma \text{ for all } g \text{ in } G\}.$$

The hypothesis $\text{cov}(Y) \in \mathbf{S}_G^+$ implies that the error structure of Y satisfies a set of symmetry restrictions, namely that $\text{cov}(Y) = \text{cov}(gY)$ for all g in G.

EXAMPLES OF SYMMETRY MODELS

Suppose that

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_k \end{pmatrix}$$

with Y_i: $q \times 1$, where $Y_1, \ldots, Y_k$ represent the measurements obtained from k identical seismographs arranged in a circular pattern (e.g., around the base of Mt. St. Helens). Here, $p = qk$. It may be reasonable to assume (or to test) that $\text{cov}(Y) \equiv \Sigma$ possesses *circular block symmetry*, i.e., that

$$\text{cov}\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{pmatrix} = \text{cov}\begin{pmatrix} Y_2 \\ Y_3 \\ \vdots \\ Y_1 \end{pmatrix} = \cdots = \text{cov}\begin{pmatrix} Y_k \\ Y_1 \\ \vdots \\ Y_{k-1} \end{pmatrix},$$

or, equivalently, that $\text{cov}(Y_i, Y_j) = \text{cov}(Y_{i+r}, Y_{j+r})$ for all i, j, r (mod k). This condition may be re-expressed as

$$(2)\quad \text{cov}(Y) = \text{cov}(P^r Y), \quad r = 0, 1, \ldots, k-1,$$

where

$$P = \begin{pmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & & \\ 0 & 0 & 0 & \ddots & \\ & & & \ddots & I \\ I & 0 & 0 & \cdots & 0 \end{pmatrix} : qk \times qk$$

and I is the $q \times q$ identity matrix. (Note that it suffices to verify (2) for $r = 1$ only.) If we set $G_1 = \{I, P, \ldots, P^{k-1}\}$ (a *cyclic* group of order k) then (2) is equivalent to the condition that $\Sigma \in \mathbf{S}_{G_1}^+$. It is easy to verify that $\mathbf{S}_{G_1}^+$ consists of all positive definite symmetric matrices of the following forms:

$$k = 3{:}\ \begin{pmatrix} A & B & B' \\ B' & A & B \\ B & B' & A \end{pmatrix}, \quad A = A',$$

$$k = 4{:}\ \begin{pmatrix} A & B & C & B' \\ B' & A & B & C \\ C & B' & A & B \\ B & C & B' & A \end{pmatrix}, \quad A = A', \quad C = C',$$

etc., where A, B, C, etc., are $q \times q$. Note that it is *not* required that $B = B'$. (In the general case, it is *not* required that $\text{cov}(Y_i, Y_j) = \text{cov}(Y_j, Y_i)$.)

Some confusion appears in the literature regarding the term "circular block symmetry." This term frequently has been used to refer instead to covariance matrices of the forms

$$k = 3{:}\ \begin{pmatrix} A & B & B \\ B & A & B \\ B & B & A \end{pmatrix}, \quad A = A', \quad B = B',$$

$$k = 4{:}\ \begin{pmatrix} A & B & C & B \\ B & A & B & C \\ C & B & A & B \\ B & C & B & A \end{pmatrix},$$

$$A = A', \quad B = B', \quad C = C',$$

etc. Here the added condition that $\text{cov}(Y_i, Y_j) = \text{cov}(Y_j, Y_i)$ is imposed. Such models do not occur as $\mathbf{S}_{G_1}^+$ for the *cyclic* group G_1 of order k but rather as $\mathbf{S}_{G_2}^+$ for the *dihedral* group G_2 of order $2k$, hence more accurately should be called *dihedral block symmetry* models. Here, $G_2 = \{I, P, \ldots, P^{k-1}, Q, QP, \ldots, QP^{k-1}\}$, where

$$Q = \begin{bmatrix} 0 & & I \\ & \cdot^{\cdot^{\cdot}} & \\ I & & 0 \end{bmatrix} : qk \times qk.$$

Stated simply, G_2 is (isomorphic to) the group of all rotations and reflections that leave a regular k-gon invariant, while G_1 consists of the rotations only. The hypothesis $\Sigma \in \mathbf{S}^+_{G_1}(\Sigma \in \mathbf{S}^+_{G_2})$ states that the error structure of Y remains invariant under all rotations (all rotations and reflections) of the k seismographs among themselves.

