生物数学学报
J. Biomath.
7(1)(1992), 1—11

MIXED MODEL APPROACHES FOR ESTIMATING GENETIC VARIANCES AND COVARIANCES

Zhu Jun

(Zhejiang Agricultural University, Hangzhou 310029)

ABSTRACT

The limitations of methods for analysis of variance (ANOVA) in estimating genetic variances are discussed. Among the three methods (maximum likelihood ML, restricted maximum likelihood REML, and minimum norm quadratic unbiased estimation MINQUE) for mixed linear models, MINQUE method is presented with formulae for estimating variance components and covariances components and for predicting genetic effects. Several genetic models, which cannot be appropriately analyzed by ANOVA methods, are introduced in forms of mixed linear models. Genetic models with independent random effects can be analyzed by MINQUE (*1*) method which is a MINQUE method with all prior values setting *1*. MINQUE (*1*) method can give unbiased estimation for variance components and covariance components, and linear unbiased prediction (LUP) for genetic effects. There are more complicate genetic models for plant seeds which involve correlated random effects. MINQUE (*0*/*1*) method, which is a MINQUE method with all prior covariances setting *0* and all prior variances setting *1*, is suitable for estimating variance and covariance components in these models. Mixed model approaches have advantage over ANOVA methods for the capacity of analyzing unbalanced data and complicated models. Some problems about estimation and hypothesis test by MINQUE method are discussed.

Key words: Mixed linear models, MINQUE method, variance and covariance estimation, random effect prediction.

1991年3月20日收到

1. INTRODUCTION

Estimating genetic variance is of importance for quantitative genetic research as well as for plant and animal breeding. The simplest genetic linear model for partitioning phenotypic value can be expressed as

$$y = \mu + G + e$$

with

$$V_y = V_G + V_e,$$

where y = phenotypic value with mean μ and variance V_y, μ = population mean, G = genotypic value with mean 0 and variance V_G, e = nongenetic effect with mean 0 and variance V_e.

Partitioning of genetic variance (V_G) for additive (V_A), dominance (V_D) and epistatic (V_I) components was introduced by Fisher [6]. Cockerham [3] further subdivided the epistatic component into additive by additive (V_{AA}), additive by dominance (V_{AD}), and dominance by dominance (V_{DD}) epistatic components. Fisher [7] introduced the so-called analysis of variance (ANOVA) method for estimating variance components from balanced data. Many genetic models were proposed for estimating genetic variance components based on ANOVA methods. Among various mating designs, the nested (Design I) and factorial (Design II) designs [2], and the diallel mating designs [8-10] are most used by plant and animal breeders. Variance components in these models can be estimated by ANOVA methods. ANOVA methods are easy to use for balanced data. Under the normality assumption ANOVA estimators are uniformly minimum variance unbiased. Therefore these methods are still widely used by breeders and geneticists until now.

There are several limitations of ANOVA methods. Unbalanced data cannot be appropriately analyzed by ANOVA methods. Genetic variance components for some complicated models are not estimable by ANOVA methods. These limitations could be removed by the new estimation approaches for mixed linear models when maximum likelihood (ML) method [9], restricted maximum likelihood (REML) method [13], and minimum norm quadratic unbiased estimation (MINQUE) method [14] were developed. Among the mixed model approaches MINQUE method has advantages of simple computaton, no requirement for normality distribution, and unbiased estimation as compared to ML and REML methods.

