CONSTRUCTION OF SPLIT PLOT DESIGN FOR ESTIMATING VARIANCE COMPONENT

Oluwole Nuga, MSc
Bells University of Technology, Nigeria
G.N Amahia, PhD
University of Ibadan, Nigeria
Fakorede Aderogba
Bells University of Technology

Abstract

The work is aimed at studying design effect for the maximum likelihood estimators of variance components in a split plot design. The study used the general linear model with one whole plot factor and one sub-plot factor and assumed that both factor effects are random variables. The main problem studied is how to assign a given number of whole plots with equal sizes to the level of the whole plot factor in a way that will form a balanced one way design. The work introduced a method of classifying the five variance components to make comparison and presentation meaningful. The resulting optimal designs depend on the true proportional value of the variance components.

Keywords: Local Optimality, D-Optimality, Whole plot, Sub plot, Variance Component

Introduction

Optimal designs for variance components model have been discussed fairly in experiment that are ran in a completely random order. Most of the published work dates back to the 60’s and 70’s and have been restricted to specific models namely, one-way random model, the two-way crossed classification random model and the two way nested model. R.L Anderson and many of his co-workers are the main contributors to the design area during that period (Anderson 1975, 1981). . For the one way model, Hammersly (1949), Crump (1954), Anderson& Crump (1967) were some of the earliest authors. Hammersly (1949) showed that for a fixed N, the variance $Var(\sigma_{\alpha}^{2})$ is minimized by allocating an equal number n, of observation to each class where $n = \dfrac{N\rho + N + 1}{N\rho + 2}$, since this formula may not yield an integer value, it was suggested that the closest integer value for n be chosen. Crump (1954) and Anderson & Crump (1967) showed that for fixed K and N, $Var(\sigma_\alpha^2)$ is minimized when $n_i = n = N/a$ for all i. The optimal value for a in this case is given as $a_1 = \dfrac{N(N\rho+2)}{N(\rho+1)+1} = \dfrac{N}{n}$

Other authors are Kussmaul & Anderson (1967), Thompson and Anderson (1975), Herrendofer (1979), Murkerjue &Huda (1988), Giovagnoli & Sebastiani (1989), Norell (2006). Norell (2006) studied design effect for the one way random model of the maximum likelihood estimators.

The construction of optimal design for the two way crossed models seems to have been considered first by Gaylor (1960). He considered the problem of optimal designs to estimate variance components using the fitting constant method of estimation of variance components for the unbalanced data. Bush (1962) and Bush and Anderson (1963), HIrotsu(1966), Mostafa (1967) are some of the other contributors to the designing experiment using the two way random model.

Some pioneering articles that address the problem of estimating variance components in a nested classification are Bainbridge (1965) Prairie (1962), Prairie and Anderson (1962), Bainbridge (1965), they proposed designs that systematically spread the information in the experiment more equally among the variance components. Goldsmith and Gaylor (1970) carried out extensive investigation on optimal designs for estimating variance components in a completely random nested classification. Delgado (1999) defined a class of unbalanced design for estimating variance component in the three stage nested classification using the ANOVA method of estimation.

Loeza-Serrano. S and A .Donev (2012) constructed D- optimal design for variance components estimation in a three stage crossed and nested classification.

For experiments that include both crossed and nested factor in the same model, no assumption of a complete random model has been made. Work that design experiment for variance component estimation are based on the linear mixed effect model .Beverly (1981) , Ankenman, Liu, Karr, and Picka (2001) and Aviles and Pinheiro (2001) are authors that have published work. However experiments that complete randomization of order of runs is not feasible or might be too expensive to use is performed using split plot models.

Split Plot Design

Split plot designs initially developed by Fisher (1925) for use in agricultural experiments are basically the modified form of randomized block designs. These designs are used in situations where complete randomization of runs within block is not possible. These designs are used widely in industrial experiments, experiments where one set of factors may require a large amount of experimental materials(Whole Plot factors), while another set of factors might be applied to smaller experimental materials (Sub Plot factors). Another situation that leads to the use of split plot designs in industrial experiment is when there exist one or more factors called hard to change factors, that are expensive or time consuming to change level settings (WP factors) and the other factors (SP factors), whose level settings are easier to change are called easy to change factors.

In general split plot design can be used for any experimental situation that involves two different types of experimental unit (large and small), randomly assigned independently at the two different levels.

