Statistics and Applications {ISSN 2454-7395 (online)}
Special Issue in Memory of Prof. C R Rao
Volume 22, No. 3, 2024 (New Series), pp 19–27
http://www.ssca.org.in/journal

Employing Rao Theorems in Mixed Effects Growth Curves

Samaradasa Weerahandi
X-Techniques, 23 Chestnut Street, Edison, NJ 08817

Received: 09 April 2024; Revised: 10 June 2024; Accepted: 12 June 2024

Abstract

This article is motivated by the author's pleasant experience when late Professor Rao helped validate an assertion made in Weerahandi and Berger (1999). Additional implications of Rao (1967) in Growth Curve Models under compound symmetric covariance structure are also presented. The inferences are made using the generalized p-value approach. Desired further research are also discussed.

Key words: Rao's covariance structure; Generalized inference; Compound symmetry; Generalized Hotelling T^2; Parametric bootstrap.

AMS Subject Classifications: 62K05, 05B05

1. Introduction

The author of this article had the pleasure of having a short chat with Professor Rao many years ago on the sidelines at few conferences in New Jersey and in Europe. At that time I had no idea that this world class researcher would care about a favor requested by a mediocre researcher like me. Now, to briefly describe the experience, first consider the following problem in the context of Mixed Effects models in Growth Curves, which is a particular problem involving repeated measures.

Consider the linear mixed effects growth curve model based on observations from n subjects

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta_i} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon_i} \qquad \text{for } i = 1, \ldots, n, \tag{1}$$

where $\mathbf{y}_i$ is the $T \times 1$ vector of responses from ith subject, $\mathbf{X}_i$ and $\mathbf{Z}_i$ are known design matrices of dimension $T \times p$ and $T \times q$, respectively, $\boldsymbol{\beta_i}$ is a vector of fixed effects, and the random effects, $\mathbf{b}_i$ and the error vector $\boldsymbol{\epsilon}$, jointly and independently distributed as

$$\mathbf{b}_i \sim \mathbf{N}_q(\mathbf{0}, \boldsymbol{\Psi}) \tag{2}$$

and

$$\boldsymbol{\epsilon}_i \sim \mathbf{N}_T(\mathbf{0}, \boldsymbol{\Lambda}_i)\ ,$$

Corresponding Author: Samaradasa Weerahandi
Email: Weerahandi@aol.com

where $\mathbf{\Lambda}_i$ is a within-subject covariance matrix of dimension $T \times T$ and $\mathbf{\Psi}$ is usually a between-subject covariance matrix of dimension $q \times q$. The model can also be rewritten in the form of a structured covariance matrix as

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta_i} + \mathbf{e}_i, \tag{3}$$

where

$$\mathbf{e}_i \sim \mathbf{N}_T(\mathbf{0}, \mathbf{\Lambda}_i + \mathbf{Z}_i \mathbf{\Psi} \mathbf{Z}_i').$$

When the covariance matrix of a growth curves model has a special structure, classical approaches do not provide exact solutions to inference problems even for a situation of a single growth curve. In this article we concentrate on the case of one group of subjects when the covariances follow compound symmetric structure, which is also known as the intraclass correlation structure.

2. Case of intraclass correlation structure

Weerahandi and Berger (1999) considered the particular case of one group of subjects when the covariance matrix is compound symmetric. In this section, we will concentrate on the distributional results providing details of Professor Rao's contribution. To do so, consider the simple growth curve model

$$Y_{it} = \alpha_i + \mathbf{X}_t' \boldsymbol{\beta} + \epsilon_{it}, \tag{4}$$

where $\mathbf{X}_t'$ is the $p \times 1$ design vector, $\boldsymbol{\beta}$ is a $p \times 1$ vector of parameters common for all subjects, α_i is a random effect due to subjects, and ϵ_{it} is the error term. In particular, when one deals with polynomial growth curves, the design matrix is of the form