2. MINQUE METHOD FOR MODELS WITH INDEPENDENT RANDOM EFFECTS

Cockerham and Weir [5] proposed a bio-model of diallel crosses. The

bio-model provides a way for estimating maternal and paternal variance components. If other higher-order interaction effects are not included in the model, the mean observation in the k-th block of the cross between maternal line i and paternal line j can be partitioned as

$$y_{ijk}=\mu+n_i+n_j+t_{ij}+m_i+p_j+b_k+\epsilon_{ijk}$$

with

$$\sigma_y^2=2\sigma_n^2+\sigma_t^2+\sigma_m^2+\sigma_p^2+\sigma_b^2+\sigma_e^2,$$

where y_{ijk} is the average phenotypic value of individuals from line $i\times$line j in block k; μ is the fixed population mean; n_i is the effect of nuclear contribution of maternal line i, $n_i\sim(0,\sigma_n^2)$; n_j is the effect of nuclear contribution of paternal line j, $n_j\sim(0,\sigma_n^2)$; t_{ij} is the interaction effect of nuclear contributions of lines $i\times j$, $t_{ij}\sim(0,\sigma_t^2)$; m_i is the extranuclear maternal effect of line i, $m_i\sim(0,\sigma_m^2)$; p_j is the extranuclear paternal effect of line j, $p_j\sim(0,\sigma_p^2)$; b_k is the effect of block k, $b_k\sim(0,\sigma_b^2$; ϵ_{ijk} is the residual effect, $\epsilon_{ijk}\sim(0,\sigma_e^2)$.

If the parents are inbred lines, genetic variance components can be estimated as $V_A=2\sigma_n^2$, $V_D=\sigma_t^2$, $V_M=\sigma_m^2$, $V_P=\sigma_p^2$, $V_e=\sigma_e^2$. Cockerham and Weir [5] derived ANOVA procedure for estimating genetic variance components for the bio-model. By the ANOVA method maternal variance component and paternall variance component still cannot be estimated separately. The genetic effects are not estimable no matter what kinds of side conditions are imposed. Only balanced data for method 3 of diallel mating (excluding all parent lines) can be analyzed by the ANOVA method. There are no formulae available for method 1 (including parent lines) and for half diallel mating (method 2 and method 4). All these problems can be solved by the mixed model approaches.

The bio-model of diallel crosses can be written as a matrix form of the mixed model

$$\underset{\sim}{y}=\underset{\sim}{1}\mu+\underset{\sim}{U}_n\underset{\sim}{e}_n+\underset{\sim}{U}_t\underset{\sim}{e}_t+\underset{\sim}{U}_m\underset{\sim}{e}_m+\underset{\sim}{U}_p\underset{\sim}{e}_p+\underset{\sim}{U}_b\underset{\sim}{e}_b+\underset{\sim}{e}. \qquad (2.1)$$

$$=\underset{\sim}{1}\mu+\sum_{u=1}^{5}\underset{\sim}{U}_u\underset{\sim}{e}_u+\underset{\sim}{U}_e\underset{\sim}{e}_e\sim\left(\underset{\sim}{\mu}=\underset{\sim}{1}\mu,\ \underset{\sim}{V}=\sum_{u=1}^{5}\sigma_u^2\underset{\sim}{U}_u\underset{\sim}{U}_u+\sigma_e^2\underset{\sim}{I}\right),$$

where $\underset{\sim}{U}_u$ is an incidence matrix of known constants, $\underset{\sim}{e}_u$ is a vector of independent random variables with $E(\underset{\sim}{e}_u) = \underset{\sim}{0}$ and $\mathrm{var}(\underset{\sim}{e}_u) = \sigma_u^2 \underset{\sim}{I}$, and $\underset{\sim}{U}_6 = \underset{\sim}{I}$ is an identity matrix. Variance components of the mixed model can be estimatod by Solving the following MINQUE equations for $U, v = 1, 2, \cdots, 6$:

$$[\mathrm{tr}(\underset{\sim}{U}_u' \underset{\sim}{Q} \underset{\sim}{U}_v \underset{\sim}{U}_v' \underset{\sim}{Q}_a \underset{\sim}{U}_u)] \hat{\underset{\sim}{\sigma}}^2 = [\underset{\sim}{y}' \underset{\sim}{Q}_a \underset{\sim}{U}_u \underset{\sim}{U}_u' \underset{\sim}{Q}_a \underset{\sim}{y}], \qquad (2.2)$$

where

$$\underset{\sim}{V}_a = \sum_{u=1}^{6} \alpha_u \underset{\sim}{U}_u \underset{\sim}{U}_u' \quad \text{with inverse } \underset{\sim}{V}_a^{-1},$$

$$\underset{\sim}{Q}_a = \underset{\sim}{V}_a^{-1} - \underset{\sim}{V}_a^{-1} \underset{\sim}{1} (\underset{\sim}{1}' \underset{\sim}{V}_a^{-1} \underset{\sim}{1})^{-1'} \underset{\sim}{V}_a^{-1}.$$