The optimal design for split-plot experiments has received attention by Goos and Vandebroek (2001b, 2003a, 2004). The work was based on a more complex design structure that used the first and second order polynomial model to represent the response, but in general the work used a linear mixed effect model which assumed fixed effects for the settings of the whole plot factors and sub-plot factors and two variance components associated with the whole plot error and sub-plot error.

The work is aimed at constructing D optimal designs for maximum likelihood estimators in a split plot experiment with one whole plot (WP factors) and one sub plot (SP factor), with the assumption of random effect for both factors.

Model and variance structure

The model equation for the split plot design with one WP factor (Factor A) and one SP factor (Factor B) can be written as,

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_{ik} + e_{ijk} \qquad i = 12.......a\ ,\ \ j = 1.....b\ ,\ k = 1.......r_i\ ,$$

r_i is the number of whole plot at i_{th} level of factor A, but $r_1 = r_2 ..... = r_a = r$ since equal number of whole plot are allocated to whole plot factor A . There are ar whole plot of equal sizes available.

y_{ijk} is the response at the k_{th} replicate of the i_{th} level of factor A and the j_{th} level of factor B, μ is the general mean, α_i is the effect of the i_{th} level of factor A , β_j is the effect of the j_{th} level of factor B. $(\alpha\beta)_{ij}$ is the interaction effect of the i_{th} level of factor A and the j_{th} level of factor B.

γ_{ik} is the error term of the k_{th} replicate of the i_{th} level of factor A (WP error term), e_{ijk} is the error term corresponding to individual y_{ijk} (SP error term). The random variables
α_i, β_j $(\alpha\beta)_{ij}$ γ_{ik} e_{ijk} and are assumed to be normally distributed with zero mean variance $\sigma_\alpha^2, \sigma_\beta^2, \sigma_{\alpha\beta}^2, \sigma_\gamma^2, \sigma_e^2,$ respectively. In matrix form the model can be written as,

$$Y = \mu 1 + Z_1\gamma_1 + Z_2\gamma_2 + .Z_3\gamma_3 + Z_4\gamma_4 + Z_5\gamma_5$$

is a vector of abk observations, μ is the overall mean, Z_i is an indicator matrix associated with the *ith* variance component, γ_i is a vector of normally distributed random effects associated with the *ith* variance component such that $w_i \approx N(0, \sigma_i^{\,2} I)$. The variance matrix of observations can be written as,

$$Var(Y) = V = \sum_{i=0}^{5} \sigma_i^2 Z_i Z_i'$$

The Z are defined as the kroneker product as follows,

$$Z_0 = \mathrm{I}_a \otimes \mathrm{I}_b \otimes \mathrm{I}_r \otimes \mathrm{I}_n \qquad Z_0 Z_0' = \mathrm{I}_a \otimes \mathrm{I}_b \otimes \mathrm{I}_r \otimes \mathrm{I}_n$$

$$Z_1 = \mathrm{I}_a \otimes 1_b \otimes 1_r \otimes 1_n \qquad Z_1 Z_1' = \mathrm{I}_a \otimes J_b \otimes J_r \otimes J_n$$

$$Z_2 = 1_a \otimes \mathrm{I}_b \otimes 1_r \otimes 1_n \qquad Z_2 Z_2' = J_a \otimes \mathrm{I}_b \otimes J_r \otimes J_n$$

$$Z_3 = \mathrm{I}_a \otimes \mathrm{I}_b \otimes 1_r \otimes 1_n \qquad Z_3 Z_3' = \mathrm{I}_a \otimes \mathrm{I}_b \otimes J_r \otimes J_n$$

$$Z_4 = \mathrm{I}_a \otimes 1_b \otimes \mathrm{I}_r \otimes 1_n \qquad Z_4 Z_4' = \mathrm{I}_a \otimes J_b \otimes \mathrm{I}_r \otimes J_n$$

Large Sample Variance of the Maximum Likelihood Estimators

Maximum likelihood estimates of variance components cannot be obtain explicitly except for some balanced data, but their large sample asymptotic dispersion matrix can be derived. It is known that the large sample asymptotic dispersion matrix of the maximum likelihood estimators for any model is the inverse of the information matrix. This matrix is the negative of the expected value of the second order partial derivatives (Hessian Matrix) with respect to the parameters of the log-likelihood function.