We also mention the model of *complete block symmetry*, i.e., $\text{cov}(Y_i) = \text{cov}(Y_j)$ and $\text{cov}(Y_i, Y_j) = \text{cov}(Y_l, Y_m)$ for all $i \neq j$, $l \neq m$. This arises as $\mathbf{S}^+_{G_3}$, $G_3 = \{\Pi \otimes I_q \mid \Pi \text{ a } k \times k \text{ permutation matrix}\}$ (order $k!$), where $\otimes$ denotes the Kronecker product. Here, $\mathbf{S}^+_{G_3}$ consists of all covariance matrices of the form

$$\begin{bmatrix} A & B & & \cdots & B \\ B & A & & & \vdots \\ & & \ddots & & \\ \vdots & & & A & B \\ B & \cdots & & B & A \end{bmatrix}, \quad A = A', \quad B = B'.$$

The hypothesis $\Sigma \in \mathbf{S}^+_{G_3}$ implies that the error structure remains invariant under *all* permutations of the k seismographs.

Another important example of a group symmetry model is the *q-variate complex normal distribution*, more precisely, the *2q-variate real normal distribution with complex covariance structure*. This statistical model consists of all p-variate (real) normal distributions $N_p(\mu, \Sigma)$ with $p = 2q$ and Σ of the form

$$(3) \qquad \Sigma = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}, \quad A = A', \quad B = -B',$$

where A, B are $q \times q$. (Note that $A + iB$ is a complex positive definite Hermitian matrix whenever Σ in (3) is a real positive definite symmetric matrix.) It is easily verified that Σ has the form (3) if and only if $\Sigma \in \mathbf{S}^+_{G_4}$, where

$$G_4 = \{\pm I_p, \pm J_p\},$$
$$J_p = \begin{pmatrix} 0 & -I_q \\ I_q & 0 \end{pmatrix}.$$

From this point of view, the complex multivariate normal distribution is a special case of the real multivariate normal distribution with a certain group symmetry covariance structure.

Finally, it is easy to see that the covariance models occurring in problems (a), (b) and (c) above are also symmetry models:

$$(4) \qquad \begin{aligned} H_0^{(a)}&: \Sigma \in \mathbf{S}^+_{G(a)}, & H_1^{(a)}&: \Sigma \in \mathbf{S}^+_{\{I_p\}}, \\ H_0^{(b)}&: \Sigma \in \mathbf{S}^+_{G(b)}, & H_1^{(b)}&: \Sigma \in \mathbf{S}^+_{\{I_p\}}, \\ H_0^{(c)}&: \Sigma \in \mathbf{S}^+_{G(c)}, & H_1^{(c)}&: \Sigma \in \mathbf{S}^+_{G(a)}, \end{aligned}$$

where

$$G(a) = \left\{ \begin{pmatrix} \pm I_{p_1} & 0 \\ 0 & \pm I_{p_2} \end{pmatrix} \right\},$$

$G(b) = \{D\Pi \mid D = \text{diag}(\pm 1, \ldots, \pm 1), \Pi \text{ a } p \times p \text{ permutation matrix}\}$, and

$$G(c) = \left\{ \begin{pmatrix} \pm I_{p_1} & 0 \\ 0 & \pm I_{p_2} \end{pmatrix}, \begin{pmatrix} 0 & \pm I_{p_1} \\ \pm I_{p_2} & 0 \end{pmatrix} \right\}.$$

Here, $\#(G(a)) = 4$, $\#(G(b)) = 2^p \cdot p!$, $\#(G(c)) = 8$, where $\#(G)$ denotes the order of G. (Note that $\mathbf{S}^+_{\{I_p\}} = \{\text{all } p \times p \text{ pd matrices}\}$ and recall that $p_1 = p_2$ in (c).)

MLE AND LRT FOR SYMMETRY MODELS

If symmetries are known to be present, then sharper statistical inferences can be obtained, e.g. more accurate estimates of Σ, more powerful tests concerning Σ and/or related population mean vectors. Numerous articles dealing with particular cases (i.e., with particular groups G) have appeared in the literature. Without a common framework, however, such models have arisen and have been treated on an ad hoc basis, with the result that common structure is ignored and important problems overlooked. For example, the natural problem of testing dihedral block symmetry vs. circular block symmetry (recall the preceding examples) has not yet been posed or studied, but is a special case of the general testing problem

$$(5) \qquad H_0: \Sigma \in \mathbf{S}^+_{G_0} \quad \text{vs.} \quad H_1: \Sigma \in \mathbf{S}^+_{G_1},$$

where G_0, G_1 are finite groups of $p \times p$ orthogonal matrices such that $G_0 \supseteq G_1$ (hence, $\mathbf{S}_{G_0} \subseteq \mathbf{S}^+_{G_1}$). Similarly, from (4), each of the testing problems (a), (b) and (c) mentioned above is a special case of (5).