Although the estimates $\hat{\underset{\sim}{\sigma}}^2$ depend on prior values $\underset{\sim}{\alpha}$, they are unbiased, provided that the choice of $\underset{\sim}{\alpha}$ does not depend on the data. Because $\underset{\sim}{\alpha}$ is a vector of known values, variance components can be estimated non-iteratively. The prior values $\underset{\sim}{\alpha}$ may be chosen from prior experiments, from iterations or theoretical considerations. MINQUE(1) as a MINQUE method with $\underset{\sim}{\alpha} = \underset{\sim}{1}$ can be employed for estimating variance components.

Methods of estimating covariance components were proposed for the MINQUE procedure[15]. Those methods involve extensive computations and have been put to little use in practice. A much simpler MINQUE procedure for estimating covariance components can be derived for the mixed model (2.1). The covariance matrix of two variables $\underset{\sim}{y}_a$ and $\underset{\sim}{y}_b$ with equal design matrices is $\underset{\sim}{V}_{a/b} = \sum_{v=1}^{6} \sigma_{av/bv} \underset{\sim}{U}_v \underset{\sim}{U}_v'$ where $\sigma_{av/bv}$ is the covariance component for the v-th random factor. By MINQUE theory the expectation of the quadratic function $\underset{\sim}{y}_a' \underset{\sim}{Q}_a \underset{\sim}{U}_u \underset{\sim}{U}_u' \underset{\sim}{Q}_a \underset{\sim}{y}_b$ is

$$\mathrm{tr}(\underset{\sim}{Q}_a \underset{\sim}{U}_u \underset{\sim}{U}_u' \underset{\sim}{Q}_a \underset{\sim}{V}_{a/b}) = \sum_{v=1}^{6} \sigma_{av/bv}\, \mathrm{tr}(\underset{\sim}{U}_u' \underset{\sim}{Q}_a \underset{\sim}{U}_v \underset{\sim}{U}_v' \underset{\sim}{Q}_a \underset{\sim}{U}_u).$$

Invariant and unbiased estimators of covariance components $\hat{\underset{\sim}{\sigma}}_{a/b}$ can then be obtained by solving the following system of equations

$$[\mathrm{tr}(\underset{\sim}{U}_u' \underset{\sim}{Q}_a \underset{\sim}{U}_v \underset{\sim}{U}_v' \underset{\sim}{Q}_a \underset{\sim}{U}_u)] \hat{\underset{\sim}{\sigma}}_{a/b} = [\underset{\sim}{y}_a' \underset{\sim}{Q}_a \underset{\sim}{U}_u \underset{\sim}{U}_u' \underset{\sim}{Q}_a \underset{\sim}{y}_b] \qquad (2.3)$$

The matrices [tr ($U_u' Q_s U_s U_s' Q_s U_u$)] and $U_u Q_s$ are the same for both variance and covariance estimation. Therefore they can be stored for later recall to estimate variances and covariances for multiple traits.

In animal breeding, estimated variance components are often used for predicting genetic merits in selection programs. For the mixed model (2.1), the best linear unbiased prediction(BLUP [11]) for the u-th vector of random genetic effects can be obtained by,

$$\hat{u}(\sigma_u^2) = \sigma_u^2 U_u' V^{-1} (y - l \hat{\mu}) = \sigma_u^2 U_u' Q y ,$$

where $\hat{\mu} = (l' V^{-1} l)^{-1} l' V^{-1} y$,

$Q = V^{-1} - V^{-1} l (l' V^{-1} l)^{-1} l' V^{-1}$.

