

中国博士专著·农业领域

Monographs by China's Ph.D.s · Agricultural

遗传模型分析方法

Analysis Methods for Genetic Models

中华人民共和国农业部教育司 主编

朱军 著

中国农业出版社

内容提要

本书系统地介绍了近一二十年来在数量遗传学领域的一些新发展，尤其是混合线性模型的统计分析方法和广义遗传模型的建模原理。所介绍的统计分析方法包括 ANOVA 法、ML 和 REML 法、MINQUE 法，但是侧重于 MINQUE 方法的阐述。详细地介绍了方差和协方差估算、遗传效应值预测的一些新的统计方法。所介绍的遗传模型，不但包括常用的加性-显性模型和上位性遗传模型，还包括了最近几年创立的动物性连锁模型和种子遗传模型。在附录中介绍了各种遗传分析软件的使用方法。本书所介绍的这些新遗传模型和分析方法，全然不同于国内外数量遗传学专著所介绍的那些传统内容。

本书可作为遗传育种及数量遗传研究方向的硕士生和博士生课程教材，也可供统计学、遗传学、数量遗传学以及遗传育种的教学和科研工作者参考。

<<中国博士专著·农业领域>>编辑委员会

主任委员 毛达如 王红谊

副主任委员 韩惠鹏 陈凤秀

委员 (按姓氏笔划为序)

王红谊 王树安 云泽民

毛达如 石元春 刘佩瑛

李扬汉 李阜棣 扬天桥

扬传任 余 群 张 沅

陈凤秀 陆龙英 庞雄飞

郭智奇 盖钧镒 阎隆飞

曾士迈 韩惠鹏 景士西

蔡同一 蔡宝祥 蔡盛林

樊德方

出版者的话

科技发展靠人才。我国目前的中青年博士就是跨世纪人才中的一部分。他们当中的一些人有的已成为有关学科的学术骨干，有的已在其学科领域获重大突破，被老一代科学家公认为很有前途的学科带头人。更为可喜的是，我国高技术研究发展战略计划——“863计划”的各个领域中，也吸收一批中青年学者挑起大梁。毋庸置疑，编辑出版他们的学术著作也将一项继往开来、承前启后的跨世纪系统工程。因为他们的学术著作不仅是老一辈科学家思想体系和学术成就的继承，也是本世纪末和21世纪科技人才脱颖而出的询路之石。

《中国博士专著·农业领域》丛书的出版，必将对我国农业科技事业的发展起着不可估量的作用。

我们的农学博士们大多还很年轻，反映在他们著作中的一些观点、原理和方法是否成熟，表现出的学术研究水平究竟如何，还需要在问世后接受广大读者的检验。我们殷切期待着广大高等农业院校师生、研究生以及农业科技和管理人员提出宝贵意见，使之日臻完善。

1994年7月

序 一

我国自 1981 年建立学位制度以来，已陆续授予近万人以博士学位；同时，留学政策开放后又有大批青年学子出国深造，现已有不少人回国工作。不论国内或国外培养的博士都学有所成，业有所专，在各个学科领域作出了重要贡献。农业领域与其他领域一样，已有一批学有专长的博士在学术上有很深的造诣。中国农业出版社邀请了一部分在农业领域工作的年轻博士撰写《中国博士专著·农业领域》丛书。这部丛书将反映出我国培养博士制度的成就，并反映出我国跨世纪人才的成长，是一件意义十分重大、值得特别庆贺的事。

本丛书选题广泛，涉及农业领域的各个方面，概括了农业经济及政策、农业生物技术、作物及园艺理论、畜牧兽医基础、动植物资源以及农业高新技术应用等。这些专著的作者均为年轻博士，他们思路开阔，取材新颖，能反映当前国内外农业领域有关学科的最新成就，并阐述了作者在各自学科中的学术观点，有其独到的见解。因此，本书的出版必将对提高我国农业科学及农业生产技术起到推动作用。今后本丛书还将继续出版各学科的专著，预期本丛书将成为我国农业领域的重要学术著作。

中国科学院院士

国务院学位委员会委员 阎隆飞

中国博士后科学基金会理事

中国农业大学植物生理学教授

1994 年 7 月

序 二

自从 80 年代初我国建立学位制度以来,1984 年农科第一位博士生毕业,至今全国高等农业院校和中国农科院系统共培养了六百余名学有专长的博士研究生,同时,通过国家公派,国内外双边联合培养等多种渠道,在国外又培养了一批国内力量薄弱、新兴学科的博士,有些已学成回国。这些中青年博士在国内科研、教育、管理岗位上崭露头角,脱颖而出,成为农科各学科的学术和业务骨干,其中不少人已成为新的学科带头人,有四人被国家教育委员会评为 21 世纪青年学科带头人(农科约占全国总数的 10%)。他们勤奋好学,治学严谨,成绩卓著,并在其学术领域中获得显著进展和突破。为了展示成果,促进学术交流,推动学科发展,繁荣科学,鼓励中青年博士著书立说,发展新的观点、新的理论、新的技术、新的成果,农业部教育司与中国农业出版社共同出版《中国博士专著·农业领域》系列丛书。通过专著的出版,可以对我国博士学术研究成果进行一次公开的展示和检验。这批专著是通过广泛征集选题,再由全国知名教授、博士生导师组成的编委会审定的,这一过程也是对 21 世纪科技队伍和中青年学术队伍人才库的建立过程,为繁荣农业科技、教育、出版事业作好人才储备。

祝贺《中国博士专著·农业领域》系列丛书的出版。

感谢中国农业出版社、丛书编委会和为本专著出版付出辛勤劳动的作者和工作人员。

中国农业大学校长、植物营养学教授

《中国博士专著·农业领域》 毛达如
丛书编委会主任

1994 年 7 月

前 言

自 Fisher 于 1925 年提出方差分析方法(ANOVA 方法)以来,数量遗传学家运用方差分析的原理,已发展了许多实用的遗传模型。至今仍有不少遗传模型被遗传育种工作者广泛使用,比如北卡罗来纳设计 (NC 和 NC)、双列杂交遗传模型。方差分析统计方法的不断完善和线性模型在数量遗传分析中的运用,曾极大地推动了数量遗传学的发展。由于经典数量遗传分析的统计分析基础是 ANOVA 方法,因此存在一些固有的局限性,它不能无偏分析有不规则缺失的非平衡数据,也无法分析具有生物学意义的复杂遗传模型。数量性状分析对 ANOVA 方法的依赖性已成为阻碍数量遗传进一步发展的制约因素。

70 年代初统计学家创立了一套崭新的统计分析方法——混合线性模型分析方法(mixed linear model approaches)。这些统计分析方法克服了 ANOVA 方法的局限性,不但可以无偏分析有不规则缺失的非平衡数据,而且还能分析各种复杂的遗传模型。统计分析方法的突破极大地推动了数量遗传学科的新发展。1980 年 Cockerham 提出了广义遗传模型的建模原理,为数量遗传学家建立各种复杂的遗传模型奠定了理论基础。

80 年代末我在美国北卡罗来纳州立大学统计系攻读博士学位及博士后研究期间,曾在导师 Bruce S. Weir 教授的指导下,从事混合线性模型统计分析方法的理论研究工作。我们提出了无偏估算成对性状协方差分量以及无偏预测随机因素效应值的统计公式,并发展了一些适用于分析动植物数量性状的遗传模型。我于 1990 年留学回国后,结合我国的育种实践,又提出了一些新的遗传模型和新的统计分析方法。

虽然国内外已出版了不少数量遗传学专著,总结经典数量遗传学的研究成果,但是尚没有一本系统介绍以混合线性模型分析方法和广义遗传模型原理为基础的数量遗传学专著。近年来,我一直为浙江农业大学数量遗传研究方向的硕士生和博士生开设遗传模型分析方法学位课程,介绍国际上近一二十年来数量遗传学科发展的新动态、以及发展新遗传模型和统计分析方法的原理和方法,培养学生能应用数量遗传和线性模型的基本原理,设计新的遗传模型和发展新的分析公式。本书就是在这门课程的备课笔记基础上,并总结我们近年来的一些科研成果而完稿的。我希望这本书中介绍的一些新的统计分析方法和新的遗传模型,能对从事数量遗传研究或教

学的科学工作者有所启发。

全书共分十二章和二个附录。前四章介绍了数量遗传和统计分析的基础知识。由于混合线性模型分析方法建立在矩阵运算的基础上，而这些知识不属于数量遗传学的内容，因此在附录 A 中简要地介绍了一些矩阵运算的原理和方法。在第五章至第八章中，详细地介绍各种常用的遗传模型，并附有相应的分析实例。在第九章和第十章中，分别介绍发育数量遗传分析和杂种优势分析的一些新方法。第十一章介绍遗传率和遗传相关分析的一些新概念。在第十二章中介绍了基于混合线性模型原理的基因型评价和品种区域试验分析方法。书中介绍的一些新的遗传模型及其相应的统计分析方法，必需借助计算机才能运算。我用 C 语言编写了所有的分析软件，附录 B 中介绍了各种软件的使用方法。

本书介绍的大部分内容都是近年来才提出的新方法和新模型，还有待进一步发展。因此不免有解释不妥或不完善之处，恳请读者批评指正，以便今后修改补充。

我感谢季道藩教授引导我踏进了数量遗传学的门槛，也感谢 Bruce S. Weir 教授为我提供了全额奖学金，使我能在美国著名的北卡罗来纳州立大学统计系学习统计学和数量遗传学的深奥理论，从而步入数量遗传研究的前沿。没有他们的培养和教诲，我不可能完成这本书的撰写。我还要感谢国家自然科学基金和国家教委跨世纪人才专项基金对我从事数量遗传基础理论研究的支持。正是这些研究使我们能在统计学和遗传学研究领域取得大量的突破性成果，并将其总结在本书中。

朱军

1996 年 4 月

目 录

第一章 数量性状的特征和统计分析原理

第二章 近交系数和共同祖先系数

第一节 亲缘关系的度量

第二节 近交系数的计算

第三节 遗传协方差的计算

第三章 广义遗传模型

第一节 Cokerham的广义遗传模型

第二节 分析种子性状的广义遗传模型

第三节 包括基因型与环境互作效应的广义遗传模型

第四章 遗传模型的常用统计分析方法

第一节 遗传方差和协方差估算

1. ANOVA法
2. ML或REML法
3. MINQUE法

第二节 基因效应分析

1. 最小二乘估计法
2. 线性无偏预测法

第三节 几种蒙特卡罗模拟分析方法的效益比较

1. 方差分量估算效益的比较
2. 遗传效应预测效益的比较

第五章 加性-显性遗传模型分析

第一节 简单加性-显性遗传模型

1. 巢式交配设计的遗传分析方法
2. 析因交配设计的遗传分析方法
3. 双列杂交设计的遗传分析方法

第二节 加性-显性及与环境互作的遗传模型

1. 巢式交配设计的遗传分析方法
2. 析因交配设计的遗传分析方法
3. 双列杂交设计的遗传分析方法

第六章 加性-显性-上位性遗传模型分析

第一节 简单加性-显性-上位性遗传模型

第二节 加性-显性-上位性及与环境互作的遗传模型

第七章 包括母体效应的遗传模型分析

第一节 加性-显性-母体效应遗传模型

1. Griffing的双列杂交模型
2. Cockerham和Weir的生物模型(Bio-model)
3. 两个遗传模型的蒙特卡罗模拟比较
4. 加性-显性-母体效应遗传模型(ADM模型)

第二节 包括性连锁效应和母体效应的动物模型

1. ADLM遗传模型
2. ADLAmDm遗传模型

第八章 包括细胞质和母体效应的遗传模型分析

第一节 植物种子和动物幼畜的遗传模型

1. 加性-显性遗传模型的分析方法
2. 加性-显性及环境互作的遗传模型分析方法

第二节 二倍体胚和三倍体胚乳的种子遗传模型

第九章 发育性状的遗传分析

第一节 发育性状的净遗传效应及条件方差分量

第二节 发育性状的遗传分析方法

第三节 发育性状的分析实例

第十章 杂种优势的遗传分析

第一节 作物农艺性状的杂种优势分析

第二节 种子品质性状的杂种优势分析

第十一章 遗传率和遗传相关在育种上的应用

第一节 遗传率和选择响应

第二节 遗传相关和间接选择

第十二章 多年份、多试点品种区域试验的分析方法

第一节 单一性状的分析方法

第二节 综合性状的分析方法

附录A: 矩阵运算的原理和方法

A1. 矩阵的简单运算

A2. 线性独立性和矩阵的秩

A3. 矩阵求逆和广义逆矩阵

A4. 随机向量和矩阵

A5. 二次型的分布

附录B: 数量遗传的统计分析软件

B1. 双列杂交和杂种优势的遗传分析软件

B2. 种子性状的遗传分析软件

B3. 发育数量性状的遗传分析软件

B4. 多年份、多试点品种区域试验的统计分析软件

第一章 数量性状的特征和统计分析原理

遗传学所分析的生物性状可以分为两大类。一类是质量性状(qualitative character)，一般表现为不连续变异。质量性状主要受若干主基因控制，受环境影响较小。在杂种后代分离群体中，可以根据个体的性状表现把群体分成若干组。质量性状可用经典遗传分析的方法进行分析。在对质量性状进行研究时，常将群体中的表现型或基因型分类，采用群体遗传分析的方法估算不同基因型或基因的频率。另一类是数量性状(quantitative character)，表现为连续变异。数量性状通常受许多微效基因控制，并同时受到环境因素较大的影响。在杂种后代分离群体中，个体性状表现的变异是由基因分离和环境效应随机影响的综合作用造成的。因而无法用经典遗传分析的方法对数量性状进行分组分析。只能采用统计分析的方法，对特定遗传群体的数量性状进行遗传分析，区分遗传变异和环境变异。这种遗传分析方法称为数量遗传分析方法。

数量遗传研究最常用的统计分析方法是估算遗传群体的统计参数(parameter)，如平均数(mean)、方差(variance)和协方差(covariance)。如果有一个基因型纯合的品种，对于产量这个数量性状而言，每个单株的遗传组成中决定产量的基因型值是一个常数(constant)。如果不存在影响每个单株产量的环境效应，所有单株将有相同的产量表现，其值等于群体平均数 μ 。实际上单株产量除了受基因型影响以外，还在很大程度上受到环境因素的影响。因而群体中每个单株的产量表现都不尽相同，呈现连续的变异。统计上常用方差 σ^2 来度量群体的变异程度。对于绝大多数数量性状，其表现值呈正态分布(normal distribution)。任何正态分布都由两个统计参数决定，它们是群体平均数 μ 和方差 σ^2 。常用 $N(\mu, \sigma^2)$ 来表示正态分

布具有群体平均数 μ 和方差 σ^2 。

在这个产量例子中，群体平均数 μ 度量了品种的基因型值，方差 σ^2 则度量了环境效应所造成的变异程度。任何一个品种可能种植的植株数目是无穷的，人们不能观测分析所有可能种植的植株产量表现。因而无法直接获得群体平均数 μ 和方差 σ^2 。采用统计分析的方法，只需对一些样本植株进行观测分析，便可以估算群体平均数 μ 和方差 σ^2 。如果调查了 n 株植株，各株的产量测定值为 $y_1, y_2, \dots, y_i, \dots, y_n$ ，群体平均数 μ 可由样本平均数 (Sample Mean) $\hat{\mu}$ 估算，

$$\hat{\mu} = (y_1 + y_2 + \dots + y_i + \dots + y_n) / n = (1/n) \sum_{i=1}^n y_i$$

用样本方差 (sample variance) $\hat{\sigma}^2$ 可以估算群体方差 σ^2

$$\begin{aligned} \hat{\sigma}^2 &= [1/(n-1)][(y_1 - \hat{\mu})^2 + (y_2 - \hat{\mu})^2 + \dots + (y_i - \hat{\mu})^2 + \dots \\ &\quad + (y_n - \hat{\mu})^2] \\ &= \frac{1}{n-1} \sum_{i=1}^n (y_i - \hat{\mu})^2 \end{aligned}$$

样本平均数 $\hat{\mu}$ 和样本方差 $\hat{\sigma}^2$ 是群体平均数 μ 和方差 σ^2 的无偏估计值。即这些样本估计量 (estimator) 的期望值等于群体参数， $E(\hat{\mu}) = \mu$ ， $E(\hat{\sigma}^2) = \sigma^2$ 。

由于存在基因的连锁和基因的一因多效，生物体的不同数量性状之间常存在不同程度的相互关联。在统计分析中常用协方差来度量这种相互关联的变异程度。如果某遗传群体有两个相互关联的数量性状，即性状 X 和性状 Y，这两个性状的协方差 σ_{XY} 可用样本协方差 $\hat{\sigma}_{XY}$ (sample covariance) 来估算，

$$\hat{\sigma}_{XY} = [1/(n-1)] \sum_{i=1}^n (y_{i(X)} - \hat{\mu}_{(X)})(y_{i(Y)} - \hat{\mu}_{(Y)})$$

其中 $y_{i(X)}$ 和 $y_{i(Y)}$ 分别是性状X和性状Y的第*i*项观测值， $\hat{\mu}_{(X)}$ 和 $\hat{\mu}_{(Y)}$ 则分别是两个性状的样本平均数。

在统计分析中，群体平均数度量了群体中所有个体的平均表现，而群体方差和协方差则度量群体中个体的变异程度。数量遗传学是一门研究生物数量性状变异的遗传规律的学科。因此对数量性状方差和协方差的估算和分析是数量遗传分析的基础。数量遗传学运用统计分析的方法，研究生物体所表现的变异(即表现型变异)中归因于遗传效应 (genetic effect) 和环境效应 (environment effect) 的分量 (component)，并进一步分解遗传变异中基因效应 (gene effect) 的变异分量以及环境变异中的分量。

Johannsen(1909)通过对不同纯系的菜豆粒重变异研究，发现性状的表现型变异归因于纯系之间的变异和纯系内的变异。前者可以稳定遗传，属于遗传变异 (genetic variation); 后者是环境因素造成的，不能遗传，故称环境变异 (environment variation)。如果某性状的表现受到遗传效应 (G) 和环境随机误差效应 (e) 的影响，某个体的表现型值 y 可以用一个简单的线性公式表示， $y = \mu + G + e$ 。在这个简式中 μ 是群体平均数，为固定效应 (fixed effect); G 是遗传效应， e 是环境误差效应， G 和 e 均为随机效应 (random effect)。因为环境中的随机误差效应与生物体内的遗传效应是相互独立的两种效应，所以

$$\begin{aligned} \text{Var}(y) &= \text{Var}(\mu + G + e) = \text{Var}(\mu) + \text{Var}(G) + \text{Var}(e) \\ &= \text{Var}(G) + \text{Var}(e) = V_G + V_e \end{aligned}$$

这表明性状的表现型方差 (V_P) 是遗传方差 (V_G) 和环境误差方差 (V_e) 的和，

$$V_P = V_G + V_e$$

East(1916)首先采用分离世代(F_2)和不分离世代(P_1 、 P_2 、 F_1)的表现型方差,估算烟草花器长度的遗传方差。杂种二代(F_2)是基因型分离的世代,其表现型方差包括了遗传方差和环境机误方差, $V_{F_2} = V_G + V_e$ 。而亲本(P_1 、 P_2)和杂种一代(F_1)是基因型不分离的世代,其表现型方差只包括环境机误方差, $(V_{P_1} + V_{P_2} + V_{F_1})/3 = V_e$ 。因此遗传方差可以由这四个世代的表現型方差估算,

$$\hat{V}_G = \hat{V}_{F_2} - (\hat{V}_{P_1} + \hat{V}_{P_2} + \hat{V}_{F_1})/3$$

对于成对性状而言,性状的表现型协方差(C_P)是遗传协方差(C_G)和机误协方差(C_e)的和, $C_P = C_G + C_e$ 。如果求得以上四个遗传世代成对性状表现型协方差,遗传协方差可按式估算,

$$\hat{C}_G = \hat{C}_{F_2} - (\hat{C}_{P_1} + \hat{C}_{P_2} + \hat{C}_{F_1})/3$$

遗传效应是生物体内控制数量性状的各种基因效应相互作用的综合结果,因而遗传方差和遗传协方差是各种遗传变异分量的组成。数量遗传分析的一项重要任务就是采用统计方法分析特定的遗传群体,进一步将遗传方差和遗传协方差分解为归因于基因不同效应(例如加性效应、显性效应、上位性效应等)的遗传方差分量(genetic variance components)和遗传协方差分量(genetic covariance components)。

第二章 近交系数和共同祖先系数

第一节 亲缘关系的度量

在数量遗传研究中，常需要按一定的试验设计做一些杂交组配。在这些杂交试验中，遗传材料的个体之间可能存在不同程度的亲缘关系。为了估算遗传方差分量和遗传协方差分量，必须首先了解遗传试验中不同个体之间亲缘关系的程度。

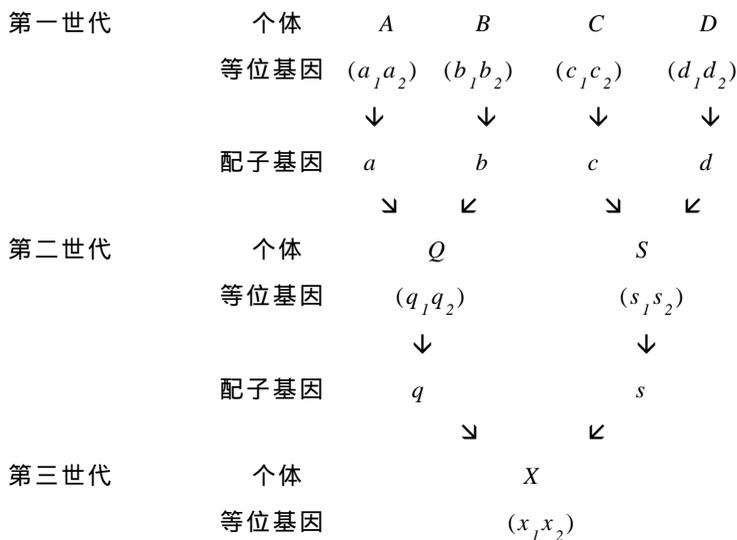


图2.1 三个世代的系谱关系图

图2.1是三个世代六个个体的系谱，并注明了每个个体的等位基因(allele)和配子基因。在第一世代中有四个亲本个体(A、B、C、D)。亲本A在某一基因位点上有等位基因 a_1 和 a_2 ，A产生的单倍体配子具有基因 a_1 或 a_2 的机会各为50%，可以用a表示。其它个体也用相同的原理标记其等位基因和配子基因。亲本A和B交配产生后代Q。个体Q有等位基因 q_1 和 q_2 ，并产生具有基因q的配子。亲本C和D交配产

生后代 S 。在第二世代中，亲本 Q 和 S 交配产生后代 X 。个体 X 具有等位基因 x_1 和 x_2 。

第一世代中的亲本之间可能存在不同程度的亲缘关系，这将导致第二世代中的个体 Q 和 S 之间也可能存在一些亲缘关系。在数量遗传和群体遗传中常用共同祖先系数(coefficient of coancestry)来度量两个个体之间的亲缘程度。个体 Q 和个体 S 之间的共同祖先系数 θ_{QS} 定义为，

$$\theta_{QS} = \Pr(q \quad s)$$

其中 $\Pr(q \quad s)$ 表示个体 Q 的配子所携带的基因 q 与个体 S 的配子所携带的基因 s 来源于同一祖先某个基因的概率。

因为基因 q 可能是等位基因 q_1 或 q_2 ，基因 s 也可能是等位基因 s_1 或 s_2 ，共同祖先系数 θ_{QS} 实际上是个体 Q 和个体 S 的等位基因来源于共同祖先某个基因的概率，或者是个体 Q 的亲本(A, B)和个体 S 的亲本(C, D)之间共同祖先系数的平均值。共同祖先系数 θ_{QS} 可以用个体 Q 的亲本和个体 S 的亲本共同祖先系数来表示，

$$\begin{aligned}\theta_{QS} &= \frac{1}{4} [\Pr(q_1 \quad s_1) + \Pr(q_1 \quad s_2) + \Pr(q_2 \quad s_1) + \Pr(q_2 \quad s_2)] \\ &= \frac{1}{4} [\Pr(a \quad c) + \Pr(a \quad d) + \Pr(b \quad c) + \Pr(b \quad d)] \\ &= \frac{1}{4} (\theta_{AC} + \theta_{AD} + \theta_{BC} + \theta_{BD})\end{aligned}$$

在实际应用时，常可采用不同的计算方法，

$$\begin{aligned}\theta_{QS} &= \frac{1}{2} (\theta_{AS} + \theta_{BS}) \\ &= \frac{1}{2} \left[\frac{1}{2} (\theta_{AC} + \theta_{AD}) + \frac{1}{2} (\theta_{BC} + \theta_{BD}) \right] \\ &= \frac{1}{4} (\theta_{AC} + \theta_{AD} + \theta_{BC} + \theta_{BD})\end{aligned}$$

常用近交系数(inbreeding coefficient)度量个体内某个基因位点

(locus)上两个等位基因来源于共同祖先某个基因的概率。以图2.1的系谱为例，个体X的近交系数 F_X 定义为

$$F_X = \Pr(x_1 = x_2) = \Pr(q = s) = \theta_{QS}$$

由此可见某个体的近交系数等于该个体的双亲之间的共同祖先系数。可以用这一关系式来推导近交系数值。现以遗传研究中常用的几种交配类型，举例说明近交系数的计算方法。

第二节 近交系数的计算

例2.1.半同胞交配(half-sib mating, 简称HS)

在图2.2中如果亲本A、B和C没有亲缘关系，半同胞Q和S的后代X的近交系数 F_X 可以按下式计算，

$F_X = \Pr(x_1 = x_2)$ $= \Pr(q = s) = \theta_{QS}$ $(c_1 c_2)$ $= \frac{1}{4} [\Pr(a = c) + \Pr(a = b) + \Pr(b = c) + \Pr(b = b)]$ $= \frac{1}{4} (\theta_{AC} + \theta_{AB} + \theta_{BC} + \theta_{BB})$ $= \frac{1}{4} (0 + 0 + 0 + \theta_{BB})$ $= \frac{1}{4} \theta_{BB}$ $= \frac{1}{16} [\Pr(b_1 = b_1) + \Pr(b_2 = b_2) + 2\Pr(b_1 = b_2)]$ $= \frac{1}{16} (1 + 1 + 2 F_B) = \frac{1}{8} (1 + F_B)$	<table style="margin: auto;"> <tr> <td style="padding: 5px;">A</td> <td style="padding: 5px;">B</td> <td style="padding: 5px;">C</td> </tr> <tr> <td style="padding: 5px;">$(a_1 a_2)$</td> <td style="padding: 5px;">$(b_1 b_2)$</td> <td></td> </tr> <tr> <td style="padding: 5px;">↓</td> <td style="padding: 5px;">↙ ↘</td> <td style="padding: 5px;">↓</td> </tr> <tr> <td style="padding: 5px;">a</td> <td style="padding: 5px;">b</td> <td style="padding: 5px;">b</td> </tr> <tr> <td style="padding: 5px;">↘</td> <td style="padding: 5px;">↙ ↘</td> <td style="padding: 5px;">↙</td> </tr> <tr> <td></td> <td style="padding: 5px;">Q</td> <td style="padding: 5px;">S</td> </tr> <tr> <td></td> <td style="padding: 5px;">$(q_1 q_2)$</td> <td style="padding: 5px;">$(s_1 s_2)$</td> </tr> <tr> <td></td> <td style="padding: 5px;">↓</td> <td style="padding: 5px;">↓</td> </tr> <tr> <td></td> <td style="padding: 5px;">q</td> <td style="padding: 5px;">s</td> </tr> <tr> <td></td> <td style="padding: 5px;">↘</td> <td style="padding: 5px;">↙</td> </tr> <tr> <td></td> <td colspan="2" style="padding: 5px;">X</td> </tr> <tr> <td></td> <td colspan="2" style="padding: 5px;">$(x_1 x_2)$</td> </tr> </table>	A	B	C	$(a_1 a_2)$	$(b_1 b_2)$		↓	↙ ↘	↓	a	b	b	↘	↙ ↘	↙		Q	S		$(q_1 q_2)$	$(s_1 s_2)$		↓	↓		q	s		↘	↙		X			$(x_1 x_2)$	
A	B	C																																			
$(a_1 a_2)$	$(b_1 b_2)$																																				
↓	↙ ↘	↓																																			
a	b	b																																			
↘	↙ ↘	↙																																			
	Q	S																																			
	$(q_1 q_2)$	$(s_1 s_2)$																																			
	↓	↓																																			
	q	s																																			
	↘	↙																																			
	X																																				
	$(x_1 x_2)$																																				

图2.2 半同胞交配

系谱图

例2.2.全同胞交配(full-sib mating, 简称FS)

在图2.3中如果亲本A和B没有亲缘关系, 全同胞Q和S的后代X的近交系数 F_X 可以按下式计算,

$$\begin{aligned}
 F_X &= \Pr(x_1 = x_2) \\
 &= \Pr(q = s) = \theta_{QS} \\
 &= \frac{1}{4} [\Pr(a = a) + 2\Pr(a = b) + \Pr(b = b)] \\
 &= \frac{1}{4} (\theta_{AA} + \theta_{AB} + \theta_{AB} + \theta_{BB}) \\
 &= \frac{1}{4} (\theta_{AA} + \theta_{BB}) \\
 &= \frac{1}{8} (2 + F_A + F_B)
 \end{aligned}$$

图2.3 全同胞交配

系谱图

例2.3.自交(selfing, 简称S)

图2.4中X是亲本Q的自交后代, X的近交系数 F_X 可按下式计算,

$$\begin{aligned}
 F_X &= \Pr(x_1 = x_2) \\
 &= \Pr(q = q) \\
 &= \theta_{QQ} \\
 &= \frac{1}{2} (1 + F_Q)
 \end{aligned}$$

图2.4 自交系谱

图

第三节 遗传协方差的计算

如果不同个体之间存在亲缘关系，那么个体之间可能在遗传上具有不同程度的相似表现。在数量遗传分析中，常用遗传协方差度量个体间的遗传相似性。

在遗传分量估算中，常需要计算不同交配世代个体间的协方差分量。估算近亲个体之间遗传协方差分量时，不但需要计算个体之间的共同祖先系数及亲本的近交系数，还要求计算个体之间某基因位点上两个等位基因同时来源于共同祖先的概率。

如果个体 Q 的双亲是 A 和 B ，而个体 S 的双亲是 C 和 D (见图2.1)，个体 Q 和 S 分别具有等位基因 (q_1q_2) 和 (s_1s_2) ，则概率 Δ_{QS} 可由下式定义，

$$\begin{aligned}\Delta_{QS} &= \Pr(q_1 \ s_1 \ q_2 \ s_2) + \Pr(q_1 \ s_2 \ q_2 \ s_1) \\ &= \Pr(q_1 \ s_1)\Pr(q_2 \ s_2) + \Pr(q_1 \ s_2)\Pr(q_2 \ s_1) \\ &= (\theta_{AC}\theta_{BD} + \theta_{AD}\theta_{BC})\end{aligned}$$

任何个体之间(如 Q 与 S)的加性和显性协方差分量可由以下通式求得，

$$\text{Cov}(Q, S) = 2\theta_{QS}V_A + \Delta_{QS}V_D$$

如果个体 Q 的双亲是 A 和 B ，而个体 S 的双亲是 C 和 D (见图2.1)，那么个体 Q 和 S 之间协方差的遗传分量可以分解为，

$$\text{Cov}(Q, S) = \frac{1}{2}(\theta_{AC} + \theta_{BD} + \theta_{AD} + \theta_{BC})V_A + (\theta_{AC}\theta_{BD} + \theta_{AD}\theta_{BC})V_D$$

在遗传分量估算中，常需要计算半同胞和全同胞的协方差分量。例2.1中半同胞 Q 和 S 的协方差的计算公式是

$$\text{Cov}(Q, S) = C_{HS} = \frac{1}{4}(1 + F_B)V_A,$$

其中 C_{HS} 是半同胞协方差， V_A 是加性遗传方差分量。

例2.2中全同胞Q和S的协方差的计算公式是

$$\text{Cov}(Q, S) = C_{FS} = \frac{1}{4}(2 + F_A + F_B)V_A + \frac{1}{4}(1 + F_A)(1 + F_B)V_D,$$

其中 C_{FS} 是全同胞协方差, V_D 是显性遗传方差分量。

在数量遗传分析中常需要计算杂交亲本及其子代的协方差。杂交亲本与子代的协方差可按下式计算,

$$\begin{aligned} \text{Cov}(A, Q) &= 2\theta_{AQ}V_A + \Delta_{AQ}V_D && \begin{matrix} X & & Y \\ & \searrow & \swarrow \\ & & A & & B \end{matrix} \\ &= 2\left(\frac{1}{2}\right)(\theta_{AA} + \theta_{AB})V_A && \\ &\quad + (\theta_{XA}\theta_{YB} + \theta_{XB}\theta_{YA})V_D && \\ &= 2\left(\frac{1}{4}\right)(1 + F_A)V_A + (0 + 0)V_D && \begin{matrix} & \searrow & \swarrow \\ & & Q \end{matrix} \\ &= \frac{1}{2}(1 + F_A)V_A && \\ &= \text{Cov}(B, Q) \end{aligned}$$

$F_1(A \text{ 或 } B)$ 与 F_2 子代(Q) 的协方差计算公式是

$$\begin{aligned} \text{Cov}(A, Q) &= 2\theta_{AQ}V_A + \Delta_{AQ}V_D && \begin{matrix} X & & Y & & X \\ & & Y & & \\ & \searrow & \swarrow & & \searrow & \swarrow \\ & & A & & & A \end{matrix} \\ &= 2\left(\frac{1}{2}\right)(\theta_{AA} + \theta_{AA})V_A && \\ &\quad + (\theta_{XA}\theta_{YA} + \theta_{XA}\theta_{YA})V_D && \\ &= 2\left(\frac{1}{2}\right)(1 + F_A)V_A + 2\theta_{XA}\theta_{YA}V_D && \begin{matrix} & \searrow & \swarrow \\ & & Q \end{matrix} \\ &= (1 + F_A)V_A + 2\left[\frac{1}{2}(\theta_{XX} + \theta_{XY}) \times \frac{1}{2}(\theta_{XY} + \theta_{YY})\right]V_D && \\ &= (1 + F_A)V_A + \frac{1}{8}(1 + F_X)(1 + F_Y)V_D && \\ &= \text{Cov}(B, Q) \end{aligned}$$

轮回亲本与回交后代的协方差计算:

$$\text{Cov}(A, Q) = 2\theta_{AQ}V_A + \Delta_{AQ}V_D$$

X Y

其中

$$\theta_{AQ} = \frac{1}{2}(\theta_{AA} + \theta_{AC})$$

A B

$$= \frac{1}{2}[\theta_{AA} + \frac{1}{2}(\theta_{AA} + \theta_{AB})]$$

A C

$$= \frac{1}{4}(3\theta_{AA} + \theta_{AB})$$

A C

$$= \frac{3}{8}(1 + F_A)$$

Q

$$AQ = (\theta_{XA}\theta_{YC} + \theta_{YA}\theta_{XC})$$

$$= [\theta_{XA}(\frac{1}{2})(\theta_{YA} + \theta_{YB}) + \theta_{YA}(\frac{1}{2})(\theta_{XA} + \theta_{XB})]$$

$$= \frac{1}{2}(\theta_{XA}\theta_{YA} + \theta_{XA}\theta_{YB} + \theta_{YA}\theta_{XA} + \theta_{YA}\theta_{XB})$$

$$= (\theta_{XA}\theta_{YA})$$

$$= [\frac{1}{4}(1 + F_X) \times \frac{1}{4}(1 + F_Y)]$$

$$= \frac{1}{16}(1 + F_X)(1 + F_Y)$$

所以 $\text{Cov}(A, Q) = 2[\frac{3}{8}(1 + F_A)V_A] + \frac{1}{16}(1 + F_X)(1 + F_Y)V_D$

$$= \frac{3}{4}(1 + F_A)V_A + \frac{1}{16}(1 + F_X)(1 + F_Y)V_D$$

第三章 广义遗传模型

第一节 Cockerham的广义遗传模型

如果试验在一年一地实施、控制某性状的遗传效应只包括基因的加性效应和显性效应，那么纯合亲本*i*和纯合亲本*j*的 F_1 个体*k*的性状观察值*y*可以由以下遗传模型表示，

$$\begin{aligned} y_{ijk} &= \mu + G_{ij} + e_{ijk} \\ &= \mu + A_i + A_j + D_{ij} + e_{ijk} \end{aligned}$$

其中 μ 是群体平均数， G_{ij} 是归因于亲本*i*和亲本*j*的基因遗传效应值， A_i (或 A_j)是归因于亲本*i*(或亲本*j*)基因的加性效应值， D_{ij} 是 F_1 (*ij*) 中来源于亲本*i*基因和亲本*j*基因的显性效应值， e_{ijk} 是 F_1 个体*k*的环境机误。

由一组随机的纯合亲本衍生的各世代个体的遗传效应值(G)都可以用一个广义遗传模型分解其不同的基因效应(Cockerham, 1980)。当一个遗传群体的遗传变异归因于若干不同的纯合亲本来源的基因效应，由这个遗传群体衍生的任何个体的基因型值*G*可由以下式子表示，

$$G = \sum_i \alpha_i A_i + \sum_i \sum_{j \geq i} \delta_{ij} D_{ij} \quad (3.1)$$

其中 α_i 是加性效应值的系数值 ($\sum_i \alpha_i = 2$)，

δ_{ij} 是显性效应值的系数值 ($\sum_i \sum_{j \geq i} \delta_{ij} = 1$)。

根据这一广义遗传模型，可以写出任何交配世代的基因效应值。由纯合亲本 *i*、*j* 和 *k* 产生的三交种 $i \times (j \times k)$ 的世代平均数，其总遗传效应值 $G_{i(jk)}$ 可以分解为，

$$G_{i(jk)} = A_i + \frac{1}{2} A_j + \frac{1}{2} A_k + \frac{1}{2} D_{ij} + \frac{1}{2} D_{ik}$$

由杂种一代 F_1 与纯合亲本 *i* 回交产生的回交一代 $i \times (i \times j)$ 平均数，其总遗传效应

值 $G_{i(ij)}$ 可以分解为，

$$G_{i(ij)} = \frac{3}{2}A_i + \frac{1}{2}A_j + \frac{1}{2}D_{ii} + \frac{1}{2}D_{ij}$$

采用同样的原理，可以写出不同世代自交后代的平均数。由于 F_1 是不分离的基因型， F_1 个体的自交或个体间随机杂交都能产生具有相同遗传组成的杂种二代。

杂种二代的总遗传效应值 G_{ij}^S 可分解为

$$G_{ij}^S = A_i + A_j + \frac{1}{4}D_{ii} + \frac{1}{4}D_{jj} + \frac{1}{2}D_{ij}$$

由纯合亲本 i 、 j 和 k 产生的三交种群体 $i \times (j \times k)$ 中有频率相等的二种基因型(G_{ij} 和 G_{ik})，其自交一代平均数的基因效应为

$$G_{i(jk)}^S = A_i + \frac{1}{2}A_j + \frac{1}{2}A_k + \frac{1}{4}D_{ii} + \frac{1}{8}D_{jj} + \frac{1}{8}D_{kk} + \frac{1}{4}D_{ij} + \frac{1}{4}D_{ik}$$

对于某些性状，遗传效应除了有加性效应和显性效应以外，还可能包括上位性效应(epistatic effect)。Fisher(1918)首先提出将总遗传效应分解为加性效应(A)、显性效应(D)和上位性效应(I)。Cockerham(1954)进一步把上位性效应分解为加性与加性的上位性效应(AA)、加性与显性的上位性效应(AD)和显性与显性的性上位效应(DD)。因而总遗传效应值可以进一步扩展，包括加 \times 加上位、加 \times 显上位和显 \times 显上位三种上位性效应，

$$G = \sum_i \alpha_i A_i + \sum_i \sum_{j \geq i} \delta_{ij} D_{ij} + (\sum_i \alpha_i A_i)^2 + (\sum_i \alpha_i A_i)(\sum_i \sum_{j \geq i} \delta_{ij} D_{ij}) + (\sum_i \sum_{j \geq i} \delta_{ij} D_{ij})^2$$

加性 显性 加 \times 加上位性 加 \times 显上位性 显 \times 显上位性

由此可以写出不同遗传群体总遗传效应值的包括上位性效应的分解公式。由于加 \times 显上位性和显 \times 显上位性比较复杂，在动植物遗传分析中一般均比其它效应弱，因此常可忽略不计。加 \times 加上位性是可经选择固定的遗传效应，常有必要与加性和显性效应一齐分析。包括加性、显性和加 \times 加上位性效应的遗传分量为，

$$G = \sum_i \alpha_i A_i + \sum_i \sum_{j \geq i} \delta_{ij} D_{ij} + (\sum_i \alpha_i A_i)^2 \tag{3.2}$$

加性 显性 加 \times 加上位性

现举例说明一些世代遗传效应的分解，

$$\text{杂种一代: } G_{ij} = (A_i + A_j) + D_{ij} + (AA_{ii} + AA_{jj} + 2AA_{ij})$$

加性 显性 加 × 加上位性

$$\text{三交种一代: } G_{i(jk)} =$$

$$(A_i + \frac{1}{2}A_j + \frac{1}{2}A_k) + (\frac{1}{2}D_{ij} + \frac{1}{2}D_{ik}) + (AA_{ii} + \frac{1}{4}AA_{jj} + \frac{1}{4}AA_{kk} + AA_{ji} + AA_{ik} + \frac{1}{2}AA_{jk})$$

加性 显性 加 × 加上位性

第二节 分析种子性状的广义遗传模型

农作物品质育种的深入发展，促进了种子数量性状的遗传研究。种子的营养物质由母体植株所提供，因此种子的数量性状表现可能会同时受到种子核基因和母体植株核基因两套遗传体系的控制。另外细胞质基因也可能通过控制叶绿体(或线粒体)的合成而影响植株的光合(或呼吸)作用，从而间接控制种子性状的表现。当分析种子性状时，Cockerham(1980)的广义遗传模型需要进一步扩展。Zhu和Weir(1994a)提出了包括种子核基因、细胞质基因和母体核基因遗传效应的广义遗传模型，

$$G = G_O + G_C + G_M \tag{3.3}$$

上式中的三项遗传效应可进一步分解为，

$$\text{直接遗传效应 } G_O = \sum_i \tau_i A_i + \sum_i \sum_{j \geq i} \delta_{ij} D_{ij}$$

$$\text{细胞质遗传效应 } G_C = \sum_i C_i$$

$$\text{母体遗传效应 } G_M = \sum_i \tau_{m_i} A_{m_i} + \sum_i \sum_{j \geq i} \delta_{m_{ij}} D_{m_{ij}}$$

其中 A_i 是种子核基因的直接加性效应， D_{ij} 是种子核基因的直接显性效应； C_i 是细胞质基因效应； A_{m_i} 是母体植株核基因的母体加性效应， $D_{m_{ij}}$ 是母体植株核基因的母

体显性效应。

种子性状的广义遗传模型，为发展具有生物学意义的双子叶作物二倍体种子模型(朱军，1992；Zhu和Weir，1994a)和单子叶作物三倍体胚乳模型(朱军，1992；Zhu和Weir，1994b；朱军和许馥华，1994)奠定了基础。

作物二倍体种子模型可以包括三个遗传世代，各组合世代遗传效应的分解为，

$$\text{亲本}P_i\text{遗传效应分量} = 2A_i + D_{ii} + C_i + 2Am_i + Dm_{ii}$$

杂种一代 F_{1ij} 遗传效应分量

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

杂种二代 F_{2ij} 遗传效应分量

$$= A_i + A_j + 0.25D_{ii} + 0.25D_{jj} + 0.5D_{ij} \\ + C_i + Am_i + Am_j + Dm_{ij}$$

分析作物三倍体胚乳性状的遗传模型可以只包括三个遗传世代，各组合世代遗传效应可分解为，

$$\text{亲本}P_i\text{遗传效应分量} = 3A_i + 3D_{ii} + C_i + 2Am_i + Dm_{ii}$$

杂种一代 F_{1ij} 遗传效应分量

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

杂种二代 F_{2ij} 遗传效应分量

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

第三节 包括基因型与环境互作效应的广义遗传模型

以上所介绍的广义遗传模型及其扩展模型都基于遗传假设：不存在基因型与环境的互作效应。由于生物体的基因是在环境条件作用下表达的，环境条件的不同

可能导致基因表达方式或程度的差异。已有许多研究表明,不少数量性状存在基因型与环境的互作效应。如果互作确实存在,则需要不同的环境条件下实施遗传实验,并采用包括互作效应的遗传模型,才能无偏地分析基因效应以及基因与环境的互作效应。

因此,广义遗传模型需要在原有的基因效应基础上进一步扩展,再包括基因各项效应与环境的互作效应。当遗传试验在多个环境条件下实施时,某遗传群体在一种环境下的性状平均观察值 y 可以由以下遗传模型表示,

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

其中 μ 是群体平均数, E 是环境效应, G 是基因的遗传效应值, GE 是基因与环境的互作效应值, e 是机误。

Cockerham的广义遗传模型(1980)可扩展为 $G + GE$ (朱军,1994),基因遗传效应值 G 的定义如模型(3.1)所示,基因与环境互作效应值 GE 的定义为,

$$GE = \sum_h \sum_i \alpha_{hi} AE_{hi} + \sum_h \sum_i \sum_{j \geq i} \beta_{hij} DE_{hij} \quad (3.4)$$

其中 AE_{hi} 是加性 \times 环境互作效应, DE_{hij} 是显性 \times 环境互作效应。

根据加性-显性遗传模型,亲本、 F_1 和 F_2 在环境 h 中的基因型与环境互作效应为,

$$\text{亲本和杂种一代: } GE_{hij} = AE_{hi} + AE_{hj} + DE_{hij}$$

$$\text{交种二代: } GE_{hij} = AE_{hi} + AE_{hj} + 0.25DE_{hii} + 0.25DE_{hjj} + 0.5DE_{hij}$$

当广义遗传模型(3.2)扩展为 $G + GE$ 时,基因遗传效应值 G 的定义如模型(3.2)所示,基因与环境互作效应值 GE 可分解为加性 \times 环境互作(AE)、显性 \times 环境互作(DE)和加 \times 加上位性 \times 环境互作(AAE)等效应的遗传分量,

$$GE = \sum_h \sum_i \alpha_{hi} AE_{hi} + \sum_h \sum_i \sum_{j \geq i} \beta_{hij} DE_{hij} + \left(\sum_h \sum_i \alpha_{hi} AE_{hi} \right)^2 \quad (3.5)$$

根据广义遗传模型(3.5),可以利用双列杂交的三个世代(亲本、 F_1 和 F_2)在若干环境下的遗传实验资料,分析加性 \times 环境、显性 \times 环境和加 \times 加 \times 环境的各项遗传

效应。亲本、 F_1 和 F_2 的遗传效应 G_{ij} 的分解已有定义(朱军, 1992), 基因效应与环境交互效应 GE_{hij} 的分解因组合及世代而异,

亲本($i=j$)和杂种一代($i \neq j$):

$$GE_{ij} = (AE_{hi} + AE_{hj}) + DE_{hij} + (AAE_{hii} + AAE_{hij} + 2AAE_{hij})$$

杂种二代:

$$GE_{hij} = (AE_{hi} + AE_{hj}) + (0.25DE_{hii} + 0.25DE_{hij} + 0.5DE_{hij}) + (AAE_{hii} + AAE_{hij} + 2AAE_{hij})$$

如分析种子性状的广义遗传模型(3.3)作相应的扩展($G + GE$), 包括基因型与环境的交互效应 GE , 交互总效应 GE 是三项分量的组成(朱军, 1994),

$$GE = G_OE + G_C E + G_M E \quad (3.6)$$

上式中的三项交互效应可进一步分解为,

$$\text{直接交互效应 } G_OE = \sum_h \sum_i h_i AE_{hi} + \sum_h \sum_i \sum_{j \geq i} h_{ij} DE_{hij}$$

$$\text{细胞质交互效应 } G_C E = \sum_h \sum_i \lambda_{hi} CE_{hi}$$

$$\text{母体交互效应 } G_M E = \sum_h \sum_i \alpha_{m_{hi}} AmE_{hi} + \sum_h \sum_i \sum_{j \geq i} \beta_{m_{hij}} DmE_{hij}$$

其中 AE_{hi} 是种子核基因的直接加性 × 环境交互效应, DE_{hij} 是种子核基因的直接显性 × 环境交互效应; CE_{hi} 是细胞质基因 × 环境交互效应; AmE_{hi} 是母体植株核基因的母体加性 × 环境交互效应, DmE_{hij} 是母体植株核基因的母体显性 × 环境交互效应。

对于双子叶作物种子数量性状, 母本 i 与父本 j 的第 k 种交配类型在环境 h 内的基因与环境交互效应值 GE_{hijk} 的分解如下式所示。

亲本 P_i 交互效应分量($k = 0$)

$$GE_{hii0} = 2AE_{hi} + DE_{hii} + CE_{hi} + 2AmE_{hi} + DmE_{hii}$$

杂种一代 $F_{1ij}(P_i \times P_j)$ 交互效应分量($k = 1$)

$$GE_{hij1} = AE_{hi} + AE_{hj} + DE_{hij} + CE_{hi} + 2AmE_{hi} + DmE_{hii}$$

杂种二代 F_{2ij} 互作效应分量($k=2$)

$$GE_{hij2} = AE_{hi} + AE_{hj} + 0.25DE_{hii} + 0.25DE_{hij} + 0.5DE_{hij} \\ + CE_{hi} + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

而对于单子叶作物胚乳数量性状，各遗传世代基因与环境互作效应值 GE_{hijk} 的分解如下式所示。

亲本 P_i 互作效应分量($k=0$)

$$GE_{hii0} = 3AE_{hi} + 3DE_{hii} + CE_{hi} + 2AmE_{hi} + DmE_{hii}$$

杂种一代 F_{1ij} ($P_i \times P_j$)互作效应分量($k=1$)

$$GE_{hij1} = 2AE_{hi} + AE_{hj} + DE_{hii} + 2DE_{hij} + CE_{hi} + 2AmE_{hi} + DmE_{hii}$$

杂种二代 F_{2ij} 互作效应分量($k=2$)

$$GE_{hij2} = 1.5AE_{hi} + 1.5AE_{hj} + DE_{hii} + DE_{hij} + DE_{hij} \\ + CE_{hi} + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

根据广义遗传模型(3.6)所定义的各项遗传参数，利用若干环境下遗传实验(在田间种植亲本和 F_1 ，由其自交产生亲本种子和 F_2 种子，亲本间杂交可获得 F_1 种子)的资料，便可有效地分析控制种子性状的三套遗传体系的基因效应及其与环境的互作效应。

在广义遗传模型中遗传效应可以定义为随机效应，也可以定义为固定效应。环境效应通常指试验方法、试点或年份，设置的数目一般比较少，故可定为固定效应。如果遗传材料是一组随机样本，试验目的主要是分析性状的遗传变异表现，则将遗传效应定义为随机效应，并且遗传 \times 环境互作效应也是随机效应。这时可以估算各项随机效应的方差分量或协方差分量。如果试验材料是一组特定的遗传材料，试验目的主要是分析不同遗传材料的特定遗传表现，可将遗传效应定义为固定效应，并且遗传 \times 环境互作效应也是固定效应。这时应该估算各项随机效应的数值。

任何广义遗传模型都可以用一个混合线性模型(mixed linear model)表示，

$$\begin{aligned} \mathbf{y} &= \mathbf{X}_1 \mathbf{b}_1 + \mathbf{X}_2 \mathbf{b}_2 + \cdots + \mathbf{X}_n \mathbf{b}_n + \mathbf{U}_1 \mathbf{e}_1 + \mathbf{U}_2 \mathbf{e}_2 + \cdots + \mathbf{U}_m \mathbf{e}_m \\ &= \sum_{i=1}^n \mathbf{X}_i \mathbf{b}_i + \sum_{u=1}^m \mathbf{U}_u \mathbf{e}_u \\ &= \mathbf{Xb} + \mathbf{Ue} \end{aligned}$$

其中 \mathbf{b}_i 是第 i 项固定效应向量, \mathbf{X}_i 是第 i 项固定效应的系数矩阵; \mathbf{e}_u 是第 u 项独立随机变量向量, 具有平均数零、方差 $\sigma_u^2 \mathbf{I}$, \mathbf{U}_u 是第 u 项随机效应的系数矩阵。

第四章 遗传模型的常用统计分析方法

第一节 遗传方差和协方差估算

1. ANOVA法

采用不同的统计分析方法，可以估算混合线性模型中的方差分量。早期采用的方法主要是方差分析(analysis of variance, 简称ANOVA)的方法。该方法具有许多优点，如运算简便、估算值无偏等。但是ANOVA法也存在一些局限性，比如不能有效地分析不规则缺失的非平衡数据(unbalanced data)，也无法分析复杂的遗传模型。尽管存在这些限制，ANOVA法仍然是目前运用最广泛的一种方法。运用ANOVA方法，一般都需要先计算各项效应的平方和(sum of squares)和均方(mean squares)。这些均方的期望值是方差分量的线性函数。设各项均方等于其期望值，可以得到方差分量线性函数的一组联立方程。解这一联立方程组，便可获得各项方差分量的无偏估算值。

现以某组遗传材料在一年多试点的资料分析为例，介绍各项方差分量的估算方法。一年多试点遗传试验统计分析模型可由下式表示，

$$y_{ijk} = \mu + E_i + G_j + GE_{ij} + B_{k(i)} + e_{ijk}$$

其中

y_{ijk} 是第 j 种遗传材料 ($j = 1, 2, \dots, g$) 在第 i 个试点 ($i = 1, 2, \dots, l$) 内第 k 个随机区组 ($k = 1, 2, \dots, b$) 中的观察值；

μ 是群体平均数，固定效应； E_i 是试点固定效应；

G_j 是基因型遗传效应， $G_j \sim (0, \sigma_G^2)$ ；

GE_{ij} 是基因型 \times 试点互作随机效应， $GE_{ij} \sim (0, \sigma_{GE}^2)$ ；

$B_{k(i)}$ 是试点内随机区组效应， $B_{k(i)} \sim (0, \sigma_B^2)$ ；

e_{ijk} 是剩余效应, $e_{ijk} \sim (0, \sigma_e^2)$ 。

以上公式可以改写为

$$y_{ijk} - \mu = E_i + G_j + GE_{ij} + B_{k(i)} + e_{ijk}$$

如果公式中的所有效应都由其估计值取代, 则可获得总效应的分解

$y_{ijk} - \bar{y}_{...}$	总效应
$= (\bar{y}_{i..} - \bar{y}_{...})$	试点效应
$+ (\bar{y}_{.j.} - \bar{y}_{...})$	基因型效应
$+ (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})$	基因型 × 试点互作效应
$+ (\bar{y}_{i.k} - \bar{y}_{i..})$	区组效应
$+ (y_{ijk} - \bar{y}_{i.k} - \bar{y}_{ij.} + \bar{y}_{i..})$	机误效应

其中 $\bar{y}_{...}$ 是试验的总平均数, $\bar{y}_{.j.}$ 是基因型 j 的平均数, $\bar{y}_{i..}$ 是试点 i 的平均数, $\bar{y}_{ij.}$ 是基因型 j 在试点 i 的平均数, $\bar{y}_{i.k}$ 是试点 i 内区组 k 的品种平均数。

将上式两边平方后对所有的试点、遗传材料和区组累加, 经化简后可得到平方和的分解项

$\sum_{i=1}^l \sum_{j=1}^g \sum_{k=1}^b (y_{i..} - \bar{y}_{...})^2$	总平方和
$= gb \sum_{i=1}^l (\bar{y}_{i..} - \bar{y}_{...})^2$	试点平方和
$+ lb \sum_{j=1}^g (\bar{y}_{.j.} - \bar{y}_{...})^2$	基因型平方和
$+ b \sum_{i=1}^l \sum_{j=1}^g (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$	基因型 × 试点互作平方和
$+ g \sum_{i=1}^l \sum_{k=1}^b (\bar{y}_{i.k} - \bar{y}_{i..})^2$	区组平方和
$+ \sum_{i=1}^l \sum_{j=1}^g \sum_{k=1}^b (y_{ijk} - \bar{y}_{i.k} - \bar{y}_{ij.} + \bar{y}_{i..})^2$	机误平方和

或者简写为

$$SSTC = SSE + SSG + SSGE + SSB + SSe$$

各项自由度的相应分解为

$$\begin{aligned} df_{TC} &= df_E + df_G + df_{GE} + df_B + df_e \\ &= (l-1) + (g-1) + (l-1)(g-1) + l(b-1) + l(g-1)(b-1) \end{aligned}$$

以上结果可以归纳为方差分析表4.1。

表4.1 一年多点品种试验的方差分析表

变异原因	自由度	平方和	均方	期望均方
试点	$l-1$	SSE	MSE	
基因型	$g-1$	SSG	MSG	$\sigma_e^2 + bl\sigma_G^2$
基因型 × 试点	$(g-1)(l-1)$	SSGE	MSGE	$\sigma_e^2 + b\sigma_{GE}^2$
区组	$l(b-1)$	SSB	MSB	$\sigma_e^2 + g\sigma_B^2$
机误	$l(g-1)(b-1)$	SSe	MSe	σ_e^2

注: g = 品种数, l = 试点数, b = 试点内区组数。

各项方差分量可由解以下等式而求得,

$$MSG = SSG/(g-1) = \hat{\sigma}_e^2 + b\hat{\sigma}_{GE}^2 + bl\hat{\sigma}_G^2$$

$$MSGE = SSGE/[(g-1)(l-1)] = \hat{\sigma}_e^2 + b\hat{\sigma}_{GE}^2$$

$$MSB = SSB/[l(b-1)] = \hat{\sigma}_e^2 + g\hat{\sigma}_B^2$$

$$MSe = SSe/[l(g-1)(b-1)] = \hat{\sigma}_e^2。$$

方差分量的无偏估计值为

$$\hat{\sigma}_e^2 = MSe$$

$$\hat{\sigma}_B^2 = (MSB - MSe)/g$$

$$\hat{\sigma}_{GE}^2 = (MSGE - MSe)/b$$

$$\hat{\sigma}_G^2 = (MSG - MSGE)/bl。$$

多点遗传试验中成对性状的各项协方差分量也可以用协方差分析的方法估

算。在ANOVA法中，通过各项平方和可以估算方差分量，由成对性状的叉积和 (sum of cross products)也可以估算协方差分量。如果有成对性状X和性状Y，它们的协方差分析表列于表4.2。

表4.2 一年多点品种试验的协方差分析表

变异原因	自由度	叉积方和	均积	期望均积
试点	$l-1$	SCPE	MCPE	
基因型	$g-1$	SCPG	MCPG	$\sigma_{e(XY)} + bl\sigma_{G(XY)}$
基因型 × 试点	$(g-1)(l-1)$	SCPGE	MCPGE	$\sigma_{e(XY)} + b\sigma_{GE(XY)}$
区组	$L(b-1)$	SCPB	MCPB	$\sigma_{e(XY)} + g\sigma_{B(XY)}$
机误	$L(g-1)(b-1)$	SCPe	MCPe	$\sigma_{e(XY)}$

注: g =品种数, l =试点数, b =试点内区组数。

表4.2中的各项叉积和可以按下式计算:

$$\text{试点叉积和 SCPE} = gb \sum_{i=1}^l (\bar{y}_{i..(X)} - \bar{y}_{... (X)}) (\bar{y}_{i..(Y)} - \bar{y}_{... (Y)})$$

$$\text{基因型叉积和 SCPG} = lb \sum_{j=1}^g (\bar{y}_{.j.(X)} - \bar{y}_{... (X)}) (\bar{y}_{.j.(Y)} - \bar{y}_{... (Y)})$$

基因型 × 试点互作叉积和 SCPGE =

$$b \sum_{i=1}^l \sum_{j=1}^g (\bar{y}_{ij.(X)} - \bar{y}_{i..(X)} - \bar{y}_{.j.(X)} + \bar{y}_{... (X)}) (\bar{y}_{ij.(Y)} - \bar{y}_{i..(Y)} - \bar{y}_{.j.(Y)} + \bar{y}_{... (Y)})$$

$$\text{区组叉积和 SCPB} = g \sum_{i=1}^l \sum_{k=1}^b (\bar{y}_{i.k(X)} - \bar{y}_{i..(X)}) (\bar{y}_{i.k(Y)} - \bar{y}_{i..(Y)})$$

机误叉积和 SCPe =

$$\sum_{i=1}^l \sum_{j=1}^g \sum_{k=1}^b (y_{ijk(X)} - \bar{y}_{i.k(X)} - \bar{y}_{ij.(X)} + \bar{y}_{i..(X)}) (y_{ijk(Y)} - \bar{y}_{i.k(Y)} - \bar{y}_{ij.(Y)} + \bar{y}_{i..(Y)})$$

表4.2中的各协项方差分量可由解以下等式而求得，

$$\text{MCPG} = \text{SCPG}/(g-1) = \hat{\sigma}_{e(XY)} + b\hat{\sigma}_{GE(XY)} + bl\hat{\sigma}_{G(XY)}$$

$$\text{MCPGE} = \text{SCPGE}/[(g-1)(l-1)] = \hat{\sigma}_{e(XY)} + b\hat{\sigma}_{GE(XY)}$$

$$\text{MCPB} = \text{SCPb}/[l(b-1)] = \hat{\sigma}_{e(\text{XY})} + g\hat{\sigma}_{b(\text{XY})}$$

$$\text{MCPe} = \text{SCPe}/[l(g-1)(b-1)] = \hat{\sigma}_{e(\text{XY})}$$

协方差分量的无偏估计值为

$$\hat{\sigma}_{e(\text{XY})} = \text{MCPe}$$

$$\hat{\sigma}_{B(\text{XY})} = (\text{MCPB} - \text{MSe})/g$$

$$\hat{\sigma}_{\text{GE}(\text{XY})} = (\text{MSGE} - \text{MSe})/b$$

$$\hat{\sigma}_{G(\text{XY})} = (\text{MSG} - \text{MSGE})/bl$$

当基因型为一组特定的遗传交配材料时，有可能进一步把总基型方差分解为加性、显性等遗传方差分量，并且基因型 × 试点互作方差也可被进一步分解。与遗传方差的分解一样，协方差也可以被分解成不同的协方差分量。以上介绍的ANOVA法虽然运算简便，但不能分析非平衡数据。Henderson(1953)提出的三种估算方差分量的ANOVA法，可以分析某些非平衡数据。其中Henderson方法3有较大的实用价值。Searle(1968)将Henderson的方法改写为以矩阵表达的公式，因而可以用混合线性模型表示。

如果用矩阵形式表示平方和，可以把平方和改写为观察值向量 \mathbf{y} 的二阶式(quadratic form)。对于混合线性模型

$$\mathbf{y} = \mathbf{Xb} + \sum_{u=1}^{m-1} \mathbf{U}_u \mathbf{e}_u + \mathbf{e}_m$$

$$\sim (\mathbf{Xb}, \mathbf{V} = \sum_{u=1}^{m-1} \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T + \sigma_m^2 \mathbf{I})$$

任何二阶式 $\mathbf{y}^T \mathbf{A} \mathbf{y}$ 的期望值都可用以下通式表示

$$\begin{aligned} E(\mathbf{y}^T \mathbf{A} \mathbf{y}) &= \text{tr}(\mathbf{A} \mathbf{V}) + \mathbf{b}^T \mathbf{X}^T \mathbf{A} \mathbf{X} \mathbf{b} \\ &= \sum_{u=1}^{m-1} \sigma_u^2 \text{tr}(\mathbf{U}_u^T \mathbf{A} \mathbf{U}_u) + \sigma_m^2 \text{tr}(\mathbf{A}) + \mathbf{b}^T \mathbf{X}^T \mathbf{A} \mathbf{X} \mathbf{b} \end{aligned}$$

由此可见二阶式 $\mathbf{y}^T \mathbf{A} \mathbf{y}$ 的期望值函数不仅包括了方差分量，还包括了观察值向量的期望值。如果能找到某个矩阵 \mathbf{A} ，使得 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 为零矩阵，便可在二阶式 $\mathbf{y}^T \mathbf{A} \mathbf{y}$ 的期望值函数中去掉 $\mathbf{b}^T \mathbf{X}^T \mathbf{A} \mathbf{X} \mathbf{b}$ 这一项，使得期望值只包括方差分量的线性函数。如果定义以下矩阵 ($j = 1, 2, \dots, m-1$)

$$\mathbf{W}_j = [\mathbf{X}; \mathbf{U}_1; \mathbf{U}_2; \dots; \mathbf{U}_j]$$

$$\mathbf{P}_j = \mathbf{W}_j (\mathbf{W}_j^T \mathbf{W}_j)^+ \mathbf{W}_j^T$$

$$\mathbf{P}_0 = \mathbf{X}_j (\mathbf{X}_j^T \mathbf{X}_j)^+ \mathbf{X}_j^T$$

其中矩阵 $(\mathbf{W}_j^T \mathbf{W}_j)^+$ 是矩阵 $(\mathbf{W}_j^T \mathbf{W}_j)$ 的广义逆， \mathbf{P}_j 和 \mathbf{P}_0 是射影矩阵 (projection matrix)。射影矩阵 \mathbf{P} 具有以下特性

$$\mathbf{P} \mathbf{X} = \mathbf{X}$$

$$\mathbf{X}^T \mathbf{P} = \mathbf{X}^T$$

$$\mathbf{P}^T = \mathbf{P}$$

$$\mathbf{P} \mathbf{P} = \mathbf{P}$$

如果定义矩阵 \mathbf{A} 是射影矩阵 \mathbf{P} 的函数，

$$\mathbf{A}_j = \mathbf{P}_j - \mathbf{P}_{j-1}, \quad j = 1, 2, \dots, m-1$$

$$\mathbf{A}_m = \mathbf{I} - \mathbf{P}_{m-1}$$

则当 $j = 1, 2, \dots, m$ 时，有

$$\mathbf{X}^T \mathbf{A}_j \mathbf{X} = \mathbf{0}$$

$$E(\mathbf{y}^T \mathbf{A}_j \mathbf{y}) = \text{tr}(\mathbf{A}_j \mathbf{V})$$

$$= \sum_{u=1}^m \sigma_u^2 \text{tr}(\mathbf{U}_u^T \mathbf{A}_j \mathbf{U}_u)$$

原混合线性模型总共有 m 项方差分量，当取 m 个不同的矩阵 \mathbf{A}_j ($j = 1, 2, \dots, m$) 时，可以得到 m 个二阶式期望值函数，

$$E(\mathbf{y}^T \mathbf{A}_1 \mathbf{y}) = \sum_{u=1}^m \sigma_u^2 \text{tr}(\mathbf{U}_u^T \mathbf{A}_1 \mathbf{U}_u)$$

$$E(\mathbf{y}^T \mathbf{A}_2 \mathbf{y}) = \sum_{u=2}^m \sigma_u^2 \text{tr}(\mathbf{U}_u^T \mathbf{A}_2 \mathbf{U}_u)$$

$$\vdots = \vdots$$

$$E(\mathbf{y}^T \mathbf{A}_m \mathbf{y}) = \sigma_m^2 \text{tr}(\mathbf{U}_m^T \mathbf{A}_m \mathbf{U}_m)$$

因此，解以下方程组可以获得方差分量的无偏估计值，

$$[\text{tr}(\mathbf{U}_i^T \mathbf{A}_j \mathbf{U}_i)] [\sigma_j^2] = [\mathbf{y}^T \mathbf{A}_j \mathbf{y}], \quad i \leq j = 1, 2, \dots, m$$

当观察值具有多变量正态分布 $\mathbf{y} \sim \text{MVN}(\mathbf{X}\mathbf{b}, \mathbf{V})$ 时，可以进一步估算出方差分量

估计值向量 $[\sigma_j^2]$ 的方差-协方差矩阵，

$$\begin{aligned} \text{Var}[\sigma_j^2] &= \mathbf{H}^{-1} \text{Var}[\mathbf{y}^T \mathbf{A}_j \mathbf{y}] \mathbf{H}^{-1T} \\ &= \mathbf{H}^{-1} [\text{tr}(\mathbf{A}_j \mathbf{V} \mathbf{A}_j \mathbf{V})] \mathbf{H}^{-1T} \end{aligned}$$

其中

$$\mathbf{H} = [\text{tr}(\mathbf{U}_i^T \mathbf{A}_j \mathbf{U}_i)].$$

Henderson方法中的二阶式 $\mathbf{y}^T \mathbf{A}_j \mathbf{y}$ 是平方和的矩阵表达式。如果成对性状X和Y的观察值向量分别为 \mathbf{y}_X 和 \mathbf{y}_Y ，利用Henderson方法也可以估算协方差分量。解以下方程组可以获得协方差分量的无偏估计值，

$$[\text{tr}(\mathbf{U}_i^T \mathbf{A}_j \mathbf{U}_i)] [\sigma_{j(XY)}] = [\mathbf{y}_X^T \mathbf{A}_j \mathbf{y}_Y], \quad i \leq j = 1, 2, \dots, m$$

Henderson的方法适用于一些简单的试验设计的方差分析，并适用于对非平衡数据的分析。但对一些特殊的线性模型，采用此法估算的方差分量可能会产生偏差。七十年代以来发展的一系列混合线性模型分析方法，包括最大似然法(maximum likelihood，简称ML法)、限制性最大似然法(restricted maximum likelihood，简称REML法)、最小范数二阶无偏估算法(minimum norm quadratic unbiased estimation，简称MINQUE法等)，可以克服Henderson方法的局限性。

2. ML或REML法

最大似然法的原理早已在统计分析中得到广泛应用，但该方法在方差分析中的应用则自七十年代开始。Hartley和Rao(1967)首先提出应用最大似然法分析混合线性模型的非平衡数据。对于多变量正态分布的随机向量，Hartley和Rao(1967)的混合模型可以改写为

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \sum_{u=1}^m \mathbf{U}_u \mathbf{e}_u$$

$$\sim \text{MVN}(\mathbf{X}\mathbf{b}, \mathbf{V} = \sum_{u=1}^m \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T)$$

该变量的自然对数似然值为

$$\ln(L) = \text{constant} - \frac{1}{2} \ln(|\mathbf{V}|) - \frac{1}{2} (\mathbf{y} - \mathbf{X}\mathbf{b})^T \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\mathbf{b})$$

由以上自然对数似然值 $\ln(L)$ 对参数 \mathbf{b} 和 σ_u^2 ($u = 1, 2, \dots, m$)求偏导，可以得到

$$\frac{\partial \ln(L)}{\partial \mathbf{b}} = \mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}\mathbf{b} - \mathbf{X}^T \mathbf{V}^{-1} \mathbf{y}$$

和($u = 1, 2, \dots, m$)

$$\frac{\partial \ln(L)}{\partial \sigma_u^2} = -\frac{1}{2} \text{tr}(\mathbf{U}_u \mathbf{U}_u^T \mathbf{V}^{-1}) + \frac{1}{2} (\mathbf{y} - \mathbf{X}\mathbf{b})^T \mathbf{V}^{-1} \mathbf{U}_u \mathbf{U}_u^T \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\mathbf{b})$$

分别设它们为零，可以获得以下似然方程组，

$$\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X}\hat{\mathbf{b}} = \mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{y}$$

和($u = 1, 2, \dots, m$)

$$\text{tr}(\mathbf{U}_u \mathbf{U}_u^T \hat{\mathbf{V}}^{-1}) = (\mathbf{y} - \mathbf{X}\hat{\mathbf{b}})^T \hat{\mathbf{V}}^{-1} \mathbf{U}_u \mathbf{U}_u^T \hat{\mathbf{V}}^{-1} (\mathbf{y} - \mathbf{X}\hat{\mathbf{b}})$$

以上公式经变换后可得，

$$\text{tr}(\mathbf{U}_u \mathbf{U}_u^T \hat{\mathbf{V}}^{-1}) = \text{tr}(\mathbf{U}_u \mathbf{U}_u^T \hat{\mathbf{V}}^{-1} \hat{\mathbf{V}} \hat{\mathbf{V}}^{-1})$$

$$= \sum_{v=1}^m \text{tr}(\mathbf{U}_u^T \hat{\mathbf{V}}^{-1} \mathbf{U}_v \mathbf{U}_v^T \hat{\mathbf{V}}^{-1} \mathbf{U}_u) \sigma_v^2$$

及

$$\begin{aligned} & (\mathbf{y} - \mathbf{X}\hat{\mathbf{b}})^T \hat{\mathbf{V}}^{-1} \mathbf{U}_u \mathbf{U}_u^T \hat{\mathbf{V}}^{-1} (\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}) \\ & = \mathbf{y}^T [\hat{\mathbf{V}}^{-1} - \hat{\mathbf{V}}^{-1} \mathbf{X} (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X}) + \mathbf{X}^T \hat{\mathbf{V}}^{-1}]^T \mathbf{U}_u \mathbf{U}_u^T [\hat{\mathbf{V}}^{-1} - \hat{\mathbf{V}}^{-1} \mathbf{X} (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X}) + \mathbf{X}^T \hat{\mathbf{V}}^{-1}] \mathbf{y} \end{aligned}$$

由此可以导出估算方差分量的ML方程组:

$$[\text{tr}(\mathbf{U}_u^T \hat{\mathbf{V}}^{-1} \mathbf{U}_v \mathbf{U}_v^T \hat{\mathbf{V}}^{-1} \mathbf{U}_u)] [\hat{\sigma}_u^2] = [\mathbf{y}^T \hat{\mathbf{Q}} \mathbf{U}_u \mathbf{U}_u^T \hat{\mathbf{Q}} \mathbf{y}]$$

其中

$$\hat{\mathbf{Q}} = \hat{\mathbf{V}}^{-1} - \hat{\mathbf{V}}^{-1} \mathbf{X} (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X}) + \mathbf{X}^T \hat{\mathbf{V}}^{-1}$$

$$\hat{\mathbf{V}} = \sum_{u=1}^m \hat{\sigma}_u^2 \mathbf{U}_u \mathbf{U}_u^T$$

由于估算方差分量 $[\sigma_u^2]$ 时需用到方差分量的估计值 $[\hat{\sigma}_u^2]$ ，因此ML方程组不可能直接用于估算方差分量。但应用以下迭代的方法，可以获得方差分量的ML估计值，

$$[\text{tr}(\mathbf{U}_u^T \hat{\mathbf{V}}_{[h]}^{-1} \mathbf{U}_v \mathbf{U}_v^T \hat{\mathbf{V}}_{[h]}^{-1} \mathbf{U}_u)] [\hat{\sigma}_{u[h+1]}^2] = [\mathbf{y}^T \hat{\mathbf{Q}}_{[h]} \mathbf{U}_u \mathbf{U}_u^T \hat{\mathbf{Q}}_{[h]} \mathbf{y}]$$

其中

$$\hat{\mathbf{Q}}_{[h]} = \hat{\mathbf{V}}_{[h]}^{-1} - \hat{\mathbf{V}}_{[h]}^{-1} \mathbf{X} (\mathbf{X}^T \hat{\mathbf{V}}_{[h]}^{-1} \mathbf{X}) + \mathbf{X}^T \hat{\mathbf{V}}_{[h]}^{-1}$$

$$\hat{\mathbf{V}}_{[h]} = \sum_{u=1}^m \hat{\sigma}_{u[h]}^2 \mathbf{U}_u \mathbf{U}_u^T, \hat{\sigma}_{u[h]}^2 \text{ 是 } \sigma_u^2 \text{ 的第 } [h] \text{ 迭代估计值。}$$

最大似然法的估算原理要求所有的方差分量估计值为非零的值，因此迭代估算时需设置限定条件 $\hat{\sigma}_u^2 \geq 0$ ($u = 1, 2, \dots, m$)。ML估计值的渐近方差-协方差矩阵为 $2 \times [\text{tr}(\mathbf{U}_u^T \mathbf{V}^{-1} \mathbf{U}_v \mathbf{U}_v^T \mathbf{V}^{-1} \mathbf{U}_u)]^{-1}$ (Searle, 1970)。

ML法的估算值受到固定效应b的影响，可能导致较严重的有偏估计。为了克服这一缺点，Paterson和Thompson (1971)提出了限制性最大似然法 (REML法)，使似然值不包括固定效应b。估算方差分量的REML方程组是:

$$[\text{tr}(\mathbf{U}_u^T \hat{\mathbf{Q}} \mathbf{U}_v \mathbf{U}_v^T \hat{\mathbf{Q}} \mathbf{U}_u)] [\hat{\sigma}_u^2] = [\mathbf{y}^T \hat{\mathbf{Q}} \mathbf{U}_u \mathbf{U}_u^T \hat{\mathbf{Q}} \mathbf{y}]$$

其中

$$\hat{\mathbf{Q}} = \hat{\mathbf{V}}^{-1} - \hat{\mathbf{V}}^{-1}\mathbf{X}(\mathbf{X}^T\hat{\mathbf{V}}^{-1}\mathbf{X})^+\mathbf{X}^T\hat{\mathbf{V}}^{-1}$$

$$\hat{\mathbf{V}} = \sum_{u=1}^m \hat{\sigma}_u^2 \mathbf{U}_u \mathbf{U}_u^T$$

方差分量的REML估计值可由下式迭代获得，

$$[\text{tr}(\mathbf{U}_u^T \hat{\mathbf{Q}}_{[h]} \mathbf{U}_v \mathbf{U}_v^T \hat{\mathbf{Q}}_{[h]} \mathbf{U}_u)] [\hat{\sigma}_{u[h+1]}^2] = [\mathbf{y}^T \hat{\mathbf{Q}}_{[h]} \mathbf{U}_u \mathbf{U}_u^T \hat{\mathbf{Q}}_{[h]} \mathbf{y}]$$

其中

$$\hat{\mathbf{Q}}_{[h]} = \hat{\mathbf{V}}_{[h]}^{-1} - \hat{\mathbf{V}}_{[h]}^{-1}\mathbf{X}(\mathbf{X}^T\hat{\mathbf{V}}_{[h]}^{-1}\mathbf{X})^+\mathbf{X}^T\hat{\mathbf{V}}_{[h]}^{-1}$$

$$\hat{\mathbf{V}}_{[h]} = \sum_{u=1}^m \hat{\sigma}_{u[h]}^2 \mathbf{U}_u \mathbf{U}_u^T, \hat{\sigma}_{u[h]}^2 \text{ 是 } \sigma_u^2 \text{ 的第 } [h] \text{ 迭代估计值。}$$

运用REML法代估算时也设置限定条件 $\hat{\sigma}_u^2 \geq 0$ ($u = 1, 2, \dots, m$)。REML估计值的渐近方差-协方差矩阵为 $2 \times [\text{tr}(\mathbf{U}_u^T \mathbf{Q} \mathbf{U}_v \mathbf{U}_v^T \mathbf{Q} \mathbf{U}_u)]^{-1}$ (Searle, 1970)。

3. MINQUE法

在分析混合线性模型的各种方法中，Rao(1971)提出的最小范数二阶无偏估算(MINQUE)法比最大似然(ML)法和限制性最大似然(REML)法等更简便和优越。MINQUE法不需要进行迭代运算，对线性模型也没有正态分布的限定。MINQUE方法的推导，是基于使欧氏范数(euclidean norm)为最小。

如果混合线性模型用矩阵简式表达

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\mathbf{b} + \mathbf{U}\mathbf{e} \\ &= \mathbf{X}\mathbf{b} + \sum_{u=1}^m \mathbf{U}_u \mathbf{e}_u \\ &= (\mathbf{X}\mathbf{b}, \quad \mathbf{V} = \sum_{u=1}^m \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T) \end{aligned}$$

其中

$$\mathbf{U} = [\mathbf{U}_1 : \mathbf{U}_2 : \dots : \mathbf{U}_{m-1} : \mathbf{U}_m]$$

$$\mathbf{U}_m = \mathbf{I}$$

$$\mathbf{e}^T = [\mathbf{e}_1^T : \mathbf{e}_2^T : \dots : \mathbf{e}_{m-1}^T : \mathbf{e}_m^T]$$

那么方差分量线性函数的无偏估计能用一个二阶函数表示

$$\sum_{u=1}^m c_u \hat{\sigma}_u^2 = \mathbf{y}^T \mathbf{A} \mathbf{y} = \mathbf{e}^T \mathbf{U}^T \mathbf{A} \mathbf{U} \mathbf{e}$$

其中

$$\mathbf{A} \mathbf{X} = \mathbf{0}$$

$$c_u = \text{tr}(\mathbf{U}_u^T \mathbf{A} \mathbf{U}_u)$$

如果假定 \mathbf{e} 是已知的，并具有 $\mathbf{e}_u \sim (0, \sigma_u^2 \mathbf{I})$ 。方差分量线性函数的估计量自然应该是 $\mathbf{e}^T \mathbf{D} \mathbf{e}$ ，其中 \mathbf{D} 是已知的对角矩阵。该假定的估计量与实际估计量之间的差异是 $\mathbf{e}^T (\mathbf{U}^T \mathbf{A} \mathbf{U} - \mathbf{D}) \mathbf{e}$ 。MINQUE法所获得的方差分量估计是使欧氏范数 $\|\mathbf{U}^T \mathbf{A} \mathbf{U} - \mathbf{D}\|^2$ 为最小值，即使 $\text{tr}[(\mathbf{U}^T \mathbf{A} \mathbf{U} - \mathbf{D})(\mathbf{U}^T \mathbf{A} \mathbf{U} - \mathbf{D})]$ 为最小值。Rao导出方差分量的MINQUE估计量可由解以下方程组而得， $u, v = 1, 2, \dots, m$

$$[\text{tr}(\mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{U}_v \mathbf{U}_v^T \mathbf{Q}_\alpha \mathbf{U}_u)] [\hat{\sigma}_u^2] = [\mathbf{y}^T \mathbf{Q}_\alpha \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y}]$$

其中

$$\mathbf{Q}_\alpha = \mathbf{V}_\alpha^{-1} - \mathbf{V}_\alpha^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}_\alpha^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}_\alpha^{-1}$$

$$\mathbf{V}_\alpha = \sum_{u=1}^m \alpha_u \mathbf{U}_u \mathbf{U}_u^T$$

$[\alpha_u]$ 是人为选择的先验值。

尽管方差分量的估计量取决于人为选择的先验值 $[\alpha_u]$ ，只要这些先验值不依赖于试验数据，MINQUE估计量仍是无偏的。选择任意的先验值 $[\alpha_u]$ ，可以求得方差分量的估计值 $[\hat{\sigma}_u^2]$ 。如果用估计值替代先验值重新估算，便可获得新的估计值。重复这一过程，直到新的估计值非常接近旧的估计值为止。这种迭代的估算方

法就是REML法，其估算结果是REML估计值。在多数情况下，REML估计值和MINQUE估计值是比较接近的。

先验值 $[\alpha_u]$ 的选择可以凭经验，或根据以往的分析结果。更简单的方法是取所有的先验值为1($\alpha_u = 1$)，这种方法称为MINQUE(1)法，所获得的方差分量的估计量是MINQUE(1)估计量。

MINQUE(1)估计量的无偏性证明如下：

设

$$\begin{aligned} \mathbf{V}_{(1)} &= \sum_{u=1}^m \mathbf{U}_u \mathbf{U}_u^T \\ \mathbf{Q}_{(1)} &= \mathbf{V}_{(1)}^{-1} - \mathbf{V}_{(1)}^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}_{(1)}^{-1} \mathbf{X})^+ \mathbf{X}^T \mathbf{V}_{(1)}^{-1} \\ \mathbf{W} &= (\mathbf{V}_{(1)}^{-1})^{1/2} \mathbf{X} \end{aligned}$$

$$\mathbf{P}_W = \mathbf{W}(\mathbf{W}^T \mathbf{W})^+ \mathbf{W}^T \text{ 是射影矩阵}$$

则有

$$\begin{aligned} \mathbf{Q}_{(1)} &= (\mathbf{V}_{(1)}^{-1})^{1/2} (\mathbf{I} - \mathbf{P}_W) (\mathbf{V}_{(1)}^{-1})^{1/2} \\ \mathbf{Q}_{(1)} \mathbf{X} &= (\mathbf{V}_{(1)}^{-1})^{1/2} (\mathbf{I} - \mathbf{P}_W) (\mathbf{V}_{(1)}^{-1})^{1/2} \mathbf{X} \\ &= (\mathbf{V}_{(1)}^{-1})^{1/2} (\mathbf{I} - \mathbf{P}_W) \mathbf{W} \\ &= (\mathbf{V}_{(1)}^{-1})^{1/2} (\mathbf{W} - \mathbf{W}) \\ &= \mathbf{0} \end{aligned}$$

因此

$$\begin{aligned} E[\mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_{(1)} \mathbf{y}] &= \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_{(1)} \mathbf{V}) + \mathbf{b}^T \mathbf{X}^T \mathbf{Q}_{(1)} \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_{(1)} \mathbf{X} \mathbf{b} \\ &= \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_{(1)} \sum_{v=1}^m \sigma_v^2 \mathbf{U}_v \mathbf{U}_v^T) \\ &= [\text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_{(1)} \mathbf{U}_v \mathbf{U}_v^T)] [\sigma_v^2] \\ &= [\text{tr}(\mathbf{U}_u^T \mathbf{Q}_{(1)} \mathbf{U}_v \mathbf{U}_v^T \mathbf{Q}_{(1)} \mathbf{U}_u)] [\sigma_u^2] \end{aligned}$$

证

毕。

朱军(1992)提出了采用MINQUE方法估算成对性状之间遗传协方差分量的算

法。如果成对性状的观察值向量分别为 \mathbf{y}_1 和 \mathbf{y}_2 ，协方差分量的线性函数可以用一个随机向量的二次型来表示，

$$\sum_{u=1}^m c_u \hat{\sigma}_{1u/2u} = \mathbf{y}_1^T \mathbf{A} \mathbf{y}_2$$

当 $\mathbf{A} = \mathbf{Q}_\alpha \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_\alpha$ 时， $\mathbf{y}_1^T \mathbf{A} \mathbf{y}_2$ 的期望值为，

$$\text{tr}(\mathbf{Q}_\alpha \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{C}) = \sum_{v=1}^m \sigma_{v/v} \text{tr}(\mathbf{Q}_\alpha \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{U}_v \mathbf{U}_v^T)$$

其中 $\text{Cov}(\mathbf{y}_1, \mathbf{y}_2^T) = \mathbf{C} = \sum_u^m \sigma_{u/u} \mathbf{U}_u \mathbf{U}_u^T$ 。

因此随机变量 \mathbf{y}_1 和 \mathbf{y}_2 的协方差分量可以由以下MINQUE方程组解得，

$$[\text{tr}(\mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{U}_v \mathbf{U}_v^T \mathbf{Q}_\alpha \mathbf{U}_u)] [\hat{\sigma}_{u/u}] = [\mathbf{y}_1^T \mathbf{Q}_\alpha \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y}_2]$$

MINQUE估计方差分量的公式可以改写为以下形式 ($u, v = 1, 2, \dots, m$)，

$$[\hat{\sigma}_u^2] = \mathbf{H}_\alpha^{-1} [\mathbf{y}^T \mathbf{Q}_\alpha \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y}]$$

其中

$$\mathbf{H}_\alpha = [\text{tr}(\mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{U}_v \mathbf{U}_v^T \mathbf{Q}_\alpha \mathbf{U}_u)]。$$

