KM 5(1982)
pag 129-150

ESTIMATION AND MODEL COMPARISONS FOR REPEATED MEASURES DATA

Charles Lewis and Carina van Knippenberg*

Rijksuniversiteit Groningen

Summary

Alternative approaches to the use of multivariate analysis of variance for repeated measures data are described and compared. Two which have been advocated in the literature should definitely not be used. Suggestions for obtaining appropriate analyses from the computer program MULTIVARIANCE are given in an appendix.

Keywords: repeated measures, growth curves, multivariate analysis of variance, weighted analysis, analysis of covariance.

* Charles Lewis, Vakgroep Statistiek en Meettheorie, FSW, Oude Boteringestr. 23, 9712 GC Groningen
Carina van Knippenberg, Verkeerskundig Studiecentrum, R.U.G., Postbus 14, Haren.

1. Introduction

Repeated observations on the same experimental units are a common occurrence in research in the social and biological sciences. In cases where the assumptions seem appropriate, multivariate analysis of variance has been widely advocated as a means of making statistical inferences from such data. Adopting this approach has been made easier by the wide availability of general purpose computer programs for multivariate statistical analysis, such as MULTIVARIANCE (Finn, 1978). Unfortunately, some of the literature which discusses the use of the technique in general, and MULTIVARIANCE in particular, is misleading and incomplete. This is primarily a problem for social science researchers, since complete and accurate details have been available in the biometric literature for some time. (See, for instance, Grizzle & Allen, 1969.)

In the second section of this paper, we describe how multivariate analysis of variance may be "correctly" used (based on likelihood considerations) to compare models for repeated measures data and to estimate parameters for such models. In the third section, with the approach we describe as a reference point, we consider some of the literature on the subject. This is not meant to be a complete review, but is primarily aimed at pointing out some of the misleading and incomplete statements referred to above.

In the fourth section, to make the issues more concrete, we provide multiple analyses of a data set used by Grizzle & Allen (1969). Finally, in an appendix, we describe some difficulties with Version VI of the MULTIVARIANCE program and how it may be used to carry out the analyses we describe in the second section.

While it is difficult to discuss multivariate analysis of variance without going into technicalities, we try to avoid doing so wherever possible. The technical details, after all, exist elsewhere. Our goal is to make the basic issues as clear as possible.

2. An approach based on likelihood considerations

For the sake of simplicity, we restrict our attention in this section to the analysis of "pure" repeated measures data. There is a single group of subjects, each of whom responds on a single dependent variable under a variety of conditions (or at several times). The conditions (or times) are the same for all subjects. In traditional univariate analysis of variance terms, this might be described as a subjects × conditions design.

We assume that the interest of the researcher is in full rank linear models for the population mean responses (and not, for instance, in the variance component "due to subjects"). Suppose that the models of interest may be arranged in a hierarchy so that, except for the first, each model may be obtained as a special case of the one before it by setting certain parameters in the more general model equal to zero. (In practice, the researcher might be interested in several such hierarchies.)

In any event, the approach we describe demands that the first model in the hierarchy be complete (or saturated, as it is sometimes called) in the sense that it provides a perfect (and unique, thanks to the full rank assumption mentioned above) description of any possible pattern of response means. This implies, among other things, that the number of parameters in this complete model will be equal to the number of repeated measures in the study.

A common example of such a hierarchy occurs when responses are obtained at p times and the interest is in finding a polynomial function of time which describes the response means. The polynomial of order p-1 provides a complete model

$$\beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_{p-1} t^{p-1} \qquad (1)$$

with the p parameters $\beta_0, \beta_1, \beta_2, \ldots, \beta_{p-1}$. Lower order polynomials provide successive models in the hierarchy and are obtained by successively setting the higher order coefficients (β_i) equal to zero.

In the multivariate analysis of repeated measures data, it is assumed that the responses of each subject form a multivariate random variable, identically and independently distributed across subjects according to a multivariate normal distribution with a nonsingular covariance matrix. This unlikely assumption represents, in fact, an important generalization of the assumptions underlying traditional univariate analysis of variance of repeated measures data.

The robustness of the analyses we describe to violations of this assumption is a critical issue which, nonetheless, lies outside the scope of this paper. We restrict ourselves to the much simpler task of describing what can be done when the multivariate normality assumption is met, recognizing (and hoping our readers recognize) that what we say may have no legitimate practical application. (Needless to say, we hope this last is not the case, and robustness studies such as that of Olson, 1974, seem to give some reason for such a hope.)