$$\mathbf{X}_t' = \left(1, t, , t^2, \ldots, t^{p-1}\right)$$

If random effects are all normally distributed, we get

$$\alpha_i \sim N(0, \sigma_\alpha^2) \tag{5}$$

and

$$\epsilon_{it} \sim N(0, \sigma_e^2),$$

where σ_α^2 and σ_e^2 are variance components of the model. Moreover, α_i and all ϵ_{it} terms are assumed to be independently distributed. Collecting data from ith subject, the model for the $T \times 1$ vector of responses, $\mathbf{Y}_i$, can be written in vector form in terms of the $T \times p$ design matrix $\mathbf{X}$ as

$$\mathbf{Y}_i = \alpha_i \mathbf{1}_T + \mathbf{X}\boldsymbol{\beta} + \epsilon_i\ , \tag{6}$$

where $\mathbf{1}_T$ is a $T \times 1$ vector of 1s. It is easily seen from (5) that $Var(Y_{it}) = \sigma_\alpha^2 + \sigma_e^2$ and that $Cov(Y_{it}, Y_{it'}) = \sigma_\alpha^2$, and hence

$$\mathbf{Y}_i \sim N_T(\mathbf{X}\boldsymbol{\beta}, \mathbf{\Sigma}) \text{ with the covariance matrix } \mathbf{\Sigma} = \sigma_\alpha^2 \mathbf{1}_T \mathbf{1}_T' + \sigma_e^2 \mathbf{I}_T \tag{7}$$

This means that the covariance matrix of the observations vector has the intraclass structure. The model (6) is a special case of model (1) with

$$\mathbf{Z}_i\mathbf{\Psi}_i\mathbf{Z}_i' = \sigma_\alpha^2 \mathbf{1}_T\mathbf{1}_T' \quad \text{and} \quad \mathbf{\Lambda}_i = \sigma_e^2\mathbf{I}_T$$

Being a matrix with intraclass structure, the inverse of $\mathbf{\Sigma}$ is also an intraclass matrix. More specifically,

$$\mathbf{\Sigma}^{-1} = \sigma_e^{-2}\left[\mathbf{I}_T - \frac{\sigma_\alpha^2}{\sigma_e^2 + T\sigma_\alpha^2}\mathbf{1}_T\mathbf{1}_T'\right]. \tag{8}$$

The problem is to make inferences about the unknown parameters β and the variance components σ_α^2 and σ_e^2. It follows from (7) that the maximum likelihood estimate (MLE) of $\boldsymbol{\beta}$ is the weighted least-squares estimate (WLSE)

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\overline{\mathbf{Y}}, \tag{9}$$

which is also known as the generalized least squares estimate (GLSE) of β, where $\overline{\mathbf{Y}} = \sum \mathbf{Y}_i/N$ is a $T \times 1$ vector, where N is the number of subjects, who were observed over time.

Rao (1967) and Rao (1973) showed that, if the columns of $\mathbf{\Sigma X}$ is a subspace of the vector space spanned by the columns of $\mathbf{X}$, then the GLSE reduces to the ordinary least-squares estimate (OLSE), regardless of what $\mathbf{\Sigma}$ is. When $\mathbf{\Sigma}$ is as in (7) and the first column of $\mathbf{X}$ is a vector of 1's (i.e., an intercept term is present in the growth curve model), this condition is satisfied and consequently (9) reduces to the OLSE, A covariance matrix satisfying this condition is referred to as Rao's covariance structure; see also Ghosh and Gokhale (1987). Then, the point estimator of $\boldsymbol{\beta}$ is given by

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\overline{\mathbf{Y}} \tag{10}$$

3. Controversy about GLSE reducing to OLSE

When the author of this article submitted the manuscript underlying Weerahandi and Berger (1999) for publication in Biometrics, a referee disputed the validity of the distributional results outlined in the above section. The referee thought that Rao (1967) results do not imply that GLSE reduces to the OLSE under the compound symmetric covariances structure. When I referred to McElroy (1967) the referee did not concede and recommended rejection of manuscript. To overcome this dilemma, then I provided my own algebraic derivation, which is simpler to understand, but similar to McElroy (1967), even then the editor did not reconsider the manuscript.