当观察值具有多变量正态分布 $\mathbf{y} \sim \text{MVN}(\mathbf{Xb}, \mathbf{V})$ 时，MINQUE方差分量估计值的方差-协方差矩阵可以由下式推导而得

$$\begin{aligned} \text{Var}[\hat{\sigma}_u^2] &= \mathbf{H}_\alpha^{-1} \text{Var}[\mathbf{y}^T \mathbf{Q}_\alpha \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y}] \mathbf{H}_\alpha^{-1} \\ &= \mathbf{H}_\alpha^{-1} [2\text{tr}(\mathbf{Q}_\alpha \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{V} \mathbf{Q}_\alpha \mathbf{U}_v \mathbf{U}_v^T \mathbf{Q}_\alpha \mathbf{V})] \mathbf{H}_\alpha^{-1} \end{aligned}$$

Rao和Kleffe(1988)进一步证明当试验的样本足够大时，方差分量的MINQUE估计值的方差近似于

$$\text{Var}[\hat{\sigma}_u^2] = 2[\text{tr}(\mathbf{Q} \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q} \mathbf{U}_v \mathbf{U}_v^T \mathbf{Q} \mathbf{V})]^{-1}，$$

并且MINQUE估计值具有渐进多元正态分布

$$[\hat{\sigma}_u^2] \sim \text{MVN}([\sigma_u^2], 2\mathbf{H}^{-1})$$

其中

$$\mathbf{H} = [\text{tr}(\mathbf{U}_u^T \mathbf{Q} \mathbf{U}_v \mathbf{U}_v^T \mathbf{Q} \mathbf{U}_u)] ,$$

$$\mathbf{Q} = \mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1}$$

$$\mathbf{V} = \sum_{u=1}^m \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T$$

由此可以对方差分量估计值作线性统计检验。

对于无效假设 $H_0: \sum_{u=1}^m c_u \sigma_u^2 = \rho$, 可以用 χ^2 分布来检验。当无效假设成立

时, 统计量

$$\frac{(\sum_{u=1}^m c_u \hat{\sigma}_u^2 - \rho)^2}{2(\sum_{u=1}^m c_u^2 \hat{\mathbf{H}}_{uu}^{-1} + 2 \sum_{u=1}^{m-1} \sum_{v>u}^m c_u c_v \hat{\mathbf{H}}_{uv}^{-1})}$$

近似地具有自由度为1的 χ^2 分布, 其中

$$\hat{\mathbf{H}} = [\text{tr}(\mathbf{U}_u^T \hat{\mathbf{Q}} \mathbf{U}_v \mathbf{U}_v^T \hat{\mathbf{Q}} \mathbf{U}_u)]$$

$$\hat{\mathbf{Q}} = \hat{\mathbf{V}}^{-1} - \hat{\mathbf{V}}^{-1} \mathbf{X} (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{V}}^{-1}$$

$$\hat{\mathbf{V}} = \sum_{u=1}^m \hat{\sigma}_u^2 \mathbf{U}_u \mathbf{U}_u^T \text{ 并有逆矩阵 } \hat{\mathbf{V}}^{-1}$$

例如测验 $H_0: \sigma_u^2 = 0$ 时, 如果

$$\frac{(\hat{\sigma}_u^2)^2}{2\hat{\mathbf{H}}_{uu}^{-1}} > \chi_{(0.05; 1)}^2$$

则否定无效假设 H_0 , 推断 σ_u^2 显著大于零。

再例如测验 $H_0: \sigma_u^2 = \sigma_v^2$ 时, 如果

$$\frac{(\hat{\sigma}_u^2 - \hat{\sigma}_v^2)^2}{2(\hat{\mathbf{H}}_{uu}^{-1} + \hat{\mathbf{H}}_{vv}^{-1} - 2\hat{\mathbf{H}}_{uv}^{-1})} > \chi_{(0.05; 1)}^2$$

则否定无效假设 H_0 , 推断 σ_u^2 与 σ_v^2 存在显著差异。

根据方差分量的MINQUE估计值的近似正态分布，可以确定所有方差分量参数的近似置信区域(asymptotic confidence region)。100(1-)%置信区域由以下等式确定， $u = 1, 2, \dots, m$,

$$[\hat{\sigma}_u^2 - \sigma_u^2]^T \hat{\mathbf{H}}[\hat{\sigma}_u^2 - \sigma_u^2] \leq 2X_{(\alpha; m)}^2$$

对于某一个方差分量参数，100(1-)%置信区间可以由下式算得

$$\hat{\sigma}_u^2 - z(\alpha)\sqrt{2\hat{\mathbf{H}}_{uu}^{-1}} \leq \sigma_u^2 \leq \hat{\sigma}_u^2 + z(\alpha)\sqrt{2\hat{\mathbf{H}}_{uu}^{-1}}$$

其中 $z(\alpha)$ 是标准正态分布表上具有累计概率为(1-)的值。

方差分量估计值及其标准误也可以用Jackknife的方法(Miller, 1974; 朱军, 1992)对基因型或试验区组进行抽样而算得。如果所分析的遗传资料是基因型的平均数，无区组的观察值，常采用对基因型抽样的技术，每个世代平均数的 r 次观察值是Jackknife的重复抽样单位。当遗传试验有随机区组时，一般对试验区组进行抽样，每个随机区组中的所有观察值是Jackknife的重复抽样单位。

现以基因型重复抽样为例，介绍估算遗传参数Jackknife估计值的方法。对于遗传参数 θ (方差分量或协方差分量)，分析所有 g 个基因型观察值时得到的估计值为 $\hat{\theta}$ 。当从资料中剔除第 i 种基因型世代平均数的 r 次观察值以后，重新估算而得到估计值 $\hat{\theta}_i$ 。采用此法从完整的数据中每次剔除 r 次观察值，可以获得 g 个不同的估计值 $\hat{\theta}_i$ 。Jackknife估计值(或预测值) $\hat{\theta}_j$ 和抽样方差 $\text{Var}(\hat{\theta}_j)$ 分别为

$$\hat{\theta}_j = g\hat{\theta} - (g-1)\hat{\theta}$$

$$\text{Var}(\hat{\theta}_j) = \frac{g-1}{g} \sum_{i=1}^g (\hat{\theta}_i - \hat{\theta})^2$$

其中 $\hat{\theta} = \frac{1}{g} \sum_{i=1}^g \hat{\theta}_i$ 。 $(\hat{\theta}_j - \theta) / \sqrt{\text{Var}(\hat{\theta}_j)}$ 近似地具有自由度为 $(g-1)$ 的 t 分布。从而

可以用 t 测验对遗传参数作统计检验。

Zhu (1989)用蒙特卡罗模拟比较了估算方差分量估计值的方差及参数检验的效益。采用的两种统计分析方法分别是：MINQUE估计值方差近似估算法和

Jackknife 抽样估算法。比较的主要指标是偏差(bias)和功效值(power value)。功效值是在5%显著水准下检测出方差分量大于零的概率,该值越大表明统计检验的效益越高。由表4.1可见,MINQUE估计值方差近似估算法和Jackknife 抽样估算法估计方差分量的偏差都较小,但用近似估算法算得的方差检测出方差显著性的概率远低于或者。因此Jackknife 抽样估算法不但可以无偏地估算方差分量,还能有效地检测方差的显著性。在遗传分析中一些遗传参数(如遗传率、遗传相关系数等)估计值的抽样方差也可以用Jackknife 的方法算得,并可用 *t*-测验作统计检验。

表4.3 两种方法估算8×8双列杂交(方法3)方差分量的偏差和功效值比较

方差分量	参数值	近似估算法		抽样估算法	
		偏差	功效值	偏差	功效值
V _A	100	-2.07	0.06	7.79	0.93
V _D	30	1.87	0.73	-0.05	0.89
V _M	20	2.25	0.01	1.01	0.53
V _P	20	1.17	0.00	0.61	0.47
V _e	30	0.93	1.00	0.35	1.00

在分析一些复杂的遗传模型(如种子模型)时,随机变量之间常存在特定的相关性。对于这类混合线性模型,可以采用MINQUE(0/1)的方法(朱军,1992; Zhu和Weir, 1994a)无偏估算方差和协方差分量。现以种子遗传模型为例,阐述MINQUE(0/1)的分析方法。

如果以矩阵形式表示的种子遗传模型是,

$$\begin{aligned}
 \mathbf{y} &= \mathbf{Xb} + \mathbf{U}_A \mathbf{e}_A + \mathbf{U}_D \mathbf{e}_D + \mathbf{U}_C \mathbf{e}_C + \mathbf{U}_{Am} \mathbf{e}_{Am} + \mathbf{U}_{Dm} \mathbf{e}_{Dm} + \mathbf{U}_B \mathbf{e}_B + \mathbf{e}_e \\
 &= \mathbf{Xb} + \sum_u^7 \mathbf{U}_u \mathbf{e}_u
 \end{aligned}$$

其中 \mathbf{X} 是固定效应向量 \mathbf{b} 的系数矩阵, \mathbf{e}_u 是第 u 项随机效应向量, $\mathbf{e}_u \sim (0, \sigma_u^2 \mathbf{I})$; \mathbf{U}_u 是随机效应向量 \mathbf{e}_u 的系数矩阵($u=1, 2, \dots, 6$), $\mathbf{U}_7 = \mathbf{I}$ 是单位矩阵。

种子数量性状表现型向量 \mathbf{y} 的方差-协方差矩阵为

$$\text{Var}(\mathbf{y}) = \sigma_A^2 \mathbf{U}_A \mathbf{U}_A^T + \sigma_D^2 \mathbf{U}_D \mathbf{U}_D^T + \sigma_C^2 \mathbf{U}_C \mathbf{U}_C^T + \sigma_{Am}^2 \mathbf{U}_{Am} \mathbf{U}_{Am}^T$$

$$\begin{aligned}
 & +\sigma_{Dm}^2 \mathbf{U}_{Dm} \mathbf{U}_{Dm}^T + \sigma_B^2 \mathbf{U}_B \mathbf{U}_B^T + \sigma_{A.Am} (\mathbf{U}_A \mathbf{U}_{Am}^T + \mathbf{U}_{Am} \mathbf{U}_A^T) \\
 & +\sigma_{D.Dm} (\mathbf{U}_D \mathbf{U}_{Dm}^T + \mathbf{U}_{Dm} \mathbf{U}_D^T) + \sigma_e^2 \mathbf{I}
 \end{aligned}$$

其中 σ_A^2 是直接加性效应方差分量, σ_D^2 是直接显性效应方差分量, σ_C^2 是细胞质效应方差分量, σ_{Am}^2 是母体加性效应方差分量, σ_{Dm}^2 是母体显性效应方差分量, σ_B^2 是区组效应方差分量, $\sigma_{A.Am}$ 是直接加性效应与母体加性效应的协方差分量, $\sigma_{D.Dm}$ 是直接显性效应与母体显性效应的协方差分量, σ_e^2 是剩余效应方差分量。以上方差-协方差矩阵还可以改写为便于运算的形式

$$\begin{aligned}
 \text{Var}(\mathbf{y}) & = \sigma_A^2 \mathbf{V}_1 + \sigma_D^2 \mathbf{V}_2 + \sigma_C^2 \mathbf{V}_3 + \sigma_{Am}^2 \mathbf{V}_4 + \sigma_{Dm}^2 \mathbf{V}_5 + \sigma_B^2 \mathbf{V}_6 \\
 & \quad + \sigma_{A.Am} \mathbf{V}_7 + \sigma_{D.Dm} \mathbf{V}_8 + \sigma_e^2 \mathbf{V}_9 \\
 & = \sum_{u=1}^9 \theta_u \mathbf{V}_u
 \end{aligned}$$

其中 $\mathbf{V}_u = \mathbf{U}_u \mathbf{U}_u^T$ ($u = 1, 2, \dots, 6$)、 $\mathbf{V}_7 = (\mathbf{U}_1 \mathbf{U}_4^T + \mathbf{U}_4 \mathbf{U}_1^T)$, $\mathbf{V}_8 = (\mathbf{U}_2 \mathbf{U}_5^T + \mathbf{U}_5 \mathbf{U}_2^T)$ 、 $\mathbf{V}_9 = \mathbf{I}$; $\theta_1 = \sigma_A^2$ 、 $\theta_2 = \sigma_D^2$ 、 $\theta_3 = \sigma_C^2$ 、 $\theta_4 = \sigma_{Am}^2$ 、 $\theta_5 = \sigma_{Dm}^2$ 、 $\theta_6 = \sigma_B^2$ 、 $\theta_7 = \sigma_{A.Am}$ 、 $\theta_8 = \sigma_{D.Dm}$ 、 $\theta_9 = \sigma_e^2$ 。

采用Rao提出的MINQUE法(1971)可以无偏估算各项方差分量和协方差分量。MINQUE估计量可由解以下方程组而得, $u, v = 1, 2, \dots, 9$:

$$[\text{tr}(\mathbf{Q}_\alpha \mathbf{V}_u \mathbf{Q}_\alpha \mathbf{V}_v)] [\hat{\theta}_u] = [\mathbf{y}^T \mathbf{Q}_\alpha \mathbf{V}_u \mathbf{Q}_\alpha \mathbf{y}]$$

其中

$$\begin{aligned}
 \mathbf{Q}_\alpha & = \mathbf{V}_\alpha^{-1} - \mathbf{V}_\alpha^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}_\alpha^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}_\alpha^{-1} \\
 \mathbf{V}_\alpha & = \sum_{u=1}^9 \alpha_u \mathbf{V}_u
 \end{aligned}$$

α_u 为方差先验值($u = 1, 2, \dots, 6$ 和9)或协方差先验值($u = 7, 8$)。

MINQUE(0/1)法, 即设先验协方差为0, 先验方差为1的一种MINQUE法, 可以无偏地估算以上各项方差和协方差分量。MINQUE(0/1)估计量可由解以下方程组

而得, $u, v = 1, 2, \dots, 9$:

$$[\text{tr}(\mathbf{Q}_{(0/1)} \mathbf{V}_u \mathbf{Q}_{(0/1)} \mathbf{V}_v)] [\hat{\theta}_u] = [\mathbf{y}^T \mathbf{Q}_{(0/1)} \mathbf{V}_u \mathbf{Q}_{(0/1)} \mathbf{y}]$$

其中

$$\mathbf{Q}_{(0/1)} = \mathbf{V}_{(0/1)}^{-1} - \mathbf{V}_{(0/1)}^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}_{(0/1)}^{-1} \mathbf{X})^+ \mathbf{X}^T \mathbf{V}_{(0/1)}^{-1}$$

$$\mathbf{V}_{(0/1)} = \sum_{u=1}^7 \mathbf{U}_u \mathbf{U}_u^T$$

对于成对性状, 也可以采用MINQUE(0/1)法无偏地估算各项遗传协方差分量(朱军, 1992; Zhu和Weir, 1994a)。对于成对性状的观察值 \mathbf{y}_1 和 \mathbf{y}_2 , 其协方差矩阵为:

$$\begin{aligned} \text{Cov}(\mathbf{y}_1, \mathbf{y}_2^T) &= \sigma_{A/A} \mathbf{V}_1 + \sigma_{D/D} \mathbf{V}_2 + \sigma_{C/C} \mathbf{V}_3 + \sigma_{Am/Am} \mathbf{V}_4 + \sigma_{Dm/Dm} \mathbf{V}_5 \\ &\quad + \sigma_{B/B} \mathbf{V}_6 + \sigma_{A/Am} \mathbf{V}_7 + \sigma_{D/Dm} \mathbf{V}_8 + \sigma_{e/e} \mathbf{V}_9 \\ &= \sum_{u=1}^9 \theta_{u/u} \mathbf{V}_u \end{aligned}$$

其中 $\theta_{1/1} = \sigma_{A/A}$ 是加性协方差、 $\theta_{2/2} = \sigma_{D/D}$ 是显性协方差、 $\theta_{3/3} = \sigma_{C/C}$ 是细胞质协方差、 $\theta_{4/4} = \sigma_{Am/Am}$ 是母体加性协方差、 $\theta_{5/5} = \sigma_{Dm/Dm}$ 是母体显性协方差、 $\theta_{6/6} = \sigma_{B/B}$ 是区组协方差、 $\theta_{7/7} = \sigma_{A/Am}$ 是直接加性效应与母体加性效应协方差、 $\theta_{8/8} = \sigma_{D/Dm}$ 是直接显性效应与母体显性效应协方差、 $\theta_{9/9} = \sigma_{e/e}$ 是剩余效应协方差。

各项协方差分量可以由以下MINQUE(0/1)方程组解得,

$$[\text{tr}(\mathbf{Q}_{(0/1)} \mathbf{V}_u \mathbf{Q}_{(0/1)} \mathbf{V}_v)] [\hat{\theta}_{u/u}] = [\mathbf{y}_1^T \mathbf{Q}_{(0/1)} \mathbf{V}_u \mathbf{Q}_{(0/1)} \mathbf{y}_2]$$

其中

$$\mathbf{Q}_{(0/1)} = \mathbf{V}_{(0/1)}^{-1} - \mathbf{V}_{(0/1)}^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}_{(0/1)}^{-1} \mathbf{X})^+ \mathbf{X}^T \mathbf{V}_{(0/1)}^{-1}$$

$$\mathbf{V}_{(0/1)} = \sum_{u=1}^7 \mathbf{U}_u \mathbf{U}_u^T$$

蒙特卡罗模拟证明, 采用MINQUE(0/1)法可以无偏地估算各项方差和协方差分量(Zhu和Weir, 1994a)。

以上所介绍的估算成对性状的MINQUE方法有一个限定条件，即成对性状的遗传模型和试验设计矩阵都相等。采用以上方法只能估算农艺性状间或种子性状间的遗传协方差分量，而不能估算农艺性状与种子性状间的遗传协方差分量，因为农艺性状和种子性状具有不同的遗传模型和试验设计矩阵。

在混合线性模型研究领域，不等矩阵成对性状协方差分量的估算方法一直是统计学家试图攻克的一个难题。朱军(1993b)在这一研究领域已取得突破，首次提出了估算不同性状遗传协方差分量的新方法。运用该方法。可以无偏估算具有不同遗传模型(如种子和植株)性状之间的遗传协方差分量。

如果性状 i 的观察值向量 \mathbf{y}_i 可以用以下混合线性模型的矩阵简式表达

$$\begin{aligned}\mathbf{y}_i &= \mathbf{X}\mathbf{b}_{(i)} + \mathbf{U}_{(i)1}\mathbf{e}_{(i)1} + \mathbf{U}_{(i)2}\mathbf{e}_{(i)2} + \cdots + \mathbf{U}_{(i)k}\mathbf{e}_{(i)k} \\ &= \mathbf{X}\mathbf{b}_{(i)} + \sum_{j=1}^k \mathbf{U}_{(i)j}\mathbf{e}_{(i)j}\end{aligned}$$

其中 \mathbf{X} 是性状 i 的固定效应向量 $\mathbf{b}_{(i)}$ 的系数矩阵， $\mathbf{e}_{(i)j}$ 是性状 i 的第 j 项随机效应向量， $\mathbf{e}_{(i)j} \sim (0, \sigma_j^2\mathbf{I})$ ； $\mathbf{U}_{(i)j}$ 是性状 i 的随机效应向量 $\mathbf{e}_{(i)j}$ 的系数矩阵($j = 1, 2, \dots, k$)。

在这个混合线性模型中随机变量之间可能存在相关性，因此向量 \mathbf{y}_i 的方差矩阵包括 k 个方差分量以及一些协方差分量。如果另一个性状 i' 的观察值向量 $\mathbf{y}_{i'}$ 具有不同的遗传模型和试验设计矩阵，向量 \mathbf{y}_i 与 $\mathbf{y}_{i'}$ 的协方差是不对称的，

$$\begin{aligned}\text{Cov}(\mathbf{y}_i, \mathbf{y}_{i'}^T) &= \sigma_{1/1}\mathbf{F}_{(ii')1} + \sigma_{2/2}\mathbf{F}_{(ii')2} + \cdots + \sigma_{r/r}\mathbf{F}_{(ii')r} \\ &= \sum_{u=1}^r \sigma_{u/u}\mathbf{F}_{(ii')u} \\ \text{Cov}(\mathbf{y}_{i'}, \mathbf{y}_i^T) &= \sigma_{1/1}\mathbf{F}_{(i'i)1} + \sigma_{2/2}\mathbf{F}_{(i'i)2} + \cdots + \sigma_{r/r}\mathbf{F}_{(i'i)r} \\ &= \sum_{u=1}^r \sigma_{u/u}\mathbf{F}_{(i'i)u}\end{aligned}$$

这二个协方差矩阵之和是对称矩阵

$$\begin{aligned} \mathbf{V} &= \text{Cov}(\mathbf{y}_i, \mathbf{y}_i^T) + \text{Cov}(\mathbf{y}_{i'}, \mathbf{y}_{i'}^T) \\ &= \sum_{u=1}^r \sigma_{u/u} (\mathbf{F}_{(ii')u} + \mathbf{F}_{(i'i)u}) \\ &= \sum_{u=1}^r \sigma_{u/u} \mathbf{F}_u \end{aligned}$$

定义对称矩阵

$$\begin{aligned} \mathbf{V}_{(\alpha)} &= \sum_{u=1}^r \alpha_{u/u} (\mathbf{F}_{(ii')u} + \mathbf{F}_{(i'i)u}) \\ &= \sum_{u=1}^r \alpha_{u/u} \mathbf{F}_u \end{aligned}$$

$$\mathbf{Q}_{(\alpha)} = \mathbf{V}_{(\alpha)}^{-1} - \mathbf{V}_{(\alpha)}^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}_{(\alpha)}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}_{(\alpha)}^{-1}$$

那么随机变量二次型的期望值可由下式推得，

$$\begin{aligned} &E(\mathbf{y}_i^T \mathbf{Q}_{(\alpha)} \mathbf{F}_u \mathbf{Q}_{(\alpha)} \mathbf{y}_{i'}, \mathbf{y}_{i'}^T \mathbf{Q}_{(\alpha)} \mathbf{F}_u \mathbf{Q}_{(\alpha)} \mathbf{y}_i) \\ &= E(2\mathbf{y}_i^T \mathbf{Q}_{(\alpha)} \mathbf{F}_u \mathbf{Q}_{(\alpha)} \mathbf{y}_{i'}) \\ &= \text{tr}(\mathbf{Q}_{(\alpha)} \mathbf{F}_u \mathbf{Q}_{(\alpha)} \mathbf{V}) \\ &= \text{tr}[\mathbf{Q}_{(\alpha)} \mathbf{F}_u \mathbf{Q}_{(\alpha)} (\sum_{v=1}^r \alpha_{v/v} \mathbf{F}_v)] \\ &= [\text{tr}(\mathbf{Q}_{(\alpha)} \mathbf{F}_u \mathbf{Q}_{(\alpha)} \mathbf{F}_v)] [\alpha_{v/v}], v = 1, 2, \dots, r. \end{aligned}$$

设以上二次型函数($u = 1, 2, \dots, r$)分别等于其期望值，所有协方差分量的估计值便可以由解以下正规方程组而获得($u, v = 1, 2, \dots, r$),

$$[\text{tr}(\mathbf{Q}_{(\alpha)} \mathbf{F}_u \mathbf{Q}_{(\alpha)} \mathbf{F}_v)] [\hat{\alpha}_{v/v}] = [2\mathbf{y}_i^T \mathbf{Q}_{(\alpha)} \mathbf{F}_u \mathbf{Q}_{(\alpha)} \mathbf{y}_{i'}]$$

现以种子性状与生产种子的母体植株农艺性状的相关分析为例，说明非等矩阵的变量之间协方差的估算方法。如果以矩阵形式表达的种子性状遗传模型是，

$$\begin{aligned} \mathbf{y}_S &= \mathbf{X}\mathbf{b}_{(S)} + \mathbf{U}_A \mathbf{e}_{(S)A} + \mathbf{U}_D \mathbf{e}_{(S)D} + \mathbf{U}_C \mathbf{e}_{(S)C} + \mathbf{U}_{Am} \mathbf{e}_{(S)Am} + \mathbf{U}_{Dm} \mathbf{e}_{(S)Dm} \\ &\quad + \mathbf{U}_B \mathbf{e}_{(S)B} + \mathbf{e}_{(S)e} \\ &= \mathbf{X}\mathbf{b}_{(S)} + \sum_u \mathbf{U}_u \mathbf{e}_{(S)u} \end{aligned}$$

其中 \mathbf{X} 是种子性状固定效应向量 $\mathbf{b}_{(S)}$ 的系数矩阵, $\mathbf{e}_{(S)u}$ 是种子性状第 u 项随机效应向量, $\mathbf{e}_{(S)u} \sim (0, \sigma_{(S)u}^2 \mathbf{I})$; \mathbf{U}_u 是种子性状随机效应向量 $\mathbf{e}_{(S)u}$ 的系数矩阵 ($u=1, 2, \dots, 6$), $\mathbf{U}_7 = \mathbf{I}$ 是单位矩阵。

产生种子的母体植株的农艺性状遗传模型便是

$$\begin{aligned} \mathbf{y}_P &= \mathbf{X}\mathbf{b}_{(P)} + \mathbf{U}_C\mathbf{e}_{(P)C} + \mathbf{U}_{Am}\mathbf{e}_{(P)Am} + \mathbf{U}_{Dm}\mathbf{e}_{(P)Dm} + \mathbf{U}_B\mathbf{e}_{(P)B} + \mathbf{e}_{(P)e} \\ &= \mathbf{X}\mathbf{b}_{(P)} + \sum_u^5 \mathbf{U}_u\mathbf{e}_{(P)u} \end{aligned}$$

种子数量性状与母体植株的农艺性状的协方差为

$$\begin{aligned} \text{Cov}(\mathbf{y}_S, \mathbf{y}_P) &= \sigma_{A/Am} \mathbf{U}_A \mathbf{U}_{Am}^T + \sigma_{D/Dm} \mathbf{U}_D \mathbf{U}_{Dm}^T + \sigma_{C/C} \mathbf{U}_C \mathbf{U}_C^T \\ &\quad + \sigma_{Am/Am} \mathbf{U}_{Am} \mathbf{U}_{Am}^T + \sigma_{Dm/Dm} \mathbf{U}_{Dm} \mathbf{U}_{Dm}^T + \sigma_{B/B} \mathbf{U}_B \mathbf{U}_B^T + \sigma_{e/e} \mathbf{I} \\ \text{Cov}(\mathbf{y}_P, \mathbf{y}_S) &= \sigma_{A/Am} \mathbf{U}_{Am} \mathbf{U}_A^T + \sigma_{D/Dm} \mathbf{U}_{Dm} \mathbf{U}_D^T + \sigma_{C/C} \mathbf{U}_C \mathbf{U}_C^T \\ &\quad + \sigma_{Am/Am} \mathbf{U}_{Am} \mathbf{U}_{Am}^T + \sigma_{Dm/Dm} \mathbf{U}_{Dm} \mathbf{U}_{Dm}^T + \sigma_{B/B} \mathbf{U}_B \mathbf{U}_B^T + \sigma_{e/e} \mathbf{I} \end{aligned}$$

其中 $\sigma_{A/Am}$ 是种子性状直接加性效应与母体植株性状加性效应的协方差分量, $\sigma_{D/Dm}$ 是种子性状直接显性效应与母体植株性状显性效应的协方差分量, $\sigma_{C/C}$ 是种子性状细胞质效应与母体植株性状细胞质效应的协方差分量, $\sigma_{Am/Am}$ 是种子性状母体加性效应与母体植株性状加性效应的协方差分量, $\sigma_{Dm/Dm}$ 是种子性状母体显性效应与母体植株性状显性效应的协方差分量, $\sigma_{B/B}$ 是种子性状区组效应与母体植株性状区组效应的协方差分量, $\sigma_{e/e}$ 是种子性状剩余效应与母体植株性状剩余效应的协方差分量。以上二个协方差矩阵之和可以改写为便于运算的形式

$$\begin{aligned} \mathbf{V} &= \text{Cov}(\mathbf{y}_S, \mathbf{y}_P) + \text{Cov}(\mathbf{y}_P, \mathbf{y}_S) \\ &= \sigma_{A/Am} \mathbf{F}_1 + \sigma_{D/Dm} \mathbf{F}_2 + \sigma_{C/C} \mathbf{F}_3 + \sigma_{Am/Am} \mathbf{F}_4 + \sigma_{Dm/Dm} \mathbf{F}_5 \\ &\quad + \sigma_{B/B} \mathbf{F}_6 + \sigma_{e/e} \mathbf{F}_7 \\ &= \sum_{u=1}^7 \sigma_{u/u} \mathbf{F}_u \end{aligned}$$

其中 $\mathbf{F}_1 = (\mathbf{U}_A \mathbf{U}_{Am}^T + \mathbf{U}_{Am} \mathbf{U}_A^T)$ 、 $\mathbf{F}_2 = (\mathbf{U}_D \mathbf{U}_{Dm}^T + \mathbf{U}_{Dm} \mathbf{U}_D^T)$ 、 $\mathbf{F}_3 = (2\mathbf{U}_C \mathbf{U}_C^T)$ 、
 $\mathbf{F}_4 = (2\mathbf{U}_{Am} \mathbf{U}_{Am}^T)$ 、 $\mathbf{F}_5 = (2\mathbf{U}_{Dm} \mathbf{U}_{Dm}^T)$ 、 $\mathbf{F}_6 = (2\mathbf{U}_B \mathbf{U}_B^T)$ 、 $\mathbf{F}_7 = 2\mathbf{I}$ 。

采用MINQUE(0/1)法，种子数量性状与母体植株的农艺性状的协方差分量可由解以下方程组而得， $u, v = 1, 2, \dots, 7$ ：

$$[\text{tr}(\mathbf{Q}_{(0/1)} \mathbf{F}_u \mathbf{Q}_{(0/1)} \mathbf{F}_v)] [\hat{\sigma}_{u/v}] = [2\mathbf{y}_S^T \mathbf{Q}_{(0/1)} \mathbf{F}_u \mathbf{Q}_{(0/1)} \mathbf{y}_P]$$

其中

$$\mathbf{Q}_{(0/1)} = \mathbf{V}_{(0/1)}^{-1} - \mathbf{V}_{(0/1)}^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}_{(0/1)}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}_{(0/1)}^{-1}$$

$$\mathbf{V}_{(0/1)} = 2[\mathbf{U}_C \mathbf{U}_C^T + \mathbf{U}_{Am} \mathbf{U}_{Am}^T + \mathbf{U}_{Dm} \mathbf{U}_{Dm}^T + \mathbf{U}_B \mathbf{U}_B^T + \mathbf{I}]$$

采用以上方法估算非等矩阵的成对性状协方差分量时有一条约束条件，即 \mathbf{U}_A 和 \mathbf{U}_{Am} 、 \mathbf{U}_D 和 \mathbf{U}_{Dm} 列数分别相等。

第二节 基因效应分析

1. 最小二乘估计法

在回归分析中常采用最小二乘(least squares)估计法估算回归系数。数量遗传分析中也可用最小二乘法估算基因效应。当不存在基因型 \times 环境互作时，遗传模型可以用以下混合线性模型简式表示，

$$\mathbf{y} = \mathbf{Xb} + \mathbf{U}_G \mathbf{e}_G + \mathbf{e}_\varepsilon$$

其中

$$\mathbf{U}_G = [\mathbf{U}_A : \mathbf{U}_D : \dots]$$

$$\mathbf{e}_G^T = [\mathbf{e}_A^T : \mathbf{e}_D^T : \dots]$$

\mathbf{e}_ε 是随机机误， $\mathbf{e}_\varepsilon \sim \text{MVN}(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I})$ 。

采用回归分析时一般要求回归系数为固定效应，所以在估算基因效应时需将遗传效应定义为固定效应。当 \mathbf{e}_G 是固定效应时，以上遗传模型可以该写为，

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\mathbf{b} + \mathbf{U}_G \mathbf{e}_G + \mathbf{e}_\varepsilon \\ &= \mathbf{W}\boldsymbol{\beta} + \mathbf{e}_\varepsilon \\ &\sim \text{MVN}(\mathbf{W}\boldsymbol{\beta}, \sigma_\varepsilon^2 \mathbf{I}) \end{aligned}$$

其中

$$\begin{aligned} \mathbf{W} &= [\mathbf{X}; \mathbf{U}_G] \\ \boldsymbol{\beta}^T &= [\mathbf{b}^T; \mathbf{e}_G^T] \end{aligned}$$

固定效应 的估计值可由以下正规方程组解得，

$$(\mathbf{W}^T \mathbf{W})\boldsymbol{\beta} = \mathbf{W}^T \mathbf{y}$$

当常数矩阵 \mathbf{W} 是满秩(full rank)的情况下，存在逆矩阵 $(\mathbf{W}^T \mathbf{W})^{-1}$ ，固定效应 的最佳无偏估计值为

$$\hat{\boldsymbol{\beta}} = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{y}$$

此时

$$\begin{aligned} E(\hat{\boldsymbol{\beta}}) &= E[(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{y}] \\ &= (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T E[\mathbf{y}] \\ &= (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{W}\boldsymbol{\beta} \\ &= \boldsymbol{\beta} \end{aligned}$$

$$\begin{aligned} \text{Var}(\hat{\boldsymbol{\beta}}) &= \text{Var}[(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{y}] \\ &= (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \text{Var}[\mathbf{y}] \mathbf{W} (\mathbf{W}^T \mathbf{W})^{-1} \\ &= (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T (\sigma_\varepsilon^2 \mathbf{I}) \mathbf{W} (\mathbf{W}^T \mathbf{W})^{-1} \\ &= \sigma_\varepsilon^2 (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{W} (\mathbf{W}^T \mathbf{W})^{-1} \\ &= \sigma_\varepsilon^2 (\mathbf{W}^T \mathbf{W})^{-1} \end{aligned}$$

$$\hat{\boldsymbol{\beta}} \sim \text{MVN}(\boldsymbol{\beta}, \sigma_\varepsilon^2 (\mathbf{W}^T \mathbf{W})^{-1})$$

在遗传分析的模型中常数矩阵 \mathbf{W} 一般都不是满秩的，因而不存在 $(\mathbf{W}^T \mathbf{W})$ 的逆矩阵。采用 $(\mathbf{W}^T \mathbf{W})$ 的广义逆矩阵 $(\mathbf{W}^T \mathbf{W})^-$ ，可以获得正规方程组的一组解，

$$\hat{\beta} = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{y}$$

在这里 $\hat{\beta}$ 只是正规方程组的一组解，并不是回归参数的估计值。在遗传分析中，如果不对遗传模型中的环境效应和基因效应值设定必要的约束条件，固定效应一般是不可估计的(unestimable)。

尽管回归参数不可估计，但是回归参数的函数 $\mathbf{W}\hat{\beta}$ 是可以被无偏估计的。

$$\begin{aligned} \mathbf{W}\hat{\beta} &= [\mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T] \mathbf{y} \\ &= \mathbf{P}_W \mathbf{y} \end{aligned}$$

$$\begin{aligned} E(\mathbf{W}\hat{\beta}) &= E([\mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T] \mathbf{y}) \\ &= \mathbf{P}_W E(\mathbf{y}) \\ &= \mathbf{P}_W \mathbf{W}\beta \\ &= \mathbf{W}\beta \\ &= E(\mathbf{y}) \end{aligned}$$

$$\begin{aligned} \text{Var}(\mathbf{W}\hat{\beta}) &= \text{Var}([\mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T] \mathbf{y}) \\ &= \mathbf{P}_W \text{Var}(\mathbf{y}) \mathbf{P}_W \\ &= \mathbf{P}_W (\sigma_\varepsilon^2 \mathbf{I}) \mathbf{P}_W \\ &= \sigma_\varepsilon^2 \mathbf{P}_W \mathbf{P}_W \\ &= \sigma_\varepsilon^2 \mathbf{P}_W \end{aligned}$$

$$\mathbf{W}\hat{\beta} \sim \text{MVN}(\mathbf{W}\beta, \sigma_\varepsilon^2 \mathbf{P}_W)$$

这表明每个观察值的环境效应和基因效应的总值是可以无偏估计的。如果 \mathbf{a} 是任一常数向量， $\mathbf{a}^T \mathbf{W}\hat{\beta} = \mathbf{c}^T \beta$ 也是可以无偏估计的。 $\mathbf{c}^T \beta$ 是可估计线性函数，它具有以下一些特征：

对于任意 $\hat{\beta}$ ， $\mathbf{c}^T \hat{\beta}$ 是不变量。

如果回归模型的正规方程组有一组解 $\hat{\beta} = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{y}$ ，另一组解是

$$\hat{\beta}^o = [(\mathbf{W}^T \mathbf{W})^{-1} + (\mathbf{I} - (\mathbf{W}^T \mathbf{W})^{-1} (\mathbf{W}^T \mathbf{W}) \mathbf{A})] \mathbf{W}^T \mathbf{y}$$

其中 \mathbf{A} 是任意矩阵。那么

$$\begin{aligned} \mathbf{c}^T \hat{\beta} &= \mathbf{c}^T (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{y} \\ &= \mathbf{a}^T \mathbf{W} (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{y} \\ &= \mathbf{a}^T \mathbf{P}_W \mathbf{y} \end{aligned}$$

$$\begin{aligned} \mathbf{c}^T \hat{\beta}^o &= \mathbf{c}^T [(\mathbf{W}^T \mathbf{W})^{-1} + (\mathbf{I} - (\mathbf{W}^T \mathbf{W})^{-1} (\mathbf{W}^T \mathbf{W}) \mathbf{A})] \mathbf{W}^T \mathbf{y} \\ &= \mathbf{a}^T \mathbf{W} [(\mathbf{W}^T \mathbf{W})^{-1} + (\mathbf{I} - (\mathbf{W}^T \mathbf{W})^{-1} (\mathbf{W}^T \mathbf{W}) \mathbf{A})] \mathbf{W}^T \mathbf{y} \\ &= \mathbf{a}^T (\mathbf{P}_W + \mathbf{0}) \mathbf{y} \\ &= \mathbf{a}^T \mathbf{P}_W \mathbf{y} \\ &= \mathbf{c}^T \hat{\beta} \end{aligned}$$

对于任意 $\hat{\beta}$, $\mathbf{c}^T \hat{\beta}$ 是 $\mathbf{c}^T \beta$ 的最佳线性无偏估计 (best linear unbiased estimation, 简称BLUE)。这是线性回归分析中著名的 Gauss - Markov 定理。

$$\begin{aligned} E(\mathbf{c}^T \hat{\beta}) &= E(\mathbf{c}^T (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{y}) \\ &= \mathbf{a}^T \mathbf{W} (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T E(\mathbf{y}) \\ &= \mathbf{a}^T \mathbf{P}_W \mathbf{W} \beta \\ &= \mathbf{a}^T \mathbf{W} \beta \\ &= \mathbf{c}^T \beta \end{aligned}$$

所以 $\mathbf{c}^T \hat{\beta}$ 是 $\mathbf{c}^T \beta$ 的线性无偏估计。

$$\begin{aligned} \text{Var}(\mathbf{c}^T \hat{\beta}) &= \text{Var}(\mathbf{c}^T (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{y}) \\ &= \text{Var}(\mathbf{a}^T \mathbf{W} (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{y}) \\ &= \text{Var}(\mathbf{a}^T \mathbf{P}_W \mathbf{y}) \\ &= \mathbf{a}^T \mathbf{P}_W [\text{Var}(\mathbf{y})] \mathbf{P}_W \mathbf{a} \\ &= \mathbf{a}^T \mathbf{P}_W [\sigma_\varepsilon^2 \mathbf{I}] \mathbf{P}_W \mathbf{a} \\ &= \sigma_\varepsilon^2 \mathbf{a}^T \mathbf{P}_W \mathbf{P}_W \mathbf{a} \\ &= \sigma_\varepsilon^2 \mathbf{a}^T \mathbf{P}_W \mathbf{a} \\ &= \sigma_\varepsilon^2 \mathbf{c}^T (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{c} \end{aligned}$$

如果设 $\mathbf{a}^T \mathbf{y}$ 为线性函数 $\mathbf{c}^T \beta$ 的另一个线性无偏估计 (LUE) , $\mathbf{a}^T \mathbf{W} = \mathbf{c}^T$,

$$\begin{aligned}\text{Var}(\mathbf{a}^T \mathbf{y}) &= \text{Var}(\mathbf{a}^T \mathbf{y} - \mathbf{c}^T \hat{\beta} + \mathbf{c}^T \hat{\beta}) \\ &= \text{Var}(\mathbf{a}^T \mathbf{y} - \mathbf{c}^T \hat{\beta}) + \text{Var}(\mathbf{c}^T \hat{\beta}) + 2\text{Cov}(\mathbf{a}^T \mathbf{y} - \mathbf{c}^T \hat{\beta}, \mathbf{c}^T \hat{\beta}) \\ &= \text{Var}(\mathbf{a}^T \mathbf{y} - \mathbf{c}^T \hat{\beta}) + \text{Var}(\mathbf{c}^T \hat{\beta}) \geq \text{Var}(\mathbf{c}^T \hat{\beta})\end{aligned}$$

因为

$$\begin{aligned}\text{Cov}(\mathbf{a}^T \mathbf{y} - \mathbf{c}^T \hat{\beta}, \mathbf{c}^T \hat{\beta}) &= \text{Cov}(\mathbf{a}^T \mathbf{y} - \mathbf{a}^T \mathbf{P}_W \mathbf{y}, \mathbf{a}^T \mathbf{P}_W \mathbf{y}) \\ &= \sigma_\varepsilon^2 \mathbf{a}^T (\mathbf{I} - \mathbf{P}_W) \mathbf{P}_W \mathbf{a} \\ &= \mathbf{0}\end{aligned}$$

所以 $\mathbf{c}^T \hat{\beta}$ 是 $\mathbf{c}^T \beta$ 的一致最小方差的估计。一致最小方差的线性无偏估计就是所谓的最佳线性无偏估计。

遗传效应之间的差异可以用 $\mathbf{c}^T \beta$ 求得，并可进行相应的显著性测验。当

$$\left| \frac{\mathbf{c}^T \hat{\beta}}{\sqrt{\text{Var}(\mathbf{c}^T \hat{\beta})}} \right| > z(\alpha)$$

时，可以推测遗传效应之间存在显著性差异。

在通常情况下遗传模型中的基因效应参数是不可估计的。但是通过对基因效应参数设置一些必要的约束条件，有可能使参数被无偏地估算。对于遗传模型，

$$\mathbf{y} = \mathbf{W}\beta + \mathbf{e}_\varepsilon$$

如果设置联合不可估计的约束(jointly non - estimable constraints)

$$\mathbf{C}\beta = \mathbf{0}$$

其中矩阵的秩为 $(\mathbf{W}) = r$ ， $(\mathbf{C}) = p - r = s$ ，便可以获得该约束性模型回归参数的无偏估计

$$\hat{\beta}_c = (\mathbf{W}^T \mathbf{W} + \mathbf{C}^T \mathbf{C})^{-1} \mathbf{W}^T \mathbf{y}$$

对于简单的加性-显性模型，约束条件常为 $\sum_i A_i = 0$ ，及 $\sum_j D_{ij} = 0$ 。

2. 线性无偏预测法

采用回归分析方法分析遗传效应时，必须假定遗传效应为固定效应，即所分析的遗传群体是一组特定的材料。因此，分析所获得的遗传结论也只适用于此组遗传材料。在分析遗传效应值的同时不能再估算相应的遗传方差分量。

在数量性状的遗传分析中，采用的遗传材料通常是取自某遗传群体的一组随机样本。人们感兴趣的主要是该遗传群体的遗传变异性。通过估算各项遗传方差分量，可以推断各项基因效应的变异性。在有些试验中，试验者不但需要估算方差分量，有时还希望能推断遗传模型中某些基因效应的值。遗传模型中的各项遗传效应通常是不可估计的，但是采用混合线性模型的一些分析方法，模型中的随机效应却是可预测的(predictable)。对于混合线性模型

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \sum_{u=1}^{m-1} \mathbf{U}_u \mathbf{e}_u + \mathbf{e}_m$$

$$\sim \text{MVN}(\mathbf{X}\mathbf{b}, \mathbf{V} = \sum_{u=1}^{m-1} \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T + \sigma_m^2 \mathbf{I})$$

第 u 项随机因素具有多元正态分布 $\mathbf{e}_u \sim \text{MNV}(\mathbf{0}, \sigma_u^2 \mathbf{I})$ 。Henderson(1963)提出，如果随机因素的方差分量 σ_u^2 是已知的，可以获得随机效应的最佳线性无偏预测(best linear unbiased prediction, 简称BLUP)。第 u 项随机因素效应值 \mathbf{e}_u 的最佳线性无偏预测可以由下式估算得到

$$\hat{\mathbf{e}}_u = \sigma_u^2 \mathbf{U}_u^T \mathbf{V}^{-1} (\mathbf{y} - \hat{\mathbf{X}}\mathbf{b})$$

$$= \sigma_u^2 \mathbf{U}_u^T \mathbf{Q} \mathbf{y}$$

其中

$$\mathbf{Q} = \mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1}$$

$$\mathbf{V} = \sum_{u=1}^{m-1} \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T + \sigma_m^2 \mathbf{I} \text{ 并有逆矩阵 } \mathbf{V}^{-1}$$

$$\hat{\mathbf{b}} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{y}$$

按上式预测的随机效应是依变量 \mathbf{y} 的线性函数，因而也是多元正态随机变量

$$\begin{aligned}
 E(\hat{\mathbf{e}}_u) &= E(\sigma_u^2 \mathbf{U}_u^T \mathbf{Q} \mathbf{y}) \\
 &= \sigma_u^2 \mathbf{U}_u^T \mathbf{Q} E(\mathbf{y}) \\
 &= \sigma_u^2 \mathbf{U}_u^T \mathbf{Q} \mathbf{X} \mathbf{b} \\
 &= \sigma_u^2 \mathbf{U}_u^T \mathbf{0} \mathbf{b} \\
 &= \mathbf{0}
 \end{aligned}$$

因为

$$\begin{aligned}
 \mathbf{Q} \mathbf{X} &= \mathbf{V}^{-1/2} (\mathbf{I} - \mathbf{P}_W) \mathbf{V}^{-1/2} \mathbf{X} \\
 &= \mathbf{V}^{-1/2} (\mathbf{I} - \mathbf{P}_W) \mathbf{W} \\
 &= \mathbf{V}^{-1/2} (\mathbf{W} - \mathbf{W}) \\
 &= \mathbf{0}
 \end{aligned}$$

其中

$$\begin{aligned}
 \mathbf{W} &= \mathbf{V}^{-1/2} \mathbf{X} \\
 \mathbf{P}_W &= \mathbf{W} (\mathbf{W}^T \mathbf{W}) \mathbf{W}^T
 \end{aligned}$$

还可以进一步推导最佳线性无偏预测值 $\hat{\mathbf{e}}_u$ 的方差，

$$\begin{aligned}
 \text{Var}(\hat{\mathbf{e}}_u) &= \text{Var}(\sigma_u^2 \mathbf{U}_u^T \mathbf{Q} \mathbf{y}) \\
 &= (\sigma_u^2)^2 \mathbf{U}_u^T \mathbf{Q} \text{Var}(\mathbf{y}) \mathbf{Q} \mathbf{U}_u \\
 &= (\sigma_u^2)^2 \mathbf{U}_u^T \mathbf{Q} \mathbf{V} \mathbf{Q} \mathbf{U}_u \\
 &= (\sigma_u^2)^2 \mathbf{U}_u^T \mathbf{Q} \mathbf{U}_u
 \end{aligned}$$

因而 $\hat{\mathbf{e}}_u$ 具有多变量正态分布

$$\hat{\mathbf{e}}_u \sim \text{MVN}(\mathbf{0}, \sigma_u^4 \mathbf{U}_u^T \mathbf{Q} \mathbf{U}_u)$$

随机效应的预测值是参数的无偏预测，并且在所有的线性无偏预测中具有最小方差。

在运用混合线性模型分析具体试验数据时，方差分量 σ_u^2 是需要被估算的。因而实际上并不可能获得随机效应值的最佳线性无偏预测。在实际应用时，常采用方差分量的估计值 $\hat{\sigma}_u^2$ 代替其参数，由以下公式对随机效应进行预测

$$\hat{\mathbf{e}}_u = \hat{\sigma}_u^2 \mathbf{U}_u^T \hat{\mathbf{Q}} \mathbf{y}$$

其中

$$\hat{\mathbf{Q}} = \hat{\mathbf{V}}^{-1} - \hat{\mathbf{V}}^{-1} \mathbf{X} (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{V}}^{-1}$$

$$\hat{\mathbf{V}} = \sum_{u=1}^{m-1} \hat{\sigma}_u^2 \mathbf{U}_u \mathbf{U}_u^T + \hat{\sigma}_m^2 \mathbf{I} \text{ 并有逆矩阵 } \hat{\mathbf{V}}^{-1}$$

这时的预测值已不是真正的BLUP预测值，而是所谓的“BLUP”预测值，因为它即不是依变量的线性函数，也不能保证是无偏的估计值了。

朱军(1992)建议采用运算较为简便的MINQUE法，预测随机效应值。如果在用MINQUE法估算方差分量时选用了先验值 $[\alpha_u]$ ，随机效应的MINQUE法预测值便为

$$\hat{\mathbf{e}}_u = \alpha_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y}$$

其中

$$\mathbf{Q}_\alpha = \mathbf{V}_\alpha^{-1} - \mathbf{V}_\alpha^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}_\alpha^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}_\alpha^{-1}$$

$$\mathbf{V}_\alpha = \sum_{u=1}^{m-1} \alpha_u \mathbf{U}_u \mathbf{U}_u^T + \alpha_m \mathbf{I} \text{ 并有逆矩阵 } \mathbf{V}_\alpha^{-1}$$

如果先验值 $[\alpha_u]$ 的选择不取决于所分析的数据，该预测值是观察值向量 \mathbf{y} 的线性函数，也是随机效应的无偏预测值。所以这一方法称为线性无偏预测法(linear unbiased prediction, 简称LUP)。 $\hat{\mathbf{e}}_u$ 无偏性的证明如下：

$$\begin{aligned} E(\hat{\mathbf{e}}_u) &= E(\alpha_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y}) \\ &= \alpha_u \mathbf{U}_u^T \mathbf{Q}_\alpha E(\mathbf{y}) \\ &= \alpha_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{X} \mathbf{b} \\ &= \alpha_u \mathbf{U}_u^T \mathbf{0} \mathbf{b} \\ &= \mathbf{0} \end{aligned}$$

可用同样方法证明 $\mathbf{Q}_\alpha \mathbf{X} = \mathbf{0}$ 。还可以进一步推导线性无偏预测值 $\hat{\mathbf{e}}_u$ 的方差，

$$\begin{aligned}\text{Var}(\hat{\mathbf{e}}_u) &= \text{Var}(\alpha_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y}) \\ &= (\alpha_u)^2 \mathbf{U}_u^T \mathbf{Q}_\alpha \text{Var}(\mathbf{y}) \mathbf{Q}_\alpha \mathbf{U}_u \\ &= \alpha_u^2 \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{V} \mathbf{Q}_\alpha \mathbf{U}_u\end{aligned}$$

因而 $\hat{\mathbf{e}}_u$ 具有多变量正态分布

$$\hat{\mathbf{e}}_u \sim \text{MVN}(\mathbf{0}, \alpha_u^2 \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{V} \mathbf{Q}_\alpha \mathbf{U}_u)$$

如果运用MINQUE(1), 设定所有先验值为1 [$\alpha_u = 1$], 便可以在估算方差分量的同时获得随机效应的MINQUE(1)预测值。

仔细比较MINQUE法估算方差分量的正规方程组, 可以发现预测随机效应的MINQUE公式的主要部分 $\mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y}$ 是MINQUE方程组右边的组成部分。因而在用MINQUE法估算混合线性模型中的方差分量时, 不必进行很多的额外运算, 便可获得随机效应的MINQUE法预测值。

采用LUP法虽然可以获得随机效应的无偏预测值, 但这些预测值的变异幅度往往不够大。朱军(1993)建议采用以下公式求调整的无偏预测(adjusted unbiased prediction, 简称AUP)。可以使随机效应预测值的均值和方差都具有无偏性。第 u 项遗传因素的效应值向量 \mathbf{e}_u 的AUP预测值为

$$\hat{\mathbf{e}}_u = \kappa_u (\alpha_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y})$$

其中 $\alpha_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y}$ 是 \mathbf{e}_u 的线性无偏预测(LUP),

κ_u 是预测的调整系数,

$$\kappa_u = \sqrt{(n_u - 1) \hat{\sigma}_u^2 / (\alpha_u^2 \mathbf{y}^T \mathbf{Q}_\alpha \mathbf{U}_u \mathbf{U}_u^T \mathbf{Q}_\alpha \mathbf{y})},$$

当 $\hat{\sigma}_u^2 < 0$ 时, 取 $\hat{\sigma}_u^2 = 0$,

n_u 是系数矩阵 \mathbf{U}_u 的列数, $\hat{\sigma}_u^2$ 是第 u 项方差分量的估计值。

现以一个简单的加性-显性遗传模型为例, 说明运用LUP法和AUP法预测基因型值的过程。当运用MINQUE(1)法估算遗传方差分量时, 可以同时获得加性效应和显性效应的无偏预测。LUP预测值为

$$\text{加性效应向量 } \hat{\mathbf{e}}_A = \mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{y}$$

$$\text{显性效应向量 } \hat{\mathbf{e}}_D = \mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{y}$$

加性效应和显性效应的AUP预测值为

$$\text{加性效应向量 } \hat{\mathbf{e}}_A = \kappa_A \mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{y}$$

$$\text{其中 } \kappa_A = \sqrt{(n_A - 1) \hat{\sigma}_A^2 / (\mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{y})}$$

$$\text{显性效应向量 } \hat{\mathbf{e}}_D = \kappa_D \mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{y}$$

$$\text{其中 } \kappa_D = \sqrt{(n_D - 1) \hat{\sigma}_D^2 / (\mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{y})}$$

在种子遗传模型中如果随机变量 \mathbf{e}_u 与变量 \mathbf{e}_v 存在协方差 $\sigma_{u,v} \mathbf{I}$ ，随机变量 \mathbf{e}_u 的BLUP预测公式为，

$$\hat{\mathbf{e}}_u = (\sigma_u^2 \mathbf{U}_u^T + \sigma_{u,v} \mathbf{U}_v^T) \mathbf{Q} \mathbf{y}$$

其中

$$\mathbf{Q} = \mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^+ \mathbf{X}^T \mathbf{V}^{-1}$$

$$\mathbf{V} = \sum_u \theta_u \mathbf{V}_u$$

在实际应用时以参数的估计值 $\hat{\theta}_u$ 代入上式，可以获得“BLUP”的预测值。

如果采用MINQUE(0/1)分析种子(或幼畜)的数量性状，各项遗传效应 \mathbf{e}_u ，(u=1, 2, ..., 5)的LUP预测值可由以下公式算得，

$$\hat{\mathbf{e}}_u = \mathbf{U}_u^T \mathbf{Q}_{(0/1)} \mathbf{y}$$

其中

$$\mathbf{Q}_{(0/1)} = \mathbf{V}_{(0/1)}^{-1} - \mathbf{V}_{(0/1)}^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}_{(0/1)}^{-1} \mathbf{X})^+ \mathbf{X}^T \mathbf{V}_{(0/1)}^{-1}$$

$$\mathbf{V}_{(0/1)} = \sum_{u=1}^6 \mathbf{U}_u \mathbf{U}_u^T$$

虽然LUP值是遗传效应值的线性无偏预测，但由LUP预测值估计的遗传方差

分量一般均略小于方差分量的MINQUE估计值。为了纠正这一偏差，朱军(1993)提出了调整无偏预测方法，简称AUP法。第 u 项遗传因素的效应值向量 \mathbf{e}_u 的AUP预测值为

$$\hat{\mathbf{e}}_u = \kappa \mathbf{U}_u^T \mathbf{Q}_{(0/1)} \mathbf{y}$$

其中调整系数 $\kappa = \sqrt{(n_u - 1) \hat{\sigma}_u^2 / (\mathbf{y}^T \mathbf{Q}_{(0/1)} \mathbf{V}_u \mathbf{Q}_{(0/1)} \mathbf{y})}$ ， n_u 是系数矩阵 \mathbf{U}_u 的列数， $\hat{\sigma}_u^2$ 是第 u 项方差分量的估计值， $\hat{\sigma}_u^2 > 0$ 。

第三节 几种蒙特卡罗模拟分析方法的效益比较

1. 方差分量估算效益的比较

采用蒙特卡罗模拟(Monte Carlo simulation)可以比较不同统计分析方法的效益(Swallow和Monahan, 1984)。朱军(1989)用蒙特卡罗模拟比较了Henderson方法3(HEND3)、ML、REML、MINQUE(1)和MINQUE()等五种分析方法估算遗传方差分量的效益。MINQUE()设置所有的先验值为参数值($\alpha_u = \sigma_u^2$)，它是具有最小方差的无偏估算方法(Rao, 1972)，因此在所有介绍的混合线型模型分析方法中是最佳的一种估算方法。

采用的遗传模型是Cockerham和Weir(1977)提出的双列杂交生物模型(bio-model)(见第七章第一节)，根据100次蒙特卡罗模拟结果，计算了平均的方差估计值($\hat{\sigma}^2$ 或 \hat{V})、参数估计的偏差(Bias = 估计值—参数值)、均方误差(MSE = Bias² + Var(估计值))。在相同参数情况下，均方误差越小，表明估计的效益越高。

表4.4是6亲本双列杂交(方法3)平衡数据的分析结果。由表4.4可知，REML、MINQUE(1)和MINQUE()三种方法对平衡数据方差分量的估算效益是相同的。HEND3对 V_A 和 V_D 的估算结果与前三种方法一致，但对 V_M 和 V_P 的估算导致了极大的偏差。由于bio-model的 V_M 和 V_P 方差分量是无法用ANOVA方法无偏估算的

(Cockerham和Weir, 1977), 而Henderson方法是仿效ANOVA方法的, 因此也不能估算 V_M 和 V_P 方差分量。但是在巢式设计或析因设计的蒙特卡罗模拟比较中, HEND3与REML、MINQUE(1)和MINQUE()一样可以无偏地估算各项方差分量(朱军, 1989)。由ML法所估算的第一项方差分量有极其严重的负偏差, 因此ML方不是一种有效的方差分量估算方法。在巢式设计或析因设计的蒙特卡罗模拟比较中, ML法也有偏地低估了第一项方差分量(朱军, 1989)。

表4.4 用五种方法估算6个亲本双列杂交(方法3)平衡数据的遗传方差分量的偏差(Bias)和均方误差(MSE)的比较

估算方法	$V_A = 100$		$V_D = 30$		$V_M = 20$		$V_P = 20$	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
HEND3	5.49	6430.3	2.61	438.3	$>10^5$	$>10^{11}$	$>10^4$	$>10^9$
ML	-42.61	5270.4	4.81	450.8	8.72	818.4	1.80	567.5
REML	5.49	6430.3	2.61	438.3	3.03	833.6	2.19	636.3
MINQUE(1)	5.49	6430.3	2.61	438.3	3.03	833.6	2.19	636.3
MINQUE(0)	5.49	6430.3	2.61	438.3	3.03	833.6	2.19	636.3

$V_e = 30$, Bias = -1.33, MSE = 153.9.

表4.5 三种方法估算10个亲本双列杂交(方法3)非平衡数据的遗传方差分量和遗传率的偏差(Bias)和均方误差(MSE)的比较

参数	真值	REML		MINQUE(1)		MINQUE(0)	
		Bias	MSE	Bias	MSE	Bias	MSE
V_A	100	-3.58	3522.1	-3.24	3562.9	-3.29	3574.5
V_D	30	-2.30	254.2	-2.34	255.7	-2.33	254.5
V_M	20	4.96	450.4	5.08	464.0	5.11	464.0
V_P	20	-1.09	369.1	-0.98	379.2	-1.00	381.7
V_e	30	0.33	78.8	0.32	78.2	0.31	77.8
h_N^2 (%)	50	-5.39	352.9	-5.37	351.3	-5.42	354.5

表4.5是10个亲本双列杂交(方法3)非平衡数据的分析结果。该设计假设亲本1—6之间不能相互杂交, 另外还缺失(1×9)和(1×10)的正反杂交组合。由于HEND3和ML不能有效地分析双列杂交的平衡数据, 因此在非平衡数据的模拟分析将不包括这二种方法。由表4.5可知, REML、MINQUE(1)和MINQUE()三种方法对非平衡数据方差分量的估算效益是很相似的。

在表4.6和表4.7, 进一步比较了REML、MINQUE(1)和MINQUE()三种方法对不同方差及协方差参数值的估算效益, 这可检验统计分析方法的稳健性(robustness)。仍然采用Cockerham和Weir (1977) 提出的生物模型(bio-model), 根据200次蒙特卡罗模拟结果, 比较了不同参数条件下的参数估计的偏差(Bias)和均方误差(MSE)。由表4.6可知, REML、MINQUE(1)和MINQUE()三种方法对非平衡数据方差分量的估算效益不尽相同的。当方差分量为非零值时, 这三种方法均可获得无偏的估计值, 分析效益相似。但当有些方差分量(如 V_M 和 V_P)为零时, MINQUE(1)和MINQUE()可以获得接近参数真值的估计, 而REML则总是过高地估计零参数。这种偏差是由于REML在估算参数时采用了估计值必须大于零的限定。在估算成对性状间的协方差分量时, REML法无须设计限定条件, 因此REML与MINQUE(1)和MINQUE()一样, 可以无偏地估算协方差分量。以上蒙特卡罗模拟结果表明, MINQUE(1)的估算效益与最小方差无偏估算方法MINQUE()相似, 当方差分量参数值为零时MINQUE(1)比REML更能有效地无偏估计参数。MINQUE法比REML法的另一个优点是无需作迭代运算。因此, 我们认为在混合模型的分析中可以采用MINQUE法估算方差分量和协方差分量。

表4.6 三种方法估算十个亲本双列杂交 (bio-model) 非平衡数据的遗传方差分量的偏差(Bias)和均方误差(MSE)的比较

参数	真值	MINQUE(0)		MINQUE(1)		REML	
		Bias	MSE	Bias	MSE	Bias	MSE
σ_n^2	50	0.78	1279	0.70	1267	1.00	1260
σ_t^2	30	0.86	158	0.88	159	0.85	161
σ_m^2	20	1.60	506	1.56	508	2.88	449
σ_p^2	20	-1.89	362	-1.92	361	-0.51	286
σ_e^2	30	0.00	13	0.00	13	-0.01	13
σ_n^2	50	2.20	1105	2.13	1096	2.33	1105
σ_t^2	30	0.86	158	0.88	159	0.83	159
σ_m^2	20	0.24	315	0.15	319	0.61	304
σ_p^2	0	-0.01	184	-0.02	187	5.07	96
σ_e^2	30	-0.00	13	0.01	13	-0.00	13
σ_n^2	50	2.37	878	2.33	872	2.53	877
σ_t^2	30	0.87	158	0.88	159	0.81	158
σ_m^2	0	0.16	21	0.30	32	2.22	17
σ_p^2	0	-0.10	23	-0.24	31	2.07	15
σ_e^2	30	0.01	13	0.01	13	-0.01	13

表4.7 三种方法估算十个亲本双列杂交 (bio-model) 非平衡数据的遗传协方差分量的偏差(Bias)和均方误差(MSE)的比较

参数	真值	MINQUE(0)		MINQUE(1)		REML	
		Bias	MSE	Bias	MSE	Bias	MSE
$\sigma_{n/n}$	25	-0.15	669	-0.15	668	1.11	563
$\sigma_{t/t}$	15	0.79	73	0.78	74	1.72	96
$\sigma_{m/m}$	10	0.26	239	0.24	240	2.33	172
$\sigma_{p/p}$	10	-0.37	281	-0.35	282	2.06	212
$\sigma_{e/e}$	15	0.01	10	0.01	10	0.01	10
$\sigma_{n/n}$	25	0.11	521	0.08	524	0.79	461
$\sigma_{t/t}$	15	0.76	73	0.77	73	1.42	104
$\sigma_{m/m}$	10	0.06	157	0.05	159	1.14	128
$\sigma_{p/p}$	0	-0.27	98	-0.31	99	3.91	55
$\sigma_{e/e}$	15	0.01	10	0.01	10	0.00	10
$\sigma_{n/n}$	25	0.34	490	0.38	491	1.36	429
$\sigma_{t/t}$	15	0.76	73	0.76	73	1.14	90
$\sigma_{m/m}$	0	0.44	12	0.47	16	1.68	10
$\sigma_{p/p}$	0	-0.49	12	-0.57	15	1.17	5
$\sigma_{e/e}$	15	0.00	9	0.01	10	-0.00	10

2. 遗传效应预测效益的比较

混合模型中随机效应预测效益的比较有三条指标： 预测值的平均偏差是否为零($\text{Bias} = \mathbf{1}^T(\hat{\mathbf{e}} - \mathbf{e})/n = \mathbf{1}^T\hat{\mathbf{e}}/n$)； 用预测值计算的方差($\text{Var.} = \frac{\hat{\mathbf{e}}^T\hat{\mathbf{e}}}{df}$)是否接近方差参数； 模型中随机遗传效应的预测值向量 $\hat{\mathbf{e}}$ 与实际值向量 $\tilde{\mathbf{e}}$ 之间的距离($\text{Dis.} = \|\hat{\mathbf{e}} - \tilde{\mathbf{e}}\|$)较小。采用蒙特卡罗模拟的方法比较了二种随机效应预测方法 (AUP和“BLUP”)与BLUP法的预测效益。模拟所采用的遗传设计与表4.6和表4.7模拟试

验的相同。对于四项遗传效应，三种预测方法的预测值偏差绝对值均为 $10^{-5} \sim 10^{-7}$ ，表明这三种方法都可获得无偏的预测值。

表4.8 三种方法预测十亲本双列杂交 (bio-model) 非平衡数据的遗传效应的方差(Var.)和距离(Diatance)的比较($\sigma_n^2 = 50, \sigma_t^2 = 30$)

Effect	BLUP $\hat{e}_{u(\theta)}$		AUP $\hat{e}_{u(l)}$		“BLUP” $\hat{e}_{u(\hat{\theta})}$	
	Var.	Distance	Var.	Distance	Var.	Distance
$\sigma_m^2 = 20, \sigma_p^2 = 20$						
n	36.7	12.6	50.9	13.4	39.9	13.9
t	18.7	18.3	30.8	19.4	19.6	18.8
m	10.7	10.2	22.7	12.2	16.4	11.6
p	10.2	10.0	19.4	11.6	13.2	11.4
$\sigma_m^2 = 20, \sigma_p^2 = 0$						
n	43.5	10.3	52.1	12.1	40.7	11.6
t	18.9	18.2	30.9	19.4	19.6	18.7
m	16.9	6.9	20.5	9.6	13.4	8.3
p	0.0	0.0	5.0	4.2	3.0	3.0
$\sigma_m^2 = 0, \sigma_p^2 = 0$						
n	44.2	10.0	52.3	10.3	43.7	10.5
t	18.9	18.2	30.8	19.3	19.5	18.6
m	0.0	0.0	2.27	2.93	0.60	1.43
p	0.0	0.0	2.03	2.62	1.08	3.48

表4.8进一步比较了BLUP、AUP和“BLUP”三种方法的预测值方差(Var.)及距离(Dis.)。对于所有的遗传效应，用BLUP值 $\hat{e}_{u(\theta)}$ 或“BLUP”值 $\hat{e}_{u(\hat{\theta})}$ 计算的方差均低于其参数值。这表明，遗传效应的BLUP(或“BLUP”)绝对值比实际遗传效应的绝对值小。由表4.8可见，采用调整无偏预测法所获得的遗传效应AUP值，其方差是模型方差参数的无偏估计，因而遗传效应的AUP绝对值与实际遗传效应的绝对值非常接近。

遗传效应的预测值与实际值之间的距离是度量预测有效性的最重要指标，距

离越小，表明预测值越接近实际值。在所有线性无偏预测中，BLUP值的距离应该是最小的。在本试验中各项遗传效应AUP值的距离都非常接近(略大于)BLUP值的距离。这表明遗传效应AUP值与BLUP值几乎是同等有效的。在实际应用时，由于方差参数是未知的，因而BLUP值是无法算得的，只能获得所谓的“BLUP”。

“BLUP”值的计算需要大量的迭代运算，而在估算遗传方差分量的同时，不需太多的计算便可获得AUP值。AUP值虽然不是遗传效应的线性预测，却是无偏和有效的预测。由此可以认为，AUP法比“BLUP”更适用于预测混合线性模型的随机效应。

第五章 加性-显性遗传模型分析

第一节 简单加性-显性遗传模型

对总遗传效应(G)最简单的分解,是把它分解为加性遗传效应(A)和显性遗传效应(D)两个分量。因此总遗传方差(V_G)也可以分解为加性遗传方差分量(V_A)和显性遗传方差分量(V_D)两部分。表现型方差(V_P)的分量组成是 $V_P = V_G + V_e = V_A + V_D + V_e$ 。采用数量遗传分析的方法对遗传设计的方差组成进行分析,可以估算加性和显性遗传方差分量。在动、植物遗传分析中最常用的遗传设计有巢式交配设计(nested mating design)、析因交配设计(factorial mating design)和双列杂交设计(diallel mating design) (Hallauer 和 Miranda, 1981)。

1. 巢式交配设计的遗传分析方法

巢式交配设计又称北卡罗来纳设计(NC design),适用于雌性个体只能与一个雄性个体交配产生若干后代,但一个雄性个体却可以与若干雌性个体交配。如果有 s 个雄性亲本(父本),每个雄性亲本与其 d 个不同的雌性亲本(母本)交配。第 i 个父本与其第 j 个母本交配产生了 r 个后代,个体 k 的表现型值可以用以下线性模型表示,

$$y_{ijk} = \mu + s_i + d_{ij} + w_{ijk} \quad (5.1)$$

其中 μ 是群体平均数,固定效应;

s_i 是父本效应, $s_i \sim N(0, \sigma_s^2)$;

d_{ij} 是父本内的母本效应, $d_{ij} \sim N(0, \sigma_{d(s)}^2)$;

w_{ijk} 是家系内的个体效应, $w_{ijk} \sim N(0, \sigma_w^2)$ 。

因而个体的性状表现型值也具有正态分布

$$y_{ijk} \sim N(\mu, \sigma_s^2 + \sigma_{d(s)}^2 + \sigma_w^2)$$

用巢式试验设计的方差分析方法,可以估算以上三项方差分量。方差分析的变异来源、自由度、平方和、均方和期望均方列于表5.1。

表5.1 巢式交配设计的方差分析表

变异来源	自由度	平方和	均方	期望均方
父本间	$s - 1$	SSS	MSS	$\sigma_w^2 + r\sigma_{d(s)}^2 + rd\sigma_s^2$
母本间	$s(d - 1)$	SSD	MSD	$\sigma_w^2 + r\sigma_{d(s)}^2$
个体间	$sd(r - 1)$	SSW	MSW	σ_w^2

表5.1中的各项平方和可以按下式计算:

$$\text{父本间平方和 SSS} = rd \sum_{i=1}^s (\bar{y}_{i..} - \bar{y}_{...})^2$$

$$\text{父本内母本间平方和 SSD} = r \sum_{i=1}^s \sum_{j=1}^d (\bar{y}_{ij.} - \bar{y}_{i..})^2$$

$$\text{家系内个体间平方和 SSW} = \sum_{i=1}^s \sum_{j=1}^d \sum_{k=1}^r (y_{ijk} - \bar{y}_{ij.})^2$$

其中 $\bar{y}_{...}$ 是试验的总平均数, $\bar{y}_{i..}$ 是第 i 个父本的所有后代平均数, $\bar{y}_{ij.}$ 是与第 i 个父本交配的第 j 个母本的后代平均数。

将计算出的各项均方等于其期望均方的估计值, 解以下三项等式,

$$\text{MSS} = \text{SSs}/(s - 1) = \hat{\sigma}_w^2 + r\hat{\sigma}_{d(s)}^2 + rd\hat{\sigma}_s^2$$

$$\text{MSD} = \text{SSd}/[s(d - 1)] = \hat{\sigma}_w^2 + r\hat{\sigma}_{d(s)}^2$$

$$\text{MSW} = \text{SSw}/[sd(r - 1)] = \hat{\sigma}_w^2$$

便可以求得方差分量的无偏估计值 $\hat{\sigma}_w^2 = \text{MSW}$, $\hat{\sigma}_{d(s)}^2 = (\text{MSD} - \text{MSW})/r$,

$$\hat{\sigma}_s^2 = (\text{MSS} - \text{MSD})/rd.$$

以上所求得的是巢式交配试验设计的方差分量估计值, 还不是所需要的遗传方差分量估计值。通过分析同胞间协方差的遗传组成, 可以间接地估算遗传方差分量。在巢式交配设计中, 同一家系内的个体是全同胞, 而同一父系内不同家系间的个体是半同胞。如果不存在上位性效应及母体效应, 根据线性模型(5.1), 半同

胞之间的遗传协方差组成是

$$\begin{aligned}\text{Cov}(y_{ijk}, y_{ij'k'}) &= \text{Cov}(\mu + s_i + d_{ij} + w_{ijk}, \mu + s_i + d_{ij'} + w_{ij'k'}) \\ &= \text{Cov}(s_i, s_i) \\ &= \sigma_s^2 \\ &= C_{HS} \\ &= \frac{1}{4}(1+F)V_A\end{aligned}$$

全同胞之间的遗传协方差组成为

$$\begin{aligned}\text{Cov}(y_{ijk}, y_{ij'k'}) &= \text{Cov}(\mu + s_i + d_{ij} + w_{ijk}, \mu + s_i + d_{ij'} + w_{ij'k'}) \\ &= \text{Cov}(s_i + d_{ij}, s_i + d_{ij'}) \\ &= \sigma_s^2 + \sigma_{d(s)}^2 \\ &= C_{FS} \\ &= \frac{1}{2}(1+F)V_A + \frac{1}{4}(1+F)^2V_D\end{aligned}$$

其中F是父母本的近交系数。因此可以按下式估算遗传方差分量，

$$\text{加性方差分量 } V_A = 4\sigma_s^2 / (1+F) = 4(\text{MSS} - \text{MSD}) / [rd(1+F)]$$

$$\text{显性方差分量 } V_D = 4(\hat{\sigma}_{d(s)}^2 - \hat{\sigma}_s^2) / (1+F)^2$$

$$= 4[(d+1)\text{MSD} - \text{MSS} - d\text{MSW}] / [rd(1+F)^2]$$

近交系数F的取值决定于巢式交配设计中父本和母本遗传材料的近交情况。如果亲

本是随机交配的动物或异花授粉的植物，近交系数F常取值0。于是 $\hat{V}_A = 4\hat{\sigma}_s^2$ ，

$\hat{V}_D = 4(\hat{\sigma}_{d(s)}^2 - \hat{\sigma}_s^2)$ 。自花授粉植物的纯系材料作亲本时，则取F值为1，因而

$$\hat{V}_A = 2\hat{\sigma}_s^2, \quad \hat{V}_D = \hat{\sigma}_{d(s)}^2 - \hat{\sigma}_s^2。$$

如果在巢式试验设计中研究了多对性状，可以用协方差分析的方法估算成对性状的遗传协方差分量。现将性状X和性状Y协方差分析的变异来源、自由度、叉积和、均积和期望均积列于表5.2。

表5.2 巢式交配设计的协方差分析表

变异来源	自由度	叉积和	均积	期望均积
父本间	$s - 1$	SCPS	MCPS	$\sigma_w(XY) + r\sigma_{d(s)}(XY) + rd\sigma_{s(XY)}$
母本间	$s(d - 1)$	SCPD	MCPD	$\sigma_w(XY) + r\sigma_{d(s)}(XY)$
个体间	$sd(r - 1)$	SCPW	MCPW	$\sigma_w(XY)$

可按下式计算表5.2中的各项叉积和:

$$\text{父本间叉积和 SCPS} = rd \sum_{i=1}^s (\bar{y}_{i..(X)} - \bar{y}_{... (X)}) (\bar{y}_{i..(Y)} - \bar{y}_{... (Y)})$$

$$\text{父本内母本间叉积和 SCPD} = r \sum_{i=1}^s \sum_{j=1}^d (\bar{y}_{ij.(X)} - \bar{y}_{i..(X)}) (\bar{y}_{ij.(Y)} - \bar{y}_{i..(Y)})$$

$$\text{家系内个体间叉积和 SCPW} = \sum_{i=1}^s \sum_{j=1}^d \sum_{k=1}^r (y_{ijk(X)} - \bar{y}_{ij.(X)}) (y_{ijk(Y)} - \bar{y}_{ij.(Y)})$$

其中 $\bar{y}_{... (X)}$ 和 $\bar{y}_{... (Y)}$ 分别是性状X和性状Y的总平均数, $\bar{y}_{i..(X)}$ 和 $\bar{y}_{i..(Y)}$ 分别是第*i*个父本后代性状X和性状Y的平均数, $\bar{y}_{ij.(X)}$ 和 $\bar{y}_{ij.(Y)}$ 分别是第*i*个父本与第*j*个母本交配的后代性状X和性状Y的平均数。

如将计算出的各项均积等于它们期望均积的估计值,

$$\text{MCPS} = \text{SCPS} / (s - 1) = \hat{\sigma}_w(XY) + r\hat{\sigma}_{d(s)}(XY) + rd\hat{\sigma}_s(XY)$$

$$\text{MCPD} = \text{SCPD} / [s(d - 1)] = \hat{\sigma}_w(XY) + r\hat{\sigma}_{d(s)}(XY)$$

$$\text{MCPW} = \text{SCPW} / [sd(r - 1)] = \hat{\sigma}_w(XY)$$

然后解以上等式便可以求得协方差分量的无偏估计值 $\hat{\sigma}_w(XY) = \text{MCPW}$, $\hat{\sigma}_{d(s)}(XY)$

$= (\text{MCPD} - \text{MCPW}) / r$, $\hat{\sigma}_s(XY) = (\text{MCPS} - \text{MCPD}) / rd$ 。用以上估计值和已知的亲

本近交系数, 可以进一步估算遗传协方差分量。加性协方差分量 \hat{C}_A

$= 4\hat{\sigma}_{s(XY)} / (1+F) = 4(\text{MCPS} - \text{MCPD}) / [rd(1+F)]$, 显性协方差分量 $\hat{C}_D =$

$$4(\hat{\sigma}_{d(s)(XY)} - \hat{\sigma}_{s(XY)}) / (1+F)^2 = 4[(d+1)MCPD - MCPS - dMCPW] / [rd(1+F)^2]$$

采用混合线性模型的分析方法也可以无偏地估算巢式交配设计的各项方差分量和协方差分量。如果用混合线性模型矩阵形式表示巢式交配设计的遗传模型，则有

$$\begin{aligned} \mathbf{y} &= \mathbf{1}\mu + \mathbf{U}_s \mathbf{e}_s + \mathbf{U}_{d(s)} \mathbf{e}_{d(s)} + \mathbf{e}_w \\ &= \mathbf{1}\mu + \sum_{u=1}^3 \mathbf{U}_u \mathbf{e}_u \\ &\sim (\mathbf{1}\mu, \mathbf{V} = \sum_{u=1}^3 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T) \end{aligned}$$

其中 \mathbf{e}_s 是父本效应向量， $\mathbf{e}_s \sim (0, \sigma_s^2 \mathbf{I})$ ， \mathbf{U}_s 是父本效应的系数矩阵； $\mathbf{e}_{d(s)}$ 是父本内的母本效应向量， $\mathbf{e}_{d(s)} \sim (0, \sigma_{d(s)}^2 \mathbf{I})$ ， $\mathbf{U}_{d(s)}$ 是父本内母本效应的系数矩阵； \mathbf{e}_w 是剩余效应向量， $\mathbf{e}_w \sim (0, \sigma_w^2 \mathbf{I})$ 。

采用MINQUE(1)法，可以无偏估算交配设计的各项方差分量和协方差分量，从而可以进一步估算遗传方差分量和遗传协方差分量。

遗传试验的表现型值 \mathbf{y} 具有方差-协方差矩阵

$$\text{Var}(\mathbf{y}) = \mathbf{V} = \sigma_s^2 \mathbf{U}_s \mathbf{U}_s^T + \sigma_{d(s)}^2 \mathbf{U}_{d(s)} \mathbf{U}_{d(s)}^T + \sigma_w^2 \mathbf{I} = \sum_{u=1}^3 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T$$

方差分量的无偏估计值可解以下MINQUE(1)方程组求得，

$$\begin{aligned} \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_s^2 + \text{tr}(\mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)} \mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)}) \hat{\sigma}_{d(s)}^2 + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_w^2 \\ = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{y} \\ \text{tr}(\mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)}) \hat{\sigma}_s^2 + \text{tr}(\mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)} \mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)}) \hat{\sigma}_{d(s)}^2 \\ + \text{tr}(\mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_{d(s)}) \hat{\sigma}_w^2 = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)} \mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{y} \\ \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)}) \hat{\sigma}_s^2 + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_{d(s)} \mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)}) \hat{\sigma}_{d(s)}^2 + \text{tr}(\mathbf{Q}_{(1)} \mathbf{Q}_{(1)}) \hat{\sigma}_w^2 \\ = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{y} \end{aligned}$$

成对性状 y_1 和 y_2 之间协方差分量的无偏估计可由以下MINQUE(1)方程组解得，

$$\begin{aligned} & \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{s/s} + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)} \mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{d(s)/d(s)} + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{w/w} \\ & \quad = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{y}_2 \\ & \text{tr}(\mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)}) \hat{\sigma}_{s/s} + \text{tr}(\mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)} \mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)}) \hat{\sigma}_{d(s)/d(s)} \\ & \quad + \text{tr}(\mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_{d(s)}) \hat{\sigma}_{w/w} = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{U}_{d(s)} \mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)} \mathbf{y}_2 \\ & \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)}) \hat{\sigma}_{s/s} + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_{d(s)} \mathbf{U}_{d(s)}^T \mathbf{Q}_{(1)}) \hat{\sigma}_{d(s)/d(s)} + \text{tr}(\mathbf{Q}_{(1)} \mathbf{Q}_{(1)}) \hat{\sigma}_{w/w} \\ & \quad = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{y}_2 \end{aligned}$$

2. 析因交配设计的遗传分析方法

有些植物雌性个体上有许多花器，因而可以同时与若干雄性交配，在同一植株上产生不同组合的后代。这样的遗传材料适用于采用析因交配设计的分析方法，估算遗传方差分量和遗传协方差分量。析因交配设计又称北卡罗来纳设计 (NC design)。如果有 s 个雄性亲本(父本)与 d 个雌性亲本(母本)交配，每个组合产生了 r 个后代，第 i 个父本与第 j 个母本杂交组合中个体 k 的表现型值可以用以下线性模型表示，

$$y_{ijk} = \mu + s_i + d_j + sd_{ij} + w_{ijk} \quad (5.2)$$

其中 μ 是群体平均数，固定效应；

s_i 是父本效应， $s_i \sim N(0, \sigma_s^2)$ ；

d_j 是母本效应， $d_j \sim N(0, \sigma_d^2)$ ；

sd_{ij} 是父母本互作效应， $sd_{ij} \sim N(0, \sigma_{sd}^2)$ ；

w_{ijk} 是组合内的个体效应， $w_{ijk} \sim N(0, \sigma_w^2)$ 。

因而个体的性状表现型值具有正态分布

$$y_{ijk} \sim N(\mu, \sigma_s^2 + \sigma_d^2 + \sigma_{sd}^2 + \sigma_w^2)。$$

析因交配设计的方差分析表列于表5.3。

表5.3 析因交配设计的方差分析表

变异来源	自由度	平方和	均方	期望均方
父本间	$s - 1$	SSS	MSS	$\sigma_w^2 + r\sigma_{sd}^2 + rd\sigma_s^2$
母本间	$d - 1$	SSD	MSD	$\sigma_w^2 + r\sigma_{sd}^2 + rs\sigma_d^2$
父 × 母互作	$(s - 1)(d - 1)$	SSSD	MSSD	$\sigma_w^2 + r\sigma_{sd}^2$
个体间	$sd(r - 1)$	SSW	MSW	σ_w^2

表5.3中的各项平方和可以按下式计算:

$$\text{父本间平方和 SSS} = rd \sum_{i=1}^s (\bar{y}_{i..} - \bar{y}_{...})^2$$

$$\text{母本间平方和 SSD} = rs \sum_{j=1}^d (\bar{y}_{.j.} - \bar{y}_{...})^2$$

$$\text{父} \times \text{母互作平方和 SSSD} = r \sum_{i=1}^s \sum_{j=1}^d (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$$

$$\text{组合内个体间平方和 SSW} = \sum_{i=1}^s \sum_{j=1}^d \sum_{k=1}^r (y_{ijk} - \bar{y}_{ij.})^2$$

其中 $\bar{y}_{...}$ 是试验的总平均数, $\bar{y}_{i..}$ 是父本 i 的所有后代平均数, $\bar{y}_{.j.}$ 是母本 j 的所有后代平均数, $\bar{y}_{ij.}$ 父本 i 与母本 j 交配的后代平均数.

析因交配设计中的各项方差分量可由解以下四项等式而求得,

$$MSS = SSS / (s - 1) = \hat{\sigma}_w^2 + r\hat{\sigma}_{sd}^2 + rd\hat{\sigma}_s^2$$

$$MSD = SSD / (d - 1) = \hat{\sigma}_w^2 + r\hat{\sigma}_{sd}^2 + rs\hat{\sigma}_d^2$$

$$MSSD = SSSD / [(s - 1)(d - 1)] = \hat{\sigma}_w^2 + r\hat{\sigma}_{sd}^2$$

$$MSW = SSW / [sd(r - 1)] = \hat{\sigma}_w^2$$

方差分量的无偏估计值为 $\hat{\sigma}_w^2 = MSW$, $\hat{\sigma}_{sd}^2 = (MSSD - MSW) / r$, $\hat{\sigma}_d^2 = (MSD - MSSD) / rs$, $\hat{\sigma}_s^2 = (MSS - MSSD) / rd$.