The first problem we consider is that of comparing, via hypothesis testing, two models in the hierarchy described above. The null hypothesis in this case may be thought of as the proposition that the more restricted model is correct (provides a perfect description of the population response means). The alternative is that the more general of the two models being compared is necessary to describe the means completely.

More specifically, suppose we can divide the parameters of the complete model into three sets:

> $\underset{\sim}{\alpha}$, the parameters which appear in the more restricted model,
>
> $\underset{\sim}{\beta}$, the parameters which appear in the more general model, but not in the more restricted one, and
>
> $\underset{\sim}{\gamma}$, the parameters which appear in the complete model only.

(In the case where the more general model happens to be the complete model, there will be no parameters $\underset{\sim}{\gamma}$.)

In these terms, the hypothesis testing problem can be stated as follows: in the complete model, test the hypothesis that $\underset{\sim}{\beta}=\underset{\sim}{0}$, given that $\underset{\sim}{\gamma}=\underset{\sim}{0}$ but making no assumptions about $\underset{\sim}{\alpha}$. Thus $\underset{\sim}{\gamma}=\underset{\sim}{0}$ is assumed, and $\underset{\sim}{\beta}=\underset{\sim}{0}$ is tested under this assumption.

The most common way to approach this problem involves an initial linear transformation of the repeated measures to a new set of dependent variables. The transformation we use here is the one which transforms the population response means into the parameters of the complete model. As was noted earlier, this transformation is unique. It also has the property that the population means of the new variables are identical to the parameters of the complete model. Thus we may partition the new variables into three sets, corresponding to the three sets of parameters defined above. We denote these sets of variables by $\underset{\sim}{a}$, $\underset{\sim}{b}$, and $\underset{\sim}{c}$, with population means $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$, and $\underset{\sim}{\gamma}$, respectively.

A direct test of the hypothesis $\underset{\sim}{\beta}=\underset{\sim}{0}$ applies Hotelling's $\underline{T}^2$ to the new variable set $\underset{\sim}{b}$ to test if their population means are all zero.

This is sometimes referred to as an "unweighted analysis." Formally, as a likelihood ratio test, it compares the complete model (parameters $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$, and $\underset{\sim}{\gamma}$) with the model (not in our hierarchy) having only parameters $\underset{\sim}{\alpha}$ and $\underset{\sim}{\gamma}$. Of course, if $\underset{\sim}{\gamma}=\underset{\sim}{0}$ this is just the comparison we want (namely, $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ with $\underset{\sim}{\alpha}$ only). Thus this unweighted analysis certainly provides a valid answer to our problem. It may, however, be conceptually unsatisfying in the sense that it makes no use of the assumption $\underset{\sim}{\gamma}=\underset{\sim}{0}$. Our main purpose in this section is to describe an approach which does use this assumption.

We now turn our attention to the likelihood ratio test for comparing the model having only $\underset{\sim}{\alpha}$ with the model which includes both $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$. In this test, the variables $\underset{\sim}{b}$ are again used as dependent variables. The set of variables $\underset{\sim}{c}$ (with population means $\underset{\sim}{\gamma}$) are, however, now used as covariates. Thus we test the hypothesis that the population means of the variables $\underset{\sim}{b}$, "adjusted" for the covariates $\underset{\sim}{c}$, are all zero. Since the population means of $\underset{\sim}{c}$ ($\underset{\sim}{\gamma}$) are all assumed to be zero, this new hypothesis is identical to the one we wanted to test, namely, $\underset{\sim}{\beta}=\underset{\sim}{0}$. Again, a form of Hotelling's $\underline{T}^2$ provides the appropriate test statistic.

Formally, the test we have just described is a one-sample multivariate analysis of covariance, which tests the hypothesis that the "constant term" (vector of adjusted grand means) is zero. Some caution must be used here, however, since many descriptions of the analysis of covariance begin by subtracting the grand means of the covariates from the individual covariate scores. When this is done, the grand means of the dependent variables are not adjusted for the covariates, whereas, in the analysis we have described, this <u>does</u> occur. (For a discussion of this point as it affects the use of the program MULTIVARIANCE, see the Appendix.)