The BLUP needs known variances. Since the true variances are unknown in practice, unbiased estimates could be used in prediction. The unknown variances can also be replaced by prior values from prior experiment or from reasonable guesses. Therefore the genetic effects can be predicted by choosing prior values α as in the case of MINQUE method, $\hat{u}(\alpha) = \alpha_u U_u' Q y$.

Monte Carlo simulations showed that variance components and covariance components can be efficiently estimated with negligible bias by the MINQUE (1) procedure. Even unbalanced data from diallel mating can be estimated as efficiently as balanced data if they have similar experiment sizes. The MINQUE (1) procedure can give liner unbiased prediction (LUP) for random genetic effects almost as good as the BLUP.

All the linear models for ANOVA methods have a restriction that coefficients for the random effects should be integer (most being 1 or 0). But some real genetic models with biological accuracy are mixed linear models with decimal coefficients. These models cannot be analyzed by the ANOVA methods. Mixed model approaches are the only procedures which can give unbiased estimation for variance conponents and covariance components .

In quantitative genetic analysis epistatic variance components are more difficult to estimate than additive and dominance variance components by the ANOVA method. When applying Cockerham's general genetic model[4], Zhu[16] proposed mating designs of modified diallel crosses for estimating additive by additive epistatic variance component. One of the modified diallel mating systems consists of F_1's and F_2's. The other mating system

involves backcrosses(BC's)in place of F_2's. The assumptions of the genetic model are (*1*) regular diploid segregation; (*2*) randomly selected inbred parents; (*3*) no linkage; (*4*) no extranuclear effects; (*5*) no genotype by environment interaction. The model can be expressed as

$$y_{ijk}=\mu+G_{ij}+b_k+\text{error},$$

where the total genotypic effect G_{ij} depends on the specific genetic entry; for F_1 of line $i\times$ line j:

$$G_{ij}(F_{1ij})=A_i+A_j+D_{ij}+AA_{ii}+AA_{jj}+2AA_{ij}$$

for F_2 of line $i\times$ line j:

$$G_{ij}(F_{2ij})=A_i+A_j+0.25D_{ii}+0.25D_{jj}+0.5D_{ij}+AA_{ii}+AA_{jj}+2AA_{ij}$$

and for backcross (BC) of $F_{1ij}\times$ line j:

$$G_{ij}(\text{BC})=0.5A_i+1.5A_j+0.5D_{ij}+0.5D_{ij}+0.25AA_{ii}+2.25AA_{jj}+1.5AA_{ij}.$$

In this genetic model A_i is the cumulative additive effect from line i; A_j is the cumulative additive effect from line j; D_{ij} is the cumulative dominance effect from the cross of line $i\times$ line j; and AA_{ij} is the cumulative epistatic effect from line $i\times$ line j.

The mixed linear model can be written in a matrix form for all the entries of F_1's, and F_2's (or BC's),

$$\underset{\sim}{y}=\underset{\sim}{l}\mu+U_A\underset{\sim}{e}_A+U_D\underset{\sim}{e}_D+U_{AA}\underset{\sim}{e}_{AA}+U_b\underset{\sim}{e}_b+\underset{\sim}{e}, \qquad (2.4)$$

with the variance-covariance matrix for entries of the mating design

$$var(\underset{\sim}{y})=U_AU_A'\sigma_A^2+U_DU_D'\sigma_D^2+U_{AA}U_{AA}'\sigma_{AA}^2+U_bU_b'\sigma_b^2+I\sigma^2.$$

All the effects except the constant μ are independent random effects. Genetic variance components can be directly estimated by $V_A=2\sigma_A^2$, $V_D=\sigma_D^2$, $V_{AA}=4\sigma_{AA}^2$, and $V_e=\sigma^2$, without using covariances of relatives. As for the diallel crosses, four methods can be chosen for the modified diallel crosses. In the modified diallel mating system, the segregating generation F_2 is included for self-pollinated crops or BC for cross-pollinated crops and animals. Monte Carlo simulations showed that variance components could be well estimated by six-parent modified diallel crosses.