由半同胞和全同胞协方差的遗传组成，可以推导出以下遗传方差分量的估算公式，

$$\hat{V}_A = 2(\hat{\sigma}_s^2 + \hat{\sigma}_d^2)/(1+F), \quad \hat{V}_D = 4\hat{\sigma}_{sd}^2/(1+F)^2,$$

近交系数F的取值决定于析因交配设计中父本和母本遗传材料的近交情况。如果亲本是随机交配的遗传群体，近交系数F常取值0，于是加性方差分量

$$\hat{V}_A = 2(\hat{\sigma}_s^2 + \hat{\sigma}_d^2), \quad \text{显性方差分量 } \hat{V}_D = 4\hat{\sigma}_{sd}^2。$$

自交的纯系材料作亲本时，则取F

值为1，因而加性方差分量 $\hat{V}_A = \hat{\sigma}_s^2 + \hat{\sigma}_d^2$ ，显性方差分量 $\hat{V}_D = \hat{\sigma}_{sd}^2$ 。

可用协方差分析的方法估算析因交配设计成对性状之间的遗传协方差分量。

现将性状X和性状Y的协方差分析表列于表5.4。

表5.4 析因交配设计的协方差分析表

变异来源	自由度	叉积和	均积	期望均积
父本间	$s-1$	SCPS	MCPS	$\sigma_w(XY) + r\sigma_{sd}(XY) + rd\sigma_s(XY)$
母本间	$d-1$	SCPD	MCPD	$\sigma_w(XY) + r\sigma_{sd}(XY) + rs\sigma_d(XY)$
父 × 母互作	$(s-1)(d-1)$	SCPSD	MCPSD	$\sigma_w(XY) + r\sigma_{sd}(XY)$
个体间	$sd(r-1)$	SCPW	MCPW	$\sigma_w(XY)$

表5.4中的各项叉积和可以按下式计算：

$$\text{父本间叉积和 SCPS} = rd \sum_{i=1}^s (\bar{y}_{i.(X)} - \bar{y}_{...(X)}) (\bar{y}_{i..(Y)} - \bar{y}_{...(Y)})$$

$$\text{母本间叉积和 SCPD} = rs \sum_{j=1}^d (\bar{y}_{.j.(X)} - \bar{y}_{...(X)}) (\bar{y}_{.j..(Y)} - \bar{y}_{...(Y)})$$

父 × 母互作叉积和 SCPSD =

$$r \sum_{i=1}^s \sum_{j=1}^d (\bar{y}_{ij.(X)} - \bar{y}_{i..(X)} - \bar{y}_{.j.(X)} + \bar{y}_{...(X)}) (\bar{y}_{ij..(Y)} - \bar{y}_{i..(Y)} - \bar{y}_{.j..(Y)} + \bar{y}_{...(Y)})$$

$$\text{组合内个体间叉积和 SCPW} = \sum_{i=1}^s \sum_{j=1}^d \sum_{k=1}^r (y_{ijk(X)} - \bar{y}_{ij.(X)}) (y_{ijk(Y)} - \bar{y}_{ij..(Y)})$$

其中 $\bar{y}_{...(X)}$ 和 $\bar{y}_{...(Y)}$ 分别是性状X和性状Y的总平均数， $\bar{y}_{i..(X)}$ 和 $\bar{y}_{i..(Y)}$ 分别是父本/所有后代性状X和性状Y的平均数， $\bar{y}_{.j.(X)}$ 和 $\bar{y}_{.j..(Y)}$ 分别是母本/所有后代性状X

和性状y的平均数， $\bar{y}_{ij.(X)}$ 和 $\bar{y}_{ij.(Y)}$ 分别是父本i与母本j交配后代性状X和性状Y的平均数。

将计算出的各项叉积和除以其自由度得到均积，然后再估算析因交配设计的协方差分量，

$$\hat{\sigma}_{w(XY)} = \text{MCPW}$$

$$\hat{\sigma}_{sd(XY)} = (\text{MCPsd} - \text{MCPW})/r$$

$$\hat{\sigma}_{d(XY)} = (\text{MCPd} - \text{MCPsd})/rs$$

$$\hat{\sigma}_{s(XY)} = (\text{MCPs} - \text{MCPsd})/rd。$$

可以按以下公式估算遗传协方差分量，

$$\text{加性协方差分量 } \hat{C}_A = 2(\hat{\sigma}_{s(XY)} + \hat{\sigma}_{d(XY)})/(1+F),$$

$$\text{显性协方差分量 } \hat{C}_D = 4\hat{\sigma}_{sd(XY)}/(1+F)^2。$$

采用混合线性模型的分析方法，也可以无偏地估算析因交配设计的各项方差分量和协方差分量。如果用混合线性模型矩阵形式表示析因交配设计的遗传模型，则有

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{U}_s\mathbf{e}_s + \mathbf{U}_d\mathbf{e}_d + \mathbf{U}_{sd}\mathbf{e}_{sd} + \mathbf{e}_w$$

$$= \mathbf{1}\mu + \sum_{u=1}^4 \mathbf{U}_u\mathbf{e}_u$$

$$\sim (\mathbf{1}\mu, \mathbf{V} = \sum_{u=1}^4 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T)$$

其中 \mathbf{e}_s 是父本效应向量， $\mathbf{e}_s \sim (0, \sigma_s^2 \mathbf{I})$ ， \mathbf{U}_s 是父本效应的系数矩阵； \mathbf{e}_d 是母本效应向量， $\mathbf{e}_d \sim (0, \sigma_d^2 \mathbf{I})$ ， \mathbf{U}_d 母本效应的系数矩阵； \mathbf{e}_{sd} 是父母本互作效应向量， $\mathbf{e}_{sd} \sim (0, \sigma_{sd}^2 \mathbf{I})$ ， \mathbf{U}_{sd} 是父母本互作效应的系数矩阵； \mathbf{e}_w 是剩余效应向量， $\mathbf{e}_w \sim (0, \sigma_w^2 \mathbf{I})$ 。

采用MINQUE(1)法，可以无偏估算析因交配设计的各项方差分量和协方差分量，从而可以进一步估算遗传方差分量和遗传协方差分量。

遗传试验的表现型值 \mathbf{y} 具有方差-协方差矩阵

$$\text{Var}(\mathbf{y}) = \mathbf{V} = \sigma_s^2 \mathbf{U}_s \mathbf{U}_s^T + \sigma_d^2 \mathbf{U}_d \mathbf{U}_d^T + \sigma_{sd}^2 \mathbf{U}_{sd} \mathbf{U}_{sd}^T + \sigma_w^2 \mathbf{I} = \sum_{u=1}^4 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T$$

方差分量的无偏估计值可解以下MINQUE(1)方程组求得，

$$\text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_s^2 + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_d \mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_d^2 + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_{sd} \mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{sd}^2 + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_w^2 = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{y}$$

$$\begin{aligned} & \text{tr}(\mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_d) \hat{\sigma}_s^2 + \text{tr}(\mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_d \mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_d) \hat{\sigma}_d^2 \\ & + \text{tr}(\mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_{sd} \mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_d) \hat{\sigma}_{sd}^2 + \text{tr}(\mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_d) \hat{\sigma}_w^2 = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_d \mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{y} \\ & \text{tr}(\mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_{sd}) \hat{\sigma}_s^2 + \text{tr}(\mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_d \mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_{sd}) \hat{\sigma}_d^2 \\ & + \text{tr}(\mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_{sd} \mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_{sd}) \hat{\sigma}_{sd}^2 + \text{tr}(\mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_{sd}) \hat{\sigma}_w^2 = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_{sd} \mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{y} \\ & \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)}) \hat{\sigma}_s^2 + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_d \mathbf{U}_d^T \mathbf{Q}_{(1)}) \hat{\sigma}_d^2 + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_{sd} \mathbf{U}_{sd}^T \mathbf{Q}_{(1)}) \hat{\sigma}_{sd}^2 \\ & + \text{tr}(\mathbf{Q}_{(1)} \mathbf{Q}_{(1)}) \hat{\sigma}_w^2 = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{y} \end{aligned}$$

成对性状 y_1 和 y_2 之间协方差分量的无偏估计可由以下MINQUE(1)方程组解得，

$$\begin{aligned} & \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{s/s} + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_d \mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{d/d} + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_{sd} \mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{sd/sd} \\ & + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{w/w} = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{y}_2 \\ & \text{tr}(\mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_d) \hat{\sigma}_{s/s} + \text{tr}(\mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_d \mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_d) \hat{\sigma}_{d/d} \\ & + \text{tr}(\mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_{sd} \mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_d) \hat{\sigma}_{sd/sd} + \text{tr}(\mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_d) \hat{\sigma}_{w/w} = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{U}_d \mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{y}_2 \\ & \text{tr}(\mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_{sd}) \hat{\sigma}_{s/s} + \text{tr}(\mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_d \mathbf{U}_d^T \mathbf{Q}_{(1)} \mathbf{U}_{sd}) \hat{\sigma}_{d/d} \\ & + \text{tr}(\mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_{sd} \mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{U}_{sd}) \hat{\sigma}_{sd/sd} + \text{tr}(\mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_{sd}) \hat{\sigma}_{w/w} = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{U}_{sd} \mathbf{U}_{sd}^T \mathbf{Q}_{(1)} \mathbf{y}_2 \\ & \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)}) \hat{\sigma}_{s/s} + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_d \mathbf{U}_d^T \mathbf{Q}_{(1)}) \hat{\sigma}_{d/d} \\ & + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_{sd} \mathbf{U}_{sd}^T \mathbf{Q}_{(1)}) \hat{\sigma}_{sd/sd} + \text{tr}(\mathbf{Q}_{(1)} \mathbf{Q}_{(1)}) \hat{\sigma}_{w/w} = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{y}_2 \end{aligned}$$

3. 双列杂交设计的遗传分析方法

如果遗传材料既可作为父本又可作为母本参与杂交，则可采用双列杂交设计估算遗传方差分量和遗传协方差分量。Gri ffi ng(1956)介绍了四种双列交配方法。

当母体效应不存在时，可选用方法4，即 ρ 个亲本之间作全部 $\rho(\rho - 1)/2$ 个正交组合。

如果双列杂交试验用 ρ 个亲本做了 $(\rho(\rho - 1)/2)$ 个正交杂交组合，有 b 个完全随机区组，母本 i 与父本 $j(j > i)$ 的杂交组合在第 k 个区组中的平均表现型值可用以下线性模型表示，

$$y_{ijk} = \mu + g_i + g_j + s_{ij} + b_k + e_{ijk} \quad (5.3)$$

其中 μ 是群体平均数；

g_i 或 g_j 是一般配合力效应， g_i 或 $g_j \sim N(0, \sigma_g^2)$ ；

s_{ij} 是特殊配合力效应， $s_{ij} \sim N(0, \sigma_s^2)$ ；

b_k 是区组效应， $b_k \sim N(0, \sigma_b^2)$ ；

e_{ijk} 是剩余效应， $e_{ijk} \sim N(0, \sigma_e^2)$ 。

在这个线性模型中，群体平均数 μ 是固定效应，其它效应都是随机效应。双列杂交组合的表现型方差 $\text{Var}(y_{ijk})$ 为

$$\sigma_y^2 = 2\sigma_g^2 + \sigma_s^2 + \sigma_b^2 + \sigma_e^2$$

一般配合力方差和特殊配合力方差可以用方差分析的方法估算。

5.5 双列杂交设计(方法4)的方差分析表

变异来源	自由度	平方和	均方	期望均方
区组间	$b - 1$	SSB	MSB	
一般配合力	$\rho - 1$	SSG	MSG	$\sigma_e^2 + b\sigma_s^2 + b(\rho - 2)\sigma_g^2$
特殊配合力	$\rho(\rho - 3)/2$	SSS	MSS	$\sigma_e^2 + b\sigma_s^2$
机误	df_e	SSe	MSe	σ_e^2

注: $df_e = (b - 1)[\rho(\rho - 1)/2 - 1]$ 。

如果设 $S1 = [b(\rho - 1)^2/(\rho - 2)] \sum_{i=1}^{\rho} (\bar{y}_{i..})^2$ ，各项平方和可按下式计算：

区组间平方和 $SSB = [\rho(\rho - 1)/2] \sum_{k=1}^b (\bar{y}_{..k} - \bar{y}_{...})^2$

一般配合力平方和 $SSG = S1 - [b\rho(\rho - 1)^2/(\rho - 2)] (\bar{y}_{...})^2$

$$\text{特殊配合力平方和 } SSS = b \sum_{i=1}^{p-1} \sum_{j>i}^p (\bar{y}_{ij.})^2 - S1$$

$$+ \{b\rho^2(\rho - 1)/[2(\rho - 2)]\}(\bar{y}_{...})^2$$

$$\text{机误平方和 } SSe = \sum_{i=1}^{p-1} \sum_{j>i}^p \sum_{k=1}^b (y_{ijk})^2 - SSG - SSS - SSB$$

其中 $\bar{y}_{...}$ 是试验的总平均数, $\bar{y}_{..k}$ 是第 k 个区组的所有杂交组合平均数, $\bar{y}_{i..}$ 是以第 i 个亲本作母本或父本的所有组合平均数, $\bar{y}_{ij.}$ 是组合 $(i \times j)$ 的区组平均数。

双列杂交设计中的各项方差分量可由解以下三项等式而求得,

$$MSG = SSG / (\rho - 1) = \hat{\sigma}_e^2 + b\hat{\sigma}_s^2 + b(\rho - 2)\hat{\sigma}_g^2$$

$$MSS = 2SSS / [\rho(\rho - 3)] = \hat{\sigma}_e^2 + b\hat{\sigma}_s^2$$

$$MSe = SSe / df_e = \hat{\sigma}_e^2$$

方差分量的无偏估计值为, $\hat{\sigma}_e^2 = MSe$, $\hat{\sigma}_s^2 = (MSS - MSe) / b$, $\hat{\sigma}_g^2 = (MSG - MSS) / [b(\rho - 2)]$ 。

如果不存在上位性效应, 可由半同胞和全同胞协方差的遗传组成推导出遗传方差分量的估算公式, 加性方差分量 $\hat{V}_A = 4\hat{\sigma}_g^2 / (1 + F)$, 显性方差分量

$\hat{V}_D = 4\hat{\sigma}_s^2 / (1 + F)$ 。在实际应用时, 通常选用基因型纯合的材料作为双列杂交的亲本, 其遗传方差分量可按下式分别直接估算, $\hat{V}_A = 2\hat{\sigma}_g^2$, $\hat{V}_D = \hat{\sigma}_s^2$ 。

也可采用协方差分析的方法, 估算双列杂交设计成对性状 X 和 Y 之间的遗传协方差分量。

5.6 双列杂交设计(方法4)的协方差分析表

变异来源	自由度	叉积和	均积	期望均积
区组间	$b-1$	SCP _B	MCP _B	
一般配合力	$\rho-1$	SCP _G	MCP _G	$\sigma_e(XY) + b\sigma_s(XY) + b(\rho-2)\sigma_g(XY)$
特殊配合力	$\rho(\rho-3)/2$	SCP _S	MCPS	$\sigma_e(XY) + b\sigma_s(XY)$
机误	df_e	SCP _e	MCP _e	$\sigma_e(XY)$

注: $df_e = (b - 1)[\rho(\rho - 1) / 2 - 1]$ 。

$$\text{设 } C1 = [b(\rho - 1)^2 / (\rho - 2)] \sum_{i=1}^p (\bar{y}_{i..(X)}) (\bar{y}_{i..(Y)})$$

$$C2 = b \sum_{i=1}^{p-1} \sum_{j>i}^p (\bar{y}_{ij..(X)}) (\bar{y}_{ij..(Y)})$$

表5.6中的各项叉积和可按下式计算:

$$\text{区组间叉积和SCPB} = [\rho(\rho - 1) / 2] \sum_{k=1}^b (\bar{y}_{..k(X)} - \bar{y}_{... (X)}) (\bar{y}_{..k(Y)} - \bar{y}_{... (Y)})$$

$$\text{一般配合力叉积和SCPG} = C1 - [b\rho(\rho - 1)^2 / (\rho - 2)] (\bar{y}_{... (X)}) (\bar{y}_{... (Y)})$$

$$\text{特殊配合力叉积和SCPS} = C2 - C1$$

$$+ b\rho^2(\rho - 1) / [2(\rho - 2)] (\bar{y}_{... (X)}) (\bar{y}_{... (Y)})$$

$$\text{机误叉积和SCPe} = \sum_{i=1}^{p-1} \sum_{j>i}^p \sum_{k=1}^b (y_{ijk(X)}) (y_{ijk(Y)}) - \text{SSG} - \text{SSS} - \text{SSB}$$

解以下三项等式可求得双列杂交设计中的各项协方差分量，

$$\text{MCPG} = \text{SCPG} / (\rho - 1) = \hat{\sigma}_{e(XY)} + b\hat{\sigma}_{s(XY)} + b(\rho - 2)\hat{\sigma}_{g(XY)}$$

$$\text{MCPS} = 2\text{SCPS} / [\rho(\rho - 3)] = \hat{\sigma}_{e(XY)} + b\hat{\sigma}_{s(XY)}$$

$$\text{MCPe} = \text{SCPe} / \text{df}_e = \hat{\sigma}_{e(XY)}$$

协方差分量的估计值为 $\hat{\sigma}_{e(XY)} = \text{MCPW}$ ， $\hat{\sigma}_{s(XY)} = (\text{MCPS} - \text{MCPW}) / b$ ，

$$\hat{\sigma}_{g(XY)} = (\text{MCPg} - \text{MCPS}) / [b(\rho - 2)]。$$

遗传协方差分量的估算公式是，加性协方差分量 $\hat{C}_A = 4\sigma_{g(XY)} / (1 + F)$ ，显性协方差分量 $\hat{C}_D = 4\sigma_{s(XY)} / (1 + F)^2$ 。双列杂交的亲本通常都是纯系材料，近交系数F取值为0，因此 $\hat{C}_A = 2\sigma_{g(XY)}$ ， $\hat{C}_D = \sigma_{s(XY)}$ 。

以上所介绍的双列分析方法是按Cockerham(1963)的遗传方差分析模型推导的，与Griffing(1956)原文的公式略有不同。前者基于性状观察值的小区平均数为分析基础，而后者基于小区中单株观察值为分析基础。

采用混合线性模型的分析方法，可以无偏地估算双列杂交设计的各项方差分量和协方差分量。如果用混合线性模型矩阵形式表示双列杂交设计的遗传模型，则有

$$\begin{aligned} \mathbf{y} &= \mathbf{1}\mu + \mathbf{U}_g\mathbf{e}_g + \mathbf{U}_s\mathbf{e}_s + \mathbf{U}_b\mathbf{e}_b + \mathbf{e}_\varepsilon \\ &= \mathbf{1}\mu + \sum_{u=1}^4 \mathbf{U}_u\mathbf{e}_u \\ &\sim (\mathbf{1}\mu, \mathbf{V} = \sum_{u=1}^4 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T) \end{aligned}$$

其中 \mathbf{e}_g 是一般配合力效应向量， $\mathbf{e}_g \sim (\mathbf{0}, \sigma_g^2 \mathbf{I})$;

\mathbf{U}_g 是一般配合力效应的系数矩阵;

\mathbf{e}_s 是特殊配合力效应向量， $\mathbf{e}_s \sim (\mathbf{0}, \sigma_s^2 \mathbf{I})$;

\mathbf{U}_s 是特殊配合力效应的系数矩阵;

\mathbf{e}_b 是区组效应向量， $\mathbf{e}_b \sim (\mathbf{0}, \sigma_b^2 \mathbf{I})$;

\mathbf{U}_b 是区组效应的系数矩阵;

\mathbf{e}_ε 是剩余效应向量， $\mathbf{e}_\varepsilon \sim (\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I})$ 。

采用MINQUE(1)法，可以无偏估算双列杂交设计的各项方差分量和协方差分量，从而可以进一步估算遗传方差分量和遗传协方差分量。

遗传试验的表现型值 \mathbf{y} 具有方差-协方差矩阵

$$\text{Var}(\mathbf{y}) = \mathbf{V} = \sigma_g^2 \mathbf{U}_g \mathbf{U}_g^T + \sigma_s^2 \mathbf{U}_s \mathbf{U}_s^T + \sigma_b^2 \mathbf{U}_b \mathbf{U}_b^T + \sigma_\varepsilon^2 \mathbf{I} = \sum_{u=1}^4 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T$$

方差分量的无偏估计值可解以下MINQUE(1)方程组求得，

$$\begin{aligned} &\text{tr}(\mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_g \mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_g) \hat{\sigma}_g^2 + \text{tr}(\mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_g) \hat{\sigma}_s^2 + \text{tr}(\mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_b \mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_g) \hat{\sigma}_b^2 \\ &\quad + \text{tr}(\mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_g) \hat{\sigma}_\varepsilon^2 = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_g \mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{y} \\ &\text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_g \mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_g^2 + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_s^2 \\ &\quad + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_b \mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_b^2 + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_\varepsilon^2 = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{y} \end{aligned}$$

$$\begin{aligned} & \text{tr}(\mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_g \mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_b) \hat{\sigma}_g^2 + \text{tr}(\mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_b) \hat{\sigma}_s^2 \\ & + \text{tr}(\mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_b \mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_b) \hat{\sigma}_b^2 + \text{tr}(\mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_b) \hat{\sigma}_\varepsilon^2 = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_b \mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{y} \\ & \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_g \mathbf{U}_g^T \mathbf{Q}_{(1)}) \hat{\sigma}_g^2 + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)}) \hat{\sigma}_s^2 \\ & + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_b \mathbf{U}_b^T \mathbf{Q}_{(1)}) \hat{\sigma}_b^2 + \text{tr}(\mathbf{Q}_{(1)} \mathbf{Q}_{(1)}) \hat{\sigma}_\varepsilon^2 = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{y} \end{aligned}$$

成对性状 y_1 和 y_2 之间协方差分量的无偏估计可由以下MINQUE(1)方程组解得，

$$\begin{aligned} & \text{tr}(\mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_g \mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_g) \hat{\sigma}_{g/g} + \text{tr}(\mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_g) \hat{\sigma}_{s/s} + \text{tr}(\mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_b \mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_g) \hat{\sigma}_{b/b} \\ & + \text{tr}(\mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_g) \hat{\sigma}_{\varepsilon/\varepsilon} = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{U}_g \mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{y}_2 \\ & \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_g \mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{g/g} + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{s/s} \\ & + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_b \mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{b/b} + \text{tr}(\mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_s) \hat{\sigma}_{\varepsilon/\varepsilon} = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{y}_2 \\ & \text{tr}(\mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_g \mathbf{U}_g^T \mathbf{Q}_{(1)} \mathbf{U}_b) \hat{\sigma}_{g/g} + \text{tr}(\mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)} \mathbf{U}_b) \hat{\sigma}_{s/s} \\ & + \text{tr}(\mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_b \mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{U}_b) \hat{\sigma}_{b/b} + \text{tr}(\mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_b) \hat{\sigma}_{\varepsilon/\varepsilon} = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{U}_b \mathbf{U}_b^T \mathbf{Q}_{(1)} \mathbf{y}_2 \\ & \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_g \mathbf{U}_g^T \mathbf{Q}_{(1)}) \hat{\sigma}_{g/g} + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_s \mathbf{U}_s^T \mathbf{Q}_{(1)}) \hat{\sigma}_{s/s} \\ & + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_b \mathbf{U}_b^T \mathbf{Q}_{(1)}) \hat{\sigma}_{b/b} + \text{tr}(\mathbf{Q}_{(1)} \mathbf{Q}_{(1)}) \hat{\sigma}_{\varepsilon/\varepsilon} = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{y}_2 \end{aligned}$$

以上所介绍的双列杂交遗传模型及方差分析方法是通过一般配合力及特殊配合力方差分量，间接估算加性方差和显性方差分量。而且对双列杂交的组配有严格的要求。朱军等(1993)提出了运用采用MINQUE(1)法直接估算加性、显性方差分量的加性-显性遗传模型(简称AD模型)。

AD模型的遗传假设是：具有正常二倍体分裂，参与杂交的纯合亲本是遗传群体的一个随机样本，不存在母体效应及基因的上位性效应，不存在基因型与环境的互作。如果双列杂交试验有 n 个亲本及其正反杂交组合 ($n(n-1)$ 个) 和 r 个完全随机区组，母本 i 与父本 j 的杂交组合在第 k 个区组中的平均表现型值可用以下线性模型表示

$$y_{ijk} = \mu + A_i + A_j + D_{ij} + B_k + \varepsilon_{ijk}$$

其中 μ 是群体平均数；

$$A_i \text{ 或 } A_j \text{ 是累加的加性效应, } A_i \text{ 或 } A_j \sim (0, \sigma_A^2);$$

D_{ij} 是累加的显性效应, $D_{ij} \sim (0, \sigma_D^2)$;

B_k 是区组效应, $B_k \sim (0, \sigma_B^2)$;

ε_{ijk} 是剩余效应, $\varepsilon_{ijk} \sim (0, \sigma_e^2)$ 。

在这个线性模型中, 群体平均数 μ 是固定效应, 其它效应都是随机效应。双列杂交组合的表现型方差 $\text{Var}(y_{ijk})$ 为

$$\sigma_y^2 = 2\sigma_A^2 + \sigma_D^2 + \sigma_B^2 + \sigma_e^2。$$

其遗传方差分量可按下式分别直接估算, $V_A = 2\sigma_A^2$, $V_D = \sigma_D^2$ 。

如果用混合线性模型矩阵形式表示双列杂交设计的AD遗传模型, 则有

$$\begin{aligned} \mathbf{y} &= \mathbf{1}\mu + \mathbf{U}_A\mathbf{e}_A + \mathbf{U}_D\mathbf{e}_D + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_\varepsilon \\ &= \mathbf{1}\mu + \sum_{u=1}^4 \mathbf{U}_u\mathbf{e}_u \\ &\sim (\mathbf{1}\mu, \mathbf{V} = \sum_{u=1}^4 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T) \end{aligned}$$

其中 \mathbf{e}_A 是累加的加性效应向量, $\mathbf{e}_A \sim (\mathbf{0}, \sigma_A^2 \mathbf{I})$;

\mathbf{U}_A 是累加的加性效应的系数矩阵;

\mathbf{e}_D 是累加的显性效应向量, $\mathbf{e}_D \sim (\mathbf{0}, \sigma_D^2 \mathbf{I})$;

\mathbf{U}_D 是累加的显性效应的系数矩阵;

\mathbf{e}_B 是区组效应向量, $\mathbf{e}_B \sim (\mathbf{0}, \sigma_B^2 \mathbf{I})$;

\mathbf{U}_B 是区组效应的系数矩阵;

\mathbf{e}_ε 是剩余效应向量, $\mathbf{e}_\varepsilon \sim (\mathbf{0}, \sigma_e^2 \mathbf{I})$ 。

采用MINQUE法, 可以无偏地分析以上两个模型的平衡数据和有缺失的非平衡数据。

现以一个 3×3 的完全双列杂交为例, 说明如何运用MINQUE(1)法估算一个简单加性-显性遗传模型中的遗传方差分量。如果没有正反交效应, 第 i 个亲本与第 j 个亲本的后代平均观察值可以用以下线性模型表示 ($i=1, 2, 3$; $j=1, 2, 3$),

$$y_{ij} = \mu + A_i + A_j + D_{ij} + \varepsilon_{ij}$$

其中 μ 是群体平均数，固定效应

A_i (或 A_j) 是第 i (或 j) 个亲本的加性效应, A_i (或 A_j) $\sim (0, \sigma_A^2)$

D_{ij} 是亲本 $i \times$ 亲本 j 后代的显性效应, $D_{ij} \sim (0, \sigma_D^2)$

ε_{ij} 是剩余效应, $\varepsilon_{ij} \sim (0, \sigma_e^2)$

以上模型可以用矩阵形式表示

$$\begin{aligned} \mathbf{y} &= \mathbf{1}\mu + \mathbf{U}_A \mathbf{e}_A + \mathbf{U}_D \mathbf{e}_D + \mathbf{e}_\varepsilon \\ &= \mathbf{1}\mu + \sum_{u=1}^3 \mathbf{U}_u \mathbf{e}_u \end{aligned}$$

其中 \mathbf{e}_A 是加性效应向量, $\mathbf{e}_A \sim (\mathbf{0}, \sigma_A^2 \mathbf{I})$;

\mathbf{U}_A 是加性效应的系数矩阵;

\mathbf{e}_D 是显性效应向量, $\mathbf{e}_D \sim (\mathbf{0}, \sigma_D^2 \mathbf{I})$,

\mathbf{U}_D 是显性效应的系数矩阵;

\mathbf{e}_ε 是剩余效应向量, $\mathbf{e}_\varepsilon \sim (\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I})$ 。

具有方差-协方差矩阵

$$\text{Var}(\mathbf{y}) = \mathbf{V} = \sigma_A^2 \mathbf{U}_A \mathbf{U}_A^T + \sigma_D^2 \mathbf{U}_D \mathbf{U}_D^T + \sigma_e^2 \mathbf{I} = \sum_{u=1}^3 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T$$

对于一组非平衡的数据，组合平均数的向量可以表达为

$$\begin{bmatrix} Y_{12} \\ Y_{13} \\ Y_{21} \\ Y_{23} \\ Y_{31} \\ Y_{33} \end{bmatrix} = \begin{bmatrix} \mu \\ \mu \\ \mu \\ \mu \\ \mu \\ \mu \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} D_{12} \\ D_{13} \\ D_{23} \\ D_{33} \end{bmatrix} + \begin{bmatrix} e_{12} \\ e_{13} \\ e_{21} \\ e_{23} \\ e_{31} \\ e_{33} \end{bmatrix}$$

这一组数据的方差-协方差矩阵是

$$\mathbf{V} = \sigma_A^2 \begin{bmatrix} 2 & 1 & 2 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 & 2 & 2 \\ 2 & 1 & 2 & 1 & 1 & 0 \\ 1 & 1 & 1 & 2 & 1 & 2 \\ 1 & 2 & 1 & 1 & 2 & 2 \\ 0 & 2 & 0 & 2 & 2 & 4 \end{bmatrix} + \sigma_D^2 \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} + \sigma_e^2 \mathbf{I}$$

当所有的先验值都设为 $\alpha_{ii} = 1$ ，即MINQUE(1)估算方差分量时，该试验设计的 \mathbf{Q}_α 矩阵为

$$\mathbf{Q}_\alpha = \begin{bmatrix} .5663 & -.0414 & -.4337 & -.0691 & -.0414 & .0193 \\ -.0414 & .5884 & -.0414 & -.0193 & -.4116 & -.0746 \\ -.4337 & -.0414 & .5663 & -.0691 & -.0414 & .0193 \\ -.0691 & -.0193 & -.0691 & .3011 & -.0193 & -.1243 \\ -.0414 & -.4116 & -.0414 & -.0193 & .5888 & -.0746 \\ .0193 & -.0746 & .0193 & -.1243 & -.0746 & .2348 \end{bmatrix}$$

由以下MINQUE(1)方程组，可以求得方差分量的无偏估计值，

$$\begin{aligned} & \text{tr}(\mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{U}_A) \hat{\sigma}_A^2 + \text{tr}(\mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{U}_A) \hat{\sigma}_D^2 + \text{tr}(\mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_A) \hat{\sigma}_e^2 \\ & \quad = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{y} \\ & \text{tr}(\mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{U}_D) \hat{\sigma}_A^2 + \text{tr}(\mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{U}_D) \hat{\sigma}_D^2 + \text{tr}(\mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_D) \hat{\sigma}_e^2 \\ & \quad = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{y} \\ & \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_{(1)}) \hat{\sigma}_A^2 + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_{(1)}) \hat{\sigma}_D^2 + \text{tr}(\mathbf{Q}_{(1)} \mathbf{Q}_{(1)}) \hat{\sigma}_e^2 \\ & \quad = \mathbf{y}^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{y} \end{aligned}$$

对于以上具体的3×3双列杂交，估算方差分量的MINQUE(1)方程组是

$$\begin{bmatrix} .5359 & .2707 & .1934 \\ .2707 & .5156 & .3683 \\ .1934 & .3683 & 2.2836 \end{bmatrix} \begin{bmatrix} \sigma_A^2 \\ \sigma_D^2 \\ \sigma_e^2 \end{bmatrix} = \begin{bmatrix} \mathbf{y}^T \mathbf{Q}_\alpha \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_\alpha \mathbf{y} \\ \mathbf{y}^T \mathbf{Q}_\alpha \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_\alpha \mathbf{y} \\ \mathbf{y}^T \mathbf{Q}_\alpha \mathbf{Q}_\alpha \mathbf{y} \end{bmatrix}$$

如果某性状的六个观察值是 $\mathbf{y}^T = [28.90 \quad 30.93 \quad 28.38 \quad 21.60 \quad 27.31 \quad 27.23]$ ，以上方程组右边向量的转置是 $[3.6209 \quad 5.2051 \quad 11.3284]$ 。解以上方程组可以得到方差分量的三个估计值 $\hat{\sigma}_A^2 = 2.2555$ ， $\hat{\sigma}_D^2 = 6.2197$ ， $\hat{\sigma}_e^2 = 3.7666$ 。如果另

一个性状的六个观察值是 [12.63 14.98 13.02 11.40 13.22 12.89]，以上方程组右边向量的转置是 [0.4832 0.4831 1.9847]。解以上方程组可以得到方差分量的三个估计值 $\hat{\sigma}_A^2 = 0.5830$ ， $\hat{\sigma}_D^2 = 0.0511$ ， $\hat{\sigma}_e^2 = 0.8115$ 。

成对性状 y_1 和 y_2 之间协方差分量的无偏估计可由以下MINQUE(1)方程组解得，

$$\begin{aligned} & \text{tr}(\mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{U}_A) \hat{\sigma}_{A/A} + \text{tr}(\mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{U}_A) \hat{\sigma}_{D/D} + \text{tr}(\mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_A) \hat{\sigma}_{\varepsilon/\varepsilon} \\ & \quad = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{y}_2 \\ & \text{tr}(\mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_{(1)} \mathbf{U}_D) \hat{\sigma}_{A/A} + \text{tr}(\mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{U}_D) \hat{\sigma}_{D/D} + \text{tr}(\mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{U}_D) \hat{\sigma}_{\varepsilon/\varepsilon} \\ & \quad = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_{(1)} \mathbf{y}_2 \\ & \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_{(1)}) \hat{\sigma}_{A/A} + \text{tr}(\mathbf{Q}_{(1)} \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_{(1)}) \hat{\sigma}_D + \text{tr}(\mathbf{Q}_{(1)} \mathbf{Q}_{(1)}) \hat{\sigma}_{\varepsilon/\varepsilon} \\ & \quad = \mathbf{y}_1^T \mathbf{Q}_{(1)} \mathbf{Q}_{(1)} \mathbf{y}_2 \end{aligned}$$

对于这个3×3双列杂交，估算协方差分量的MINQUE(1)方程组是

$$\begin{bmatrix} .5359 & .2707 & .1934 \\ .2707 & .5156 & .3683 \\ .1934 & .3683 & 2.2836 \end{bmatrix} \begin{bmatrix} \sigma_{A/A} \\ \sigma_{D/D} \\ \sigma_{e/e} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1^T \mathbf{Q}_\alpha \mathbf{U}_A \mathbf{U}_A^T \mathbf{Q}_\alpha \mathbf{y}_2 \\ \mathbf{y}_1^T \mathbf{Q}_\alpha \mathbf{U}_D \mathbf{U}_D^T \mathbf{Q}_\alpha \mathbf{y}_2 \\ \mathbf{y}_1^T \mathbf{Q}_\alpha \mathbf{Q}_\alpha \mathbf{y}_2 \end{bmatrix}$$

以上方程组右边向量的转置是 [1.2773 1.2697 4.1970]。解以上方程组可以得到协方差分量的三个估计值 $\hat{\sigma}_{A/A} = 1.5509$ ， $\hat{\sigma}_{D/D} = 0.4849$ ， $\hat{\sigma}_{e/e} = 1.6284$ 。

第二节 加性-显性与环境互作的遗传模型

基因的加性和显性效应在各种环境下可能表现不一致。如果遗传试验在不同年份或地点实施,便可以估算加性和显性效应与环境的互作。这时性状的表现型方差(V_P)的分量组成是 $V_P = V_G + V_{GE} + V_e = V_A + V_D + V_{AE} + V_{DE} + V_e$ 。动、植物遗传分析中常用的遗传设计(如巢式交配设计、析因交配设计和双列杂交设计等),可用于间接估算加性和显性效应与环境的互作方差分量和协方差分量。

1. 巢式交配设计的遗传分析方法

巢式交配设计的遗传试验如果在不同环境下实施,第*i*个父本与其第*j*个母本交配的后代*k*在环境*h*中的表现型值可以用以下线性模型表示,

$$y_{hijk} = \mu + e_h + s_i + d_{ij} + es_{hi} + ed_{hij} + w_{hijk} \quad (5.4)$$

其中 μ 是群体平均数,固定效应;

e_h 是环境效应,固定效应;

s_i 是父本效应, $s_i \sim N(0, \sigma_s^2)$;

d_{ij} 是父本内的母本效应, $d_{ij} \sim N(0, \sigma_{d(s)}^2)$;

es_{hi} 是环境与父本的互作效应, $es_{hi} \sim N(0, \sigma_{es}^2)$;

ed_{hij} 是环境与父本内的母本的互作效应, $ed_{hij} \sim N(0, \sigma_{ed(s)}^2)$;

w_{hijk} 是家系内的个体在环境*h*中的效应, $w_{hijk} \sim N(0, \sigma_w^2)$ 。

因而个体的性状表现型值也具有正态分布

$$y_{hijk} \sim N(\mu + e_h, \sigma_s^2 + \sigma_{d(s)}^2 + \sigma_{es}^2 + \sigma_{ed(s)}^2 + \sigma_w^2)。$$

用多试点巢式试验设计的方差分析方法,可以估算以上五项方差分量。方差分析的变异来源、自由度、均方和期望均方列于表5.7。

5.7 巢式交配设计的方差分析表

变异来源	自由度	均方	期望均方
环境间	e-1	MSE	
父本间	s-1	MSS	$\sigma_w^2 + r\sigma_{ed(s)}^2 + rd\sigma_{es}^2 + re\sigma_{d(s)}^2 + red\sigma_s^2$
母本间	s(d-1)	MSD	$\sigma_w^2 + r\sigma_{ed(s)}^2 + re\sigma_{d(s)}^2$
环境 × 父本	(e-1)(s-1)	MSES	$\sigma_w^2 + r\sigma_{ed(s)}^2 + rd\sigma_{es}^2$
环境 × 母本	s(e-1)(d-1)	MSED	$\sigma_w^2 + r\sigma_{ed(s)}^2$
个体间	esd(r-1)	MSW	σ_w^2

表5.7中的各项平方和可以按下式计算:

$$\text{环境间平方和SSE} = rsd \sum_{h=1}^e (\bar{y}_{h...} - \bar{y}_{...})^2$$

$$\text{父本间平方和SSS} = red \sum_{i=1}^s (\bar{y}_{i..} - \bar{y}_{...})^2$$

$$\text{父本内母本间平方和SSD} = re \sum_{i=1}^s \sum_{j=1}^d (\bar{y}_{ij.} - \bar{y}_{i..})^2$$

$$\text{环境} \times \text{父本间平方和SSES} = rd \sum_{h=1}^e \sum_{i=1}^s (\bar{y}_{hi..} - \bar{y}_{h...} - \bar{y}_{i..} + \bar{y}_{...})^2$$

环境 × 父本内母本间平方和SSED

$$= r \sum_{h=1}^e \sum_{i=1}^s \sum_{j=1}^d (\bar{y}_{hij.} - \bar{y}_{h..} - \bar{y}_{ij.} + \bar{y}_{i..})^2$$

个体间平方和SSW

$$= \sum_{h=1}^e \sum_{i=1}^s \sum_{j=1}^d \sum_{k=1}^r (y_{hijk} - \bar{y}_{...})^2 - \text{SSE} - \text{SSS} - \text{SSD} - \text{SSES} - \text{SSED}$$

各项平方和除以其相应的自由度，可以算到各项均方。将计算出的各项均方等于其期望均方的估计值，解以下五项等式，

$$\text{MSS} = \hat{\sigma}_w^2 + r\hat{\sigma}_{ed(s)}^2 + rd\hat{\sigma}_{es}^2 + re\hat{\sigma}_{d(s)}^2 + red\hat{\sigma}_s^2$$

$$\text{MSD} = \hat{\sigma}_w^2 + r\hat{\sigma}_{\text{ed(s)}}^2 + re\hat{\sigma}_{\text{d(s)}}^2$$

$$\text{MSES} = \hat{\sigma}_w^2 + r\hat{\sigma}_{\text{ed(s)}}^2 + rd\hat{\sigma}_{\text{es}}^2$$

$$\text{MSED} = \hat{\sigma}_w^2 + r\hat{\sigma}_{\text{ed(s)}}^2$$

$$\text{MSW} = \hat{\sigma}_w^2$$

便可以求得方差分量的无偏估计值,

$$\hat{\sigma}_w^2 = \text{MSW}$$

$$\hat{\sigma}_{\text{ed(s)}}^2 = (\text{MSED} - \text{MSW})/r$$

$$\hat{\sigma}_{\text{es}}^2 = (\text{MSES} - \text{MSED})/rd$$

$$\hat{\sigma}_{\text{d(s)}}^2 = (\text{MSD} - \text{MSED})/re$$

$$\hat{\sigma}_s^2 = (\text{MSS} - \text{MSES} - \text{MSD} + \text{MSED})/red$$

巢式交配试验设计的方差分量估计值,可用于估算遗传方差分量。如果亲本是随机交配的动物或异花授粉的植物(近交系数 $F=0$),于是 $\hat{V}_A = 4\hat{\sigma}_s^2$, $\hat{V}_D = 4(\hat{\sigma}_{\text{d(s)}}^2 - \hat{\sigma}_s^2)$, $\hat{V}_{\text{AE}} = 4\hat{\sigma}_{\text{es}}^2$, $\hat{V}_{\text{DE}} = 4(\hat{\sigma}_{\text{ed(s)}}^2 - \hat{\sigma}_{\text{es}}^2)$ 。自花授粉植物的纯系材料作亲本时,取 F 值为1,因而 $\hat{V}_A = 2\hat{\sigma}_s^2$, $\hat{V}_D = \hat{\sigma}_{\text{d(s)}}^2 - \hat{\sigma}_s^2$, $\hat{V}_{\text{AE}} = 2\hat{\sigma}_{\text{es}}^2$, $\hat{V}_{\text{DE}} = \hat{\sigma}_{\text{ed(s)}}^2 - \hat{\sigma}_{\text{es}}^2$ 。

对于成对性状,通过计算各项效应的叉积和及均积,可以估算相应的遗传协方差分量。现将性状 X 和性状 Y 协方差分析的变异来源、自由度、均积和期望均积列于表5.8。

表5.8 巢式交配设计的协方差分析表

变异来源	自由度	均积	期望均积
环境间	$e-1$	MCPE	
父本间	$s-1$	MCPS	$\sigma_{w(XY)} + r\sigma_{ed(s)(XY)} + rd\sigma_{es(XY)} + re\sigma_{d(s)(XY)} + red\sigma_{s(XY)}$
母本间	$s(d-1)$	MCPD	$\sigma_{w(XY)} + r\sigma_{ed(s)(XY)} + re\sigma_{d(s)(XY)}$
环境 × 父本	$(e-1)(s-1)$	MCPES	$\sigma_{w(XY)} + r\sigma_{ed(s)(XY)} + rd\sigma_{es(XY)}$
环境 × 母本	$s(e-1)(d-1)$	MCPED	$\sigma_{w(XY)} + r\sigma_{ed(s)(XY)}$
个体间	$esd(r-1)$	MCPW	$\sigma_{w(XY)}$

将计算出的各项均积等于它们期望均积的估计值，然后可以求得协方差分量的无偏估计值：

$$\hat{\sigma}_{w(XY)} = \text{MCPW}$$

$$\hat{\sigma}_{ed(s)(XY)} = (\text{MCPED} - \text{MCPW})/r$$

$$\hat{\sigma}_{es(XY)} = (\text{MCPES} - \text{MCPED})/rd$$

$$\hat{\sigma}_{d(s)(XY)} = (\text{MCPD} - \text{MCPED})/re$$

$$\hat{\sigma}_{s(XY)} = (\text{MCPS} - \text{MCPES} - \text{MCPD} + \text{MCPED})/red$$

用以上估计值，可以进一步估算遗传协方差分量。当亲本是随机异交的动、植物时， $\hat{C}_A = 4\hat{\sigma}_{s(XY)}$ ， $\hat{C}_D = 4(\hat{\sigma}_{d(s)(XY)} - \tilde{\sigma}_{s(XY)})$ ， $\hat{C}_{AE} = 4\hat{\sigma}_{es(XY)}$ ， $\hat{C}_{DE} = 4(\hat{\sigma}_{ed(s)(XY)} - \hat{\sigma}_{es(XY)})$ 。对于自花授粉植物的纯系材料， $\hat{C}_A = 2\hat{\sigma}_{s(XY)}$ ， $\hat{C}_D = \hat{\sigma}_{d(s)(XY)} - \tilde{\sigma}_{s(XY)}$ ， $\hat{C}_{AE} = 2\hat{\sigma}_{es(XY)}$ ， $\hat{C}_{DE} = \hat{\sigma}_{ed(s)(XY)} - \hat{\sigma}_{es(XY)}$ 。

多年份或多试点巢式交配设计的各项方差分量和协方差分量，也可以采用混合线性模型的方式进行分析。用混合线性模型矩阵形式表示的巢式交配设计遗传模型是，

$$\mathbf{y} = \mathbf{Xb} + \mathbf{U}_S\mathbf{e}_S + \mathbf{U}_{D(S)}\mathbf{e}_{D(S)} + \mathbf{U}_{ES}\mathbf{e}_{ES} + \mathbf{U}_{ED(S)}\mathbf{e}_{ED(S)} + \mathbf{e}_w$$

$$= \mathbf{Xb} + \sum_{u=1}^5 \mathbf{U}_u \mathbf{e}_u$$

$$\sim (\mathbf{Xb}, \mathbf{V} = \sum_{u=1}^5 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T)$$

其中 \mathbf{b} 是环境固定效应向量， \mathbf{X} 是环境效应的系数矩阵；

\mathbf{e}_S 是父本效应向量， $\mathbf{e}_S \sim (\mathbf{0}, \sigma_S^2 \mathbf{I})$ ；

\mathbf{U}_S 是父本效应的系数矩阵；

$\mathbf{e}_{D(S)}$ 是父本内的母本效应向量， $\mathbf{e}_{D(S)} \sim (\mathbf{0}, \sigma_{d(s)}^2 \mathbf{I})$ ；

$\mathbf{U}_{D(S)}$ 是父本内母本效应的系数矩阵；

\mathbf{e}_{ES} 是环境 \times 父本互作效应向量， $\mathbf{e}_{ES} \sim (\mathbf{0}, \sigma_{es}^2 \mathbf{I})$ ；

\mathbf{U}_{ES} 是环境 \times 父本互作效应的系数矩阵；

$\mathbf{e}_{ED(S)}$ 是环境 \times 母本互作效应向量， $\mathbf{e}_{ED(S)} \sim (\mathbf{0}, \sigma_{ed(s)}^2 \mathbf{I})$ ；

$\mathbf{U}_{ED(S)}$ 是环境 \times 母本互作效应的系数矩阵；

\mathbf{e}_w 是剩余效应向量， $\mathbf{e}_w \sim (\mathbf{0}, \sigma_w^2 \mathbf{I})$ 。

采用MINQUE(1)法，可以无偏估算交配设计的各项方差分量和协方差分量，从而可以进一步估算遗传方差分量和遗传协方差分量。

2. 析因交配设计的遗传分析方法

如果在不同环境下实施析因交配设计的遗传试验，第 i 个父本与第 j 个母本交配的后代 k 在环境 h 中的表现型值可以用以下线性模型表示，

$$y_{hijk} = \mu + e_h + s_i + d_j + sd_{ij} + es_{hi} + ed_{hj} + esd_{hij} + w_{hijk} \quad (5.5)$$

其中 μ 是群体平均数，固定效应；

e_h 是环境效应，固定效应；

s_i 是父本效应， $s_i \sim N(0, \sigma_s^2)$ ；

d_j 是母本效应， $d_j \sim N(0, \sigma_d^2)$ ；

sd_{ij} 是父本 \times 母本互作效应， $sd_{ij} \sim N(0, \sigma_{sd}^2)$ ；

es_{hi} 是环境 × 父本的互作效应, $es_{hi} \sim N(0, \sigma_{es}^2)$;

ed_{hj} 是环境 × 母本的互作效应, $ed_{hj} \sim N(0, \sigma_{ed}^2)$;

esd_{hij} 是环境 × 父本 × 母本的互作效应, $esd_{hij} \sim N(0, \sigma_{esd}^2)$;

w_{hijk} 是个体在环境h中的效应, $w_{hijk} \sim N(0, \sigma_w^2)$ 。

因而个体的性状表现型值具有正态分布

$$y_{hijk} \sim N(\mu + e_h, \sigma_s^2 + \sigma_d^2 + \sigma_{sd}^2 + \sigma_{es}^2 + \sigma_{ed}^2 + \sigma_{esd}^2 + \sigma_w^2)$$

用多试点析因试验设计的方差分析方法, 可以估算以上七项方差分量。

方差分析的变异来源、自由度、均方和期望均方列于表5.9。

表5.9 析因交配设计的方差分析表

变异来源	自由度	均方	期望均方
环境间	$e-1$	MSE	
父本间	$s-1$	MSS	$\sigma_w^2 + r\sigma_{esd}^2 + rd\sigma_{es}^2 + re\sigma_{sd}^2 + red\sigma_s^2$
母本间	$d-1$	MSD	$\sigma_w^2 + r\sigma_{esd}^2 + rs\sigma_{ed}^2 + re\sigma_{sd}^2 + res\sigma_d^2$
父本 × 母本	$(s-1)(d-1)$	MSSD	$\sigma_w^2 + r\sigma_{esd}^2 + re\sigma_{sd}^2$
环境 × 父本	$(e-1)(s-1)$	MSES	$\sigma_w^2 + r\sigma_{esd}^2 + rd\sigma_{es}^2$
环境 × 母本	$(e-1)(d-1)$	MSED	$\sigma_w^2 + r\sigma_{esd}^2 + rs\sigma_{ed}^2$
环境 × 父本 × 母本	$(e-1)(s-1)(d-1)$	MSES D	$\sigma_w^2 + r\sigma_{esd}^2$
个体间	$e(sd-1)(r-1)$	MSW	σ_w^2

表5.9中的各项平方和可以按下式计算:

$$\text{环境间平方和SSE} = rsd \sum_{h=1}^e (\bar{y}_{h...} - \bar{y}_{...})^2$$

$$\text{父本间平方和SSS} = red \sum_{i=1}^s (\bar{y}_{i..} - \bar{y}_{...})^2$$

$$\text{母本间平方和SSD} = res \sum_{j=1}^d (\bar{y}_{.j.} - \bar{y}_{...})^2$$

$$\text{父本} \times \text{母本平方和SSSD} = re \sum_{i=1}^s \sum_{j=1}^d (\bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{...})^2$$

$$\text{环境} \times \text{父本间平方和SSES} = rd \sum_{h=1}^e \sum_{i=1}^s (\bar{y}_{hi.} - \bar{y}_{h...} - \bar{y}_{i.} + \bar{y}_{...})^2$$

$$\text{环境} \times \text{母本间平方和SSED} = rs \sum_{h=1}^e \sum_{j=1}^d (\bar{y}_{h.j} - \bar{y}_{h...} - \bar{y}_{.j} + \bar{y}_{...})^2$$

环境 × 父本 × 母本平方和SSESD

$$= r \sum_{h=1}^e \sum_{i=1}^s \sum_{j=1}^d (\bar{y}_{hij} - \bar{y}_{hi.} - \bar{y}_{h.j} - \bar{y}_{ij} + \bar{y}_{h...} + \bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{...})^2$$

$$\begin{aligned} \text{个体间平方和SSW} &= \sum_{h=1}^e \sum_{i=1}^s \sum_{j=1}^d \sum_{k=1}^r (y_{hijk} - \bar{y}_{...})^2 - \text{SSE} - \text{SSS} - \text{SSD} \\ &\quad - \text{SSSD} - \text{SSES} - \text{SSED} - \text{SSESD} \end{aligned}$$

平方和除以其相应的自由度，可以算到各项均方。将计算出的各项均方等于其期望均方的估计值，解以下七项等式，

$$\text{MSS} = \hat{\sigma}_w^2 + r\hat{\sigma}_{\text{esd}}^2 + rd\hat{\sigma}_{\text{es}}^2 + re\hat{\sigma}_{\text{sd}}^2 + red\hat{\sigma}_s^2$$

$$\text{MSD} = \hat{\sigma}_w^2 + rs\hat{\sigma}_{\text{ed}}^2 + r\hat{\sigma}_{\text{esd}}^2 + re\hat{\sigma}_{\text{sd}}^2 + res\hat{\sigma}_d^2$$

$$\text{MSSD} = \hat{\sigma}_w^2 + r\hat{\sigma}_{\text{esd}}^2 + re\hat{\sigma}_{\text{sd}}^2$$

$$\text{MSES} = \hat{\sigma}_w^2 + r\hat{\sigma}_{\text{esd}}^2 + rd\hat{\sigma}_{\text{es}}^2$$

$$\text{MSED} = \hat{\sigma}_w^2 + r\hat{\sigma}_{\text{esd}}^2 + rs\hat{\sigma}_{\text{ed}}^2$$

$$\text{MSESD} = \hat{\sigma}_w^2 + r\hat{\sigma}_{\text{esd}}^2$$

$$\text{MSW} = \hat{\sigma}_w^2$$

可求得方差分量的无偏估计值 $\hat{\sigma}_w^2 = \text{MSW}$, $\hat{\sigma}_{\text{esd}}^2 = (\text{MSED} - \text{MSW})/r$, $\hat{\sigma}_{\text{ed}}^2 = (\text{MSED} - \text{MSESD})/rs$, $\hat{\sigma}_{\text{es}}^2 = (\text{MSES} - \text{MSESD})/rd$, $\hat{\sigma}_{\text{sd}}^2 = (\text{MSSD} - \text{MSESD})/re$, $\hat{\sigma}_d^2 = (\text{MSD} - \text{MSSD} - \text{MSED} + \text{MSESD})/res$, $\hat{\sigma}_s^2 = (\text{MSS} - \text{MSSD} - \text{MSES} + \text{MSESD})/red$ 。

析因交配试验设计的方差分量估计值，可用于估算遗传方差分量。如果

亲本是随机交配的遗传群体(近交系数 $F=0$)，加性方差分量 $\hat{V}_A = 2(\hat{\sigma}_s^2 + \hat{\sigma}_d^2)$ ，显性方差分量 $\hat{V}_D = 4\hat{\sigma}_{sd}^2$ ，加性 \times 环境互作方差分量 $\hat{V}_{AE} = 2(\hat{\sigma}_{es}^2 + \hat{\sigma}_{ed}^2)$ ，显性 \times 环境互作方差分量 $\hat{V}_{DE} = 4\hat{\sigma}_{esd}^2$ 。当采用自交的纯系材料作亲本时，则取 F 值为 1，因而加性方差分量 $\hat{V}_A = \hat{\sigma}_s^2 + \hat{\sigma}_d^2$ ，显性方差分量 $\hat{V}_D = \hat{\sigma}_{sd}^2$ 。加性 \times 环境互作方差分量 $\hat{V}_{AE} = \hat{\sigma}_{es}^2 + \hat{\sigma}_{ed}^2$ ，显性 \times 环境互作协方差分量 $\hat{V}_{DE} = \hat{\sigma}_{esd}^2$ 。

对于成对性状，通过计算各项效应的叉积和以及均积，便可以估算相应的遗传协方差分量。现将性状 X 和性状 Y 协方差分析的变异来源、自由度、均积和期望均积列于表 5.10。

表 5.10 析因交配设计的协方差分析表

变异来源	自由度	均积	期望均积
环境间	$e-1$	MCPE	
父本间	$s-1$	MCPS	$\sigma_{w(XY)} + r\sigma_{esd(XY)} + rd\sigma_{es(XY)} + re\sigma_{sd(XY)} + red\sigma_{s(XY)}$
母本间	$d-1$	MCPD	$\sigma_{w(XY)} + r\sigma_{esd(XY)} + rs\sigma_{ed(XY)} + re\sigma_{sd(XY)} + res\sigma_{d(XY)}$
父本 \times 母本	$(s-1)(d-1)$	MCPSD	$\sigma_{w(XY)} + r\sigma_{esd(XY)} + re\sigma_{sd(XY)}$
环境 \times 父本	$(e-1)(s-1)$	MCPES	$\sigma_{w(XY)} + r\sigma_{esd(XY)} + rd\sigma_{es(XY)}$
环境 \times 母本	$(e-1)(d-1)$	MCPED	$\sigma_{w(XY)} + r\sigma_{esd(XY)} + rs\sigma_{ed(XY)}$
环境 \times 父本 \times 母本	$(e-1)(s-1)(d-1)$	MCPESD	$\sigma_{w(XY)} + r\sigma_{esd(XY)}$
个体间	$e(sd-1)(r-1)$	MCPW	$\sigma_{w(XY)}$

将各项均积等于它们期望均积的估计值，便可以求得协方差分量的无偏估计值 $\hat{\sigma}_{w(XY)} = \text{MCPW}$ ， $\hat{\sigma}_{esd(XY)} = (\text{MCPED} - \text{MCPW})/r$ ， $\hat{\sigma}_{ed(XY)} = (\text{MCPED} - \text{MCPESD})/rs$ ， $\hat{\sigma}_{es(XY)} = (\text{MCPES} - \text{MCPESD})/rd$ ， $\hat{\sigma}_{sd(XY)} = (\text{MCPSD} - \text{MCPESD})/re$ ， $\hat{\sigma}_{d(XY)} = (\text{MCPD} - \text{MCPSD} - \text{MCPED} +$

MCPESD)/res, $\hat{\sigma}_{s(XY)} = (\text{MCPS} - \text{MCPSD} - \text{MCPES} + \text{MCPESD})/\text{red}$ 。

用以上估计值, 可以进一步估算遗传协方差分量。对于随机交配的遗传群体, 加性协方差分量 $\hat{C}_A = 2(\hat{\sigma}_{s(XY)} + \hat{\sigma}_{d(XY)})$, 显性协方差分量 $\hat{C}_D = 4\hat{\sigma}_{sd(XY)}$, 加性 × 环境互作协方差分量 $\hat{C}_{AE} = 2(\hat{\sigma}_{es(XY)} + \hat{\sigma}_{ed(XY)})$, 显性 × 环境互作协方差分量 $\hat{C}_{DE} = 4\hat{\sigma}_{esd(XY)}$ 。对于自花授粉植物的纯系材料, 加性协方差分量 $\hat{C}_A = \hat{\sigma}_{s(XY)} + \hat{\sigma}_{d(XY)}$, 显性协方差分量 $\hat{C}_D = \hat{\sigma}_{sd(XY)}$, 加性 × 环境互作协方差分量 $\hat{C}_{AE} = \hat{\sigma}_{es(XY)} + \hat{\sigma}_{ed(XY)}$, 显性 × 环境互作协方差分量 $\hat{C}_{DE} = \hat{\sigma}_{esd(XY)}$ 。

多年份或多试点析因交配设计的各项方差分量和协方差分量, 也可以采用混合线性模型的方式进行分析。可用混合线性模型矩阵形式表示析因交配设计的遗传模型,

$$\begin{aligned} \mathbf{y} &= \mathbf{Xb} + \mathbf{U}_S \mathbf{e}_S + \mathbf{U}_D \mathbf{e}_D + \mathbf{U}_{SD} \mathbf{e}_{SD} + \mathbf{U}_{ES} \mathbf{e}_{ES} \\ &\quad + \mathbf{U}_{ED} \mathbf{e}_{ED} + \mathbf{U}_{ESD} \mathbf{e}_{ESD} + \mathbf{e}_w \\ &= \mathbf{Xb} + \sum_{u=1}^7 \mathbf{U}_u \mathbf{e}_u \\ &\sim (\mathbf{Xb}, \mathbf{V} = \sum_{u=1}^7 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T) \end{aligned}$$

其中 \mathbf{b} 是环境固定效应向量, \mathbf{X} 是环境效应的系数矩阵;

\mathbf{e}_S 是父本效应向量, $\mathbf{e}_S \sim (0, \sigma_s^2 \mathbf{I})$;

\mathbf{U}_S 是父本效应的系数矩阵;

\mathbf{e}_D 是母本效应向量, $\mathbf{e}_D \sim (0, \sigma_d^2 \mathbf{I})$;

\mathbf{U}_D 是母本效应的系数矩阵;

\mathbf{e}_{SD} 是父本 × 母本互作效应向量, $\mathbf{e}_{SD} \sim (0, \sigma_{sd}^2 \mathbf{I})$;

\mathbf{U}_{SD} 是父本 × 母本互作效应的系数矩阵;

\mathbf{e}_{ES} 是环境 × 父本互作效应向量, $\mathbf{e}_{ES} \sim (0, \sigma_{es}^2 \mathbf{I})$;

U_{ES} 是环境 × 父本互作效应的系数矩阵;

e_{ED} 是环境 × 母本互作效应向量, $e_{ED} \sim (0, \sigma_{ed}^2 \mathbf{I})$;

U_{ED} 是环境 × 母本互作效应的系数矩阵;

e_{ESD} 是环境 × 父本 × 母本互作效应向量, $e_{ESD} \sim (0, \sigma_{esd}^2 \mathbf{I})$;

U_{ESD} 是环境 × 父本 × 母本互作效应的系数矩阵;

e_w 是剩余效应向量, $e_w \sim (0, \sigma_w^2 \mathbf{I})$ 。

采用MINQUE(1)法, 可以无偏估算交配设计的各项方差分量和协方差分量, 从而可以进一步估算遗传方差分量和遗传协方差分量。

3. 双列杂交设计的遗传分析方法

如果双列杂交试验(方法4)在多年份或多试点实施, 可以分析配合力与环境的互作表现。当试验包括 e 个环境, p 个亲本的正交杂交组合 ($p(p-1)/2$ 个), b 个完全随机区组时, 母本 i 与父本 $j(j>i)$ 的杂交组合在第 h 个环境内的第 k 个区组中平均表现型值可用以下线性模型表示,

$$y_{hijk} = \mu + e_h + g_i + g_j + s_{ij} + ge_{hi} + ge_{hj} + se_{hij} + b_{hk} + \varepsilon_{hijk} \quad (5.6)$$

其中 μ 是群体平均数, 固定效应;

e_h 是环境效应, 固定效应;

g_i 或 g_j 是一般配合力效应, g_i 或 $g_j \sim N(0, \sigma_g^2)$;

s_{ij} 是特殊配合力效应, $s_{ij} \sim N(0, \sigma_s^2)$;

ge_{hi} 或 ge_{hj} 是一般配合力 × 环境互作效应,

$$ge_{hi} \text{ 或 } ge_{hj} \sim N(0, \sigma_{ge}^2);$$

se_{hij} 是特殊配合力 × 环境互作效应, $se_{hij} \sim N(0, \sigma_{se}^2)$;

b_{hk} 是环境内区组效应, $b_{hk} \sim N(0, \sigma_b^2)$;

ε_{hijk} 是剩余效应, $\varepsilon_{hijk} \sim N(0, \sigma_\varepsilon^2)$ 。

双列杂交组合的表现型方差 $\text{Var}(y_{hijk})$ 为

$$\sigma_y^2 = 2\sigma_g^2 + \sigma_s^2 + 2\sigma_{ge}^2 + \sigma_{se}^2 + \sigma_b^2 + \sigma_\varepsilon^2。$$

配合力方差分量可以用方差分析的方法估算。

由表5.11可以估算双列杂交设计中的各项方差分量，并可进一步估算各项遗传方差分量。试验设计的方差分量估计值，可用于估算遗传方差分量。如果亲本是随机交配的遗传群体(近交系数 $F=0$)，加性方差分量 $\hat{V}_A = 4\hat{\sigma}_g^2$ ，显性方差分量 $\hat{V}_D = 4\hat{\sigma}_s^2$ ，加性×环境互作方差分量 $\hat{V}_{AE} = 4\hat{\sigma}_{ge}^2$ ，显性×环境互作方差分量 $\hat{V}_{DE} = 4\hat{\sigma}_{se}^2$ 。当采用自交的纯系材料作亲本时，则取 F 值为1，因而加性方差分量 $\hat{V}_A = 2\sigma_g^2$ ，显性方差分量 $\hat{V}_D = \sigma_s^2$ 。加性×环境互作方差分量 $\hat{V}_{AE} = 2\hat{\sigma}_{ge}^2$ ，显性×环境互作协方差分量 $\hat{V}_{DE} = \hat{\sigma}_{se}^2$ 。

表5.11 双列杂交设计(方法4)的方差分析表

变异来源	自由度	均方	期望均方
环境间	$e-1$	MSE	
环境内区组间	$e(b-1)$	MSB	
一般配合力	$p-1$	MSG	$\sigma_g^2 + b\sigma_{se}^2 + be\sigma_s^2 + b(p-2)\sigma_{ge}^2 + be(p-2)\sigma_g^2$
特殊配合力	$p(p-3)/2$	MSS	$\sigma_g^2 + b\sigma_{se}^2 + be\sigma_s^2$
一般配合力×环境	$(e-1)(p-1)$	MSG	$\sigma_g^2 + b\sigma_{se}^2 + b(p-2)\sigma_{ge}^2$
特殊配合力×环境	$p(e-1)(p-3)/2$	MSS	$\sigma_g^2 + b\sigma_{se}^2$
机误	df_e	MSe	σ_e^2

注: $df_e=e(b-1)[p(p-1)/2-1]$ 。

也可采用协方差分析的方法，估算成对性状 X 和 Y 之间的遗传协方差分量。

由表5.12可以估算双列杂交设计中的各项协方差分量，并可进一步估算各项遗传协方差分量。用获得的协方差分量估计值，可以进一步估算遗传协方差分量。对于随机交配的遗传群体，加性协方差分量 $\hat{C}_A = 4\hat{\sigma}_{g(XY)}$ ，显性协方差分量 $\hat{C}_D = 4\hat{\sigma}_{s(XY)}$ ，加性×环境互作协方差分量 $\hat{C}_{AE} = 4\hat{\sigma}_{ge(XY)}$ ，显性×环境互作协方差分量 $\hat{C}_{DE} = 4\hat{\sigma}_{se(XY)}$ 。对于自花授粉植物的纯系材料，加性协方差分量 $\hat{C}_A = 2\hat{\sigma}_{g(XY)}$ ，显性协方差分量 $\hat{C}_D = \hat{\sigma}_{s(XY)}$ ，加性×环境互

作协方差分量 $\hat{C}_{AE} = 2\hat{\sigma}_{ge(XY)}$ ，显性 × 环境互作协方差分量 $\hat{C}_{DE} = \hat{\sigma}_{se(XY)}$ 。

表5.12 双列杂交设计(方法4)的协方差分析表

变异来源	自由度	均积	期望均积
环境间	$e-1$	MCPE	
环境内区组间	$e(b-1)$	MCPB	
一般配合力	$p-1$	MCP	$\sigma_{\varepsilon(XY)} + b\sigma_{se(XY)} + be\sigma_{s(XY)}$ $+ b(p-2)\sigma_{ge(XY)} + be(p-2)\sigma_{g(XY)}$
特殊配合力	$p(p-3)/2$	MCPS	$\sigma_{\varepsilon(XY)} + b\sigma_{se(XY)} + be\sigma_{s(XY)}$
一般配合力 × 环境	$(e-1)(p-1)$	MCPG	$\sigma_{\varepsilon(XY)} + b\sigma_{se(XY)} + b(p-2)\sigma_{ge(XY)}$
特殊配合力 × 环境	$p(e-1)(p-3)/2$	MCPS	$\sigma_{\varepsilon(XY)} + b\sigma_{se(XY)}$
机误	df_e	MCPe	$\sigma_{\varepsilon(XY)}$

注: $df_e = e(b-1)[p(p-1)/2 - 1]$ 。

以上所介绍的双列分析方法只适用于平衡数据的分析。采用混合线性模型的分析方法，可以无偏地估算有不规则缺失的双列杂交试验的各项方差分量和协方差分量。如果用混合线性模型矩阵形式表示双列杂交设计的遗传模型，则有

$$\begin{aligned} \mathbf{y} &= \mathbf{Xb} + \mathbf{U}_G \mathbf{e}_G + \mathbf{U}_S \mathbf{e}_S + \mathbf{U}_{GE} \mathbf{e}_{GE} + \mathbf{U}_{SE} \mathbf{e}_{SG} + \mathbf{U}_B \mathbf{e}_B + \mathbf{e}_\varepsilon \\ &= \mathbf{Xb} + \sum_{u=1}^6 \mathbf{U}_u \mathbf{e}_u \\ &\sim (\mathbf{Xb}, \mathbf{V} = \sum_{u=1}^6 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T) \end{aligned}$$

其中 \mathbf{b} 是群体均值和环境固定效应向量， \mathbf{X} 是群体均值和环境效应的系数矩阵；

\mathbf{e}_G 是一般配合力效应向量， $\mathbf{e}_G \sim (\mathbf{0}, \sigma_g^2 \mathbf{I})$ ；

\mathbf{U}_G 是一般配合力效应的系数矩阵；

\mathbf{e}_S 是特殊配合力效应向量， $\mathbf{e}_S \sim (\mathbf{0}, \sigma_s^2 \mathbf{I})$ ；

\mathbf{U}_S 是特殊配合力效应的系数矩阵；

\mathbf{e}_{GE} 是一般配合力 × 环境互作效应向量， $\mathbf{e}_{GE} \sim (\mathbf{0}, \sigma_{ge}^2 \mathbf{I})$ ；

U_{GE} 是一般配合力 \times 环境互作效应的系数矩阵;

e_{SE} 是特殊配合力 \times 环境互作效应向量, $e_{SE} \sim (\mathbf{0}, \sigma_{se}^2 \mathbf{I})$;

U_{SE} 是特殊配合力效应的系数矩阵;

e_B 是区组效应向量, $e_B \sim (\mathbf{0}, \sigma_B^2 \mathbf{I})$;

U_B 是区组效应的系数矩阵;

e_ε 是剩余效应向量, $e_\varepsilon \sim (\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I})$ 。

采用MINQUE(1)法, 可以无偏估算双列杂交设计的各项方差分量和协方差分量, 从而可以进一步估算遗传方差分量和遗传协方差分量。

第六章 加性-显性-上位性遗传模型分析

第一节 简单加性-显性-上位性遗传模型

数量性状由微效多基因所控制。根据简单加性-显性遗传模型的假设,不同基因座位的遗传效应是累加的,非等位基因之间不存在互作效应(即上位性效应)。但是有些性状的非等位基因之间可能存在不同程度的互作效应,因而有必要进一步分析这些性状的上位性效应。当在遗传模型中包括上位性效应时,总遗传方差可以分解为,

$$\begin{aligned}V_G &= V_A + V_D + V_I \\ &= V_A + V_D + V_{AA} + V_{DD} + V_{AD}\end{aligned}$$

上位性遗传方差分量的估算比一般的方差分量估算更困难。数量遗传分析结果表明,大多数性状的上位性遗传方差不如加性和显性遗传方差重要(Hallauer和Mirana, 1981)。在上位性遗传方差中,加性与加性的上位性遗传方差(V_{AA})是主要分量。与加性效应一样,加性与加性的上位性效应是可以通过选择而稳定遗传的。因此估算加性与加性的上位性遗传方差和遗传协方差,对于育种和选择有重要的意义。

运用Cockerham(1980)的广义遗传模型,朱军(1989, 1992)提出了估算加性与加性的上位性遗传方差分量的遗传分析方法。采用MINQUE(1)法,分析包括杂种一代和杂种二代或回交一代的双列杂交遗传资料,可以在估算加性和显性遗传方差分量的同时,无偏地估算加性与加性的上位性遗传方差分量。

如果从某遗传群体中随机抽取含有 p 个纯合亲本的样本,并以这些亲本作双列杂交,产生一组正交杂交组合 F_1 (组合数 $\leq p(p-1)/2$ 个),由 F_1

自交而产生 F_2 。当试验材料有亲本($k = 0$)和母本 i 与父本 j 杂交组合的两个杂种世代($k = 1, 2$)，试验采用 b 个完全随机区组时，不同基因型在第 l 个区组中的平均表现型值可用以下线性模型表示，

亲本 $P(ii)$ ($k=0$)

$$y_{ijkl} = \mu + 2A_i + D_{ii} + 4AA_{ii} + B_l + e_{ijkl}$$

杂种一代 $F_1(ij)$ ($k=1$)

$$y_{ijkl} = \mu + A_i + A_j + D_{ij} + AA_{ii} + AA_{jj} + 2AA_{ij} + B_l + e_{ijkl}$$

杂种二代 $F_2(ij)$ ($k=2$)

$$y_{ijkl} = \mu + A_i + A_j + 0.25D_{ii} + 0.25D_{jj} + 0.5D_{ij} + AA_{ii} + AA_{jj} + 2AA_{ij} + B_l + e_{ijkl}$$

其中 μ 是群体平均数，固定效应；

A_i 或 A_j 是加性效应， A_i 或 $A_j \sim (0, \sigma_A^2)$ ；

D_{ii} 、 D_{jj} 或 D_{ij} 是显性效应， D_{ii} 、 D_{jj} 或 $D_{ij} \sim (0, \sigma_D^2)$ ；

AA_{ii} 、 AA_{jj} 或 AA_{ij} 是加性与加性的上位性效应，

AA_{ii} 、 AA_{jj} 或 $AA_{ij} \sim (0, \sigma_{AA}^2)$ ；

B_l 是区组效应， $B_l \sim (0, \sigma_B^2)$ ；

e_{ijkl} 是剩余效应， $e_{ijkl} \sim (0, \sigma_e^2)$ 。

在这个线性模型中，除了群体平均数 μ 是固定效应外，其它各项效应都是随机效应。表现型方差的分量组成为

$$\sigma_y^2 = 2\sigma_A^2 + \sigma_D^2 + 4\sigma_{AA}^2 + \sigma_B^2 + \sigma_e^2$$

对于动物或一些异花授粉植物， F_2 世代可以由 F_1 的个体随机交配而产生。但是有时无法获得某些生物群体的 F_2 个体，可以用 F_1 与亲本杂交产生的回交一代(BC) ($k = 3$)替代 F_2 。回交一代 $F_1(ij) \times P_j$ ($k = 3$)的遗传

模型是

$$y_{ijkl} = \mu + 0.5A_i + 1.5A_j + 0.5D_{jj} + 0.5D_{ij} + 0.25AA_{ii} + 2.25AA_{jj} + 1.5AA_{ij} + B_l + e_{ijkl}$$

由以上分析加性-显性-上位性的遗传模型可见，遗传材料中包括了F₂或BC等基因型分离的世代，因此模型中加性效应、加性×加性的上位性效应或显性效应的系数出现了小数。凡是系数出现小数的模型，都不能用传统的方差分析方法进行分析。

如果用混合线性模型矩阵形式表示加性-显性-上位性遗传模型，则有

$$\begin{aligned} \mathbf{y} &= \mathbf{1}\mu + \mathbf{U}_A\mathbf{e}_A + \mathbf{U}_D\mathbf{e}_D + \mathbf{U}_{AA}\mathbf{e}_{AA} + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_\varepsilon \\ &= \mathbf{1}\mu + \sum_{u=1}^5 \mathbf{U}_u\mathbf{e}_u \\ &\sim (\mathbf{1}\mu, \mathbf{V} = \sum_{u=1}^5 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T) \end{aligned}$$

其中 μ 是群体平均数固定效应， $\mathbf{1}$ 是常数1的系数向量；

\mathbf{e}_A 是加性效应向量， $\mathbf{e}_A \sim (0, \sigma_A^2 \mathbf{I})$ ；

\mathbf{U}_A 是加性效应的系数矩阵；

\mathbf{e}_D 是显性效应向量， $\mathbf{e}_D \sim (0, \sigma_D^2 \mathbf{I})$ ；

\mathbf{U}_D 是显性效应的系数矩阵；

\mathbf{e}_{AA} 是加性×加性的上位性效应向量， $\mathbf{e}_{AA} \sim (0, \sigma_{AA}^2 \mathbf{I})$ ；

\mathbf{U}_{AA} 是加性×加性上位性效应的系数矩阵；

\mathbf{e}_B 是区组效应向量， $\mathbf{e}_B \sim (0, \sigma_B^2 \mathbf{I})$ ；

\mathbf{U}_B 是区组效应的系数矩阵；

\mathbf{e}_ε 是剩余效应向量， $\mathbf{e}_\varepsilon \sim (0, \sigma_\varepsilon^2 \mathbf{I})$ 。

采用MINQUE(1)法可以无偏地估算加性-显性-上位性遗传模型中的各项方差分量(朱军, 1992)，从而获得遗传方差分量的无偏估计值，

$\hat{V}_A = 2\hat{\sigma}_A^2$, $\hat{V}_D = \hat{\sigma}_D^2$, $\hat{V}_{AA} = 4\hat{\sigma}_{AA}^2$ 。对于成对性状X和Y, 也可以采用MINQUE(1)法无偏估算各项遗传协方差分量, $\hat{C}_A = 2\hat{\sigma}_{A(XY)}$, $\hat{C}_D = \hat{\sigma}_{D(XY)}$, $\hat{C}_{AA} = 4\hat{\sigma}_{AA(XY)}$ 。在估算方差分量和协方差分量的同时, 还可以采用LUP法或AUP法无偏预测各项遗传效应值。根据遗传效应值的预测, 可以评价亲本及其组合的育种价值。

第二节 加性-显性-上位性及与环境互作的遗传模型

如果影响数量性状的基因上位性效应在不同环境下表现不一致, 即表明存在上位性与环境的互作效应。这时需要在不同环境条件下重复进行遗传试验, 以便分析上位性与环境的互作效应。加性-显性-上位性与环境互作的遗传模型也无法用传统的方差分析法进行分析。采用混合线性模型的分析方法, 可以有效地分析上位性与环境的互作效应。

如果用混合线性模型矩阵形式表示加性-显性-上位性与环境互作的遗传模型, 则有

$$\begin{aligned} \mathbf{y} &= \mathbf{Xb} + \mathbf{U}_A \mathbf{e}_A + \mathbf{U}_D \mathbf{e}_D + \mathbf{U}_{AA} \mathbf{e}_{AA} + \mathbf{U}_{AE} \mathbf{e}_{AE} + \mathbf{U}_{DE} \mathbf{e}_{DE} \\ &\quad + \mathbf{U}_{AAE} \mathbf{e}_{AAE} + \mathbf{U}_B \mathbf{e}_B + \mathbf{e}_\varepsilon \\ &= \mathbf{Xb} + \sum_{u=1}^8 \mathbf{U}_u \mathbf{e}_u \\ &\sim (\mathbf{Xb}, \mathbf{V} = \sum_{u=1}^8 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T) \end{aligned}$$

其中 \mathbf{b} 是环境固定效应, \mathbf{X} 是环境效应的系数矩阵;

\mathbf{e}_A 是加性效应向量, $\mathbf{e}_A \sim (0, \sigma_A^2 \mathbf{I})$;

\mathbf{U}_A 是加性效应的系数矩阵;

\mathbf{e}_D 是显性效应向量, $\mathbf{e}_D \sim (0, \sigma_D^2 \mathbf{I})$;

\mathbf{U}_D 是显性效应的系数矩阵;

\mathbf{e}_{AA} 是加性×加性的上位性效应向量, $\mathbf{e}_{AA} \sim (0, \sigma_{AA}^2 \mathbf{I})$;

\mathbf{U}_{AA} 是加性×加性上位性效应的系数矩阵;

\mathbf{e}_{AE} 是加性×环境互作效应向量, $\mathbf{e}_{AE} \sim (0, \sigma_{AE}^2 \mathbf{I})$;

\mathbf{U}_{AE} 是加性×环境互作效应的系数矩阵;

\mathbf{e}_{DE} 是显性×环境互作效应向量, $\mathbf{e}_{DE} \sim (0, \sigma_{DE}^2 \mathbf{I})$;

\mathbf{U}_{DE} 是显性×环境互作效应的系数矩阵;

\mathbf{e}_{AAE} 是加性×加性上位性×环境互作效应向量, $\mathbf{e}_{AAE} \sim (0, \sigma_{AAE}^2 \mathbf{I})$;

\mathbf{U}_{AAE} 是加性×加性上位性×环境互作效应的系数矩阵;

\mathbf{e}_B 是区组效应向量, $\mathbf{e}_B \sim (0, \sigma_B^2 \mathbf{I})$;

\mathbf{U}_B 是区组效应的系数矩阵;

\mathbf{e}_ε 是剩余效应向量, $\mathbf{e}_\varepsilon \sim (0, \sigma_\varepsilon^2 \mathbf{I})$ 。

采用MINQUE(1)法可以无偏地估算以上模型中的各项方差分量,从而获得遗传方差分量的无偏估计值, $\hat{V}_A = 2\hat{\sigma}_A^2$, $\hat{V}_D = \hat{\sigma}_D^2$, $\hat{V}_{AA} = 4\hat{\sigma}_{AA}^2$, $\hat{V}_{AE} = 2\hat{\sigma}_{AE}^2$, $\hat{V}_{DE} = \hat{\sigma}_{DE}^2$, $\hat{V}_{AAE} = 4\hat{\sigma}_{AAE}^2$ 。对于成对性状X和Y, 也可以采用MINQUE(1)法无偏地估算各项遗传协方差分量, $\hat{C}_A = 2\hat{\sigma}_{A(XY)}$, $\hat{C}_D = \hat{\sigma}_{D(XY)}$, $\hat{C}_{AA} = 4\hat{\sigma}_{AA(XY)}$, $\hat{C}_{AE} = 2\hat{\sigma}_{AE(XY)}$, $\hat{C}_{DE} = \hat{\sigma}_{DE(XY)}$, $\hat{C}_{AAE} = 4\hat{\sigma}_{AAE(XY)}$ 。在估算方差分量和协方差分量的同时, 还能采用LUP法或AUP法无偏预测各项遗传效应值。根据遗传效应与环境互作的预测分析, 可以评价亲本及其组合的在不同环境条件下的育种价值表现。

现以1992和1993年陆地棉衣分的双列杂交资料(包括10个亲本及其部分 F_1 和 F_2)遗传研究结果作为分析实例。10个品种或品系分别为:无蜜腺选系:①A226、②A160;无腺体选系:③A17;我国自育品种:④鲁棉6号、⑤中棉所12号、⑥中棉所13号、⑦徐州184、⑧泗棉2号;引种美国材料:⑨4305、⑩4318。试验在浙江农业大学实验农场进行,随机区组排列,三次重复。以亲本①~⑤作母本、亲本⑥~⑩作为父本,共配制了20个组合(见表6.3),缺失5个组合,属非平衡数据。采用MINQUE(1)法估算控制衣分的各项方差分量及其对表现型方差的比率(朱军,1992)。用AUP法(朱军,1993)预测了各性状的基因效应值。采用Jackknife的方法(朱军,1992,1993),以年份内区组为重复抽样单位,计算各项遗传参数的估计值(或预测值)及其标准误,然后采用t测验(自由度=5)对遗传参数作统计检验。

表6.1 陆地棉衣分方差分量及其对表现型方差的比率的估计值±标准误

方差	估计值 ±标准误	方差比率	估计值 ±标准误
V_A	3.412 ±0.904 **	V_A/V_P	0.271 ±0.075 **
V_D	2.730 ±0.999 *	V_D/V_P	0.217 ±0.081 *
V_{AA}	2.132 ±0.335 **	V_{AA}/V_P	0.170 ±0.024 **
V_{AE}	0.000 ±0.000	V_{AE}/V_P	0.000 ±0.000
V_{DE}	2.419 ±1.086 *	V_{DE}/V_P	0.192 ±0.071 *
V_{AAE}	0.098 ±0.044 *	V_{AAE}/V_P	0.008 ±0.004 +
V_e	1.783 ±0.247 **	V_e/V_P	0.142 ±0.021 **

+、*、** 分别达到0.10准显著水平、0.05显著水平或0.01极显著水平。

由表6.1列出的陆地棉衣分方差分量及其对表现型方差(V_P)的比率的估计值结果可以推断,控制棉花衣分性状的主要遗传效应是加性效应和显性效应,这二种遗传效应对表现型总变异的贡献分别超过20%。其它重要的遗传效应包括加×加上位性效应和显性×环境互作效应,它们对表现型总变异的贡献接近20%左右。在不同环境条件下表现稳定的基因主效应(加性、显性和加×加上位性效应)的方差比率高达65.8%,而

随机机误的方差比率为14.2%。较大的显性×环境互作效应方差分量表明，一些杂交组合在不同年份的表现可能不一致。

由于加性、显性和加×加上位性基因主效应是主要的遗传效应，因此有必要进一步分析其遗传效应值的表现，从而可以推断亲本和组合的遗传表现。表6.2列出了陆地棉10个亲本的衣分加性效应(A_i)、显性效应(D_{ii})和上位性效应(AA_{ii})的预测值及其标准误。由表6.2可见，三个推广的当家品种④鲁棉6号、⑤中棉所12号、和⑦徐州184有较高的加性效应。因此，从它们的杂种后代中较易获得高衣分的遗传材料。无蜜腺选系①A226、②A160和美国引种材料⑩4318的加性效应值较低，表明这三个材料不宜作为高衣分育种的杂交亲本。亲本①A226、②A160和⑧泗棉2号具有很低的显性效应(D_{ii})预测值，预示这三个亲本的杂种后代将有较大的自交衰退现象。加×加上位性效应的预测数值一般小于加性或显性效应的预测值，这表明加×加上位性效应对棉花衣分的作用小于加性效应和显性效应。

朱军等(1993)建议根据统计量 $-\sum_{i=1}^n \hat{D}_{ii}$ 推断显性效应的总体作用方向。为了使不同性状的基因显性效应有可比性，可以采用标准化的显性方向

$$\hat{\Delta} = \frac{1}{\sqrt{n\hat{\sigma}_D^2}} - \sum_{i=1}^n \hat{D}_{ii}$$

如果杂合显性效应有增值作用(正向杂种优势), $\Delta > 0$; 杂合显性效应表现减值作用时(负向杂种优势), $\Delta < 0$; 如果显性效应方差是零, 则所有的显性效应 D_{ii} 和 D_{ij} 都为零, 因而 $\Delta = 0$ 。这种分析显性效应作用方向的方法也

适用于加×加上位性效应: $-\sum_{i=1}^n AA_{ii}$ 可用于推断加×加上位性效应的总体作用方向, 其标准化的加×加上位性方向是

$$\hat{\Delta}_{AA} = \frac{1}{\sqrt{n\hat{\sigma}_{AA}^2}} - \sum_{i=1}^n AA_{ii}$$

如果杂合加×加上上位性(AA_{ij})有增值作用(正向杂种优势), $\Delta_{AA} > 0$; 杂合加×加上上位性效应表现减值作用时(负向杂种优势), $\Delta_{AA} < 0$; ; 如果加×加上上位性效应方差是零, 加×加上上位性效应值也都为零, 因此 $\Delta_{AA} = 0$ 。

表6.2 所分析的棉花衣分杂种优势结果表明 Δ 和 Δ_{AA} 均未显著地偏于零, 因此衣分性状不表现明显的杂种优势。

表6.2 陆地棉衣分加性效应(A_i)、显性效应(D_{ii})和上位性效应(AA_{ii})的预测值±标准误

	\hat{A}_i ±标准误		\hat{D}_{ii} ±标准误		AA_{ii} ±标准误	
$i=1$	-1.018	±0.197 **	-1.961	±0.340 **	0.210	±0.144
$i=2$	-1.459	±0.160 **	-2.202	±0.933 +	0.294	±0.168
$i=3$	0.727	±0.242 *	1.488	±0.604 +	-0.252	±0.240
$i=4$	1.173	±0.163 **	0.289	±0.566	0.440	±0.064 **
$i=5$	1.315	±0.258 **	0.626	±0.248 +	0.333	±0.092 *
$i=6$	0.668	±0.127 **	-0.771	±0.614	0.019	±0.129
$i=7$	1.599	±0.287 **	-1.015	±0.762	0.808	±0.240 *
$i=8$	-0.003	±0.149	-1.929	±0.885 +	0.603	±0.180 *
$i=9$	-0.558	±0.210 *	1.509	±0.537 *	-0.689	±0.145 **
$i=10$	-2.445	±0.287 **	0.063	±0.634	-0.976	±0.190 **

+、*、**分别达到0.10准显著水平、0.05显著水平或0.01极显著水平。

表6.3列出了陆地棉20个杂交组合的衣分显性效应(D_{ij})和上位性效应(AA_{ij})的预测值及其标准误, 这些效应值是不受环境条件变异影响的基因主效应。4个杂交组合具有较高的显性效应值, 有较强的显性杂种优势。由于棉花衣分存在较大的显性×环境交互效应, 有必要进一步分析各组合的杂种优势稳定性表现。

表6.3 陆地棉组合衣分显性效应(D_{ij})和上位性效应(AA_{ij})的预测值±标准误

编号	组合	\hat{D}_{ij} ±标准误	AA_{ij} ±标准误
1	(1×6)	4.038 ±1.069 *	-0.861 ±0.100 **
2	(1×7)	0.783 ±0.832	-0.011 ±0.186
3	(1×9)	0.210 ±0.446	-0.438 ±0.247
4	(1×10)	-1.626 ±1.096	-0.795 ±0.222 *
5	(2×6)	-2.839 ±0.803 *	1.467 ±0.152 **
6	(2×7)	1.427 ±0.648 +	-0.254 ±0.174
7	(2×8)	3.032 ±1.126 *	-1.337 ±0.256 **
8	(2×9)	-0.883 ±0.632	-0.35 ±0.198
9	(2×10)	2.946 ±1.394 +	-2.497 ±0.398 **
10	(3×7)	2.456 ±1.292	0.129 ±0.411
11	(3×9)	-0.990 ±0.693	0.190 ±0.429
12	(3×10)	-4.066 ±1.015 *	1.377 ±0.145 **
13	(4×6)	0.682 ±0.659	0.492 ±0.226 +
14	(4×8)	0.156 ±0.892	0.440 ±0.195 +
15	(4×9)	-1.855 ±1.299	0.566 ±0.213 *
16	(4×10)	1.027 ±0.924	-0.406 ±0.180 +
17	(5×7)	-1.840 ±0.541 *	1.256 ±0.154 **
18	(5×8)	0.670 ±0.785	-0.259 ±0.252
19	(5×9)	0.201 ±0.655	0.381 ±0.160 +
20	(5×10)	0.376 ±0.791	0.121 ±0.375

+、*、** 分别达到0.10准显著水平、0.05显著水平或0.01极显著水平。

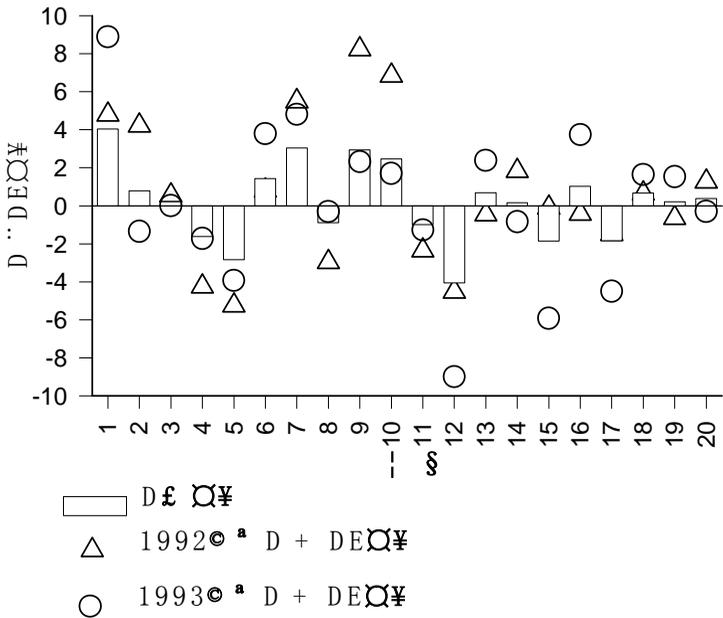


图6.1 陆地棉20个组合衣分显性主效应及显性×环境互作效应。

图6.1是陆地棉20个杂交组合的衣分显性主效应 (D_{ij}) 及1992年、1993年的显性×环境互作效应的预测值，其中柱状标记是基因的显性主效应，三角形标记是1992年的显性效应(包括显性主效应D和显性×环境互作效应DE)，园形标记是1993年的显性效应(D+DE)。由图6.1可见，虽然4个杂交组合(1)、(7)、(9)和(10)的棉花衣分具有较高的显性效应值，但 these combinations' dominant advantage stability performance is not completely the same. Combination (1) is more suitable for planting in 1993 environmental conditions (low temperature and high rainfall), while combinations (9) and (10) are the opposite, more suitable for 1992 environmental conditions (high temperature and low rainfall). Combination (7) has a better stability, showing good performance in both environments.