Returning to the general hierarchy of models discussed earlier, one common interest is to select an appropriate model from the hierarchy. Thus, in the polynomial example, the interest may be in identifying the appropriate order of the polynomial. A procedure which has been suggested for achieving this involves the sequential use of the tests described above. More concretely, as a first step, the complete model is compared with the second model in the hierarchy. Here no parameters are assumed to be zero, so the set $\underset{\sim}{\gamma}$ is empty. Consequently, there are no covariates available from the set of new variables and the two tests (unweighted analysis and analysis of covariance) are identical. In this first comparison, if the null hypothesis can be rejected at some pre-selected level, the procedure stops and the complete model is identified as the appropriate one.

If the null hypothesis cannot be rejected, then the second and third models in the hierarchy are compared. Rejection of the (new) null hypothesis means identifying the second model as appropriate. Failure to reject leads to a comparison of the third and fourth models, and so on. If none of the null hypotheses in this sequence is rejected, then we are left with the last ("simplest") model in the hierarchy.

When unweighted analysis is used , this sequential procedure involves simply testing whether the means of successive sets of transformed variables are zero in the population. For the analysis of covariance approach, the set of covariates at each step consists of the dependent variables from all previous steps.

A special case of the above occurs when each model in the hierarchy differs from the adjacent models by only one parameter. (The set of polynomial models discussed earlier falls into this category.) Here all tests become univariate. The unweighted analyses yield successive one-sample *t*-tests on the means of the transformed variables. The analyses of covariance are equivalent to a series of step-down *F*-tests on the set of transformed variables. The order in which these step-down tests are carried out is critical, and (as we shall see in the next section) has been a source of confusion in the past. The first variable in the ordering should be the one associated with the parameter appearing only in the complete model. The mean of the second variable should be the parameter appearing only in the first two models in the hierarchy, and so on. In the polynomial example, the variable representing the highest order coefficient should be first and the variable representing the linear coefficient last (given that it is not usually of interest to test whether the constant term is zero). The step-down *F*-values should then be read in this same order, stopping with the first one which exceeds the corresponding critical value.

Once an (apparently) appropriate model in the hierarchy has been identified, the question of estimating parameters in this model arises. The simplest estimates for the parameters in our final model are just the sample means of the corresponding transformed variables. Their use parallels the unweighted approach to testing. The means of these same variables, *adjusted* for all the remaining variables (used as covariates), are the maximum likelihood estimates of the final model parameters.

They are obviously associated with the analysis of covariance (or likelihood ratio) approach to model comparisons.

Using the estimated standard errors for the sample means, or the adjusted standard errors for the adjusted means, together with critical values from the appropriate _t_-distributions, interval estimates of the model parameters may be constructed in either approach. In many cases, it may be useful to substitute point estimates for the parameters in the final model and thus reconstruct estimated means for the original responses to the different conditions or times. These will typically differ from the sample means, and an examination of the differences provides an important check on the adequacy of the final model.

While the above discussion applies in principle to the study of any hierarchy of models for repeated measures, there is an important practical point to be made regarding the polynomial example. Primarily for reasons of numerical precision, it is common to work with so-called orthogonal polynomial coefficients instead of the original coefficients (β_i) shown in (1). The _i_th orthogonal coefficient is actually a linear combination of β_i, β_{i+1}, ..., and β_{p-1} . Consequently, a test that the _i_th orthogonal coefficient is zero, given that all higher order coefficients are zero, is actually a test of $\beta_i=0$ given $\beta_{i+1}= \ldots =\beta_{p-1}=0$.

Although working with orthogonal polynomial coefficients has no adverse effect on testing, it should be remembered that the estimates one obtains are _not_ estimates of the original β_i . The latter may be obtained from the estimated orthogonal coefficients via a linear transformation, but in many cases it will be more useful to estimate the original means directly. We do this in the example of Section 4.