Then, in desperation, I wrote to Professor Rao seeking help. To my surprise, in two weeks I received a letter in regular mail from Professor Rao stating something like "Weerahandi, not only your assertion is correct, but also it is valid under milder conditions and for greater class of covariance structures". When I sent the letter to the editor, she conceded and accepted the manuscript with some minor modifications. So, I am extremely grateful to late professor Rao for his support getting the article published.

4. Generalized inference

Before we review Weerahandi and Berger (1999) results of relevance, let us briefly describe the generalized tests introduced by Weerahandi (1987) and Tsui and Weerahandi (1989. In one-liner, generalized tests are based on random quantities known as Generalized Test Variables (GTV) that are functions of (i) observable random quantities, (ii) their observed values, and (iii) unknown parameters, defined in such a way that

(a). the distribution of GTV is free of unknown parameters, and

(b). at the observed sample points, the observed value of GTV will contain no unknown parameters under the null hypothesis. If a GTV is also monotonic for deviations from the null hypothesis, then it can be used to define extreme regions, on which generalized p-values can be based.

Often GTVs can be derived based on what is known as Generalized Pivotal Quantities (cf. Weerahandi (1993)), abbreviated as GPQs. To be specific, a GPQ of a parameter is also a function of (i) observable random variables, (ii) their observed values, and (iii) unknown parameters, defined in such a way that

(a). its distribution does not depend on nuisance parameters, and

(b). at the observed sample points, its observed value becomes equal to the parameter of interest.

Now getting back to the current problem, although GLSE reduce to OLSE under the compound symmetric covariance structures, even for models involving just one group of subjects, classical approach to inference fails to provide classical confidence bounds or tests of hypothesis concerning even a single parameter of the model. This is because the distribution of OLSE involves nuisance parameters. To be specific, despite the fact that GLSE is the same as the OLSE, the distribution $\widehat{\boldsymbol{\beta}}$ given by

$$\widehat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}/N) \tag{11}$$

involves the unknown variance components.

Nevertheless, Weerahandi and Berger (1999) demonstrated, how generalized tests can be constructed for testing hypotheses concerning one or more component of $\boldsymbol{\beta}$. To be specific. they considered the hypotheses on individual components of the form

$$H_0 : \beta_j \leq \beta_0$$

and provided a generalized test based on the independent sufficient statistics

$$\begin{aligned} \hat{\beta}_j &\sim N(\beta_j, \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}_{jj}/N) \quad j = 1, \cdots p \\ S_e^2 &= \sum_i \sum_t (Y_{it} - \mathbf{X}_t' \widehat{\beta} - (\overline{Y_i} - \overline{Y}))^2, \\ S_w^2 &= T \sum_i (\bar{Y}_i - \bar{Y})^2 \end{aligned} \tag{12}$$

due to Lehman (1986), where $\bar{Y}_i$ is the sample mean for i^{th} subject, $\bar{Y}$ is the sample mean of all the subjects,and $(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}_{jj}$ is the jjth element of the covariance matrix $(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$.

Let

$$S_j(\sigma_e^2, \sigma_w^2) = \frac{1}{\sqrt{N}}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})_{jj}^{1/2} \text{ and } \sigma_w^2 = \sigma_e^2 + T\sigma_\alpha^2. \tag{13}$$

The sums of squares appearing in (11) are distributed as

$$\begin{aligned} W_1 &= \frac{S_e^2}{\sigma_e^2} \sim \chi^2_{\nu_1}, \quad \text{where } \nu_1 = N(T-1) - p + 1, \text{ and} \\ W_2 &= \frac{S_w^2}{\sigma_w^2} \sim \chi^2_{\nu_2}, \quad \text{where } \nu_2 = N - 1. \end{aligned} \tag{14}$$