第七章 包括母体效应的遗传模型分析

第一节 加性-显性-母体效应遗传模型

1. Griffing 的双列杂交模型

无论是动物的幼畜还是植物的种子或幼苗，性状表现除了受其核基因的遗传效应影响外，还在不同程度上受到母体基因型的影响。有些性状可能还受到父体效应的影响。所以总的遗传效应(G)可以进一步分解为加性(A)、显性(D)、和母体(M)、父体(P)效应。总的遗传方差也包括了四项遗传方差分量， $V_G = V_A + V_D + V_M + V_P$ 。

如果双列杂交设计包括全部正反交组合时，正交和反交之间的遗传变异度量了平均的母体方差和父体方差。Griffing的双列杂交方法1和方法3可用于分析正反交变异(Griffing, 1965)。双列杂交方法3是用 p 个亲本产生 $p(p-1)$ 个正反交杂交组合。如果试验采用 b 个完全随机区组，母本 i 与父本 j 的杂交组合在第 k 个区组中的平均表现型值可用以下线性模型表示，

$$y_{ijk} = \mu + g_i + g_j + s_{ij} + r_{ij} + b_k + e_{ijk} \quad (7.1)$$

1)

其中 μ 是群体平均数，固定效应。

Griffing的双列杂交分析可以采用固定模型(模型)或随机模型(模型)。当采用随机模型时，各项随机效应值具有以下分布，一般配合力效应 g_i 或 $g_j \sim N(0, \sigma_g^2)$ ；特殊配合力效应 $s_{ij} \sim N(0, \sigma_s^2)$ ；正反交效应 $r_{ij} \sim N(0, \sigma_r^2)$ ；区组效应 $b_k \sim N(0, \sigma_b^2)$ ；剩余效应 $e_{ijk} \sim N(0, \sigma_e^2)$ 。因此

双列杂交组合的表现型方差 $\text{Var}(y_{ijk})$ 为

$$\sigma_y^2 = 2\sigma_g^2 + \sigma_s^2 + \sigma_r^2 + \sigma_b^2 + \sigma_e^2$$

对于没有缺失的平衡数据，一般配合力方差、特殊配合力方差和正反交方差可以用方差分析的方法估算。

表7.1 双列杂交设计(方法3)的方差分析表

变异来源	自由度	平方和	均方	期望均方
区组间	$b - 1$	SSb	MSb	
一般配合力	$p - 1$	SSg	MSg	$\sigma_e^2 + 2b\sigma_s^2 + 2b(p-2)\sigma_g^2$
特殊配合力	$p(p-3)/2$	SSs	MSs	$\sigma_e^2 + 2b\sigma_s^2$
正反交效应	$p(p-1)/2$	SSr	MSr	$\sigma_e^2 + 2b\sigma_r^2$
机误	df _e	SSE	MSe	σ_e^2

注: $df_e = (b-1)[p(p-1)-1]$ 。

设 $S1 = b(p-1)^2/[2(p-2)] \sum_{i=1}^p (\bar{y}_{i..} + \bar{y}_{.i})^2$ ，表7.1中的各项平方和可按下式

计算:

$$\text{区组间平方和SSb} = [p(p-1)] \sum_{k=1}^b (\bar{y}_{..k} - \bar{y}_{...})^2$$

$$\text{一般配合力平方和SSg} = S1 - [2bp(p-1)^2/(p-2)] (\bar{y}_{...})^2$$

特殊配合力平方和SSs

$$= (b/2) \sum_{i=1}^{p-1} \sum_{j>i}^p (\bar{y}_{ij.} + \bar{y}_{ji.})^2 - S1 + bp^2(p-1)/(p-2) (\bar{y}_{...})^2$$

$$\text{正反交效应平方和SSr} = (b/2) \sum_{i=1}^{p-1} \sum_{j>i}^p (\bar{y}_{ij.} - \bar{y}_{ji.})^2$$

$$\text{机误平方和SSE} = \sum_{i=1}^{p-1} \sum_{j>i}^p \sum_{k=1}^b y_{ijk}^2 - \text{SSg} - \text{SSs} - \text{SSr} - \text{SSb}$$

其中 $\bar{y}_{...}$ 是试验的总平均数， $\bar{y}_{..k}$ 是第 k 个区组的所有杂交组合平均数， $\bar{y}_{i..}$

是以第 i 个亲本作母本的所有组合平均数, $\bar{y}_{.i}$ 是以第 i 个亲本作父本的所有组合平均数, \bar{y}_{ij} 是组合 $(i \times j)$ 的区组平均数。

双列杂交设计中的各项方差分量可由解以下等式而求得,

$$MSg = SSg / (p - 1) = \hat{\sigma}_e^2 + b\hat{\sigma}_s^2 + 2b(p-2)\hat{\sigma}_g^2$$

$$MSs = 2SSs / [\rho(p-3)] = \hat{\sigma}_e^2 + 2b\hat{\sigma}_s^2$$

$$MSr = 2SSr / [\rho(p-1)] = \hat{\sigma}_e^2 + 2b\hat{\sigma}_r^2$$

$$MSe = SSe / dfe = \hat{\sigma}_e^2$$

因此方差分量的无偏估计值是 $\hat{\sigma}_e^2 = MSe$, $\hat{\sigma}_r^2 = (MSr - MSe) / 2b$, $\hat{\sigma}_s^2 = (MSs - MSe) / 2b$, $\hat{\sigma}_g^2 = (MSg - MSs) / [2b(p-2)]$ 。如果不存在上位性效应, 遗传方差分量可按下式估算, 加性方差分量 $\hat{V}_A = 4\hat{\sigma}_g^2 / (1 + F)$, 显性方差分量 $\hat{V}_D = 4\hat{\sigma}_s^2 / (1 + F)^2$, 平均母体和父体方差分量 $(1/2)(\hat{V}_M + V_p) = \hat{\sigma}_r^2$ 。

采用协方差分析的方法, 也可估算双列杂交设计(方法3)成对性状X和Y的遗传协方差分量。

表7.2 双列杂交设计(方法3)的协方差分析表

变异来源	自由度	叉积和	均积	期望均积
区组间	$b-1$	SCPb	MCPb	
一般配合力	$p-1$	SCPg	MCPg	$\sigma_{e(XY)} + 2b\sigma_{s(XY)} + 2b(p-2)\sigma_{g(XY)}$
特殊配合力	$\rho(p-3)/2$	SCPs	MCPs	$\sigma_{e(XY)} + 2b\sigma_{s(XY)}$
正反交效应	$\rho(p-1)/2$	SCPr	MCPr	$\sigma_{e(XY)} + 2b\sigma_{r(XY)}$
机误	dfe	SCPe	MCPe	$\sigma_{e(XY)}$

注: $dfe = (b-1)[\rho(p-1) - 1]$ 。

$$\text{设 } C1 = b(p-1)^2 / [2(p-2)] \sum_{i=1}^p (\bar{y}_{i..(X)} + \bar{y}_{.i.(X)}) (\bar{y}_{i..(Y)} + \bar{y}_{.i.(Y)})$$

$$C2 = (b/2) \sum_{i=1}^{p-1} \sum_{j>i}^p (\bar{y}_{ij.(X)} + \bar{y}_{ji.(X)}) (\bar{y}_{ij.(Y)} + \bar{y}_{ji.(Y)})$$

表7.2中的各项叉积和可按下列式计算:

区组间叉积和SCPb

$$= [\rho(\rho - 1)] \sum_{k=1}^b (\bar{y}_{..k(X)} - \bar{y}_{... (X)}) (\bar{y}_{..k(Y)} - \bar{y}_{... (Y)}) ,$$

一般配合力叉积和SCPg

$$= C1 - [2bp(\rho-1)^2/(\rho-2)] (\bar{y}_{... (X)}) (\bar{y}_{... (Y)}) ,$$

特殊配合力叉积和SCPs

$$= C2 - C1 + [bp^2(\rho-1)/(\rho-2)] (\bar{y}_{... (X)}) (\bar{y}_{... (Y)}) .$$

正交交互叉积和SCP_r

$$= (b/2) \sum_{i=1}^{p-1} \sum_{j>i}^p (\bar{y}_{ij.(X)} - \bar{y}_{ji.(X)}) (\bar{y}_{ij.(Y)} - \bar{y}_{ji.(Y)}) ,$$

机误叉积和SCPe

$$= \sum_{i=1}^{p-1} \sum_{j>i}^p \sum_{k=1}^b (y_{ijk(X)}) (y_{ijk(Y)}) - \text{SSg} - \text{SSs} - \text{SSr} - \text{SSb}$$

双列杂交设计中的各项协方差分量可由解以下等式而求得，

$$\text{MCPg} = \text{SCPg}/(\rho-1) = \hat{\sigma}_{e(XY)} + 2b\hat{\sigma}_{s(XY)} + 2b(\rho-2)\hat{\sigma}_{g(XY)}$$

$$\text{MCPs} = 2\text{SCPs}/[\rho(\rho-3)] = \hat{\sigma}_{e(XY)} + 2b\hat{\sigma}_{s(XY)}$$

$$\text{MCPr} = 2\text{SCP}_r/[\rho(\rho-1)] = \hat{\sigma}_{e(XY)} + 2b\hat{\sigma}_{r(XY)}$$

$$\text{MCPe} = \text{SCPe}/\text{dfe} = \hat{\sigma}_{e(XY)}$$

协方差分量的无偏估计值为， $\hat{\sigma}_{e(XY)} = \text{MCPe}$ ， $\hat{\sigma}_{r(XY)} = (\text{MCP}_r - \text{MCPe})/2b$ ，

$\hat{\sigma}_{s(XY)} = (\text{MCPs} - \text{MCPe})/2b$ ， $\hat{\sigma}_{g(XY)} = (\text{MCPg} - \text{MCPs})/[2b(\rho-2)]$ 。

遗传协方差分量的估算公式是，加性协方差分量 $\hat{C}_A =$

$4\hat{\sigma}_{g(XY)} / (1+F)$, 显性协方差分量 $\hat{C}_D = 4\hat{\sigma}_{s(XY)} / (1+F)^2$, 平均母体和父体协方差分量 $(1/2)(\hat{C}_M + \hat{C}_P) = \hat{\sigma}_{r(XY)}$ 。

有时参与双列杂交的亲本是一组特定的遗传材料, 双列分析所获得的结论只用于推断这一组材料的遗传表现。在这种情况下, Griffing的双列分析必须采用固定模型。此时可以估算各亲本的一般配合力效应值以及组合的特殊配合力效应值, 但是不能再同时估算各项遗传方差分量。

2. Cockerham 和 Weir 的生物模型(Bio-model)

Griffing的双列分析只能用正反交方差估算平均的母体和父体方差分量。如果不存在父体效应, 正反交方差只估算了一半的母体方差分量。在实际应用时, 无法确定父体效应在正反交效应中所占的分量, 因而不能正确测定实际的母体效应分量。

Cockerham和Weir(1977)提出了一个具有生物学意义的双列分析模型, 把正反交效应分解为母体效应和父体效应。这个生物模型(Bio-model)可以用以下线性公式表示,

$$y_{ijk} = \mu + n_i + n_j + t_{ij} + m_i + p_j + b_k + e_{ijk} \quad (7.2)$$

其中 y_{ijk} 是母本 i 与父本 j 的杂交组合在第 k 个区组中的平均表现型值;

μ 是群体平均数;

n_i 或 n_j 是核基因累加效应, n_i 或 $n_j \sim N(0, \sigma_n^2)$;

t_{ij} 是核基因互作效应, $t_{ij} \sim N(0, \sigma_t^2)$;

m_i 是母体效应, $m_i \sim N(0, \sigma_m^2)$;

p_j 是父体效应, $p_j \sim N(0, \sigma_p^2)$;

b_k 是区组效应, $b_k \sim N(0, \sigma_b^2)$;

e_{ijk} 是剩余效应, $e_{ijk} \sim N(0, \sigma_e^2)$ 。

在这个线性模型中，群体平均数 μ 是固定效应，其它效应都是随机效应。表现型方差 $\text{Var}(y_{ijk})$ 为

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

Cockerham和Weir(1977)给出了无偏估算 σ_n^2 和 σ_t^2 的方差分析公式。如果亲本是纯系遗传材料，可以直接估算出加性和显性方差分量： $\hat{V}_A = 2\hat{\sigma}_n^2$ ， $\hat{V}_D = \hat{\sigma}_t^2$ 。但是他们没能提出分别估算母体方差和父体方差的公式。朱军(1989)运用蒙特卡罗模拟证明，用MINQUE法可以无偏地估算这个生物模型中的全部方差分量。

3. 两个遗传模型的蒙特卡罗模拟比较

朱军(1989)采用蒙特卡罗模拟比较了Griffing(1956)双列杂交遗传模型(方法1)与Cockerham和Weir(1977)双列杂交生物模型(Bio-model)的分析效益。

对于同一组亲本的双列杂交资料，采用Griffing(1956)的方法及Cockerham和Weir(1977)的方法分析，其估算的方差分量具有以下可比性：

$$\sigma_g^2 \sim \sigma_n^2, \quad \sigma_s^2 \sim \sigma_t^2, \quad \sigma_r^2 \sim \frac{1}{2}(\sigma_m^2 + \sigma_p^2).$$

蒙特卡罗模拟的试验群体是三次重复的8×8完全双列杂交(方法1)。Griffing(1956)的模型采用ANOVA的方法估算方差分量，Cockerham和Weir(1977)的方法则采用MINQUE(1)的方法与Jackknife抽样技术估算方差分量。根据100次蒙特卡罗模拟结果，计算了平均的方差估计值($\hat{\sigma}^2$)、参数估计的偏差和均方误差，其分析结果列于表7.1。

表7.3 两种遗传模型估算方差分量的偏差(Bias)和均方误(MSE)比较

参数	Griffing 方法1			生物模型	
	真值	Bias	MSE	Bias	MSE
$\sigma_g^2:\sigma_n^2$	50	7.23	986.9	5.09	1445.4
$\sigma_s^2:\sigma_t^2$	30	0.70	76.1	-0.53	91.8
$\sigma_r^2:\frac{1}{2}(\sigma_m^2 + \sigma_p^2)$	20	-1.76	131.6	-0.37	148.4
$-\sigma_m^2$	20			-1.59	436.1
$-\sigma_p^2$	20			1.12	445.1
$\sigma_g^2:\sigma_n^2$	50	12.10	1509.8	3.14	1418.8
$\sigma_s^2:\sigma_t^2$	30	0.40	102.2	-0.46	95.1
$\sigma_r^2:\frac{1}{2}(\sigma_m^2 + \sigma_p^2)$	20	-0.85	108.0	2.23	153.9
$-\sigma_m^2$	40			-2.49	619.7
$-\sigma_p^2$	0			7.14	182.7
$\sigma_g^2:\sigma_n^2$	50	4.92	977.8	4.68	1051.3
$\sigma_s^2:\sigma_t^2$	60	1.20	368.4	-0.80	330.7
$\sigma_r^2:\frac{1}{2}(\sigma_m^2 + \sigma_p^2)$	0	0.59	1.2	1.69	7.6
$-\sigma_m^2$	0			2.05	21.0
$-\sigma_p^2$	0			1.61	10.4

蒙特卡罗模拟结果表明：当存在核外遗传效应 ($\sigma_m^2 = \sigma_p^2 = 20$

或 $\sigma_m^2 = 20, \sigma_p^2 = 0$) 时，一般配合力方差 σ_g^2 的估算误差较大，而核加性效

应方差 σ_n^2 的估算无偏性较好。Griffing(1956)的方法可以无偏地估算平

均的核外遗传效应方差、即正反交方差 $\frac{1}{2}(\sigma_m^2 + \sigma_p^2) \sim \sigma_r^2$ 。当父母本效应都存在时 ($\sigma_m^2 = \sigma_p^2 = 20$)，Cockerham和Weir (1977)的方法可以无偏地分别估算其方差分量。但是，如果只存在母体效应而不存在父体效应 ($\sigma_m^2 = 20, \sigma_p^2 = 0$)，母体效应方差可以获得较好的估算，可是往往会过高地估算父体效应方差。因此，此时宜采用不包括父体效应的遗传模型。如果完全不存在核外遗传效应 ($\sigma_m^2 = \sigma_p^2 = 0$)，Griffing(1956)的方法或Cockerham和Weir(1977)的方法都能有效地估算核遗传效应的方差分量。

4. 加性-显性-母体效应遗传模型(ADM模型)

朱军等(1993)运用Cockerham (1980)的广义遗传模型原理，提出了可直接分析数量性状加性、显性和母体遗传效应的模型(简称ADM模型)。如果从某遗传群体中随机抽取含有 p 个纯合亲本的样本，并以这些亲本作双列杂交产生一组杂交组合 F_1 。这个遗传模型可以用以下线性公式表示，

$$y_{ijk} = \mu + A_i + A_j + D_{ij} + M_i + B_k + e_{ijk} \quad (7.3)$$

其中 y_{ijk} 是母本 i 与父本 j 的杂交组合在第 k 个区组中的平均表现型值；

μ 是群体平均数；

A_i 或 A_j 是累加的加性效应， A_i 或 $A_j \sim (0, \sigma_A^2)$ ；

D_{ij} 是累加的显性效应， $D_{ij} \sim (0, \sigma_D^2)$ ；

M_i 是累加的母体效应， $M_i \sim (0, \sigma_M^2)$ ；

B_k 是区组效应， $B_k \sim (0, \sigma_B^2)$ ；

e_{ijk} 是剩余效应， $e_{ijk} \sim (0, \sigma_e^2)$ 。

在这个线性模型中，群体平均数 μ 是固定效应，其它效应都是随机效

应。表现型方差 $\text{Var}(y_{ijk})$ 为

$$\sigma_y^2 = 2\sigma_A^2 + \sigma_D^2 + \sigma_M^2 + \sigma_B^2 + \sigma_e^2。$$

这个ADM模型中遗传效应可以用混合线性模型的分析方法分析。如果用混合线性模型矩阵形式表示加性-显性-母体效应的遗传模型，则有

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{U}_A\mathbf{e}_A + \mathbf{U}_D\mathbf{e}_D + \mathbf{U}_M\mathbf{e}_M + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_\varepsilon \quad (7.4)$$

$$= \mathbf{1}\mu + \sum_{u=1}^5 \mathbf{U}_u \mathbf{e}_u$$

$$\sim (\mathbf{1}\mu, \mathbf{V} = \sum_{u=1}^5 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T)$$

其中 μ 是群体平均数固定效应， $\mathbf{1}$ 是常数1的系数向量；

\mathbf{e}_A 是加性效应向量， $\mathbf{e}_A \sim (0, \sigma_A^2 \mathbf{I})$ ；

\mathbf{U}_A 是加性效应的系数矩阵；

\mathbf{e}_D 是显性效应向量， $\mathbf{e}_D \sim (0, \sigma_D^2 \mathbf{I})$ ；

\mathbf{U}_D 是显性效应的系数矩阵；

\mathbf{e}_M 是母体效应向量， $\mathbf{e}_M \sim (0, \sigma_M^2 \mathbf{I})$ ；

\mathbf{U}_M 是母体效应的系数矩阵；

\mathbf{e}_B 是区组效应向量， $\mathbf{e}_B \sim (0, \sigma_B^2 \mathbf{I})$ ；

\mathbf{U}_B 是区组效应的系数矩阵；

\mathbf{e}_ε 是剩余效应向量， $\mathbf{e}_\varepsilon \sim (0, \sigma_\varepsilon^2 \mathbf{I})$ 。

采用MINQUE(1)法可以无偏估算试验设计中的各项随机效应的方差分量，然后由下式直接估算出加性、显性和母体方差分量： $\hat{V}_A = 2\hat{\sigma}_A^2$ ，

$$\hat{V}_D = \hat{\sigma}_D^2, \hat{V}_M = \hat{\sigma}_M^2。$$

对于成对性状，采用MINQUE(1)法可以无偏地估算各项协方差分量(朱军，1992)，然后可以估算加性、显性和母体协方差分量： $\hat{C}_A = 2\hat{\sigma}_{A/A}$ ，

$\hat{C}_D = \hat{\sigma}_{D/D}$, $\hat{C}_M = \hat{\sigma}_{M/M}$ 。在估算遗传方差分量和遗传协方差分量的同时, 还能采用LUP法或AUP法无偏预测亲本的加性效应值和母体效应值, 以及组合的显性效应值。根据遗传效应的预测分析结果, 可以评价亲本及其组合的育种价值和杂种优势的表现。

采用Griffing(1956)的双列杂交模型分析正反交效应, 需要有全套的正反交组合, 且尚无法准确地分析母体效应。采用ADM遗传模型, 可以无偏地分析母体效应, 而且双列杂交的正反交组合并不是都必需的。在遗传交配设计中包括全套的正反交组合, 可以提高分析母体效应的功效, 有利于鉴别微小的母体效应。如果实验中没有包括反交组合, 或只包括不完整的反交组合, 也仍然可以采用ADM遗传模型分析母体效应。

朱军等(1993)曾以六个陆地棉品种(徐州142、中棉所7号、协作2号、岱字棉15号、无腺体的GL-5和无蜜腺的HG-H-12)完全双列杂交的产量性状资料, 比较了ADM模型、AD模型与Griffing配合力模型的分析结果。遗传实验包括6个亲本和30个正反交组合, 完全随机区组设计, 重复4次。采用Jackknife数值抽样技术对试验区组进行抽样, 计算方差分量的估计值、基因效应的预测值以及它们的标准误。由于有4个区组, 自由度为 $df = 4 - 1 = 3$ 。

配合力方差分量、遗传方差分量和狭义遗传率, 其估算结果列于表7.4。各性状的正反交方差或母体方差均不显著, 这表明陆地棉产量性状主要受核基因控制。因而采用AD模型分析, 只需估算较少的参数, 可望获得更有效的估计值。采用Griffing配合力模型, 所估算的一般配合力方差的标准误较大, 因而不易检验出显著的一般配合力方差; 此外, 按Griffing的方法还无法直接估算狭义遗传率的标准误。

ADM模型和AD模型所估算的遗传方差分量和狭义遗传率值很接近, 但是一些估计值的标准误略有差异, 因而导致显著性测验的结果不尽一致。ADM模型分析结果表明, 单株铃数和衣分的加性方差、单铃重的加性方差和狭义遗传率都达到5%的显著水准。可是AD模型的分析结果表明, 单株铃

数和衣分的加性方差达到1%的极显著水准，而单铃重的加性方差和狭义遗传率均不显著。四个产量性状的显性方差均为显著或极显著，单株铃数、衣分和皮棉产量的狭义遗传率极显著。

表7.4 用Griffing模型、ADM模型和AD模型估算的方差分量及其标准误

参数	单株铃数		单铃重		衣分		皮棉产量	
	估计值	标准误	估计值	标准误	估计值	标准误	估计值	标准误
Griffing模型								
σ_g^2	2.73 ± 1.87		0.04 ± 0.03		0.63 ± 0.47		37.89 ± 27.29	
σ_s^2	1.14 ± 0.57		0.05 ± 0.03		0.62 ± 0.29		32.23 ± 13.51	
σ_r^2	-0.01 ± 0.13		-0.01 ± 0.01		-0.02 ± 0.04		0.20 ± 1.65	
σ_e^2	2.79 ± 0.39		0.23 ± 0.03		1.08 ± 0.15		32.50 ± 4.49	
$h_N^2\%$	58.20	—	21.55	—	42.96	—	53.86	—
ADM模型								
σ_A^2	2.48 ± 0.67		0.03 ± 0.01		0.57 ± 0.14		36.58 ± 3.32	
σ_D^2	1.63 ± 0.46		0.08 ± 0.02		0.78 ± 0.25		43.68 ± 6.24	
σ_M^2	-0.10 ± 0.01		-0.01 ± 0.01		-0.08 ± 0.04		-0.14 ± 2.67	
σ_e^2	2.83 ± 0.22		0.23 ± 0.05		1.07 ± 0.16		32.16 ± 6.42	
$h_N^2\%$	54.12 ± 6.61		15.91 ± 5.26		39.73 ± 2.49		49.09 ± 2.63	
AD模型								
σ_A^2	2.55 ± 0.38		0.03 ± 0.01		0.57 ± 0.04		32.12 ± 6.90	
σ_D^2	1.62 ± 0.46		0.08 ± 0.02		0.78 ± 0.24		43.77 ± 6.32	
$\sigma_{e e}^2$	2.80 ± 0.21		0.22 ± 0.05		1.03 ± 0.15		32.04 ± 7.44	
$h_N^2\%$	53.73 ± 4.21		16.32 ± 8.22		38.39 ± 4.43		46.25 ± 7.04	

采用Griffing模型和AD模型分析所得的6个供试亲本品种的一般配合力、特殊配合力估计值和加性效应、显性效应预测值列于表7.5。这两种模型的分析结果表现了一致的趋势，但Griffing模型配合力估计值的标准误较小。徐州142(品种1)和中棉所7号(品种2)都是高产组合的优良亲本，但

它们使杂种 F_1 增产的遗传机制不同。徐州142能使后代有较多的成铃数和较高的衣分，而中棉所7号则主要能使 F_1 增加单铃重。如果进一步利用 F_2 的杂种优势，则不宜采用中棉所7号的杂交组合。因为中棉所7号的显性效应估计值 \hat{D}_{22} 最小，其 F_2 可能会有较严重的衰退。四个性状的显性方向均为显著或极显著的正值，这表明陆地棉杂交组合的显性基因在产量性状上起增值作用，导致产量性状表现正向的杂种优势。所有亲本皮棉产量的纯合显性效应值 D_{ii} 都是显著或极显著的负值，因此皮棉产量的显性方向为超亲正向。

Griffing的配合力模型及其分析方法以方差分析(ANOVA)为基础。该分析方法具有一些难以克服的局限性。首先它们只能分析平衡数据。对于育种实践中常出现的不规则缺失的遗传数据无法分析。育种工作者对一组双列杂交材料进行遗传分析时，既希望能够了解亲本参与杂交的组合能力，也希望能够认识性状的遗传规律。对某一特定的双列杂交材料，Griffing的分析方法不能在估算亲本配合力(采用固定模型)的同时再估算配合力方差分量和遗传率(采用随机模型)。人们一般只能用Griffing的双列分析方法估算配合力，再用Hayman(1954)的双列分析方法估算遗传方差分量和遗传率。由于这是两个不同的遗传模型和分析方法，对一组资料分析的结果可能不尽一致。采用MINQUE(1)的统计方法分析ADM模型或AD模型，可以在估算遗传方差和遗传率的同时预测遗传效应值，从而可对亲本和组合的遗传表现进行评估。

表7.5 用Griffing模型估算及用AD模型预测的遗传效应值及其标准误

参数	单株铃数		单铃重		衣分		皮棉产量	
	估计值	标准误	估计值	标准误	估计值	标准误	估计值	标准误
Griffing模型								
g_1	2.22 ± 0.05		-0.06 ± 0.01		1.02 ± 0.02		6.64 ± 0.56	
g_2	-0.98 ± 0.05		0.18 ± 0.01		0.38 ± 0.02		5.93 ± 0.56	
g_3	-0.20 ± 0.05		-0.21 ± 0.01		-0.20 ± 0.02		-2.05 ± 0.56	
g_4	0.60 ± 0.05		-0.09 ± 0.01		0.49 ± 0.02		2.48 ± 0.56	
g_5	-2.70 ± 0.05		0.37 ± 0.01		-1.47 ± 0.02		-11.03 ± 0.56	
g_6	1.07 ± 0.05		-0.19 ± 0.01		-0.22 ± 0.02		-1.96 ± 0.56	
s_{11}	-2.92 ± 0.25		-0.49 ± 0.02		0.74 ± 0.10		-7.18 ± 2.93	
s_{22}	-0.82 ± 0.25		-0.79 ± 0.02		-1.80 ± 0.10		-16.37 ± 2.93	
s_{33}	-0.84 ± 0.25		-0.19 ± 0.02		-0.63 ± 0.10		-5.98 ± 2.93	
s_{44}	-1.29 ± 0.25		-0.61 ± 0.02		-0.90 ± 0.10		-11.35 ± 2.93	
s_{55}	-2.55 ± 0.25		-0.33 ± 0.02		0.89 ± 0.10		-4.78 ± 2.93	
s_{66}	-2.31 ± 0.25		-0.25 ± 0.02		-1.78 ± 0.10		-7.62 ± 2.93	
AD模型								
A_1	1.85 ± 0.21		-0.05 ± 0.07		1.05 ± 0.17		6.08 ± 0.56	
A_2	-0.77 ± 0.40		0.15 ± 0.09		0.21 ± 0.16		4.46 ± 1.72	
A_3	-0.08 ± 0.30		-0.16 ± 0.02		-0.18 ± 0.19		-1.51 ± 0.69	
A_4	0.58 ± 0.10		-0.09 ± 0.02		0.40 ± 0.16		1.94 ± 0.19	
A_5	-2.48 ± 0.18		0.31 ± 0.06		-1.15 ± 0.05		-9.36 ± 0.65	
A_6	0.90 ± 0.31		-0.15 ± 0.04		-0.33 ± 0.10		-1.61 ± 0.65	
D_{11}	-1.84 ± 0.24		-0.41 ± 0.11		0.74 ± 0.41		-4.56 ± 0.89	
D_{22}	-0.96 ± 0.39		-0.58 ± 0.08		-1.31 ± 0.15		-11.65 ± 2.03	
D_{33}	-0.80 ± 0.20		-0.22 ± 0.15		-0.56 ± 0.14		-5.58 ± 1.29	
D_{44}	-0.97 ± 0.46		-0.50 ± 0.17		-0.61 ± 0.30		-8.63 ± 1.12	
D_{55}	-2.64 ± 0.63		-0.20 ± 0.04		0.30 ± 0.25		-6.65 ± 0.87	
D_{66}	-1.63 ± 0.34		-0.27 ± 0.06		-1.43 ± 0.20		-6.80 ± 0.61	

如果存在基因型与环境的交互，需要在多点或多年实验的基础上分析

加性、显性和母体效应以及这些效应与环境的互作。朱军(1994)列出了包括基因型与环境互作的ADM遗传模型。在环境 h 内,由亲本 i 和亲本 j 杂交产生一组杂交组合 $F_{1(i,j)}$ 在区组 k 中的遗传模型,可以用以下线性公式表示,

$$y_{hijk} = \mu + E_h + A_i + A_j + D_{ij} + M_i + AE_{hi} + AE_{hj} + DE_{hij} + ME_{hi} + B_{k(h)} + e_{hijk} \quad (7.5)$$

其中 y_{hijk} 是母本 i 与父本 j 的杂交组合在环境 h 内第 k 个区组中的平均表现型值;

μ 是群体平均数, E_h 是环境效应, 这两个参数均为固定效应;

A_i 或 A_j 是累加的加性效应, A_i 或 $A_j \sim (0, \sigma_A^2)$;

D_{ij} 是累加的显性效应, $D_{ij} \sim (0, \sigma_D^2)$;

M_i 是累加的母体效应, $M_i \sim (0, \sigma_M^2)$;

AE_{hi} 或 AE_{hj} 是加性 \times 环境互作效应, AE_{hi} 或 $AE_{hj} \sim (0, \sigma_{AE}^2)$;

DE_{hij} 是显性 \times 环境互作效应, $DE_{hij} \sim (0, \sigma_{DE}^2)$;

ME_{hi} 是母体 \times 环境互作效应, $ME_{hi} \sim (0, \sigma_{ME}^2)$;

$B_{k(h)}$ 是环境内的区组效应, $B_{k(h)} \sim (0, \sigma_B^2)$;

e_{hijk} 是剩余效应, $e_{hijk} \sim (0, \sigma_e^2)$ 。

表现型方差 $\text{Var}(y_{hijk})$ 为

$$\sigma_y^2 = 2\sigma_A^2 + \sigma_D^2 + \sigma_M^2 + 2\sigma_{AE}^2 + \sigma_{DE}^2 + \sigma_{ME}^2 + \sigma_B^2 + \sigma_e^2$$

如果用混合线性模型矩阵形式表示包括互作效应的加性-显性-母体效应的遗传模型, 则有

$$\mathbf{y} = \mathbf{Xb} + \mathbf{U}_A \mathbf{e}_A + \mathbf{U}_D \mathbf{e}_D + \mathbf{U}_M \mathbf{e}_M + \mathbf{U}_{AE} \mathbf{e}_{AE} + \mathbf{U}_{DE} \mathbf{e}_{DE} + \mathbf{U}_{ME} \mathbf{e}_{ME} + \mathbf{U}_B \mathbf{e}_B + \mathbf{e}_e \quad (7.6)$$

$$= \mathbf{Xb} + \sum_{u=1}^8 \mathbf{U}_u \mathbf{e}_u$$

$$\sim (\mathbf{Xb}, \mathbf{V} = \sum_{u=1}^8 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T)$$

其中 \mathbf{b} 是群体平均数和环境效应，固定效应参数向量；

\mathbf{X} 是固定效应的系数矩阵；

\mathbf{e}_A 是加性效应向量， $\mathbf{e}_A \sim (0, \sigma_A^2 \mathbf{I})$ ；

\mathbf{U}_A 是加性效应的系数矩阵；

\mathbf{e}_D 是显性效应向量， $\mathbf{e}_D \sim (0, \sigma_D^2 \mathbf{I})$ ；

\mathbf{U}_D 是显性效应的系数矩阵；

\mathbf{e}_M 是母体效应向量， $\mathbf{e}_M \sim (0, \sigma_M^2 \mathbf{I})$ ；

\mathbf{U}_M 是母体效应的系数矩阵；

\mathbf{e}_{AE} 是加性 × 环境互作效应向量， $\mathbf{e}_{AE} \sim (0, \sigma_{AE}^2 \mathbf{I})$ ；

\mathbf{U}_{AE} 是加性 × 环境互作效应的系数矩阵；

\mathbf{e}_{DE} 是显性 × 环境互作效应向量， $\mathbf{e}_{DE} \sim (0, \sigma_{DE}^2 \mathbf{I})$ ；

\mathbf{U}_{DE} 是显性 × 环境互作效应的系数矩阵；

\mathbf{e}_{ME} 是母体 × 环境互作效应向量， $\mathbf{e}_{ME} \sim (0, \sigma_{ME}^2 \mathbf{I})$ ；

\mathbf{U}_{ME} 是母体 × 环境互作效应的系数矩阵；

包括基因型与环境互作的ADM可以采用MINQUE(1)法无偏地估算各项遗传效应的方差分量及成对性状间遗传效应的协方差分量。在估算遗传方差分量和遗传协方差分量的同时，还能采用LUP法或AUP法无偏预测亲本的加性效应值和母体效应值、显性效应值，以及这些遗传效应与环境效应的互作。根据基因型与环境互作效应的预测分析结果，可以评价亲本及其组合在不同环境下的育种价值和杂种优势表现。

第二节 包括性连锁效应和母体效应的动物模型

1. ADLM 模型

动物的正反交差异除了由母体效应控制以外，还可能受到与性染色体有关的基因效应或性连锁效应(sex-linked effects)的影响。Eisen等(1966)曾提出了包括常染色体的基因加性效应、显性效应和母体效应、以及X(或Z)性染色体的基因效应(性连锁效应)的生物模型。虽然模型可以分析加性、显性、母体效应以及性连锁效应，但其应用受制于模型的二个隐含假设：哺乳动物雌性个体XX或鳞翅目昆虫雄性个体ZZ的二条性染色体都具有活性；所有组合的雌雄个体总体表现不存在遗传差异。

第一个隐含假设通常是不成立的。在每个细胞中的两条X(或Z)性染色体，其中有一条是随机失活的。朱军(1989, 1992)根据性染色体的随机失活理论，修正了Eisen等(1966)的动物遗传模型，提出了包括加性-显性-性连锁-母体效应的遗传模型(简称ADLM模型)。

如果由某遗传群体中随机抽取含 p 个纯合亲本相互杂交，产生一组杂交组合 F_1 的雌雄个体。母本 $i \times$ 父本 j 的异配性别后代(XY或ZW, $s = 1$)和同配性别后代(XX或ZZ, $s = 2$)在区组 r 中的平均表现可以用以下遗传模型表示，

$$y_{ijsk} = \mu + G_{ijs} + B_k + e_{ijsk} \quad (7.7)$$

其中 y_{ijsk} 是母本 i 与父本 j 的杂交组合的 s 性别的个体在第 k 个区组中的平均表现型值； μ 是群体平均数，固定效应； G_{ijs} 是母本 $i \times$ 父本 j 性别 s 的基因型效应值； B_k 是区组效应， $B_k \sim (0, \sigma_B^2)$ ； e_{ijsk} 是剩余效应，

$$e_{ijsk} \sim (0, \sigma_e^2)。$$

在这个线性模型中，基因型值 G_{ijs} 的定义因性别而异。对于异配性别后代(XY或ZW, $s = 1$)，

$$G_{ij1}^{XY} = A_i + A_j + D_{ij} + L_i + M_i$$

$$\text{或 } G_{ij1}^{ZW} = A_i + A_j + D_{ij} + L_j + M_i$$

而对于同配性别后代(XX或ZZ, $s = 2$)

$$G_{ij2}^{XX/ZZ} = A_i + A_j + D_{ij} + \frac{1}{2}L_i + \frac{1}{2}L_j + M_i$$

其中 A_i 或 A_j 是常染色体基因的加性效应, A_i 或 $A_j \sim (0, \sigma_A^2)$; D_{ij} 是常染色体基因的显性效应, $D_{ij} \sim (0, \sigma_D^2)$; L_i 或 L_j 是性连锁效应, L_i 或 $L_j \sim (0, \sigma_L^2)$; M_i 是母体基因效应, $M_i \sim (0, \sigma_M^2)$ 。

在许多情况下雌雄个体的总体表现存在明显的遗传差异。Zhu和Weir (1996)提出了一个改良的动物遗传模型, 该模型不受Eisen等(1966)提出的动物遗传模型中的二个隐含假设的限制。根据Zhu和Weir (1996)的模型, (7.7)中的基因型值 G_{ijs} 的定义与朱军(1989)的模型略有不同。对于异配性别后代(XY或ZW, $s = 1$),

$$G_{ij1}^{XY} = A_i + A_j + D_{ij} + L_{i1} + M_i$$

$$\text{或 } G_{ij1}^{ZW} = A_i + A_j + D_{ij} + L_{j1} + M_i$$

而对于同配性别后代(XX或ZZ, $s = 2$)

$$G_{ij2}^{XX/ZZ} = A_i + A_j + D_{ij} + \frac{1}{2}L_{i2} + \frac{1}{2}L_{j2} + M_i$$

其中 L_{i1} 或 L_{j1} 是异配性别的性连锁效应, L_{i2} 或 L_{j2} 是同配性别的性连锁效应, L_{i1} 、 L_{j1} 、 L_{i2} 或 $L_{j2} \sim (0, \sigma_L^2)$ 。

朱军(1989, 1992)的动物遗传模型以及Zhu和Weir (1996)的动物遗传模型中的性连锁效应的系数均出现了小数, 因此各项遗传效应的方差分量无法采用传统的ANOVA方法估算。

可以采用混合线性模型的统计分析方法无偏估算各项遗传方差分量及成对形状的协方差分量, 以及无偏预测各项遗传效应值。以上这两个动物模型(ADLM模型)都可由相同的矩阵形式表示,

$$\mathbf{y} = \mathbf{1}\boldsymbol{\mu} + \mathbf{U}_A\mathbf{e}_A + \mathbf{U}_D\mathbf{e}_D + \mathbf{U}_L\mathbf{e}_L + \mathbf{U}_M\mathbf{e}_M + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_\varepsilon$$

$$= \mathbf{1}\mu + \sum_{u=1}^6 \mathbf{U}_u \mathbf{e}_u$$

$$\sim (\mathbf{1}\mu, \mathbf{V} = \sum_{u=1}^6 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T)$$

其中 μ 是群体平均数固定效应， $\mathbf{1}$ 是常数1的系数向量；

\mathbf{e}_A 是加性效应向量， $\mathbf{e}_A \sim (0, \sigma_A^2 \mathbf{I})$ ；

\mathbf{U}_A 是加性效应的系数矩阵；

\mathbf{e}_D 是显性效应向量， $\mathbf{e}_D \sim (0, \sigma_D^2 \mathbf{I})$ ；

\mathbf{U}_D 是显性效应的系数矩阵；

\mathbf{e}_L 是性连锁效应向量， $\mathbf{e}_L \sim (0, \sigma_L^2 \mathbf{I})$ ；

\mathbf{U}_L 是性连锁效应的系数矩阵；

\mathbf{e}_M 是母体效应向量， $\mathbf{e}_M \sim (0, \sigma_M^2 \mathbf{I})$ ；

\mathbf{U}_M 是母体效应的系数矩阵；

\mathbf{e}_B 是区组效应向量， $\mathbf{e}_B \sim (0, \sigma_B^2 \mathbf{I})$ ；

\mathbf{U}_B 是区组效应的系数矩阵；

\mathbf{e}_E 是剩余效应向量， $\mathbf{e}_E \sim (0, \sigma_E^2 \mathbf{I})$ 。

采用MINQUE法可以无偏地估算动物遗传模型中的全部方差分量，从而可以直接估算出加性、显性、性连锁和母体方差分量： $\hat{V}_A = 2\hat{\sigma}_A^2$ ，

$$\hat{V}_D = \hat{\sigma}_D^2, \hat{V}_L = \hat{\sigma}_L^2, \hat{V}_M = \hat{\sigma}_M^2。$$

对于成对性状，采用MINQUE(1)法可以无偏地估算各项协方差分量 (Zhu和Weir, 1996)，然后可以估算加性、显性、性连锁和母体协方差分量： $\hat{C}_A = 2\hat{\sigma}_{A/A}$ ， $\hat{C}_D = \hat{\sigma}_{D/D}$ ， $\hat{C}_L = \hat{\sigma}_{L/L}$ ， $\hat{C}_M = \hat{\sigma}_{M/M}$ 。在估算遗传方差分

量和遗传协方差分量的同时，还能采用LUP法或AUP法无偏预测亲本的加性效应值和母体效应值，以及组合的显性效应值。根据遗传效应的预测分析结果，可以评价亲本及其组合的育种价值和杂种优势的表现。

Zhu和Weir(1996)运用ADLM模型分析了北卡罗来州立大学遗传系W. R. Atchley提供的28日龄小鼠体重和尾长遗传实验(7×7双列杂交、方法1)的资料。亲本和F₁表现型值分别为：体重雌性亲本平均16.94g，雄性亲本平均19.05g，雌性F₁平均16.74g，雄性F₁平均19.36g；尾长雌性亲本平均77.03mm，雄性亲本平均77.65mm，雌性F₁平均77.39mm，雄性F₁平均78.11mm。表现型值的差异表明28日龄小鼠体重和尾长都无明显的杂种优势，但体重存在明显的性别差异。遗传方差和协方差分量的估算结果列于表7.6。遗传分析结果表明，体重和尾长均存在较强的加性效应，但显性效应不显著。小鼠体重的性连锁效应和母体效应也已达到极显著水准。体重与尾长的正向加性协方差显著，这表明体重与尾长的加性效应存在正相关。

表7.6 出生28日的小鼠的体重和尾长的方差及协方差的估算结果(Zhu和Weir, 1996)

方差	体重	尾长	协方差	体重与尾长
	估计值±SE	估计值±SE		估计值±SE
V _A	3.74 ^{**} ± 0.89	6.02* ± 3.43	C _A	3.13* ± 1.51
V _D	-0.07 ± 0.27	0.32 ± 2.75	C _D	-0.17 ± 0.81
V _L	2.53 ^{**} ± 0.36	-1.03 ± 0.93	C _L	0.36 ± 0.93
V _M	2.15 ^{**} ± 0.75	4.22 ± 4.16	C _M	1.49 ± 1.57
V _e	0.68 ^{**} ± 0.23	6.55 ^{**} ± 2.01	C _e	1.75 ± 0.65

*、** 分别达到0.05显著水平或0.01极显著水平。

如果控制动物数量形状的基因效应与环境存在互作，则在一种环境下所分析的遗传效应可能导致较大的偏差。当确实存在基因型与环境的互作时，应采用包括基因型与环境的互作效应的ADLM遗传模型(Zhu和Weir, 1996)，分析动物的数量形状。

在多环境下实施动物遗传实验时，第 h 个环境内，纯系母本 $i \times$ 纯系父本 j 的异配性别后代(XY或ZW, $s = 1$)和同配性别后代(XX或ZZ, $s = 2$)在环境内区组 r 中的平均表现可以用以下遗传模型表示，

$$y_{hijrsk} = \mu + E_h + G_{ijs} + GE_{hijrs} + B_{k(h)} + e_{hijrsk} \quad (7.8)$$

其中 y_{hijrsk} 是环境 h 内母本 i 与父本 j 的 s 性别的个体在第 k 个区组中的平均表现型值；群体平均数 μ 和环境效应 E_h 都是固定效应； G_{ijs} 是母本 $i \times$ 父本 j 性别 s 的基因型效应值； GE_{hijrs} 是基因型 \times 环境互作效应值； $B_{k(h)}$ 是环境内区组效应， $B_{k(h)} \sim (0, \sigma_B^2)$ ； e_{hijrsk} 是剩余效应， $e_{hijrsk} \sim (0, \sigma_e^2)$ 。

在线性模型(7.8)中，基因型值 G_{ijs} 的定义与线性模型(7.7)同。基因型 \times 环境互作效应值 GE_{hijrs} 的定义因性别而异。根据Zhu和Weir (1996)的模型，对于异配性别后代(XY或ZW, $s = 1$)，

$$GE_{hij1}^{XY} = AE_{hi} + AE_{hj} + DE_{hij} + LE_{hi1} + ME_{hi}$$

$$\text{或 } GE_{hij1}^{ZW} = AE_{hi} + AE_{hj} + DE_{hij} + LE_{hj1} + ME_{hi}$$

而对于同配性别后代(XX或ZZ, $s = 2$)

$$GE_{hij2}^{XX/ZZ} = AE_{hi} + AE_{hj} + DE_{hij} + \frac{1}{2}LE_{hi2} + \frac{1}{2}LE_{hj2} + ME_{hi}$$

其中 AE_{hi} 或 AE_{hj} 是常染色体基因的加性 \times 环境互作效应， AE_{hi} 或 $AE_{hj} \sim (0, \sigma_{AE}^2)$ ； DE_{hij} 是常染色体基因的显性 \times 环境互作效应， $DE_{hij} \sim (0, \sigma_{DE}^2)$ ； LE_{hi1} 或 LE_{hj1} 是异配性别的性连锁 \times 环境互作效应， LE_{hi2} 或 LE_{hj2} 是同配性别的性连锁 \times 环境互作效应， LE_{hi1} 、 LE_{hj1} 、 LE_{hi2} 或 $LE_{hj2} \sim (0, \sigma_{LE}^2)$ ； ME_{hi} 是母体基因效应 \times 环境互作， $ME_{hi} \sim (0, \sigma_{ME}^2)$ 。

可以采用混合线性模型的统计分析方法无偏估算各项遗传方差分量及成对形状的协方差分量，以及无偏预测各项遗传效应值。包括基因型与环境互作的动物模型(ADLM模型)可以由以下矩阵形式表示，

$$\begin{aligned}
 \mathbf{y} &= \mathbf{X}\mathbf{b} + \mathbf{U}_A\mathbf{e}_A + \mathbf{U}_D\mathbf{e}_D + \mathbf{U}_L\mathbf{e}_L + \mathbf{U}_M\mathbf{e}_M \\
 &\quad + \mathbf{U}_{AE}\mathbf{e}_{AE} + \mathbf{U}_{DE}\mathbf{e}_{DE} + \mathbf{U}_{LE}\mathbf{e}_{LE} + \mathbf{U}_{ME}\mathbf{e}_{ME} + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_\varepsilon \\
 &= \mathbf{X}\mathbf{b} + \sum_{u=1}^{10} \mathbf{U}_u\mathbf{e}_u \\
 &\sim (\mathbf{X}\mathbf{b}, \mathbf{V} = \sum_{u=1}^{10} \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T)
 \end{aligned}$$

其中 \mathbf{b} 是群体平均数和环境效应的参数向量， \mathbf{X} 是固定效应系数向量；

\mathbf{e}_A 是加性效应向量， $\mathbf{e}_A \sim (0, \sigma_A^2 \mathbf{I})$ ；

\mathbf{U}_A 是加性效应的系数矩阵；

\mathbf{e}_D 是显性效应向量， $\mathbf{e}_D \sim (0, \sigma_D^2 \mathbf{I})$ ；

\mathbf{U}_D 是显性效应的系数矩阵；

\mathbf{e}_L 是性连锁效应向量， $\mathbf{e}_L \sim (0, \sigma_L^2 \mathbf{I})$ ；

\mathbf{U}_L 是性连锁效应的系数矩阵；

\mathbf{e}_M 是母体效应向量， $\mathbf{e}_M \sim (0, \sigma_M^2 \mathbf{I})$ ；

\mathbf{U}_M 是母体效应的系数矩阵；

\mathbf{e}_{AE} 是加性 × 环境互作效应向量， $\mathbf{e}_{AE} \sim (0, \sigma_{AE}^2 \mathbf{I})$ ；

\mathbf{U}_{AE} 是加性 × 环境互作效应的系数矩阵；

\mathbf{e}_{DE} 是显性 × 环境互作效应向量， $\mathbf{e}_{DE} \sim (0, \sigma_{DE}^2 \mathbf{I})$ ；

\mathbf{U}_{DE} 是显性 × 环境互作效应的系数矩阵；

\mathbf{e}_{LE} 是性连锁 × 环境互作效应向量， $\mathbf{e}_{LE} \sim (0, \sigma_{LE}^2 \mathbf{I})$ ；

\mathbf{U}_{LE} 是性连锁 × 环境互作效应的系数矩阵；

\mathbf{e}_M 是母体 × 环境互作效应向量， $\mathbf{e}_M \sim (0, \sigma_M^2 \mathbf{I})$ ；

\mathbf{U}_{ME} 是母体 × 环境互作效应的系数矩阵；

\mathbf{e}_B 是环境内的区组效应向量， $\mathbf{e}_B \sim (0, \sigma_B^2 \mathbf{I})$ ；

\mathbf{U}_B 是环境内的区组效应系数矩阵；

\mathbf{e}_ε 是剩余效应向量, $\mathbf{e}_\varepsilon \sim (0, \sigma_\varepsilon^2 \mathbf{I})$ 。

采用MINQUE(1)法可以无偏地估算动物遗传模型中的全部方差分量, 并可以直接估算出加性 × 环境互作、显性 × 环境互作、性连锁 × 环境互作和母体 × 环境互作的方差分量: $\hat{V}_{AE} = 2\hat{\sigma}_{AE}^2$, $\hat{V}_{DE} = \hat{\sigma}_{DE}^2$, $\hat{V}_{LE} = \sigma_{LE}^2$, $\hat{V}_{ME} = \hat{\sigma}_{ME}^2$ 。

对于成对性状, 采用MINQUE(1)法也可以无偏地估算各项协方差分量 (Zhu和Weir, 1995), 然后可以估算加性 × 环境互作、显性 × 环境互作、性连锁 × 环境互作和母体 × 环境互作的协方差分量: $\hat{C}_{AE} = 2\hat{\sigma}_{AE/AE}$, $\hat{C}_D = \hat{\sigma}_{DE/DE}$, $\hat{C}_L = \hat{\sigma}_{LE/LE}$, $\hat{C}_M = \hat{\sigma}_{ME/ME}$ 。在估算遗传方差分量和遗传协方差分量的同时, 还能采用LUP法或AUP法无偏预测亲本的加性 × 环境互作效应值和母体 × 环境互作效应值, 以及组合的显性 × 环境互作效应值。根据基因型 × 环境互作效应的预测分析结果, 可以评价亲本及其组合在不同环境中的育种价值和杂种优势的稳定性表现。

2. ADLAmDm 模型

Zhu和Weir(1996)提出的ADLM动物遗传模型, 可以分析亲本和 F_1 的双列杂交遗传资料, 把性连锁效应和母体效应从常染色体核遗传效应中区分出来, 但还无法进一步将母体效应区分为母体加性效应和母体显性效应两个分量。Zhu和Duan(1994)曾提出了一个动物遗传模型, 可以把母体效应(M)再分解为母体加性效应(Am)和母体显性效应(Dm)。朱军和段家龙(1994)的遗传交配设计是在Zhu和Weir(1996)的遗传交配设计中增加组合的 F_2 个体。动物的 F_2 个体由 F_1 雌雄个体杂交而获得。

当不存在基因型与环境互作效应时, 可以分析一年一地的实验资料。由某遗传群体中随机抽取含 p 个纯合亲本相互杂交, 产生一组杂交组合 F_1 的

雌雄个体， F_1 雌雄个体相互杂交产生 F_2 的雌雄个体。母本 $i \times$ 父本 j 的第 k 种世代($k = 0$ 为亲本、 $k = 1$ 为 F_1 、 $k = 2$ 为 F_2)的异配性别个体(XY或ZW, $s = 1$)和同配性别个体(XX或ZZ, $s = 2$)在区组 l 中的平均表现可以用以下遗传模型表示，

$$y_{ijkl} = \mu + G_{ijks} + B_l + e_{ijkl} \quad (7.9)$$

其中 y_{ijkl} 是基因型 G_{ijks} 在第 l 个区组中的平均表现型值； μ 是群体平均数，固定效应； G_{ijks} 是母本 $i \times$ 父本 j 的第 k 种世代性别 s 的基因型效应值； B_l 是区组效应， $B_l \sim (0, \sigma_B^2)$ ； e_{ijkl} 是剩余效应， $e_{ijkl} \sim (0, \sigma_e^2)$ 。

在这个线性模型中，基因型值 G_{ijks} 的定义因性别而异。对于亲本($k = 0$)或 F_1 ($k = 1$)，异配性别后代(XY或ZW, $s = 1$)，

$$G_{ijk1}^{XY} = A_i + A_j + D_{ij} + L_{i1} + 2Am_i + Dm_{ij}$$

$$\text{或 } G_{ijk1}^{ZW} = A_i + A_j + D_{ij} + L_{j1} + 2Am_i + Dm_{ij}$$

而同配性别后代(XX或ZZ, $s = 2$)，

$$G_{ijk2}^{XX/ZZ} = A_i + A_j + D_{ij} + \frac{1}{2}L_{i2} + \frac{1}{2}L_{j2} + 2Am_i + Dm_{ij}$$

对于 F_2 ($k = 2$)，异配性别后代(XY或ZW, $s = 1$)，

$$G_{ijk1}^{XY/ZW} = A_i + A_j + \frac{1}{4}D_{ii} + \frac{1}{4}D_{jj} + \frac{1}{2}D_{ij} + \frac{1}{2}L_{i1} + \frac{1}{2}L_{j1} + Am_i + Am_j + Dm_{ij}$$

同配性别后代(XX或ZZ, $s = 2$)，

$$G_{ijk2}^{XX} = A_i + A_j + \frac{1}{4}D_{ii} + \frac{1}{4}D_{jj} + \frac{1}{2}D_{ij} + \frac{3}{4}L_{i2} + \frac{1}{4}L_{j2} + Am_i + Am_j + Dm_{ij}$$

$$\text{或 } G_{ijk2}^{ZZ} = A_i + A_j + \frac{1}{4}D_{ii} + \frac{1}{4}D_{jj} + \frac{1}{2}D_{ij} + \frac{1}{4}L_{i2} + \frac{3}{4}L_{j2} + Am_i + Am_j + Dm_{ij}$$

其中 A_i 或 A_j 是常染色体基因的加性效应， A_i 或 $A_j \sim (0, \sigma_A^2)$ ； D_{ij} 是常染色体基因的显性效应， $D_{ij} \sim (0, \sigma_D^2)$ ； L_{i1} 或 L_{j1} 是异配性别的性连锁效应， L_{i2} 或 L_{j2} 是同配性别的性连锁效应， L_{i1} 、 L_{j1} 、 L_{i2} 或 $L_{j2} \sim (0, \sigma_L^2)$ ； Am_i 是母体基因加性效应， Am_i 或 $Am_j \sim (0, \sigma_{Am}^2)$ ； Dm_{ij} 是母体基因的显性效应， $Dm_{ij} \sim (0, \sigma_{Dm}^2)$ 。在该遗传模型中常染色体基因的加性和显性效应与母体基因的

加性和显性效应不完全是相互独立的，可能存在不同程度的协方差：加性协

方差 $\text{Cov}(A_i, Am_i) = \text{Cov}(A_j, Am_j) = \sigma_{A.Am}$, 显性协方差 $\text{Cov}(D_{ij}, Dm_{ij}) = \sigma_{D.Dm}$ 。

由于模型中遗传效应的系数出现了小数，并且遗传效应间不是相互独立的，因此各项遗传效应的方差分量无法采用传统的ANOVA方法估算。可以采用混合线性模型的统计分析方法无偏估算各项遗传方差分量及成对形状的协方差分量，以及无偏预测各项遗传效应值。ADLAmDm模型可由以下矩阵形式表示，

$$\begin{aligned} \mathbf{y} &= \mathbf{1}\mu + \mathbf{U}_A\mathbf{e}_A + \mathbf{U}_D\mathbf{e}_D + \mathbf{U}_L\mathbf{e}_L + \mathbf{U}_{Am}\mathbf{e}_{Am} + \mathbf{U}_{Dm}\mathbf{e}_{Dm} + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_\varepsilon \\ &= \mathbf{1}\mu + \sum_{u=1}^7 \mathbf{U}_u\mathbf{e}_u \end{aligned}$$

具有方差和协方差矩阵

$$\begin{aligned} \text{Var}(\mathbf{y}) &= \mathbf{V} \\ &= \sigma_A^2 \mathbf{U}_A \mathbf{U}_A^T + \sigma_D^2 \mathbf{U}_D \mathbf{U}_D^T + \sigma_L^2 \mathbf{U}_L \mathbf{U}_L^T + \sigma_{Am}^2 \mathbf{U}_{Am} \mathbf{U}_{Am}^T \\ &\quad + \sigma_{Dm}^2 \mathbf{U}_{Dm} \mathbf{U}_{Dm}^T + \sigma_B^2 \mathbf{U}_B \mathbf{U}_B^T + \sigma_{A.Am} (\mathbf{U}_A \mathbf{U}_{Am}^T + \mathbf{U}_{Am} \mathbf{U}_A^T) \\ &\quad + \sigma_{D.Dm} (\mathbf{U}_D \mathbf{U}_{Dm}^T + \mathbf{U}_{Dm} \mathbf{U}_D^T) + \sigma_e^2 \mathbf{I} \end{aligned}$$

其中 μ 是群体平均数固定效应， $\mathbf{1}$ 是常数1的系数向量；

\mathbf{e}_A 是加性效应向量， $\mathbf{e}_A \sim (0, \sigma_A^2 \mathbf{I})$ ；

\mathbf{U}_A 是加性效应的系数矩阵；

\mathbf{e}_D 是显性效应向量， $\mathbf{e}_D \sim (0, \sigma_D^2 \mathbf{I})$ ；

\mathbf{U}_D 是显性效应的系数矩阵；

\mathbf{e}_L 是性连锁效应向量， $\mathbf{e}_L \sim (0, \sigma_L^2 \mathbf{I})$ ；

\mathbf{U}_L 是性连锁效应的系数矩阵；

\mathbf{e}_{Am} 是母体加性效应向量， $\mathbf{e}_{Am} \sim (0, \sigma_{Am}^2 \mathbf{I})$ ；

\mathbf{U}_{Am} 是母体加性效应的系数矩阵；

\mathbf{e}_{Dm} 是母体显性效应向量, $\mathbf{e}_{Dm} \sim (0, \sigma_{Dm}^2 \mathbf{I})$;

\mathbf{U}_{Dm} 是母体显性效应的系数矩阵;

\mathbf{e}_B 是区组效应向量, $\mathbf{e}_B \sim (0, \sigma_B^2 \mathbf{I})$;

\mathbf{U}_B 是区组效应的系数矩阵;

\mathbf{e}_ε 是剩余效应向量, $\mathbf{e}_\varepsilon \sim (0, \sigma_\varepsilon^2 \mathbf{I})$ 。

采用MINQUE(0/1)法(朱军, 1992; Zhu和Weir, 1994a)可以无偏地估算遗传模型中的全部方差分量和协方差分量。MINQUE(0/1)法是一种MINQUE方法, 它设方差分量的先验值为1、协方差分量的先验值为0。

对于ADLAmDm遗传模型, 表现型方差的估算公式为:

$$\begin{aligned} V_P &= V_G + V_e \\ &= V_A + V_D + V_L + V_{Am} + V_{Dm} + 2C_{A.Am} + 2C_{D.Dm} + V_e \end{aligned}$$

其中基因型方差 V_G 的各项分量定义为: V_A 是加性方差、 V_D 是显性方差、 V_L 是性连锁方差、 V_{Am} 是母体加性方差、 V_{Dm} 是母体显性方差, $C_{A.Am}$ 是加性效应与母体加性效应的协方差、 $C_{D.Dm}$ 是显性效应与母体显性效应的协方差, V_e 是剩余效应方差。

各项遗传方差分量和协方差分量的估算因遗传世代而异:

$$\begin{aligned} V_P(F_1) &= 2\sigma_A^2 + \sigma_D^2 + \sigma_L^2 + 4\sigma_{Am}^2 + \sigma_{Dm}^2 + 4\sigma_{A.Am} + \sigma_e^2 \\ V_P(F_2) &= 2\sigma_A^2 + \frac{3}{8}\sigma_D^2 + \sigma_L^2 + 2\sigma_{Am}^2 + \sigma_{Dm}^2 + 4\sigma_{A.Am} + \sigma_{D.Dm} + \sigma_e^2 \end{aligned}$$

对于成对性状, 采用MINQUE(0/1)法也可以无偏地估算各项协方差分量(朱军, 1992; Zhu和Weir, 1994a), 然后可以估算成对性状表现型协方差

$$\begin{aligned} C_P &= C_G + C_e \\ &= C_A + C_D + C_L + C_{Am} + C_{Dm} + 2C_{A/Am} + 2C_{D/Dm} + C_e \end{aligned}$$

其中基因型协方差 C_G 的各项分量定义为: C_A 是加性协方差、 C_D 是协显性方差、 C_L 是性连锁协方差、 C_{Am} 是母体加性协方差、 C_{Dm} 是母体显性协方差, $C_{A/Am}$

是一个性状的加性效应与另一个性状的母体加性效应的协方差、 $C_{D/Dm}$ 是一个性状的显性效应与另一个性状的母体显性效应的协方差， C_e 是剩余效应协方差。

成对性状遗传协方差分量的估算公式是：

$$C_P(F_1) = 2\sigma_{A/A} + \sigma_{D/D} + \sigma_{L/L} + 4\sigma_{Am/Am} + \sigma_{Dm/Dm} + 4\sigma_{A/Am} + \sigma_{e/e}$$

$$C_P(F_2) = 2\sigma_{A/A} + \frac{3}{8}\sigma_{D/D} + \sigma_{L/L} + 2\sigma_{Am/Am} + \sigma_{Dm/Dm} + 4\sigma_{A/Am} + \sigma_{D/Dm} + \sigma_{ee}$$

在估算遗传方差分量和遗传协方差分量的同时，还能采用LUP法或AUP法无偏预测亲本的加性效应值和母体加性效应值，以及组合的显性效应值和母体显性效应值。根据遗传效应的预测分析结果，可以评价亲本及其组合的育种价值和杂种优势的表现。

段家龙和朱军(1994)采用ADLAmDm遗传模型，分析8个家蚕品种(明晖、72秋、苏4、苏12、大造、东海、浙一和华1)不完全双列杂交的亲本、 F_1 和 F_2 的试验资料。分析的家蚕茧质性状是全茧量、茧层量和茧层率。采用MINQUE(0/1)法估算性状的各项方差和协方差分量(朱军, 1992)，在估算成对性状遗传协方差的基础上进一步估算了各项遗传相关系数。用AUP法(朱军, 1993a)预测了各性状的基因效应值。采用Jackknife的方法(朱军, 1992, 1993a)，计算各项遗传参数的估计值(或预测值)及其标准误。以每个基因型的世代平均数作为Jackknife的重复抽样单位。采用 t 测验对遗传参数作统计检验。

由表7.7可见，全茧量同时受到直接效应和母体效应的控制，其中直接显性的作用大于直接加性的作用，母体加性的作用大于母体显性的作用。茧层量则主要受直接效应控制，直接显性的作用大于直接加性的作用。控制茧层率的主要遗传效应是母体加性效应，其次为直接加性效应。3个性状的性连锁效应都达到极显著水准，但方差分量值较小。直接效应与母体效应一般表现负的相关性。剩余效应方差的值较小，表明除所分析的遗传效应外，其它效应及误差的干扰较小。

表7.7. 家蚕茧质性状方差分量的估计值及其标准误

参数	全茧量	茧层量	茧层率
	估计值 ± 标准误	估计值 ± 标准误	估计值 ± 标准误
V_A	970.35** ± 138.48	31.34** ± 4.52	3.84** ± 0.72
V_D	1253.03** ± 231.82	60.59** ± 11.76	0.96** ± 0.24
V_L	259.60** ± 20.36	1.48** ± 0.28	2.24** ± 0.23
V_{Am}	1822.69** ± 300.76	0.00 ± 0.00	8.63** ± 2.14
V_{Dm}	232.89** ± 46.33	10.94** ± 2.22	0.99** ± 0.20
$C_{A.Am}$	-711.68** ± 129.67	0.00 ± 0.00	-2.76** ± 0.87
$C_{D.Dm}$	-168.54** ± 37.41	-7.75** ± 1.81	-0.18** ± 0.05
V_e	226.56** ± 113.41	4.90** ± 1.05	0.88** ± 0.15

** 达到0.01极显著水平。

相关分析结果(表7.8)表明,除母体加性效应以外,全茧量和茧层量在其它遗传效应上均存在极显著的正相关。因此,提高杂交组合或选系的全茧量,可以导致茧层量的增加。全茧量和茧层率的性连锁效应和母体加性效应之间存在极显著的负相关,选择全茧量高的选系,可能导致茧层率的降低。由于全茧量和茧层率的直接显性效应之间存在极显著的正相关,可望选配到全茧量和茧层率均高的杂交组合。茧层量和茧层率之间也存在极显著的正向直接和母体显性相关,因此也可能同时改良杂交组合的茧层量和茧层率。

表7.8. 家蚕茧质性状遗传分量的相关系数估计值及其标准误.

参数	全茧量&茧层量	全茧量&茧层率	茧层量&茧层率
	估计值 ± 标准误	估计值 ± 标准误	估计值 ± 标准误
r_A	0.54** ± 0.06	-0.13 ± 0.07	0.09 ± 0.07
r_D	0.54** ± 0.06	0.17** ± 0.06	0.33** ± 0.06
r_L	0.43** ± 0.07	-0.72** ± 0.05	-0.37** ± 0.06
r_{Am}	0.00 ± 0.00	-0.26** ± 0.08	0.00 ± 0.00
r_{Dm}	0.39** ± 0.06	-0.04 ± 0.07	0.18** ± 0.07
$r_{A/Am}$	0.00 ± 0.00	-0.05 ± 0.03	0.00 ± 0.00
$r_{D/Dm}$	0.22** ± 0.06	0.02 ± 0.05	0.16** ± 0.05
r_e	0.46** ± 0.07	-0.20** ± 0.08	0.27 ± 0.07

** 达到0.01极显著水平。

表7.9列出了亲本及组合的遗传效应预测值。对于全茧量而言,亲本P3和P4具有极显著的正向加性效应和雌性性连锁效应,亲本P6和P7具有极显著的正向母体加性和显性效应和雌性性连锁效应。因此,以亲本P6和P7为母本和亲本P3和P4为父本的 F_1 雌性个体的产茧量将较高。由于性染色体的效应 $\sum_{i=1}^8 L_{i1}$ 为较大的正值,这表明家蚕的雌性个体一般表现较高的产茧量。

全茧量直接效应的标准化显性方向为极显著的正值($\hat{\Delta}_O = 6.19^{**} \pm 1.04$),而母体效应的标准化显性方向为极显著的负值($\hat{\Delta}_m = -10.08^{**} \pm 1.71$),这表明家蚕的 F_1 在产茧量性状上存在极强的杂种优势,但是 F_2 则将表现严重的近交衰退。

亲本P2、P6、P7和P8的茧层量具有极显著的正向母体加性。因此,以这几个亲为母本的 F_1 个体的茧层量将较高。由于茧层量的 $\sum_{i=1}^8 L_{i1}$ 为较大的正值,这表明家蚕的雌性个体一般表现较高的茧层量。茧层量直接效应的标准化显性方向为极显著的正值($\hat{\Delta}_O = 5.38^{**} \pm 1.05$),而母体效应的标准化显性方向为极显著的负值($\hat{\Delta}_m = -9.10^{**} \pm 1.45$),这表明家蚕的 F_1 在茧层量性状上存在极强的杂种优势,而 F_2 则将表现严重的近交衰退。

茧层率的直接效应和母体效应的的标准化显性方向均不显著($\hat{\Delta}_O = 1.17 \pm 1.45$, $\hat{\Delta}_m = 0.48 \pm 1.37$),这表明家蚕 F_1 的茧层率性状不存在明显的杂种优势。由于茧层率的 $\sum_{i=1}^8 L_{i1}$ 为显著的负值,家蚕的雌性个体一般表现较低的茧层率。亲本P4具有极显著的负向加性效应和正向母体加性效应,而亲本P7却表现相反,具有显著的正向加性效应和负向母体加性效应。因此选用P4作为母本而P7作为父本有利于提高 F_1 的茧层率。

表7.9. 家蚕茧质性状遗传效应的预测值.

效应	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$
全茧量								
A_i	4.06	9.13	26.03**	20.34**	-19.60**	-10.39	-20.25	-9.37
$D_{i,i}$	-45.76	-12.85	-77.54	-56.69	-31.12	-77.18**	-81.90**	-57.46**
$L_{i,1}$	0.19	20.01**	21.98**	22.71**	18.49**	4.92	2.23	20.97*
$L_{i,2}$	-9.41	-13.73**	-8.35	-6.25	-13.14**	-25.78**	-27.77**	-7.10
Am_i	-10.40	-14.70*	-41.36**	-30.04	10.54	28.44**	41.32**	16.16
$Dm_{i,i}$	2.28	21.16**	15.09	13.87	29.85*	47.68**	55.77**	25.73*
茧层量								
A_i	3.53	0.99	3.15	1.57	-4.93*	-1.71	-2.61	0.01
$D_{i,i}$	-15.20	1.02	-17.06	-8.16	-3.16	-14.41**	-12.74**	-17.58**
$L_{i,1}$	1.31*	0.68	1.52*	0.33	0.66	0.46	0.06	0.05
$L_{i,2}$	0.15	-0.93	0.78	-0.87	-0.25	-0.92	-0.99*	-2.04*
Am_i	—	—	—	—	—	—	—	—
$Dm_{i,i}$	2.34	5.59**	4.03	4.33	2.19	13.99**	9.26*	9.03**
茧层率								
A_i	1.38**	-0.87	-0.27	-1.42**	-1.44**	0.55	1.67*	0.38
$D_{i,i}$	-0.91	1.64	-2.09	0.46	0.83	-0.46	0.62	-0.96
$L_{i,1}$	-1.30	-1.79**	-0.67	-1.73**	-1.93**	-1.00	-1.72**	-0.93
$L_{i,2}$	0.22	1.25**	2.70**	0.76	1.20*	1.67**	1.75**	1.52**
Am_i	-0.56	2.16**	0.02	1.37**	-0.92	0.01	-2.12*	0.03
$Dm_{i,i}$	0.25	0.02	-0.01	0.29	-2.03*	1.89*	-1.40	-0.29

*、** 分别达到0.05显著水平或0.01极显著水平.