3. MINQUE METHOD FOR MODELS WITH CORRELATED RANDOM EFFECTS

An understanding of inheritance for characteristics and nutrition content

of seeds is important to plant breeding for improvement of yield potential and seed quality. Since maternal plants supply assimilates for seed filling, and since the photosynthetic activity of the maternal plaant is determined by the maternal nuclear genes and cytoplasmic genes, a real genetic model should consist of maternal genetic effects and cytoplasmic effects along with direct genetic effects. Seed nuclear genes are derived partly from maternal plants. The genetic model for seed traits should include covariances between nuclear genetic effects and maternal genetic effects.

Another restriction of ANOVA methods is that all the random effects are not correlated to each other. ANOVA methods cannot handle linear models with correlated random effects. Therefore some complicated models such as the genetic models for seed traits could not be approached by ANOVA methods. MINQUE methods are good candidates for estimaton of variance and covariance components of mixed models with correlated random effects.

A genetic model is proposed for quantitative traits of diploid seeds under the assumptions of no paternal effects, constant inheritance of cytoplasmic genes through maternal lines only, no epistatic effects, and no genotype by environment interaction. Modified diallel crosses consisting of F_1's and reciprocal F_1's from a set of completely inbred lines, and backcrosses of F_1 to their two parents are used in the genetic model for diploid plant seeds with cytoplasmic and maternal effects. This model is also suitable for quantitative traits of animal offspring. The genetic models can be written as a linear model for the mean observation in the l-th block of the k-th type of genetic entry from line i and line j,

$$y_{ijkl}=\mu+G_{ijk}+b_l+\epsilon_{ijkl},$$

and the total genetic efffect G_{ijk} depends on the specific genetic entry; for F_1 ij from maternal line i × paternal line j ($k=1$):

$$G_{ij1}=A_i+A_j+D_{ij}+C_i+2Am_i+Dm_{ij},$$

for backcross BC_1 from maternal F_1 ij × paternal line i ($k=2$):

$$G_{ij2}=1.5A_i+0.5A_j+0.5D_{ii}+0.5D_{ij}+C_i+Am_i+Am_j+Dm_{ij},$$

and for backcross BC_2 from maternal F_1 ij × paternal line j ($k=3$):

$$G_{ij3}=0.5A_i+1.5A_j+0.5D_{jj}+0.5D_{ij}+C_j+Am_i+Am_j+Dm_{ij}.$$

The genetic effects in the model are defined as follows: A_i is the cumulative additive effect of direct genes from line i, $A_i\sim(0,\ \sigma_A^2)$; the cumulative dominance effect of direct genes is $D_{ij}\sim(0,\ \sigma_D^2)$; the cytoplasmic effect is $C_i\sim(0,\ \sigma_c^2)$; the cumulative additive effect of maternal genes is $Am_i\sim$

$(0, \sigma^2_{Am})$; and the cumulative dominance effect of maternal genes is $Dm_{ij} \sim (0, \sigma^2_{Dm})$. There are covariances between direct genes and maternal genes, $\mathrm{cov}(A_i, Am_i) = \sigma_{A.Am}$ and $\mathrm{cov}(D_{ij}, Dm_{ij}) = \sigma_{D.Dm}$.

This modified diallel crosses with F_1's, reciprocal F_1's and their backcrosses are suitable for cross-pollinated crops. For some self-pollinated crops, F_2 seed can be easily obtained from F_1 plants. F_2 can then be included in the genetic model, and the total genetic effect for F_{2ij} $(k=4)$ is

$$G_{ij4} = A_i + A_j + .25D_{ii} + .25D_{jj} + .5D_{ij} + C_i + Am_i + Am_j + Dm_{ij}.$$