For our data vector $y \approx N(X\beta, V)$, the likelihood function is,

$$L = L(\beta, V / y) = \frac{e^{-\frac{1}{2}(y - X\beta)' V^{-1}(y - X\beta)}}{(2)^{\frac{1}{2}N} \left| V^{\frac{1}{2}} \right|}$$

Upon taking the log of the likelihood

$$l = LogL = -\frac{1}{2} N \log 2\pi - \frac{1}{2} \log |V| - \frac{1}{2}(y - X\beta)' V^{-1}(y - X\beta)$$

The matrix is given as

$$\begin{bmatrix} \partial^2 l / \partial\beta\partial\beta' & \partial^2 l / \partial\beta\partial\sigma^{2'} \\ \partial^2 l / \partial\beta\partial\sigma^{2'} & \partial^2 l / \partial\sigma^2\partial\sigma^{2'} \end{bmatrix}$$

which gives the following

$$In\begin{bmatrix} \beta \\ \sigma^2 \end{bmatrix} = \begin{bmatrix} X' V^{-1} X & 0 \\ 0 & \frac{1}{2} tr(V^{-1} Z_i Z_i' V^{-1} Z_j Z_j') \end{bmatrix} i, j = 1........5$$

the model for this work involve only the variance components and therefore the information matrix becomes

$$In\left[\sigma^2\right] = \frac{1}{2}\left[\{_m tr(V^{-1} Z_i Z_i' V^{-1} Z_j Z_j')\}\right] i, j = 1.....5.$$

Using Searle (1970) this reduces to

$$\frac{1}{2}\left[\{_m tr(V^{-1} Z_i Z_i' V^{-1} Z_j Z_j')\}\right] = \frac{1}{2}\left[\{_m sesq(Z_i' V^{-1} Z_j)\}\right] i, j = 1.....5.$$

The inverse of v is also obtained using the results of Henderson and Searle (1979),

$$V^{-1} = \theta_0^{-1}(I_a \otimes C_b \otimes C_r \otimes C_n) + \theta_1^{-1}(C_a \otimes C_b \otimes \bar{J}_r \otimes \bar{J}_n) + \theta_2^{-1}(I_a \otimes \bar{J}_b \otimes C_r \otimes \bar{J}_n) + \theta_3^{-1}(C_a \otimes$$

$$+ \theta_4^{-1}(\bar{J}_a \otimes C_b \otimes \bar{J}_r \otimes \bar{J}_n) + \theta_5^{-1}(\bar{J}_a \otimes \bar{J}_b \otimes \bar{J}_r \otimes \bar{J}_n)$$

Where

$\theta_0 = \sigma_e^2, \theta_1 = (\sigma_e^2 + bn\sigma_\gamma^2)$, $\theta_2 = (\sigma_e^2 + rn\sigma_{\alpha\beta}^2)$, $\theta_3 = (\sigma_e^2 + rn\sigma_{\alpha\beta}^2 + arn\sigma_\beta^2)$

$\theta_4 = (\sigma_e^2 + bn\sigma_\gamma^2 + rn\sigma_{\alpha\beta}^2 + brn\sigma_\alpha^2)$, $\theta_5 = (\sigma_e^2 + bn\sigma_\gamma^2 + rn\sigma_{\alpha\beta}^2 + brn\sigma_\alpha^2 + arn\sigma_\beta^2)$

a computational information matrix can thereafter be obtained ………Searle et al (2006) pg 247.

$$\begin{bmatrix} \sigma_e^2 \\ \sigma_\gamma^2 \\ \sigma_{\alpha\beta}^2 \\ \sigma_\beta^2 \\ \sigma_\alpha \end{bmatrix} = \begin{bmatrix} t_{ee} & t_{\gamma\gamma}/bn & t_{\alpha\beta}/rn & t_{\beta\beta}/arn & t_{\alpha\alpha}/brn \\ & t_{\gamma\gamma} & t_{\alpha\alpha}/r & abrn^2/\theta_5 & t_{\alpha\alpha}/r \\ & & t_{(\alpha\beta)^2} & t_{\beta\beta}/a & t_{\alpha\alpha}/b \\ & symmetric & & t_{\beta\beta a\backslash} & ab(rn)^2/\theta_5 \\ & & & & t_{\alpha\alpha} \end{bmatrix}$$