3. Historical developments

Two early discussions of the use of multivariate analysis of variance with repeated measures are given by Bock(1963) and Finn(1969). Both advocate a sequential search through a hierarchy of models, as discussed in the previous section. Moreover, they correctly describe unweighted analysis as one approach to testing in this context. When they turn to the use of analysis of covariance (and specifically step-down *F*-tests), however, both make the critical error of using the *wrong* covariates. Thus in terms of the sets of variables $\underset{\sim}{a}$, $\underset{\sim}{b}$, and $\underset{\sim}{c}$ of Section 2, they test the means of $\underset{\sim}{b}$ adjusted for the covariates $\underset{\sim}{a}$ (instead of $\underset{\sim}{c}$). In the step-down analysis, this means that they analyze the variables in exactly the reverse of the correct order. It also means that they test the wrong hypotheses. Instead of considering values of $\underset{\sim}{\beta}$ adjusted by linear combinations of $\underset{\sim}{\gamma}$ (all elements of which are assumed to be zero), Bock and Finn test $\underset{\sim}{\beta}$ adjusted by linear combinations of $\underset{\sim}{\alpha}$ (about which nothing has been assumed). Thus, for instance, a correct rejection of such a null hypothesis could be due to non-zero elements of $\underset{\sim}{\alpha}$, rather than non-zero $\underset{\sim}{\beta}$.

When considered in this light, step-down analysis in the "wrong order" clearly has nothing to recommend it. Unfortunately, it is also described (among other places) in Bock(1975) and Finn & Mattsson (1978). Moreover, there appears to be no recognition in the literature (other than a mild comment by Roskam, 1976, p. 120) of the basic flaw in such a wrong order analysis: namely that it tests the wrong hypotheses. We find this an unfortunate state of affairs and hope the present paper will at least serve to warn researchers to avoid such analyses.

Other work in the field includes that of Potthoff & Roy (1964), who developed a weighted least squares approach to the analysis of repeated measures data (or "growth curve" data, as it is called in the biometric literature). In their treatment, one has the freedom of selecting an arbitrary (symmetric, nonsingular) matrix of weights. If one choses the identity matrix for this purpose, the result is unweighted analysis. (Now we can see the reason for the name.) The optimal choice of a weight matrix would be the inverse of the matrix of population covariances among the responses. Unfortunately, this is never available to the researcher. Khatri (1966) took a likelihood approach to the growth curve problem and showed that maximum likelihood estimates were given by expressions equivalent to using the inverse of the sample covariance matrix as weights in the approach of Potthoff & Roy. He also showed how models could be compared based on likelihood considerations.

Rao (1965, 1966, 1967) developed the analysis of covariance approach described in Section 2, where variables corresponding to parameters not in the model are used as covariates. In these papers, he also considered the possibility of using only a subset of the available covariates. The (unknown) pattern of covariances among the responses, together with the number of observations in the sample, may very well be such that better estimates and more powerful tests can be obtained by dropping some (or even all) of the potential covariates. Thus, under some circumstances, an unweighted analysis (where all covariates have been dropped) may be more desirable -- from a statistical point of view -- than a complete analysis of covariance.

We avoided these issues (for the sake of simplicity) in Section 2 and do not wish to discuss them in any detail here. We find the analysis of covariance (using all available covariates) conceptually appealing, but realize that -- especially with small sample sizes -- its use may be undesirable or even impossible. To give a simple example, imagine estimating the parameters for a quadratic model applied to responses at ten time points from a sample of five subjects. Here there are seven covariates available (ten variables minus the three parameters of the quadratic model) but, if all were used, the standard errors for our estimates would have <u>negative</u> degrees of freedom (i.e., such an analysis is impossible). With unweighted analysis, on the other hand, there would still be four error degrees of freedom.

The work of Grizzle & Allen (1969) systematically describes, compares, and expands the results of Potthoff & Roy, Khatri, and Rao. An important comparison is summed up in the following quote (Grizzle & Allen, 1969, p. 362):

> Hence, the estimates obtained by the analysis of covariance, by maximum likelihood, and by weighting inversely as the estimated variance are identical.

What this does <u>not</u> say is that the tests, standard errors, and degrees of freedom obtained from Potthoff & Roy's approach (which we may simply call a "weighted analysis") do <u>not</u> agree with those from the analysis of covariance. This is because the weighted analysis is only developed for the case when the weight matrix is constant and not a function of the data. Consequently, it makes no allowance for the fact that we have used the sample covariance matrix, and that this will vary over repeated sampling. As a result, the tests, standard errors, and degrees of freedom from a weighted analysis will only be valid in the limit as the sample size approaches infinity. Since the analysis of covariance provides exact, small sample results, these should be used in preference to those of the weighted analysis.