In order to demonstrate the utility and simplicity of this general formulation, we briefly consider MLEs and LRTs for group symmetry models. Suppose that S denotes the sample covariance matrix based on a random sample of size m from $N_p(\mu, \Sigma)$. Thus, $mS \sim W_p(\Sigma, m)$, the Wishart distribution with m degrees of freedom, $m = m - 1$. If $m \geq p$, so that S is positive definite with probability one, it is straightforward to show that under the symmetry hypothesis $\Sigma \in \mathbf{S}^+_G$, the MLE of Σ is given by the positive definite matrix

$$S_G \equiv \frac{1}{\#(G)} \sum_{g \in G} gSg', \tag{6}$$

an explicit linear function of S. Furthermore, the LRT for (5) rejects H_0 if

$$\frac{|S_{G_1}|}{|S_{G_0}|} \le c \tag{7}$$

for some constant c, $0 < c < 1$, where $|\cdot|$ denotes the determinant. It is easily checked that the LRTs derived in Anderson for problems (a), (b) and (c) are special cases of (7): in (a), if $S = (S_{ij})$ is partitioned in the same manner as Σ, then

$$S_{G(a)} = \begin{pmatrix} S_{11} & 0 \\ 0 & S_{22} \end{pmatrix}, \quad S_{\{I_p\}} = S,$$

$$\frac{|S_{\{I_p\}}|}{|S_{G(a)}|} = \frac{|S|}{|S_{11}| \cdot |S_{22}|}$$

(cf. Anderson, page 379, equation (16)); in (b),

$$S_{G(b)} = \frac{1}{p}(\operatorname{tr} S)I_p, \quad S_{\{I_p\}} = S,$$

$$\frac{|S_{\{I_p\}}|}{|S_{G(b)}|} = \frac{|S|}{[(1/p)\operatorname{tr} S]^p}$$

(cf. Anderson, page 428, equation (7)); while in (*c*), for the case of *equal* sample sizes from the two populations,

$$S_{G(c)} = \frac{1}{2}\begin{pmatrix} S_{11} + S_{22} & 0 \\ 0 & S_{11} + S_{22} \end{pmatrix},$$

$$S_{G(a)} = \begin{pmatrix} S_{11} & 0 \\ 0 & S_{22} \end{pmatrix},$$

$$\frac{|S_{G(a)}|}{|S_{G(c)}|} = \frac{|S_{11}| \cdot |S_{22}|}{|\tfrac{1}{2}(S_{11} + S_{22})|^2}$$

(cf. Anderson, page 406, equation (8)). (The LRT for problem (c) for the case of unequal sample sizes also can be expressed via (7) but requires some extra notation.)

The reader also can easily apply (6) to obtain the MLE for Σ under the symmetry models $\mathbf{S}^{+}_{G_j}$, $j = 1, \ldots, 4$, introduced above.

By using the standard theory of group representations, Andersson (1975) obtained the following characterization of group symmetry covariance models. A family $\mathbf{S}$ of $p \times p$ pd covariance matrices is a group symmetry model if and only if there exist positive integers $t, p_1, \ldots, p_t, r_1, \ldots, r_t$ and a fixed $p \times p$ orthogonal matrix Γ such that $\Gamma\mathbf{S}\Gamma'$ consists of all block diagonal covariance matrices Σ of the form

$$\Sigma = \operatorname{diag}(\underbrace{\Sigma_1, \ldots, \Sigma_1}_{r_1}; \cdots; \underbrace{\Sigma_t, \ldots, \Sigma_t}_{r_t})$$

where Σ_i is $p_i \times p_i$, $\sum_1^t p_i r_i = p$, and each Σ_i ranges over all $p_i \times p_i$ pd real covariance matrices of real, complex or quaternion structure. (Σ is of real structure if it is a real symmetric matrix; Σ is of complex structure if it is of the form (2); see Andersson (1975) for the definition of quaternion structure.) From this result, Andersson, Brøns and Jensen (unpublished) have shown that the general problem (5) of testing one group symmetry model against another may be decomposed into a finite product of problems of ten basic types (each with possible multiplicity). These ten problems are (1, 2, 3) testing independence of sets of variates in the real, complex and quaternion multivariate normal distributions; (4, 5, 6) testing equality of covariance matrices in the real, complex and quaternion cases; (7) testing for the reality of a p-dimensional complex normal distribution; (8) testing that a $2p$-dimensional real normal distribution has a p-dimensional complex structure; (9–10) the same as 7–8 with "real" and "complex" replaced by "complex" and "quaternion."

Problems 1–6 are relatively well-known, although problems 7 and 8 are nonstandard and first were treated individually by Khatri (1965) and Andersson (1978), respectively. They obtained explicit representations of the maximal invariant statistics (under appropriate transformation groups) as roots of determinantal equations involving the sample covariance matrix, together with their null (central) distributions. This in turn yields the central distributions of the LRT statistics. Recently Andersson, Brøns and Jensen (1983) developed a unified approach which simultaneously yields these results for all ten basic problems. The first results for problems 7 and 8 concerning noncentral distributions, which are needed for decision-theoretic results concerning the power functions of tests (e.g., admissibility), were obtained by Andersson and Perlman (1984).