如果基因效应与环境存在交互, 则应采用包括基因型与环境的交互效应的ADLAmDm遗传模型(Zhu和Duan, 1994), 分析动物的数量形状的母体遗传效应。在多环境下实施动物遗传实验时, 第*h*个环境内, 纯系母本*i* × 纯系父本*j*的世代*k*的异配性别后代(XY或ZW, $s = 1$)和同配性别后代(XX或ZZ,

$s = 2$) 在环境内区组 l 中的平均表现可以用以下遗传模型表示，

$$y_{hijkl} = \mu + E_h + G_{ijks} + GE_{hijks} + B_{l(h)} + e_{hijkl} \quad (7.10)$$

其中 y_{hijkl} 是环境 h 内母本 i 与父本 j 的世代 k 的 s 性别个体在第 l 个区组中的平均表现型值；群体平均数 μ 和环境效应 E_h 都是固定效应； G_{ijks} 是基因型值； GE_{hijks} 是基因型 \times 环境互作效应值； $B_{k(h)}$ 是环境内的区组效应，

$$B_{k(h)} \sim (0, \sigma_B^2) ; e_{hijkl} \text{ 是剩余效应, } e_{hijkl} \sim (0, \sigma_e^2)。$$

在线性模型(7.10)中，基因型值 G_{ijks} 的定义与线性模型(7.9)同。基因型 \times 环境互作效应值 GE_{hijks} 的定义因世代和性别而异。对于亲本 ($k = 0$) 或 F_1 ($k = 1$)，异配性别后代(XY或ZW, $s = 1$)的互作效应，

$$GE_{hijk1}^{XY} = AE_{hi} + AE_{hj} + DE_{hij} + LE_{hi1} + 2AmE_{hi} + DmE_{hii}$$

$$\text{或 } GE_{hijk1}^{ZW} = AE_{hi} + AE_{hj} + DE_{hij} + LE_{hj1} + 2AmE_{hi} + DmE_{hii}$$

而同配性别后代(XX或ZZ, $s = 2$) 的互作效应，

$$GE_{hijk2}^{XX/ZZ} = AE_{hi} + AE_{hj} + DE_{hij} + \frac{1}{2}LE_{hi2} + \frac{1}{2}LE_{hj2} + 2AmE_{hi} + DmE_{hii}$$

对于 F_2 ($k = 2$)，异配性别后代(XY或ZW, $s = 1$)的互作效应，

$$GE_{hijk1}^{XY/ZW} = AE_{hi} + AE_{hj} + \frac{1}{4}DE_{hii} + \frac{1}{4}DE_{hjj} + \frac{1}{2}DE_{hij} + \frac{1}{2}LE_{hi1} + \frac{1}{2}LE_{hj1} \\ + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

同配性别后代(XX或ZZ, $s = 2$)的互作效应，

$$GE_{hijk2}^{XX} = AE_{hi} + AE_{hj} + \frac{1}{4}DE_{hii} + \frac{1}{4}DE_{hjj} + \frac{1}{2}DE_{hij} + \frac{3}{4}LE_{hi2} + \frac{1}{4}LE_{hj2} \\ + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

$$\text{或 } GE_{hijk2}^{ZZ} = AE_{hi} + AE_{hj} + \frac{1}{4}DE_{hii} + \frac{1}{4}DE_{hjj} + \frac{1}{2}DE_{hij} + \frac{1}{4}LE_{hi2} + \frac{3}{4}LE_{hj2} \\ + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

其中 AE_{hi} 或 AE_{hj} 是常染色体基因的加性 \times 环境互作效应， AE_{hi} 或

$$AE_{hj} \sim (0, \sigma_{AE}^2) ; DE_{hij} \text{ 是常染色体基因的显性 } \times \text{ 环境互作效应，}$$

$$DE_{hij} \sim (0, \sigma_{DE}^2) ; LE_{hi1} \text{ 或 } LE_{hj1} \text{ 是异配性别的性连锁 } \times \text{ 环境互作效应，}$$

$$LE_{hi2} \text{ 或 } LE_{hj2} \text{ 是同配性别的性连锁 } \times \text{ 环境互作效应， } LE_{hi1}、LE_{hj1}、LE_{hi2}$$

或 $LE_{hj2} \sim (0, \sigma_{LE}^2)$; AmE_{hi} 或 AmE_{hj} 是母体基因加性 \times 环境互作效应 ,
 AmE_{hi} 或 $AmE_{hj} \sim (0, \sigma_{AmE}^2)$; DmE_{hij} 是母体基因的显性 \times 环境互作效
 应 , $DmE_{hij} \sim (0, \sigma_{DmE}^2)$ 。还可能存在互作效应间的协方差 : 加性 \times 环境
 互作协方差 $Cov(AE_{hi}, AmE_{hi}) = Cov(AE_{hj}, AmE_{hj}) = \sigma_{AE.AmE}$, 显性 \times 环
 境互作协方差 $Cov(DE_{hij}, DmE_{hij}) = \sigma_{D.Dm}$ 。

可以采用混合线性模型的统计分析方法无偏估算各项遗传方差分量及
 成对形状的协方差分量 , 以及无偏预测各项遗传效应值。包括基因型与环境
 互作的ADLAmDm模型可由以下矩阵形式表示 ,

$$\begin{aligned} \mathbf{y} &= \mathbf{Xb} + \mathbf{U}_A \mathbf{e}_A + \mathbf{U}_D \mathbf{e}_D + \mathbf{U}_L \mathbf{e}_L + \mathbf{U}_{Am} \mathbf{e}_{Am} + \mathbf{U}_{Dm} \mathbf{e}_{Dm} \\ &+ \mathbf{U}_{AE} \mathbf{e}_{AE} + \mathbf{U}_{DE} \mathbf{e}_{DE} + \mathbf{U}_{LE} \mathbf{e}_{LE} + \mathbf{U}_{AmE} \mathbf{e}_{AmE} + \mathbf{U}_{DmE} \mathbf{e}_{DmE} + \mathbf{U}_B \mathbf{e}_B + \mathbf{e}_\varepsilon \\ &= \mathbf{Xb} + \sum_{u=1}^{12} \mathbf{U}_u \mathbf{e}_u \end{aligned}$$

具有方差和协方差矩阵

$$\begin{aligned} \text{Var}(\mathbf{y}) &= \mathbf{V} \\ &= \sigma_A^2 \mathbf{U}_A \mathbf{U}_A^T + \sigma_D^2 \mathbf{U}_D \mathbf{U}_D^T + \sigma_L^2 \mathbf{U}_L \mathbf{U}_L^T + \sigma_{Am}^2 \mathbf{U}_{Am} \mathbf{U}_{Am}^T + \sigma_{Dm}^2 \mathbf{U}_{Dm} \mathbf{U}_{Dm}^T \\ &+ \sigma_{AE}^2 \mathbf{U}_{AE} \mathbf{U}_{AE}^T + \sigma_{DE}^2 \mathbf{U}_{DE} \mathbf{U}_{DE}^T + \sigma_{LE}^2 \mathbf{U}_{LE} \mathbf{U}_{LE}^T + \sigma_{AmE}^2 \mathbf{U}_{AmE} \mathbf{U}_{AmE}^T \\ &+ \sigma_{DmE}^2 \mathbf{U}_{DmE} \mathbf{U}_{DmE}^T + \sigma_B^2 \mathbf{U}_B \mathbf{U}_B^T + \sigma_{A.Am} (\mathbf{U}_A \mathbf{U}_{Am}^T + \mathbf{U}_{Am} \mathbf{U}_A^T) \\ &+ \sigma_{D.Dm} (\mathbf{U}_D \mathbf{U}_{Dm}^T + \mathbf{U}_{Dm} \mathbf{U}_D^T) + \sigma_{AE.AmE} (\mathbf{U}_{AE} \mathbf{U}_{AmE}^T + \mathbf{U}_{AmE} \mathbf{U}_{AE}^T) \\ &+ \sigma_{DE.DmE} (\mathbf{U}_{DE} \mathbf{U}_{DmE}^T + \mathbf{U}_{DmE} \mathbf{U}_{DE}^T) + \sigma_e^2 \mathbf{I} \end{aligned}$$

其中 \mathbf{b} 是群体平均数和环境效应的参数向量 , \mathbf{X} 是固定固定效应系数向量 ;

\mathbf{e}_A 是加性效应向量 , $\mathbf{e}_A \sim (0, \sigma_A^2 \mathbf{I})$;

\mathbf{U}_A 是加性效应的系数矩阵 ;

\mathbf{e}_D 是显性效应向量 , $\mathbf{e}_D \sim (0, \sigma_D^2 \mathbf{I})$;

\mathbf{U}_D 是显性效应的系数矩阵 ;

\mathbf{e}_L 是性连锁效应向量 , $\mathbf{e}_L \sim (0, \sigma_L^2 \mathbf{I})$;

\mathbf{U}_L 是性连锁效应的系数矩阵 ;

\mathbf{e}_{Am} 是母体加性效应向量 , $\mathbf{e}_{Am} \sim (0, \sigma_{Am}^2 \mathbf{I})$;

U_{Am} 是母体加性效应的系数矩阵;

e_{Dm} 是母体显性效应向量, $e_{Dm} \sim (0, \sigma_{Dm}^2 \mathbf{I})$;

U_{Dm} 是母体显性效应的系数矩阵;

e_{AE} 是加性 \times 环境互作效应向量, $e_{AE} \sim (0, \sigma_{AE}^2 \mathbf{I})$;

U_{AE} 是加性 \times 环境互作效应的系数矩阵;

e_{DE} 是显性 \times 环境互作效应向量, $e_{DE} \sim (0, \sigma_{DE}^2 \mathbf{I})$;

U_{DE} 是显性 \times 环境互作效应的系数矩阵;

e_{LE} 是性连锁 \times 环境互作效应向量, $e_{LE} \sim (0, \sigma_{LE}^2 \mathbf{I})$;

U_{LE} 是性连锁 \times 环境互作效应的系数矩阵;

e_{AmE} 是母体加性 \times 环境互作效应向量, $e_{AmE} \sim (0, \sigma_{AmE}^2 \mathbf{I})$;

U_{AmE} 是母体加性 \times 环境互作效应的系数矩阵;

e_{DmE} 是母体显性 \times 环境互作效应向量, $e_{DmE} \sim (0, \sigma_{DmE}^2 \mathbf{I})$;

U_{DmE} 是母体显性 \times 环境互作效应的系数矩阵;

e_B 是环境内的区组效应向量, $e_B \sim (0, \sigma_B^2 \mathbf{I})$;

U_B 是环境内的区组效应系数矩阵;

e_ε 是剩余效应向量, $e_\varepsilon \sim (0, \sigma_\varepsilon^2 \mathbf{I})$ 。

采用MINQUE(0/1)法可以无偏地估算动物遗传模型中的全部方差分量和协方差分量。对于包括基因型 \times 环境互作的ADLAmDm遗传模型,表现型方差的估算公式为:

$$\begin{aligned} V_P &= V_G + V_{GE} + V_e \\ &= (V_A + V_D + V_L + V_{Am} + V_{Dm} + 2C_{A.Am} + 2C_{D.Dm}) \\ &\quad + (V_{AE} + V_{DE} + V_{LE} + V_{AmE} + V_{DmE} + 2C_{AE.AmE} + 2C_{DE.DmE}) + V_e \end{aligned}$$

其中遗传效应与环境互作方差 V_{GE} 各项分量的定义为: V_{AE} 是加性 \times 环境互作方差、 V_{DE} 是显性 \times 环境互作方差、 V_{LE} 是性连锁 \times 环境互作方差、 V_{AmE} 是母体加性 \times 环境互作方差、 V_{DmE} 是母体显性 \times 环境互作方差、 $C_{AE.AmE}$ 是加性 \times 环境互作效应与母体加性 \times 环境互作效应的协方差、 $C_{DE.DmE}$ 是显性 \times 环境互作效应与母体显性 \times 环境互作效应的协方差。

各项方差分量和协方差分量的估算因遗传世代而异:

$$\begin{aligned} V_P(F_1) &= 2\sigma_A^2 + \sigma_D^2 + \sigma_L^2 + 4\sigma_{Am}^2 + \sigma_{Dm}^2 + 4\sigma_{A.Am} \\ &\quad + 2\sigma_{AE}^2 + \sigma_{DE}^2 + \sigma_{LE}^2 + 4\sigma_{AmE}^2 + \sigma_{DmE}^2 + 4\sigma_{AE.AmE} + \sigma_e^2 \\ V_P(F_2) &= 2\sigma_A^2 + \frac{3}{8}\sigma_D^2 + \sigma_L^2 + 2\sigma_{Am}^2 + \sigma_{Dm}^2 + 4\sigma_{A.Am} + \sigma_{D.Dm} \\ &\quad + 2\sigma_{AE}^2 + \frac{3}{8}\sigma_{DE}^2 + \sigma_{LE}^2 + 2\sigma_{AmE}^2 + \sigma_{DmE}^2 + 4\sigma_{AE.AmE} + \sigma_{DE.DmE} + \sigma_e^2 \end{aligned}$$

对于成对性状，采用MINQUE(0/1)法也可以无偏地估算各项协方差分量(朱军，1992；Zhu和Weir，1994a)，然后可以估算成对性状表现型协方差

$$C_P = C_G + C_{GE} + C_e$$

$$= (C_A + C_D + C_L + C_{Am} + C_{Dm} + 2C_{A/Am} + 2C_{D/Dm})$$

$$+ (C_{AE} + C_{DE} + C_{LE} + C_{AmE} + C_{DmE} + 2C_{AE/AmE} + 2C_{DE/DmE}) + C_e$$

其中遗传效应与环境互作协方差 C_{Ge} 各项分量的定义为： C_{AE} 是加性×环境互作协方差、 C_{DE} 是显性×环境互作协方差、 C_{LE} 是性连锁×环境互作协方差、 C_{AmE} 是母体加性×环境互作协方差、 C_{DmE} 是母体显性×环境互作协方差， $C_{AE/AmE}$ 是一个性状的加性×环境互作效应与另一个性状的母体加性×环境互作效应的协方差、 $C_{DE/DmE}$ 是一个性状的显性×环境互作效应与另一个性状的母体显性×环境互作效应的协方差。

成对性状遗传协方差分量的估算公式是：

$$C_P(F_1) = 2\sigma_{A/A} + \sigma_{D/D} + \sigma_{L/L} + 4\sigma_{Am/Am} + \sigma_{Dm/Dm} + 4\sigma_{A/Am}$$

$$+ 2\sigma_{AE/AE} + \sigma_{DE/DE} + \sigma_{LE/LE} + 4\sigma_{AmE/AmE} + \sigma_{DmE/DmE}$$

$$+ 4\sigma_{AE/AmE} + \sigma_{e/e}$$

$$C_P(F_2) = 2\sigma_{A/A} + \frac{3}{8}\sigma_{D/D} + \sigma_{L/L} + 2\sigma_{Am/Am} + \sigma_{Dm/Dm} + 4\sigma_{A/Am} + \sigma_{D/Dm}$$

$$+ 2\sigma_{AE/AE} + \frac{3}{8}\sigma_{DE/DE} + \sigma_{LE/LE} + 2\sigma_{AmE/AmE} + \sigma_{DmE/DmE} + 4\sigma_{AE/AmE}$$

$$+ \sigma_{DE/DmE} + \sigma_{ee}$$

在估算遗传方差分量和遗传协方差分量的同时，还能采用LUP法或AUP法无偏预测亲本的加性×环境互作效应值和母体×环境互作效应值，以及组合的显性×环境互作效应值。根据基因型×环境互作效应的预测分析结果，可以评价亲本及其组合在不同环境中的育种价值和杂种优势的稳定性表现。

第八章 包括细胞质和母体效应的遗传模型分析

第一节 植物种子和动物幼畜的遗传模型

植物种子生长在母体植株上，其营养物质由母体植株所提供。种子的某些数量性状表现可能会同时受到种子的核基因和母体植株的核基因两套遗传体系的控制。动物的幼畜由其母畜抚育，因此幼畜的某些性状也可能会同时受到幼畜的核基因和母畜的核基因两套遗传体系的控制。另外细胞质基因也可能通过控制叶绿体(或线粒体)的合成而影响植株的光合作用或动植物的呼吸作用，从而间接控制种子或幼畜的性状表现。

朱军等(朱军, 1992; Zhu和Weir, 1994a)运用Cockerham(1980)的广义遗传模型建模原理，提出了具有生物学意义的、包括种子(或幼畜)核基因、母体核基因和细胞质基因三套遗传效应的遗传模型。根据所提出的植物种子和动物幼畜的遗传模型，可以用若干世代的平均数同时分析种子(或幼畜)、细胞质和母体三套遗传体系的基因效应。这个遗传模型可以把总遗传效应(G)进一步分解为直接加性效应(A)、直接显性效应(D)、细胞质效应(C)、母体加性效应(A_m)和母体显性效应(D_m)。因而总遗传方差(V_G)分解为直接遗传方差(V_{Go})、细胞质方差(V_C)和母体遗传方差(V_{Gm})，并可进一步分解为五项遗传方差分量，

$$\begin{aligned} V_G &= V_{Go} + V_C + V_{Gm} \\ &= V_A + V_D + V_C + V_{Am} + V_{Dm} \end{aligned}$$

如果以下假定成立，即 不存在父体效应(paternal effect)，细胞质基因只通过母体稳定遗传，不存在上位性效应，和不存在基因型与环境的互作效应，双子叶植物种子和动物幼畜的数量性状表现型值可用以下线性模型表示。一组纯系材料作遗传交配，第*i*个母本与第*j*个父

本的第 k 种交配类型在第 l 个区组内的世代平均数表现型值 (y_{ijkl}) 可以表示为

$$y_{ijkl} = \mu + G_{ijk} + B_l + e_{ijkl} \quad (8.1)$$

其中 μ 是群体平均数, 固定效应; B_l 是区组效应, 随机效应; e_{ijkl} 是剩余机误, 随机效应; 如果试验未设完全随机区组, 而只设了重复, 则在模型中应取消区组效应。 G_{ijk} 是世代平均数的遗传效应。 G_{ijk} 的遗传组成决定于不同亲本 (i, j) 及交配类型 ($k = 0, 1, 2, \dots, 6$) 的遗传效应由以下模型确定:

亲本 P_i 的遗传效应分量 ($k = 0$)

$$G_{ii0} = 2A_i + D_{ii} + C_i + 2Am_i + Dm_{ii}$$

杂种一代 F_{1ij} ($P_i \times P_j$) 的遗传效应分量 ($k = 1$)

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

杂种二代 F_{2ij} 的遗传效应分量 ($k = 2$)

$$G_{ij2} = A_i + A_j + 0.25D_{ii} + 0.25D_{jj} + 0.5D_{ij} + C_i + Am_i + Am_j + Dm_{ij}$$

回交一代 BC_i ($F_{1ij} \times P_i$) 的遗传效应分量 ($k = 3$)

$$G_{ij3} = 1.5A_i + 0.5A_j + 0.5D_{ii} + 0.5D_{ij} + C_i + Am_i + Am_j + Dm_{ij}$$

回交一代 BC_j ($F_{1ij} \times P_j$) 的遗传效应分量 ($k = 4$)

$$G_{ij4} = 0.5A_i + 1.5A_j + 0.5D_{jj} + 0.5D_{ij} + C_i + Am_i + Am_j + Dm_{ij}$$

反回交一代 RBC_i ($P_i \times F_{1ij}$) 的遗传效应分量 ($k = 5$)

$$G_{ij5} = 1.5A_i + 0.5A_j + 0.5D_{ii} + 0.5D_{ij} + C_i + 2Am_i + Dm_{ii}$$

反回交一代 RBC_j ($P_j \times F_{1ij}$) 的遗传效应分量 ($k = 6$)

$$G_{ij6} = 0.5A_i + 1.5A_j + 0.5D_{jj} + 0.5D_{ij} + C_j + 2Am_j + Dm_{jj}$$

种子(或幼畜)遗传模型中的遗传效应都是随机效应, 直接加性效应

A_i 或 $A_j \sim (0, \sigma_A^2)$; 直接显性效应 D_{ii} 、 D_{jj} 或 $D_{ij} \sim (0, \sigma_D^2)$; 细胞质效

应 $C_i \sim (0, \sigma_C^2)$; 母体加性效应 Am_i 或 $Am_j \sim (0, \sigma_{Am}^2)$; 母体显性效应 Dm_{ii} 、 Dm_{jj} 或 $Dm_{ij} \sim (0, \sigma_{Dm}^2)$ 。

以上各项遗传效应之间并不都是相互独立的。由于二倍体种子(或幼畜)的基因一半来自其母体, 应该存在直接加性效应与母体加性效应的遗传协方差 $Cov(A_i, Am_i) = \sigma_{A.Am}$, 以及直接显性效应与母体显性效应的遗传协方差 $Cov(D_{ij}, Dm_{ij}) = \sigma_{D.Dm}$ 。

虽然列出了7个遗传世代, 但在实际分析时采用不低于3个世代的双列杂交材料, 便可获得无偏的分析结果。对于自花授粉作物, F_1 植株上结的种子是 F_2 世代。因此还可以在遗传分析中包括亲本、杂种一代 F_1 和杂种二代 F_2 。对于异花授粉作物或动物, 则可以在遗传分析中包括亲本、杂种一代 F_1 和回交一代。

种子(或幼畜)数量性状表现型方差 (V_p) 应该包括

$$V_P = V_{Go} + V_C + V_{Gm} + 2C_{Go.Gm} + V_e$$

$$= (V_A + V_D) + V_C + (V_{Am} + V_{Dm}) + 2(C_{A.Am} + C_{D.Dm}) + V_e$$

其中直接遗传方差 (V_{Go}) 的分量有直接加性方差 (V_A) 和直接显性方差 (V_D), V_c 是细胞质方差, 母体遗传方差 (V_{Gm}) 的分量有母体加性方差 (V_{Am}) 和母体显性方差 (V_{Dm})。直接遗传效应与母体遗传效应的协方差由两项协方差分量组成(加性遗传协方差 $C_{A.Am}$ 和显性遗传协方差 $C_{D.Dm}$)。

不同遗传世代的表現型方差计算公式如下:

$$V_P(F_1) = (2\sigma_A^2 + \sigma_D^2) + \sigma_C^2 + (4\sigma_{Am}^2 + \sigma_{Dm}^2) + 2(2\sigma_{A.Am}) + \sigma_e^2$$

$$V_P(F_2) = (2\sigma_A^2 + \frac{3}{8}\sigma_D^2) + \sigma_C^2 + (2\sigma_{Am}^2 + \sigma_{Dm}^2) + 2(2\sigma_{A.Am} + \frac{1}{2}\sigma_{D.Dm}) + \sigma_e^2$$

$$V_P(BC_i) = V_P(BC_j)$$

$$= (2\frac{1}{2}\sigma_A^2 + \frac{1}{2}\sigma_D^2) + \sigma_C^2 + (2\sigma_{Am}^2 + \sigma_{Dm}^2) + 2(2\sigma_{A.Am} + \frac{1}{2}\sigma_{D.Dm}) + \sigma_e^2$$

对于成对性状, 表现型协方差 (C_p) 包括的协方差分量是

$$C_P = C_{Go} + C_C + C_{Gm} + 2C_{Go/Gm} + C_e$$

$$= (C_A + C_D) + C_C + (C_{Am} + C_{Dm}) + 2(C_{A/Am} + C_{D/Dm}) + C_e$$

其中直接遗传协方差(C_{G_0})的分量有直接加性协方差(C_A)和直接显性协方差(C_D)， C_C 是细胞质协方差，母体遗传协方差(C_M)的分量有母体加性协方差(C_{A_m})和母体显性协方差(C_{D_m})。一个性状的直接遗传效应与另一个性状的母体遗传效应的协方差由两项协方差分量组成(加性遗传协方差 C_{A/A_m} 和显性遗传协方差 C_{D/D_m})。

不同遗传世代的表现型协方差计算公式如下：

$$C_P(F_1) = (2\sigma_{A/A} + \sigma_{D/D}) + \sigma_{C/C} + (4\sigma_{A_m/A_m} + \sigma_{D_m/D_m}) + 2(2\sigma_{A/A_m}) + \sigma_{e/e}$$

$$C_P(F_2) = (2\sigma_{A/A} + \frac{3}{8}\sigma_{DD}) + \sigma_{C/C} + (2\sigma_{A_m/A_m} + \sigma_{D_m/D_m}) \\ + 2(2\sigma_{A/A_m} + \frac{1}{2}\sigma_{D/D_m}) + \sigma_{e/e}$$

$$C_P(BC_i) = C_P(BC_j)$$

$$= (2\frac{1}{2}\sigma_{A/A} + \frac{1}{2}\sigma_{DD}) + \sigma_{C/C} + (2\sigma_{A_m/A_m} + \sigma_{D_m/D_m}) \\ + 2(2\sigma_{A/A_m} + \frac{1}{2}\sigma_{D/D_m}) + \sigma_{e/e}$$

在分析种子(或幼畜)数量性状时要区分两种协方差。 $C_{A.A_m}$ 、 $C_{D.D_m}$ (或

$\sigma_{A.A_m}$ 、 $\sigma_{D.D_m}$)是控制一个性状的直接加性、显性效应与控制该性状的母体加性、显性效应的协方差， C_{A/A_m} 、 C_{D/D_m} (或 σ_{A/A_m} 、 σ_{D/D_m})是控制一个性状的直接加性、显性效应与控制另一个性状的母体加性、显性效应的协方差。

单子叶植物种子的主要组成部分是胚乳。胚乳数量性状同时受到三倍体胚乳、细胞质和二倍体母体植株的基因控制，其总遗传效应(G)也可以分解为胚乳直接遗传效应(G_0)、细胞质遗传效应(C)和母体植株遗传效应(G_m) (朱军, 1992; Zhu和Weir, 1994a, 1994b)。胚乳直接遗传效应(G_0)又可被进一步分解为直接加性(A)和直接显性(D)遗传分量。母体植株遗传效应(G_m)则可进一步被分解为母体加性(A_m)、母体显性(D_m)遗传分量。基于假定 不存在父体效应， 细胞质基因只通过母体稳定遗传， 不存在上位性效应， 不存在高价显性效应(三个等位基因之间的互作)

和 不存在基因型与环境的互作效应，胚乳数量性状表现型值可用以下线性模型表示。一组纯系材料作遗传交配，第 i 个母本与第 j 个父本的第 k 种交配类型在第 l 个区组内的世代平均数值 (y_{ijkl}) 可以表示为

$$y_{ijkl} = \mu + G_{ijk} + B_l + e_{ijkl} \quad (8.2)$$

其中 μ 是群体平均数，固定效应； B_l 是区组效应，随机效应； e_{ijkl} 是剩余机误，随机效应；如果试验未设完全随机区组，而只设了重复，则在模型中应取消区组效应。 G_{ijk} 是世代平均数的遗传效应。 G_{ijk} 的遗传组成决定于不同亲本 (i, j) 及交配类型 (k)，交配类型不能少于3种 ($k = 0, 1, 2, \dots, 6$)。

亲本 P_i 的遗传效应分量 ($k = 0$)

$$G_{ii0} = 3A_i + 3D_{ii} + C_i + 2Am_i + Dm_{ii}$$

杂种一代 F_{1ij} 的 ($P_i \times P_j$) 的遗传效应分量 ($k = 1$)

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

杂种二代 F_{2ij} 的遗传效应分量 ($k = 2$)

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

回交一代 $F_{1ij} \times P_i$ 的遗传效应分量 ($k = 3$)

$$G_{ij3} = 2A_i + A_j + 1.5D_{ii} + 0.5D_{jj} + D_{ij} + C_i + Am_i + Am_j + Dm_{ij}$$

回交一代 $F_{1i} \times P_j$ 的遗传效应分量 ($k = 4$)

$$G_{ij4} = A_i + 2A_j + 0.5D_{ii} + 1.5D_{jj} + D_{ij} + C_i + Am_i + Am_j + Dm_{ij}$$

反回交一代 $P_i \times F_{1ij}$ 的遗传效应分量 ($k = 5$)

$$G_{ij5} = 2.5A_i + 0.5A_j + 2D_{ii} + D_{ij} + C_i + 2Am_i + Dm_{ii}$$

反回交一代 $P_j \times F_{1ij}$ 的遗传效应分量 ($k = 6$)

$$G_{ij6} = 0.5A_i + 2.5A_j + 2D_{jj} + D_{ij} + C_j + 2Am_j + Dm_{jj}$$

如果杂交亲本是从一个推断群体中抽取的随机样本，以上各项遗传效应分量均为随机效应。遗传效应的定义与双子叶植物相同，直接加性效

应 A_i 或 $A_j \sim (0, \sigma_A^2)$; 直接显性效应 D_{ii} 、 D_{jj} 或 $D_{ij} \sim (0, \sigma_D^2)$; 细胞质效应 $C_i \sim (0, \sigma_C^2)$; 母体加性效应 Am_i 或 $Am_j \sim (0, \sigma_{Am}^2)$; 母体显性效应 Dm_{ii} 、 Dm_{jj} 或 $Dm_{ij} \sim (0, \sigma_{Dm}^2)$ 。同时还存在直接加性效应与母体加性效应的遗传协方差 $COV(A_i, Am_i) = \sigma_{A.A_m}$, 以及直接显性效应与母体显性效应的遗传协方差 $COV(D_{ij}, Dm_{ij}) = \sigma_{D.D_m}$ 。

并不需要在—个遗传分析中同时采用以上所列出的7个遗传世代。在实际分析时,分析亲本、杂种一代 F_1 和杂种二代 F_2 三个世代的双列杂交遗传资料,便可获得无偏的分析结果。

三倍体胚乳数量性状的表现型方差(V_P)包括

$$V_P = V_{Go} + V_C + V_{Gm} + 2C_{Go.Gm} + V_e \\ = (V_A + V_D) + V_C + (V_{Am} + V_{Dm}) + 2(C_{A.Am} + C_{D.Dm}) + V_e$$

其遗传效应的方差分量估算方法与二倍体种子(或幼畜)的略微不同。不同遗传世代的表现型方差计算公式如下:

$$V_P(F_1) = (5\sigma_A^2 + 5\sigma_D^2) + \sigma_C^2 + (4\sigma_{Am}^2 + \sigma_{Dm}^2) + 2(4\sigma_{A.Am} + \sigma_{D.Dm}) + \sigma_e^2 \\ V_P(F_2) = (4\frac{1}{2}\sigma_A^2 + 3\sigma_D^2) + \sigma_C^2 + (2\sigma_{Am}^2 + \sigma_{Dm}^2) + 2(3\sigma_{A.Am} + \sigma_{D.Dm}) + \sigma_e^2 \\ V_P(BC_i) = V_P(BC_j) \\ = (5\sigma_A^2 + 3\frac{1}{2}\sigma_D^2) + \sigma_C^2 + (2\sigma_{Am}^2 + \sigma_{Dm}^2) + 2(3\sigma_{A.Am} + \sigma_{D.Dm}) + \sigma_e^2$$

对于成对性状,表现型协方差(C_P)包括的协方差分量是

$$C_P = C_{Go} + C_C + C_{Gm} + 2C_{Go/Gm} + C_e \\ = (C_A + C_D) + C_C + (C_{Am} + C_{Dm}) + 2(C_{A/Am} + C_{D/Dm}) + C_e$$

不同遗传世代的表现型协方差计算公式如下:

$$C_P(F_1) = (5\sigma_{A/A} + 5\sigma_{D/D}) + \sigma_{C/C} + (4\sigma_{Am/Am} + \sigma_{Dm/Dm}) \\ + 2(4\sigma_{A/Am}) + \sigma_{e/e}$$

$$C_P(F_2) = (4\frac{1}{2}\sigma_{A/A} + 3\sigma_{DD}) + \sigma_{C/C} + (2\sigma_{Am/Am} + \sigma_{Dm/Dm}) \\ + 2(3\sigma_{A/Am} + \sigma_{D/Dm}) + \sigma_{e/e}$$

$$C_P(BC_i) = C_P(BC_j) \\ = (5\sigma_{A/A} + 3\frac{1}{2}\sigma_{DD}) + \sigma_{C/C} + (2\sigma_{Am/Am} + \sigma_{Dm/Dm}) \\ + 2(3\sigma_{A/Am} + \sigma_{D/Dm}) + \sigma_{e/e}$$

二倍体种子(或幼畜)和三倍体胚乳的遗传实验观察值均可以用以下混合线性模型表示，

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{U}_A \mathbf{e}_A + \mathbf{U}_D \mathbf{e}_D + \mathbf{U}_C \mathbf{e}_C + \mathbf{U}_{Am} \mathbf{e}_{Am} + \mathbf{U}_{Dm} \mathbf{e}_{Dm} + \mathbf{U}_B \mathbf{e}_B + \mathbf{e}_e \\ = \mathbf{1}\mu + \sum_u^7 \mathbf{U}_u \mathbf{e}_u$$

其中常数向量 $\mathbf{1}$ 是系数为常数1的向量， \mathbf{e}_u 是第 u 项随机效应向量，

$\mathbf{e}_u \sim (0, \sigma_u^2 \mathbf{I})$ ； \mathbf{U}_u 是随机效应向量 \mathbf{e}_u 的系数矩阵 ($u = 1, 2, \dots, 6$)， $\mathbf{U}_7 = \mathbf{I}$ 是单位矩阵。

表现型向量 \mathbf{y} 的方差-协方差矩阵为

$$\text{Var}(\mathbf{y}) = \sigma_A^2 \mathbf{V}_1 + \sigma_D^2 \mathbf{V}_2 + \sigma_C^2 \mathbf{V}_3 + \sigma_{Am}^2 \mathbf{V}_4 + \sigma_{Dm}^2 \mathbf{V}_5 + \sigma_B^2 \mathbf{V}_6 \\ + \sigma_{A.Am} \mathbf{V}_7 + \sigma_{D.Dm} \mathbf{V}_8 + \sigma_e^2 \mathbf{V}_9 \\ = \sum_{u=1}^9 \theta_u \mathbf{V}_u$$

其中 $\mathbf{V}_u = \mathbf{U}_u \mathbf{U}_u^T$ ($u = 1, 2, \dots, 6$)、 $\mathbf{V}_7 = (\mathbf{U}_1 \mathbf{U}_4^T + \mathbf{U}_4 \mathbf{U}_1^T)$ ，

$\mathbf{V}_8 = (\mathbf{U}_2 \mathbf{U}_5^T + \mathbf{U}_5 \mathbf{U}_2^T)$ 、 $\mathbf{V}_9 = \mathbf{I}$ ； $\theta_1 = \sigma_A^2$ 、 $\theta_2 = \sigma_D^2$ 、 $\theta_3 = \sigma_C^2$ 、

$\theta_4 = \sigma_{Am}^2$ 、 $\theta_5 = \sigma_{Dm}^2$ 、 $\theta_6 = \sigma_B^2$ 、 $\theta_7 = \sigma_{A.Am}$ 、 $\theta_8 = \sigma_{D.Dm}$ 、 $\theta_9 = \sigma_e^2$ 。

以上混合线性模型可用MINQUE(0/1)法无偏地估算各项方差和协方差分量，也可用LUP或AUP法无偏预测各项遗传效应。

如果控制种子数量性状的基因在不同环境条件下表现不一致，则需要在多种环境条件下分析性状的遗传规律，估算基因效应与环境效应的相互作用。某一环境下种子性状的表现型值主要由环境效应(E)、基因型效应(G)和基因型 × 环境互作效应(GE)所决定。基因型 × 环境互作效应(GE)也可作相应的分解，由种子直接遗传效应与环境效应的互作(GoE)、细胞质遗传效应与环境效应的互作(CE)和母体植株遗传效应与环境效应的互作(GmE)所组成。GoE可进一步被分解为直接加性 × 环境互作效应(AE)和直接显性 × 环境互作效应(DE)遗传分量。GmE则可被分解为母体加性 × 环境互作效应(AmE)、母体显性 × 环境互作效应(DmE)遗传分量。

如果假定不存在父体效应(paternal effect)，细胞质基因只通过母体稳定遗传，不存在上位性效应，种子数量性状表现型值可用以下线性模型表示。一组纯系材料作双列杂交遗传交配，第*i*个母本与第*j*个父本的第*k*种交配类型在环境*h*内第*l*个区组中的世代平均数值(y_{hijkl})可以表示为

$$y_{hijkl} = \mu + E_h + G_{ijk} + GE_{hijk} + B_{l(h)} + e_{hijkl} \quad (8.3)$$

其中 μ 是群体平均数，固定效应； E_h 是环境效应，固定效应； $B_{l(h)}$ 是环境内区组效应，随机效应； e_{hijkl} 是剩余机误，随机效应； G_{ijk} 是遗传效应， GE_{hijk} 是遗传效应与环境效应的互作。

种子营养物质，双子叶作物主要储存在二倍体的子叶中，而单子叶作物主要储存在三倍体的胚乳中。因此 G_{ijk} 和 GE_{hijk} 的遗传组成决定于作物类型、不同亲本(*i, j*)及交配类型(*k*)。对于双子叶作物种子数量性状，各遗传世代基因效应 G_{ijk} 的定义如(8.1)，而遗传与环境互作效应 GE_{hijk} 的分解如下式所示。

亲本 P_i 的遗传 × 环境互作效应分量 ($k = 0$)

$$GE_{hi0} = 2AE_{hi} + DE_{hi} + CE_{hi} + 2AmE_{hi} + DmE_{hi}$$

杂种一代 F_{1ij} ($P_i \times P_j$) 的遗传 \times 环境互作效应分量 ($k = 1$)

$$GE_{hij1} = AE_{hi} + AE_{hj} + DE_{hij} + CE_{hi} + 2AmE_{hi} + DmE_{hii}$$

杂种二代 F_{2ij} 的遗传 \times 环境互作效应分量 ($k = 2$)

$$GE_{hij2} = AE_{hi} + AE_{hj} + 0.25DE_{hii} + 0.25DE_{hjj} + 0.5DE_{hij} + CE_{hi} \\ + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

回交一代 $F_{1ij} \times P_i$ 的遗传 \times 环境互作效应分量 ($k = 3$)

$$GE_{hij3} = 1.5AE_{hi} + 0.5AE_{hj} + 0.5DE_{hii} + 0.5DE_{hij} + CE_{hi} \\ + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

回交一代 $F_{1ij} \times P_j$ 的遗传 \times 环境互作效应分量 ($k = 4$)

$$GE_{hij4} = 0.5AE_{hi} + 1.5AE_{hj} + 0.5DE_{hjj} + 0.5DE_{hij} + CE_{hi} \\ + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

反回交一代 $P_i \times F_{1ij}$ 的遗传 \times 环境互作效应分量 ($k = 5$)

$$GE_{hij5} = 1.5AE_{hi} + 0.5AE_{hj} + 0.5DE_{hii} + 0.5DE_{hij} + CE_{hi} \\ + 2AmE_{hi} + DmE_{hii}$$

反回交一代 $P_j \times F_{1ij}$ 的遗传 \times 环境互作效应分量 ($k = 6$)

$$GE_{hij6} = 0.5AE_{hi} + 1.5AE_{hj} + 0.5DE_{hjj} + 0.5DE_{hij} + CE_{hj} \\ + 2AmE_{hj} + DmE_{hij}$$

而对于单子叶作物胚乳数量性状，各遗传世代基因效应 $G_{i,j,k}$ 的定义如(8.2)，遗传与环境互作效应 $GE_{hij,k}$ 的分解如下式所示。

亲本 P_i 的遗传 \times 环境互作效应分量 ($k = 0$)

$$GE_{hij0} = 3AE_{hi} + 3DE_{hii} + CE_{hi} + 2AmE_{hi} + DmE_{hii}$$

杂种一代 F_{1ij} ($P_i \times P_j$) 的遗传 \times 环境互作效应分量 ($k = 1$)

$$GE_{hij1} = 2AE_{hi} + AE_{hj} + DE_{hii} + 2DE_{hij} + CE_{hi} + 2AmE_{hi} \\ + DmE_{hii}$$

杂种二代 F_{2ij} 的遗传 \times 环境互作效应分量 ($k = 2$)

$$GE_{hij2} = 1.5AE_{hi} + 1.5AE_{hj} + DE_{hii} + DE_{hij} + DE_{hij} + CE_{hi} \\ + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

回交一代 $F_{1j} \times P_i$ 的遗传 \times 环境互作效应分量 ($k = 3$)

$$GE_{hij3} = 2AE_{hi} + AE_{hj} + 1.5DE_{hii} + 0.5DE_{hij} + DE_{hij} + CE_{hi} \\ + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

回交一代 $F_{1j} \times P_j$ 的遗传 \times 环境互作效应分量 ($k = 4$)

$$GE_{hij4} = AE_{hi} + 2AE_{hj} + 0.5DE_{hii} + 1.5DE_{hij} + DE_{hij} + CE_{hi} \\ + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

反回交一代 $P_i \times F_{1j}$ 的遗传 \times 环境互作效应分量 ($k = 5$)

$$GE_{hij5} = 2.5AE_{hi} + 0.5AE_{hj} + 2DE_{hii} + DE_{hij} + CE_{hi} \\ + 2AmE_{hi} + DmE_{hii}$$

反回交一代 $P_j \times F_{1j}$ 的遗传 \times 环境互作效应分量 ($k = 6$)

$$GE_{hij6} = 0.5AE_{hi} + 2.5AE_{hj} + 2DE_{hij} + DE_{hij} + CE_{hj} \\ + 2AmE_{hj} + DmE_{hij}$$

虽然列出了7个世代($k = 0, 1, 2, \dots, 6$)平均数的遗传效应分量的公式,在实际分析作物种子数量性状时,并不需要包括全套遗传材料。最简便的遗传交配设计是采用一组亲本、正反交 F_1 及其自交后代 F_2 。采用这一交配设计只需在田间种植亲本和 F_1 植株,由其自交产生亲本种子和 F_2 种子,亲本间杂交可获得 F_1 种子。

包括环境互作效应的二倍体种子(或幼畜)和三倍体胚乳的遗传资料,可以用以下混合线性模型表示,

$$\mathbf{y} = \mathbf{Xb} + \mathbf{U}_A \mathbf{e}_A + \mathbf{U}_D \mathbf{e}_D + \mathbf{U}_C \mathbf{e}_C + \mathbf{U}_{Am} \mathbf{e}_{Am} + \mathbf{U}_{Dm} \mathbf{e}_{Dm} \\ + \mathbf{U}_{AE} \mathbf{e}_{AE} + \mathbf{U}_{DE} \mathbf{e}_{DE} + \mathbf{U}_{CE} \mathbf{e}_{CE} + \mathbf{U}_{AmE} \mathbf{e}_{AmE} + \mathbf{U}_{DmE} \mathbf{e}_{DmE} + \mathbf{U}_B \mathbf{e}_B + \mathbf{e}_e \\ = \mathbf{Xb} + \sum_u^{12} \mathbf{U}_u \mathbf{e}_u$$

其中 \mathbf{X} 是固定效应向量 \mathbf{b} 的系数矩阵, \mathbf{e}_u 是第 u 项随机效应向量, $\mathbf{e}_u \sim (0, \sigma_u^2 \mathbf{I})$; \mathbf{U}_u 是随机效应向量 \mathbf{e}_u 的系数矩阵 ($u = 1, 2, \dots, 11$), $\mathbf{U}_{12} = \mathbf{I}$ 是单位矩阵。

表现型向量 \mathbf{y} 的方差-协方差矩阵为

$$\begin{aligned} \text{Var}(\mathbf{y}) &= \sigma_A^2 \mathbf{V}_1 + \sigma_D^2 \mathbf{V}_2 + \sigma_C^2 \mathbf{V}_3 + \sigma_{Am}^2 \mathbf{V}_4 + \sigma_{Dm}^2 \mathbf{V}_5 \\ &\quad + \sigma_{AE}^2 \mathbf{V}_6 + \sigma_{DE}^2 \mathbf{V}_7 + \sigma_{CE}^2 \mathbf{V}_8 + \sigma_{AmE}^2 \mathbf{V}_9 + \sigma_{DmE}^2 \mathbf{V}_{10} + \sigma_B^2 \mathbf{V}_{11} \\ &\quad + \sigma_{A.Am} \mathbf{V}_{12} + \sigma_{D.Dm} \mathbf{V}_{13} + \sigma_{AE.AmE} \mathbf{V}_{14} + \sigma_{DE.DmE} \mathbf{V}_{15} + \sigma_e^2 \mathbf{V}_{16} \\ &= \sum_{u=1}^{16} \theta_u \mathbf{V}_u \end{aligned}$$

其中 $\mathbf{V}_u = \mathbf{U}_u \mathbf{U}_u^T$ ($u = 1, 2, \dots, 11$)、 $\mathbf{V}_{12} = (\mathbf{U}_1 \mathbf{U}_4^T + \mathbf{U}_4 \mathbf{U}_1^T)$ 、 $\mathbf{V}_{13} = (\mathbf{U}_2 \mathbf{U}_5^T + \mathbf{U}_5 \mathbf{U}_2^T)$ 、 $\mathbf{V}_{14} = (\mathbf{U}_6 \mathbf{U}_9^T + \mathbf{U}_9 \mathbf{U}_6^T)$ 、 $\mathbf{V}_{15} = (\mathbf{U}_7 \mathbf{U}_{10}^T + \mathbf{U}_{10} \mathbf{U}_7^T)$ 、 $\mathbf{V}_{16} = \mathbf{I}$; $\theta_1 = \sigma_A^2$ 、 $\theta_2 = \sigma_D^2$ 、 $\theta_3 = \sigma_C^2$ 、 $\theta_4 = \sigma_{Am}^2$ 、 $\theta_5 = \sigma_{Dm}^2$ 、 $\theta_6 = \sigma_{AE}^2$ 、 $\theta_7 = \sigma_{DE}^2$ 、 $\theta_8 = \sigma_{CE}^2$ 、 $\theta_9 = \sigma_{AmE}^2$ 、 $\theta_{10} = \sigma_{DmE}^2$ 、 $\theta_{11} = \sigma_B^2$ 、 $\theta_{12} = \sigma_{A.Am}$ 、 $\theta_{13} = \sigma_{D.Dm}$ 、 $\theta_{14} = \sigma_{AE.AmE}$ 、 $\theta_{15} = \sigma_{DE.DmE}$ 、 $\theta_{16} = \sigma_e^2$ 。

可用MINQUE(0/1)法无偏地估算以上混合线性模型的各项方差和协方差分量,也可用LUP或AUP法无偏预测各项遗传效应。

现以陆地棉5个亲本及其 F_1 、 F_2 棉籽营养品质性状的三年实验资料分析,作为演示作物种子遗传模型的分析实例。遗传试验于1990~1993年在浙江农业大学实验场进行,选用5个陆地棉亲本按 4×5 杂交方式配制 F_1 。五个陆地棉品种分别为: :中棉所7号; :HG-H-12,高棉酚品种; :PD0111; :PD0458; :GL-5,无棉酚品种。于1991、1992和1993分别种植亲本和 F_1 ,随机区组设计,二次重复,小区面积为 $1.33 \times 2.5\text{m}^2$ 。亲本和 F_1 分别自交产生亲本和 F_2 种子,亲本间杂交获得 F_1 当代

种子。测定了棉籽仁油分含量、油指(百粒棉籽仁油分重量, g)、蛋白质含量、蛋白质指数(百粒棉籽仁蛋白质重量, g)、赖氨酸含量和赖氨酸指数(百粒棉籽仁赖氨酸重量, g)。

采用包括基因型与环境互作效应的作物二倍体种子遗传模型及分析方法, 分析多环境下棉籽仁营养品质的遗传特性。运用MINQUE(0/1)法无偏地估算各性状的各项遗传效应方差分量及基因效应与环境互作方差分量; 采用调整无偏预测法(AUP法)预测各项遗传效应。采用Jackknife重复抽样技术计算各估算值和预测值的标准误, 并用t测验对参数进行显著性检验。

表8.1. 亲本及F₁、F₂棉籽仁营养品质性状的平均表型值(1991-1993)

性状	亲本			F ₁			F ₂		
	1991	1992	1993	1991	1992	1993	1991	1992	1993
油分含量	34.93	34.60	35.29	35.95	35.47	36.32	36.06	36.35	36.53
油指	2.65	2.83	2.24	2.63	2.79	2.28	2.55	3.31	2.42
蛋白质含量	40.74	40.71	39.59	39.29	39.34	38.88	39.06	38.26	38.49
蛋白质指数	3.13	3.37	2.55	2.86	3.11	2.45	2.76	3.48	2.57
赖氨酸含量	0.96	0.98	1.04	0.98	0.97	1.01	0.93	0.95	0.98
赖氨酸指数	7.22	8.04	6.64	6.84	7.61	6.34	6.67	8.17	6.13

5个亲本及其F₁、F₂棉籽仁营养品质性状在3年中的平均表现型值列于表8.1。由表8.1可知, 各年的F₁棉籽仁平均油分含量均高于各自亲本, 表现出一定的正向直接杂种优势。F₂平均油分含量也均高于F₁, 这表明提供F₂种子营养物质的F₁母体植株存在一定的正向母体杂种优势。F₁蛋白质平均含量在各年中表现与油分含量相反, 均低于亲本, 蛋白质含量具有负向直接杂种优势。F₂平均蛋白质含量也低于F₁, 这说明蛋白质含量存在负向母体杂种优势。赖氨酸含量存在一定的负向母体杂种优势。其它性状在杂种优势上表现不明显。综合分析亲本、F₁和F₂种子3年的资料, 各性状的表现年份间有一定波动, 可能存在基因型与环境的互

作。

表8.2. 陆地棉种子仁营养品质性状的遗传方差和协方差分量估值

参数	油分含量	油指	蛋白质含量	蛋白质指数	赖氨酸含量	赖氨酸指数
V_A	0.000	0.060**	1.678**	0.089**	0.204**	0.300**
V_D	0.363**	0.000	0.541**	0.000	0.000	0.000
V_C	0.000	0.000	1.139**	0.000	0.527**	0.590**
V_{Am}	1.829**	0.079**	0.000	0.121**	0.000	0.651**
V_{Dm}	1.117**	0.000	1.269**	0.000	0.133**	0.000
V_{AE}	2.483**	0.000	1.747**	0.054**	0.319**	0.000
V_{DE}	0.000	0.023**	0.252**	0.020**	0.125**	0.000
V_{CE}	1.221**	0.072**	1.766**	0.118**	0.915**	1.212**
V_{AmE}	1.835**	0.000	0.814**	0.000	0.000	0.000
V_{DmE}	0.000	0.047**	0.000	0.039**	0.000	0.000
$C_{A. Am}$	0.000	0.014	0.000	0.059**	0.000	0.139
$C_{D. Dm}$	0.147	0.000	0.101	0.000	0.000	0.000
$C_{AE. AmE}$	-0.414	0.000	-1.171	0.000	0.000	0.000
$C_{DE. DmE}$	0.000	-0.012	0.000	-0.039	0.000	0.000
V_e	1.859**	0.062**	0.199**	0.057**	0.052**	0.425**
V_p	10.172**	0.347**	7.263**	0.539**	2.275**	3.457**

*、** 分别表示达到0.05和0.01的显著水平。

棉籽仁各性状的基因主效应方差分量、基因效应与环境互作方差分量及协方差分量估算值列于表8.2。由表8.2可知，各性状的直接遗传方差分量(V_A+V_D)和母体遗传方差分量($V_{Am}+V_{Dm}$)均达到极显著水平，说明种子直接基因效应和母体基因效应同时控制棉籽仁营养品质，其中油分含量、油指、蛋白质指数和赖氨酸指数的母体方差分量大于直接方差分量，表明这些性状主要受制于母体效应。蛋白质含量的遗传以直接效应为主，受细胞质效应的影响也较大，细胞质方差占基因主效应总方差的32.65%。赖氨酸含量则以细胞质效应为主。在基因主效应中，各性状均

以加性效应为主。油分含量和蛋白质含量的直接显性方差和母体显性方差均达到显著水平，表明种子杂种优势和母体杂种优势会同时存在。以上是不受环境影响的基因主效应。

本资料分析结果表明，各性状还不同程度地存在基因型效应与环境互作的作用。在各品质性状中，基因效应与环境互作方差对总方差的贡献率为：蛋白质含量63.0%，赖氨酸含量59.7%、油分含量54.5%，最小是蛋白质指数、油指和赖氨酸指数，分别为40.9%、42.9%和34.7%。其中直接加性×环境互作效应和母体加性×环境互作效应对油分含量影响较为明显，互作方差也较大。在互作方差中，以细胞质互作效应对赖氨酸含量、油指和蛋白质指数影响最为明显，细胞质×环境互作方差分别占总互作变异的50.7%、38.6%和51.1%左右，油分含量存在一定的细胞质互作效应。同时这也说明尚不能排除细胞质效应对油分含量、油指和蛋白质指数的影响。直接显性×环境互作效应和母体显性×环境互作效应对油指和蛋白质指数影响均达到极显著水平。由此表明，油分含量和蛋白质指数还可能存在着种子直接和母体互作杂种优势。蛋白质指数的显著加性协方差 ($\hat{C}_{A.Am} = 0.059^{**}$)，说明蛋白质指数的种子直接加性效应与母体加性效应有着较强的遗传正相关，这对蛋白质指数的育种选择是有利的。另外，各性状的机误差分量均达到极显著水平，环境误差、抽样误差或测定误差对各性状也有一定的影响。

表8.3只列出了5个亲本表现显著的营养品质性状的直接加性效应 (A_i)、细胞质效应 (C_i)和母体加性效应 (Am_i)的预测值。各亲本互作效应的表现将在文中讨论。

表 8.3. 五个亲本的棉籽仁营养品质性状的遗传效应预测值

性状	参数	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
油分含量						
	母体加性 A_i	0.29	-0.02	-0.13	0.92	-1.06 ⁺
油指						
	直接加性 A_i	-0.24 ⁺	0.06	0.02	-0.06	0.22 ⁺
	母体加性 Am_i	-0.05	-0.03	-0.13 ⁺	0.11	0.10
蛋白质含量						
	直接加性 A_i	-0.84 ⁺	0.04	-0.29	0.15	0.94 ⁺
	细胞质 C_i	-0.73	0.87 ⁺	-1.05	-0.38	1.29
蛋白质指数						
	直接加性 A_i	-0.26 ⁺	0.03	-0.03	0.02	0.25 ^{**}
	母体加性 Am_i	-0.05	-0.05	-0.12 ⁺	0.01	0.22 ⁺
赖氨酸含量						
	直接加性 A_i	0.03	0.02 ⁺	0.02	-0.06 ⁺	-0.01
	细胞质 C_i	0.09 ⁺	-0.08 ^{**}	-0.06	0.01	0.03 ^{**}
赖氨酸指数						
	直接加性 A_i	-0.29	0.14	0.04	-0.27	0.38 ⁺
	细胞质 C_i	0.98	-1.17	-0.32 ⁺	-0.23	0.74
	母体加性 Am_i	-0.28	-0.05	-0.27	0.13	0.47 ⁺

+、*、**分别表示达到0.10、0.05和0.01的显著水平。

由表8.3可知，5个亲本中，P1(中棉所7号)和P4(PD0458)油分含量具有较大的母体加性效应值，P5(GL-5)有最小的母体加性效应值。P1的

赖氨酸含量也具有较大的直接加性效应值和细胞质效应值。各亲本的直接加性互作效应值、细胞质互作效应值和母体互作加性效应值在各年份表现都不一致，可能与三年不同的气候环境有关。P1和P5的互作效应值在1991年均均有较好的表现。如以P1作杂交亲本，尤其作母本，有利于提高后代棉籽仁的油分含量和赖氨酸含量。利用P1作亲本，在环境适宜或特定年份下(如1991年)，改良棉籽仁的油分含量和赖氨酸含量效果更明显。虽然P5在某一环境下油分含量会有所提高，但不宜作高油分含量育种材料。各亲本的直接加性互作效应与母体加性互作效应在多数年份中表现方式相反。直接加性互作效应与母体加性互作效应可能存在负相关，这为棉籽仁油分含量的改良工作增加了难度。P5蛋白质含量和赖氨酸指数的直接加性效应和细胞质效应值均较大，且在3个互作效应中，显著互作效应预测值均表现为正值。这说明在大多数环境下，它们能有较好的表现。其它亲本的互作效应值年份间表现不稳定。这表明在多种环境下利用P5都能选配到高蛋白含量组合。五个亲本中，P5的油指和蛋白质指数和赖氨酸指数具有最大的直接加性效应和母体加性效应值。而细胞质互作效应在各年份作用方向不一致。可见，虽然P5是改良油指和蛋白质指数较好的亲本材料，但在某些年份如1992年，棉籽仁的油指和蛋白质指数仍会表现一般。其余亲本不是改良油指和蛋白质指数的理想亲本。此外5个亲本各性状受环境影响较大，均存在显著的基因型与环境互作效应值。这与前面遗传方差分量分析的结果是相吻合的。

可以采用统计量 $\hat{\Delta} = -\sum_{i=1}^n \hat{D}_{ii} / \sqrt{n\hat{\sigma}_D^2}$ 确定直接显性效应作用方向、判断种子直接杂种优势的大小(朱军等, 1993)。同样， $\hat{\Delta}_M =$

$-\sum_{i=1}^n \hat{D}m_{ii} / \sqrt{n\hat{\sigma}_{Dm}^2}$ 可确定母体显性效应作用方向和母体杂种优势的存在。

$\hat{\Delta}_{OE} = -\sum_{i=1}^n D\hat{E}_{ii} / \sqrt{n\hat{\sigma}_{DE}^2}$ 确定种子直接显性与环境互作效应作用方向和种

子显性直接与环境互作杂种优势的存在。 $\hat{\Delta}_{ME} = -\sum_{i=1}^n \hat{DmE}_{ii} / \sqrt{n\hat{\sigma}_{DmE}^2}$ 确

定母体显性与环境互作效应作用方向和母体显性与环境互作杂种优势的存在。本文分析结果发现，营养品质性状存在种子直接杂种优势和母体杂种优势及其与环境互作杂种优势。油分含量的 $\hat{\Delta} = 1.275^+$ 和 $\hat{\Delta}_M = 2.218^{**}$ ，蛋白质含量的 $\hat{\Delta} = -2.043^{**}$ 和 $\hat{\Delta}_M = -2.418^{**}$ ，这表明这两个性状均存在种子直接杂种优势和母体杂种优势，其油分含量杂种优势是正向的，蛋白质含量杂种优势是负向的。赖氨酸含量存在显著的母体负杂种优势，其 $\hat{\Delta}_M = -2.739^{**}$ 。显然同时利用油分含量和蛋白质含量的杂种优势有相当大的困难。这与前面平均表现型值分析结果相符。油脂和蛋白质含量均存在直接与环境互作杂种优势，其分别有 $\hat{\Delta}_{OE} = 2.246^+$ 、 $\hat{\Delta}_{OE} = -1.710^{**}$ 。蛋白质指数存在显著的种子直接杂种优势 ($\hat{\Delta} = -3.024^+$) 和母体互作杂种优势 ($\hat{\Delta}_{ME} = -4.053^{**}$)。以上表明，在一定环境下，选配油脂和蛋白质含量或蛋白质指数杂交组合表现出较大的杂种优势。

在种子品质育种过程中如果能通过选择物理性状间接改良棉籽营养品质性状，则可减少品质测试工作、提高育种选择效率。因此分析两者的遗传相关和各项遗传相关在实际工作中有重要的意义。表8.4列出了籽指(百粒棉籽重, g)、容重、仁/壳(仁重/壳重的比值)和仁指(百粒棉籽仁重, g)分别与含油率、油脂、蛋白质含量、蛋白质指数的各项相关系数。

表 8.4. 棉籽物理性状与营养品质性状间的遗传相关系数估计值

性状	参数	含油率	油指	蛋白质含量	蛋白质指数
籽指	Γ_P	-0.11	0.33**	0.22**	0.39**
	Γ_G	-0.14 ⁺	0.24**	0.23**	0.32**
	Γ_A	0.00	0.17 ⁺	0.18 ⁺	0.14 ⁺
	Γ_D	0.00	0.00	0.00	0.00
	Γ_C	0.00	0.00	0.00	0.00
	Γ_{Am}	-0.16 ⁺	-0.07	0.00	-0.00
	Γ_{Dm}	0.00	0.00	0.00	0.00
	Γ_{AE}	0.00	0.00	0.00	0.00
	Γ_{DE}	0.00	0.23**	-0.79**	0.12 ⁺
	Γ_{CE}	-0.06	0.51**	0.21**	0.51**
	Γ_{AmE}	0.00	0.00	0.00	0.00
	Γ_{DmE}	0.00	-0.14 [*]	0.00	-0.20**
	Γ_e	0.08	0.86**	0.01	0.91**
容重	Γ_P	-0.04	0.16 [*]	0.15**	0.20**
	Γ_G	-0.11 ⁺	0.17**	0.17**	0.23**
	Γ_A	0.00	0.47**	0.92**	0.67**
	Γ_D	0.00	0.00	0.00	0.00
	Γ_C	0.00	0.00	0.00	0.00
	Γ_{Am}	0.26**	-0.57**	0.00	-0.69**
	Γ_{Dm}	0.00	0.00	0.00	0.00
	Γ_{AE}	-0.56**	0.00	-0.12 ⁺	-0.09
	Γ_{DE}	0.00	0.30**	0.04	0.30**
	Γ_{CE}	0.00	0.00	0.00	0.00
	Γ_{AmE}	-0.54**	0.00	0.24**	0.00
	Γ_{DmE}	0.00	0.00	0.00	0.00
	Γ_e	0.25**	0.13 ⁺	0.12 ⁺	0.05

+、*、**分别表示达到0.10、0.05和0.01的显著水平。

表8.4 (续). 棉籽物理性状与营养品质性状间的遗传相关系数估计值.

性状	参数	含油率	油指	蛋白质含量	蛋白质指数
仁/壳	r_P	0.07	0.30**	0.12*	0.26**
	r_G	0.09	0.26**	0.14*	0.22**
	r_A	0.00	0.29**	0.40**	0.47**
	r_D	-0.22**	0.00	0.23**	0.00
	r_C	0.00	0.00	0.00	0.00
	r_{Am}	0.00	0.00	0.00	0.00
	r_{Dm}	0.49**	0.00	-0.33**	0.00
	r_{AE}	-0.04	0.00	0.75**	-0.59**
	r_{DE}	0.00	0.00	0.00	0.00
	r_{CE}	-0.45**	0.42**	0.54**	0.60**
	r_{AmE}	0.00	0.00	0.00	0.00
	r_{DmE}	0.00	0.12	0.00	-0.02
	r_e	-0.04	0.50**	-0.13*	0.58**
仁指	r_P	-0.08	0.37**	0.19*	0.40**
	r_G	-0.10	0.27**	0.20**	0.32**
	r_A	0.00	0.34**	0.33**	0.35**
	r_D	0.00	0.00	0.00	0.00
	r_C	0.00	0.00	0.00	0.00
	r_{Am}	-0.10	-0.04	0.00	0.01
	r_{Dm}	0.00	0.00	0.00	0.00
	r_{AE}	0.00	0.00	0.00	0.00
	r_{DE}	0.00	0.49**	-0.57**	0.38**
	r_{CE}	-0.21**	0.42**	0.28**	0.49**
	r_{AmE}	0.00	0.00	0.00	0.00
	r_{DmE}	0.00	-0.06	0.00	-0.12*
	r_e	0.04	0.91**	0.08	0.99**

+, *, ** 分别表示达到0.10、0.05和0.01的显著水平。

表 8.5. 棉籽营养品质性状与植株产量性状间的遗传协方差估值

品质性状	协方差分量	产 量 性 状			
		单株产量	单株铃数	单铃重	衣分
油分指数	C _{A/Am}	0.014	0.002	-0.001	-0.006
	C _{D/Dm}	-0.018	-0.003	0.001	0.008
	C _{C/C}	-0.857	-0.615 [*]	0.211 ^{**}	0.227 ⁺
	C _{Am/Am}	-0.581	-0.448 [*]	0.154 ^{**}	0.149
	C _{Dm/Dm}	-0.379	-0.271 [*]	0.093 ^{**}	0.101 ⁺
	C _{AE/AmE}	-0.015	0.020	-0.008	0.014
	C _{DE/DmE}	0.006	-0.048	0.018	-0.017
	C _{CE/CE}	-0.936	-0.807 ⁺	0.280 ^{**}	0.219
	C _{AmE/AmE}	-0.568	-0.394 [*]	0.134 ^{**}	0.154 ⁺
	C _{DmE/DmE}	-0.328	-0.295	0.103 ⁺	0.074
	C _{e/e}	0.044	0.014	0.001	0.008
蛋白质指数	C _{A/Am}	0.013	0.003	-0.000	-0.004
	C _{D/Dm}	-0.018	-0.003	0.000	0.005
	C _{C/C}	-0.703	-0.522 ⁺	0.195 ^{**}	0.240 [*]
	C _{Am/Am}	-0.480	-0.381	0.144 ^{**}	0.164 [*]
	C _{Dm/Dm}	-0.314	-0.231 ⁺	0.086 ^{**}	0.106 [*]
	C _{AE/AmE}	-0.021	0.014	-0.008	0.004
	C _{DE/DmE}	0.023	-0.034	0.019	-0.001
	C _{CE/CE}	-0.771	-0.686	0.264 ^{**}	0.265
	C _{AmE/AmE}	-0.472	-0.336 ⁺	0.124 ^{**}	0.158 ^{**}
	C _{DmE/DmE}	-0.263	-0.248	0.098 [*]	0.094
	C _{e/e}	-0.045	-0.015	-0.003	-0.008

+、*、** 分别表示达到0.10、0.05和0.01的显著水平。

采用朱军(1993)提出的统计分析方法可以分析多环境下种子品质性状与植株产量性状间的遗传相关性。由于控制种子性状的遗传效应(G)由种子的直接加性效应(A)、直接显性效应(D)、细胞质效应(C)、母体植株加性效应(Am)和母体植株显性效应(Dm)等各项遗传分量组成,而控制植株的农艺性状的遗传效应(G)则由细胞质效应(C)、植株加性效应(Am)和植株显性

效应(Dm)等分量组成。因此存在以下各项遗传协方差分量：种子直接加性效应与母体加性效应间的加性协方差 $C_{A/Am}$ 、种子直接显性效应与母体显性效应间的显性协方差 $C_{D/Dm}$ 、种子细胞质效应与母体细胞质效应间的细胞质协方差 $C_{C/C}$ 、种子的母体加性效应与农艺性状的加性效应间的母体加性协方差 $C_{Am/Am}$ 、种子的母体显性效应与农艺性状的显性效应间的母体显性协方差 $C_{Dm/Dm}$ 。当存在基因型×环境互作效应(GE)时，所有这些遗传效应与环境效应都存在相应的互作分量：种子直接加性互作效应(AE)与母体加性互作效应(AmE)间的加性互作协方差 $C_{AE/AmE}$ 、种子直接显性互作效应(DE)与母体显性互作效应(DmE)间的显性互作协方差 $C_{DE/DmE}$ 、种子细胞质互作效应(CE)与母体细胞质互作效应(CE)间的细胞质互作协方差 $C_{CE/CE}$ 、种子的母体加性互作效(AmE)应与农艺性状的加性互作效应间(AmE)的母体加性互作协方差 $C_{AmE/AmE}$ 、种子的母体显性互作效应(DmE)与农艺性状的显性互作效应(DmE)间的母体显性互作协方差 $C_{DmE/DmE}$ 。另外，还存在随机误差之间的协方差 $C_{e/e}$ 。

棉籽油分指数、蛋白质指数与棉花产量构成因素的遗传协方差分量表列于8.5。油分指数和蛋白质指数与单株产量不存在显著的各项遗传相关，它们的各项遗传协方差均未达到显著水平。由此表明，对油分指数和蛋白质指数与单株产量可分别进行改良。

油分指数与单株铃数存在较大的遗传相关性，其细胞质协方差、母体加性协方差、母体显性协方差以及细胞质互作协方差、母体加性互作协方差均为显著的负值。这说明，控制单株铃数的细胞质效应与影响油分指数的细胞质效应存在负向的遗传关系，导致单株铃数增加的植株加性效应、显性效应却可使棉籽油分指数降低，但在某些环境下表现可能有所差异。蛋白质指数与单株铃数间的主要遗传相关分量与油分指数的表现相似。由此表明，在杂种早代选择单株铃数较少的植株，有可能获得棉籽油分指数和蛋白质指数较高的后代。

油分指数和蛋白质指数与单铃重存在显著的细胞质相关、母体加性相关、母体显性相关以及细胞质互作相关、母体加性互作相关。它们间均具有较强的细胞质和细胞质互作遗传相关。其中油分指数、蛋白质指数与单铃重的母体显性互作协方差也达到显著水平。这表明加强早代大铃选育，不难达到对单铃重和油分指数和蛋白质指数的同步改良，改良效果可能会因环境变化而有所不同。此外，铃重的强优势组合往往具有高油分和蛋白质指数，在某些环境中，组合优势表现尤为突出。

衣分与油分指数的母体加性互作协方差达到显著水平。显然，可在特定环境下通过早代间接选择衣分来改良油分指数，在某些环境下表现为衣分高的材料，其棉籽油分指数也高。衣分与油分指数的母体显性协方差也达到显著水平。由于衣分主要以加性遗传为主，存在微弱的杂种优势，因而利用衣分杂种优势间接改善油分指数的实际意义不大。蛋白质指数与衣分遗传相关性表现与油分指数的表现相似，但还存在较大的细胞质协方差和母体加性协方差，选择高衣分材料可望同时改良蛋白质指数。

第二节 二倍体胚和三倍体胚乳的种子遗传模型

以上所介绍的二倍体种子遗传模型包括了胚及子叶、细胞质、母体植株三套遗传体系的基因效应，没有包括胚乳的遗传效应；而胚乳遗传模型只包括了胚乳、细胞质、母体植株三套遗传体系的基因效应，没有包括种胚的遗传效应。种子的胚和胚乳是单子叶植物种子的主要组成部分。因此，有必要发展一个新的种子遗传模型，它包括种子的胚、胚乳、细胞质和母体植株四套遗传体系的基因效应。这个遗传模型可以把总遗传效应(G)进一步分解为胚直接加性效应(A_o)和胚直接显性效应(D_o)、胚乳直接加性效应(A_e)和胚乳直接显性效应(D_e)、细胞质效应(C)、母体加性效应(A_m)和母体显性效应(D_m)。因而总遗传方差(V_G)可分解为胚直接遗传方差(V_{G_o})、胚乳直接遗传方差(V_{G_e})、细胞质方差(V_C)和母体遗传方差(V_{G_m})，并可进一步分解为七项遗传方差分量，

$$V_G = V_{Go} + V_{Ge} + V_C + V_{Gm}$$

$$= V_{Ao} + V_{Do} + V_{Ae} + V_{De} + V_C + V_{Am} + V_{Dm}$$

如果以下遗传假设成立：不存在父体效应，细胞质基因只通过母体稳定遗传，不存在上位性效应，不存在高价显性效应(三个等位基因之间的互作)、不存在基因型与环境的互作效应，单子叶植物种子数量性状表现型值可用以下线性模型表示。一组纯系材料作遗传交配，第*i*个母本与第*j*个父本的第*k*种交配类型在第*l*个区组内的世代平均数表现型值(y_{ijkl})可以表示为

$$y_{ijkl} = \mu + G_{ijk} + B_l + e_{ijkl} \quad (8.4)$$

其中 μ 是群体平均数，固定效应； B_l 是区组效应，随机效应； e_{ijkl} 是剩余机误，随机效应；如果试验未设完全随机区组，而只设了重复，则在模型中应取消区组效应。 G_{ijk} 是世代平均数的遗传效应。 G_{ijk} 的遗传组成决定于不同亲本(*i, j*)及交配类型(*k*)。几种常用的交配类型($k = 0, 1, 2, \dots, 6$)的遗传效应由以下模型确定：

亲本 P_i 遗传效应分量 ($k = 0$)

$$G_{ii0} = 2A_{oi} + D_{oii} + 3A_{ei} + 3D_{eii} + C_i + 2A_{mi} + D_{mii}$$

杂种一代 F_{1ij} ($P_i \times P_j$) 遗传效应分量 ($k = 1$)

$$G_{ij1} = A_{oi} + A_{oj} + D_{oij} + 2A_{ei} + A_{ej} + D_{eii} + 2D_{eij} + C_i + 2A_{mi} + D_{mii}$$

杂种二代 F_{2ij} 遗传效应分量 ($k = 2$)

$$G_{ij2} = A_{oi} + A_{oj} + 0.25D_{oii} + 0.25D_{ojj} + 0.5D_{oij} + 1.5A_{ei} + 1.5A_{ej}$$

$$+ D_{eii} + D_{eij} + D_{eij} + C_i + A_{mi} + A_{mj} + D_{mij}$$

回交一代 BC_i ($F_{1ij} \times P_i$) 遗传效应分量 ($k = 3$)

$$G_{ij4} = 1.5A_{oi} + 0.5A_{oj} + 0.5D_{oii} + 0.5D_{oij} + 2A_{ei} + A_{ej}$$

$$+ 1.5D_{eii} + 0.5D_{eij} + D_{eij} + C_i + A_{mi} + A_{mj} + D_{mij}$$

回交一代 BC_j ($F_{1ij} \times P_j$) 遗传效应分量 ($k = 4$)

$$G_{ij3} = 0.5A_{oi} + 1.5A_{oj} + 0.5D_{oij} + 0.5D_{oij} + A_{ei} + 2A_{ej} \\ + 0.5De_{ii} + 1.5De_{jj} + De_{ij} + C_i + Am_i + Am_j + Dm_{ij}$$

反回交一代RBC_i (P_i × F_{1*ij*})遗传效应分量 (k = 5)

$$G_{ij6} = 1.5A_{oi} + 0.5A_{oj} + 0.5D_{oii} + 0.5D_{oij} + 2.5A_{ei} + 0.5A_{ej} \\ + 2De_{ii} + De_{ij} + C_i + 2Am_i + Dm_{ii}$$

反回交一代RBC_j (P_j × F_{1*ij*})遗传效应分量 (k = 6)

$$G_{ij5} = 0.5A_{oi} + 1.5A_{oj} + 0.5D_{oij} + 0.5D_{oij} + 0.5A_{ei} + 2.5A_{ej} \\ + 2De_{jj} + De_{ij} + C_j + 2Am_j + Dm_{jj}$$

其中胚直接加性效应 A_{oi} 或 $A_{oj} \sim (0, \sigma_{Ao}^2)$; 胚直接显性效应 D_{oii} 、 D_{oij} 或 $D_{oij} \sim (0, \sigma_{Do}^2)$; 胚乳直接加性效应 A_{ei} 或 $A_{ej} \sim (0, \sigma_{Ae}^2)$; 胚乳直接显性效应 De_{ii} 、 De_{jj} 或 $De_{ij} \sim (0, \sigma_{De}^2)$; 细胞质效应 $C_i \sim (0, \sigma_C^2)$; 母体加性效应 Am_i 或 $Am_j \sim (0, \sigma_{Am}^2)$; 母体显性效应 Dm_{ii} 、 Dm_{jj} 或 $Dm_{ij} \sim (0, \sigma_{Dm}^2)$ 。

虽然列出了7个遗传世代，但在实际分析时可以只采用一组亲本双列杂交的遗传材料，包括亲本、杂种一代F₁、杂种二代F₂和二个回交世代。

种子数量性状表现型方差 (V_P) 应该包括

$$V_P = V_{Go} + V_{Ge} + V_C + V_{Gm} + 2C_{Go.Ge} + 2C_{Go.Gm} + 2C_{Ge.Gm} + V_e \\ = (V_{Ao} + V_{Do}) + (V_{Ae} + V_{De}) + V_C + (V_{Am} + V_{Dm}) \\ + 2(C_{Ao.Ae} + C_{Do.De}) + 2(C_{Ao.Am} + C_{Do.Dm}) + 2(C_{Ae.Am} + C_{De.Dm}) + V_e$$

其中胚直接遗传方差 (V_{Go}) 的分量有胚直接加性方差 (V_{Ao}) 和胚直接显性方差 (V_{Do})，胚乳直接遗传方差 (V_{Ge}) 的分量有胚乳直接加性方差 (V_{Ae}) 和胚乳直接显性方差 (V_{De})， V_C 是细胞质方差，母体遗传方差 (V_{Gm}) 的分量有母体加性方差 (V_{Am}) 和母体显性方差 (V_{Dm})。胚直接遗传效应与胚乳遗传效应的协方差由两项协方差分量组成 (加性遗传协方差 $C_{Ao.Ae}$ 和显性遗传协方差 $C_{Do.De}$)。胚直接遗传效应与母体遗传效应的协方差由两项协方差分量组成 (加性遗传协方差 $C_{Ao.Am}$ 和显性遗传协方差 $C_{Do.Dm}$)。胚乳直接遗传效应与母体遗传效应的协方差由两项协方差分量组成 (加性遗传协方差 $C_{Ae.Am}$ 和显性遗传协方差

$C_{D_e, Dm}$)。

不同遗传世代的表現型方差计算公式如下：

$$\begin{aligned}
 V_P(F_1) &= (2\sigma_{A_o}^2 + \sigma_{D_o}^2) + (5\sigma_{A_e}^2 + 5\sigma_{D_e}^2) + \sigma_C^2 + (4\sigma_{A_m}^2 + \sigma_{D_m}^2) \\
 &\quad + 2(3\sigma_{A_o, A_e} + 2\sigma_{D_o, D_e}) + 2(2\sigma_{A_o, A_m}) + 2(4\sigma_{A_e, A_m} + \sigma_{D_e, D_m}) + \sigma_e^2 \\
 V_P(F_2) &= (2\sigma_{A_o}^2 + \frac{3}{8}\sigma_{D_o}^2) + (4\frac{1}{2}\sigma_{A_e}^2 + 3\sigma_{D_e}^2) + \sigma_C^2 + (2\sigma_{A_m}^2 + \sigma_{D_m}^2) \\
 &\quad + 2(3\sigma_{A_o, A_e} + \sigma_{D_o, D_e}) + 2(2\sigma_{A_o, A_m} + \frac{1}{2}\sigma_{D_o, D_m}) + 2(3\sigma_{A_e, A_m} + \sigma_{D_e, D_m}) + \sigma_e^2 \\
 V_P(BC_i) &= V_P(BC_j) \\
 &= (2\frac{1}{2}\sigma_{A_o}^2 + \frac{1}{2}\sigma_{D_o}^2) + (5\sigma_{A_e}^2 + 3\frac{1}{2}\sigma_{D_e}^2) + \sigma_C^2 + (2\sigma_{A_m}^2 + \sigma_{D_m}^2) \\
 &\quad + 2(3\frac{1}{2}\sigma_{A_o, A_e} + 1\frac{1}{4}\sigma_{D_o, D_e}) + 2(2\sigma_{A_o, A_m} + \frac{1}{2}\sigma_{D_o, D_m}) \\
 &\quad + 2(3\sigma_{A_e, A_m} + \sigma_{D_e, D_m}) + \sigma_e^2
 \end{aligned}$$

对于成对性状，表現型协方差 (C_r) 包括的协方差分量是

$$\begin{aligned}
 C_P &= C_{G_o} + C_{G_e} + C_C + C_{G_m} + 2C_{G_o/G_e} + 2C_{G_o/G_m} + 2C_{G_e/G_m} + C_e \\
 &= (C_{A_o} + C_{D_o}) + (C_{A_e} + C_{D_e}) + C_C + (C_{A_m} + C_{D_m}) \\
 &\quad + 2(C_{A_o/A_e} + C_{D_o/D_e}) + 2(C_{A_o/A_m} + C_{D_o/D_m}) \\
 &\quad + 2(C_{A_e/A_m} + C_{D_e/D_m}) + C_e
 \end{aligned}$$

不同遗传世代的表現型协方差计算公式如下：

$$\begin{aligned}
 C_P(F_1) &= (2\sigma_{A_o/A_o} + \sigma_{D_o/D_o}) + (5\sigma_{A_e/A_e} + 5\sigma_{D_e/D_e}) + \sigma_{C/C} + (4\sigma_{A_m/A_m} + \sigma_{D_m/D_m}) \\
 &\quad + 2(3\sigma_{A_o/A_e} + 2\sigma_{D_o/D_e}) + 2(2\sigma_{A_o/A_m}) + 2(4\sigma_{A_e/A_m} + \sigma_{D_e/D_m}) + \sigma_{e/e} \\
 C_P(F_2) &= (2\sigma_{A_o/A_o} + \frac{3}{8}\sigma_{D_o/D_o}) + (4\frac{1}{2}\sigma_{A_e/A_e} + 3\sigma_{D_e/D_e}) + \sigma_{C/C} + (2\sigma_{A_m/A_m} + \sigma_{D_m/D_m}) \\
 &\quad + 2(3\sigma_{A_o/A_e} + \sigma_{D_o/D_e}) + 2(2\sigma_{A_o/A_m} + \frac{1}{2}\sigma_{D_o/D_m}) + 2(3\sigma_{A_e/A_m} + \sigma_{D_e/D_m}) + \sigma_{e/e} \\
 C_P(BC_i) &= C_P(BC_j) \\
 &= (2\frac{1}{2}\sigma_{A_o/A_o} + \frac{1}{2}\sigma_{D_o/D_o}) + (5\sigma_{A_e/A_e} + 3\frac{1}{2}\sigma_{D_e/D_e}) + \sigma_{C/C} + (2\sigma_{A_m/A_m} + \sigma_{D_m/D_m}) \\
 &\quad + 2(3\frac{1}{2}\sigma_{A_o/A_e} + 1\frac{1}{4}\sigma_{D_o/D_e}) + 2(2\sigma_{A_o/A_m} + \frac{1}{2}\sigma_{D_o/D_m}) \\
 &\quad + 2(3\sigma_{A_e/A_m} + \sigma_{D_e/D_m}) + \sigma_{e/e}
 \end{aligned}$$

单子叶种子数量性状遗传研究的实验观察值可以用以下混合线性模型表示，

$$\begin{aligned} \mathbf{y} &= \mathbf{1}\mu + \mathbf{U}_{Ao}\mathbf{e}_{Ao} + \mathbf{U}_{Do}\mathbf{e}_{Do} + \mathbf{U}_{Ae}\mathbf{e}_{Ae} + \mathbf{U}_{De}\mathbf{e}_{De} + \mathbf{U}_C\mathbf{e}_C \\ &\quad + \mathbf{U}_{Am}\mathbf{e}_{Am} + \mathbf{U}_{Dm}\mathbf{e}_{Dm} + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_e \\ &= \mathbf{1}\mu + \sum_u^9 \mathbf{U}_u\mathbf{e}_u \end{aligned}$$

其中常数向量 $\mathbf{1}$ 是的系数为1的向量， \mathbf{e}_u 是第 u 项随机效应向量，

$\mathbf{e}_u \sim (0, \sigma_u^2\mathbf{I})$ ； \mathbf{U}_u 是随机效应向量 \mathbf{e}_u 的系数矩阵 ($u = 1, 2, \dots, 8$)， $\mathbf{U}_9 = \mathbf{I}$ 是单位矩阵。

表现型向量 \mathbf{y} 的方差-协方差矩阵为

$$\begin{aligned} \text{Var}(\mathbf{y}) &= \sigma_{Ao}^2 \mathbf{V}_1 + \sigma_{Do}^2 \mathbf{V}_2 + \sigma_{Ae}^2 \mathbf{V}_3 + \sigma_{De}^2 \mathbf{V}_4 + \sigma_C^2 \mathbf{V}_5 \\ &\quad + \sigma_{Am}^2 \mathbf{V}_6 + \sigma_{Dm}^2 \mathbf{V}_7 + \sigma_B^2 \mathbf{V}_8 + \sigma_{Ao.Ae} \mathbf{V}_9 + \sigma_{Do.De} \mathbf{V}_{10} \\ &\quad + \sigma_{Ao.Am} \mathbf{V}_{11} + \sigma_{Do.Dm} \mathbf{V}_{12} + \sigma_{Ae.Am} \mathbf{V}_{13} + \sigma_{De.Dm} \mathbf{V}_{14} + \sigma_e^2 \mathbf{V}_{15} \\ &= \sum_{u=1}^{15} \theta_u \mathbf{V}_u \end{aligned}$$

其中 $\mathbf{V}_u = \mathbf{U}_u \mathbf{U}_u^T$ ($u = 1, 2, \dots, 8$)、 $\mathbf{V}_9 = (\mathbf{U}_1 \mathbf{U}_3^T + \mathbf{U}_3 \mathbf{U}_1^T)$ ，

$\mathbf{V}_{10} = (\mathbf{U}_2 \mathbf{U}_4^T + \mathbf{U}_4 \mathbf{U}_2^T)$ 、 $\mathbf{V}_{11} = (\mathbf{U}_1 \mathbf{U}_6^T + \mathbf{U}_6 \mathbf{U}_1^T)$ ， $\mathbf{V}_{12} = (\mathbf{U}_2 \mathbf{U}_7^T + \mathbf{U}_7 \mathbf{U}_2^T)$ 、

$\mathbf{V}_{13} = (\mathbf{U}_3 \mathbf{U}_6^T + \mathbf{U}_6 \mathbf{U}_3^T)$ ， $\mathbf{V}_{14} = (\mathbf{U}_4 \mathbf{U}_7^T + \mathbf{U}_7 \mathbf{U}_4^T)$ 、 $\mathbf{V}_{15} = \mathbf{I}$ ； $\theta_1 = \sigma_{Ao}^2$ 、

$\theta_2 = \sigma_{Do}^2$ 、 $\theta_3 = \sigma_{Ae}^2$ 、 $\theta_4 = \sigma_{De}^2$ 、 $\theta_5 = \sigma_C^2$ 、 $\theta_6 = \sigma_{Am}^2$ 、 $\theta_7 = \sigma_{Dm}^2$ 、 $\theta_8 = \sigma_B^2$ 、

$\theta_9 = \sigma_{Ao.Ae}$ 、 $\theta_{10} = \sigma_{Do.De}$ 、 $\theta_{11} = \sigma_{Ao.Am}$ 、 $\theta_{12} = \sigma_{Do.Dm}$ 、 $\theta_{13} = \sigma_{Ae.Am}$ 、

$\theta_{14} = \sigma_{De.Dm}$ 、 $\theta_{15} = \sigma_e^2$ 。

以上混合线性模型可用MINQUE(0/1)法无偏地估算各项方差和协方差分量，也可用LUP或AUP法无偏预测各项遗传效应。

如果控制单子叶种子数量性状的基因在不同环境条件下表现不一致，需要在多种环境条件下分析性状的遗传规律，估算基因效应与环境效应的

相互作用。某一环境下单子叶种子性状的表现型值主要由环境效应(E)、基因型效应(G)和基因型 × 环境互作效应(GE)所决定。基因型 × 环境互作效应(GE)可作相应的分解,由胚种子直接遗传效应与环境效应的互作(GoE)、胚乳种子直接遗传效应与环境效应的互作(GeE)、细胞质遗传效应与环境效应的互作(CE)和母体植株遗传效应与环境效应的互作(GmE)所组成。GoE可进一步被分解为胚直接加性 × 环境互作效应(AoE)和胚直接显性 × 环境互作效应(DoE)遗传分量。GeE分解为胚乳直接加性 × 环境互作效应(AeE)和胚乳直接显性 × 环境互作效应(DeE)遗传分量。GmE则可被分解为母体加性 × 环境互作效应(AmE)、母体显性 × 环境互作效应(DmE)遗传分量。

如果假定 不存在父体效应(paternal effect), 细胞质基因只通过母体稳定遗传, 不存在上位性效应, 种子数量性状表现型值可用以下线性模型表示。一组纯系材料作双列杂交遗传交配, 第*i*个母本与第*j*个父本的第*k*种交配类型在环境*h*内第*l*个区组中的世代平均数值(y_{hijkl})可以表示为

$$y_{hijkl} = \mu + E_h + G_{ijk} + GE_{hijk} + B_{l(h)} + e_{hijkl} \quad (8.5)$$

其中 μ 是群体平均数, 固定效应; E_h 是环境效应, 固定效应; $B_{l(h)}$ 是环境内区组效应, 随机效应; e_{hijkl} 是剩余误差, 随机效应; G_{ijk} 是遗传效应, GE_{hijk} 是遗传效应与环境效应的互作。

对于单子叶作物种子数量性状, 各遗传世代基因效应 G_{ijk} 的定义如(8.4), 而遗传与环境互作效应 GE_{hijk} 的分解如下式所示。

亲本 P_i 遗传效应分量 ($k = 0$)

$$GE_{hii0} = 2AoE_{hi} + DoE_{hii} + 3AeE_{hi} + 3DeE_{hii} + CE_{hi} + 2AmE_{hi} + DmE_{hii}$$

杂种一代 F_{1ij} ($P_i \times P_j$) 遗传效应分量 ($k = 1$)

$$GE_{hij1} = AoE_{hi} + AoE_{hj} + DoE_{hij} + 2AeE_{hi} + AeE_{hj} + DeE_{hii} + 2DeE_{hij} + CE_{hi} + 2AmE_{hi} + DmE_{hii}$$

杂种二代 F_{2ij} 遗传效应分量 ($k = 2$)

$$GE_{hij2} = AoE_{hi} + AoE_{hj} + 0.25DoE_{hii} + 0.25DoE_{hij} + 0.5DoE_{hij} \\ + 1.5AeE_{hi} + 1.5AeE_{hj} + DeE_{hii} + DeE_{hij} + DeE_{hij} \\ + CE_{hi} + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

回交一代 $BC_i (F_{1ij} \times P_i)$ 遗传效应分量 ($k = 3$)

$$GE_{hij4} = 1.5AoE_{hi} + 0.5AoE_{hj} + 0.5DoE_{hii} + 0.5DoE_{hij} \\ + 2AeE_{hi} + AeE_{hj} + 1.5DeE_{hii} + 0.5DeE_{hij} + DeE_{hij} \\ + CE_{hi} + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

回交一代 $BC_j (F_{1ij} \times P_j)$ 遗传效应分量 ($k = 4$)

$$GE_{hij3} = 0.5AoE_{hi} + 1.5AoE_{hj} + 0.5DoE_{hij} + 0.5DoE_{hij} \\ + AeE_{hi} + 2AeE_{hj} + 0.5DeE_{hii} + 1.5DeE_{hij} + DeE_{hij} \\ + CE_{hi} + AmE_{hi} + AmE_{hj} + DmE_{hij}$$

反回交一代 $RBC_i (P_i \times F_{1ij})$ 遗传效应分量 ($k = 5$)

$$GE_{hij6} = 1.5AoE_{hi} + 0.5AoE_{hj} + 0.5DoE_{hii} + 0.5DoE_{hij} \\ + 2.5AeE_{hi} + 0.5AeE_{hj} + 2DeE_{hii} + DeE_{hij} + CE_{hi} \\ + 2AmE_{hi} + DmE_{hii}$$

反回交一代 $RBC_j (P_j \times F_{1ij})$ 遗传效应分量 ($k = 6$)

$$GE_{hij5} = 0.5AoE_{hi} + 1.5AoE_{hj} + 0.5DoE_{hij} + 0.5DoE_{hij} \\ + 0.5AeE_{hi} + 2.5AeE_{hj} + 2DeE_{hij} + DeE_{hij} + CE_{hj} \\ + 2AmE_{hj} + DmE_{hij}$$

其中胚直接加性×环境互作效应 AoE_{hi} 或 $AoE_{hj} \sim (0, \sigma_{AoE}^2)$; 胚直接显性×环境互作效应 DoE_{hii} 、 DoE_{hij} 或 $DoE_{hij} \sim (0, \sigma_{DoE}^2)$; 胚乳直接加性×环境互作效应 AeE_{hi} 或 $AeE_{hj} \sim (0, \sigma_{AeE}^2)$; 胚乳直接显性×环境互作效应 DeE_{hii} 、 DeE_{hij} 或 $DeE_{hij} \sim (0, \sigma_{DeE}^2)$; 细胞质×环境互作效应 $CE_{hi} \sim (0, \sigma_{CE}^2)$; 母体加性×环境互作效应 AmE_{hi} 或 $AmE_{hj} \sim (0, \sigma_{AmE}^2)$; 母体显性×环境互作效应 DmE_{hii} 、 DmE_{hij} 或 $DmE_{hij} \sim (0, \sigma_{DmE}^2)$ 。

在实际分析时可以只采用一组亲本双列杂交的遗传材料，包括亲本、杂种一代 F_1 和杂种二代 F_2 和二个回交世代在多环境下实施试验。

种子数量性状表现型方差 (V_P) 应该包括

$$\begin{aligned} V_P &= V_{Go} + V_{Ge} + V_C + V_{Gm} + V_{GoE} + V_{GeE} + V_{CE} + V_{GmE} \\ &\quad + 2C_{Go.Ge} + 2C_{Go.Gm} + 2C_{Ge.Gm} + 2C_{GoE.GeE} + 2C_{GoE.GmE} \\ &\quad + 2C_{GeE.GmE} + V_e \\ &= (V_{Ao} + V_{Do}) + (V_{Ae} + V_{De}) + V_C + (V_{Am} + V_{Dm}) \\ &\quad + (V_{AoE} + V_{DoE}) + (V_{AeE} + V_{DeE}) + V_{CE} + (V_{AmE} + V_{DmE}) \\ &\quad + 2(C_{Ao.Ae} + C_{Do.De}) + 2(C_{Ao.Am} + C_{Do.Dm}) + 2(C_{Ae.Am} + C_{De.Dm}) \\ &\quad + 2(C_{AoE.AeE} + C_{DoE.DeE}) + 2(C_{AoE.AmE} + C_{DoE.DmE}) \\ &\quad + 2(C_{AeE.AmE} + C_{DeE.DmE}) + V_e \end{aligned}$$

不同遗传世代的表现型方差计算公式如下：

$$\begin{aligned} V_P(F_1) &= (2\sigma_{Ao}^2 + \sigma_{Do}^2) + (5\sigma_{Ae}^2 + 5\sigma_{De}^2) + \sigma_C^2 + (4\sigma_{Am}^2 + \sigma_{Dm}^2) \\ &\quad + (2\sigma_{AoE}^2 + \sigma_{DoE}^2) + (5\sigma_{AeE}^2 + 5\sigma_{DeE}^2) + \sigma_{CE}^2 + (4\sigma_{AmE}^2 + \sigma_{DmE}^2) \\ &\quad + 2(3\sigma_{Ao.Ae} + 2\sigma_{Do.De}) + 2(2\sigma_{Ao.Am}) + 2(4\sigma_{Ae.Am} + \sigma_{De.Dm}) \\ &\quad + 2(3\sigma_{AoE.AeE} + 2\sigma_{DoE.DeE}) + 2(2\sigma_{AoE.AmE}) \\ &\quad + 2(4\sigma_{AeE.AmE} + \sigma_{DeE.DmE}) + \sigma_e^2 \end{aligned}$$

$$\begin{aligned} V_P(F_2) &= (2\sigma_{Ao}^2 + \frac{3}{8}\sigma_{Do}^2) + (4\frac{1}{2}\sigma_{Ae}^2 + 3\sigma_{De}^2) + \sigma_C^2 + (2\sigma_{Am}^2 + \sigma_{Dm}^2) \\ &\quad + (2\sigma_{AoE}^2 + \frac{3}{8}\sigma_{DoE}^2) + (4\frac{1}{2}\sigma_{AeE}^2 + 3\sigma_{DeE}^2) + \sigma_{CE}^2 + (2\sigma_{AmE}^2 + \sigma_{DmE}^2) \\ &\quad + 2(3\sigma_{Ao.Ae} + \sigma_{Do.De}) + 2(2\sigma_{Ao.Am} + \frac{1}{2}\sigma_{Do.Dm}) + 2(3\sigma_{Ae.Am} + \sigma_{De.Dm}) \\ &\quad + 2(3\sigma_{AoE.AeE} + \sigma_{DoE.DeE}) + 2(2\sigma_{AoE.AmE} + \frac{1}{2}\sigma_{DoE.DmE}) \\ &\quad + 2(3\sigma_{AeE.AmE} + \sigma_{DeE.DmE}) + \sigma_e^2 \end{aligned}$$

$$\begin{aligned} V_P(BC_i) &= V_P(BC_j) \\ &= (2\frac{1}{2}\sigma_{Ao}^2 + \frac{1}{2}\sigma_{Do}^2) + (5\sigma_{Ae}^2 + 3\frac{1}{2}\sigma_{De}^2) + \sigma_C^2 + (2\sigma_{Am}^2 + \sigma_{Dm}^2) \\ &\quad + (2\frac{1}{2}\sigma_{AoE}^2 + \frac{1}{2}\sigma_{DoE}^2) + (5\sigma_{AeE}^2 + 3\frac{1}{2}\sigma_{DeE}^2) + \sigma_{CE}^2 + (2\sigma_{AmE}^2 + \sigma_{DmE}^2) \\ &\quad + 2(3\frac{1}{2}\sigma_{Ao.Ae} + 1\frac{1}{4}\sigma_{Do.De}) + 2(2\sigma_{Ao.Am} + \frac{1}{2}\sigma_{Do.Dm}) + 2(3\sigma_{Ae.Am} + \sigma_{De.Dm}) \\ &\quad + 2(3\frac{1}{2}\sigma_{AoE.AeE} + 1\frac{1}{4}\sigma_{DoE.DeE}) + 2(2\sigma_{AoE.AmE} + \frac{1}{2}\sigma_{DoE.DmE}) \\ &\quad + 2(3\sigma_{AeE.AmE} + \sigma_{DeE.DmE}) + \sigma_e^2 \end{aligned}$$

对于成对性状，表现型协方差 (C_p) 包括的协方差分量是

$$\begin{aligned}
 C_p &= C_{Go} + C_{Ge} + C_C + C_{Gm} + C_{GoE} + C_{GeE} + C_{CE} + C_{GmE} \\
 &\quad + 2C_{Go/Ge} + 2C_{Go/Gm} + 2C_{Ge/Gm} + 2C_{GoE/GeE} + 2C_{GoE/GmE} \\
 &\quad + 2C_{GeE/GmE} + C_e \\
 &= (C_{Ao} + C_{Do}) + (C_{Ae} + C_{De}) + C_C + (C_{Am} + C_{Dm}) \\
 &\quad + (C_{AoE} + C_{DoE}) + (C_{AeE} + C_{DeE}) + C_{CE} + (C_{AmE} + C_{DmE}) \\
 &\quad + 2(C_{Ao/Ae} + C_{Do/De}) + 2(C_{Ao/Am} + C_{Do/Dm}) + 2(C_{Ae/Am} + C_{De/Dm}) \\
 &\quad + 2(C_{AoE/AeE} + C_{DoE/DeE}) + 2(C_{AoE/AmE} + C_{DoE/DmE}) \\
 &\quad + 2(C_{AeE/AmE} + C_{DeE/DmE}) + C_e
 \end{aligned}$$