This point arises in connection with a recent article (Bock, 1979) on the analysis of time-structured data which includes a section on multivariate analysis of repeated measures. It is also an issue for an option in MULTIVARIANCE, Version VI, and the accompanying discussion in the User's Guide (Finn,1978). Bock (1979) describes both the unweighted and weighted analyses of repeated measures data, and correctly notes that the latter does not give exact tests for finite samples. There is, however, no mention of analysis of covariance as a solution to this problem. Similarly, weighted analysis is available for the analysis of repeated measures in MULTIVARIANCE, Version VI, and the user is warned that it is only an asymptotically valid procedure (Finn, 1978, p. 48). The only discussion of analysis of covariance for repeated measures analysis does not occur in connection with the weighted analysis (Finn, 1978, p. 44), and, more seriously, still advises using what we have seen to be the wrong covariates (!)

Thus another goal for this paper is to inform researchers who find the idea of a weighted analysis attractive that they can obtain the same point estimates, together with exact tests, standard errors, and degrees of freedom by taking the analysis of covariance approach. (Again, see the Appendix for more details on doing this with MULTIVARIANCE.)

4. Example illustrating various analyses

To make some of the points in the previous sections more concrete, we reanalyzed a data set used by Grizzle & Allen (1969, p.359, Table 1). The data come from an unpublished medical study and are repeated measurements of coronary sinus potassium (in mil equivalents per liter) made on four groups of dogs following coronary occlusion. Measurements are made for each dog at 1, 3, 5, 7, 9, 11, and 13 minutes after occlusion (for a total of seven repeated measures). The dogs in the first group ($n_1 = 9$) serve as a control. Those in the remaining three receive various treatments prior to coronary occlusion, as follows:

2. extrinsic cardiac denervation three weeks prior ($n_2 = 10$),
3. extrinsic cardiac denervation immediately prior ($n_3 = 8$), and
4. bilateral thoracic sympathectomy and stellectomy three weeks prior ($n_4 = 9$).

We assume that the effects of the experimental treatments are unequal to zero, so the parameters for these effects were always included in the models that were compared. Several analyses were performed:

1. unweighted analysis (no assumptions about parameters not included in the models),
2. step-down tests in the wrong order,
3. step-down tests in the correct order,
4. weighted analysis (with the inverse of the sample within-groups covariance matrix used for weights), and
5. multivariate analysis of covariance.

Test results of the first four analyses are presented in Table 1, up to the first significant result. (We have left the exact choice of significance level purposely vague. The choice is only practically relevant in this example for the unweighted analysis, and here we have simply used .05 per test.) The goal of the analyses in Table 1 is to identify an appropriate polynomial model for the main effect of time (upper half of table) and for the time × treatment interaction. Note that the sextic term tests are identical for the weighted, unweighted, and correct-order step-down analyses. This will always be true for model comparisons involving the complete model (in this case the sixth order polynomial). The wrong-order sextic tests differ from the others because the linear through quintic terms have been used as covariates.

Further examination of Table 1 reveals that the remaining _F_-values from the weighted analysis are consistently slightly larger than the correct-order step-down _F_'s. The extent of agreement is an indication of the quality of the approximation implicit in the weighted approach.

TABLE 1. Test results for illustration using four methods of analysis

Constant given Treatments	unweighted			wrong order			right order			weighted		
	F	df	p<	F	df	p<	F	df	p<	F	df	p<
sextic	.02	1,32	.89	.03	1,27	.88	.02	1,32	.89	.02	1,32	.89
quintic	2.69	1,32	.12	2.41	1,28	.14	2.60	1,31	.12	2.68	1,32	.12
quartic	.00	1,32	.97	3.32	1,29	.08	.04	1,30	.84	.05	1,32	.83
cubic	4.07	1,32	.06	.01	1,30	.93	15.76	1,29	.01	18.89	1,32	.01
quadratic	14.28	1,32	.01	14.97	1,31	.01	----			----		
linear	----			----			----			----		
Treatments given Constant	**unweighted**			**wrong order**			**right order**			**weighted**		
	F	df	p<	F	df	p<	F	df	p<	F	df	p<
sextic	.24	3,32	.87	1.36	3,27	.28	.24	3,32	.87	.24	3,32	.87
quintic	.50	3,32	.69	1.11	3,28	.37	.48	3,31	.70	.50	3,32	.69
quartic	.60	3,32	.62	.38	3,29	.77	.63	3,30	.61	.69	3,32	.57
cubic	3.42	3,32	.03	1.30	3,30	.30	8.01	3,29	.01	9.07	3,32	.01
quadratic	----			2.13	3,31	.12	----			----		
linear	----			4.97	3,32	.01	----			----		