Then by taking the generalized approach to inference, Weerahandi and Berger (1999) showed that

$$p = \Pr\left(\frac{Z}{\sqrt{W/\nu}} \geq \sqrt{\nu}\frac{(b_0 - \beta_j)}{S_j(\frac{s_e^2}{B}, \frac{s_w^2}{1-B})}\right), \tag{15}$$

is a generalized p-value appropriate for testing the above null hypothesis, where

$$B \sim Beta(\nu_1/2, \nu_2/2) \text{ and } W = W_1 + W_2 \sim \chi^2_\nu : \quad \nu = \nu_1 + \nu_2 = NT - p$$

Although, they did not address the problem of interval estimation, one can construct generalized confidence Intervals on any single component using the Generalized Confidence interval approach suggested by Weerahandi (1993). Using the notion of Generalized Pivotal Quantity, one can also construct generalized confidence ellipsoids for few components of interest, as we demonstrate in the next section.

Taking that approach one can tackle problems involving more complicated compound symmetric covariance structures and number of groups of subjects, in a one-way layout setting as Chi and Weerahandi (1998) did. The Weerahandi and Berger (1999) results itself can be extended to make inferences on a number of regression coefficients, as we further discuss in the following sections.

5. Generalized inference on a vector of coefficients

Weerahandi and Berger (1999) results can be extended way beyond the problem they considered. Confining to the distributional results concerning Rao's covariance structure, consider the problem of constructing confidence regions on a subset of $\boldsymbol{\beta}$, say $\boldsymbol{\beta}_j$, a sub vector of $\boldsymbol{\beta}$, or $\boldsymbol{\beta}$ itself. The generalized inference on $\boldsymbol{\beta}_j$ can be constructed based on the foregoing distributional results along with the following:

$$\widehat{\boldsymbol{\beta}_j} \sim N(\boldsymbol{\beta}_j, (\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})_{jj}^{-1}/N), \tag{16}$$

where $(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})_{jj}$ is the jj^{th} subset of $(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})$ corresponding to the $\boldsymbol{\beta}$ coefficients of interest. Assuming positive definite covariance matrices, we can standardize (16) as

$$\mathbf{Z} = \sqrt{N}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})_{jj}^{1/2})\left(\widehat{\boldsymbol{\beta}_j} - \boldsymbol{\beta}_j\right) \sim N(\mathbf{0}, \mathbf{I}), \tag{17}$$

Now it is evident that, if the covariance matrix $\mathbf{\Sigma}$ were known, testing of hypotheses concerning $\boldsymbol{\beta}_j$ or confidence ellipsoids can be constructed using the χ^2 statistic,

$$\begin{aligned}\tilde{H} &= N\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_j\right)'\left(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})_{jj}\right)\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_j\right)\\ &= \left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_j\right)'(S_j^2(\sigma_e^2,\sigma_w^2)_{jj})\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_j\right)\sim\chi^2_{p_j} \qquad (18)\end{aligned}$$

where p_j is the dimension of $\boldsymbol{\beta}_j$

5.1. Hypothesis testing

Typically the covariance matrix is unknown, and in that case, the generalized inferences can be performed based on the *generalized Hotelling* T^2 statistic

$$H=\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_j\right)'(S_j^2(s_e^2/W1,s_w^2/W2)_{jj})\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_j\right), \qquad (19)$$

because

$$\frac{s_e^2}{W_1}\text{ is a GPQ for }\sigma_e^2\text{ and }\frac{s_w^2}{W_2}\text{ is a GPQ for }\sigma_w^2. \qquad (20)$$

First, to perform hypotheses testing concerning sub-vectors of coefficients $\boldsymbol{\beta}_j$, consider null hypotheses of the form

$$H_0:\boldsymbol{\beta}_j=\boldsymbol{\beta}_0,$$

where $\boldsymbol{\beta}_0$ is a certain hypothesized value. Under the null hypothesis, we get from (17)

$$\mathbf{Z}\sqrt{N}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})_{jj}^{1/2})\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_0\right)=\mathbf{Z}S_j(\sigma_e^2,\sigma_w^2)_{jj},\text{ where }\mathbf{Z}\sim N(\mathbf{0},\mathbf{I}). \qquad (21)$$