$$t_{ee} = \frac{v_e}{\theta_0^2} + \frac{v_{\gamma\gamma}}{\theta_1^2} + \frac{t_{\alpha\beta}}{(rn)^2} \qquad t_{\alpha\alpha} = b^2(rn)^2\left(\frac{v_\alpha}{\theta_4^2} + \frac{1}{\theta_5^2}\right)$$

$$t_{\beta\beta} = a^2(rn)^2\left(\frac{v_\beta}{\theta_3^2} + \frac{1}{\theta_5^2}\right)$$

$$, t_{\alpha\beta} = (rn)^2\left(\frac{v_{\alpha\beta}}{\theta_2^2} + \frac{v_\beta}{\theta_3^2} + \frac{v_\alpha}{\theta_4^2} + \frac{1}{\theta_5^2}\right) \qquad , t_{\gamma\gamma} = (bn)^2\left(\frac{v_\gamma}{\theta_1^2} + \frac{v_\alpha}{\theta_4^2} + \frac{1}{\theta_5^2}\right)$$

Where

$$v_e = a(b-1)(r-1), v_\gamma = a(r-1), v_{\alpha\beta} = (a-1)(b-1), v_\beta = b-1, v_\alpha = a-1$$

Design optimality and generation

No closed form analytical expression is available for the variance covariance matrix in this linear random effect model; we examine optimality using the asymptotic variance covariance matrix. A design from a group of designs with the same number and sizes of whole plots is said to be optimal if it minimizes an optimality criterion related to the variance- covariance matrix of the parameter estimates. Equivalently, we seek the design that maximizes an optimality criterion related to the information matrix of the five variance components. Optimal design in a linear random effects model depends on the relative size of the true values of the variance components, and we will not be able investigate optimality unless an assumption is made on the true values of the variance components. Since optimality for such models is similar to that of nonlinear models, we borrow an idea from optimization on theory of nonlinear models and use the local optimality. The five variance components were classified into two sets, the first set consist of the main effects and Interaction variance components (MIVC), which consist of σ_α^2, σ_β^2, and $\sigma_{\alpha\beta}^2$. The second set include the whole plot and sub plot error variance components (WSEVC) σ_γ^2 and σ_e^2.

The work will initially assign proportional value of variance components to the two sets in such a way that the sum equals one, and thereafter distribute proportional value to each set based on initially allocation. As an example

MIVC= $\sigma_{\alpha}^{2}+\sigma_{\beta}^{2}+\sigma_{\alpha\beta}^{2}=0.5$ WSEVC= $\sigma_{\gamma}^{2}+\sigma_{e}^{2}=0.5$, sub assigning MIVC and WSEVC, we have

$\sigma_{\alpha}^{2}=0.05,\ \sigma_{\beta}^{2}=0.40,\ \sigma_{\alpha\beta}^{2}=0.05$ and WSEVC $=\sigma_{\gamma}^{2}=0.42,\ \sigma_{e}^{2}=0.08$.

As stated earlier, the optimality of the design depends on the proportion of the true value of the variance components and not the total variance components. Proportional value of the variance components will be assigned to each of the five variance components. We employed this approach to make comparison easy. We wish to state that the proportional value of variance components used for comparison in the work is by no means exhaustive, what is required is the knowledge of the true values of variance components. Our approach however, will enable us make some statements about the choice of optimal design for this model.

A MATLAB code was written in the context of the information matrix of section (2.1) in such a way that enumerated design for a particular number and sizes of whole plot can be compared based on any configuration of the true values of the variance components.

Design Generation

The work intend to assign $ar = R$ whole plots of equal size randomly to the level of the whole plot factor such that equal number of whole plots is assigned to individual levels of whole plot factor. Each level of the sub plot factor is applied once within each whole plot and one observation is measured within each sub plot. The resulting design structure is balanced.

The number of possible designs (design space) equals the total number of ways to partition $ar=\sum_{i}^{a} r_i$ subject to $r_i = r_2 = ... = r_a$ and $2 \le a < ar$. The assignment of $ar = R$ whole plots to the levels of whole plot factor formed a balanced one way design.

As an example, consider six (6) whole plots of size two, the size of the whole plot equals the number of levels available for sub plot factor B(In this case b=2). The number of possible designs equals the number of ways of partitioning 6 i.e. a=2, r=3 [2, 2, 2] and a=3 r=2 [3, 3], we list the possible designs for various number of whole plot between 6 and 20 that satisfy the condition for balancedness.