Comparison of the tests for the cubic term (both main effect and interaction) from the unweighted and right-order analyses reveals the apparently greater power of the latter (assuming the cubic model is appropriate) in this example. As expected, the wrong-order tests bear little relation to the others, with the exception of the unweighted test of the quadratic main effect of time. Here the only difference between the two is that the wrong-order analysis uses the linear variable as a covariate and, apparently, the adjustment due to this variable was minimal.

In this example, the results of the step-down analysis (in the correct order) point to the choice of a cubic polynomial model for both main effect and interaction. Therefore, we continue by estimating the parameters (and the associated standard errors) for that model. In a complete model, the parameter estimates and standard errors are identical for the weighted, unweighted and covariance analyses (since there are no covariates). However, if we estimate the parameters and standard errors in a more restricted model (such as the cubic polynomial in our example), the three methods will yield different results. (The wrong-order analysis will not be considered further in this discussion.)

In Table 2 the differences among the three methods are illustrated for two selected parameters of our final model. Examination of the point estimates reveals the identity of the values produced by weighted analysis and the analysis of covariance. The point estimates from the unweighted method, on the other hand, are identical to the values produced for these parameters by all three methods when fitting the complete model. The slightly smaller standard errors and larger degrees of freedom resulting from the weighted analysis(compared to the analysis of covariance results) are a consequence of the failure to take into account the variability introduced by the covariance adjustment.

TABLE 2. Parameter estimates in the cubic model using three methods

Constant term for comparing groups 1 and 3

Analysis	point estimate	standard error	df
unweighted	-1.26	.66	32
weighted	-1.50	.62	32
covariance	-1.50	.67	29

Quadratic term for comparing groups 1 and 4

Analysis	point estimate	standard error	df
unweighted	.20	.22	32
weighted	.34	.15	32
covariance	.34	.16	29

A major disadvantage of the transformation of the original responses to variables representing orthogonal polynomial coefficients (as occurred in this example) is that the resulting parameter estimates are not readily interpretable. The MULTIVARIANCE program (Version VI) allows the computation of estimated means when working with an incomplete model. These estimated means can then be compared with the original sample means as a further assessment of the fit of the model. In Table 3, the estimated means from the cubic model and the residuals (observed minus estimated) are presented. There appear to be no serious discrepancies.

TABLE 3. Estimated response means based on the cubic model (with residuals from observed sample means)

Time	Group 1	Group 2	Group 3	Group 4
1	4.14(-.03)	3.58(-.04)	3.55(.05)	3.57(.07)
3	4.19(-.01)	3.65(-.02)	3.70(.03)	3.74(.04)
5	4.51(.00)	3.66(-.04)	3.99(.21)	3.87(.14)
7	4.90(-.13)	3.63(-.07)	4.30(.14)	3.96(.11)
9	5.19(-.12)	3.58(-.02)	4.51(-.01)	4.01(-.03)
11	5.16(.06)	3.51(-.01)	4.52(.01)	4.01(.06)
13	4.64(.08)	3.46(.00)	4.21(.05)	3.98(.06)

In Figure 1, the estimated means are plotted.

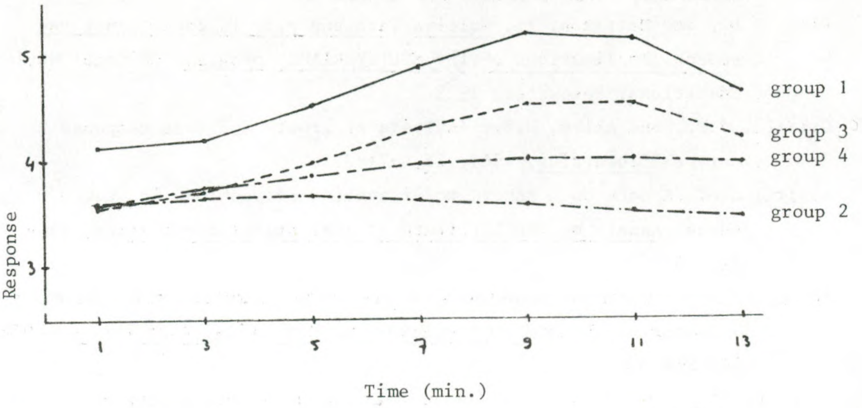

FIGURE 1. Estimated response means based on the cubic model

REFERENCES

Bock, R.D., Multivariate analysis of repeated measurements. In C.W. Harris (ed.), Problems in measuring change. Madison, Wisconsin: the University of Wisconsin Press, 1963, 85-103.