By taking advantage of the two results (21) and (18), we can define a potential GTV, a *generalized Hotelling* T^2 as

$$\begin{aligned}T^2 &= \left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_0\right)'S_j(\sigma_e^2,\sigma_w^2)_{jj}\left((S_j^2(\frac{s_e^2}{W1},\frac{s_w^2}{W2})_{jj}\right)^{-1}S_j(\sigma_e^2,\sigma_w^2)_{jj}\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_0\right)\\ &= \mathbf{Z}'\left(S_j^2(\frac{s_e^2}{W1},\frac{s_w^2}{W2})_{jj}\right)^{-1}\mathbf{Z}. \qquad (22)\end{aligned}$$

The above random quantity is indeed a GTV, because (i) it is distributed free of unknown parameters, (ii) being a Hotelling T^2 type statistic, it tends to increase for deviations from the null hypothesis, (iii) its observed value $\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_j\right)'\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_0\right)$ is free of nuisance parameters, namely the unknown variances. Therefore, the random quantity defined by (22) is indeed a valid GTV. Therefore, the hypothesis can be tested based on the generalized p-value

$$p = Pr(\mathbf{Z}'(S_j^2(\frac{s_e^2}{W_1},\frac{s_w^2}{W_2})_{jj}^{-1}\mathbf{Z})>\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_0\right)'\left(\widehat{\boldsymbol{\beta}_j}-\boldsymbol{\beta}_0\right)).$$

The p-value is easily computed by numerical integration or Monte Carlo integration, as we further describe below.

5.2. Confidence regions

Generalized confidence ellipsoids for $\boldsymbol{\beta}_j$ are easily computed based on the GPQ corresponding to the above GTV. For example, the $100\gamma\%$ generalized regions is constructed as follows. First, find the cdf of T^2 as

$$\begin{aligned} F_T(t) &= Pr\left(\left(\widehat{\boldsymbol{\beta}_j} - \boldsymbol{\beta}_j\right)' S_j^2((\sigma_e^2, \sigma_w^2)_{jj}^{1/2} (S_j^2(\frac{s_e^2}{W_1}, \frac{s_w^2}{W_2})_{jj}^{-1}) S_j^2((\sigma_e^2, \sigma_w^2)_{jj}^{1/2} \left(\widehat{\boldsymbol{\beta}_j} - \boldsymbol{\beta}_j\right) \leq t\right) \\ &= \left(\mathbf{Z}'(S_j^2(\frac{s_e^2}{W1}, \frac{s_w^2}{W2})_{jj}^{-1})\mathbf{Z} \leq t\right). \end{aligned} \tag{23}$$

Then, find the quantile q_γ such that $F_T(q_\gamma) = \gamma$.

The generalized confidence ellipsoid for $\boldsymbol{\beta}_j$ implied by the above results is

$$\left(\widehat{\boldsymbol{\beta}_j} - \boldsymbol{\beta}_j\right)' \left(\widehat{\boldsymbol{\beta}_j} - \boldsymbol{\beta}_j\right) \leq q_\gamma,$$

because at the observed sample points, mid terms of (23) cancel out, The computation is carried out as follows:

(a). Generate large number, say M, samples from $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I})$,

(b). Generate M random numbers from $\chi^2_{\nu_1}$ and $\chi^2_{\nu_2}$,

(c). Compute and sort the values of $\left(\mathbf{Z}'(S_j^2(\frac{s_e^2}{W1}, \frac{s_w^2}{W2})_{jj})\mathbf{Z}\right)$,

(d). Estimate the quantile q_γ as the $M\gamma^{th}$ value of the sorted data.

(e). Construct the generalized ellipsoid using the above formula.

6. Discussion

Further research is necessary to extend forgoing results to more complicated models and hypotheses. Of particular interest is RANOVA (repeated measures ANOVA) and RMANOVA (repeated measures MANOVA) type models involving a number of groups of subjects. Growth curve models involving a number of groups of subjects is a particular case of RMANOVA. Chi and Weerahandi (1998) provided some preliminary results on RMANOVA and provided some guidance on how to handle such problems as multiple comparisons, but did not directly address them. Moreover, there is a need to extend such results to Two-Way RMANOVA, when there are two factors of interest, say treatments groups and groups of subjects characterized by some subject attributes.