Genetic models are also suggested for quantitative traits of triploid endosperm with nuclear genetic effects, cytoplasmic effects, and maternal genetic effects. If high-order dominance effects of three-allele interaction are negligible or not much concerned in practice, a reduced genetic model can be employed. The reduced genetic model can be written as a linear model for the mean observation in the l-th block of the k-th type of genetic entry from lines i, j:

$$y_{ijkl} = \mu + G_{ijk} + b_l + e_{ijkl},$$

where the total genetic effect G_{ijk} depends on specific genetic entry of endosperms. For inbred line P_i $(i=j)$ and F_{1ij} $(i \neq j)$ from maternal $P_i \times$ paternal P_j $(k=1)$:

$$G_{ij1} = 2A_i + A_j + D_{ii} + 2D_{ij} + C_i + 2Am_i + Dm_{ij},$$

for F_{2ij} $(k=2)$:

$$G_{ij2} = 1.5A_i + 1.5A_j + D_{ii} + D_{jj} + D_{ij} + C_i + Am_i + Am_j + Dm_{ij}.$$

These two genetic models can be written in a matrix form for all entries of the mating design,

$$\underset{\sim}{y} = \underset{\sim}{1}\mu + \underset{\sim}{U}_A \underset{\sim}{e}_A + \underset{\sim}{U}_D \underset{\sim}{e}_D + \underset{\sim}{U}_C \underset{\sim}{e}_C + \underset{\sim}{U}_{Am} \underset{\sim}{e}_{Am} + \underset{\sim}{U}_{Dm} \underset{\sim}{e}_{Dm} + \underset{\sim}{U}_b \underset{\sim}{e}_b + \underset{\sim}{e}_* = \underset{\sim}{1}\mu + \sum_{u=1}^{7} \underset{\sim}{U}_u \underset{\sim}{e}_u \qquad (3.1)$$

with the variance-covariance matrix

$$\begin{aligned}\mathrm{var}(y) = {} & \sigma^2_A \underset{\sim}{U}_A \underset{\sim}{U}'_A + \sigma^2_D \underset{\sim}{U}_D \underset{\sim}{U}'_D + \sigma^2_C \underset{\sim}{U}_C \underset{\sim}{U}'_C \\ & + \sigma^2_{Am} \underset{\sim}{U}_{Am} \underset{\sim}{U}'_{Am} + \sigma^2_{Dm} \underset{\sim}{U}_{Dm} \underset{\sim}{U}'_{Dm} + \sigma^2_b \underset{\sim}{U}_b \underset{\sim}{U}'_b \\ & + \sigma_{A.Am}(\underset{\sim}{U}_A \underset{\sim}{U}'_{Am} + \underset{\sim}{U}_{Am} \underset{\sim}{U}'_A) + \sigma_{D.Dm}) \underset{\sim}{U}_D \underset{\sim}{U}'_{Dm} + \underset{\sim}{U}_{Dm} \underset{\sim}{U}'_D) \\ & + \sigma^2_e \underset{\sim}{I} = \sum_{u=1}^{9} \theta_u \underset{\sim}{V}_u = \underset{\sim}{V}.\end{aligned}$$

where $\underset{\sim}{U}_u$ is the known incidence matrix relating to the random vector $e_u \sim (\underset{\sim}{0}, \sigma_u^2 \underset{\sim}{I})$ for $u=1, 2, \cdots, 7$; $\underset{\sim}{V}_u = \underset{\sim}{U}_u \underset{\sim}{U}_u'$ for $u=1, 2, \cdots, 6$; $\underset{\sim}{V}_7 = (\underset{\sim}{U}_1 \underset{\sim}{U}_4' + \underset{\sim}{U}_4 \underset{\sim}{U}_1')$, $\underset{\sim}{V}_8 = (\underset{\sim}{U}_2 \underset{\sim}{U}_5' + \underset{\sim}{U}_5 \underset{\sim}{U}_2')$, and $\underset{\sim}{V}_9 = \underset{\sim}{U}_7 = \underset{\sim}{I}$. For diploid seeds in modified diallel crosses from inbred lines, nuclear genetic variance components can be estimated by $V_A = 2\sigma_A^2$, $V_D = \sigma_D^2$. Maternal genetic variance components are $V_{Am} = 2\sigma_{Am}^2$, $V_{Dm} = \sigma_{Dm}^2$. Genetic variance components for triploid endosperm can be estimated by $V_A = 3\sigma_A^2$, $V_D = 3\sigma_D^2$, $V_C = \sigma_C^2$, $V_{Am} = 2\sigma_{Am}^2$ and $V_{Dm} = \sigma_{Dm}^2$.