Bock, R.D., Multivariate statistical methods in behavioral research. New York: McGraw-Hill, 1975.

Bock, R.D., Univariate and multivariate analysis of time-structured data. In J.R. Nesselroade & P.B. Baltes (eds.), Longitudinal research in the study of behavior and development. New York: Academic Press, 1979, 199-231.

Finn, J.D., Multivariate analysis of repeated measures data, Multivariate Behavioral Research, 1969, 4, 391-413.

Finn, J.D., MULTIVARIANCE User's Guide. Chicago: National Educational Resources, 1978. (Version VI, Release 2)

Finn, J.D., and Mattsson, I., Multivariate analysis in educational research. Applications of the MULTIVARIANCE program. Chicago: National Educational Resources, 1978.

Grizzle, J.E., and Allen, D.M., Analysis of growth and dose response curves. Biometrics, 1969, 25, 371-381.

Khatri, C.G., A note on a MANOVA model applied to problems in growth curve. Annals of the Institute of Statistical Mathematics, 1966, 18, 75-86.

Olson, C.L., Comparative robustness of six tests in multivariate analysis of variance. Journal of the American Statistical Association, 1974, 69, 894-908.

Potthoff, R.R., & Roy, S.N., A generalized multivariance analysis of variance model useful especially for growth curve problems. Biometrika, 1964, 51, 313-326.

Rao, C.R., The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. Biometrika, 1965, 52, 447-458.

Rao, C.R., Covariance adjustment and related problems in multivariate analysis. In P.R. Krishnaiah (ed.), Multivariate Analysis, New York: Academic Press, 1966, 87-103.

Rao, C.R., Least squares theory using an estimated dispersion matrix and its application to measurement of signals. Proceedings of the Fifth Berkeley Symposium. Mathematical Statistics and Probability, Vol.I, 1967, 355-472.

Roskam, E.E., Multivariate analysis of change and growth: critical review and perspectives. In D.N.M. de Gruijter & L.J.T. van der Kamp (eds.) Advances in psychological and educational measurement, London: John Wiley & Sons, 1976, 111-133.

APPENDIX

The analyses in Section 4 were carried out with the help of the MULTIVARIANCE program, Version VI. In working with this program to get results such as ours,one must take account of a few peculiarities.

In the MULTIVARIANCE program, the sample design (groups or treatments in our case) and the design on the dependent variables (repeated measures over time in our case) are separately specified. When one specifies an incomplete model for the design on the dependent variables (the cubic model in our case), the point estimates of the constant in the sample design (corresponding to the main effects for the repeated measures) are not adjusted for the covariates when an analysis of covariance is performed, whereas in a weighted analysis with MULTIVARIANCE they are. However, the corresponding standard errors from the analysis of covariance are adjusted for the covariates. Thus the standard errors are what we want, but are not appropriate for the unadjusted estimates given by the program. As we have already mentioned, the estimated standard errors and degrees of freedom from the weighted analysis are only asymptotically valid.

Finally, to obtain estimated means of the original responses in MULTIVARIANCE, one must use the weighted analysis option. The estimated means obtained with the analysis of covariance are of the transformed variables only.

With these observations in mind, we may now construct the following outline for carrying out the sort of analysis described in Section 2 and illustrated in Section 4:

1. Carry out a step-down analysis of the transformed variables (starting with the complete model and working down the hierarchy) to assess the complexity of the appropriate model.
2. After making a decision about the appropriate model to be fitted, carry out a multivariate analysis of covariance on the variables associated with parameters in the chosen model, using the remaining variables as covariates . From this analysis, standard errors and degrees of freedom can be used, but not the point estimates (at least not those for the repeated measures main effects).
3. Carry out a weighted analysis, specifying for the design on the dependent variables the chosen model. Use only the point estimates of the parameters and the estimated response means from this analysis.