One may also consider other approaches to inference, such as the Parametric Bootstrap (PB) approach and the Generalized Fiducial (GF) approach. However, it should be noted, as argued by Ananda et al. (2022), that in most applications, these two methods tend to be subsets of the generalized inference approach. In other words, the latter can reproduce or beat PB based tests and GF based tests, as shown by Ananda et al. (2022).

Kurata (1998) provided a generalization of Rao's Covariance Structure. The results in that article provided distribution theory necessary to tackle greater class of applications combined with generalized approach to inference to handle nuisance parameters.

In a slightly different context, Ghosh and Sinha (1980) studied the criterion robustness of the standard likelihood ratio test (LRT) under the multivariate normal regression model and also the inference robustness of the same test under the univariate set up for certain non-normal distributions of errors. Restricting attention to the normal distribution of errors in the context of univariate regression models, they derived conditions on the design matrix under which the usual LRT of a linear hypothesis (under homoscedasticity of errors) remains valid if the errors have an intraclass covariance structure. The conditions hold in the case of some standard designs. For further related results, the reader is referred to Rao (1967), Zyskind (1967), and Mukhopadhyay and Sinha (1980).

Acknowledgements

I am grateful to the Editors for their guidance and counsel. I also wish to thank the reviewers for their valuable comments and suggestions.

References

Ananda, M. A., Dag, O., and Weerahandi, S. (2023). Heteroscedastic two-way ANOVA under constraints. *Communications in Statistics-Theory and Methods*, **52**, 8207–8222.

Chi, E. M. and Weerahandi, S. (1998). Comparing treatments under growth curve models: exact tests using generalized p-values. *Journal of Statistical Planning and Inference*, **71**, 179–189.

Ghosh, M. and Sinha, B. K. (1980). On the robustness of least squares procedures in regression models. *Journal of Multivariate Analysis*, **10**, 332–342.

Ghosh, S. and Gokhale, D. V. (1987). Estimation and tests for departures from rao-structured covariance matrices. *Biometrical Journal*, **29**, 269–275.

Kurata, H. (1998). A generalization of Rao's covariance structure with applications to several linear models. *Journal of Multivariate Analysis*, **67**, 297–305.

Lehman, E. L. (1986). *Testing Statistical Hypothesis*. John Wiley and Sons, Inc., New York.

McElroy, F. W. (1967). A necessary and sufficient condition that ordinary least-squares estimators be best linear unbiased. *Journal of the American Statistical Association*, **62**, 1302–1304.

Mukhopadhyay, B. B. and Sinha Bikas, K. (1980). A note on the result of M. Ghosh and Bimal K. Sinha: "On the robustness of least squares procedures in regression models". *Calcutta Statistical Association*, **29**, 169–171.

Rao, C. R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 355–372. Berkeley.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York.

Tsui, K. and Weerahandi, S. (1989). Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. *Journal of the American Statistical Association*, **84**, 602–607.

Weerahandi, S. (1987). Testing regression equality with unequal variances. *Econometrica: Journal of the Econometric Society*, **1**, 1211–1215.

Weerahandi, S. (1993). Generalized confidence intervals. *Journal of the American Statistical Association*, **88**, 899–905.

Weerahandi, S. (2004). *Generalized Inference in Repeated Measures: Exact Methods in MANOVA and Mixed Models*. Wiley.

Weerahandi, S. and Berger, V. W. (1999). Exact inference for growth curves with intraclass correlation structure. *Biometrics*, **55**, 921–924.

Xu, L. and Tian, M. (2016). Parametric bootstrap inferences for panel data models. *Communications in Statistics-Theory and Methods*, **46**, 5579–5594.

Zyskind, G. (1967). On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *The Annals of Mathematical Statistics*, **38**, 1092–1110.