MINQUE ($0/1$) procedure is a MINQUE method with 0 for all the prior covariances and 1 for all the prior variances. The MINQUE ($0/1$) procedure is efficient and unbiased for estimating variance and covariance components by mixed models with correlated random factors. Estimators of variance and covariance components $\underset{\sim}{\hat{\theta}}$ for these genetic models can be obtained by solving the following MINQUE ($0/1$) equations for $u, v=1, 2, \cdots, 9$:

$$[\mathrm{tr}(\underset{\sim}{Q}_{(0/1)} \underset{\sim}{V}_u \underset{\sim}{Q}_{(0/1)} \underset{\sim}{V}_v)]\ \underset{\sim}{\hat{\theta}} = [\underset{\sim}{y}' \underset{\sim}{Q}_{(0/1)} \underset{\sim}{V}_u \underset{\sim}{Q}_{(0/1)} \underset{\sim}{y}], \qquad (3.2)$$

where

$$\underset{\sim}{V}_{(0/1)} = \sum_{u=1}^{6} \underset{\sim}{U}_u \underset{\sim}{U}_u' + \underset{\sim}{I}, \text{ with inverse } \underset{\sim}{V}_{(0/1)}^{-1},$$

$$\underset{\sim}{Q}_{(0/1)} = \underset{\sim}{V}_{(0/1)}^{-1} - \underset{\sim}{V}_{(0/1)}^{-1} \underset{\sim}{1} (\underset{\sim}{1}' \underset{\sim}{V}_{(0/1)}^{-1} \underset{\sim}{1})^{-1} \underset{\sim}{1}' \underset{\sim}{V}_{(0/1)}^{-1}.$$

The invariant and unbiased estimators of covariance components $\underset{\sim}{\hat{\theta}}_{a/b}$ between trait a and trait b with equal design matrices can also be obtained by solving the following system of equations for $u, v=1, 2, \cdots, 9$:

$$[\mathrm{tr}(\underset{\sim}{Q}_{(0/1)} \underset{\sim}{V}_u \underset{\sim}{Q}_{(0/1)} \underset{\sim}{V}_v)]\ \underset{\sim}{\hat{\theta}}_{a/b} = [\underset{\sim}{y}_a' \underset{\sim}{Q}_{(0/1)} \underset{\sim}{V}_u \underset{\sim}{Q}_{(0/1)} \underset{\sim}{y}_b], \qquad (3.3)$$

where $\underset{\sim}{y}_a$ and $\underset{\sim}{y}_b$ are vectors of mean observations for two traits. By the MINQUE ($0/1$) procedure, a linear unbiased prediction (LUP) for the u-th random genetic factor can be obtained by $\underset{\sim}{\hat{c}}_{u(0/1)} = \underset{\sim}{U}_u' \underset{\sim}{Q}_{(0/1)} \underset{\sim}{y}$ for $u \leqslant 5$.

4. DISCUSSION

Mixed models can be approaches by three major methods ML, REML, and MINQUE. ML tends to give biased estimates for some variance

components. In most cases REML give estimates similar to MINQUE estimates. Both ML and REML require iterations. For mixed models with balanced data several iterations may be enough. But in some rare cases iterations might not couverge. MINQUE method is a non-iterative procedure. MINQUE method is superior to ML and REML for its simple computation and unbiased estimation.

As compared with ANOVA methods, ML, REML, and MINQUE methods involve matrix algebra with much more complicated computation. It is only recent event of applying these methods in estimating genetic variance components. When more genetic models with biological accuracy are developed and applied in plant and animal breeding, inheritance of quantitative traits will be better understood.