不同遗传世代的表現型协方差计算公式如下：

$$\begin{aligned}
 C_p(F_1) &= (2\sigma_{Ao/Ao} + \sigma_{Do/Do}) + (5\sigma_{Ae/Ae} + 5\sigma_{De/De}) + \sigma_{C/C} + (4\sigma_{Am/Am} + \sigma_{Dm/dm}) \\
 &\quad + (2\sigma_{AoE/AoE} + \sigma_{DoE/DoE}) + (5\sigma_{AeE/AeE} + 5\sigma_{DeE/DeE}) + \sigma_{CE/CE} \\
 &\quad + (4\sigma_{AmE/AmE} + \sigma_{DmE/DmE}) + 2(3\sigma_{Ao/Ae} + 2\sigma_{Do/De}) + 2(2\sigma_{Ao/Am}) \\
 &\quad + 2(4\sigma_{Ae/Am} + \sigma_{De/Dm}) + 2(3\sigma_{AoE/AeE} + 2\sigma_{DoE/DeE}) + 2(2\sigma_{AoE/AmE}) \\
 &\quad + 2(4\sigma_{AeE/AmE} + \sigma_{DeE/DmE}) + \sigma_{e/e} \\
 C_p(F_2) &= (2\sigma_{Ao/Ao} + \frac{3}{8}\sigma_{Do/Do}) + (4\frac{1}{2}\sigma_{Ae/Ae} + 3\sigma_{De/De}) + \sigma_{C/C} \\
 &\quad + (2\sigma_{Am/Am} + \sigma_{Dm/Dm}) + (2\sigma_{AoE/AoE} + \frac{3}{8}\sigma_{DoE/DoE}) + (4\frac{1}{2}\sigma_{AeE/AeE} + 3\sigma_{DeE/DeE}) \\
 &\quad + \sigma_{CE/CE} + (2\sigma_{AmE/AmE} + \sigma_{DmE/DmE}) + 2(3\sigma_{Ao/Ae} + \sigma_{Do/De}) \\
 &\quad + 2(2\sigma_{Ao/Am} + \frac{1}{2}\sigma_{Do/Dm}) + 2(3\sigma_{Ae/Am} + \sigma_{De/Dm}) + 2(3\sigma_{AoE/AeE} + \sigma_{DoE/DeE}) \\
 &\quad + 2(2\sigma_{AoE/AmE} + \frac{1}{2}\sigma_{DoE/DmE}) + 2(3\sigma_{AeE/AmE} + \sigma_{DeE/DmE}) + \sigma_{e/e} \\
 C_p(BC_i) &= C_p(BC_j) \\
 &= (2\frac{1}{2}\sigma_{Ao/Ao} + \frac{1}{2}\sigma_{Do/Do}) + (5\sigma_{Ae/Ae} + 3\frac{1}{2}\sigma_{De/De}) + \sigma_{CC} + (2\sigma_{Am/Am} + \sigma_{Dm/Dm}) \\
 &\quad + (2\frac{1}{2}\sigma_{AoE/AoE} + \frac{1}{2}\sigma_{DoE/DoE}) + (5\sigma_{AeE/AeE} + 3\frac{1}{2}\sigma_{DeE/DeE}) + \sigma_{CE/CE} \\
 &\quad + (2\sigma_{AmE/AmE} + \sigma_{DmE/DmE}) + 2(3\frac{1}{2}\sigma_{Ao/Ae} + 1\frac{1}{4}\sigma_{Do/De}) + 2(2\sigma_{Ao/Am} + \frac{1}{2}\sigma_{Do/Dm}) \\
 &\quad + 2(3\sigma_{Ae/Am} + \sigma_{De/Dm}) + 2(3\frac{1}{2}\sigma_{AoE/AeE} + 1\frac{1}{4}\sigma_{DoE/DeE}) \\
 &\quad + 2(2\sigma_{AoE/AmE} + \frac{1}{2}\sigma_{DoE/DmE}) + 2(3\sigma_{AeE/AmE} + \sigma_{DeE/DmE}) + \sigma_{e/e}
 \end{aligned}$$

单子叶种子数量性状在多环境下的遗传研究实验观察值可以用以下混

合线性模型表示，

$$\begin{aligned}
 \mathbf{y} &= \mathbf{X}\mathbf{b} + \mathbf{U}_{Ao}\mathbf{e}_{Ao} + \mathbf{U}_{Do}\mathbf{e}_{Do} + \mathbf{U}_{Ae}\mathbf{e}_{Ae} + \mathbf{U}_{De}\mathbf{e}_{De} + \mathbf{U}_C\mathbf{e}_C \\
 &\quad + \mathbf{U}_{Am}\mathbf{e}_{Am} + \mathbf{U}_{Dm}\mathbf{e}_{Dm} \\
 &\quad + \mathbf{U}_{AoE}\mathbf{e}_{AoE} + \mathbf{U}_{DoE}\mathbf{e}_{DoE} + \mathbf{U}_{AeE}\mathbf{e}_{AeE} + \mathbf{U}_{DeE}\mathbf{e}_{DeE} + \mathbf{U}_{CE}\mathbf{e}_{CE} \\
 &\quad + \mathbf{U}_{AmE}\mathbf{e}_{AmE} + \mathbf{U}_{DmE}\mathbf{e}_{DmE} + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_e \\
 &= \mathbf{X}\mathbf{b} + \sum_u^{16} \mathbf{U}_u \mathbf{e}_u
 \end{aligned}$$

表现型向量 \mathbf{y} 的方差×协方差矩阵为

$$\begin{aligned}
 \text{Var}(\mathbf{y}) &= \sigma_{Ao}^2 \mathbf{V}_1 + \sigma_{Do}^2 \mathbf{V}_2 + \sigma_{Ae}^2 \mathbf{V}_3 + \sigma_{De}^2 \mathbf{V}_4 + \sigma_C^2 \mathbf{V}_5 \\
 &\quad + \sigma_{Am}^2 \mathbf{V}_6 + \sigma_{Dm}^2 \mathbf{V}_7 + \sigma_{AoE}^2 \mathbf{V}_8 + \sigma_{DoE}^2 \mathbf{V}_9 + \sigma_{AeE}^2 \mathbf{V}_{10} \\
 &\quad + \sigma_{DeE}^2 \mathbf{V}_{11} + \sigma_{CE}^2 \mathbf{V}_{12} + \sigma_{AmE}^2 \mathbf{V}_{13} + \sigma_{DmE}^2 \mathbf{V}_{14} + \sigma_B^2 \mathbf{V}_{15} \\
 &\quad + \sigma_{Ao.Ae} \mathbf{V}_{16} + \sigma_{Do.De} \mathbf{V}_{17} + \sigma_{Ao.Am} \mathbf{V}_{18} + \sigma_{Do.Dm} \mathbf{V}_{19} \\
 &\quad + \sigma_{Ae.Am} \mathbf{V}_{20} + \sigma_{De.Dm} \mathbf{V}_{21} + \sigma_{AoE.AeE} \mathbf{V}_{22} + \sigma_{DoE.DeE} \mathbf{V}_{23} \\
 &\quad + \sigma_{AoE.AmE} \mathbf{V}_{24} + \sigma_{DoE.DmE} \mathbf{V}_{25} + \sigma_{AeE.AmE} \mathbf{V}_{26} \\
 &\quad + \sigma_{DeE.DmE} \mathbf{V}_{27} + \sigma_e^2 \mathbf{V}_{28} \\
 &= \sum_{u=1}^{28} \theta_u \mathbf{V}_u
 \end{aligned}$$

其中 $\mathbf{V}_u = \mathbf{U}_u \mathbf{U}_u^T$ ($u = 1, 2, \dots, 15$)、 $\mathbf{V}_{16} = (\mathbf{U}_1 \mathbf{U}_3^T + \mathbf{U}_3 \mathbf{U}_1^T)$ 、
 $\mathbf{V}_{17} = (\mathbf{U}_2 \mathbf{U}_4^T + \mathbf{U}_4 \mathbf{U}_2^T)$ 、 $\mathbf{V}_{18} = (\mathbf{U}_1 \mathbf{U}_6^T + \mathbf{U}_6 \mathbf{U}_1^T)$ 、
 $\mathbf{V}_{19} = (\mathbf{U}_2 \mathbf{U}_7^T + \mathbf{U}_7 \mathbf{U}_2^T)$ 、 $\mathbf{V}_{20} = (\mathbf{U}_3 \mathbf{U}_6^T + \mathbf{U}_6 \mathbf{U}_3^T)$ 、 $\mathbf{V}_{21} = (\mathbf{U}_4 \mathbf{U}_7^T + \mathbf{U}_7 \mathbf{U}_4^T)$ 、
 $\mathbf{V}_{22} = (\mathbf{U}_8 \mathbf{U}_{10}^T + \mathbf{U}_{10} \mathbf{U}_8^T)$ 、 $\mathbf{V}_{23} = (\mathbf{U}_9 \mathbf{U}_{11}^T + \mathbf{U}_{11} \mathbf{U}_9^T)$ 、
 $\mathbf{V}_{24} = (\mathbf{U}_8 \mathbf{U}_{13}^T + \mathbf{U}_{13} \mathbf{U}_8^T)$ 、 $\mathbf{V}_{25} = (\mathbf{U}_9 \mathbf{U}_{14}^T + \mathbf{U}_{14} \mathbf{U}_9^T)$ 、
 $\mathbf{V}_{26} = (\mathbf{U}_{10} \mathbf{U}_{13}^T + \mathbf{U}_{13} \mathbf{U}_{10}^T)$ 、 $\mathbf{V}_{27} = (\mathbf{U}_{11} \mathbf{U}_{14}^T + \mathbf{U}_{14} \mathbf{U}_{11}^T)$ 、 $\mathbf{V}_{28} = \mathbf{I}$ ；
 $\theta_1 = \sigma_{Ao}^2$ 、 $\theta_2 = \sigma_{Do}^2$ 、 $\theta_3 = \sigma_{Ae}^2$ 、 $\theta_4 = \sigma_{De}^2$ 、 $\theta_5 = \sigma_C^2$ 、 $\theta_6 = \sigma_{Am}^2$ 、
 $\theta_7 = \sigma_{Dm}^2$ 、 $\theta_8 = \sigma_{AoE}^2$ 、 $\theta_9 = \sigma_{DoE}^2$ 、 $\theta_{10} = \sigma_{AeE}^2$ 、 $\theta_{11} = \sigma_{DeE}^2$ 、 $\theta_{12} = \sigma_{CE}^2$ 、
 $\theta_{13} = \sigma_{AmE}^2$ 、 $\theta_{14} = \sigma_{DmE}^2$ 、 $\theta_{15} = \sigma_B^2$ 、 $\theta_{16} = \sigma_{Ao.Ae}$ 、 $\theta_{17} = \sigma_{Do.De}$ 、
 $\theta_{18} = \sigma_{Ao.Am}$ 、 $\theta_{19} = \sigma_{Do.Dm}$ 、 $\theta_{20} = \sigma_{Ae.Am}$ 、 $\theta_{21} = \sigma_{De.Dm}$ 、 $\theta_{22} = \sigma_{AoE.AeE}$ 、

$$\theta_{23} = \sigma_{\text{DoE.DeE}}、\theta_{24} = \sigma_{\text{AoE.AmE}}、\theta_{25} = \sigma_{\text{DoE.DmE}}、\theta_{26} = \sigma_{\text{AcE.AmE}}、$$

$$\theta_{27} = \sigma_{\text{DeE.DmE}}、\theta_{28} = \sigma_{\text{e}}^2。$$

以上混合线性模型可用MINQUE(0/1)法无偏地估算各项方差和协方差分量，也可用LUP或AUP法无偏预测各项遗传效应。

第九章 发育性状的遗传分析

第一节 发育性状的净遗传效应及条件方差分量

发育遗传学中有关数量性状的资料分析，常需要研究某一段时间内($t-1$) $\rightarrow t$ 的基因效应的表现。研究基因在特定时刻的表达及其对数量性状的影响，是发育数量遗传学的一项重要研究内容。

分析连续发育阶段的遗传规律目前常采用两种方法：

(1)分析发育性状在 t 时刻的遗传变异性，其方差分量 $\sigma_{(t)}^2$ 度量了从时间 $0 \rightarrow t$ 全发育阶段的变异性。

发育遗传常分析数量性状在 t 时刻($t=1, 2, \dots$)的观察值向量 $\mathbf{y}_{(t)}$ ，该变量可用以下混合线性模型表达并具有多变量正态分布：

$$\begin{aligned}\mathbf{y}_{(t)} &= \mathbf{X}\mathbf{b}_{(t)} + \sum_{u=1}^m \mathbf{U}_u \mathbf{e}_{u(t)} \\ &= \mathbf{X}\mathbf{b}_{(t)} + \mathbf{U}\mathbf{e}_{(t)} \\ &\sim \text{MVN}(\mathbf{X}\mathbf{b}_{(t)}, \mathbf{V}_{(t)})\end{aligned}$$

其中 $\mathbf{b}_{(t)}$ 是在 t 时刻的固定效应向量， \mathbf{X} 是固定效应的系数矩阵； $\mathbf{e}_{u(t)}$ 是在 t 时刻第 u 项独立随机变量向量，具有平均数零、方差 $\sigma_{u(t)}^2 \mathbf{I}$ 。 \mathbf{U}_u 是第 u 项随机效应的系数矩阵。

采用混合线性模型的分析方法(朱军, 1992)，可以无偏地估算以上统计模型中的各项随机效应的方差分量 $\sigma_{u(t)}^2$ 。如果把数量性状在($t-1$)时刻的察值向量 $\mathbf{y}_{(t-1)}$ 与该性状在 t 时刻的察值向量 $\mathbf{y}_{(t)}$ 当作成对性状分析，则可以无偏估算在($t-1$)时刻的随机效应 $\mathbf{e}_{u(t-1)}$ 与 t 时刻的随机效应 $\mathbf{e}_{u(t)}$ 的协方差分量 $\sigma_{u(t-1)(t)} = \sigma_{u(t)(t-1)}$ 。

(2)分析发育性状在($t-1$)和 t 两个时间的观察值差值。

在发育遗传研究中，常需要分析从 $(t-1) \rightarrow t$ 这段时间内生物体所表现的遗传变异性。通常的方法是分析从 $(t-1) \rightarrow t$ 这段时间内增加的表现型值，即前后两个时间的观察值的差值 $\mathbf{y}_{(d)} = \mathbf{y}_{(t)} - \mathbf{y}_{(t-1)}$ 。由于发育性状在前后两个阶段的表现存在相关性，变量 $\mathbf{y}_{(t)}$ 的变异性在一定程度上受到变量 $\mathbf{y}_{(t-1)}$ 的影响。这两个变量的协方差矩阵为

$$\text{Cov}(\mathbf{y}_{(t)}, \mathbf{y}_{(t-1)}^T) = \mathbf{C}_{(t, t-1)} = \mathbf{C}_{(t-1, t)}$$

差值 $\mathbf{y}_{(d)} = \mathbf{y}_{(t)} - \mathbf{y}_{(t-1)}$ 的方差是

$$\begin{aligned} \text{Var}(\mathbf{y}_{(d)}) &= \text{Var}(\mathbf{y}_{(t)}) + \text{Var}(\mathbf{y}_{(t-1)}) - 2\text{Cov}(\mathbf{y}_{(t)}, \mathbf{y}_{(t-1)}^T) \\ &= \mathbf{V}_{(t)} + \mathbf{V}_{(t-1)} - 2\mathbf{C}_{(t, t-1)} \end{aligned}$$

采用这种方法所分析的方差分量 $\sigma_{u(d)}^2$ 并未真正度量从 $(t-1) \rightarrow t$ 这段时间的变异性，因为

$$\sigma_{u(d)}^2 = \sigma_{u(t)}^2 + \sigma_{u(t-1)}^2 - 2\sigma_{u(t, t-1)}$$

以上两种传统的方法都不能有效地分析某一段时间的基因表达效应。在多变量统计分析中，条件方差(conditional variance)可度量给定某一变量观察值时的条件变量(conditional variable)的变异性。条件变量 $\mathbf{y}_{(t)}|\mathbf{y}_{(t-1)}$ 与该性状在 $(t-1)$ 时刻的表现型值 $\mathbf{y}_{(t-1)}$ 是相互独立的，因此条件遗传效应及其条件方差度量了从 $(t-1) \rightarrow t$ 这段时间的净遗传变异性。Zhu(1995)运用混合线性模型的分析原理，提出了估算条件遗传方差分量和预测条件遗传效应值的统计分析方法。

给定 $t-1$ 时刻的表现型值 $\mathbf{y}_{(t-1)}$ ， t 时刻的条件表现型值 $\mathbf{y}_{(t)}|\mathbf{y}_{(t-1)}$ 是条件随机变量，

$$\begin{aligned} \mathbf{y}_{(t)}|\mathbf{y}_{(t-1)} &= \mathbf{X}\mathbf{b}_{(t|t-1)} + \sum_{u=1}^m \mathbf{U}_u \mathbf{e}_{u(t|t-1)} \\ &= \mathbf{X}\mathbf{b}_{(t|t-1)} + \mathbf{U}\mathbf{e}_{(t|t-1)} \\ &\sim \text{MVN}(\mathbf{X}\mathbf{b}_{(t|t-1)}, \mathbf{V}_{(t|t-1)}) \end{aligned}$$

其中 $\mathbf{b}_{(t|t-1)}$ 是在 t 时刻的条件固定效应向量, \mathbf{X} 是条件固定效应的系数矩阵; $\mathbf{e}_{u(t|t-1)}$ 是在 t 时刻第 u 项条件随机变量向量, 具有平均数零、方差 $\sigma_{u(t|t-1)}^2 \mathbf{I}$ 。 \mathbf{U}_u 是第 u 项条件随机效应的系数矩阵。

条件随机变量 $\mathbf{y}_{(t)}|\mathbf{y}_{(t-1)}$ 具有条件方差矩阵,

$$\mathbf{V}_{(t|t-1)} = \mathbf{V}_{(t)} - \mathbf{C}_{(t-1,t)} \mathbf{V}_{(t-1)}^{-1} \mathbf{C}_{(t,t-1)}$$

该条件方差矩阵可以由下式直接估算

$$\hat{\mathbf{V}}_{(t|t-1)} = \hat{\mathbf{V}}_{(t)} - \hat{\mathbf{C}}_{(t-1,t)} \hat{\mathbf{V}}_{(t-1)}^{-1} \hat{\mathbf{C}}_{(t,t-1)}$$

其中方差矩阵和协方差矩阵的无偏估计可以由 MINQUE 法算得。但是无法直接获得各项遗传效应的条件方差分量 $\sigma_{u(t|t-1)}^2$, 也不能直接得到条件随机变量 $\mathbf{e}_{u(t|t-1)}$ 的预测值。

第二节 发育性状的遗传分析方法

Zhu(1995)建议采用间接估算的方法, 估算混合线性模型中条件随机效应的条件方差分量。按下式可得到新的变量,

$$\mathbf{y}^{(*)} = \mathbf{y}_{(t)} - \mathbf{C}_{(t-1,t)} \mathbf{V}_{(t-1)}^{-1} (\mathbf{y}_{(t-1)} - \mathbf{X}\mathbf{b}_{(t-1)})$$

新变量 $\mathbf{y}^{(*)}$ 的方差等于条件随机变量 $\mathbf{y}_{(t)}|\mathbf{y}_{(t-1)}$ 的条件方差矩阵。

$$\begin{aligned} \text{Var}(\mathbf{y}^{(*)}) &= \mathbf{V}^{(*)} \\ &= \mathbf{V}_{(t)} - \mathbf{C}_{(t-1)(t)} \mathbf{V}_{(t-1)}^{-1} \mathbf{C}_{(t)(t-1)} \\ &= \mathbf{V}_{(t|t-1)} \end{aligned}$$

可以证明 $\mathbf{y}_{(*)}$ 与 $t-1$ 时刻的表现型值 $\mathbf{y}_{(t-1)}$ 是相互独立的, 其协方差 $\text{Cov}(\mathbf{y}_{(*)}, \mathbf{y}_{(t-1)})$ 为零矩阵。

变量 $\mathbf{y}_{(*)}$ 的估算需要知道方差矩阵和协方差矩阵, 而这些参数是未知的。在实际应用时, 可用方差矩阵和协方差矩阵的无偏估计值代替参数值计算变量 $\mathbf{y}_{(*)}$ 。然后按以下模型

$$\begin{aligned}\mathbf{y}_{(*)} &= \mathbf{X}\mathbf{b}_{(*)} + \sum_{u=1}^m \mathbf{U}_u \mathbf{e}_{u(*)} \\ &= \mathbf{X}\mathbf{b}_{(*)} + \mathbf{U}\mathbf{e}_{(*)} \\ &\sim \text{MVN}(\mathbf{X}\mathbf{b}_{(*)}, \mathbf{V}_{(*)})\end{aligned}$$

估算模型中的各项条件方差分量 $\sigma_{u(*)}^2$, 并可采用 LUP 法(朱军, 1992)或 AUP 法(朱军, 1993)预测条件随机效应值 $\mathbf{e}_{u(*)}$ 。

各项条件方差分量 $\sigma_{u(*)}^2$ 度量了从时间 $(t-1) \rightarrow t$ 阶段的净随机效应 $\mathbf{e}_{u(*)}$ 的变异程度, 即净变异分量。 $\sigma_{u(*)}^2 = \sigma_{u(t|t-1)}^2$ 及 $\mathbf{e}_{u(*)} = \mathbf{e}_{u(t|t-1)}$ 是等量的。

在数量性状的发育遗传分析中, 常需要研究从发育阶段 t 到最终时刻 (f) 这一段时间内 $(t \rightarrow f)$ 基因表达的净效应。如果按下式定义新的变量,

$$\mathbf{y}_{(*)} = \mathbf{y}_{(f)} - \mathbf{C}_{(t,f)} \mathbf{V}_{(t)}^{-1} (\mathbf{y}_{(t)} - \mathbf{X}\mathbf{b}_{(t)})$$

新变量 $\mathbf{y}_{(*)}$ 的方差等于条件随机变量 $\mathbf{y}_{(f)} | \mathbf{y}_{(t)}$ 的条件方差矩阵,

$$\begin{aligned}\text{Var}(\mathbf{y}_{(*)}) &= \mathbf{V}_{(*)} \\ &= \mathbf{V}_{(f)} - \mathbf{C}_{(t,f)} \mathbf{V}_{(t)}^{-1} \mathbf{C}_{(f,t)} \\ &= \mathbf{V}_{(f|t)}\end{aligned}$$

其条件方差矩阵中的各项条件方差分量 $\sigma_{u(*)}^2$ 度量了从时间 t 到最终时刻 f 的净随机效应 $\mathbf{e}_{u(*)}$ 的变异程度, 即净变异分量。 $\sigma_{u(*)}^2 = \sigma_{u(f|t)}^2$ 及 $\mathbf{e}_{u(*)} =$

$\mathbf{e}_{u(f|t)}$ 是等量的。

第三节 发育性状的分析实例

棉花最终结铃数是整个花铃期成铃的总和。棉花某一时期的铃数决定于前一时期的铃数以及当时的遗传和环境影响。本节以陆地棉的两年遗传实验资料，分析其单株成铃数的发育遗传规律，作为分析应用实例。

采用包括与环境互作效应的加性—显性遗传模型，分析了陆地棉 4 个亲本(P1=中棉所 7 号、P2=岱字棉 15、P3=无腺体的 GL—15、P4=无蜜腺的 HG—H—12)的两年(1981, 1985)双列杂交不同时期单株成铃数的资料，估算了各项遗传方差分量。每年从 7 月 5 日—9 月 3 日每隔五天调查各小区的单株成铃数。

棉花单株成铃数在 t 时刻($t=1, 2, \dots, f$)的观察值向量 $\mathbf{y}_{(t)}$ 的遗传模型是

$$\begin{aligned} \mathbf{y}_{(t)} &= \mathbf{X}\mathbf{b}_{(t)} + \mathbf{U}_A \mathbf{e}_{A(t)} + \mathbf{U}_D \mathbf{e}_{D(t)} + \mathbf{U}_{AE} \mathbf{e}_{AE(t)} + \mathbf{U}_{DE} \mathbf{e}_{DE(t)} + \mathbf{U}_B \mathbf{e}_{B(t)} + \mathbf{e}_{\varepsilon(t)} \\ &= \mathbf{X}\mathbf{b}_{(t)} + \sum_{u=1}^6 \mathbf{U}_u \mathbf{e}_{u(t)} \\ &\sim \text{MVN}(\mathbf{X}\mathbf{b}_{(t)}, \mathbf{V}_{(t)} = \sum_{u=1}^6 \sigma_{u(t)}^2 \mathbf{U}_u \mathbf{U}_u^T) \end{aligned}$$

其中 $\mathbf{b}_{(t)}$ 是 t 时刻的环境固定效应， \mathbf{X} 是环境效应的系数矩阵；

$\mathbf{e}_{A(t)}$ 是 t 时刻的加性效应向量， $\mathbf{e}_{A(t)} \sim (0, \sigma_{A(t)}^2 \mathbf{I})$ ；

\mathbf{U}_A 是加性效应的系数矩阵；

$\mathbf{e}_{D(t)}$ 是 t 时刻的显性效应向量， $\mathbf{e}_{D(t)} \sim (0, \sigma_{D(t)}^2 \mathbf{I})$ ；

\mathbf{U}_D 是显性效应的系数矩阵；

$\mathbf{e}_{AE(t)}$ 是 t 时刻的加性×环境互作效应向量， $\mathbf{e}_{AE(t)} \sim (0, \sigma_{AE(t)}^2 \mathbf{I})$ ；

\mathbf{U}_{AE} 是加性×环境互作效应的系数矩阵;

$\mathbf{e}_{DE(t)}$ 是 t 时刻的显性×环境互作效应向量, $\mathbf{e}_{DE(t)} \sim (0, \sigma_{DE(t)}^2 \mathbf{I})$;

\mathbf{U}_{DE} 是显性×环境互作效应的系数矩阵;

$\mathbf{e}_{B(t)}$ 是 t 时刻的区组效应向量, $\mathbf{e}_{B(t)} \sim (0, \sigma_{B(t)}^2 \mathbf{I})$;

\mathbf{U}_B 是区组效应的系数矩阵;

$\mathbf{e}_{\mathcal{E}(t)}$ 是 t 时刻的剩余效应向量, $\mathbf{e}_{\mathcal{E}(t)} \sim (0, \sigma_{\mathcal{E}(t)}^2 \mathbf{I})$ 。

采用 MINQUE(1)法可以无偏地估算不同时期的遗传方差分量 $\hat{\sigma}_{u(t)}^2$ 以及间隔五天的遗传协方差分量 $\hat{\sigma}_{u(t-1,t)}$ 。并可进一步估算得到方差矩阵和协方差矩阵,

$$\hat{\mathbf{V}}_{(t-1)} = \sum_{u=1}^6 \hat{\sigma}_{u(t-1)}^2 \mathbf{U}_u \mathbf{U}_u^T$$

$$\hat{\mathbf{C}}_{(t,t-1)} = \hat{\mathbf{C}}_{(t-1,t)} = \sum_{u=1}^6 \hat{\sigma}_{u(t,t-1)} \mathbf{U}_u \mathbf{U}_u^T$$

然后计算新变量

$$\mathbf{y}_{(*)} = \mathbf{y}_{(t)} - \hat{\mathbf{C}}_{(t-1)(t)} \hat{\mathbf{V}}_{(t-1)}^{-1} (\mathbf{y}_{(t-1)} - \mathbf{X} \hat{\mathbf{b}}_{(t-1)})$$

其中 $\hat{\mathbf{b}}$ 由年份效应的平均数算得。

按照以下遗传模型

$$\begin{aligned} \mathbf{y}_{(*)} &= \mathbf{X} \mathbf{b}_{(*)} + \mathbf{U}_A \mathbf{e}_{A(*)} + \mathbf{U}_D \mathbf{e}_{D(*)} + \mathbf{U}_{AE} \mathbf{e}_{AE(*)} + \mathbf{U}_{DE} \mathbf{e}_{DE(*)} + \mathbf{U}_B \mathbf{e}_{B(*)} + \mathbf{e}_{\mathcal{E}(*)} \\ &= \mathbf{X} \mathbf{b}_{(*)} + \sum_{u=1}^6 \mathbf{U}_u \mathbf{e}_{u(*)} \\ &\sim \text{MVN}(\mathbf{X} \mathbf{b}_{(*)}, \mathbf{V}_{(t)}) = \sum_{u=1}^6 \sigma_{u(*)}^2 \mathbf{U}_u \mathbf{U}_u^T \end{aligned}$$

采用 MINQUE 法估算各项遗传方差分量 $\hat{\sigma}_{u(*)}^2$ 作为条件方差分量 $\sigma_{u(t|t-1)}^2$ 的估计值, 并采用 AUP 法获得随机效应预测值 $\hat{\mathbf{e}}_{u(*)}$ 作为条件遗传效应 $\mathbf{e}_{u(t|t-1)}$ 的

预测。

表 9.1 列出了陆地棉 4 个亲本及其 F_1 不同发育时期的平均单株成铃数。早熟亲本 P3 及其 F_1 在早期有较多的结铃数，但进入中后期结铃数锐减。晚熟亲本 P2 及其 F_1 在早期则结铃较少，中后期结铃数增加较多。中熟亲本及其 F_1 的结铃表现也略不同于早、晚熟材料。由此可见，各种遗传材料在不同时期的基因表达可能不完全相同，有待进一步作遗传分析。

表 9.1 陆地棉 4 个亲本及其 F_1 不同发育时期的平均单株成铃数

日期	P1	$F_{1(1)}$	P2	$F_{1(2)}$	P3	$F_{1(3)}$	P4	$F_{1(4)}$
7/05	0.51	0.73	0.07	0.47	0.96	1.09	0.44	0.73
7/10	1.48	1.95	0.73	1.95	2.07	1.95	1.47	1.95
7/15	2.65	3.57	1.44	3.49	2.63	4.49	2.43	3.81
7/20	3.20	5.03	2.13	5.13	3.16	6.69	3.13	5.64
7/25	4.05	5.51	2.87	5.90	3.60	7.49	4.43	6.36
7/30	5.25	6.01	4.23	6.60	4.43	8.00	6.34	6.82
8/04	7.26	7.67	6.87	8.25	5.02	9.12	9.66	8.55
8/09	9.88	9.25	10.14	10.06	5.56	10.38	12.38	10.77
8/14	12.34	10.73	12.86	12.00	6.89	11.57	14.73	12.59
8/19	13.30	11.77	14.67	13.33	7.69	12.21	15.90	13.71
8/24	14.23	12.11	15.87	13.80	8.53	12.50	16.70	14.16
8/29	15.10	12.46	16.80	14.30	9.10	12.71	17.19	14.45
9/03	15.91	12.70	17.58	14.70	9.46	12.75	17.55	14.66

陆地棉不同发育时期单株成铃数的方差分量和条件方差分量的估计值列于表 9.2。

方差分量的分析结果表明，在不同发育时期基因对单株成铃数的效应表现不同。在整个发育时期基因型方差($V_G = V_A + V_D$)明显地大于基因型 \times 环境互作方差($V_{GE} = V_{AE} + V_{DE}$)。在初花期(7 月 5 日)，显性方差($\hat{V}_D =$

0.17 ± 0.04)和加性方差($\hat{V}_A = 0.11 \pm 0.03$)是主要的变异分量。 V_A 逐渐下降, 7月15日已降为零, 并且直到8月9日被测到, 然后一直呈上升趋势。 V_D 先渐增加, 至7月30日达到最高值($\hat{V}_D = 4.00 \pm 1.04$), 然后渐渐地下降。在棉花发育的早期, 显性方差是遗传变异的主要分量, 但是到了中后期加性方差成了主要的分量。在立秋前后, 存在显著的加性×环境互作方差, 而显性×环境互作方差则在8月下旬较明显。

表 9.2 陆地棉不同发育时期单株成铃数的方差和条件方差分量

月/日	V_A	V_D	V_{AE}	V_{DE}	$V_{A(t t-1)}$	$V_{D(t t-1)}$	$V_{AE(t t-1)}$	$V_{DE(t t-1)}$
7/05	0.11 **	0.17 **	0.02 *	0.08 *	—	—	—	—
7/10	0.19 *	0.51 *	0.07	0.24 *	0.00	0.03 *	0.02 *	0.00
7/15	0.00	1.26 **	0.18 *	0.29 *	0.00	0.06 *	0.01 *	0.00
7/20	0.00	2.88 **	0.33 **	0.13 *	0.01	0.00	0.03 *	0.00
7/25	0.00	3.79 **	0.56 *	0.52	0.00	0.03	0.20 *	0.07 *
7/30	0.00	4.00 **	1.08 *	0.56	0.11 *	0.00	0.02	0.09 *
8/04	0.00	3.59 *	1.77 *	0.55	0.60 **	0.06	0.44 *	0.00
8/09	2.25 *	3.81 *	2.21 *	0.61 *	0.53 **	0.07 *	0.17 *	0.15 *
8/14	4.47 **	3.73 *	1.18 *	1.82 **	0.10 *	0.11	0.00	0.07
8/19	6.27 **	3.51 *	1.06 *	1.52 **	0.07 *	0.05 *	0.00	0.04 *
8/24	7.26 **	2.90 *	1.28 *	2.72 *	0.00	0.05	0.03 *	0.03 *
8/29	7.78 **	2.63 *	1.11 *	2.56 *	0.01 *	0.03 **	0.03 *	0.00
9/03	8.18 **	2.73 *	1.09	1.64 **	0.02 **	0.00	0.00	0.08 *

*、**分别表示达到0.05和0.01的显著水平。

条件方差分量的分析结果揭示了在某一段时期(5天内)基因的遗传效应表现。在初花期(7月10-15日), 未测出条件加性方差 $V_{A(t|t-1)}$, 但是条件显性方差 $V_{D(t|t-1)}$ 却较强(7月10日是 0.03 ± 0.01 ; 7月15日是 $0.06 \pm$

0.02)。这揭示了棉花早期加性方差分量下降而显性方差分量上升的原因。在7月20日-8月4日这一段时期内,已检测不到显著的条件显性方差分量,表明显性效应的基因表达已经关闭。但是显性方差分量 V_D 却持续增加至7月30日。从7月30日-8月19日存在显著的条件加性方差 $V_{A(t|t-1)}$,表明在这一时期加性效应的基因已经表达。条件加性方差的高峰值出现在8月4日($\hat{V}_{A(t|t-1)} = 0.60 \pm 0.12$),而此时加性方差 V_A 尚无法检测到。

条件方差分量的分析结果揭示了基因的动态表达过程,也有效地解释了棉花发育后期显性方差逐渐减小、加性方差持续增加的遗传现象。在能检测到普通加性方差 V_A 以前的10天,条件加性方差 $V_{A(t|t-1)}$ 已经能被显著地检测出,这具有重要的发育遗传学的意义。它揭示了发育数量遗传的一个规律:在控制数量性状的基因效应还未被传统的数量遗传分析方法检测出以前,基因已经开始表达。而这种特定阶段的基因表达,能够较早期地被条件方差分析检测到。

在8月4日单株成铃数的条件加性方差 $V_{A(t|t-1)}$ 和条件加性 \times 环境互作方差 $V_{AE(t|t-1)}$ 占总变异的70%左右,但是这时还未测到普通的加性及互作方差。条件加性 \times 环境互作方差 $V_{AE(t|t-1)}$ 在7月下旬至8月上旬也达显著(7月25日是 0.20 ± 0.08 ;8月4日是 0.44 ± 0.17 ;8月9日是 0.17 ± 0.07)。这表明,在1981和1985,两年不同的环境条件对基因加性效应的表达及对棉花单株成铃数有影响。

为了比较陆地棉不同发育时期的基因效应对最终($f = 9月3日$)单株成铃数的遗传影响,图9.1列出了陆地棉单株成铃数在不同发育时期的条件方差分量($V_{(f|t)}$)的柱状图以及9月3日方差分量($V_{(f)}$)的柱状图(最右边)。7/05至7/30的条件方差分量($V_{(9/03|7/05-30)}$)逐渐增大,然后急剧地减小。8/14以后,条件方差分量($V_{(9/03|8/14-29)}$)已降至接近零。这表明8/14以后的基因净效应对棉花单株最终成铃数的额外遗传贡献已极小。最大的条件方差分量值($V_{(9/03|7/25-30)}$)出现在7/25至7/30这段时期,揭示了用常

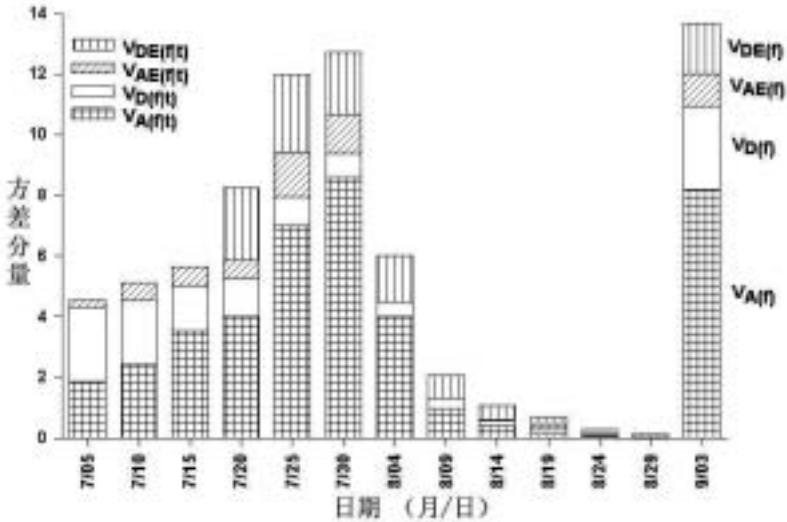


图 9.1 陆地棉最终($f = 9$ 月 3 日)单株成铃数的方差分量($V_{(f)}$)及

规分析方法所无法获得的遗传现象：在 7 月下旬陆地棉的基因表达对单株最终成铃数遗传变异的影响很小。综合表 9.2 和图 9.1 的分析结果可以发现，立秋前后(8 月上旬)是棉花发育的一个重要阶段，在这期间有许多新的基因表达(较大的条件方差分量 $V_{(t|t-1)}$)，并且这些基因表达的净遗传效应应对单株最终成铃数有很大的遗传影响(很小的条件方差分量 $V_{(f|t)}$)。

陆地棉不同发育时期单株成铃数的加性效应(A_i)和条件加性效应($A_{i(t|t-1)}$)的预测值列于(表 9.3)。在所分析的四个亲本中，早熟亲本(P3)在前期单株成铃较多，中后期则明显减少。其加性效应值 7 月 10 日前为显著的正值，8 月 9 日后为显著的负值。晚熟亲本(P2)的表现相反，单株成铃加性效应在 7 月 10 日前为显著的负值，8 月 9 日后为显著的正值。所有的亲本在 7 月 15 日至 8 月 4 日期间不存在加性遗传效应，但以后 P3 的加性效应值为极显著的负值，而 P2 和 P4 的加性效应则逐渐增加，并达到

显著和极显著水准。P1 在整个花铃期都未测到显著的加性效应。

表 9.3 陆地棉不同发育时期单株成铃数的加性效应和条件加性效应的预测值。

日期	A_1	A_2	A_3	A_4	$A_{1(t t-1)}$	$A_{2(t t-1)}$	$A_{3(t t-1)}$	$A_{4(t t-1)}$
7/05	-0.02	-0.27**	0.30**	0.00	—	—	—	—
7/10	-0.01	-0.35*	0.32*	0.04	0.00	0.00	0.00	0.00
7/15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7/20	0.00	0.00	0.00	0.00	-0.14	0.04	0.03	0.06
7/25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7/30	0.00	0.00	0.00	0.00	-0.14	0.12*	-0.29*	0.30*
8/04	0.00	0.00	0.00	0.00	0.01	0.44**	-0.79**	0.35*
8/09	-0.14	0.38*	-1.31**	1.07	0.17	0.66**	-0.56**	-0.27
8/14	-0.21	0.81*	-2.06**	1.46*	0.09*	0.27*	-0.14	-0.21
8/19	-0.30	1.18**	-2.47**	1.60*	-0.04	0.25	-0.03	-0.18*
8/24	-0.33	1.40**	-2.67**	1.60*	0.00	0.00	0.00	0.00
8/29	-0.29	1.54**	-2.78**	1.52*	0.04	0.06	-0.02	-0.08**
9/03	-0.21	1.64**	-2.87**	1.44*	0.13	0.07	-0.05	-0.14**

*、**分别表示达到 0.05 和 0.01 的显著水平。

早在 7 月 30 日至 8 月 4 日期间已测到三个亲本的条件加性效应值 $A_{i(t|t-1)}$, P3 为显著的负值, 而 P2 和 P4 为显著的正值, 尽管这时的普通加性效应值都是零。因此, 数量基因的加性效应已经在 P3、P2 和 P4 三个亲本中表达, 并能被条件遗传效应分析所检测, 但是却要在 10 天以后才能被传统的遗传效应分析方法测出。在棉花成铃中后期, P2 和 P4 的普通加性效应的表现较一致。但是条件加性效应分析结果表明, 这两个亲本的基因表达在中后期也是不同的。 $A_{2(t|t-1)}$ 和 $A_{4(t|t-1)}$ 在 7 月 30 日至 8 月 4 日期间的条件加性效应都是正值, 但以后 $A_{2(t|t-1)}$ 继续为正值, 而 $A_{4(t|t-1)}$ 却为负值。

以上分析表明，条件遗传效应和条件遗传方差分量的分析比传统的、非条件的遗传分析更能有效地揭示数量性状的发育遗传规律。

第十章 杂种优势的遗传分析

第一节 作物农艺性状的杂种优势分析

1. 加性-显性遗传模型的分析方法

数量遗传分析的一项重要任务是：通过估算性状的遗传方差分量及成对性状的遗传协方差分量，了解性状的遗传规律。除此以外，还需要通过预测遗传效应值及亲本和杂交组合的基因型值，评价亲本育种价值和杂交组合的杂种优势。

作物杂种优势利用是提高产量和品质的一项有效的重要措施。杂种优势的遗传分析以往主要采用 Griffing 的配合力遗传模型，估算亲本的一般配合力及杂交组合的特殊配合力。但 Griffing 的方法只适用于亲本和 F_1 世代。对 F_2 杂种优势的研究，目前主要通过一些间接指标(如亲本的亲缘关系、因子距离、中亲值等)与 F_2 表现型值的相关性预测 F_2 的杂种优势，但是这些间接预测的效果并不理想。由于作物农艺性状受环境机误影响较大， F_2 表现型值并不一定能完全反映组合的基因型值。因而只有直接对组合的基因型值进行分析，才能排除环境机误的干扰。

玉米、高粱和水稻等作物已成功地利用雄性不育系制种，在生产上大面积种植 F_1 。但目前有些作物由于尚不能经济地产生大量 F_1 种子，因而在生产上大面积地利用 F_1 的杂种优势还有困难。杂种优势的研究结果表明，陆地棉、冬小麦、大麦、油菜和芝麻等作物的一些重要农艺性状均存在明显的 F_2 杂种优势。如能直接利用杂种 F_2 代，可以解决 F_1 的供种问题，从而将能在生产上大面积利用杂种优势。

分析分离世代 F_2 的杂种优势表现，一般都需要种植 F_2 植株。由于 F_2 是分离世代，需要较大的群体，因而会增加遗传试验的费用和难度。对其它杂种后代的遗传研究则困难更大。如能由不分离世代(亲本和 F_1)直接预测 F_2 或其它杂种后代的基因型值，就可以提高试验效益，缩短试

验年限。

朱军(1993a)提出了杂种优势分析的新方法。该方法以加性-显性遗传模型分析亲本和 F_1 的双列杂交遗传资料,从而无偏预测基因加性效应值和显性效应值,并进一步预测 F_1 和其它杂种后代基因型值及其杂种优势。

在随机区组 k 内种植的亲本 i 和亲本 j 交配的后代,其表现型值的平均数 y_{ijkl} 可用以下线性模型表示

$$y_{ijkl} = \mu + G_{ij} + B_k + e_{ijk} \quad (10.)$$

1)

其中 μ 是群体平均数;

G_{ij} 是遗传效应值,由不同的基因效应分量所组成;

B_k 是区组效应, $B_k \sim (0, \sigma_B^2)$;

e_{ijk} 是剩余效应, $e_{ijk} \sim (0, \sigma_e^2)$ 。

如果供试材料的遗传表现符合以下四项遗传假设: 具有正常二倍体分离, 参与杂交的纯合亲本是遗传群体的一个随机样本, 不存在基因的上位性效应, 不存在基因型与环境的互动, 则遗传效应值 G_{ijk} 可进一步按以下不同的世代分解为加性效应和显性效应两项基因效应分量。

亲本 P_i (或 P_j)基因效应分量($i = j$)

$$G_{ii} = 2A_i + D_{ii}$$

杂种一代 F_{1ij} 基因效应分量

$$G_{ij} = A_i + A_j + D_{ij}$$

其中 A_i 或 A_j 是累加的加性效应, A_i 或 $A_j \sim (0, \sigma_A^2)$; D_{ii} 、 D_{jj} 或 D_{ij} 是累加的显性效应, D_{ii} 、 D_{jj} 或 $D_{ij} \sim (0, \sigma_D^2)$ 。

根据以上加性-显性遗传模型，可以由只包括亲本和 F_1 世代的双列杂交，预测加性效应值(A_i 和 A_j)和显性效应值(纯合显性 D_{ii} 、 D_{jj} 和杂合显性 D_{ij})。然后用这些基因效应值，再预测亲本、 F_1 或其它杂种世代的基因型值。

获得基因加性和显性效应的无偏预测值，然后可按下式无偏地预测各世代的遗传效应值：

$$\text{亲本 } G(P_i) = 2A_i + D_{ii}$$

$$\text{杂种一代 } G(F_{1ij}) = A_i + A_j + D_{ij}$$

$$\text{杂种二代 } G(F_{2ij}) = A_i + A_j + \frac{1}{4}D_{ii} + \frac{1}{4}D_{jj} + \frac{1}{2}D_{ij}$$

$$\text{杂种三代 } G(F_{3ij}) = A_i + A_j + \frac{3}{8}D_{ii} + \frac{3}{8}D_{jj} + \frac{1}{4}D_{ij}$$

$$\vdots \qquad \qquad \qquad \vdots$$

各世代的基因型值预测公式是($\hat{\mu} + \hat{G}$)，其中群体平均数估计值 $\hat{\mu}$ 可以由遗传试验的全部观察值的平均数估算。

假定亲本 i 为优亲，它与劣亲本 j 的遗传差异的绝对值为

$$\sigma_G = |2(A_i - A_j) + (D_{ii} - D_{jj})|$$

那么 F_1 的平均优势($H_M(F_1)$)和超亲优势($H_B(F_1)$)为：

$$\begin{aligned} H_M(F_1) &= G(F_{1ij}) - \frac{1}{2}[G(P_i) + G(P_j)] \\ &= [D_{ij} - \frac{1}{2}(D_{ii} + D_{jj})] \end{aligned}$$

$$\begin{aligned} H_B(F_1) &= G(F_{1ij}) - G(P_i) \\ &= [(A_j - A_i) + (D_{ij} - D_{ii})] \\ &= H_M(F_1) - \frac{1}{2}\sigma_G \end{aligned}$$

由 F_1 衍生的其它各世代杂种优势的预测值都是 $H_M(F_1)$ 和 σ 的函数。杂交种的自交后代平均优势可按以下公式预测。

杂种二代平均优势：

$$\begin{aligned} H_M(F_2) &= G(F_2) - \frac{1}{2} [G(P_i) + G(P_j)] \\ &= \frac{1}{2} [D_{ij} - \frac{1}{2} (D_{ii} + D_{jj})] \\ &= \frac{1}{2} H_M(F_1) \end{aligned}$$

杂种三代平均优势:

$$\begin{aligned} H_M(F_3) &= G(F_3) - \frac{1}{2} [G(P_i) + G(P_j)] \\ &= \frac{1}{4} [D_{ij} - \frac{1}{2} (D_{ii} + D_{jj})] \\ &= \frac{1}{4} H_M(F_1) \end{aligned}$$

杂交种的自交后代超亲优势的预测公式是，

杂种二代超亲优势:

$$\begin{aligned} H_B(F_2) &= G(F_2) - G(P_i) \\ &= \frac{1}{2} [2(A_j - A_i) + (D_{ij} + \frac{1}{2} D_{jj} - \frac{3}{2} D_{ii})] \\ &= \frac{1}{2} H_M(F_1) - \frac{1}{2} \sigma_G \end{aligned}$$

杂种三代超亲优势:

$$\begin{aligned} H_B(F_3) &= G(F_3) - G(P_i) \\ &= \frac{1}{4} [4(A_j - A_i) + (D_{ij} + \frac{3}{8} D_{jj} - \frac{3}{8} D_{ii})] \\ &= \frac{1}{4} H_M(F_1) - \frac{1}{2} \sigma_G \end{aligned}$$

依次类推，可以获得杂种 n 世代 F_n 的优势预测公式

杂种 F_n 世代的平均优势：

$$\begin{aligned} H_M(F_n) &= G(F_n) - \frac{1}{2} [G(P_i) + G(P_j)] \\ &= (\frac{1}{2})^{n-1} H_M(F_1) \end{aligned}$$

杂种 F_n 世代的超亲优势：

$$\begin{aligned} H_B(F_n) &= G(F_n) - G(P_i) \\ &= (\frac{1}{2})^{n-1} H_M(F_1) - \frac{1}{2} \sigma_G \end{aligned}$$

由杂种优势预测公式可见，随着世代数的增加自交后代的平均优势和超亲优势，每代递减 $\frac{1}{2} H_M(F_1)$ 。能否在杂种生产上利用 F_1 的自交后代，并

非完全由杂种优势的递减速率所决定。尽管杂种 F_1 自交后代的基因型值总是不如 F_1 的基因型值，但是如果组合的优亲是生产推广品种，并且自交后代仍然存在一定程度的正向超亲优势，就有可能在生产上继续利用该组合的自交后代。在实际应用时可以事先确定必须达到的正向超亲优势程度 β ，即要求 $H_B(F_n) \geq \beta$ 。然后根据所确定的 β 值，由以下关系式预测生产上可利用的杂种世代数 n ，

$$\frac{H_M(F_1)}{2\beta + \sigma_G} \geq 2^{(n-2)}$$

该关系式表明，杂种后代超亲优势不低于 β 所维持的世代数 n ，取决于 F_1 平均优势和双亲遗传差异。 $H_M(F_1)$ 越大，或 σ 越小，则正向超亲优势维持的年限越长。为了能获得使用年限较长的杂交组合，宜选用表现较小差异的优良品种作为双亲。如果 $H_M(F_1) > 0$ ，可以由下式直接预测世代数 n ，

$$n = \frac{\log[H_M(F_1)/(2\beta + \sigma_G)]}{\log(2)} + 2$$

当比值 $H_M(F_1)/(2\beta + \sigma)$ 分别达到 1、2、4 或 8 时，正向超亲优势维持的世代数分别为 2、3、4 或 5。如果某杂交组合的优亲是生产推广品种，那么根据其产量和品质性状的 F_1 平均优势及双亲的差异，便可以预测该组合可能在生产上连续使用的年限。 $H_M(F_1) \geq 2\beta + \sigma_G$ 是预测能否在生产上利用 F_2 的一项指标。在实际应用公式计算 n 时，育种者往往只有为数很少的组合，因而无法作遗传分析预测基因型值。此时可采用以下简易公式，近似地预测某组合超亲优势大于零的世代数，

$$n \approx \frac{\log[(F_1 - MP)/DP]}{\log(2)} + 2$$

其中 F_1 为组合的表型值，MP 为双亲表型平均值，DP 是双亲表型平均差异值(正值)。

以上所介绍的是杂种优势绝对值的分析方法。由于受到性状测量单

位的影响，不同性状之间的杂种优势绝对值没有可比性。在实际应用时常采用杂种优势的相对值，即平均优势

$$(F_1 - MP) / MP \approx \frac{D_{ij} - \frac{1}{2}(D_{ii} + D_{jj})}{\mu + A_i + A_j + \frac{1}{2}(D_{ii} + D_{jj})}$$

或超亲优势

$$(F_1 - BP) / BP \approx \frac{(A_j - A_i) + (D_{ij} - D_{ii})}{\mu + 2A_i + D_{ii}}$$

其中 MP 为双亲表型平均值，BP 为优亲表型平均值。由于分母中含有加性效应，因此这些杂种优势的统计量并不能真正度量杂种优势的基因效应 $[D_{ij} - (D_{ii} + D_{jj})/2]$ 。朱军等(1993)建议用群体平均优势 (H_{PM}) 来度量杂种优势的基因效应，

$$\begin{aligned} H_{PM}(F_1) &= \frac{1}{\mu} H_M \\ &= \frac{1}{\mu} [D_{ij} - \frac{1}{2}(D_{ii} + D_{jj})] \end{aligned}$$

它是杂种优势基因效应 $\Delta_G = [D_{ij} - \frac{1}{2}(D_{ii} + D_{jj})]$ 的简单函数。在实际应用时可用群体平均数 μ 的无偏估计值 $\hat{\mu}$ (双列杂交所有亲本和组合的算术平均数) 来估算群体平均优势和群体超亲优势。

如果用群体超亲优势 (\hat{H}_{PB})，则可以更好地度量超亲优势，

$$\begin{aligned} H_{PB}(F_1) &= \frac{1}{\mu} H_B \\ &= \frac{1}{\mu} (H_M - \frac{1}{2}\sigma) \\ &= \frac{1}{\mu} H_M - \frac{1}{2\mu}\sigma \\ &= H_{PM} - \frac{1}{2}\delta_G \end{aligned}$$

其中 $\delta_G = \sigma_G / \mu$ 是亲本的相对遗传差异。

群体超亲优势是平均优势基因效应和双亲遗传差异 σ_G 的函数。群体平均优势大于群体超亲优势，其差值反映了亲本之间的遗传差异。不同世代 ($n = 1, 2, 3, \dots$) 群体杂种优势的预测可以按以下公式求得：

杂种 F_n 世代的群体平均优势：

$$\begin{aligned} H_{PM}(F_n) &= \frac{1}{\mu} H_M(F_n) \\ &= \left(\frac{1}{2}\right)^{n-1} H_{PM}(F_1) \end{aligned}$$

杂种 F_n 世代的群体超亲优势：

$$\begin{aligned} H_{PB}(F_n) &= \frac{1}{\mu} H_B(F_n) \\ &= \left(\frac{1}{2}\right)^{n-1} H_{PM}(F_1) - \frac{1}{2} \delta_G \end{aligned}$$

可以根据事先确定必须达到的正向群体超亲优势程度 α ，即要求 $H_{PB}(F_n) \geq \alpha$ ， α 可以取 0.05 (5% 正向群体超亲优势) 或 0.10 (10% 正向群体超亲优势)。然后根据所确定的 α 值，由以下关系式预测生产上可利用的杂种世代数 n ，

$$\frac{H_{PM}(F_1)}{2\alpha + \delta_G} \geq 2^{(n-2)}$$

如果 $H_{PM}(F_1) > 0$ ，可以由下式直接预测世代数 n ，

$$n = \frac{\log [H_{PM}(F_1) / (2\alpha + \delta_G)]}{\log(2)} + 2$$

当比值 $H_{PM}(F_1) / (2\alpha + \delta_G)$ 分别达到 1、2、4 或 8 时，正向超亲优势维持的世代数分别为 2、3、4 或 5。

朱军 (1993a) 按加性-显性遗传模型，分析了 6 个陆地棉品种 (徐州 142、中棉所 7 号、协作 2 号、岱字棉 15 号、无腺体的 GL-5、无蜜腺的 HG-H-12) 完全双列杂交的单株铃数和单铃重两个产量性状。该试验包括 6 个亲本和 30 个正反交组合，完全随机区组设计，重复四次。采用 Jackknife 数值抽样技术对试验区组进行抽样，计算方差

分量的估计值、基因型值和群体超亲优势的预测值以及它们的标准误。

棉花产量组成因素单株铃数和单铃重的方差分量分析结果表明，单株铃数主要受基因加性效应影响，加性遗传方差 V_A 约为显性遗传方差 V_D 的三倍；单铃重则主要受基因显性效应影响，加性遗传方差 V_A 不显著，显性遗传方差 V_D 达到 1% 极显著水准。

单株铃数和单铃重的 15 个组合的 F_2 基因型值和超亲优势预测以及预期的世代数 ($n = 0.05$ 时) 列于表 10.1，括号内是各项预测值的标准误 SE 。由预测值 $\pm t_{0.05} SE$ 可得到关于真实参数的 95% 置信区间，从而可以比较不同组合的参数之间是否存在真实差异。各组合预期的世代数 n 可以大致预测连续种植该组合的杂种 n 代后，群体超亲优势仍将不低于 5%。

在单株铃数上， F_2 群体超亲优势预测值较高的组合是 (4×6) 和 (1×6) ，其 $H_{PB}(F_2)$ (及 95% 置信区间) 分别为 0.116(0.046~0.186) 和 0.056(-0.027~0.139)。虽然这两个组合的超亲优势预测值都超过 5%，但置信区间的分析表明组合 (4×6) 的 F_2 显著超过优亲 4.6% 以上；而组合 (1×6) 的 F_2 并未显著超过优亲。各组合预期的世代数与群体超亲优势预测值是吻合的。当 $n = 2$ 时， F_2 群体超亲优势预测值约为 5%。组合 (4×6) 的预期世代数约为 3，可以推断该组合 F_3 的单株铃数仍可能超过优亲 5%。这两个组合也有较高的基因型值，分别为 16.34(14.08~18.60) 和 17.27(15.04~19.50)。置信区间的分析表明，这两个组合的 F_2 基因型值差异不显著。单铃重 F_2 群体超亲优势预测值较高的组合是 (2×4) 和 (1×2) 。这两个组合的预期世代数都不小于 2。单铃重共有 4 个组合的预期世代数达到了 2，而单株铃数却只有两个组合。这是由于单铃重主要受显性效应控制，因而表现较强的超亲优势；而单株铃数存在较强的加性效应，超亲优势相对较弱。15 个组合单株铃数的 F_1 平均群体超亲优势为 1.6%，而单铃重的则高达 6.7%。

表 10.1 棉花杂交组合单株铃数和单铃重基因型值和超亲优势的预测

以及预期的世代数

组合	单株铃数				单铃重			
	F_{2ij}	(SE)	HPB	(SE) n	F_{2ij}	(SE)	HPB	(SE) n
(1×2)) 4	15.5	(0.56)	-0.06	(0.053) 0.9	5.3	(0.06)	0.052	(0.028) 2.0
(1×3)) 3	16.4	(0.57)	0.000	(0.037) 1.4	4.8	(0.11)	0.018	(0.017) 0.6
(1×4)) 5	17.1	(0.22)	0.049	(0.017) 1.9	5.0	(0.07)	0.049	(0.017) 2.0
(1×5)) 0	12.8	(0.17)	-0.25	(0.030) 0.0	5.6	(0.07)	-0.02	(0.032) 0.9
(1×6)) 7	17.2	(0.70)	0.056	(0.026) 2.1	5.0	(0.06)	0.041	(0.012) 1.7
(2×3)) 7	13.2	(0.29)	-0.03	(0.050) 0.5	5.1	(0.07)	0.028	(0.024) 1.4
(2×4)) 9	13.7	(0.46)	-0.06	(0.035) 0.0	5.3	(0.06)	0.073	(0.026) 2.3
(2×5))	9.96	(0.51)	-0.12	(0.047) 0.2	5.6	(0.14)	-0.00	(0.031) 1.0
(2×6)) 0	13.8	(0.27)	-0.08	(0.059) < 0.6	5.2	(0.21)	0.049	(0.025) 2.0
(3×4)) 8	13.7	(0.87)	-0.07	(0.059) 0.0	4.9	(0.09)	0.029	(0.023) 1.5
(3×5)) 0	11.2	(0.55)	-0.15	(0.051) 0.4	5.5	(0.06)	-0.04	(0.031) 0.5
(3×6))	15.0	(0.78)	0.020	(0.032) 1.4	4.8	(0.10)	0.024	(0.012) 0.0

)	4))	5	9))	9		
(4×5	11.9	(0.26	-0.19	(0.029	0.	5.3	(0.07	-0.06	(0.032	0.
)	2)	2)	3	8)	5)	1
(4×6	16.3	(0.71	0.116	(0.022	3.	4.9	(0.12	0.033	(0.016	1.
)	4))	2	1))	5		
(5×6	12.0	(0.57	-0.17	(0.047	0.	5.3	(0.22	-0.07	(0.029	< 0
)	5)	6)	4	5)	1)	

由表 10.1 可见, 杂种优势最高的组合并不一定是基因型值最高的组合。单株铃数 F_2 群体超亲优势预测值最高的组合是(4 × 6), 其 F_2 基因型值却位居第四。单铃重 F_2 群体超亲优势预测值最高的组合(2 × 4), 其 F_2 基因型值也只位居第四。因此评价组合的优劣, 应同时考虑组合的基因型值和杂种优势。单株铃数和单铃重两个性状 F_2 优势的表现不完全一致。这表明, 有必要根据作物产量和品质各项性状的综合分析结果, 全面评价杂交组合的优劣。生产上利用的优势组合必须在主要产量组成因素和品质性状方面均表现较好。

2. 加性-显性及环境互作的遗传模型分析方法

如果存在基因型与环境的互作, 杂种优势将包括两个分量: 其一是不因环境影响而变的基因型优势, 其二是因环境不同而变异的环境互作优势。

在环境 h 内, 由亲本 i 和亲本 j 杂交产生的后代在区组 k 中的遗传模型, 可以用以下线性公式表示,

$$y_{hijk} = \mu + E_h + G_{ij} + GE_{hij} + B_{hk} + e_{hijk} \quad (10.)$$

2)

其中 y_{hijk} 是母本 i 与父本 j 的组合在环境 h 内第 k 个区组中的平均表现型值;

μ 是群体平均数, E_h 是环境效应, 这两个参数均为固定效应; B_{hk} 是环境内的区组效应, $B_{hk} \sim (0, \sigma_B^2)$; e_{hijk} 是剩余效应, $e_{hijk} \sim (0, \sigma_e^2)$ 。

G_{ij} 的定义如(10.1), GE_{hij} 的分解如下:

亲本 P_i (或 P_j) 的基因型 \times 环境互作效应分量 ($i = j$)

$$GE_{hij} = 2AE_{hi} + DE_{hii}$$

杂种一代 F_{1ij} 的基因型 \times 环境互作效应分量

$$GE_{hij} = AE_{hi} + AE_{hj} + DE_{hij}$$

其中 AE_{hi} 或 AE_{hj} 是加性 \times 环境互作效应, AE_{hi} 或 $AE_{hj} \sim (0, \sigma_{AE}^2)$; DE_{hii} 、 DE_{hjj} 或 DE_{hij} 是累加的显性效应, DE_{hii} 、 DE_{hjj} 或 $DE_{hij} \sim (0, \sigma_{DE}^2)$ 。

根据以上加性-显性及环境互作的遗传模型, 可以由只包括亲本和 F_1 世代的双列杂交, 预测加性效应值(A_i 和 A_j)、显性效应值(纯合显性 D_{ii} 、 D_{jj} 和杂合显性 D_{ij})以及加性 \times 环境互作效应值(AE_{hi} 和 AE_{hj})、显性 \times 环境互作效应值(纯合显性 \times 环境互作 DE_{hii} 、 DE_{hjj} 和杂合显性 \times 环境互作 DE_{hij})。然后用这些基因效应值和基因型 \times 环境互作效应值, 再预测亲本、 F_1 或其它杂种世代的遗传表现。

获得基因加性 \times 环境互作和显性 \times 环境互作效应的无偏预测值, 然后可按式无偏地预测各世代的基因型 \times 环境互作效应值:

$$GE(P_i) = 2AE_{hi} + DE_{hii}$$

$$GE(F_{1ij}) = AE_{hi} + AE_{hj} + DE_{hij}$$

$$GE(F_{2ij}) = AE_{hi} + AE_{hj} + \frac{1}{4}DE_{hii} + \frac{1}{4}DE_{hjj} + \frac{1}{2}DE_{hij}$$

$$GE(F_{3ij}) = AE_{hi} + AE_{hj} + \frac{3}{8}DE_{hii} + \frac{3}{8}DE_{hjj} + \frac{1}{4}DE_{hij}$$

⋮

假定亲本 i 为优亲, 它与劣亲本 j 在环境 h 中的遗传差异绝对值为

$$\varpi_{GE} = |2(AE_{hi} - AE_{hj}) + (DE_{hii} - DE_{hjj})|$$

那么 F_1 的平均优势互作离差 ($H_{ME}(F_1)$) 和超亲优势互作离差 ($H_{BE}(F_1)$) 为：

$$\begin{aligned} H_{ME}(F_1) &= GE(F_{1ij}) - \frac{1}{2}[GE(P_i) + GE(P_j)] \\ &= DE_{hij} - \frac{1}{2}(DE_{hii} + DE_{hjj}) \end{aligned}$$

$$\begin{aligned} H_{BE}(F_1) &= GE(F_{1ij}) - GE(P_i) \\ &= (AE_{hj} - AE_{hi}) + (DE_{hij} - DE_{hii}) \\ &= H_{ME}(F_1) - \frac{1}{2}\varpi_{GE} \end{aligned}$$

由 F_1 衍生的其它各世代杂种优势互作离差的预测值都是 $H_{ME}(F_1)$ 和 ϖ_{GE} 的函数。杂交种的自交后代平均优势互作离差可按以下公式预测。
杂种二代平均优势互作离差：

$$\begin{aligned} H_{ME}(F_2) &= GE(F_2) - \frac{1}{2}[GE(P_i) + GE(P_j)] \\ &= \frac{1}{2}[DE_{hij} - \frac{1}{2}(DE_{hii} + DE_{hjj})] \\ &= \frac{1}{2}H_{ME}(F_1) \end{aligned}$$

杂种三代平均优势互作离差：

$$\begin{aligned} H_{ME}(F_3) &= GE(F_3) - \frac{1}{2}[GE(P_i) + GE(P_j)] \\ &= \frac{1}{4}[DE_{hij} - \frac{1}{2}(DE_{hii} + DE_{hjj})] \\ &= \frac{1}{4}H_{ME}(F_1) \end{aligned}$$

杂交种的自交后代超亲优势互作离差的预测公式是，

杂种二代超亲优势互作离差：

$$\begin{aligned} H_{BE}(F_2) &= GE(F_2) - GE(P_i) \\ &= \frac{1}{2}[2(AE_{hj} - AE_{hi}) + (DE_{hij} + \frac{1}{2}DE_{hjj} - \frac{3}{2}DE_{hii})] \\ &= \frac{1}{2}H_{ME}(F_1) - \frac{1}{2}\varpi_{GE} \end{aligned}$$

杂种三代超亲优势互作离差：

$$\begin{aligned} H_{BE}(F_3) &= GE(F_3) - GE(P_i) \\ &= \frac{1}{4}[4(AE_{hj} - AE_{hi}) + (DE_{hij} + \frac{3}{8}DE_{hij} - \frac{3}{8}DE_{hii})] \\ &= \frac{1}{4}H_{ME}(F_1) - \frac{1}{2}\varpi_{GE} \end{aligned}$$

依次类推，可以获得杂种 n 世代 F_n 的优势互作离差预测公式

杂种 F_n 世代的平均优势互作离差：

$$\begin{aligned} H_{ME}(F_n) &= GE(F_n) - \frac{1}{2}[GE(P_i) + GE(P_j)] \\ &= (\frac{1}{2})^{n-1} H_{ME}(F_1) \end{aligned}$$

杂种 F_n 世代的超亲优势互作离差：

$$\begin{aligned} H_{BE}(F_n) &= GE(F_n) - GE(P_i) \\ &= (\frac{1}{2})^{n-1} H_{ME}(F_1) - \frac{1}{2}\varpi_{GE} \end{aligned}$$

为了便于比较不同性状的杂种优势互作离差，可以采用群体平均优势互作离差 (H_{PME}) 来度量杂种优势的基因 \times 环境互作效应，

$$\begin{aligned} H_{PME}(F_1) &= \frac{1}{\mu} H_{ME} \\ &= \frac{1}{\mu} [DE_{hij} - \frac{1}{2}(DE_{hii} + DE_{hjj})] \end{aligned}$$

它是杂种优势基因 \times 环境互作效应 $\Delta_{GE} = [DE_{hij} - \frac{1}{2}(DE_{hii} + DE_{hjj})]$ 的简单函数。

如果用群体超亲优势 (\hat{H}_{PB})，则可以更好地度量超亲优势互作离差，

$$\begin{aligned} H_{PBE}(F_1) &= \frac{1}{\mu} H_{BE} \\ &= \frac{1}{\mu} (H_{ME} - \frac{1}{2}\varpi_{GE}) \\ &= \frac{1}{\mu} H_{ME} - \frac{1}{2\mu}\varpi_{GE} \\ &= H_{PME} - \frac{1}{2}\delta_{GE} \end{aligned}$$

其中 $\delta_{GE} = \varpi_{GE} / \mu$ 是亲本在环境 h 中的相对遗传差异绝对值。

不同世代($n = 1, 2, 3, \dots$)群体杂种优势互作离差的预测可以按下式求得：

杂种 F_n 世代的群体平均优势互作离差：

$$\begin{aligned} H_{PME}(F_n) &= \frac{1}{\mu} H_{ME}(F_n) \\ &= \left(\frac{1}{2}\right)^{n-1} H_{PME}(F_1) \end{aligned}$$

杂种 F_n 世代的群体超亲优势互作离差：

$$\begin{aligned} H_{PBE}(F_n) &= \frac{1}{\mu} H_{BE}(F_n) \\ &= \left(\frac{1}{2}\right)^{n-1} H_{PME}(F_1) - \frac{1}{2} \delta_{GE} \end{aligned}$$

在环境互作的遗传分析中,环境一般是指年份或地点的生态大环境。这些大环境是人类无法控制的,随年份和地点而有较大的波动。因此,一般没有必要预测在特定环境下杂交组合可使用的世代数。

杂种优势的互作离差度量了某一环境条件下(年份或地点)优势偏离平均数的值。某一环境条件下的杂种优势表现由二部分分量组成：平均优势 = $H_M + H_{ME}$ ；超亲优势 = $H_B + H_{BE}$ 。第一部分是由基因主效应决定的杂种优势,度量了杂种优势在不同环境下的稳定性；第二部分是由基因型 \times 环境互作效应决定的杂种优势,反映了杂种优势在特定环境下的波动性。

现以 1992 至 1993 年 10 个陆地棉亲本的双列杂交的遗传资料分析实例,说明互作杂种优势的分析及其在育种上的应用。10 个品种或品系分别为：无蜜腺选系：A226、A160；无腺体选系：A17；我国自育品种：鲁棉 6 号、中棉所 12 号、中棉所 13 号、徐州 184、泗棉 2 号；引种美国材料：4305、4318。试验在浙江农业大学实验农场进行。以亲本 ~ 作母本、亲本 ~ 作为父本,共配制了 20 个组合。

4 个 F_2 高产组合产量性状的遗传表现列于表 10.2。分析结果表明, F_2 产量性状的表现型值与 F_2 基因型值间的表现并不完全一致。其中, F_2

皮棉产量表现型值与基因型值间存在较大的差异，其它各产量性状的表现型值与基因型值则较为接近。其原因是表现型值与环境 and 机误有关，皮棉产量受环境和机误影响较大。3个杂交组合(4×6)、(4×8)和(5×8)F₂皮棉产量的群体超亲优势较大，其预计的世代数大于2，即这3个杂交组合F₂皮棉产量在多种环境中群体超亲优势平均表现大于5%。组合(4×6)F₂皮棉产量群体超亲优势和群体平均优势较为接近，约为14%，其主要原因是该组合双亲的相对遗传差异都较小。组合(4×6)和(5×7)F₂单株铃数基因型值较大，分别为14.03和14.63个，但这两个组合F₂单株铃数杂种优势较小。组合(4×8)和(5×8)F₂单株铃数基因型值也较大，分别为14.00和13.21个，且这两个组合F₂群体平均优势和超亲优势都较大，平均优势分别为9%和6%，超亲优势较为接近，约为5%左右，预计的世代数约为2。(4×6)、(5×7)和(4×8)组合F₂单铃重的基因型值接近，分别为4.69g、4.75g和4.70g。组合(5×8)F₂单铃重则表现出一定的正向群体平均优势和群体超亲优势。这4个组合F₂衣分的基因型值均较大，衣分超过40%以上。其中(4×6)和(4×8)组合F₂衣分均表现显著的群体平均优势，F₂超亲优势显著小于0。组合(5×7)和(5×8)F₂衣分平均优势不显著，群体超亲优势显著小于0。

表 10.2 四个单交 F₂ 高产组合的产量性状表现型值、基因型值、杂种优势和世代数

性状	参数	组合(4×6)	组合(4×8)	组合(5×7)	组合(5×8)
产量	P(F ₂) (kg)	52.15	50.78	59.12	53.51
	μ+G (kg)	61.89	64.16	66.77	62.26
	H _{PM}	0.15**	0.14**	0.13**	0.14**
	H _{PB}	0.10*	0.14**	0.01	0.11*
	n	2.53	3.34	1.60	2.78
铃数	P(F ₂) (个/株)	13.18	12.94	12.61	13.10
	μ+G (个/株)	14.03	14.00	14.63	13.21
	H _{PM}	0.02	0.09	0.06	0.06
	H _{PB}	-0.01	0.04	-0.04	0.05
	n	1.09	1.91	0.55	2.01
单铃重	P(F ₂) (g)	4.44	4.59	4.65	4.40
	μ+G (g)	4.69	4.70	4.75	4.86
	H _{PM}	0.00	-0.00	0.01	0.04*
	H _{PB}	-0.02	-0.01	-0.02	0.03
	n	0.00	0.00	0.00	1.85
衣分	P(F ₂)(%)	41.11	41.00	43.11	39.72
	μ+G (%)	40.01	41.11	42.37	40.55
	H _{PM}	0.02 ⁺	0.01 ⁺	0.01	-0.01
	H _{PB}	-0.02**	-0.02**	-0.01**	-0.03**
	n	0.00	0.00	0.00	0.00

+、*、** 分别表示 0.10、0.05 和 0.01 的显著水平。

表 10.3 四个陆地棉 F₂ 高产组合产量性状在两年中的遗传表现

性状	参数	组合(4×6)		组合(4×8)		组合(5×7)		组合(5×8)	
		1992	1993	1992	1993	1992	1993	1992	1993
产量	P(F ₂) (kg)	58.15	46.15	54.12	47.44	62.72	55.52	62.24	44.78
	μ+G+GE (kg)	65.04	60.48	67.16	64.08	67.71	70.21	63.98	62.25
	H _{PM} +H _{PME}	0.32	0.14	0.29	0.14	0.24	0.17	0.30	0.12
	H _{PB} +H _{PBE}	0.27	0.10	0.27	0.14	0.10	0.09	0.22	0.10
铃数	P(F ₂) (个/株)	14.58	11.77	14.70	11.18	14.53	10.69	16.10	10.09
	μ+G+GE (个/株)	14.59	13.85	13.55	14.86	15.52	14.92	12.99	13.17
	H _{PM} +H _{PM}	0.01	0.09	0.11	0.16	0.10	0.05	0.10	0.08
	H _{PB} +H _{PBE}	0.00	0.02	0.05	0.10	-0.02	-0.04	0.06	0.05
单铃重	P(F ₂) (g)	4.47	4.42	4.90	4.27	4.91	4.39	4.48	4.32
	μ+G+GE (g)	4.57	4.64	4.72	4.57	4.53	4.93	4.65	5.10
	H _{PM} +H _{PM}	-0.01	0.01	0.00	-0.01	0.05	-0.01	0.05	0.08
	H _{PB} +H _{PBE}	-0.04	0.00	-0.03	-0.03	0.03	-0.06	0.02	0.07
衣分	P(F ₂) (%)	40.77	41.45	39.72	42.27	41.53	44.69	39.90	39.53
	μ+G+GE (%)	40.72	40.95	40.86	41.23	42.37	42.60	40.34	40.10
	H _{PM} +H _{PM}	0.01	0.04	0.02	0.01	0.01	-0.01	0.01	-0.03
	H _{PB} +H _{PBE}	-0.01	-0.02	-0.00	-0.01	0.00	-0.01	-0.01	-0.05

由于各性状都存在不同程度的基因型 \times 环境的互动，因此各组合 F_2 基因型值和杂种优势在年份间的表现不一致。为进一步分析各高产组合 F_2 杂种优势的稳定性表现，表 10.3 列出了四个高产组合 F_2 产量性状在两年中的表现型值、基因型值和杂种优势。由于表现型值与环境 and 机误效应有关，而基因型值则不包含这些效应，因此， F_2 产量性状在某环境下的基因型值与表现型值间表现为不同程度的差异(见表 10.3)。其中皮棉产量、单株铃数和前期收花率差异较大，衣分和单铃重差异较小。 F_2 产量性状基因型值在两年中的表现不完全一致。四个单交组合 F_2 皮棉产量的基因型值在 1992 年的环境条件下一般略大些。其它性状的基因型值在两年中的表现较不一致。例如，1992 年组合(4 \times 6)和(5 \times 7) F_2 单株铃数基因型值大于 1993 年，而组合(4 \times 8)和(5 \times 8) F_2 单株铃数基因型值略小于 1993 年。其它性状表现与单株铃数类似。由表 10.3 可知， F_2 产量性状的杂种优势在两年中的表现也不完全一致。例如，四个组合在 1992 年中 F_2 皮棉产量的杂种优势大于 1993 年。组合(4 \times 6)和(4 \times 8) F_2 单株铃数在 1992 年的杂种优势小于 1993 年，而组合(5 \times 7)和(5 \times 8) F_2 单株铃数的优势在两年间的表现则相反。由此表明，不同组合 F_2 在不同环境中杂种优势的稳定性表现有所不同。因此，对产量性状遗传规律的研究需要强调在多个环境中进行。

与表 10.2 相比，组合 F_2 在两年中的基因型值或杂种优势的平均值与基因效应的基因型值和杂种优势并不完全一致。例如，组合(5 \times 7)在两年中的基因型值和杂种优势的平均值都略大于其基因效应的基因型值和杂种优势。其它 F_2 产量性状在不同年份内的平均遗传表现与基因效应的表现也不完全一致。由此可知，对这些性状杂种优势的遗传规律的研究还需要在更多的环境中进行。

第二节 种子品质性状的杂种优势分析

如果控制植物种子数量性状遗传表现的遗传体系有三套：种子的直接基因效应、细胞质基因效应、母体植株基因效应。种子品质性状杂种优势的表现也将同时受制于这三套遗传体系的基因效应。

当不存在环境互作效应时，种子的数量性状遗传表现受总遗传效应控制，-总遗传效应值可以分解为三个分量 ($G = G_O + C + G_M$)。双子叶植物的二倍体种子 ($G = G_O(2n) + C + G_M$) 与单子叶植物的三倍体胚乳 ($G = G_O(3n) + C + G_M$) 在计算遗传效应值时，其直接遗传效应值的计算方法有很大的不同，但细胞质和母体效应值的计算公式相同。

采用混合线性模型的分析方法分析种子性状遗传资料，可以获得各项遗传效应的无偏预测值，然后可进一步预测各种杂交世代的遗传表现以及杂种优势。

按以下公式可以预测亲本及母本 i 与父本 j 杂交不同世代的遗传效应值：

G	$G_O(2n)$	$G_O(3n)$	C	G_M
$G(P_i)$	$2A_i + D_{ii}$	$3A_i + 3D_{ii}$	C_i	$2Am_i + Dm_{ii}$
$G(F_1)$	$A_i + A_j + D_{ij}$	$2A_i + A_j + D_{ii} + 2D_{ij}$	C_i	$2Am_i + Dm_{ii}$
$G(F_2)$	$A_i + A_j + \frac{1}{4}D_{ii}$ $+ \frac{1}{4}D_{jj} + \frac{1}{2}D_{ij}$	$\frac{3}{2}A_i + \frac{3}{2}A_j + D_{ii}$ $+ D_{jj} + D_{ij}$	C_i	$Am_i + Am_j + Dm_{ij}$

不存在环境互作时，数量性状的杂种优势是杂种优势基因主效应 (Δ_G) 和亲本遗传差异值 (ω_G) 的函数。种子性状的杂种优势基因效应有二种：直接基因效应的 $\Delta_O = D_{ij} - \frac{1}{2}(D_{ii} + D_{jj})$ 和母体基因效应的 $\Delta_M = Dm_{ij} - \frac{1}{2}(Dm_{ii} + Dm_{jj})$ 。亲本遗传差异值 ($\omega_G = G(P_i) - G(P_j)$) 包括

三个分量,

$$\begin{aligned}\varpi_G &= \varpi_O + \varpi_C + \varpi_M \\ &= [G_O(P_i) - G_O(P_j)] + [G_C(P_i) - G_C(P_j)] + [G_M(P_i) - G_M(P_j)]\end{aligned}$$

其遗传组成为

$$\begin{aligned}\varpi_O(2n) &= 2(A_i - A_j) + (D_{ii} - D_{jj}) \\ \varpi_O(3n) &= 3(A_i - A_j) + 3(D_{ii} - D_{jj}) \\ \varpi_C &= C_i - C_j \\ \varpi_M &= 2(Am_i - Am_j) + (Dm_{ii} - Dm_{jj})\end{aligned}$$

种子数量性状的杂优分析, 常可估算两种优势: 平均优势, 即杂交种子遗传表现超过双亲平均数的优势,

$$\begin{aligned}H_M(F_n) &= G(F_n) - \frac{1}{2}[G(P_i) + G(P_j)] \\ &= H_{MO} + \frac{1}{2}\varpi_C + H_{MM}\end{aligned}$$

由于母体植株的细胞质基因和核基因对种子性状的影响较大, 因此在利用杂种优势时常采用遗传表现优良的材料作为母本。杂交种子遗传表现超过母体亲本的优势, 可以定义为杂交种子的超亲优势,

$$\begin{aligned}H_F(F_n) &= G(F_n) - G(P_i) \\ &= H_{FO} + H_{FM}\end{aligned}$$

其中 H_{MO} 是归因于直接效应的平均优势、 H_{MM} 是归因于母体效应的平均优势; H_{FO} 是归因于直接效应的超亲优势、 H_{FM} 是归因于母体效应的超亲优势。

这两种优势都可以用杂种优势基因效应 (Δ_G) 和亲本遗传差异值 (ϖ_G) 分量的函数表示。平均优势的遗传分量是,

H_M	$H_{MO}(2n)$	$H_{MO}(3n)$	H_{MM}
$H_M(F_1)$	Δ_O	$2\Delta_O + \frac{1}{6}\varpi_O$	$\frac{1}{2}\varpi_M$
$H_M(F_2)$	$\frac{1}{2}\Delta_O$	Δ_O	Δ_M
⋮	⋮	⋮	⋮

$H_M(F_n)$	$(\frac{1}{2})^{n-1} \Delta_O$	$(\frac{1}{2})^{n-2} \Delta_O$	$(\frac{1}{2})^{n-2} \Delta_M$
------------	--------------------------------	--------------------------------	--------------------------------

超亲优势的遗传分量是，

H_F	$H_{FO}(2n)$	$H_{FO}(3n)$	H_{FM}
$H_F(F_1)$	$\Delta_O - \frac{1}{2} \varpi_O$	$2\Delta_O - \frac{1}{3} \varpi_O$	—
$H_F(F_2)$	$\frac{1}{2} \Delta_O - \frac{1}{2} \varpi_O$	$\Delta_O - \frac{1}{2} \varpi_O$	$\Delta_M - \frac{1}{2} \varpi_M$
⋮	⋮	⋮	⋮
$H_F(F_n)$	$(\frac{1}{2})^{n-1} \Delta_O - \frac{1}{2} \varpi_O$	$(\frac{1}{2})^{n-2} \Delta_O - \frac{1}{2} \varpi_O$	$(\frac{1}{2})^{n-2} \Delta_M - \frac{1}{2} \varpi_M$

当存在环境互作效应时，种子的数量性状遗传表现受遗传效应及环境互作效应的控制。其中遗传效应是不因环境变化而改变的遗传主效应，环境互作效应是在特定环境条件下基因表达的遗传效应。基因型 × 环境互作效应值也可以分解为三个分量 ($GE = GoE + CE + GmE$)。可以按以下公式预测亲本及母本 i 与父本 j 杂交的不同世代在环境 h 中的基因型 × 环境互作效应值：

GE	GoE(2n)	GoE(3n)	CE	GmE
$GE(P_1)$	$2AE_{hi} + DE_{hii}$	$3AE_{hi} + 3DE_{hii}$	CE_{hi}	$2AmE_{hi} + DmE_{hii}$
$GE(F_1)$	$AE_{hi} + AE_{hj} + DE_{hij}$	$3AE_{hi} + AE_{hj} + DE_{hii} + 2DE_{hij}$	CE_{hi}	$2AmE_{hi} + DmE_{hii}$
$GE(F_2)$	$AE_{hi} + AE_{hj} + \frac{1}{4}DE_{hii} + \frac{1}{4}DE_{hij} + \frac{1}{2}DE_{hij}$	$\frac{3}{2}AE_{hi} + \frac{3}{2}AE_{hj} + DE_{hii} + DE_{hij} + DE_{hij}$	CE_{hi}	$AmE_{hi} + AmE_{hj} + DmE_{hij}$

种子性状的杂种优势基因 × 环境互作效应有二种：直接基因 × 环境互作效应的 $\Delta_{OE} = DE_{hij} - \frac{1}{2}(DE_{hii} + DE_{hjj})$ 和母体基因 × 环境互作效应的

$\Delta_{ME} = DmE_{hij} - \frac{1}{2}(DmE_{hii} + DmE_{hjj})$ 。这种杂合显性效应用于纯合显性效应的优势度是显性效应在环境 h 中的特定表现，其因环境条件的变异而表现不同。

当存在环境互作时，在某一环境中亲本之间的差异包括两部分差异：一部分是不因环境改变而变化的遗传主效应的差异，即 $\varpi_G = G(P_i) - G(P_j)$ ；另一部分是只在环境 h 中才表现的遗传差异，即互作差异值 $\varpi_{GE} = GE(P_i) - GE(P_j)$ ，它也包括三个互作分量，

$$\begin{aligned}\varpi_{GE} &= \varpi_{OE} + \varpi_{CE} + \varpi_{ME} \\ &= [G_OE(P_i) - G_OE(P_j)] + [G_CE(P_i) - G_CE(P_j)] + [G_ME(P_i) - G_ME(P_j)]\end{aligned}$$

其遗传组成为

$$\begin{aligned}\varpi_{OE}(2n) &= 2(AE_{hi} - AE_{hj}) + (DE_{hii} - DE_{hjj}) \\ \varpi_{OE}(3n) &= 3(AE_{hi} - AE_{hj}) + 3(DE_{hii} - DE_{hjj}) \\ \varpi_{CE} &= CE_{hi} - CE_{hj} \\ \varpi_{ME} &= 2(AmE_{hi} - AmE_{hj}) + (DmE_{hii} - DmE_{hjj})\end{aligned}$$

遗传主效应的杂种优势是不因环境改变而变化的优势表现。当基因型与环境效应存在互作时，还有一部分的杂种优势是由基因效应在特定环境影响下的表现所造成的。这部分杂种优势的分量是优势的互作离差。平均优势互作离差的计算公式是，

$$\begin{aligned}H_{ME}(F_n) &= GE(F_n) - \frac{1}{2}[GE(P_i) + GE(P_j)] \\ &= H_{MOE} + \frac{1}{2}\varpi_{CE} + H_{MME}\end{aligned}$$

超亲优势互作离差计算公式是，

$$\begin{aligned}H_{FE}(F_n) &= GE(F_n) - GE(P_i) \\ &= H_{FOE} + H_{FME}\end{aligned}$$

其中 H_{MOE} 是归因于直接效应 \times 环境互作效应的平均优势、 H_{MME} 是归因于母体效应 \times 环境互作效应的平均优势； H_{FOE} 是归因于直接效应 \times 环境互作效应的超亲优势、 H_{FME} 是归因于母体效应 \times 环境互作效应的超亲优势。

这两种优势都可以用杂种优势的基因 × 环境互作效应 (Δ_{GE}) 和亲本互作差异值 (ϖ_{GE}) 分量的函数表示。平均优势互作离差的分量是，

H_{ME}	$H_{MOE}(2n)$	$H_{MOE}(3n)$	H_{MME}
$H_{ME}(F_1)$	Δ_{OE}	$2\Delta_{OE} + \frac{1}{6}\varpi_{OE}$	$\frac{1}{2}\varpi_{ME}$
$H_{ME}(F_2)$	$\frac{1}{2}\Delta_{OE}$	Δ_{OE}	Δ_{ME}
⋮	⋮	⋮	⋮
$H_{ME}(F_n)$	$(\frac{1}{2})^{n-1}\Delta_{OE}$	$(\frac{1}{2})^{n-2}\Delta_{OE}$	$(\frac{1}{2})^{n-2}\Delta_{ME}$