Lest the reader be left with the impression that group symmetry covariance models lie exclusively within the realm of hypothesis testing and invariance, I also point to a Bayesian analysis of these models in a recent paper by Consonni and Dawid (1985).

An interesting characterization of a class of covariance models slightly larger than the class of symmetry models recently has been obtained by Jensen (1988).

Finally, it should be noted that the general theory of Andersson, Brøns and Jensen (unpublished) is not confined to group symmetry covariance models alone, but also treats models under which related structure is imposed simultaneously on the mean vectors. Such models include both the classical multivariate linear regression (≡ MANOVA) model and extensions wherein the covariance structure is assumed to satisfy additional symmetry conditions.

CONCLUDING REMARKS

Although my comments have been confined to classical multivariate analysis (i.e., the multivariate normal distribution and linear models), their implications are equally relevant to the broader spectrum of multivariate analysis so capably surveyed by Mark Schervish (indeed, to the entire field of statistical theory). The recent explosion of new statistical models and techniques (e.g., nonparametric, nonlinear, graph-theoretical) in multivariate analysis presents an important opportunity, in fact, an obligation, for multivariate theorists to determine the statistically and *mathematically* meaningful classes of such models. I stress "mathematically" meaningful in order to emphasize that recognition of the *precise mathematical structure* (e.g., the group symmetry of covariance models) of a class of statistical models is essential for an accurate characterization and unified analysis of the class. Conversely, this approach might lead to the recognition that a proposed class of models is not formulated in a mathematically precise way, which in turn might suggest an alternative class with more desirable properties.

Another example of an important and successful application of this approach is the characterization and analysis of orthogonal analysis of variance and variance component designs in terms of their underlying lattice structure. A good survey of this subject may be found in Tjur (1984), the accompanying discussions by Bailey, Speed and Wynn, and the references, in particular Jensen (1979). The forthcoming paper by Andersson (1987) concerning more general orthogonal designs of linear models presents a definitive treatment of the lattice-theoretic formulation of such schemes.

The characterizations of covariance models as group symmetry models and of orthogonal ANOVA designs as lattice-ordered designs demonstrate the benefits of precise mathematical formulations of multivariate models. Our knowledge of many other branches of multivariate analysis, such as factor analysis, path analysis and econometric models, will be substantially enhanced by this approach. As Schervish points out in Section 10.1, for example, graph theory is playing an increasingly important role in path analysis. Such developments are very welcome and should be pursued vigorously.

ACKNOWLEDGMENT

This work was supported in part by National Science Foundation Research Grant DMS 86-03489.

ADDITIONAL REFERENCES

ANDERSSON, S. A. (1975). Invariant normal models. *Ann. Statist.* **3** 132–154.

ANDERSSON, S. A. (1978). Canonical correlations with respect to a complex structure. Preprint No. 13, Institute of Mathematical Statistics, Univ. Copenhagen.

ANDERSSON, S. A. (1987). The lattice structure of orthogonal linear models and orthogonal variance component models. Technical Report, Institute of Mathematical Statistics, Univ. Copenhagen.

ANDERSSON, S. A., BRØNS, H. and JENSEN, S. T. (unpublished). An algebraic theory of normal statistical models. Lectures notes (in Danish), Institute of Mathematical Statistics, Univ. Copenhagen.

ANDERSSON, S. A., BRØNS, H. and JENSEN, S. T. (1983). Distribution of eigenvalues in multivariate statistical analysis. *Ann. Statist.* **11** 392–415.

ANDERSSON, S. A. and PERLMAN, M. D. (1984). Two testing problems relating the real and complex multivariate normal distributions. *J. Multivariate Anal.* **14** 21–51.

CONSONNI, G. and DAWID, A. P. (1985). Decomposition and Bayesian analysis of invariant normal linear models. *Linear Algebra Appl.* **70** 21–49.

EATON, M. L. (1983). *Multivariate Statistics: A Vector Space Approach.* Wiley, New York.

JENSEN, S. T. (1979). Varianskomponentmodeller i fuldstnendigt balancerede forsøg. Technical Report, Institute of Mathematical Statistics, Univ. Copenhagen.

JENSEN, S. T. (1988). Covariance hypotheses which are linear in both the covariance and the inverse covariance. *Ann. Statist.* **16**. To appear.

KHATRI, C. G. (1965). A test for reality of a covariance matrix in a certain complex Gaussian distribution. *Ann. Math. Statist.* **27** 713–721.

TJUR, T. (1984). Analysis of variance models in orthogonal designs (with discussion). *Internat. Statist. Rev.* **52** 33–81.