By the MINQUE method negative estimates of variance components are not unusual. When using constrains of setting all the negative estimates to be zero, unbiaseIness is no longer guaranteed. Although MINQUE method works well in general, not all mixed models can be well estimated by MINQUE method. If a genetic mating design results in a mixed model with some incidence design matrices being functions of other matrices, the left matrix in (2.2) or (3.2) will be singular. Hence variance components are not estimable without using general inverse. Some unbalanced data structures may also result in singular matrices in (2.2) or (3.2). Therefore new genetic models should be proved by Monte Carlo simulatons for unbiasedness and efficiency before they can be used in practice. Some unbalanced mating design are also needed for checking if variance components are estimable.

When variance components and covariance components have been estimated, statistic tests are usually followed for detecting their significance. Sampling variances of estimates are needed for hypothesis tests. Asymptotic sampling variances of estimates can be obtained by formulae of ML, REML, and MINQUE methods. A much simpler procedure of estmation of sampling variances is jackknifing [12]. If θ is an estimate of a genetic parameter from a sample of K observations, and $\theta_{(k)}$ is the estimate resulting when obesrvation k is omitted, then the k-th pseudovalue is $J_k(\hat{\theta}) = K\hat{\theta} - (K-1)\hat{\theta}_{(k)}$. The jackknife estimator $J(\hat{\theta})$ of parameter θ is the mean of the pseudovalues. If K is not large, $(J(\theta)-\theta)/SE(J(\theta))$ is approximately distributed as a t-distribution with $(K-1)$ degrees of freedom.

REFERENCES

(1) Comstock, R. E. and Robinson, H. F. (1952). Estimation of average dominance of genes. In *Heterosis*. Gowan, J. W. (editor), Iowa State University Press, Ames, Iowa.

(2) Coekerham, C. C. (1954). An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* 39: 859–882.

(3) Cockerham, C.C. (1980). Random and fixed effects in plant genetics. *Theor. Appl. Genet.* 56: 119–131.

(4) Cockerham, C.C. and Weir, B.S. (1977). Quadratic analyses of reciprocal crosses. *Biometrics* 33: 187–203.

(5) Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinb.* 52: 399–433.

(6) Fisher, R.A. (1925). *statistical Methods for Research Workers*. 1st ed. Oliver & Boyd. Edinburgh and London.

(7) Griffing, B. (1956). Concept of general and specific combining ability in relation to diallel crossing systems. *Aust. J. Biol. Sci.* 9: 463–493.

(8) Hallauer, A.R. and Miranda, J. B. (1981) *Quantitative Genetics in Maize Breeding*. Iowa State University Press, Ames, Iowa.

(9) Hartley, H.D. and. Rao, J. N. K. (1967). Maximum-likelihood estimation for the mixdd analysis of variance model. *Biometrika* 54: 93–108.

(10) Hayman, B. I. (1954). The theory and analysis of diallel crosses. *Genetics* 39: 789–809.

(11) Henderson, C.R. (1963). Selection index and expected genetic advance. In: Hanson, W. D., Robinson, H. F. (eds.) *Statistcal Genetics and Plant Breeding. Natl. Acad. Sci. -N R C* 982: 141–163.

(12) Miller, R. G. (1974). The jackknife–a review. *Biometrika* 61: 1–15.

(13) Patterson, H. D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* 58: 545–554.

(14) Rao, C. R. (1971). Estimation of variance and covariance components–MINQUE theory. *J. Multivar. Anal.* 1: 257–275.

(15) Rao, C. R. and Kleffe, J. (1980). Estimation of variance components. In *Handbook of Statistics*. vol. 1, 1–40. Krishnaiah, P. R. (ed.). New York.

(16) Zhu Jun, (1989). Estimation of genetic variance components in the general mixed model. Ph.D. Dissertation. North Carolina State Univ. Raleigh, N.C.