Number of whole plot	Possible Designs	Number of whole plot	Possible Designs
6	[a=3, r=2][a=2, r=3]	15	[a=3,r=5][a=5,r=3]
8	[a=4, r=2][a=2,r=4]	16	[a=2,r=8][a=8, r=2][a=4,r=4]
10	[a=5,r=2][a=2,r=5]	18	[a=2,r=9][a=9,r=2][a=3,r=6][a=6,r=3]
12	[a=6,r=2][a=2,r=6] [a=3,r=4][a=4,r=3]	20	[a=2,r=10][a=10,r=2][a=4,r=5][a=5,r=4]
14	[a=7, r=2][a=2, r=7]		

An Algorithm

(1) List the possible design for a fixed number and size of whole plot (Generate the design space)
(2) Specify the available information about individual variance components. i.e. proportional value available to MIVC and WSEVC. Such that MIVC+WSEVC=1
(3) Redistribute the proportional value above to the variance components within each set.
(4) Calculate the criterion of optimality(D-optimal) for all design in the design space using the MATLAB programme code and identify the D-optimal design
(5) By making no changes to MIVC and based on the D- optimal design above, one can obtain the regions where some other designs in the design space is optimal. i.e. When $a \leq r$ based on the optimal design identified in (4) increasing the proportional value of σ_{γ}^{2} by 0.01 and decreasing the proportional value of σ_{e}^{2} by 0.01 in sequence until another design in the **design** space is optimal. When $a \geq r$ based on the optimal design obtain in (4) decreasing the proportional value of σ_{γ}^{2} by 0.01 and increasing the proportional value of σ_{e}^{2} by 0.01 in sequence until another design in the design space is optimal.

Example: Consider the situation where we have 6 whole plot of size 2, following the steps in the algorithm the list of designs in the design space is {a=2, b=2, r=3} and {a=3. b=2, r=3}

MIVC=0.5 and WSEVC=0.5 by (2) and applying (3) above individual variance components as,

$$\{\sigma_{\alpha}^{2} = 0.05, \sigma_{\beta}^{2} = 0.40, \sigma_{\alpha\beta}^{2} = 0.05, \sigma_{\gamma}^{2} = 0.30, \sigma_{e}^{2} = 0.20\}$$

Using the MATLAB code, the D-optimal is {a=2, b=2, r=3}. For this design we increase the proportion value of the whole plot error from 0.30 and reduce the proportional value of the sub plot error from 0.20 in sequences by 0.01, there is a change in the D-optimal design at a certain configuration. In general for a fixed MIVC, the range of proportional value for which the two designs are optimal is given below.

Designs	Whole plot	Sub plot
{a=2, b=2, r=3}.	$0.01 \le \sigma_{\gamma}^{2} \le 0.42$	$0.08 \le \sigma_{e}^{2} \le 0.49$
{a=3, b=2, r=2}.	$0.43 \le \sigma_{\gamma}^{2} \le 0.49$	$0.01 \le \sigma_{e}^{2} \le 0.07$

D- Optimal designs for some selected number and sizes of whole plot are presented in the appendix.

Conclusion

After empirically comparing designs with the same number and sizes of whole plot, these two wide-ranging statements can be made.

(1) For increasing proportional value of $\sigma_{\alpha\beta}^{2}$ and σ_{α}^{2}, designs wit $a \ge r$ is optimal.

(2) Optimal design is robust to increase in value of the sizes of the whole plot. i.e. If a design is optimal for a smaller size of whole plot then the same design is optimal for higher values of the whole plot sizes except for negative determinant optimal design.

The extension of the work to situations where there is an unbalanced one way structure in the assignment of whole plot to the whole plot factor is a subject of current research by the authors.