超亲优势互作离差的分量是，

H_{FE}	$H_{FOE}(2n)$	$H_{FOE}(3n)$	H_{FME}
$H_{FE}(F_1)$	$\Delta_{OE} - \frac{1}{2}\varpi_{OE}$	$2\Delta_{OE} - \frac{1}{3}\varpi_{OE}$	—
$H_{FE}(F_2)$	$\frac{1}{2}\Delta_{OE} - \frac{1}{2}\varpi_{OE}$	$\Delta_{OE} - \frac{1}{2}\varpi_{OE}$	$\Delta_{ME} - \frac{1}{2}\varpi_{ME}$
⋮	⋮	⋮	⋮
$H_{FE}(F_n)$	$(\frac{1}{2})^{n-1}\Delta_{OE} - \frac{1}{2}\varpi_{OE}$	$(\frac{1}{2})^{n-2}\Delta_{OE} - \frac{1}{2}\varpi_{OE}$	$(\frac{1}{2})^{n-2}\Delta_{ME} - \frac{1}{2}\varpi_{ME}$

在分析种子数量形状的杂种优势遗传规律时，也可以预测杂交种子的性状表现超过母体亲本的世代数。这种预测是基于对遗传主效应的分析，其预测结果适用于对各种环境下杂种优势表现的推断。如果要求超亲优势达到 $H_F \geq \beta$ ，杂种世代数的预测公式为：

双子叶作物种子的世代数

$$n \leq 2 + \frac{\log[(\frac{1}{2}\Delta_O + \Delta_M) / (\beta + \frac{1}{2}\varpi_O + \frac{1}{2}\varpi_M)]}{\log(\beta)}$$

单子叶作物胚乳的世代数

$$n \leq 2 + \frac{\log[(\Delta_O + \Delta_M) / (\beta + \frac{1}{2}\varpi_O + \frac{1}{2}\varpi_M)]}{\log(\beta)}$$

现以 F_2 世代为例，介绍二倍体种子数量性状遗传表现以及杂种优势的预测方法。如果存在基因型 \times 环境互作效应，分析应该包括二方面：遗传主效应及其杂种优势的分析，环境互作效应及其杂种优势互作离差的分析。

遗传主效应值的预测公式是，

$$\begin{aligned} G(F_2) &= G_O(F_2) + C(F_2) + G_M(F_2) \\ &= (A_i + A_j + \frac{1}{4}D_{ii} + \frac{1}{4}D_{jj} + \frac{1}{2}D_{ij}) + C_i + (Am_i + Am_j + \frac{1}{4}Dm_{ij}) \end{aligned}$$

F_2 平均优势的预测公式是，

$$\begin{aligned} H_M(F_2) &= G(F_2) - \frac{1}{2}[G(P_i) + G(P_j)] \\ &= \frac{1}{2}\Delta_O + \frac{1}{2}\varpi_C + \frac{1}{2}\Delta_M \\ &= \frac{1}{2}[D_{ij} - \frac{1}{2}(D_{ii} + D_{jj})] + \frac{1}{2}(C_i + C_j) + [Dm_{ij} - \frac{1}{2}(Dm_{ii} + Dm_{jj})] \end{aligned}$$

F_2 超亲优势的预测公式是，

$$\begin{aligned} H_F(F_2) &= G(F_2) - G(P_i) \\ &= \frac{1}{2}(\Delta_O - \varpi_O) + (\Delta_M - \frac{1}{2}\varpi_M) \\ &= \frac{1}{2}\{[D_{ij} - \frac{1}{2}(D_{ii} + D_{jj})] - [(A_i - A_j) - (D_{ii} - D_{jj})]\} \\ &\quad + \{[Dm_{ij} - \frac{1}{2}(Dm_{ii} + Dm_{jj})] - \frac{1}{2}[(Am_i - Am_j) - (Dm_{ii} - Dm_{jj})]\} \end{aligned}$$

以上估算的是适用于各种环境条件的遗传效应值和杂种优势值。当预测在环境 h 中母本 i 与父本 j 杂交的 F_2 种子遗传表现和杂种优势表现时，需要在主效应的预测值上再加互作离差：

$$F_2 \text{ 种子遗传表现} = G(F_2) + GE(F_2)$$

$$F_2 \text{ 平均优势表现} = H_M(F_2) + H_{ME}(F_2)$$

$$F_2 \text{ 超亲优势表现} = H_F(F_2) + H_{FE}(F_2)$$

其中各项互作离差的计算公式是，

$$\begin{aligned}
 GE(F_2) &= G_OE(F_2) + CE(F_2) + G_ME(F_2) \\
 &= (AE_{hi} + AE_{hj} + \frac{1}{4}DE_{hii} + \frac{1}{4}DE_{hjj} + \frac{1}{2}DE_{hij}) + CE_i \\
 &\quad + (AmE_{hi} + AmE_{hj} + \frac{1}{4}DmE_{hij}) \\
 H_{ME}(F_2) &= GE(F_2) - \frac{1}{2}[GE(P_i) + GE(P_j)] \\
 &= \frac{1}{2}\Delta_{OE} + \frac{1}{2}\varpi_{CE} + \frac{1}{2}\Delta_{ME} \\
 &= \frac{1}{2}[DE_{hij} - \frac{1}{2}(DE_{hii} + DE_{hjj})] + \frac{1}{2}(CE_{hi} + CE_{hj}) \\
 &\quad + [DmE_{hij} - \frac{1}{2}(DmE_{hii} + DmE_{hjj})] \\
 H_{FE}(F_2) &= GE(F_2) - GE(P_i) \\
 &= \frac{1}{2}(\Delta_{OE} - \varpi_{OE}) + (\Delta_{ME} - \frac{1}{2}\varpi_{ME}) \\
 &= \frac{1}{2}[DE_{hij} - \frac{1}{2}(DE_{hii} + DE_{hjj})] \\
 &\quad - \frac{1}{2}[(AE_{hi} - AE_{hj}) - (DE_{hii} - DE_{hjj})] \\
 &\quad + [DmE_{hij} - \frac{1}{2}(DmE_{hii} + DmE_{hjj})] \\
 &\quad - \frac{1}{2}[(AmE_{hi} - AmE_{hj}) - (DmE_{hii} - DmE_{hjj})]
 \end{aligned}$$

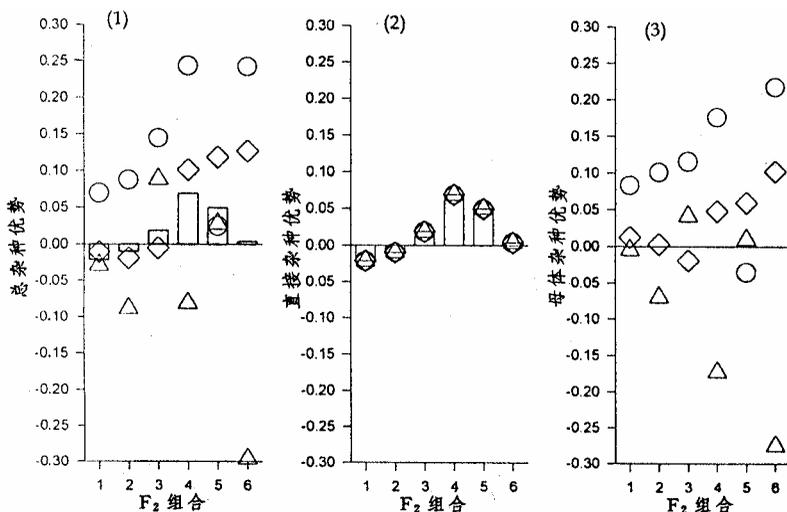


图 10.1 陆地棉 6 个组合 F_2 种子仁油分含量的群体平均优势图
 □ 主效应的杂种优势 △ 杂种优势在 1991 ◇ 杂种优势在 1992
 ○ 杂种优势在 1993

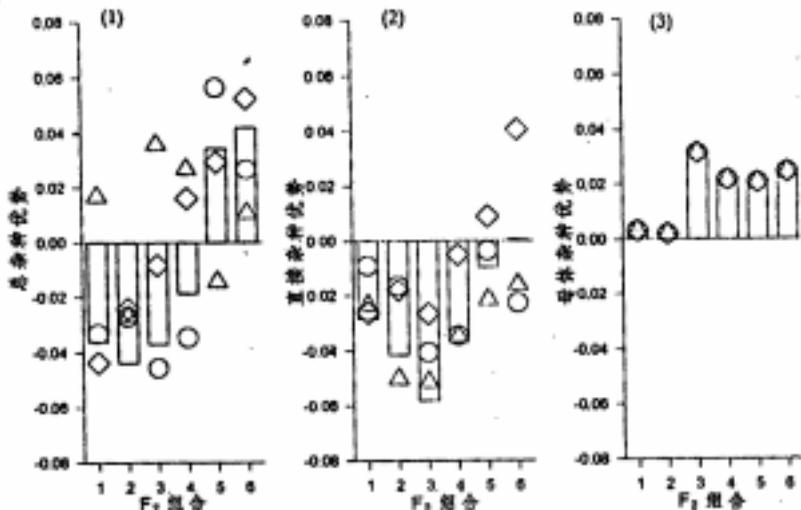


图 10.2 陆地棉 6 个组合 F_2 种子仁蛋白质含量的群体平均优势图

□ 主效应的杂种优势 △ 杂种优势在 1991 ◇ 杂种优势在 1992
○ 杂种优势在 1993

现以陆地棉 5 个亲本及其 F_1 、 F_2 和回交后代棉子仁油分含量和蛋白质含量的三年实验资料分析，作为演示作物种子性状杂种优势的分析实例。遗传试验于 1990~1993 年在浙江农业大学实验场进行，五个陆地棉品种分别为：：中棉所 7 号；：HG-H-12，高棉酚品种；：PD0111；：PD0458；：GL-5，无棉酚品种。分析的杂交组合包括：(1×3)、(1×4)、(2×3)、(2×4)、(3×5)和(4×5)。采用包括基因型与环境互作效应的作物二倍体种子遗传模型及分析方法，分析多环境下棉子营养品质的遗传特性。运用 MINQUE(0/1)法无偏地估算各性状的各项遗传效应方差分量及基因效应与环境互作方差分量；采用调整无偏预测法(AUP 法)预测各项遗传效应和杂种优势值。采用 Jackknife 重复抽样技术计算各估算值和预测值的标准误，并用 t 测验对参数进行显著性检验。

图 10.1 和图 10.2 分别列出了陆地棉 6 个杂交组合 F_2 种子仁油分

含量和蛋白含量的三种群体平均优势，(1)总的优势、(2)直接优势和(3)母体优势。其中柱状标记表示由基因主效应估算的优势，三角形标记表示杂种优势在 1991 年的遗传表现(由基因主效应和 GE 互作效应组成)，菱形标记表示杂种优势在 1992 年的遗传表现，园形标记表示杂种优势在 1993 年的遗传表现。由图 10.1 和图 10.2 可见，棉花种子仁油分含量和蛋白含量的群体平均优势同时受到基因主效应和 GE 互作效应的控制，蛋白含量的直接群体平均优势也受到基因主效应和 GE 互作效应的控制。油分含量的直接群体平均优势和蛋白含量的母体群体平均优势主要受基因主效应控制，油分含量的母体群体平均优势完全受 GE 互作效应控制。就各组合的群体平均优势而言，组合(2×4)和(3×5)的油分含量的主效应优势较高($\geq 5\%$)，但是(2×4)的互作优势大于(3×5)的，表明(2×4)的杂种优势稳定性差于(3×5)。在所分析的 6 个组合中，(4×5)的杂种优势稳定性最差，环境条件对其优势的表现影响很大。具体分析该组合的互作优势可知，环境因素主要通过影响该组合母体植株的基因表达而改变杂种平均优势的遗传表现。组合(3×5)和(4×5)的蛋白质含量表现较高的群体平均杂种优势，(4×5)的群体平均杂种优势稳定性优于(3×5)。组合(2×3)和(2×4)的蛋白质含量的直接平均优势为负值，但其母体平均优势则为正值，这种直接基因效应和母体植株基因效应的不一致表现导致了总平均优势的降低。

第十一章 遗传率和遗传相关在育种上的应用

第一节 遗传率和选择响应

运用数量遗传分析的方法，估算性状的各种遗传效应分量的变异量，了解各种变量的相对大小，对于选择育种有重要的指导意义。对主要受环境因素影响的性状，采用栽培措施改进环境条件，常能使遗传群体获得较好的性状表现。一个性状的表现型变异如果主要决定于遗传效应，则运用遗传育种的手段改良群体的遗传组成，可望取得预期的效果。在遗传的变异中，如果加性效应的变异是主要分量，在早期分离世代进行选择常能获得较好的预期效果；如果是非加性效应为主，则杂种优势的利用有较大的潜力。

遗传率是度量性状的遗传变异占表现型变异相对比率的重要遗传参数。广义遗传率(H^2)是总的遗传方差占表现型方差的比率， $H^2 = V_G/V_P$ 。狭义遗传率(h^2)一般定义为加性遗传方差占表现型方差的比率， $h^2 = V_A/V_P$ 。这一定义适用于只受加性效应和显性效应控制的那些性状，而且假设不存在基因效应×环境互作效应的互作。

现已明确数量性状不仅受加性、显性效应的控制，还可能受上位性效应或母体遗传效应的控制。另外，基因的表达也在不同程度上受到环境因素的调控，而存在基因型×环境的互作效应。因此，为了适应遗传体系的复杂性和选择育种的实际需要，遗传率的概念也应作相应的修正。

导致群体表现型发生变异的遗传原因主要有二部分：(1)遗传主效应产生的普通遗传变异(general genetic variation)，由遗传方差(V_G)来度量；(2)基因型×环境的互作效应产生的互作遗传变异(GE interaction variation)，由基因型×环境的互作方差(V_{GE})来度量。因此，遗传率也可以分解为二个分量：普通遗传率(general heritability)和互作遗传率

(interaction heretability)。这种分解适用于广义遗传率(H^2)和狭义遗传率(h^2)。

广义遗传率(H^2)的组成是,

$$H^2 = H_G^2 + H_{GE}^2$$

其中 H_G^2 是普通广义遗传率(general heretability in the broad sense), 定义为遗传方差占表现型方差的比率; H_{GE}^2 是互作广义遗传率(interaction heretability in the broad sense), 是基因型 \times 环境互作方差占表现型方差的比率。

狭义遗传率(h^2)的组成是,

$$h^2 = h_G^2 + h_{GE}^2$$

其中 h_G^2 是普通狭义遗传率(general heretability in the narrow sense), 定义为有累加性遗传效应的方差占表现型方差的比率; h_{GE}^2 是互作狭义遗传率(interaction heretability in the narrow sense), 定义为有累加性的基因效应与环境效应互作的方差占表现型方差的比率。有累加性的遗传效应是可以被选择所固定的遗传效应, 除了包括加性效应以外, 还可根据遗传模型的不同, 包括加性 \times 加性的上位性效应、母体加性效应等。

对于选择而言, 普通狭义遗传率和互作狭义遗传率都是有效的。根据普通狭义遗传率 h_G^2 而预测的选择效益适用于不同环境条件下的选择, 有较为广泛的应用价值; 而根据互作狭义遗传率 h_{GE}^2 预测的选择效果, 只度量了在某一特定环境条件下的选择效益偏差。如果需要估计在某一环境下进行选择总效益, 应该采用总的狭义遗传率 ($h_G^2 + h_{GE}^2$) 进行预测。

在实际应用时, 由于不可能测得遗传方差分量的真值, 因而也不可能得到遗传率的参数值。但是通过设计遗传试验, 采用数量遗传分析的方法, 可以获得遗传率的无偏估计值。由于估算遗传方差分量的遗传模型和分析方法常不同, 遗传率的估算方法也因此而不同。如果遗传方差主要由加性

和显性方差分量所组成，可以采用巢式交配设计、析因交配设计或双列杂交设计等遗传试验，估算加性和显性方差分量。然后用下列公式估算遗传率，

$$H_G^2 = (V_A + V_D) / V_P$$

$$H_{GE}^2 = (V_{AE} + V_{DE}) / V_P$$

$$h_G^2 = V_A / V_P$$

$$h_{GE}^2 = V_{AE} / V_P$$

其中表现型方差的计算公式是 $V_P = (V_A + V_D) + (V_{AE} + V_{DE}) + V_e$ 。如果不存在基因型×环境的互作效应，所有的互作方差分量均为零，互作遗传率也为零。此时总遗传率与普通遗传率是等量的 ($H^2 = H_G^2, h^2 = h_G^2$)。

植物种子和动物幼畜性状如果同时受到直接遗传效应和母体遗传效应的影响，对幼畜和植物种子基因型的选择和对母体基因型的选择都是有效的。因而遗传率的估算应同时考虑直接基因效应和母体基因效应，以及它们之间的相互关系。由于植物种子和动物幼畜性状同时受到三套遗传体系(种子或幼畜核基因、细胞质基因和母体核基因)的控制，而且这些遗传效应可能存在不同程度的环境互作效应，因此遗传率的计算较为复杂。广义遗传率分量的估算公式是：

$$\begin{aligned} H_G^2 &= H_O^2 + H_C^2 + H_M^2 \\ &= (V_{Go} + C_{Go.Gm}) / V_P + V_C / V_P + (V_{Gm} + C_{Go.Gm}) / V_P \\ &= (V_A + V_D + C_{A.Am} + C_{D.Dm}) / V_P + V_C / V_P \\ &\quad + (V_{Am} + V_{Dm} + C_{A.Am} + C_{D.Dm}) / V_P \end{aligned}$$

$$\begin{aligned} H_{GE}^2 &= H_{OE}^2 + H_{CE}^2 + H_{ME}^2 \\ &= (V_{GoE} + C_{GoE.GmE}) / V_P + V_{CE} / V_P + (V_{GmE} + C_{GoE.GmE}) / V_P \\ &= (V_{AE} + V_{DE} + C_{AE.AmE} + C_{DE.DmE}) / V_P + V_{CE} / V_P \\ &\quad + (V_{AmE} + V_{DmE} + C_{AE.AmE} + C_{DE.DmE}) / V_P \end{aligned}$$

其中 H_O^2 是归因于直接遗传效应的普通广义直接遗传率， H_C^2 是归因于细胞质遗传效应的普通细胞质遗传率， H_M^2 是归因于母体植株遗传效应的普通

广义母体遗传率； H_{OE}^2 是归因于直接遗传效应×环境效应的互作广义直接遗传率， H_{CE}^2 是归因于细胞质遗传效应×环境效应的互作细胞质遗传率， H_{ME}^2 是归因于母体植株遗传效应的互作广义母体遗传率；存在环境互作效应时，表现型方差的计算公式是：

$$\begin{aligned} V_P &= (V_{Go} + V_C + V_{Gm} + 2C_{Go.Gm}) + (V_{GoE} + V_{CE} + V_{GmE} + 2C_{GoE.GmE}) + V_e \\ &= (V_A + V_D + V_C + V_{Am} + V_{Dm} + 2C_{A.Am} + 2C_{D.Dm}) \\ &\quad + (V_{AE} + V_{DE} + V_{CE} + V_{AmE} + V_{DmE} + 2C_{AE.AmE} + 2C_{DE.DmE}) + V_e \end{aligned}$$

狭义遗传率分量的估算公式是：

$$\begin{aligned} h_G^2 &= h_O^2 + h_C^2 + h_M^2 \\ &= (V_A + C_{A.Am}) / V_P + V_C / V_P + (V_{Am} + C_{A.Am}) / V_P \end{aligned}$$

$$\begin{aligned} h_{GE}^2 &= h_{OE}^2 + h_{CE}^2 + h_{ME}^2 \\ &= (V_{AE} + C_{AE.AmE}) / V_P + V_{CE} / V_P + (V_{AmE} + C_{AE.AmE}) / V_P \end{aligned}$$

其中 h_O^2 是归因于直接加性效应的普通狭义直接遗传率， H_C^2 是归因于细胞质遗传效应的普通细胞质遗传率， H_M^2 是归因于母体植株加性效应的普通狭义母体遗传率； H_{OE}^2 是归因于直接加性效应×环境效应的互作狭义直接遗传率， H_{CE}^2 是归因于细胞质遗传效应×环境效应的互作细胞质遗传率， H_{ME}^2 是归因于母体植株加性效应的互作狭义母体遗传率。

如果不存在环境互作效应，所有的互作分量均为零，此时表现型方差的计算公式是：

$$\begin{aligned} V_P &= V_{Go} + V_C + V_{Gm} + 2C_{Go.Gm} + V_e \\ &= V_A + V_D + V_C + V_{Am} + V_{Dm} + 2C_{A.Am} + 2C_{D.Dm} + V_e \end{aligned}$$

生物成熟个体的一些性状常不受母体遗传效应影响，但可能受到上位性效应的影响。在上位性效应的变异中，加性与加性的上位性效应是可以遗传，并经选择可被固定的变异。在估算狭义遗传力时应包括这一部分的分量。当用简单的加性和显性遗传模型分析数量性状时，估算所得到

的加性方差分量可能已经包含有加性与加性的上位性方差分量。如果能够估算得到全部上位性效应的方差分量，遗传率的估算公式是

$$\begin{aligned}H_G^2 &= (V_A + V_D + V_{AA} + V_{DD} + V_{AD}) / V_P \\H_{GE}^2 &= (V_{AE} + V_{DE} + V_{AAE} + V_{DDE} + V_{ADE}) / V_P \\h_G^2 &= (V_A + V_{AA}) / V_P \\h_{GE}^2 &= (V_{AE} + V_{AAE}) / V_P\end{aligned}$$

存在环境互作效应时，表现型方差的计算公式是：

$$V_P = (V_A + V_D + V_{AA} + V_{DD} + V_{AD}) + (V_{AE} + V_{DE} + V_{AAE} + V_{DDE} + V_{ADE}) + V_e$$

当环境互作效应不存在时，所有的互作分量均为零，表现型方差的计算公式是：

$$V_P = V_A + V_D + V_{AA} + V_{DD} + V_{AD} + V_e$$

在实际应用时，常假设显性×显性、加性×显性的上位性效应可以忽略不计，只估算上位性效应中加性×加性的分量，这时遗传率的估算公式应是

$$\begin{aligned}H_G^2 &= (V_A + V_D + V_{AA}) / V_P \\H_{GE}^2 &= (V_{AE} + V_{DE} + V_{AAE}) / V_P \\h_G^2 &= (V_A + V_{AA}) / V_P \\h_{GE}^2 &= (V_{AE} + V_{AAE}) / V_P\end{aligned}$$

表现型方差的计算公式是：

$$V_P = (V_A + V_D + V_{AA}) + (V_{AE} + V_{DE} + V_{AAE}) + V_e$$

如果环境互作效应不存在，表现型方差的计算公式是：

$$V_P = V_A + V_D + V_{AA} + V_e$$

狭义遗传率可用于预测分离世代中对个体进行选择所能获得的选择响应(response to selection)。如果分离世代的群体平均数为 μ ，所有入选单植的平均数为 μ_S ，其差异是 $S = \mu_S - \mu$ 。选择响应是入选单株后代的平均表现与未经选择的群体平均数 μ 的差异， $R = h^2(\mu_S - \mu)$ 。选择差

($S = \mu_S - \mu$) 决定于选择强度 i 和群体的表现型变异 V_P , 即 $S = i\sqrt{V_P}$ 。因而 $R = ih^2\sqrt{V_P}$ 。当狭义遗传率为 0 时, 无论当代选择的差异 S 有多大, 后代的性状表现不会有任何改良。狭义遗传率越高, 选择的遗传效果就越好。

由于遗传率有若干分量, 基于遗传率而估计的选择响应也包括响应的分量: (1) 普通选择响应 (general response, R_G), 这一部分响应是归因于累加性的遗传主效应的, 因此这种选择在各种环境下都是有效的; (2) 互作选择响应 (interaction response, R_{GE}), 这一部分响应是归因于累加性的效应 \times 环境效应的互作, 因此只适用于特定环境下的选择。在某一环境下的总选择响应为 $R = R_G + R_{GE}$ 。如果某性状 R_G 的比重大, 预示异地(或异年)选择的有效性高, 即在一地(或一年)的环境下对该性状选择所获得的响应, 也适用于其它环境; 反之性状 R_{GE} 的比重大, 说明该性状的改良只有在特定环境下选择才有效, 异地(或异年)的选择效益较差。

选择响应的计算方法因遗传模型而异。对于加性和显性的遗传模型, 普通选择响应的估算公式是,

$$\begin{aligned} R_G &= ih_G^2\sqrt{V_P} \\ &= i\frac{V_A}{\sqrt{V_P}} \\ &= i\sqrt{h_G^2V_A} \end{aligned}$$

互作选择响应的估算公式是,

$$\begin{aligned} R_{GE} &= ih_{GE}^2\sqrt{V_P} \\ &= i\frac{V_{AE}}{\sqrt{V_P}} \\ &= i\sqrt{h_{GE}^2V_A} \end{aligned}$$

由于植物种子和动物幼畜性状同时受到直接加性效应、细胞质效应和

母体加性效应的影响,对幼畜和植物种子基因型的选择和对母体基因型(包括细胞质基因和母体植株核基因)的选择都是有效的。因而,这些性状的选择响应包括了更多的分量。普通选择响应(R_G)由普通直接响应(R_O)、普通细胞质响应(R_C)和普通母体响应(R_M)三部分组成($R_G = R_O + R_C + R_M$)。互作选择响应(R_{GE})也由互作直接响应(R_{OE})、互作细胞质响应(R_{CE})和互作母体响应(R_{ME})三部分组成($R_{GE} = R_{OE} + R_{CE} + R_{ME}$)。普通选择响应各项分量的估算方法如下:

普通直接响应	普通细胞质响应	普通母体响应
$R_O = ih_O^2 \sqrt{V_P}$ $= i \frac{V_A + C_{A.Am}}{\sqrt{V_P}}$ $= i \sqrt{h_O^2 (V_A + C_{A.Am})}$	$R_C = ih_C^2 \sqrt{V_P}$ $= i \frac{V_C}{\sqrt{V_P}}$ $= i \sqrt{h_C^2 V_A}$	$R_M = ih_M^2 \sqrt{V_P}$ $= i \frac{V_{Am} + C_{A.Am}}{\sqrt{V_P}}$ $= i \sqrt{h_M^2 (V_{Am} + C_{A.Am})}$

互作选择响应各项分量的估算方法如下:

互作直接响应	互作细胞质响应	互作母体响应
$R_{OE} = ih_{OE}^2 \sqrt{V_P}$ $= i \frac{V_{AE} + C_{AE.AmE}}{\sqrt{V_P}}$ $= i \sqrt{h_{OE}^2 (V_{AE} + C_{AE.AmE})}$	$R_{CE} = ih_{CE}^2 \sqrt{V_P}$ $= i \frac{V_{CE}}{\sqrt{V_P}}$ $= i \sqrt{h_{CE}^2 V_A}$	$R_{ME} = ih_{ME}^2 \sqrt{V_P}$ $= i \frac{V_{AmE} + C_{AE.AmE}}{\sqrt{V_P}}$ $= i \sqrt{h_{ME}^2 (V_{AmE} + C_{AE.AmE})}$

以上选择响应的公式表明:对种子性状的选择效益受到种子直接加性效应与母体植株加性效应的遗传协方差 $C_{A.Am}$ 的影响。如果协方差为正值,对种子或母体植株进行选择效益较高;但是一些性状的协方差 $C_{A.Am}$ 为负值,对这类性状的选择改良的效益较差。

第二节 遗传相关和间接选择

如果成对性状 X 和 Y 之间存在不同程度的相关(表现型相关 r_P), 这种相关性是由性状之间的遗传效应相关(遗传相关 r_G)、基因型×环境互作效应相关(互作相关 r_{GE})和环境效应的相关(环境相关 r_E)所决定的。性状 X 和性状 Y 的相关分别定义为,

$$\text{表现型相关 } r_P = C_P / \sqrt{V_{P(X)}V_{P(Y)}};$$

$$\text{遗传相关 } r_G = C_G / \sqrt{V_{G(X)}V_{G(Y)}};$$

$$\text{互作相关 } r_{GE} = C_{GE} / \sqrt{V_{GE(X)}V_{GE(Y)}};$$

$$\text{环境相关 } r_E = C_E / \sqrt{V_{E(X)}V_{E(Y)}}。$$

正如表现型方差可以被分解为不同效应的方差分量一样, 表现型相关也可被进一步分解。但表现型相关不是各项相关分量的简单相加。它是相关分量被遗传率加权的和,

$$r_P = \sqrt{H_{G(X)}^2} \sqrt{H_{G(Y)}^2} r_G + \sqrt{H_{GE(X)}^2} \sqrt{H_{GE(Y)}^2} r_{GE} \\ + \sqrt{H_{E(X)}^2} \sqrt{H_{E(Y)}^2} r_E$$

其中 $H_{E(X)}^2 = 1 - H_{G(X)}^2 - H_{GE(X)}^2$, $H_{E(Y)}^2 = 1 - H_{G(Y)}^2 - H_{GE(Y)}^2$ 。

当总遗传效应再被进一步分解为加性效应和显性效应分量时, 还可以分析以下相关分量:

$$\text{加性相关 } r_A = C_A / \sqrt{V_{A(X)}V_{A(Y)}};$$

$$\text{显性相关 } r_D = C_D / \sqrt{V_{D(X)}V_{D(Y)}};$$

$$\text{加性} \times \text{环境互作相关 } r_{AE} = C_{AE} / \sqrt{V_{AE(X)}V_{AE(Y)}};$$

$$\text{显性} \times \text{环境互作相关 } r_{DE} = C_{DE} / \sqrt{V_{DE(X)}V_{DE(Y)}}。$$

这时表现型相关的分解包括了 5 项相关分量，

$$r_P = \sqrt{h_{G(X)}^2} \sqrt{h_{G(Y)}^2} r_A + \sqrt{H_{G(X)}^2 - h_{G(X)}^2} \sqrt{H_{G(Y)}^2 - h_{G(Y)}^2} r_D \\ + \sqrt{h_{GE(X)}^2} \sqrt{h_{GE(Y)}^2} r_{AE} + \sqrt{H_{GE(X)}^2 - h_{GE(X)}^2} \sqrt{H_{GE(Y)}^2 - h_{GE(Y)}^2} r_{DE} \\ + \sqrt{H_{E(X)}^2} \sqrt{H_{E(Y)}^2} r_E$$

其中 $H_{E(X)}^2 = 1 - H_{G(X)}^2 - H_{GE(X)}^2$, $H_{E(Y)}^2 = 1 - H_{G(Y)}^2 - H_{GE(Y)}^2$ 。

种子性状间的相关分量则更多，它包括以下各种效应的相关：

直接加性相关 $r_A = C_A / \sqrt{V_{A(X)}V_{A(Y)}}$ ；

直接显性相关 $r_D = C_D / \sqrt{V_{D(X)}V_{D(Y)}}$ ；

细胞质相关 $r_C = C_C / \sqrt{V_{C(X)}V_{C(Y)}}$ ；

母体加性相关 $r_{Am} = C_{Am} / \sqrt{V_{Am(X)}V_{Am(Y)}}$ ；

母体显性相关 $r_{Dm} = C_{Dm} / \sqrt{V_{Dm(X)}V_{Dm(Y)}}$ ；

直接加性×环境互作相关 $r_{AE} = C_{AE} / \sqrt{V_{AE(X)}V_{AE(Y)}}$ ；

直接显性×环境互作相关 $r_{DE} = C_{DE} / \sqrt{V_{DE(X)}V_{DE(Y)}}$ ；

细胞质×环境互作相关 $r_{CE} = C_{CE} / \sqrt{V_{CE(X)}V_{CE(Y)}}$ ；

母体加性×环境互作相关 $r_{AmE} = C_{AmE} / \sqrt{V_{AmE(X)}V_{AmE(Y)}}$ ；

母体显性×环境互作相关 $r_{DmE} = C_{DmE} / \sqrt{V_{DmE(X)}V_{DmE(Y)}}$ ；

剩余效应相关 $r_e = C_e / \sqrt{V_{e(X)}V_{e(Y)}}$ 。

不同遗传体系之间的相关效应也可能存在相关性。如种子品质性状(S)与植株农艺性状(P)间可能存在一些遗传相关性：

种子直接加性与植株加性相关 $r_{A/Am} = C_{A/Am} / \sqrt{V_{A(S)}V_{Am(P)}}$ ；

种子直接显性与植株显性相关 $r_{D/Dm} = C_{D/Dm} / \sqrt{V_{D(S)}V_{Dm(P)}}$;

种子母体细胞质与植株细胞质相关 $r_{C/C} = C_{C/C} / \sqrt{V_{C(S)}V_{C(P)}}$;

种子母体加性与植株加性相关 $r_{Am/Am} = C_{Am/Am} / \sqrt{V_{Am(S)}V_{Am(P)}}$;

种子母体显性与植株显性相关 $r_{Dm/Dm} = C_{Dm/Dm} / \sqrt{V_{Dm(S)}V_{Dm(P)}}$ 。

性状之间相关性的遗传分析，对于选择育种具有重要的指导意义。有些重要农艺性状(如产量)的狭义遗传率较低，对这些性状进行直接选择常不能取得预期的效果。另有一些性状在被选个体上根本无法测得，因此不可能进行直接选择。如对公畜进行选择是动物改良育种的一项重要内容，但是在雄性个体上不能对产乳量等性状进行直接选择。在育种实践中可对以上这两类性状进行间接选择。

当对性状 X 进行选择时，性状 X 的选择响应是 $R_X = ih_{(X)}^2 \sqrt{V_{P(X)}}$ 。如果性状 X 和性状 Y 存在相关性，对性状 X 的选择可能会间接地影响到性状 Y。性状 Y 的相关响应(correlated response)决定于性状 X 的选择响应和两性状间加性协方差占性状 X 加性方差的比率， $CR_Y = (C_A/V_{A(X)})R_X$ 。相关响应还能以不同的方式表示，

$$\begin{aligned} CR_Y &= ir_A \sqrt{h_{(X)}^2} \sqrt{V_{A(Y)}} \\ &= ir_A \sqrt{h_{(X)}^2} \sqrt{h_{(Y)}^2} \sqrt{V_{P(Y)}} \end{aligned}$$

如果也可以对性状 Y 进行直接选择，其选择响应为 $R_Y = ih_{(Y)}^2 \sqrt{V_{P(Y)}}$ 。那么间接选择的相对遗传效益决定于相关响应对选择响应的比率可由下式计算，

$$\begin{aligned}
 CR_Y / R_Y &= \frac{ir_A \sqrt{h^2_{(X)}} \sqrt{h^2_{(Y)}} \sqrt{V_{P(Y)}}}{ih^2_{P(Y)} \sqrt{V_{P(Y)}}} \\
 &= \frac{r_A \sqrt{h^2_{(X)}}}{\sqrt{h^2_{(Y)}}}
 \end{aligned}$$

因此只有当 $r_A \sqrt{h^2_{(X)}}$ 大于 $\sqrt{h^2_{(Y)}}$ 时, 间接选择的遗传效益才能超过直接选择的。当育种的目标性状遗传率较低时, 选择与其有较高加性相关的其它性状, 较易取得育种成果。

在作物品质性状的改良中, 如果农艺性状与种子品质性状之间存在遗传相关性, 可以通过选择农艺性状(P)而间接种子改良品质性状(S)。间接选择的相关响应可以由以下公式估算,

$$CR_P = i(C_{A/Am} + C_{Am/Am}) / \sqrt{V_{P(P)}}$$

其中 $C_{A/Am}$ 是种子品质性状直接加性效应与农艺性状植株加性效应的协方差, $C_{Am/Am}$ 是种子品质性状母体加性效应与农艺性状植株加性效应的协方差, $V_{P(P)}$ 是植株农艺性状的表现型方差。

第十二章 多年份、多试点品种区域试验的分析方法

第一节 单一性状的分析方法

数量遗传的分析结果应用于育种改良,可指导品种的选育。品种区域试验是农作物选育过程中评定参试品种和推广优良品种的一个重要环节。农作物品种需要通过区域试验鉴定,才能确定其在多年份和多试点的农艺性状表现。由于作物农艺性状大多是数量性状,其表现除了受品种基因型影响外,还受环境因素的影响。运用统计分析的方法分析区域试验的资料,可以排除各种非遗传因素的干扰,从而直接评价品种的基因型效应。区域试验的平衡资料(balanced data)一般可采用传统的方差分析(ANOVA)的方法进行分析。这种方法有许多优点:计算简便;可用 F 测验作统计检验;对平衡数据的分析能获得最佳无偏估算。但是该方法最大的缺点是不能有效地分析有缺失的非平衡数据(unbalanced data)。在区域试验实施过程中,由于每年参试品种可能不同,某些试点的数据可能缺失,传统的ANOVA方法不能对多年份的区域试验非平衡资料进行综合分析。对于有缺失的非平衡数据,采用一些流行的统计分析软件(如SAS的GLM),虽可算得变异的均方值,但因它们不是相互独立的,因而无法对处理效应作 F 测验。朱军、许馥华和赖鸣冈(1993)运用混合线性模型的分析原理,提出了作物区域试验非平衡资料的统计分析方法。该分析方法不需要估算均方,通过直接估算各项随机效应的方差分量,并进一步估算品种平均数间差异或线性对比值及其标准误,从而进行相应的统计检验。他们还提出了用数值重复抽样方法计算回归参数的估计值及其标准误,以评价品种的稳定性。

如果有 g 个品种参加多年份联合区域试验,各年份和各试点的区组数为 r 。每年的参试品种可能不尽相同,某些年份的试点也可能不尽相同(即在某年中可能缺失了一些试点的资料),这将导致多年份联合区域试验的资料成为非平衡数据。根据联合区域试验常规设计,第 h 个品种在第 i 年中第 j 个试点内第 k

个随机区组的表现型值 y_{hijk} 可用以下线性模型表示,

($h = 1, 2, \dots, g; i = 1, 2, \dots, n_h; j = 1, 2, \dots, n_{hi}; k = 1, 2, \dots, r$),

$$y_{hijk} = G_h + Y_i + L_j + YL_{ij} + GY_{hi} + GL_{hj} + GYL_{hij} + B_{k(ij)} + e_{hijk} \quad (12.1)$$

其中 G_h 是品种固定效应, $\frac{1}{g} \sum_{h=1}^g G_h = \mu$, μ 为群体平均数;

Y_i 是年份随机效应, $Y_i \sim N(0, \sigma_Y^2)$;

L_j 是试点随机效应, $L_j \sim N(0, \sigma_L^2)$;

YL_{ij} 是年份 \times 试点互作随机效应, $YL_{ij} \sim N(0, \sigma_{YL}^2)$;

GY_{hi} 是品种 \times 年份互作随机效应, $GY_{hi} \sim N(0, \sigma_{GY}^2)$;

GL_{hj} 是品种 \times 试点互作随机效应, $GL_{hj} \sim N(0, \sigma_{GL}^2)$;

GYL_{hij} 是品种 \times 年份 \times 试点互作随机效应, $GYL_{hij} \sim N(0, \sigma_{GYL}^2)$;

$B_{k(ij)}$ 是年份中试点内随机区组效应, $B_{k(ij)} \sim N(0, \sigma_B^2)$;

e_{hijk} 是随机机误效应, $e_{hijk} \sim N(0, \sigma_e^2)$ 。

这是一个混合线性模型, 其中除了品种效应为固定效应外, 其它效应均为随机效应。因而第 h 个品种在第 i 年中第 j 个试点内第 k 个随机区组的表现型值的正态分布是:

$$y_{hijk} \sim N(G_h, \sigma_Y^2 + \sigma_L^2 + \sigma_{YL}^2 + \sigma_{GY}^2 + \sigma_{GL}^2 + \sigma_{GYL}^2 + \sigma_B^2 + \sigma_e^2)$$

作物品种联合区域试验所调查的各种性状, 往往只有产量性状是以小区为单位的(区组数 $r > 1$), 其它性状通常以试验点为单位(区组数 $r = 1$)。为了便于对各性状进行分析, 可对所有的性状以试点为单位进行统计分析。当性状有区组观察值时, 可用区组平均数估计该试点观察值。第 h 个品种在第 i 年中第 j 个试点的表型值(区组数 $r = 1$)或试点内 r 个区组($r > 1$)平均数的线性模型是,

$$y_{hij} = G_h + Y_i + L_j + YL_{ij}^* + GY_{hi} + GL_{hj} + GYL_{hij}^* \quad (12.2)$$

其中 $YL_{ij}^* = YL_{ij} + \frac{1}{r} \sum_{k=1}^r B_{k(ij)}$, $YL_{ij}^* \sim N(0, \sigma_{YL}^2 + \frac{1}{r} \sigma_B^2)$ 或 $\sim N(0, \sigma_{YL^*}^2)$;

$$GYL_{hij}^* = GYL_{hij} + \frac{1}{r} \sum_{k=1}^r e_{hijk}, \quad GYL_{hij}^* \sim N(0, \sigma_{GYL}^2 + \frac{1}{r} \sigma_e^2) \text{ 或 } \sim N(0, \sigma_{GYL^*}^2)$$

因而第 h 个品种在第 i 年中第 j 个试点内的区组平均数具有以下正态分布

$$y_{hij} \sim N(G_h, \sigma_Y^2 + \sigma_L^2 + \sigma_{YL^*}^2 + \sigma_{GY}^2 + \sigma_{GL}^2 + \sigma_{GYL^*}^2)$$

如果将作物品种在各试点观察值的线性模型以矩阵形式表示, 多年份联合区域试验的资料可以用以下混合线性模型表示,

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\mathbf{g} + \mathbf{U}_Y\mathbf{e}_Y + \mathbf{U}_L\mathbf{e}_L + \mathbf{U}_{YL^*}\mathbf{e}_{YL^*} + \mathbf{U}_{GY}\mathbf{e}_{GY} + \mathbf{U}_{GL}\mathbf{e}_{GL} + \mathbf{e}_{GYL^*} \\ &= \mathbf{X}\mathbf{g} + \sum_{u=1}^5 \mathbf{U}_u\mathbf{e}_u + \mathbf{U}_6\mathbf{e}_6 \end{aligned} \quad (12.3)$$

其中 \mathbf{y} 是 $(n \times 1)$ 的观察值向量, \mathbf{g} 是品种固定效应的向量 ($\mathbf{g} = [G_h]$), \mathbf{X} 是品种固定效应的系数矩阵; \mathbf{e}_u 是第 u 项独立随机变量的向量, 具有平均数零、方差 $\sigma_u^2 \mathbf{I}$, \mathbf{U}_u 是第 u 项随机效应的系数矩阵; $\mathbf{U}_6 = \mathbf{I}$ 是单位矩阵。

观察值向量 \mathbf{y} 具有以下多变量多元正态分布

$$\mathbf{y} \sim \text{MVN}(\mathbf{X}\mathbf{g}, \sum_{u=1}^6 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T)$$

采用最小范数二阶无偏估算法(MINQUE法), 可以无偏地估算混合线性模型中的各项方差分量。估算区域试验混合线性模型中的方差分量, 可以揭示环境效应(年份、试点、年份 \times 试点)和基因型 \times 环境互作效应(品种 \times 年份、品种 \times 试点)对性状表现的影响。方差分量估计值可直接用于估算品种效应线性对比的标准误。

可以采用不同的统计方法估算以上混合线性模型中的品种固定效应 (G_h)。如果观察值向量 \mathbf{y} 的方差分量真值 σ_u^2 是可知的, 固定效应的最佳线性无偏估算值可由广义最小二乘法算得。但是方差分量真值实际上是不可知的。

如果采用方差分量的估计量 $\hat{\sigma}_u^2$ ，则固定效应的广义最小二乘法估算值便失去了无偏性，也不再是线性的了。最简单的估算方法是采用最小二乘法，由品种的平均数估算品种固定效应，

$$\hat{G}_h = \bar{y}_{h..} = \frac{1}{n_h} \sum_{i=1}^{n_h} \left(\frac{1}{n_{hi}} \sum_{j=1}^{n_{hi}} y_{hij} \right) \quad (12.5)$$

其中 n_h 是品种 h 参试的年份数， n_{hi} 是品种 h 在第 i 年份中的试点数。第 h 个品种在所有年份的总试点数是 $n_h = \sum_{i=1}^{n_h} n_{hi}$ 。品种 h 的平均数可以用一个常数转置向量 \mathbf{m}_h^T 与观察值向量 \mathbf{y} 的乘积来表示， $\hat{G}_h = \mathbf{m}_h^T \mathbf{y}$ 。常数向量 \mathbf{m} 有 n 个系数，其对应于品种 h 观察值的系数为 $\frac{1}{n_h}$ ，其它系数均为 0。品种固定效应的估算值是线性无偏的。

参试品种的线性对比(linear contrast)是国际上广泛应用的一种品种间比较方法，它可以在二个品种间或不同组类的品种之间进行比较。参试品种的任何线性对比可以用品种平均数的线性函数表示为，

$$C = \sum_{h=1}^g c_h \bar{y}_{h..} = \sum_{h=1}^g c_h \mathbf{m}_h^T \mathbf{y} = \left(\sum_{h=1}^g c_h \mathbf{m}_h^T \right) \mathbf{y} \quad (12.6)$$

如果设系数转置向量 $\mathbf{c}^T = \sum_{h=1}^g c_h \mathbf{m}_h^T$ ，则线性对比值 C 可以用观察值向量 \mathbf{y} 的线性函数来表示， $C = \mathbf{c}^T \mathbf{y}$ 。线性对比值 C 是正态分布，具有平均数

$$\mathbf{c}^T \mathbf{Xg} = \sum_{h=1}^g c_h G_h$$

和方差

$$\sigma^2(C) = \mathbf{c}^T \left(\sum_{u=1}^6 \sigma_u^2 \mathbf{U}_u \mathbf{U}_u^T \right) \mathbf{c}$$

如果设 $a_u = (\mathbf{c}^T \mathbf{U}_u)(\mathbf{U}_u^T \mathbf{c})$, 则线性对比值 C 的方差可由下式估算

$$\hat{\sigma}^2(C) = \sum_{u=1}^6 a_u \hat{\sigma}_u^2 \quad (12.7)$$

当试验数据平衡时(品种数 = g , 年份数 = n_Y , 试点数 = n_L), 该表达式便可简化为

$$\hat{\sigma}^2(C) = \frac{1}{n_Y n_L} (n_L \hat{\sigma}_{GY}^2 + n_Y \hat{\sigma}_{GL}^2 + \hat{\sigma}_{GYL*}^2) \sum_{h=1}^g c_h^2$$

该值等于以区组平均数模型(12.2) (品种、年份和试点三因素方差分析)所估计的品种线性对比 C 的方差

$$\hat{\sigma}^2(C) = \frac{1}{n_Y n_L} (MS_{GY} + MS_{GL} - MS_{GYL*}) \sum_{h=1}^g c_h^2$$

其中 MS_{GY} 是品种 \times 年份互作均方值, MS_{GL} 是品种 \times 试点互作均方值, MS_{GYL*} 是剩余效应均方值。

由于品种区域试验的群体一般都较大, 因而可以采用 z 测验检验以下广义统计假设,

$$\text{无效假设 } H_0: \sum_{h=1}^g c_h G_h = 0$$

$$\text{备择假设 } H_1: \sum_{h=1}^g c_h G_h \neq 0$$

如果估算值 $C/\sqrt{\hat{\sigma}^2(C)}$ 的绝对值大于 $z(\alpha)$, 便可在 α 显著水平上否定无效假设 H_0 , 从而接受备择假设 H_1 。

线性对比系数必须附合限定条件 $\sum_{h=1}^g c_h = 0$ 。广义统计假设的具体内容取决于线性对比系数 c_h 的选择。现以8个品种的区域试验为例, 说明系数 c_h 的选择方法。8个品种的平均数为 $[\bar{y}_{1.}, \bar{y}_{2.}, \dots, \bar{y}_{8.}]$, 它们分别是品种固定效应 $[G_1, G_2, \dots, G_8]$ 的无偏估计。如果要检验品种 $h (h \neq 8)$ 和品种8(对照)之间的差异、

即检验 $H_0: G_h = G_8$, 可以设 $c_h = 1$ 、 $c_8 = -1$, 而其它系数均为0、因而

$$C = \bar{y}_{h..} - \bar{y}_{8..}$$

如果希望检验所有育成品种($h < 8$)的平均表现和对照品种($h = 8$)之间是否存在差异, 可以设 $c_h = 1$ ($h < 8$)、 $c_8 = -7$, 因而

$$C = \sum_{h=1}^7 \bar{y}_{7..} - 7\bar{y}_{8..}$$

由于品种的线性对比包括了品种间或不同品种类别间的比较, 品种区域试验宜采用这种比较方法。采用混合线性模型方法分析平衡数据时, 能获得与传统的方差分析方法一致的分析结果。

为了评价品种的稳定性表现, 常采用简单回归分析的方法估算品种对环境指数的回归参数。对以模型(12.2)所表示的区域试验非平衡数据, 回归分析的依变量是品种 h 在各年份和各试点的平均表现 y_{hij} , 自变量则是各年份和各试点的环境指数(I_{ij})。环境指数通常由所有品种在该年份和该试点的平均表现估算,

$$I_{ij} = \frac{1}{n_{ij}} \sum_{h=1}^{n_{ij}} y_{hij} \quad (12.8)$$

其中 n_{ij} 是第 i 年中第 j 个试点内所包括的参试品种, $n_{ij} \leq$ 总参试品种数 g 。由于环境指数由品种平均数估计而得, 它不再是固定常量, 而是随机变量, 并具有正态分布

$$I_{ij} \sim N(\mu, \sigma_Y^2 + \sigma_L^2 + \sigma_{YL*}^2 + \frac{1}{n_{ij}}(\sigma_{GY}^2 + \sigma_{GL}^2 + \sigma_{GYL*}^2))$$

随机变量 y_{hij} 和 I_{ij} 也不是相互独立的。在这种情况下, 不能采用常规的回归分析方法估算回归参数估计值的方差, 从而也无法对回归参数作相应的统计推断。

可以采用Jackknife方法计算回归参数的估计值及其标准误。以每个试点的观察值作为Jackknife的重复抽样单位。对于某项回归参数 Φ (回归截距 a 、回归斜率 b 或相关系数 r), 分析所有观察值时可以得到估计值 $\hat{\Phi}$ 。当从资料中

剔除观察值 y_{hij} 和 I_{ij} 以后, 重新估算则可得到新的估计值 $\hat{\Phi}_{(ij)}$ 。采用此法从完整的数据中每次剔除一个观察值 y_{hij} 和 I_{ij} ($i=1,2,\dots,n_h; j=1,2,\dots,n_{hi}$), 可以获得 n_h 个不同的估计值 $\hat{\Phi}_{(ij)}$ 。回归参数的 Jackknife 估计值 $\hat{\Phi}_J$ 和抽样方差 $\text{Var}(\hat{\Phi}_J)$ 分别为

$$\hat{\Phi}_J = n_h \hat{\Phi} - (n_h - 1) \hat{\Phi}. \quad (12.9)$$

$$\text{Var}(\hat{\Phi}_J) = \frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} (\hat{\Phi}_{(ij)} - \hat{\Phi})^2$$

其中 $\hat{\Phi} = \frac{1}{n_h} \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} \hat{\Phi}_{(ij)}$ 。

回归参数估计值 $\hat{\Phi}_J$ 的标准误 $\text{SE}(\hat{\Phi}_J) = \sqrt{\text{Var}(\hat{\Phi}_J)}$, 统计量

$$\frac{\hat{\Phi}_J - \Phi_J}{\text{SE}(\hat{\Phi}_J)}$$

近似地具有自由度为 $(n_h - 1)$ 的 t 分布。从而可以用 t 测验对回归参数作统计检验, 或者由 $\hat{\Phi}_J \pm t_{(\alpha, n_h - 1)} \times \text{SE}(\hat{\Phi}_J)$ 确定回归参数的 $100(1 - \alpha)\%$ 置信区间。

朱军、许馥华和赖鸣冈(1993)在论文中分析了1989年至1990年黄河流域棉花品种区域试验(8个品种、24个区试点)的霜前皮棉产量(公斤/亩)资料。由于1989年临汾试点缺失了一个品种(中206)的资料, 1990年因气候异常菏泽试点资料不能使用, 因而这两年的联合区域试验资料属于非平衡数据, 不能用ANOVA方法对两年资料合并分析。通常的方法是: 删除1989年临汾试点其它七个品种的资料, 用ANOVA方法分别分析两年的资料; 或者删除1989年临汾和菏泽试点以及删除1989年临汾试点的资料, 分析其余22个试点的两年平衡数据。而采用混合显性模型的分析方法, 则不必删除任何数据。

影响霜前皮棉产量的随机效应的方差分量估计值分别为: 年份 $\hat{\sigma}_Y^2 =$

4.80, 试点 $\hat{\sigma}_L^2 = 7.36$, 年份 \times 试点 $\hat{\sigma}_{YL}^2 = 127.06$, 品种 \times 年份 $\hat{\sigma}_{GY}^2 = 0.64$, 品种 \times 试点 $\hat{\sigma}_{GL}^2 = 8.67$, 及剩余效应 $\hat{\sigma}_{GYL}^2 = 34.00$ 。其中年份 \times 试点互作方差分量值最大, 这表明24个试点的产量表现在1989年和1990年很不一致。8个品种及其代号分别为(1)冀84-25、(2)运1729、(3)平28、(4)石3409、(5)冀植17、(6)中206、(7)鲁155、(8)中12。它们的平均霜前皮棉产量和5%显著性测验结果如下:

冀84-25	运1729	平28	石3409	冀植17	中206	鲁155	中12(CK)
50.3	58.2	64.9	60.9	59.5	67.0	57.5	65.4
		A			A		A
	B		B	B			
	C			C		C	

这表明, 中206和平28与对照中12的霜前皮棉产量差异不显著, 其它参试品种的霜前皮棉产量都显著低于对照。

由于8个品种的育种单位不尽相同, 通过设置不同的线性对比系数 $[c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8]$, 可以比较不同单位选送品种间的产量差异。如果要比较平28和鲁155两个品种与中国农科院棉花研究所(中棉所)育成的中206和中12的平均差异, 可设线性对比系数为 $[0, 0, 1, 0, 0, -1, 1, -1]$ 。此时线性对比值 $C = -10.04$, 其标准误 $SE = 2.39$ 。估算值 $C \div SE = -4.20$, 其绝对值大于 $\alpha (0.01)$ 。这表明这两个品种平均霜前皮棉产量极显著地低于中棉所育成的品种。设线性对比系数为 $[2, 0, 0, 2, 2, -3, 0, -3]$, 可以比较冀84-25、石3409和冀植17三个品种与中棉所育成品种中206和中12的平均差异。由于线性对比值 $C = -55.60$, 标准误 $SE = 6.55$ 。估算值 $C \div SE = -8.48$, 其绝对值大于 $\alpha (0.01)$, 因此这三个品种平均霜前皮棉产量也极显著地低于与中棉所育成品种。

8个参试品种的稳定性分析结果列于表12.1。霜前皮棉产量对环境指数的回归分析表明:平28和鲁155的回归斜率显著小于1.0(置信区间上限 < 1.0), 这

两个品种对环境指数的反应不敏感。平28的相关系数最低(置信区间上限 < 0.83), 因而品种平28还表现一致性较差; 其它6个品种的回归截距与0.0无显著差异(置信区间下限 < 0.0 < 置信区间上限), 回归斜率与1无显著差异(置信区间下限 < 1.00 < 置信区间上限), 这些品种在稳定性方面表现相似。

表12.1 棉花8个品种霜前皮棉产量对环境指数回归分析的参数估计值和95%置信区间

品种	截距		斜率		相关系数	
	估计值	置信区间	估计值	置信区间	估计值	置信区间
冀84-25	-4.3	-13.3 ~ 4.6	0.90	0.75 ~ 1.06	0.86	0.77 ~ 0.94
运1729	-5.3	-13.3 ~ 2.7	1.04	0.93 ~ 1.17	0.94	0.91 ~ 0.97
平28	21.7	8.5 ~ 34.9	0.72	0.50 ~ 0.93	0.70	0.56 ~ 0.83
石3409	-7.0	-16.0 ~ 2.0	1.12	0.98 ~ 1.26	0.94	0.90 ~ 0.98
冀植17	-8.0	-16.7 ~ 0.7	1.12	0.98 ~ 1.25	0.94	0.90 ~ 0.97
中206	-1.3	-11.0 ~ 8.5	1.13	0.98 ~ 1.28	0.94	0.90 ~ 0.97
鲁155	7.1	-2.5 ~ 16.6	0.83	0.68 ~ 0.99	0.85	0.75 ~ 0.96
中12(CK)	-3.4	-12.3 ~ 5.4	1.13	0.99 ~ 1.28	0.94	0.91 ~ 0.97

第二节 综合性状的分析方法

农作物品种的经济价值取决于其产量、生育期、品质和抗性等性状的综合表现。作物区域试验的规范程序一般都包括了对这些性状的调查, 然后借助于统计分析综合评价参试品种。目前国内主要采用模糊数学的方法, 对区域试验参试品种进行综合分析。区域试验资料的模糊综合评价, 可以对参试品种的综合表现进行位次排列, 但无法对品种之间的差异作适当的统计检验。模糊综合评价方法应用于农艺性状综合评价时, 需要人为地将具有连续分布

的数量性状观察值转换成非连续分布的隶属函数值。这种转换使隶属函数值丧失了原始数据所具有的统计分布，因而其合理性还有待探讨。国外则主要采用经典的多变量方差分析(MANOVA)方法，对综合性状表现进行线性对比(linear contrast)及统计检验。但是这些方法尚不能有效地分析不规则缺失的非平衡数据。

朱军、赖鸣冈和许馥华(1993)运用混合线性模型的分析原理，提出了作物多年份区域试验非平衡数据的综合性状分析方法。该分析方法通过直接估算各项随机效应的方差和协方差分量，然后估算品种综合性状线性对比值及其标准误，从而进行相应的统计检验。还提出了用数值重复抽样方法计算综合性状平均表现对综合环境指数的回归参数的估计值及其标准误，以评价品种的综合稳定性。

如果多年份品种区域试验共有 g 个参试品种，参试品种的评价取决于对 t 个农艺性状的综合分析。由于区域试验往往只对产量等少数性状统计各区组资料，而对品质、抗性及其它农艺性状只分析各试点资料。故对不同性状作综合分析时，宜以年份内的各试点为单位分析区域试验资料。当每个试点有 r 个区组的数据时，可以将区组值平均，求得试点平均值。各年份所有试点都有统一的对照品种，但是每年的其它参试品种可能不尽相同，某些年份的试点也可能会有缺失资料。首先计算各年份所有试点对照品种的每个性状总平均数，然后将每个性状的所有观察值转换成统一对照的百分率(除以该性状对照总平均数，再乘以100)，以消除性状之间的度量差别。如果把原始数据转换成对其它品种或群体平均数的百分率，预期可以得到相同的分析结果。由以下线性模型表示性状 f 的第 h 个品种在第 i 年中第 j 个试点的表型值(区组数 $r=1$)或试点内 r 个区组($r \geq 1$)的平均值

$$y_{hij(f)} = G_{h(f)} + Y_{i(f)} + L_{j(f)} + YL_{ij(f)}^* + GY_{hi(f)} + GL_{hj(f)} + GYL_{hij(f)}^* \quad (12.10)$$

其中 $G_{h(f)}$ 是性状 f 的品种固定效应， $\frac{1}{g} \sum_{h=1}^g G_{h(f)} = \mu_{(f)}$ ， $\mu_{(f)}$ 为性状 f 的群体平均

数; 性状 f 的其它各项随机效应(包括 $YL^*_{ij(f)}$ 和 $GYL^*_{hij(f)}$)的定义与单一性状分析模型(12.2)相同。

随机变量 $y_{hij(f)}$ 具有以下正态分布

$$y_{hij(f)} \sim N(G_{h(f)}, \sigma_{Y(f)}^2 + \sigma_{L(f)}^2 + \sigma_{YL^*(f)}^2 + \sigma_{GY(f)}^2 + \sigma_{GL(f)}^2 + \sigma_{GYL^*(f)}^2)$$

如果以矩阵形式表示性状 f 在各试点的平均值, 多年份联合区域试验的各性状资料可以用以下混合线性模型表示,

$$\begin{aligned} \mathbf{y}_{(f)} &= \mathbf{X}\mathbf{g}_{(f)} + \mathbf{U}_Y\mathbf{e}_{Y(f)} + \mathbf{U}_L\mathbf{e}_{L(f)} + \mathbf{U}_{YL^*}\mathbf{e}_{YL^*(f)} + \mathbf{U}_{GY}\mathbf{e}_{GY(f)} \\ &\quad + \mathbf{U}_{GL}\mathbf{e}_{GL(f)} + \mathbf{e}_{GYL^*(f)} \end{aligned} \quad (12.11)$$

$$= \mathbf{X}\mathbf{g}_{(f)} + \sum_{u=1}^5 \mathbf{U}_u\mathbf{e}_{u(f)} + \mathbf{U}_6\mathbf{e}_{6(f)}$$

其中 $\mathbf{g}_{(f)}$ 是性状 f 的品种固定效应向量, \mathbf{X} 是品种固定效应的系数矩阵; $\mathbf{e}_{u(f)}$ 是性状 f 的第 u 项独立随机变量向量, 具有平均数零、方差 $\sigma_{u(f)}^2\mathbf{I}$, \mathbf{U}_u 是第 u 项随机效应的系数矩阵; $\mathbf{U}_6 = \mathbf{I}$ 是单位矩阵。

$\mathbf{y}_{(f)}$ 具有多变量多元正态分布, 其期望值为

$$E(\mathbf{y}_{(f)}) = \mathbf{X}\mathbf{g}_{(f)}$$

方差矩阵为

$$\text{Var}(\mathbf{y}_{(f)}) = \sum_{u=1}^6 \sigma_{u(f)}^2 \mathbf{U}_u \mathbf{U}_u^T$$

性状 f 的观察值向量 $\mathbf{y}_{(f)}$ 与性状 f' 的观察值向量 $\mathbf{y}_{(f')}$ 之间的协方差矩阵为

$$\text{Car}(\mathbf{y}_{(f)}, \mathbf{y}_{(f')}) = \sum_{u=1}^6 \sigma_{u(f f')} \mathbf{U}_u \mathbf{U}_u^T$$

其中 $\sigma_{(f f')}$ 是性状 f 和性状 f' 的第 u 项协方差分量。

采用最小范数二阶无偏估算法(即MINQUE法), 可以无偏地估算混合线性模型中的各项方差和协方差分量。估算区域试验混合线性模型中的协方差分量, 可以揭示环境效应(年份、试点、年份 \times 试点)和基因型 \times 环境互作效应(品种 \times 年份、品种 \times 试点)对性状 f 和性状 f' 表现的共同影响。方差分量($f=f'$)

和协方差分量($f \neq f'$) 估计值将用于估算品种综合性状线性对比的标准误。

可以采用最小二乘法, 由各性状的品种平均数估算其固定效应,

$$\hat{G}_{h(f)} = \bar{y}_{h..(f)} = \frac{1}{n_h} \sum_{i=1}^{n_h} \left(\frac{1}{n_{hi}} \sum_{j=1}^{n_{hi}} y_{hij(f)} \right) \quad (12.13)$$

其中 n_h 是品种 h 参试的年份数, n_{hi} 是品种 h 在第 i 年份中的试点数。第 h 个品种在所有年份的总试点数是 $n_h = \sum_{i=1}^{n_h} n_{hi}$ 。第 h 个品种性状 f 的平均数可以用一个常

数转置向量 \mathbf{m}_h^T 与观察值向量 $\mathbf{y}_{(f)}$ 的乘积来表示,

$$\hat{G}_{h(f)} = \mathbf{m}_h^T \mathbf{y}_{(f)}$$

其中 \mathbf{m}_h 是一个($n \times 1$)的向量, 其对应于第 h 个品种分析值的系数为 $\frac{1}{n_h}$, 其它

系数均为0。

性状 f 的品种线性对比可以用该性状各品种平均数的线性函数表示为,

$$C_{(f)} = \sum_{h=1}^g c_h \bar{y}_{h..(f)} = \sum_{h=1}^g c_h \mathbf{m}_h^T \mathbf{y}_{(f)} = \left(\sum_{h=1}^g c_h \mathbf{m}_h^T \right) \mathbf{y}_{(f)} \quad (12.14)$$

其中 c_h 是线性对比系数, 必须附合限定条件 $\sum_{h=1}^g c_h = 0$ 。

如果设系数转置向量 $\mathbf{c}^T = \sum_{h=1}^g c_h \mathbf{m}_h^T$, 则性状 f 的线性对比值 $C_{(f)}$ 可以用观

察值向量 $\mathbf{y}_{(f)}$ 的线性函数来表示, $C_{(f)} = \mathbf{c}^T \mathbf{y}_{(f)}$ 。综合性状的线性对比值 C_W 由各性状的线性对比值 $C(f)$ 加权和求得,

$$C_W = \sum_{f=1}^t w_f C_{(f)}$$

其中 w_f 是对性状 f 的加权系数, $\sum_{f=1}^t w_f = 1$ 。 C_W 实际上是加权的综合性状指数

$$W_h = \sum_{f=1}^t w_f \bar{y}_{h..(f)}$$

的线性对比,

$$\begin{aligned} C_W &= \sum_{f=1}^t \sum_{h=1}^g w_f c_h \bar{y}_{h..(f)} \\ &= \sum_{f=1}^t c_h \left(\sum_{h=1}^g w_f \bar{y}_{h..(f)} \right) \\ &= \sum_{f=1}^t c_h W_h \end{aligned} \tag{12.15}$$

综合性状的线性对比值 C_W 为正态分布, 具有平均数

$$\sum_{f=1}^t c_h \left(\sum_{h=1}^g w_f G_{h(f)} \right)$$

和方差

$$\begin{aligned} \sigma^2(C_W) &= \sum_{f=1}^t w_f^2 \sigma^2(C_{(f)}) + 2 \sum_{f=1}^{t-1} \sum_{f'=f+1}^t w_f w_{f'} \sigma(C_{(f)}, C_{(f')}) \\ &= \sum_{u=1}^g a_u \left[\sum_{f=1}^t w_f^2 \sigma_{u(f)}^2 + 2 \sum_{f=1}^{t-1} \sum_{f'=f+1}^t w_f w_{f'} \sigma_{(f,f')} \right] \end{aligned}$$

其中 $\sigma_{u(f)}^2$ 是性状 f 的第 u 项方差分量, $\sigma_{u(f,f')}$ 是性状 f 和性状 f' 的第 u 项协方差分量, $a_u = (\mathbf{c}^T \mathbf{U}_u)(\mathbf{U}_u^T \mathbf{c})$ 。用方差分量和协方差分量的估计值代入以上方差公式, 便可获得综合性状线性对比值 C_W 的方差估计值 $\hat{\sigma}^2(C_W)$ 。

可以采用标准正态分布的 z 测验检验以下广义统计假设,

$$\text{无效假设 } H_0: \sum_{h=1}^g c_h \left(\sum_{f=1}^t w_f G_{h(f)} \right) = 0$$

$$\text{备择假设 } H_1: \sum_{h=1}^g c_h \left(\sum_{f=1}^t w_f G_{h(f)} \right) \neq 0$$

如果 $C_W / \sqrt{\hat{\sigma}^2(C_W)}$ 的绝对值大于 $z(\alpha)$, 便可在 α 显著水平上否定无效假设 H_0 , 从而接受备择假设 H_1 。广义统计假设的具体内容取决于线性对比系数 c_h 的选择。

为了综合评价品种的稳定性表现, 可以采用简单回归分析的方法估算品种综合性状表现对综合环境指数的回归参数。回归分析的依变量是品种 h 在各年份和各试点的综合性状平均表现 $y_{hij(\cdot)}$,

$$y_{hij(\cdot)} = \sum_{f=1}^t w_f y_{hij(f)}$$

自变量则是各年份和各试点的综合环境指数 ($I_{ij(\cdot)}$),

$$\begin{aligned} I_{ij(\cdot)} &= \sum_{f=1}^t w_{(f)} I_{ij(f)} \\ &= \frac{1}{n_{ij}} \sum_{f=1}^t w_f \left(\sum_{h=1}^{n_{ij}} y_{hij(f)} \right) \end{aligned}$$

其中 n_{ij} 是第 i 年中第 j 个试点内所包括的参试品种, $n_{ij} \leq$ 总参试品种数 g 。

由于综合环境指数由品种各性状加权平均估计而得, 它不再是固定常量, 而是正态随机变量, 具有均值

$$\sum_{f=1}^t w_f \mu_{(f)}$$

和方差

$$\sum_{f=1}^t \sum_{f'=1}^t w_f w_{f'} [\sigma_{Y(f f')} + \sigma_{L(f f')} + \sigma_{YL^*(f f')} + \frac{1}{n_{ij}} (\sigma_{GY(f f')} + \sigma_{GL(f f')} + \sigma_{GYL^*(f f')})]$$

由于随机变量 $y_{hij(\cdot)}$ 和 $I_{ij(\cdot)}$ 不是相互独立的, 可以采用 Jackknife 方法, 以各区试点作为重复抽样单位, 计算回归参数的估计值及其标准误。然后用 t 测验对回归参数作统计检验, 或者确定回归参数的 $100(1-\alpha)\%$ 置信区间。

朱军、赖鸣冈和许馥华(1993)以黄河流域1989-1990年棉花区域试验的纤维比强度和霜前皮棉产量为例, 比较了单一性状分析与综合性状分析的结果。纤维比强度是以试点为单位分析的数据, 而霜前皮棉产量则是四个区组的平均数。这两个性状的资料缺失相同。在综合分析以前, 首先将原始数据分别除以联合区域试验对照品种(中12)的纤维比强度和霜前皮棉产量的平均值, 再乘以100。把原始数据转换成对照的百分率以后分析。

随机效应的方差和协方差分析结果(表12.2)表明, 影响纤维比强度的随机因素主要是剩余效应, 其次是年份与试点的互作; 霜前皮棉产量则主要受年份与试点的互作影响, 其次是剩余效应。1989年和1990年棉花纤维比强度和霜前皮棉产量在24个试点的产量不一致。两个性状的环境效应(年份、地点、年份×地点)和剩余效应的协方差分量估计值为负值, 这表明环境因素对两个性状的作用可能不一致。纤维比强度的一项方差分量估计值也是负数 ($\hat{\sigma}_{GY}^2 = -0.85$), 可以当作零。

表12.2 纤维比强度(1)与霜前皮棉产量(2)的方差分量和协方差分量估计值

方差分量	纤维比强度	霜前皮棉产量	协方差分量	强度与产量
σ_Y^2	0.43	11.22	$\sigma_{Y(12)}$	-4.70
σ_L^2	2.37	17.20	$\sigma_{L(12)}$	-2.85
σ_{YL}^2*	8.73	297.19	$\sigma_{YL*(12)}$	-0.22
σ_{GY}^2	-0.85	1.49	$\sigma_{GY(12)}$	0.68
σ_{GL}^2	0.63	20.27	$\sigma_{GL(12)}$	6.05
σ_{GYL}^2*	32.11	79.52	$\sigma_{GYL*(12)}$	-3.76

经过数据转换以后, 品种平均数有了改变, 但是平均数间差异的显著性测验结果不会改变。品种纤维比强度平均数间差异的5%显著性测验结果表

明，所有7个参试品种的纤维比强度都显著超过了对照品种(中12)，而中206、平28、冀84-25和鲁155四个品种的强度最高。品种霜前皮棉产量的显著性测验结果表明，没有一个参试品种的产量显著超过对照品种。当综合评价品种的品质和产量性状时，两个品种(中206和平28)显著超过了对照品种，一个品种(冀84-25)显著差于对照品种。综合性状的品种间差异的5%显著性测验结果如下，

冀84-25	运1729	平28	石3409	冀植17	中206	鲁155	中12(CK)
		A			A		
		B		B		B	B
		C	C	C			C
D	D		D				

对(平28和鲁155)以及(冀84-25、石3409和冀植17)与中国农科院棉花研究所育成的(中206和中12)所做的线性对比结果表明(表12.3)，品种平28和鲁155的平均纤维比强度极显著地超过了中棉所育成的品种，但平均霜前皮棉产量却极显著地低于中棉所育成的品种。这两组品种之间在纤维强度和皮棉产量的综合表现上则不存在显著差异。对于冀84-25、石3409和冀植17三个品种，它们的平均纤维比强度不低于中棉所育成品种中206和中12的，但是平均霜前皮棉产量以及强度和产量综合表现方面却极显著地低于中棉所的育成品种。

表12.3 纤维比强度、霜前皮棉产量和综合表现的线性对比所估算的z值(C/SE)

线性对比	强度	产量	强度和产量
(平28、鲁155)			
与(中206、中12)	4.41	-4.20	-0.82
(冀84-25、石3409、冀植17)			
与(中206、中12)	-0.61	-8.48	-6.39

表12.4 棉花8个品种纤维比强度和综合性状对环境指数回归分析的参数估计值

品种	截距			斜率			相关系数		
	估计值	置信区间		估计值	置信区间		估计值	置信区间	
冀84-25	25.4	-7.7	~58.4	0.79	0.48	~1.10	0.57	0.40	~0.74
运1729	21.5	-29.1	~72.0	0.78	0.23	~1.26	0.47	0.26	~0.68
平28	-41.2	-76.5	~-5.8	1.42	1.09	~1.76	0.74	0.61	~0.88
石3409	-2.5	-34.7	~29.7	0.97	0.67	~1.28	0.66	0.49	~0.82
冀植17	18.4	-22.4	~59.3	0.82	0.44	~1.20	0.57	0.31	~0.83
中206	14.5	-27.6	~56.5	0.91	0.51	~1.31	0.61	0.39	~0.82
鲁155	-14.2	-65.7	~37.2	1.16	0.68	~1.64	0.60	0.39	~0.82
中12(CK)	-19.5	-51.9	~13.0	1.13	0.83	~1.43	0.55	0.29	~0.82
综合性状									
冀84-25	-1.2	-24.3	~22.0	0.97	0.73	~1.20	0.83	0.73	~0.93
运1729	-2.5	-22.9	~17.9	1.00	0.80	~1.20	0.86	0.77	~0.95
平28	21.4	-6.9	~49.6	0.84	0.55	~1.11	0.66	0.50	~0.82
石3409	-10.1	-24.8	~4.7	1.06	0.93	~1.22	0.90	0.84	~0.94
冀植17	-7.2	-27.4	~13.0	1.06	0.86	~1.26	0.89	0.83	~0.94
中206	-6.8	-24.0	~10.4	1.14	0.97	~1.30	0.90	0.84	~0.96
鲁155	18.0	-3.8	~39.7	0.82	0.61	~1.03	0.77	0.61	~0.93
中12(CK)	-12.5	-31.2	~6.1	1.12	0.94	~1.30	0.87	0.77	~0.98

参试品种霜前皮棉产量稳定性分析所估算的回归截距和相关系数值与数据未转换时的一致，纤维比强度和综合性状的稳定性分析结果列于表12.4。平28纤维比强度回归斜率显著大于1(置信区间下限 >1)，回归截距显著小于0(置信区间上限 < 0)，因而表现了与霜前皮棉产量的稳定性完全相反的趋势。

其它品种在纤维比强度的稳定性表现上不存在显著的差异。综合性状的稳定性分析表明，所有品种在回归截距和回归斜率两个指标上差异不显著(置信区间相互重叠)，平28的相关系数 ($0.50 \leq r \leq 0.82$)显著小于石3409、冀植17和中206的相关系数($r = 0.83$)。由于纤维比强度和霜前皮棉产量的年份 \times 试点的交互方差偏大(表12.2)，试点的环境指数在各年份表现不一致，因而表12.4中的回归参数置信区间较大，不易测出显著性。

附录 A: 矩阵运算的原理和方法

在现代统计分析方法中常运用矩阵运算,尤其是混合线性模型的统计分析更是建立在矩阵运算的基础上。以矩阵表示线性模型有较大的通用性,也便于编写计算机运算程序。在本附录中将简要地介绍矩阵的基本概念及运算方法。

A1. 矩阵的简单运算

以行列安排元素的矩形数组称为矩阵(matrix)。矩阵常用大写黑体字母 $\mathbf{A} \cdots \mathbf{Z}$ 或其两维元素 $[a_{ij}] \cdots [z_{ij}]$ 表示。矩阵 \mathbf{A} 具有 n 行和 m 列元素,

$$\mathbf{A}_{n \times m} = [a_{ij}] = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1m} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{im} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nm} \end{bmatrix}$$

如果矩阵 \mathbf{A} 的元素 a_{ij} 是常量,则 \mathbf{A} 为常量矩阵;当元素 a_{ij} 是变量时, \mathbf{A} 为变量矩阵。

行数等于列数的矩阵是方阵。如果方阵的对角线元素 d_{ii} 不全为零,而非对角线的元素 $d_{ij} (i \neq j)$ 全为零,则该矩阵为对角矩阵(diagonal matrix)。对角矩阵 \mathbf{D} 可以有以下几种表示方法,

$$\mathbf{D}_{n \times n} = [d_{ij}] = \begin{bmatrix} d_{11} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & d_{ii} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & d_{nn} \end{bmatrix}$$

对角线元素 d_{ii} 全等的对角矩阵, 称为纯量矩阵(scalar matrix)。对角线元素 d_{ii} 全为 1 的纯量矩阵, 称为单位矩阵(identity matrix)。单位矩阵 \mathbf{I} 的几种表示方法是,

$$\mathbf{I}_{n \times n} = [i_{ij}] = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

只有一列元素的数组称为列向量, 或简称向量(vector)。向量常用小写黑体字母 $\mathbf{a} \cdots \mathbf{z}$ 或其一维元素 $[a_i] \cdots [z_i]$ 表示。

$$\mathbf{a}_{n \times 1} = [a_i] = \begin{bmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{bmatrix}$$

矩阵的简单运算包括加减及乘。

$$\text{矩阵相加: } \mathbf{A}_{n \times m} + \mathbf{B}_{n \times m} = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [c_{ij}] = \mathbf{C}_{n \times m}$$

$$\text{矩阵相减: } \mathbf{A}_{n \times m} - \mathbf{B}_{n \times m} = [a_{ij}] - [b_{ij}] = [a_{ij} - b_{ij}] = [c_{ij}] = \mathbf{C}_{n \times m}$$

$$\text{矩阵相乘: } \mathbf{A}_{n \times m} \mathbf{B}_{m \times p} = [a_{ij}][b_{jk}] = \left[\sum_{j=1}^m a_{ij} b_{jk} \right] = [c_{ik}] = \mathbf{C}_{n \times p}$$

在矩阵相乘运算时， $\mathbf{AB} \neq \mathbf{BA}$ 。

一个矩阵的转置(transpose)也是一个矩阵。矩阵 \mathbf{A} 的转置常用 \mathbf{A}^T 表示。

如果 $\mathbf{A} = [a_{ij}]_{n \times m}$, 矩阵 \mathbf{A} 的转置矩阵为 $\mathbf{A}^T = [a_{ij}]^T = [a_{ji}]$ 。当

$$\mathbf{A} = \begin{bmatrix} 6 & 1 \\ 1 & 2 \\ 4 & 4 \end{bmatrix}_{3 \times 2} \text{ 时, 矩阵 } \mathbf{A} \text{ 的转置矩阵为 } \mathbf{A}^T = \begin{bmatrix} 6 & 1 & 4 \\ 1 & 2 & 4 \end{bmatrix}_{2 \times 3}.$$

如果一个方阵等于它的转置 $\mathbf{A} = [a_{ij}] = \mathbf{A}^T = [a_{ji}]$, 则该方阵为对称矩阵(symmetric matrix)。

$$\text{列向量 } \mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{bmatrix}_{n \times 1} \text{ 的转置是一个行向量 } \mathbf{a}^T = [a_1 \quad \cdots \quad a_j \quad \cdots \quad a_n]_{1 \times n}$$

矩阵加减及乘积的转置可按以下公式算得，

$$(\mathbf{A} + \mathbf{B} + \mathbf{C})^T = \mathbf{A}^T + \mathbf{B}^T + \mathbf{C}^T$$

$$(\mathbf{AB})^T = ([a_{ij}] [a_{jk}])^T = \left[\sum_{j=1} a_{ij} b_{jk} \right]^T = \left[\sum_{j=1} a_{ji} b_{kj} \right] = \left[\sum_{j=1} b_{kj} a_{ji} \right] = \mathbf{B}^T \mathbf{A}^T$$

$$(\mathbf{ABC})^T = \mathbf{C}^T \mathbf{B}^T \mathbf{A}^T$$

$$(\mathbf{A}^T)^T = \mathbf{A}$$

矩阵对角线元素 a_{ii} 的和，称为矩阵的迹(trace)，记为 $\text{tr}(\mathbf{A})$ ，即

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii}$$

矩阵的求迹运算法则是

$$\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$$

$$\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$$

$$\text{tr}(\mathbf{AA}^T) = \sum_{i=1}^n \sum_{j=1}^m a_{ij}^2$$

A2. 线性独立性和矩阵的秩

如果矩阵 \mathbf{A} 由 m 个列向量组成 $\mathbf{A}_{n \times m} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_m]$, 当存在非零的常数向量 $\mathbf{c}^T = [c_1 \ c_2 \ \cdots \ c_m] \neq \mathbf{0}$, 使得 $\mathbf{A} \mathbf{c} = \mathbf{0}$, 即 $\sum_{j=1}^m c_j \mathbf{a}_j = \mathbf{0}$, 且 c_j 不全为零, 那么矩阵 \mathbf{A} 的列向量为线性相关(linear dependence)。对于矩阵

$$\mathbf{A}_{3 \times 4} = \begin{bmatrix} 1 & 2 & 5 & 1 \\ 2 & 2 & 10 & 6 \\ 3 & 4 & 15 & 1 \end{bmatrix} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3 \ \mathbf{a}_4]$$

可以找到一个向量 \mathbf{c} ($\mathbf{c}^T = [5 \ 0 \ -1 \ 0]$), 使得

$$\mathbf{A} \mathbf{c} = \begin{bmatrix} 1 & 2 & 5 & 1 \\ 2 & 2 & 10 & 6 \\ 3 & 4 & 15 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 0 \\ -1 \\ 0 \end{bmatrix} = 5 \times \mathbf{a}_1 + 0 \times \mathbf{a}_2 - 1 \times \mathbf{a}_3 + 0 \times \mathbf{a}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0}$$

因此矩阵 \mathbf{A} 的列向量是线性相关的。

如果只存在零向量 $\mathbf{c} = \mathbf{0}$, 使 $\mathbf{A} \mathbf{c} = \mathbf{0}$, 那么矩阵 \mathbf{A} 的列向量为线性独立 (linear independence)。

矩阵相互独立的向量(列或行)数目, 称为矩阵的秩(rank of matrix)。矩阵的秩是不变的。矩阵 \mathbf{A} 的秩以 $\rho(\mathbf{A})$ 表示。矩阵的秩运算法则是,

$$\rho(\mathbf{A}_{n \times m}) \leq \min(n, m)$$

$$\rho(\mathbf{AB}) \leq \min\{\rho(\mathbf{A}), \rho(\mathbf{B})\}$$

$$\rho(\mathbf{A+B}) \leq \rho(\mathbf{A:B}) \leq [\rho(\mathbf{A}) + \rho(\mathbf{B})]$$

如果 $\mathbf{AA} = \mathbf{A}$ ，那么 \mathbf{A} 是幂等矩阵(idempotent matrix)。对于幂等矩阵

\mathbf{A} ，有以下性质

$$tr(\mathbf{A}) = \rho(\mathbf{A})$$

$$\mathbf{A}^k = \mathbf{A}$$

$\mathbf{I} - \mathbf{A}$ 也是幂等矩阵。

例如

$$\mathbf{AA} = \mathbf{A} = \begin{bmatrix} 0.5 & 0.5 & 0.0 \\ 0.5 & 0.5 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\rho(\mathbf{A}) = 2 = tr(\mathbf{A})$$

A3. 矩阵求逆和广义逆矩阵

在矩阵运算中，如果一个矩阵 \mathbf{A}^{-1} 与矩阵 \mathbf{A} 相乘的积是单位矩阵

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{AA}^{-1}$$

那么 \mathbf{A}^{-1} 就称为 \mathbf{A} 的逆矩阵(inverse of matrix)。例如

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{A}^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}, \quad \mathbf{AA}^{-1} = \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

矩阵的其它运算法则是，

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$$

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$$

$$\mathbf{D}(d_{ii})^{-1} = \mathbf{D}\left(\frac{1}{d_{ii}}\right)$$

如果 $\rho(\mathbf{A}) = r = n$, 矩阵 \mathbf{A} 是满秩(full rank)矩阵, 存在 \mathbf{A}^{-1} 。如果

$\rho(\mathbf{A}) = r < n$, \mathbf{A}^{-1} 不存在。

对于任何矩阵, 无论是否满秩, 都可以找到其广义逆(generalized inverse)。如果 $\mathbf{AA}^{-}\mathbf{A} = \mathbf{A}$, 那么 \mathbf{A}^{-} 是矩阵 \mathbf{A} 的一个广义逆。矩阵 \mathbf{A} 的广义逆数目是无穷的。

对于任意矩阵 \mathbf{Z} ,

$\mathbf{A}^{-} + (\mathbf{I} - \mathbf{A}^{-}\mathbf{A})\mathbf{Z}$ 也是矩阵 \mathbf{A} 的一个广义逆。

例如

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{A}^{-} = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{2}{3} \end{bmatrix}, \quad \mathbf{AA}^{-} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

因此 $\mathbf{AA}^{-}\mathbf{A} = \mathbf{A}$ 。

$$\text{如果 } \mathbf{Z} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{A}^{-} + (\mathbf{I} - \mathbf{A}^{-}\mathbf{A})\mathbf{Z} = \begin{bmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ 0 & \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{2} & -\frac{1}{6} & \frac{2}{3} \end{bmatrix}$$

也是矩阵 \mathbf{A} 的一个广义逆。

对于任何矩阵 \mathbf{A} , 由奇异值分解(singular value decomposition)可以得到

$$\mathbf{A} = \mathbf{P}^{-1} \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{Q}^{-1}$$

矩阵 \mathbf{A} 的一个广义逆可以由下式得到

$$\mathbf{A}^{-} = \mathbf{Q} \begin{bmatrix} \mathbf{D}^{-1} & \mathbf{E} \\ \mathbf{B} & \mathbf{C} \end{bmatrix} \mathbf{P}$$

其中 \mathbf{D} 是对角线矩阵, \mathbf{B} , \mathbf{C} , \mathbf{E} 可以是任何值的矩阵。

对于任何矩阵, 都可以找到唯一的一个广义逆, 称为 Mose-Penrose 广义逆。对于矩阵 \mathbf{A} , 其相应的 Mose-Penrose 广义逆以 \mathbf{A}^{+} 表示。该广义逆满足以下全部条件:

$$\mathbf{A}\mathbf{A}^{+}\mathbf{A} = \mathbf{A}, \mathbf{A}^{+}\mathbf{A}\mathbf{A}^{+} = \mathbf{A}^{+}, (\mathbf{A}^{+}\mathbf{A})^T = (\mathbf{A}^{+}\mathbf{A}), (\mathbf{A}\mathbf{A}^{+})^T = (\mathbf{A}\mathbf{A}^{+})。$$

矩阵 \mathbf{A}^{+} 可以由下式求得,

$$\mathbf{A}^{+} = \mathbf{Q} \begin{bmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{P}$$