Appendix

Number Of Whole Plot	Size	Initial proportion	Prop of variance components	D-Optimal designs	Range for optimality for the D-optimal design
6	2	MIVC=0.5 WSEVC=0.5	$\{\sigma_{\alpha}^{2} = 0.05, \sigma_{\beta}^{2} = 0.40, \sigma_{\alpha\beta}^{2} = 0$ $\sigma_{\gamma}^{2} = 0.30, \sigma_{e}^{2} = 0.20\}$	{a=2,b=2, r=3}.	$0.08 \le \sigma_{e}^{2} \le 0.49$ $0.01 \le \sigma_{\gamma}^{2} \le 0.42$
	5			{a=2,b=5, r=3}.	$0.02 \le \sigma_{e}^{2} \le 0.49$ $0.01 \le \sigma_{\gamma}^{2} \le 0.48$
12	2	MIVC=0.95 WSEV=0.05	$\{\sigma_{\alpha}^{2} = 0.05, \sigma_{\beta}^{2} = 0.85, \sigma_{\alpha\beta}^{2} = 0$ $\sigma_{\gamma}^{2} = 0.01, \sigma_{e}^{2} = 0.04\}$	{a=3,b=2, r=4}.	$0.03 \le \sigma_{e}^{2} \le 0.04$ $0.01 \le \sigma_{\gamma}^{2} \le 0.02$
	5			{a=6,b=5, r=2}.	$0.01 \le \sigma_{\gamma}^{2} \le 0.04$ $0.01 \le \sigma_{e}^{2} \le 0.04$

For the ranges of value that were not captured by the optimal design, some other designs in the design space are optimal. For example when R=12 and b=2, for the same proportional value in MIVC,

{a=6,b=2, r=2} is the D-optimal design when $0.03 \leq \sigma_{\gamma}^{2} \leq 0.04$, $0.01 \leq \sigma_{e}^{2} \leq 0.02$

When R=12 and b=2, for same MIVC, no other range of value exist for other designs to be optimal {a=6,b=5, r=2} is exhaustive for all possible range of WSEVC.

References:

Anderson R.L and Crump P.P (1967), Comparisons of designs and estimation procedures for estimating parameters in two stage nested process, Technometrics, 9, 499-516

Ankenman, B. E., Liu, H., Karr, A. F. and Picka, J. D. (2001). A class of experimental designs for estimating a response surface and variance components. *Technometrics* **44**, 45-54.

Ankermann B. E., Avilles, A.I ., Pinheiro J.C., (2003), Optimal Design for Mixed Effects Models With Two Random Nested Factors, Statistica Sinica,13, 385-401

Bainbridge T.R (1965), Staggered nested designs for estimating Variance components, Indutr. Quality Control, 22-1, 12-20

Crump S.L (1951) The present Status of variance components Analysis, Biometrics, 7, 1-16

Delgado, J., Iyer, H., 1999. Search for optimal designs in a three stage nested random model. Statistics and Computing 9, 187–193.

Giovagnoli& Sebastiani (1989), Experimental Design for mean and variance estimation in variance component models, Computational Statistics. Data Analysis.8, 21-28

Goldsmith, C. H. and Gaylor, D. W. (1970). Three stage nested designs for estimating variance components. *Technometrics* **12**, 487-498.

Goos, P., Vandebroek, M., 2003, D-optimal Split-Plot Designs with Given Numbers and Sizes of Whole Plots, Technometrics, 45, 235-245

Goos, P., Vandebroek, M., 2001, Optimal Split-Plot Designs, Journal of Quality Technology,

Graybill, F.A (1976) Theory and Application of Linear Models ,Duxbury, North Scituate, MA.

Hammersley J.M (1949), The Unbiased estimate and standard error of interclass Varince. Metron, 15, 189-205

Khuri A.I.,(2000), Designs for Variance Components Estimation, International Statistica Review68, 311-322

Loeza S, Donev A.N (2014). Construction of Experimental Design to estimate Variance Components. Computational Statistics and Data Analysis 71, 1168–1177.

Norell L. (2006) Optimal Designs for Maximum Likelihood Estimators in One way Random Model, U.U.D.M report 2006: 24, Department of Mathematics, Upsalla University

Sahai H, Ojeda M., 2004 Analysis of Variance for Random Models. Vol.1. Theory,Applications and Data Analysis, Birhauser Boston.

Sahai H, Ojeda M., 2004 Analysis of Variance for Random Models. Vol.2. Theory,Applications and Data Analysis, Birhauser Boston.

Searle, S. R., Casella, G. and McCulloch, C. E. (1992). *Variance Components*. Wiley, NewYork

Searle S.R, Harold V. Anderson (1979) Dispersion matrices for Variance Components Models, Journal of American Statistical Association. Vol. 74,No 366,465-470