在线性模型分析中常用到射影矩阵(projection matrix)。对于任意矩阵 \mathbf{X} , 有 $\mathbf{P}_X = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$, \mathbf{P}_X 是 \mathbf{X} 的射影矩阵。射影矩阵有以下特性:

$$\mathbf{P}_X\mathbf{X} = \mathbf{X}, \mathbf{X}^T\mathbf{P}_X = \mathbf{X}^T, (\mathbf{P}_X)^T = \mathbf{P}_X, \mathbf{P}_X\mathbf{P}_X = \mathbf{P}_X。$$

A4. 随机向量和矩阵

如果矩阵或向量中的元素都是常量(constant), 则称为常量矩阵或常量向量。当矩阵或向量中的元素都是变量(random Variable)时, 则是随机矩阵或随机向量。

设 \mathbf{x} 、 \mathbf{y} 为随机向量, \mathbf{a} 、 \mathbf{b} 、 \mathbf{c} 等为常量向量; \mathbf{X} 、 \mathbf{Y} 为随机矩阵, \mathbf{A} 、 \mathbf{B} 、 \mathbf{C} 等为常量矩阵。

对随机矩阵或向量求期望值, 就是对其所有元素求期望值,

$$E(\mathbf{X}) = [E(x_{ij})], E(\mathbf{x}) = [E(x_i)]。$$

常量矩阵或向量的期望值就是原矩阵或向量，

$$E(\mathbf{A}) = \mathbf{A}, \quad E(\mathbf{a}) = \mathbf{a}.$$

以下是对矩阵或向量的函数求期望值，

$$E(\mathbf{X} + \mathbf{Y}) = E(\mathbf{X}) + E(\mathbf{Y})$$

$$E(\mathbf{AX}) = \mathbf{A}E(\mathbf{X})$$

$$E(\mathbf{XB}) = E(\mathbf{X})\mathbf{B}$$

$$E(\mathbf{AX} + \mathbf{Y}) = \mathbf{A}E(\mathbf{X}) + E(\mathbf{Y})$$

$$E(\mathbf{AX} + \mathbf{BY}) = \mathbf{A}E(\mathbf{X}) + \mathbf{B}E(\mathbf{Y})$$

在多变量分析中，常需对随机向量求方差(Variance)或协方差(coVariance)。对随机向量 \mathbf{x} 求方差定义为，

$$\begin{aligned} \text{Var}(\mathbf{x}) &= E\{[\mathbf{x} - E(\mathbf{x})][\mathbf{x} - E(\mathbf{x})]^T\} \\ &= [\sigma_{ij}] \\ &= \mathbf{V} \text{ 或 } \mathbf{V}_x \end{aligned}$$

其中 $\sigma_{ii} = \sigma_i^2 = \text{Var}(x_i)$ 是变量 x_i 的方差， $\sigma_{ij} = \sigma_{ji} = \text{Cov}(x_i, x_j) = \text{Cov}(x_j, x_i)$ 是变量 x_i 和 x_j 的协方差 ($i \neq j$)。

对随机向量 \mathbf{x} 和 \mathbf{y} 求协方差定义为，

$$\begin{aligned} \text{Cov}(\mathbf{x}, \mathbf{y}^T) &= E\{[\mathbf{x} - E(\mathbf{x})][\mathbf{y} - E(\mathbf{y})]^T\} \\ &= [\sigma_{ij(XY)}] \\ &= \mathbf{C} \text{ 或 } \mathbf{V}_{XY} \end{aligned}$$

其中 $\sigma_{ij} = \sigma_{ji} = \text{Cov}(x_i, y_j) = \text{Cov}(x_j, y_i)$ 是变量 x_i 和 y_j 的协方差。

\mathbf{V} 和 \mathbf{C} 都是对称矩阵。常量向量的方差和协方差均为零。

以下是对随机向量的函数求方差或协方差的基本运算法则，

$$\text{Var}(\mathbf{y} + \mathbf{a}) = \text{Var}(\mathbf{y})$$

$$\text{Var}(\mathbf{a}^T \mathbf{y}) = \mathbf{a}^T \mathbf{V} \mathbf{a}$$

$$\text{Var}(\mathbf{A} \mathbf{y}) = \mathbf{A} \mathbf{V} \mathbf{A}^T$$

$$\begin{aligned}\text{Var}(\mathbf{x} + \mathbf{y}) &= \text{Var}(\mathbf{x}) + \text{Var}(\mathbf{y}) + 2\text{Cov}(\mathbf{x}, \mathbf{y}) \\ &= \mathbf{V}_X + \mathbf{V}_Y + 2\mathbf{C}\end{aligned}$$

$$\text{Cov}(\mathbf{a}^T \mathbf{x}, \mathbf{b}^T \mathbf{y}) = \mathbf{a}^T \mathbf{C} \mathbf{b}$$

$$\text{Cov}(\mathbf{A} \mathbf{x}, \mathbf{B} \mathbf{y}) = \mathbf{A} \mathbf{C} \mathbf{B}^T$$

$$\text{Cov}(\mathbf{x}, \mathbf{a} + \mathbf{B} \mathbf{y}) = \mathbf{C} \mathbf{B}^T$$

A5. 二次型的分布

在多元回归分析及方差分析中常用二次型(quadric form)表示平方和。设

$$\mathbf{x} \sim \text{MVN}(\boldsymbol{\mu}_X, \mathbf{V}_X)$$

则

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

是随机变量 \mathbf{x} 的二次型, 其中 \mathbf{A} 是对称常量矩阵。 $\mathbf{x}^T \mathbf{A} \mathbf{y}$ 是随机变量 \mathbf{x} 和 \mathbf{y} 的二次型, 常用来表示协方差分析的叉积和。二次型 $\mathbf{x}^T \mathbf{A} \mathbf{x}$ 的平均数和方差分别为,

$$E(\mathbf{x}^T \mathbf{A} \mathbf{x}) = \text{tr}(\mathbf{A} \mathbf{V}_X) + \boldsymbol{\mu}_X^T \mathbf{A} \boldsymbol{\mu}_X$$

$$E(\mathbf{x}^T \mathbf{A} \mathbf{y}) = \text{tr}(\mathbf{A} \mathbf{C}) + \boldsymbol{\mu}_X^T \mathbf{A} \boldsymbol{\mu}_Y$$

$$\text{Var}(\mathbf{x}^T \mathbf{A} \mathbf{x}) = 2\text{tr}(\mathbf{A} \mathbf{V}_X)^2 + 4 \boldsymbol{\mu}_X^T \mathbf{A} \mathbf{V}_X \mathbf{A} \boldsymbol{\mu}_X$$

$$\text{Cov}(\mathbf{x}, \mathbf{x}^T \mathbf{A} \mathbf{x}) = 2 \mathbf{V}_X \mathbf{A} \boldsymbol{\mu}_X$$

$$\text{Cov}(\mathbf{x}^T \mathbf{A} \mathbf{x}, \mathbf{x}^T \mathbf{B} \mathbf{x}) = 2\text{tr}(\mathbf{A} \mathbf{V}_X \mathbf{B} \mathbf{V}_X) + 4 \boldsymbol{\mu}_X^T \mathbf{A} \mathbf{V}_X \mathbf{B} \boldsymbol{\mu}_X$$

在多元回归分析及方差分析中常需要辨别平方和的独立性、即随机变量的二次型独立性。如果 $\mathbf{A} \mathbf{V}_X \mathbf{B} = \mathbf{0}$, 那么 $\mathbf{x}^T \mathbf{A} \mathbf{x}$ 和 $\mathbf{x}^T \mathbf{B} \mathbf{x}$ 相互独立。

附录 B: 数量遗传的统计分析软件

本书所介绍的混合线性模型分析方法可以无偏地分析各种复杂的遗传模型以及有不规则缺失的非平衡数据,但是这些分析方法需要借助于复杂的矩阵运算。作者采用C语言编写了本书所介绍的统计方法和遗传模型的电脑程序,并编译成可以在DOS环境下运行的16位执行程序和在WINDOW'95环境下运行的32位执行程序。当运用分析软件时,如果直接输入执行文件名,软件将按英文提示,并输出英文提示的结果。如果在输入执行文件名后,再带一空格和英文字母C(如GENVAR1R C),则可按中文(Chinese)提示,并输出中文提示的结果。

浙江农业大学农学系对外提供以下三套软件:(1)双列杂交分析和杂种优势分析;(2)种子性状遗传分析;(3)多年份、多试点品种区域试验的统计分析。每套软件优惠价90元(包括邮费)。需要购买软件者,请直接与浙江农业大学农学系遗传育种教研组联系。

B1. 双列杂交和杂种优势的遗传分析软件

a 建立原始数据文件

在实施遗传分析前,必须首先在程序运行的子目录中建立所需分析的原始数据文件(一般以.TXT作为后缀),文件以文本格式存储。如果需要采用AD模型、ADM模型、ADAA模型,建立的数据文件文本格式应如下所示:

环境	母本	父本	世代	区组	皮棉产量	铃数	铃重	衣分
1	1	1	0	1	28.48	12.1	4.494	36.75
1	1	1	0	2	30.45	7.86	4.397	36.21
1	1	1	0	3	47.28	11.22	4.359	38.5
1	1	6	1	1	73.61	19.3	4.346	39.22
1	1	6	1	2	62.42	13.67	4.498	38.24
1	1	6	1	3	83.55	17.1	4.599	40.42

⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2	9	9	0	1	60.32	8.67	5.38	38.7
2	9	9	0	2	24.14	7.75	4.65	37.21
2	9	9	0	3	44.21	8.67	3.828	36.38
2	10	10	0	1	45.87	12.89	5.056	37.32
2	10	10	0	2	40.63	10.88	4.322	33.51
2	10	10	0	3	33.25	13	4.175	34.33

数据文件应该包括五列试验设计代号以及若干列所需分析的性状数据。试验设计代号的前五列分别是：(1)环境、(2)母本、(3)父本、(4)世代、(5)区组。在世代代号第四列中，亲本的代号为0； F_1 的代号为1； F_2 的代号为2；其余4列中最小的常数为1。最大的常数值在第一列为总环境数，第二、三列为总亲本数，第五列为总区组或重复数，这4列的代号应以顺序编排。

在数据文件中不能有断行。每一列的第一行是描述性的文字，可由中文或英文字母组成，但不能夹带空格。所有的列与列之间都以空格隔开，空格数目不限。性状的列数就是所分析的性状数，不能超过100个。

如果需要采用ADLM模型或ADLAmDm模型分析动物的性连锁效应和母体效应，建立数据文件时应在区组列后增列性别列，其格式如下所示：

环境	母本	父本	世代	区组	性别	14天体重	21天体重	28天体重
1	1	1	0	1	1	7.576	10.269	16.763
1	1	1	0	1	2	7.571	9.764	15.531
1	1	2	1	1	1	8.262	11.578	17.843
1	1	2	1	1	2	9.164	11.732	17.039
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1	7	6	1	1	1	10.94	14.575	22.548
1	7	6	1	1	2	10.713	13.319	17.499
1	7	7	0	1	1	9.008	12.998	20.454

1 7 7 0 1 2 8.72 12.285 16.987

在这个原始数据文件中，环境列和区组列的系数都是 1，表明这个双列杂交遗传试验只在一个环境中实施，并且未设重复(区组数是 1)。在性别列中，系数 1 是同配性别(XX 或 ZZ)的代号，系数 2 是异配性别(XY 或 ZW)的代号。

b 建立遗传模型系数矩阵文件

在运用遗传分析软件以前，首先需要选定所用的遗传模型，然后根据原始数据文件建立该遗传模型的系数矩阵文件。

执行 GENAD 软件可以生成加性-显性(AD)遗传模型的系数矩阵。GENADM 软件则可以生成加性-显性-母体(ADM)遗传模型的系数矩阵。运行 GENADE 软件可以生成加性-显性-上位性(ADAA)遗传模型的系数矩阵。使用 GENSEX 软件可以生成包括性连锁和母体效应的 ADLM 遗传模型的系数矩阵。

在电脑中输入“GENAD C”(或“GENADM C”、或“GENADE C”、或“GENSEX C”)执行命令时，显示器上会出现以下中文提示(如果运行软件不带“C”则显示英文提示)：

中文提示	英文提示	键盘回应
请输入数据文件名:	Input Name of Your Data File:	输入数据文件名
是否需要包括区组效应?是 Y/否 N	Do You Have Block Effects Within Location? Y/N	输入 y 或 n

在运行 GENADM 软件时，将根据二倍体种子模型或三倍体胚乳模型生成加性-显性-母体性遗传模型的系数矩阵。因此在软件运行过程中，显示器上会出现以下提示：

中文提示：“您是否分析三倍体胚乳性状？Y/N”

英文提示：“Do You Analyze Triploid Endosperm? Y/N”

如果所分析的数据是三倍体胚乳性状，则在键盘上输入“y”，否则输入“n”。

系数矩阵建模软件运行完毕后会生成2个文件。一个是以“.MAT"为后缀的文本文件，储存了遗传模型系数矩阵的有关信息；另一个是以“.DAT"为后缀的非文本文件，储存了所分析的性状数据。

c 分析方差分量和遗传效应值

当分析方差分量和遗传效应时，需要首先选定 Jackknife 的抽样方法，以基因型为抽样单位还是以区组为抽样单位。如果遗传试验未设重复，只能以基因型为抽样单位计算方差估计值和遗传效应预测值的标准误。可以运行“GENVARIC”，获得中文的提示和分析结果；或运行“GENVARIC”，获得英文的提示和分析结果。

当采用完全随机区组设计，并具有3—4次重复时，可以考虑采用以区组为抽样单位的 Jackknife 方法，运行“GENVARIC”或“GENVAR1R”，可以获得中文或英文的提示和分析结果。

当运行方差和遗传效应分析软件时，显示器上会出现一些提示，使用者可根据提示输入键盘回应。

中文提示	英文提示	键盘回应
请输入数据文件名：	Input Name of Your Data File:	输入数据文件名
您的亲本类型是自交的(输入1)或异交的(输入0)：	What Kind of Parents Do You Use? Input 1 for Inbred or 0 for Outbred:	输入1或0
选择预测遗传效应的方法！用LUP法或AUP法？L/O	Choosing Prediction Method! Using LUP or AUP? L/O	输入L或O
请输入各项线性对比测验的亲本系数：第一组亲本为1，第二组亲本为-1，其它的亲本为0。	Input Coefficients for Each Parent: 1 for the First Group, -1 for the Second Group, 0 for Others.	输入1或0

其中 LUP 法是线性无偏预测法，AUP 是调整无偏预测法。线性对比一般是用于比较二类亲本的主效应，其中一类亲本的系数可选定为 1，另一类亲本的系数则选为 1，其余不参与比较的亲本系数为 0。如果所比较的二类亲本数目不一致，电脑软件会自动判别调整其相应的系数。

如果以基因型为抽样单位(执行 GENVAR1C C 或 GENVAR1C)，还会出现以下提示：

中文提示: 每次 Jackknife 抽样的数目 = 请输入一个 ≥ 1 的整数，按回车键。

英文提示: Input Sampling Number for the Jackknife Procedure:

如某试验有 n 次重复观察值($n \geq 1$)，则可输入 n ，每次剔除 n 个观察值，便可以达到每次删除一个基因型的目的。

方差和遗传效应的分析结果将自动存入以“.VAR”为后缀的文本文件。现以一个中文输出的结果为例，说明各项分析的排列格式。所分析的原始数据是 10 个亲本部分双列杂交的 2 年试验资料，存放在“cotdata.txt”文件中。运行“GENVAR1R C”后，分析结果自动存放在“cotdata.var”文本文件中。该文件可以直接打印，也可以调入文字编辑软件或数据库软件重新编辑。

结果文件的第一部分列出了一些描述性的信息。电脑输出结果	注解
性状数 4	所分析的性状数目
方差分量数 6	每一个性状的方差分量数目
自由度 5	2 年试验，每年 3 次重复， $df = (2 \times 3) - 1 = 5$
文件名是: cotdata.var	存放分析结果的文件名是“cotdata.var”
Date and Time for Analysis: Thu Apr 18 09:55:13 1996	分析的日期、时间、年份
采用 MINQUE(1)法估算方差分量 <GENVAR1REXE>	使用的分析软件是“GENVAR1R.EXE”

Jackknifing Over Block Conducted for Estimating S.E.	以区组为抽样单位估算标准误
采用 AUP 法预测遗传效应值.	预测遗传效应的方法

结果文件的第二部分列出所采用的统计分析方法的参考文献:

1. 朱军: 1992 估算遗传方差和协方差的混合模型方法. 生物数学学报 7(1): 1-11.
2. 朱军: 1993 作物杂种后代基因型值和杂种优势的预测方法. 生物数学学报 8(1): 32-44.
3. 朱军、季道藩、许馥华: 1993 作物品种间杂种优势遗传分析的新方法. 遗传学报 20(3): 262-271.

结果文件列出以下遗传主效应的线性对比系数, 表明该试验比较了亲本(1, 2, 3, 4)与亲本(7, 8, 9, 10)的主效应:

各项线性对比系数:

$$+<1> +<2> +<3> +<4> -<7> -<8> -<9> -<10>$$

统计分析输出的结果文件, 其主体部分是方差分量的估算结果和遗传效应值的预测结果。这些分析结果以各个性状的先后次序, 安排于5列: (1) 参数名、(2)估计值或预测值、(3)由Jackknife方法计算的标准误、(4)显著性测验的概率值(P-值)、(5)显著性程度(N=不显著, S += 10%显著, S * = 5%显著, S ** = 1%显著)。在实际文本文件中, 每一列之间都是以逗号“,”分隔的。当文件读入《Excel》软件时, 可以转换成列表形式。

在本示例所分析的第一个性状是皮棉产量。

方差分量的分析结果如下:

方差分量	估计值	标准误	P-值	显著性
(1): 加性 Var/2	21.6271	14.0983	0.0928	S +
(2): 显性 Var/1	68.4253	23.7064	0.0172	S *
(3): 加性×环境 Var/2	9.53654	3.16095	0.0148	S *

(4): 显性×环境 Var/1	70.4685	29.3341	0.0307	S *
(6): 机误 Var/1	124.145	19.1582	0.000652	S **
(7): Var(Pheno.)	294.202	20.5885	1.51E-05	S **

各项方差分量与表现型方差的比率的分析结果如下:

方差分量比率	估计值	标准误	P-值	显著性
(1): 加性 Var/Vp	0.073511	0.050279	0.102	NS
(2): 显性 Var/Vp	0.232579	0.084286	0.0199	S *
(3): 加性×环境 Var/Vp	0.032415	0.01116	0.0168	S *
(4): 显性×环境 Var/Vp	0.239524	0.083088	0.0172	S *
(6): 机误 Var/Vp	0.421971	0.051214	0.000215	S **

遗传率的分析结果(狭义遗传率 $h_N^2 = V_A / V_P$, 广义遗传率 $h_B^2 = (V_A + V_D) / V_P$, 互作狭义遗传率 $h_{NE}^2 = V_{AE} / V_P$, 互作广义遗传率 $h_{BE}^2 = (V_{AE} + V_{DE}) / V_P$)如下:

遗传率	估计值	标准误	P-值	显著性
(7): Heritability(N)	0.073511	0.050279	0.102	NS
(8): Heritability(B)	0.30609	0.068448	0.00328	S **
(9): Heritability(NE)	0.032415	0.01116	0.0168	S *
(10): Heritability(BE)	0.271939	0.075752	0.00786	S **

以下是各亲本的加性遗传效应值的分析结果, 线性对比是($\sum_{i=1}^4 A_i$ 与 $\sum_{i=7}^{10} A_i$):

加性 随机效应	预测值	标准误	P-值	显著性
A 1	-3.46097	1.524053	0.0724	S +
A 2	-2.20566	1.129524	0.108	NS
⋮	⋮	⋮	⋮	⋮

A 9	-1.5362	1.468908	0.344	NS
A 10	-4.5971	2.190961	0.09	S +
线性对比:	-3.89625	4.73581	0.448	NS

以下是亲本和杂交组合的显性遗传效应值的分析结果，线性对比是

$$(\Delta_G = \sum_{i=1}^{10} D_{ii} / \sqrt{10\sigma_D^2}):$$

显性 随机效应	预测值	标准误	P-值	显著性
D 11	-15.5891	3.364966	0.00567	S **
D 22	-9.12342	2.495194	0.0146	S *
⋮	⋮	⋮	⋮	⋮
D 1010	-12.4199	2.772691	0.00652	S **
D 16	10.81904	7.082453	0.187	NS
⋮	⋮	⋮	⋮	⋮
D 59	9.621489	5.013774	0.113	NS
D 510	4.878508	3.736286	0.248	NS
线性对比:	4.26043	1.311	0.0227	S *

以下是各亲本在2年环境下的加性×环境交互遗传效应值的分析结果，线性对比是

$$(\sum_{h=1}^2 \sum_{l=1}^4 AE_{hl} \text{ 与 } \sum_{h=1}^2 \sum_{l=7}^{10} AE_{hl}):$$

加性×环境 随机效应	预测值	标准误	P-值	显著性
AE 1 in E1	-1.99174	1.199916	0.158	NS
AE 2 in E1	-4.03367	2.375612	0.15	NS
⋮	⋮	⋮	⋮	⋮
AE 10 in E1	-1.96789	0.724412	0.0419	S *
AE 1 in E2	-0.03108	0.862752	0.973	NS

AE 2 in E2	3.14068	1.972908	0.172	NS
⋮	⋮	⋮	⋮	⋮
AE 10 in E2	-0.93733	0.868184	0.33	NS
线性对比:	-3.25518e-05	1.44E-05	0.073	S +

以下是各亲本和杂交组合在2年环境下的显性×环境交互遗传效应值的分析结

果, 线性对比是($\Delta_{GE} = \sum_{h=1}^2 \sum_{i=1}^{10} DE_{hii} / \sqrt{20\sigma_{DE}^2}$):

显性×环境 随机效应	预测值	标准误	P-值	显著性
DE 11 in E1	-17.1336	4.760151	0.0156	S *
DE 22 in E1	-13.0464	5.184964	0.0534	S +
⋮	⋮	⋮	⋮	⋮
DE 1010 in E1	-20.4783	6.305558	0.0228	S *
DE 16 in E1	3.308065	4.252789	0.472	NS
⋮	⋮	⋮	⋮	⋮
DE 59 in E1	11.66291	8.829389	0.244	NS
DE 510 in E1	8.849879	4.200734	0.089	S +
⋮	⋮	⋮	⋮	⋮
DE 59 in E2	-2.78511	4.126813	0.53	NS
DE 510 in E2	-2.8954	5.252942	0.605	NS
线性对比:	3.04978	1.78762	0.149	NS

各个环境效应的平均数:

Fixed Effect <1>	66.7211			
Fixed Effect <2>	52.3632			

d 分析杂种优势

如果分析方差分量和遗传效应时选用的遗传模型是 AD 模型,可以进一步运行“GENHET1R”或“GENHET1C”预测杂种后代的杂种优势表现。运行杂种优势分析软件时,显示器出现的提示及所需要的键盘回应与分析方差分量和遗传效应时的相似。软件执行完毕后,杂种优势的分析结果将自动存放在以“.PRE”为后缀的文本文件中。

现以 10 个亲本部分双列杂交的 2 年试验分析结果为例,说明分析项目的排列。原始数据存放在“cotdata.txt”文件中。运行“GENHET1R C”后,分析结果自动存放在“cotdata.pre”文本文件中。该文件可以直接打印,也可以调入文字编辑软件或数据库软件重新编辑。

杂种优势的分析结果文件,其前面部分与以上所介绍的方差分量和遗传效应的分析结果相似,均列出了一些描述性的信息、所需的参考文献,另外还包括了方差分量的估算结果。

分析结果的主体部分是杂种优势的分析结果。这些分析结果以各个性状的先后次序,安排于 7 列:(1)主效应或互作效应、(2)预测值、(3)由 Jackknife 方法计算的标准误、(4)显著性测验的概率值(P-值)、(5)显著性程度(N=不显著, S + = 10%显著, S * = 5%显著, S ** = 1%显著)、(6)主效应+互作效应之和、(7)平均数。在实际文本文件中,每一列之间都是以逗号“,”分隔的。当文件读入《Excel》软件时,可以转换成列表形式。

现以第一个性状(皮棉产量)为例,说明杂种优势的分析结果排列格式。当遗传试验在多环境下实施时,杂种优势可以分为互作杂种优势和主效应杂种优势。分析软件将预测每一个组合的这二种优势以及它们之和。

如果分析的是多环境下实施的资料,将首先列出每一个组合在各环境下的互作杂种优势的预测结果。在第一列(互作效应列)中包括了 6 个项目:
(1) Pre(F1) (GE) = F_1 在某环境下的 GE 预测值;
(2) Pre(F2) (GE) = F2 在某环境下的 GE 预测值;
(3) Hpm(F1) (GE) = F_1 在某环境下的群体平均优势互作离差的预测值;
(4) Delta (GE) = 双亲在某环境中的相对互作差异绝对值;
(5)

Hpb(F1) (GE) = F₁在某环境下的群体超亲优势互作离差的预测值；(6) Hpb(F2) (GE) = F₂在某环境下的群体超亲优势互作离差的预测值。第六列(主效应+互作效应列)中也包括了6个相应的项目：(1) Pre(F1) (G+GE) = F₁在某环境下的预测值；(2) Pre(F2) (G+GE) = F₂在某环境下的预测值；(3) Hpm(F1) (G+GE) = F₁在某环境下的群体平均优势预测值；(4) Delta (G+GE) = 双亲在某环境中的相对遗传和互作差异绝对值；(5) Hpb(F1) (G+GE) = F₁在某环境下的群体超亲优势预测值；(6) Hpb(F2) (G+GE) = F₂在某环境下的群体超亲优势预测值。

以下是组合1(Cross<1>), 即 P₁×P₆ (Entry <1> * Entry <6>) 在第1年(in Environ. <1>)中的分析结果:

互作效应	预测值	标准误	P-值	显著性	主效应+互作效应	平均数
Pre(F1) (GE)	1.5213	5.62628	0.798	NS	Pre(F1) (G+GE)	69.6251
Pre(F2) (GE)	-6.76092	3.35076	0.0997	S +	Pre(F2) (G+GE)	49.4333
Hpm(F1) (GE)	0.277955	0.126026	0.0785	S +	Hpm(F1) (G+GE)	0.679488
Delta (GE)	0.213748	0.063186	0.0196	S *	Delta (G+GE)	0.461633
Hpb(F1) (GE)	0.171081	0.147398	0.298	NS	Hpb(F1) (G+GE)	0.448671
Hpb(F2) (GE)	0.032103	0.085974	0.724	NS	Hpb(F2) (G+GE)	0.108927

分析结果也计算了在环境1类似条件下, 超亲优势高于5%的世代数($n \approx 2$):

Generation n for Hpb > .05 in Environ. <1> = 2.274820

环境1中所有的组合杂种优势分析结束后, 接着列出各组合在环境2中的优势表现。

以下是组合1(Cross<1>), 即 P₁×P₆ (Entry <1> * Entry <6>) 在第2年(in Environ. <1>)中的分析结果:

互作效应	预测值	标准误	P-值	显著性	主效应+互作效应	平均数
Pre(F1) (GE)	10.1144	7.65888	0.244	NS	Pre(F1) (G+GE)	78.2182

Pre(F2) (GE)	4.9251	3.53571	0.222	NS	Pre(F2) (G+GE)	61.1193
Hpm(F1) (GE)	0.174243	0.143058	0.278	NS	Hpm(F1) (G+GE)	0.575776
Delta (GE)	0.0329832	0.017271	0.114	NS	Delta (G+GE)	0.280868
Hpb(F1) (GE)	0.157752	0.143561	0.322	NS	Hpb(F1) (G+GE)	0.435342
Hpb(F2) (GE)	0.07063	0.07229	0.373	NS	Hpb(F2) (G+GE)	0.147454

在与第2年类似的环境条件下，该组合的超亲优势高于5%的世代数($n \approx 2$):

Generation n for Hpb > .05 in Environ. <2> = 2.596217

在互作杂种优势的预测结果之后，输出主效应的杂种优势分析结果，也是安排在以逗号分隔的7列中。在第一列(主效应列)中包括了6个项目：(1) Pre(F1) = F_1 在各种环境下的基因型预测值；(2) Pre(F2) = F_2 在各种环境下的基因型预测值；(3) Hpm(F1) = F_1 在各种环境下的群体平均优势预测值；(4) Delta = 双亲在各种环境中的相对遗传差异绝对值；(5) Hpb(F1) = F_1 在各种环境下的群体超亲优势预测值；(6) Hpb(F2) = F_2 在各种环境下的群体超亲优势预测值。第六列(主效应+互作效应列)中也包括了6个相应的项目：(1) Mean of Pre(F1) (G+GE) = F_1 在所分析的特定环境下的平均预测值；(2) Mean of Pre(F2) (G+GE) = F_2 在所分析的特定环境下的平均预测值；(3) Mean of Hpm(F1) (G+GE) = F_1 在所分析的特定环境下的群体平均优势平均预测值；(4) Mean of Delta (G+GE) = 双亲在所分析的特定环境中的相对遗传和互作平均差异绝对值；(5) Mean of Hpb(F1) (G+GE) = F_1 在所分析的特定环境下的群体超亲优势平均预测值；(6) Mean of Hpb(F2) (G+GE) = F_2 在所分析的特定环境下的群体超亲优势平均预测值。

以下是组合1(Cross<1>), 即 $P_1 \times P_6$ (Entry <1> * Entry <6>) 在2年中的平均杂种优势分析结果:

主效应	预测值	标准误	P-值	显著性	主效应+互作效应	平均数
Pre(F1)	68.1038	10.0833	0.435	NS	Mean of Pre(F1)(G+GE)	73.9216

Pre(F2)	56.1942	5.59041	0.575	NS	Mean of Pre(F2) (G+GE)	55.2763
Hpm(F1)	0.401533	0.171161	0.0659	S +	Mean of Hpm(F1) (G+GE)	0.627632
Delta	0.247884	0.098417	0.0533	S +	Mean of Delta (G+GE)	0.37125
Hpb(F1)	0.27759	0.207529	0.239	NS	Mean of Hpb(F1) (G+GE)	0.442006
Hpb(F2)	0.076824	0.124181	0.563	NS	Mean of Hpb(F2) (G+GE)	0.128191

计算了组合1在2年的超亲优势高于5%的世代数($n \approx 2$):

n for Hpb > .05 = 2.206909

e 分析协方差分量及相关系数

分析了双列杂交遗传资料的方差分量和遗传效应值以后,便可以进一步估算成对性状之间的遗传协方差分量及其相关系数。此时,不必再重新运行生成遗传模型系数矩阵的软件。由于计算相关系数需要利用各性状的方差分量估计值,因此协方差分析一定要安排在方差分析或杂种优势分析以后进行。

分析协方差分量前,需要首先选定 Jackknife 的抽样方法。由于协方差分析是在方差分析的基础上进行的,其抽样方法应该与方差分析时选定的抽样方法应该一致。

如果选用以基因型为抽样单位,可以运行“GENCOVIC C”,获得中文的提示和分析结果。或运行“GENCOVIC”,获得英文的提示和分析结果。当采用以区组为抽样单位时,运行“GENCOVIC”或“GENCOVIC”,可以获得中文或英文的提示和分析结果。

当运行协方差和遗传相关分析软件时,显示器上会出现一些与方差分析类似的提示,使用者可根据提示输入键盘回应。

中文提示	英文提示	键盘回应
请输入数据文件名:	Input Name of Your Data File:	输入数据文件名

您的亲本类型是自交的(输入 1)或 异交的(输入 0):	What Kind of Parents Do You Use? Input 1 for Inbred or 0 for Outbred:	输入 1 或 0
---------------------------------	--	----------

如果以基因型为抽样单位(执行 GENCOVIC C 或 GENCOVIC), 还会出现以下提示:

中文提示: 每次 Jackknife 抽样的数目 = 请输入一个 ≥ 1 的整数, 按回车键。

英文提示: Input Sampling Number for the Jackknife Procedure:

如某试验有 n 次重复观察值($n \geq 1$), 则可输入 n , 每次剔除 n 个观察值, 便可以达到每次删除一个基因型的目的。

协方差分量和遗传相关系数的分析结果将自动存入以“.COR”为后缀的文本文件。现以一个中文输出的结果为例, 说明各项分析的排列格式。所分析的原始数据是 10 个亲本部分双列杂交的 2 年试验资料, 存放在“cotdata.txt”文件中。运行“GENCOVIC C”后, 分析结果自动存放在“cotdata.cor”文本文件中。该文件可以直接打印, 也可以调入文字编辑软件或数据库软件重新编辑。

结果文件的第一部分列出了一些描述性的信息。

电脑输出结果	注解
性状数 4	所分析的性状数目
方差分量数 6	每一个性状的方差分量数目
自由度 5	2 年试验, 每年 3 次重复, $df = (2 \times 3) - 1 = 5$
文件名是: cotdata.cor	存放分析结果的文件名是“cotdata.cor”
Date and Time for Analysis: Fri Apr 12 19:01:33 1996	分析的日期、时间、年份
采用 MINQUE(1)法估算方差分量 <GENCOVIC.EXE>.	使用的分析软件是“GENCOVIC.EXE”

结果文件的第二部分列出所采用的统计分析方法的参考文献:

- | |
|---|
| 1. 朱军: 1992 估算遗传方差和协方差的混合模型方法. 生物数学学报 7(1): 1-11. |
|---|

2. 朱军: 1993 作物杂种后代基因型值和杂种优势的预测方法. 生物数学学报 8(1): 32-44.
3. 朱军、季道藩、许馥华: 1993 作物品种间杂种优势遗传分析的新方法. 遗传学报 20(3): 262-271.

协方差分量和遗传相关系数的估算结果以各成对性状的先后次序, 安排于5列: (1)参数名、(2)估计值或预测值、(3)由Jackknife方法计算的标准误、(4)显著性测验的概率值(P-值)、(5)显著性程度(N=不显著, S + = 10%显著, S * = 5%显著, S ** = 1%显著)。在实际文本文件中, 每一列之间都是以逗号“,”分隔的。当文件读入《Excel》软件时, 可以转换成列表形式。

在本示例中, 分析的第一对性状是皮棉产量与铃数。

协方差分量的分析结果如下:

协方差分量	估计值	标准误	P-值	显著性
加性 Cov	4.6005	3.20194	0.21	NS
显性 Cov	3.45533	3.42203	0.359	NS
加性×环境 Cov	-2.14329	3.16374	0.528	NS
显性×环境 Cov	6.11469	5.23638	0.296	NS
机误 Cov	9.6097	3.65591	0.0466	S *
表现型Cov 2	21.6369	2.99862	0.000797	S **
基因型Cov 1	12.0272	3.42249	0.017	S *

遗传相关系数的分析结果如下：

遗传相关分量	估计值	标准误	P-值	显著性
加性 Cor	0.990147	0.210819	0.00535	S **
显性 Cor	0.5275	0.183228	0.0346	S *
加性×环境 Cor	-0.92227	0.258199	0.016	S *
显性×环境 Cor	0.515504	0.241706	0.0861	S +
机误 Cor	0.437561	0.108527	0.01	S *
表现型Cor 2	0.443961	0.057074	0.000562	S **
基因型Cor 1	0.450667	0.122637	0.0144	S *

B2. 种子性状的遗传分析软件

a 建立原始数据文件

在实施种子性状的遗传分析前,必须首先在程序运行的子目录中建立所需分析的原始数据文件(一般以.TXT作为后缀),文件以文本格式存储。采用二倍体种子模型或三倍体胚乳模型,建立的数据文件文本格式都相同,与附录B1所介绍的数据文件相似,一个样板文件“seeddata.txt”的格式如下所示:

环境	母本	父本	基因型	区组	油分含量	蛋白质含量	赖氨酸含量
1	1	1	0	1	37.009	37.816	1.197
1	1	1	0	2	37.708	37.318	1.176
1	1	3	1	1	36.858	37.222	1.228
1	1	3	1	2	36.178	37.835	1.32
1	1	3	2	1	37.359	38.083	1.088
1	1	3	2	2	34.764	38.515	1.102
1	1	3	3	1	34.649	38.99	1.101
1	1	3	3	2	35.242	38.34	1.139
1	1	3	4	1	35.29	38.69	1.183
1	1	3	4	2	33.918	39.04	1.122
1	1	3	5	1	33.622	37.83	1.181
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2	4	5	1	2	34.8	39.043	0.96
2	4	5	2	1	33.775	38.79	0.937
2	4	5	2	2	35.425	38.322	0.947
2	4	5	3	1	34.03	40.07	0.972
2	4	5	3	2	34.925	40.36	0.988
2	4	5	4	1	35.12	39.32	0.918
2	4	5	4	2	37.135	39.62	0.938
2	4	5	5	1	36.855	39.61	0.916

2	4	5	5	2	33.335	39.1	1.021
2	4	5	6	1	31.545	42.4	1.065
2	4	5	6	2	32.12	42.89	1.048
2	5	5	0	1	31.205	44.038	1.016

数据文件应该包括五列试验设计代号以及若干列所需分析的性状数据。试验设计代号的前五列分别是：(1)环境、(2)母本、(3)父本、(4)基因型、(5)区组。在基因型代号的第四列中，亲本的代号为 0； $F_1=(P_i \times P_j)$ 的代号为 1； $F_2=(F_1 \times F_1)$ 的代号为 2； $BC1=(F_1 \times P_i)$ 的代号为 3； $BC2(F_1 \times P_j)$ 的代号为 4； $RBC1(P_i \times F_1)$ 的代号为 5； $RBC2(P_j \times F_1)$ 的代号为 6；其余 4 列中最小的常数为 1。最大的常数值在第一列为总环境数，第二、三列为总亲本数，第五列为总区组或重复数，这 4 列的代号应以顺序编排。

在数据文件中不能有断行。每一列的第一行是描述性的文字，可由中文或英文字母组成，但不能夹带空格。所有的列与列之间都以空格隔开，空格数目不限。性状的列数就是所分析的性状数，不能超过 100 个。

b 建立遗传模型系数矩阵文件

在运用种子性状遗传分析的软件以前，首先需要选定所用的遗传模型，根据原始数据文件建立该遗传模型的系数矩阵文件。

执行“GENDIPLD”(或“GENTRIPL”)软件可以生成二倍体种子(或三倍体胚乳)遗传模型(包括直接遗传效应、细胞质效应、母体遗传效应三套遗传体系)的系数矩阵文件。运行“GENEMBC”软件可以生成二倍体胚和三倍体胚乳的遗传模型(包括胚遗传效应、胚乳遗传效应、细胞质效应、母体遗传效应四套遗传体系)的系数矩阵文件。

在电脑中输入“GENDIPLD C”(或“GENTRIPL C”)执行命令时，显示器上会出现以下中文提示(如果运行软件不带“C”则显示英文提示)：

中文提示	英文提示	键盘回应
请输入数据文件名:	Input Name of Your Data File:	输入数据文件名
是否需要包括区组效应?是 Y/否 N	Do You Have Block Effects Within Location? Y/N	输入 y 或 n

系数矩阵建模软件运行完毕后会自动产生 2 个文件。一个是以“.MAT”为后缀的文本文件，储存了遗传模型系数矩阵的有关信息；另一个是以“.DAT”为后缀的非文本文件，储存了所分析的性状数据。

c 分析方差分量和遗传效应值

当分析方差分量和遗传效应时，需要首先选定 Jackknife 的抽样方法(以基因型为抽样单位还是以区组为抽样单位)。如果种子遗传试验未设区组，只能以基因型为抽样单位计算方差估计值和遗传效应预测值的标准误。可以运行“GENVAR0C C”，获得中文的提示和分析结果。或运行“GENVAR0C”，获得英文的提示和分析结果。

当采用完全随机区组设计，并具有 3—4 次重复时，可以考虑采用以区组为抽样单位的 Jackknife 方法，运行“GENVAR0R C”或“GENVAR0R”，可以获得中文或英文的提示和分析结果。

当运行方差和遗传效应分析软件时，显示器上会出现一些提示，使用者可根据提示输入键盘回应。

中文提示	英文提示	键盘回应
请输入数据文件名:	Input Name of Your Data File:	输入数据文件名
选择预测遗传效应的方法!用 LUP 法或 AUP 法?L/O	Chosing Prediction Method! Using LUP or AUP? L/O	输入 L 或 O
请输入各项线性对比测验的亲本系数: 第一组亲本为 1, 第二组亲本为-1, 其它的亲本为 0.	Input Coefficients for Each Parent: 1 for the First Group, -1 for the Second Group, 0 for Others.	输入 1 或 0

其中 LUP 法是线性无偏预测法，AUP 是调整无偏预测法。线性对比一般是用于比较二类亲本的主效应，其中一类亲本的系数可选定为 1，另一类亲本的系数则选为-1，其余不参与比较的亲本系数为 0。如果所比较的二类亲本数目不一致，电脑软件会自动判别调整其相应的系数。

如果以基因型为抽样单位(执行 GENVAR1C C 或 GENVAR1C)，还会出现以下提示：

中文提示: 每次 Jackknife 抽样的数目 = 请输入一个 ≥ 1 的整数，按回车键。
英文提示: Input Sampling Number for the Jackknife Procedure:

如某试验有 n 次重复观察值($n \geq 1$)，则可输入 n ，每次剔除 n 个观察值，便可以达到每次删除一个基因型的目的。

方差和遗传效应的分析结果将自动存入以“.VAR”为后缀的文本文件。现以一个中文输出的结果为例，说明各项分析的排列格式。所分析的原始数据是 5 个亲本部分双列杂交的 2 年试验资料，存放在“seeddata.txt”文件中。运行“GENVAR0R C”后，分析结果自动存放在“seeddata.var”文本文件中。该文件可以直接打印，也可以调入文字编辑软件或数据库软件重新编辑。

结果文件的第一部分列出了一些描述性的信息。

电脑输出结果	注解
性状数 4	所分析的性状数目
方差分量数 16	每一个性状的方差分量数目
自由度 3	2 年试验，每年 2 次重复， $df = (2 \times 2) - 1 = 3$
文件名是: seeddata.var	存放分析结果的文件名是“seeddata.var”
分析的日期和时间: Fri Apr 12 19:46:19 1996	分析的日期、时间、年份
采用 MINQUE(0/1)法估算方差分量 <GENVAR0R.EXE>.	使用的分析软件是“GENVAR0R.EXE”
以区组内的所有基因型作为	以区组为抽样单位估算标准误

Jackknifing 抽样单位估算标准误 S.E.	
采用 AUP 法预测遗传效应值.	预测遗传效应的方法

结果文件的第二部分列出所采用的统计分析方法的参考文献:

1. 朱军:1992 估算遗传方差和协方差的混合模型方法. 生物数学学报7(1): 1-11.
2. Zhu J. and Weir B. S. 1994. Analysis of cytoplasmic and maternal effects: I. a genetic model for diploid plant seeds and animals. Theoretical and Applied Genetics, 89:153-159.
3. Zhu J. and Weir B. S. 1994. Analysis of cytoplasmic and maternal effects: II. genetic models for triploid endosperms. Theoretical and Applied Genetics, 89:160-166.
4. 朱军:1994. 广义遗传模型与数量遗传分析新方法. 浙农大学报, 1994, 20(6): 551-559

结果文件列出以下遗传主效应的线性对比系数, 表明该试验比较了亲本(1, 2, 3, 4)与亲本(5)的主效应:

各项线性对比系数:

$$+ \langle 1 \rangle + \langle 2 \rangle + \langle 3 \rangle + \langle 4 \rangle - \langle 5 \rangle$$

方差和协方差分量的估算结果和遗传效应值的预测结果以各个性状的先后次序, 安排于 6 列: (1)参数名、(2)估计值或预测值、(3)由 Jackknife 方法计算的标准误、(4)显著性测验的概率值(P-值)、(5)试验方差分量转换成遗传方差分量是的系数、(6)显著性程度(N=不显著, S + = 10%显著, S * = 5%显著, S ** = 1%显著)。在实际文本文件中, 每一列之间都是以逗号“,”分隔的。当文件读入《Excel》软件时, 可以转换成列表形式。

在本示例中, 分析的第二个性状是蛋白质含量。

方差分量的分析结果如下:

方差和协方差分量	估计值	标准误	P-值	转换系数	显著性
直接加性方差	0	0	0.5	/2	NS
直接显性方差	0.359292	0.110393	0.0237	/0.375	S *
细胞质方差	0.464484	0.293978	0.106	/1	NS

母体加性方差	0.501025	0.193809	0.0407	/2	S *
母体显性方差	0.194611	0.145471	0.137	/1	NS
直接加性×环境方差	1.8995	0.726423	0.0397	/2	S *
直接显性×环境方差	0.014127	0.014127	0.196	/0.375	NS
细胞质×环境方差	0.322919	0.254568	0.147	/1	NS
母体加性×环境方差	0	0	0.5	/2	NS
母体显性×环境方差	0	0	0.5	/1	NS
直接与母体加性协方差	0	7.40113	0.5	/2	NS
直接与母体显性协方差	0.235159	0.126107	0.0795	/0.5	S +
直接与母体加性×环境协方差	0	0	0.5	/2	NS
直接与母体显性×环境协方差	0	0.708284	0.5	/0.5	NS
机误方差	0.37263	0.138993	0.0375	/1	S *
Var(Phenotype)	4.59891	0.566978	0.00196	S **	

遗传率的分析结果(直接遗传率 $h_O^2 = (V_A + C_{A/Am}) / V_P$, 母体遗传率

$h_M^2 = (V_{Am} + C_{A/Am}) / V_P$, 细胞质遗传率 $h_C^2 = V_C / V_P$; 直接环境互作遗传率

$h_{OE}^2 = (V_{AE} + C_{AE/AmE}) / V_P$, 母体环境互作遗传率 $h_{ME}^2 = (V_{AmE} + C_{AE/AmE}) / V_P$,

细胞质环境互作遗传率 $h_{CE}^2 = V_{CE} / V_P$)如下:

遗传率	估计值	标准误	P-值	
直接遗传率	0	0	0.5	NS
母体遗传率	0.108944	0	0.5	NS
细胞质遗传率	0.100999	0	0.5	NS
直接环境互作遗传率	0.413033	0.186071	0.0565	S +
母体环境互作遗传率	0	0	0.5	NS
细胞质环境互作遗传率	0.070217	0.04254	0.0987	S +

由于直接加性方差个分量的估计值是零，故未列出直接加性遗传效应值的预测

值。以下是直接显性遗传效应值的分析结果，线性对比是($\Delta_O = \sum_{i=1}^5 D_{ii} / \sqrt{5\sigma_D^2}$):

直接显性效应	预测值	标准误	P-值	显著性
D 11	0.547432	0.18278	0.0579	S +
D 22	2.180858	0.744069	0.061	S +
⋮	⋮	⋮	⋮	⋮
D 55	0.480126	0.209982	0.106	NS
D 13	-0.74064	0.298604	0.0892	S +
⋮	⋮	⋮	⋮	⋮
D 35	0.436091	0.365388	0.318	NS
D 45	0.161175	0.260651	0.58	NS
Heterosis <Delta>	-2.46371	1.17553	0.127	NS

细胞质和母体显性的方差分量估计值大于零，分析软件列出了细胞质效应的母体显性效应的预测值。但是由于这2项方差分量都不显著，因此可不讨论其遗传效应的表现。母体加性方差的估计值大于零，并达到显著水准，但是所有的母体加性效应预测值都不显著。这是因为方差分量的显著性测验采用一尾t检验，而遗传效应采用二尾t检验，因此遗传效应的检验不易测出显著性。尽管各项效应值不显著，但是母体加性效应的线性对比($\sum_{i=1}^4 Am_i$ 对 Am_5)达到极显著水准，表明亲本1-4的平均母体加性效应小于亲本5的母体加性效应。

母体加性效应	预测值	标准误	P-值	显著性
Am 1	0.099494	0.625297	0.884	NS
Am 2	-0.44404	0.210127	0.125	NS

母体加性效应	预测值	标准误	P-值	显著性
Am 3	0.145227	0.343691	0.701	NS
Am 4	0.002339	0.150455	0.989	NS
Am 5	0.196598	0.848744	0.832	NS
线性对比Cont	-0.87866	0.039181	0.000194	S **

直接加性×环境互作的方差分量显著大于零，所以可以进一步讨论其遗传效应值

的分析结果，线性对比是($\sum_{h=1}^2 \sum_{i=1}^4 AE_{hi}$ 与 $\sum_{h=1}^2 A_{h5}$):

直接加性×环境机效应	预测值	标准误	P-值	显著性
AE 1 in E1	-2.25038	0.847936	0.0767	S +
AE 2 in E1	0.842852	0.295491	0.065	S +
AE 3 in E1	-0.92438	0.298704	0.0535	S +
AE 4 in E1	-0.31552	0.119913	0.0782	S +
AE 5 in E1	2.647465	0.935531	0.0662	S +
AE 1 in E2	1.132239	0.541647	0.128	NS
AE 2 in E2	-0.24435	0.07294	0.0441	S *
AE 3 in E2	0.270377	0.149802	0.169	NS
AE 4 in E2	-0.03184	0.019724	0.205	NS
AE 5 in E2	-1.12675	0.407051	0.0697	S +
Cont	-2.46734	1.41769	0.180164	NS

其它遗传效应预测值都不显著，因此可不必列表讨论。

d 分析杂种优势

运行“GENHET0R”或“GENHET0C”可以预测种子性状各项效应的杂种优势表现。运行杂种优势分析软件时，显示器出现的提示及所需要的键

盘回应，与分析方差分量和遗传效应时的相似。软件执行完毕后，杂种优势的分析结果将自动存放在以“.PRE"为后缀的文本文件中。

现以5个亲本部分双列杂交的2年试验分析结果为例，说明分析项目的排列。原始数据存放在“seeddata.txt”文件中。运行“GENHET1RC”后，分析结果自动存放在“seeddata.pre”文本文件中。该文件可以直接打印，也可以调入文字编辑软件或数据库软件重新编辑。

杂种优势的分析结果文件前面部分与以上介绍的方差分量和遗传效应分析结果相似，均列出了一些描述性的信息、所需的参考文献，另外还包括了方差分量的估算结果。

运行种子性状遗传效应和杂种优势分析，其结果以各个性状及各杂交组合的先后次序，安排于50列。在实际文本文件中，每一列之间都是以逗号“,”分隔的。当文件读入《Excel》软件时，可以转换成列表形式。

现以棉子仁蛋白质含量的分析结果为例，说明各列的内容。如果遗传试验在多环境下实施，分析软件将先列出遗传主效应的分析结果，然后是互作效应的结果。对于主效应和互作效应的分析结果，前2列分别是：(1)杂交组合的序号、(2)杂交组合的双亲代号。随后是各项效应的分析结果，以4列为一组，分别为：(1)遗传效应或杂种优势的预测值、(2)由Jackknife方法计算的标准误、(3)显著性测验的概率值(P-值)、(4)显著性程度(N=不显著，S+ = 10%显著，S* = 5%显著，S** = 1%显著)。

在所列出的各项遗传效应预测结果中，第3列是总遗传效应的预测值， $G(\text{total}) = G_o + C + G_m$ ；第7列是直接遗传效应的预测值， $G(o) = G_o = A + D$ ；第11列是母体遗传效应的预测值， $G(m) = G_m = A_m + D_m$ ；第15列是细胞质效应的预测值， $G(c) = C$ ，在本示例中均不显著，故未列出。

1组合序号	2组合的双亲	3预测值	4标准误	5P-值	6显著性
No.	Cross	G(total)	S.E.	P-value	Signif.
Cross<1>	Entry<1*3>	-2.68247	0.227101	0.0013	S**
Cross<2>	Entry<1*4>	15.0901	0.480678	7.10E-05	S**

Cross<3>	Entry<2*3>	22.4081	0.521269	2.77E-05	S**
Cross<4>	Entry<2*4>	48.9093	0.968425	1.71E-05	S**
Cross<5>	Entry<3*5>	32.8641	0.345434	2.56E-06	S**
Cross<6>	Entry<4*5>	-5.56417	0.704548	0.00423	S**

7 预测值	8 标准误	9 P-值	10 显著性	11 预测值	12 标准误	13 P-值	14 显著性
G(o)	S.E.	P-value	Signif.	G(m)	S.E.	P-value	Signif.
0.0375377	0.0576178	0.561	NS	-2.55192	0.151555	0.000456	S**
-0.337213	0.133989	0.0864	S+	15.5954	0.450401	5.30E-05	S**
-0.449859	0.177461	0.0851	S+	21.9376	0.238677	2.84E-06	S**
0.210773	0.0638692	0.0457	S*	47.7781	0.461654	1.99E-06	S**
0.609075	0.223069	0.0719	S+	32.9076	0.23419	7.95E-07	S**
0.475617	0.180932	0.0784	S+	-5.62821	0.597849	0.00254	S**

分析结果接着列出F₂种子各项平均杂种优势的预测结果。第19列是总平均优势的预测值， $Hm(\text{total}) = Hm(o) + Hm(c) + Hm(m)$ ；第23列是直接平均优势的预测值， $Hm(o) = Go(F_2) - Go(MP)$ ；第27列是母体平均优势的预测值， $Hm(m) = Gm(F_2) - Gm(MP)$ ；第31列是细胞质平均优势的预测值， $Hm(c) = C(F_2) - C(MP)$ 。

19 预测值	20 标准误	21 P-值	22 显著性	23 预测值	24 标准误	25 P-值	26 显著性
Hm(total)	S.E.	P-value	Signif.	Hm(o)	S.E.	P-value	Signif.
0.0715796	0.0112699	0.0079	S**	-0.0199596	0.00651556	0.0548	S+
0.583347	0.0162887	4.79E-05	S**	-0.0297769	0.00882318	0.0433	S*
1.22929	0.0124105	2.27E-06	S**	-0.0534092	0.0171956	0.053	S+
1.96494	0.0160484	1.20E-06	S**	-0.0366696	0.0117641	0.0526	S+
1.58906	0.0165484	2.49E-06	S**	-0.0044369	0.00519674	0.456	NS
0.659318	0.0213649	7.48E-05	S**	-0.0080652	0.00422055	0.152	NS

27 预测值	28 标准误	29 P-值	30 显著性	31 预测值	32 标准误	33 P-值	34 显著性
Hm(m)	S.E.	P-value	Signif.	Hm(c)	S.E.	P-value	Signif.
0.0853254	0.00375842	0.000187	S**	0.0062138	0.00560354	0.348	NS
0.610002	0.00891096	6.87E-06	S**	0.00312258	0.0061189	0.645	NS
1.26252	0.00636747	2.83E-07	S**	0.0201734	0.00904547	0.112	NS
1.98453	0.0145554	8.70E-07	S**	0.0170822	0.00691758	0.0901	S+
1.60587	0.0106279	6.39E-07	S**	-0.0123694	0.00827829	0.232	NS
0.676661	0.0147393	2.28E-05	S**	-0.0092782	0.00745558	0.302	NS

分析结果还列出 F_2 种子各项超母体亲本杂种优势的预测结果。第35列是总超亲优势的预测值, $Hf(\text{total}) = Hf(o) + Hf(c) + Hf(m)$; 第39列是直接超亲优势的预测值, $Hf(o) = G_o(F_2) - G_o(P_i)$; 第43列是母体超亲优势的预测值, $Hm(m) = G_m(F_2) - G_m(P_i)$; 第47列是种子杂种优势超过母体亲本的世代数, $F_n > G(P_i)$ 。

35 预测值	36 标准误	37 P-值	38 显著性	39 预测值	40 标准误	41 P-值	42 显著性
Hf(total)	S.E.	P-value	Signif.	Hf(o)	S.E.	P-value	Signif.
-0.301605	0.00809174	4.25E-05	S**	-0.0130784	0.00606681	0.12	NS
0.154249	0.0136497	0.00149	S**	-0.0226905	0.00787359	0.0634	S+
1.37025	0.0199858	6.84E-06	S**	-0.0674761	0.0235055	0.064	S+
2.04999	0.0188357	1.71E-06	S**	-0.0505313	0.0177591	0.0654	S+
1.35651	0.00191024	6.11E-09	S**	-0.0121813	0.00389528	0.0522	S+
0.482678	0.0157307	7.60E-05	S**	-0.0160148	0.00438548	0.0354	S*

43 预测值	44 标准误	45 P-值	46 显著性	47 预测值	48标准误	49 P-值	50 显著性
Hf(m)	S.E.	P-value	Signif.	Num of Gen.	S.E.	P-value	Signif.
-0.288527	0.00412647	6.45E-06	S**	-0.673339	0	2	NS
0.176939	0.00732419	0.000155	S**	2.28572	0	2	NS
1.43773	0.00630073	1.86E-07	S**	0	0	2	NS
2.10052	0.0170427	1.18E-06	S**	0	0	2	NS
1.36869	0.00334911	3.23E-08	S**	4.44095	0	2	NS
0.498693	0.0123026	3.30E-05	S**	3.50285	0.24322	0.000726	S**

互作遗传效应分析和互作杂种优势的分析，其结果排列格式与主效应的相似。各列都对应于其主效应的列，行的安排是按环境或年份列出各杂交组合的预测结果。以下是种子蛋白质含量的互作效应分析的部分结果。

1组合序号	2组合亲本	3预测值	4标准误	5P-值	6显著性
Env.	Cross	GE(total)	S.E.	P-value	Signif.
Env.<1>	Entry<1*3>	-4.28479	1.20092	0.0376	S*
Env.<1>	Entry<1*4>	-3.71671	0.99816	0.0337	S*
Env.<1>	Entry<2*3>	0.610528	0.534332	0.336	NS
Env.<1>	Entry<2*4>	1.182	0.325105	0.0358	S*
Env.<1>	Entry<3*5>	2.40498	1.04898	0.106	NS
Env.<1>	Entry<4*5>	1.88274	0.932676	0.137	NS
Env.<2>	Entry<1*3>	1.96818	0.518334	0.0321	S*
Env.<2>	Entry<1*4>	1.89416	0.447302	0.0241	S*
Env.<2>	Entry<2*3>	0.146679	0.277535	0.634	NS
Env.<2>	Entry<2*4>	-0.400455	0.203393	0.144	NS
Env.<2>	Entry<3*5>	-0.697402	0.398223	0.178	NS
Env.<2>	Entry<4*5>	-0.444829	0.341695	0.284	NS

7 预测值	8 标准误	9 P-值	10显著性	15 预测值	16 标准误	17 P-值	18 显著性
GE(o)	S.E.	P-value	Signif.	GE(c)	S.E.	P-value	Signif.
-2.97814	1.18131	0.0861	S+	-1.30664	0.158028	0.00371	S**
-2.41007	0.986531	0.0923	S+	-1.30664	0.158028	0.00371	S**
-0.00961975	0.109119	0.935	NS	0.620148	0.448824	0.261	NS
0.561848	0.226501	0.0892	S+	0.620148	0.448824	0.261	NS
1.56899	0.649211	0.0944	S+	0.83599	0.486179	0.184	NS
2.0778	0.848092	0.0917	S+	-0.195056	0.168662	0.331	NS
1.14798	0.414162	0.0695	S+	0.820199	0.345008	0.0979	S+
1.07396	0.430693	0.0882	S+	0.820199	0.345008	0.0979	S+
0.143073	0.141595	0.387	NS	0.00360589	0.140366	0.981	NS
-0.404061	0.149468	0.0736	S+	0.00360589	0.140366	0.981	NS
-0.928936	0.400147	0.103	NS	0.231534	0.0582384	0.0285	S*
-1.09263	0.41383	0.0776	S+	0.647798	0.25377	0.0837	S+

19 预测值	20 标准误	21 P-值	22显著性	23 预测值	24 标准误	25 P-值	26 显著性
HmE(total)	S.E.	P-value	Signif.	HmE(o)	S.E.	P-value	Signif.
-0.0236444	0.00760691	0.053	S+	0.00383409	0.00161814	0.0986	S+
-0.0138134	0.00385147	0.0371	S*	0.00044232	0.000953488	0.674	NS
0.00253297	0.0114844	0.84	NS	0.00530107	0.00144844	0.0352	S*
0.012451	0.00749186	0.195	NS	0.00199633	0.00036949	0.0124	S*
0.0100646	0.00698887	0.245	NS	-7.31E-05	0.0025947	0.979	NS
-0.0080701	0.00625206	0.287	NS	-0.00498504	0.00517753	0.407	NS
0.0110036	0.00457453	0.0954	S+	0.00345418	0.00156049	0.114	NS
0.0147041	0.00511948	0.0639	S+	0.0124931	0.00575505	0.118	NS
0.0140162	0.00828375	0.189	NS	0.0169393	0.00787799	0.121	NS
0.00558154	0.00462745	0.314	NS	0.0138431	0.00613663	0.109	NS
0.0303783	0.00488105	0.00836	S**	0.00556949	0.00269687	0.131	NS
0.0424555	0.0103998	0.0266	S*	0.0123083	0.00611515	0.138	NS

31 预测值	32 标准误	33 P-值	34显著性	35 预测值	36 标准误	37 P-值	38显著性
HmE(c)	S.E.	P-value	Signif.	HfE(total)	S.E.	P-value	Signif.
-0.0274785	0.00763085	0.0367	S *	0.0368884	0.0151163	0.0925	S +
-0.0142557	0.00410701	0.0403	S *	0.0514591	0.0198579	0.081	S +
-0.0027681	0.0119467	0.832	NS	-0.0363182	0.0157055	0.104	NS
0.0104547	0.00758079	0.262	NS	-0.0216605	0.0100145	0.119	NS
0.0101377	0.00734431	0.261	NS	0.0874107	0.0349264	0.0875	S +
-0.00308507	0.0032182	0.408	NS	0.0645362	0.0262955	0.0913	S +
0.00754943	0.00513833	0.238	NS	-0.0207948	0.0100411	0.13	NS
0.00221099	0.00733549	0.783	NS	-0.0226933	0.0088217	0.0823	S +
-0.0029231	0.0024919	0.325	NS	0.0319505	0.0119852	0.076	S +
-0.00826154	0.00413758	0.14	NS	0.0179169	0.0070408	0.0843	S +
0.0248088	0.00307068	0.00396	S **	-0.0255682	0.0105636	0.0941	S +
0.0301472	0.00459268	0.00719	S **	-0.00789199	0.0065786	0.316	NS

39 预测值	40 标准误	41 P-值	42 显著性
HfE(o)	S.E.	P-value	Signif.
0.0368884	0.0151163	0.0925	S +
0.0514591	0.0198579	0.081	S +
-0.0363182	0.0157055	0.104	NS
-0.0216605	0.0100145	0.119	NS
0.0874107	0.0349264	0.0875	S +
0.0645362	0.0262955	0.0913	S +
-0.0207948	0.0100411	0.13	NS
-0.0226933	0.0088217	0.0823	S +
0.0319505	0.0119852	0.076	S +
0.0179169	0.00704081	0.0843	S +
-0.0255682	0.0105636	0.0941	S +
-0.00789199	0.00657863	0.316	NS

e 分析协方差分量及相关系数

分析了种子性状遗传资料的方差分量和遗传效应值以后,还可进一步估算成对性状之间的遗传协方差分量及其相关系数。此时,不必再重新产生遗传模型的系数矩阵。由于计算相关系数需要利用各性状的方差分量估计值,因此协方差分析一定要安排在方差分析或杂种优势分析以后进行。

分析协方差分量前,需要首先选定 Jackknife 的抽样方法。由于协方差分析是在方差分析的基础上进行的,其抽样方法应该与方差分析时选定的抽样方法一致。

如果选用以基因型为抽样单位,可以运行“GENCOV0C C”,获得中文的提示和分析结果。或运行“GENCOV0C”,获得英文的提示和分析结果。当采用以区组为抽样单位时,运行“GENCOV0R C”或“GENCOV0R”,可以获得中文或英文的提示和分析结果。

当运行协方差和遗传相关分析软件时,显示器上会出现一些与方差分析类似的提示,使用者可根据提示输入键盘回应。

中文提示	英文提示	键盘回应
请输入数据文件名:	Input Name of Your Data File:	输入数据文件名

如果以区组为抽样单位(执行 GENCOV0C C 或 GENCOV0C),还会出现以下提示:

中文提示: 每次 Jackknife 抽样的数目 = 请输入一个 ≥ 1 的整数, 按回车键。 英文提示: Input Sampling Number for the Jackknife Procedure:
--

如某试验有 n 次重复观察值($n \geq 1$), 则可输入 n , 每次剔除 n 个观察值, 便可以达到每次删除一个基因型的目的。

协方差分量和遗传相关系数的分析结果将自动存入以“.COR”为后缀的文本文件。现以一个中文输出的结果为例, 说明各项分析的排列格式。所分析的原始数据是 5 个亲本部分双列杂交的 2 年试验资料, 存放在

“seeddata.txt”文件中。运行“GENVAR0R C”后，分析结果自动存放在“seeddata.cor”文本文件中。该文件可以直接打印，也可以调入文字编辑软件或数据库软件重新编辑。

结果文件的第一部分列出了一些描述性的信息。

电脑输出结果	注解
性状数 4	所分析的性状数目
方差分量数 16	每一个性状的方差分量数目
自由度 3	2年试验，每年2次重复， $df = (2 \times 2) - 1 = 3$
文件名是: seeddata.cor	存放分析结果的文件名是“seeddata.cor”
分析的日期和时间: Fri Apr 12 19:46:19 1996	分析的日期、时间、年份
采用 MINQUE(0/1)法估算方差分量 <GENCOV0R.EXE>.	使用的分析软件是“GENCOV0R.EXE”
以区组内的所有基因型作为 Jackknifing 抽样单位估算标准误 S.E.	以区组为抽样单位估算标准误

结果文件的第二部分列出所采用的统计分析方法的参考文献:

1. 朱军: 1992 估算遗传方差和协方差的混合模型方法. 生物数学学报 7(1): 1-11.
2. Zhu J. and Weir B. S. 1994. Analysis of cytoplasmic and maternal effects: I. a genetic model for diploid plant seeds and animals. Theoretical and Applied Genetics, 89:153-159.
3. Zhu J. and Weir B. S. 1994. Analysis of cytoplasmic and maternal effects: II. genetic models for triploid endosperms. Theoretical and Applied Genetics, 89:160-166.
4. 朱军: 1994. 广义遗传模型与数量遗传分析新方法. 浙农大学学报, 1994, 20(6): 551-559

协方差分量和遗传相关系数的估算结果以各成对性状的先后次序，安排于5列：(1)参数名、(2)估计值或预测值、(3)由Jackknife方法计算的标准误、(4)显著性测验的概率值(P-值)、(5)显著性程度(N=不显著，S + = 10%显著，

S * = 5% 显著, S ** = 1% 显著)。在实际文本文件中, 每一列之间都是以逗号 “,” 分隔的。当文件读入《Excel》软件时, 可以转换成列表形式。

现以 蛋白质含量 和 赖氨酸含量的相关性分析为例, 说明各项结果的排列格式。协方差分量的分析结果如下:

协方差分量	估计值	标准误(S.E.)	P-值	显著性
直接加性 Cov	-0.1843	0.215691	0.456	NS
直接显性 Cov	0.066486	0.063008	0.369	NS
细胞质 Cov	-0.0233	0.033297	0.534	NS
母体加性 Cov	-0.02944	0.048731	0.588	NS
母体显性 Cov	-26.2275	26.2268	0.391	NS
直接加性×环境 Cov	-0.02249	0.056294	0.716	NS
直接显性×环境 Cov	-0.01261	0.006875	0.164	NS
细胞质×环境 Cov	-0.0699	0.08233	0.458	NS
母体加性×环境 Cov	-0.031	0.010384	0.0583	S +
母体显性×环境 Cov	-0.02977	0.024479	0.311	NS
直接与母体加性 Cov	-0.11872	0.084014	0.253	NS
直接与母体显性 Cov	-0.01246	0.018956	0.558	NS
直接与母体加性×环境 Cov	0.076812	0.028808	0.0759	S +
直接与母体显性×环境 Cov	0.012806	0.015716	0.475	NS
机误 Cov	-0.75562	0.753783	0.39	NS
表现型Cov 2	-27.4026	27.3984	0.391	NS
基因型Cov 1	-26.647	28.1522	0.414	NS

遗传相关系数的分析结果如下：

遗传相关分量	估计值	标准误(S.E.)	P-值	显著性
直接加性 Cor	0	0	1	NS
直接显性 Cor	0	0	1	NS
细胞质 Cor	-0.30894	0.227676	0.268	NS
母体加性 Cor	0	0	1	NS
母体显性 Cor	-1	0.335526	0.0586	S +
直接加性×环境 Cor	-0.28697	0.408248	0.533	NS
直接显性×环境 Cor	-1	0.058874	0.000444	S **
细胞质×环境 Cor	-1	0.288675	0.0405	S *
母体加性×环境 Cor	0	0	1	NS
母体显性×环境 Cor	0	0	1	NS
机误 Cor	0	0.089705	1	NS
表现型Cor 2	0	0.064576	1	NS
基因型Cor 1	-1	0.122539	0.00385	S **

B3. 发育数量性状的遗传分析软件

a 建立原始数据文件

分析数量性状的发育遗传资料前,必须首先在程序运行的子目录中建立所需分析的原始数据文件(一般以“.TXT”作为后缀),文件以文本格式存储。数据文件的格式与 B1 和 B2 中介绍的类似,所不同的是分析的性状一般都按时间序列排列。一个小鼠 7×7 双列杂交的数据文件“micedata.txt”的格式如下所示:

环境	母本	父本	基因型	区组	性别	14天体重	21天体重	28天体重
1	1	1	0	1	1	7.576	10.269	16.763
1	1	1	0	1	2	7.571	9.764	15.531
1	1	2	1	1	1	8.262	11.578	17.843
1	1	2	1	1	2	9.164	11.732	17.039
1	1	3	1	1	1	7.824	10.947	17.639
1	1	3	1	1	2	7.645	11.038	15.987
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1	7	6	1	1	1	10.94	14.575	22.548
1	7	6	1	1	2	10.713	13.319	17.499
1	7	7	0	1	1	9.008	12.998	20.454
1	7	7	0	1	2	8.72	12.285	16.987

数据文件应该包括 6 列试验设计代号以及若干列所需分析的性状数据。试验设计代号的前 6 列分别是:(1)环境、(2)母本、(3)父本、(4)基因型、(5)区组、(6)性别,1=同配性别(XX 或 ZZ),2=异配性别(XY 或 ZW)。在基因型代号的第四列中,亲本的代号为 0; $F_1=(P_i \times P_j)$ 的代号为 1; $F_2=(F_1 \times F_1)$ 的代号为 2; 其余 5 列中最小的常数为 1。最大的常数值在第一列为总环境数,第二、三列为总亲本数,第五列为总区组或重复数,第六列为 2。

在数据文件中不能有断行。每一列的第一行是描述性的文字,可由中文

或英文字母组成，但不能夹带空格。所有的列与列之间都以空格隔开，空格数目不限。性状的列数就是所分析的性状数，不能超过 100 个。

b 建立遗传模型系数矩阵文件

在运用发育数量性状遗传分析的软件以前，首先需要选定所用的遗传模型，根据原始数据文件建立该遗传模型的系数矩阵文件。

执行 GENAD 软件可以生成加性-显性(AD)遗传模型的系数矩阵。GENADM 软件则可以生成加性-显性-母体(ADM)遗传模型的系数矩阵。运行 GENADE 软件可以生成加性-显性-上位性(ADAA)遗传模型的系数矩阵。使用 GENSEX 软件可以生成包括性连锁和母体效应的 ADLM 遗传模型的系数矩阵。执行 GENDIPLD(或 GENTRIPL)软件可以生成二倍体种子(或三倍体胚乳)遗传模型(包括直接遗传效应、细胞质效应、母体遗传效应三套遗传体系)的系数矩阵文件。

在电脑中输入“GENAD C”(或“GENADM C”、或“GENADE C”、或“GENSEX C”、或“GENDIPLD C”、或“GENTRIPL C”)执行命令时，显示器上会出现以下中文提示(如果运行软件不带“C”则显示英文提示)：

中文提示	英文提示	键盘回应
请输入数据文件名:	Input Name of Your Data File:	输入数据文件名
是否需要包括区组效应?是 Y/否 N	Do You Have Block Effects Within Location? Y/N	输入 y 或 n

在运行 GENADM 软件时，将根据二倍体种子模型或三倍体胚乳模型生成加性-显性-母体性遗传模型的系数矩阵。因此在软件运行过程中，显示器上会出现以下提示：

中文提示：“您是否分析三倍体胚乳性状？ Y/N ”

英文提示：“Do You Analyze Triploid Endosperm? Y/N ”

如果所分析的数据是三倍体胚乳性状，则在键盘上输入“y”，否则输入“n”。

系数矩阵建模软件运行完毕后会自动产生 2 个文件。一个是以“.MAT”为后缀的文本文件，储存了遗传模型系数矩阵的有关信息；另一个是以“.DAT”为后缀的非文本文件，储存了所分析的性状数据。

然后采用条件变量的分析软件，产生新的条件变量数据文件。对于形态性状，运行“GENCOND1 C”或“GENCOND1”，便可按 MINQUE(1)的方法产生以“.DOC”为后缀的条件变量数据文件；对于种子性状，运行“GENCOND0 C”或“GENCOND0”，便可按 MINQUE(0/1)的方法产生以“.DOC”为后缀的条件变量数据文件。

运行条件变量分析软件时，显示器上会出现一些提示，使用者可根据提示输入键盘回应。

中文提示	英文提示	键盘回应
请输入数据文件名:	Input Name of Your Data File:	输入数据文件名
请输入条件性状的步长=	Input Step for Conditional Trait=	输入 1 或大于 1 的数

如果所分析的条件变量是 $(t|t-c)$ ，输入的步长为 c 。在通常情况下 c 的取值为 1。

现以小鼠发育性状的分析为示例，先运行“GENSEX”软件分析原始数据文件“micedata.txt”，建立遗传模型系数矩阵。然后如果运行“GENCOND1 C”软件，由原始数据文件“micedata.txt”中的 3 个序列性状，产生 2 个条件变量的数据，并将其与试验设计系数矩阵一起存入名为“micedata.doc”的条件变量数据文件。其格式与原始数据文件类似，如下所示：

环境	母本	父本	基因型	区组	性别	21天体重	14天体重	28天体重	21天体重
1	1	1	0	1	1	11.7961		18.0367	
1	1	1	0	1	2	11.2503		18.4868	
1	1	2	1	1	1	12.3308		17.5387	
1	1	2	1	1	2	11.5515		16.7771	
1	1	3	1	1	1	12.1425		18.034	
1	1	3	1	1	2	12.3583		17.6551	
⋮	⋮	⋮	⋮	⋮	⋮	⋮		⋮	
1	7	6	1	1	1	12.6555		19.1797	
1	7	6	1	1	2	11.6095		18.3046	
1	7	7	0	1	1	13.0576		18.751	
1	7	7	0	1	2	12.5652		18.0486	

其中第 7 列是小鼠的条件体重(给定 14 天体重的表现型值, 小鼠 21 天的条件体重); 第 8 列是给定 21 天体重的表现型值, 小鼠 28 天的条件体重。

条件变量数据文件 “micedata.doc” 产生以后, 必须再运行一次建立遗传模型系数矩阵的软件。在本示例中运行 “GENSEX”, 重新产生分析条件方差所需要的遗传模型系数矩阵。

c 分析条件方差分量和条件遗传效应值

一旦发育数量遗传分析所需要条件变量数据文件和遗传模型系数矩阵生成以后(电脑中将产生 3 个新文件, 分别以 “.DOC”、“.MAT”、“.DAT” 为后缀), 便可以采用 B1 和 B2 中介绍的 MINQUE(1)和 MINQUE(0/1)方法, 分析条件方差分量和条件遗传效应值。

对于形态性状, 可以采用 “GENVAR1R” 或 “GENVAR1C” 估算条件方差分量和预测条件遗传效应值; 而对于种子性状, 则可采用 “GENVAR0R” 或 “GENVAR0C”。具体分析步骤参见 B1 和 B2 的有关内容。

在本示例中, 采用 “GENVAR1C C” 分析了 “micedata.doc” 的条件

变量数据。以基因型为抽样单位，每次删除 2 个观察值(即一个基因型的 2 种性别观察值)。条件方差和条件遗传效应的分析结果将自动存入

“micedata.var”文本文件。该文件可以直接打印，也可以调入文字编辑软件或数据库软件重新编辑。

条件方差分量的估算结果和条件遗传效应值的预测结果以各个条件性状的先后次序，安排于6列：(1)参数名、(2)估计值或预测值、(3)由Jackknife方法计算的标准误、(4)显著性测验的概率值(P-值)、(5)试验方差分量转换成遗传方差分量是的系数、(6)显著性程度(N=不显著，S += 10%显著，S * = 5%显著，S ** = 1%显著)。在实际文本文件中，每一列之间都是以逗号“，”分隔的。当文件读入《Excel》软件时，可以转换成列表形式。

在小鼠发育遗传试验中，分析的第一个条件性状是“21天体重|14天体重”。由于所分析的是条件变量，其估算的方差分量便自然是条件方差分量。方差分量的分析结果如下：

方差分量	估计值	标准误	P-值	显著性
(1): 加性 Var/2	0.358569	0.0622	2.86E-07	S **
(2): 显性 Var/1	0.145085	0.042003	0.000582	S **
(3): 性连锁 Var/1	0.141082	0.02329	1.02E-07	S **
(4): 母体 Var/1	0.219497	0.043327	3.22E-06	S **
(5): 机误 Var/1	0.032656	0.009209	0.000442	S **
(6): Var(Pheno.)	0.896889	0.089699	1.27E-13	S **

各项条件方差分量与条件表现型方差的比率的分析结果如下：

方差分量比率	估计值	标准误	P-值	显著性
(1): 加性 Var/Vp	0.399792	0.039301	7.21E-14	S **
(2): 显性 Var/Vp	0.161765	0.042266	0.000187	S **
(3): 性连锁 Var/Vp	0.157301	0.023682	1.30E-08	S **

(4): 母体 Var/Vp	0.244731	0.030787	1.31E-10	S **
(5): 机误 Var/Vp	0.036411	0.020865	0.0437	S *

以下是各亲本条件加性遗传效应值的分析结果，线性对比是($\sum_{i=1}^4 A_i$ 与 $\sum_{i=5}^7 A_i$):

加性 随机效应	预测值	标准误	P-值	显著性
A 1	-0.028726	0.112744	0.8	NS
A 2	-0.804693	0.152988	3.31E-06	S **
⋮	⋮	⋮	⋮	⋮
A 6	0.404305	0.113754	0.000863	S **
A 7	0.254116	0.10399	0.0183	S *
线性对比/SE	-3.81983	1.67582	0.0271	S *

以下是各亲本和杂交组合的条件显性遗传效应值的分析结果，线性对比是

$$(\Delta = \sum_{i=1}^7 D_{ii} / \sqrt{7\sigma_D^2}):$$

显性 随机效应	预测值	标准误	P-值	显著性
D 11	-1.011021	0.609232	0.104	NS
D 22	-0.448581	0.376608	0.239	NS
⋮	⋮	⋮	⋮	⋮
D 77	0.526068	0.385318	0.179	NS
D 12	0.855106	0.418915	0.0467	S *
⋮	⋮	⋮	⋮	⋮
D 75	1.455256	0.901917	0.113	NS
D 76	-0.748783	0.425608	0.0849	S +
线性对比/SE	-0.321369	1.56538	0.838	NS

以下是条件性连锁遗传效应值的分析结果，线性对比是($\sum_{i=1}^7 L_{i1}$ 与 $\sum_{i=5}^7 L_{i2}$):

性连锁 随机效应	预测值	标准误	P-值	显著性
L 1 for Sex1	0.40394	0.172956	0.0237	S *
L 2 for Sex1	0.359347	0.088017	0.000168	S **
⋮	⋮	⋮	⋮	⋮
L 6 for Sex1	-0.268785	0.116739	0.0257	S *
L 7 for Sex1	0.347687	0.070943	1.13E-05	S **
⋮	⋮	⋮	⋮	⋮
L 6 for Sex2	-0.816437	0.144588	8.65E-07	S **
L 7 for Sex2	-0.469901	0.097211	1.42E-05	S **
线性对比/SE	3.77967	0.425516	1.06E-11	S **

以下是各亲本的条件母体遗传效应值的分析结果，线性对比是($\sum_{i=1}^4 M_i$ 与 $\sum_{i=5}^7 M_i$):

母体 随机效应	预测值	标准误	P-值	显著性
M 1	-0.213643	0.086276	0.0169	S *
M 2	0.606899	0.161767	0.000473	S **
⋮	⋮	⋮	⋮	⋮
M 6	-0.488591	0.195357	0.0158	S *
M 7	-0.071571	0.187113	0.704	NS
线性对比/SE	3.08206	2.07216	0.143	NS

B4. 多年份、多试点品种区域试验的统计分析软件

a 建立原始数据文件

分析数量性状区域试验资料前,必须首先在程序运行的子目录中建立所需分析的原始数据文件(一般以“.TXT”作为后缀),文件以文本格式存储。

建立的数据文件文本格式应如下所示:

品种	年份	试点	区组	单铃重	衣分	籽指
1	1	1	1	7.215	39.25	9.95
1	1	2	1	5.41	42	10.3
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1	1	9	1	4.86	39.05	11.6
1	1	10	1	6.23	40.65	11.65
⋮	⋮	⋮	⋮	⋮	⋮	⋮
10	4	1	1	5.485	33.05	11.45
10	4	2	1	5.65	35.2	9.3
⋮	⋮	⋮	⋮	⋮	⋮	⋮
10	4	5	1	4.38	40.6	10.25
10	4	6	1	5.945	32.8	12

数据文件应该包括4列试验设计代号以及若干列所需分析的性状数据。试验设计代号的前4列分别是:(1)品种、(2)年份、(3)试点、(4)区组。在数据文件中不能有断行。每一列的第一行是描述性的文字,可由中文或英文字母组成,但不能夹带空格。所有的列与列之间都以空格隔开,空格数目不限。性状的列数就是所分析的性状数,不能超过100个。

b 建立区试分析的试验设计系数矩阵文件

在运用区域试验分析软件以前，首先运行“GENTEST”软件，根据原始数据文件建立该遗传模型的系数矩阵文件。

在电脑中输入“GENTEST C”执行命令时，显示器上会出现以下中文提示 (如果运行软件不带“C”则显示英文提示)：

中文提示	英文提示	键盘回应
请输入数据文件名:	Input Name of Your Data File:	输入数据文件名

系数矩阵建模软件运行完毕后会自动产生 2 个文件。一个是以“.MAT”为后缀的文本文件，储存了试验设计系数矩阵的有关信息；另一个是以“.DAT”为后缀的非文本文件，储存了所分析的性状数据。

c 单一性状的分析

区域试验的资料通常包括若干农艺性状，采用“GENTESTM”可以分析各单一性状的品种表现。

在本示例中，采用“GENTESTM C”分析了“testdata.txt”的区试数据。当运行分析软件时，显示器上会出现一些提示，使用者可根据提示输入键盘回应。

中文提示	英文提示	键盘回应
请输入数据文件名:	Input Name of Your Data File:	输入数据文件名
是否把参试品种观察值转换成对照的百分率? 输入Y或N	Do You Want Ck Transformation? Y/N	输入 y 或 n
您需要作多少次线性对比测验? 请输入数字	How Many Linear Contrasing Tests Do You Want to Do?	输入所需的线性对比测验数
再输入线性对比测验每个品种	Input Coefficients for Each Variety: 1	

的系数：第一组品种为 1，第二组品种为-1，其它的品种为 0.	for the First Group, -1 for the Second Group, 0 for Others.	
---------------------------------	---	--

本软件可以根据个人的需要，确定线性对比测验的数目。线性对比一般是用于比较二类亲本的主效应，其中一类亲本的系数可选定为 1，另一类亲本的系数则选为 -1，其余不参与比较的亲本系数为 0。如果所比较的二类亲本数目不一致，电脑软件会自动判别调整其相应的系数。

在输入各亲本的系数以前，显示器会出现以下提示：

中文提示：请稍等片刻，需要输入线性对比的系数！

英文提示：Please Waiting for Input Coefficients!

表明电脑需要运行片刻，以完成矩阵的求逆。如果试验数据很大，可能需要较长时间。当电脑结束矩阵求逆后，再按显示器提示，输入品种线性对比所需要的各品种的系数。

单一性状区试分析结果将自动存入以“.VAR”为后缀的文本文件。现以一个中文输出的结果为例，说明各项分析的排列格式。所分析的原始数据存放在“testdata.txt”文件中。运行“GENTESTM C”后，分析结果自动存放在“testdata.var”文本文件中。该文件可以直接打印，也可以调入文字编辑软件或数据库软件重新编辑。

结果文件的第一部分列出了一些描述性的信息。

电脑输出结果	注解
性状数 3	所分析的性状数目
方差分量数 6	年份、试点、年份×试点、品种×年份、品种×试点、剩余效应等 6 项方差分量
采用 MINQUE(1)法估算方差分量 <GENTESTM.EXE>.	使用的分析软件是“GENTESTM.EXE”

结果文件的第二部分列出所采用的统计分析方法的参考文献:

1. 朱军等: 浙江农业大学学报 (1993) 19(1)7-13.
 〈作物区域试验非平衡数据的统计分析方法:单一性状的分析〉
2. 朱军等: 浙江农业大学学报 (1993) 19(3)241-247.
 〈作物区域试验非平衡数据的统计分析方法:综合性状的分析〉

结果文件列出以下了品种线性对比系数,

+ < 1 > + < 2 > - < 3 > - < 4 > - < 5 >

表明该试验比较了品种(1, 2)与亲本(3, 4, 5)的平均数。

区试分析的主体部分是各性状的分析结果。首先列出6个随机效应方差的估算结果。6个随机效应分别是: (1)Y=年份效应、(2)L=试点效应、(3)YL=年份×地点交互效应、(4)GY=品种×年份交互效应、(5)GL=品种×试点交互效应、(6)e=随机机误效应。

性状单铃重的统计分析结果

Estimate of Var(Y) =	0.0654411
Estimate of Var(L) =	0.291648
Estimate of Var(YL) =	0.146901
Estimate of Var(GY) =	0.00458665
Estimate of Var(GL) =	0.0109432
Estimate of Var(e) =	0.113743

随后列出了各品种的平均数及其显著性测验结果(小写字母为5%显著水准测验,大写字母为1%极显著水准测验):

位次	品种代号	平均数	显著性测验
(10)	V 7	5.5400	a A
(9)	V 1	5.5845	ab AB
(8)	V 8	5.7746	ab ABC
(7)	V 6	5.8167	ab c ABCD
(6)	V10	5.8627	b c BCDE
(5)	V 4	5.8950	b c d BCDE
(4)	V 5	6.0269	c de BCDE
(3)	V 2	6.2375	c de CDE
(2)	V 3	6.3586	c de DE
(1)	V 9	6.4317	e E

对于每一项线性对比测验,首先列出了对比的二组品种代号。对于本示例,比较了(V1, V2)与(V3, V4, V5)的平均数间的差异,结果如下:

线性对比	C-值	标准误<S.E>	u-值
Contrast 1	-1.095121	0.892828	1.226576

区试分析还列出了品种稳定性测验的分析结果,先是简要的提示:

品种的稳定性测验:

截距 a 和相关 r 的置信区间如包括 0.0 则不显著
斜率 b 的置信区间如包括 1.0 则不显著

然后列出了各个品种的稳定性参数的估计量、标准误和 95%置信区间。以下是第一个品种单铃重的估算结果:

截距 $a=0.731$	标准误 S.E.= 0.345	置信区间从 0.054 到 1.408
斜率 $b=0.856$	标准误 S.E.= 0.068	置信区间从 0.722 到 0.989
相关 $r=0.927$	标准误 S.E.= 0.063	置信区间从 0.803 到 1.051

d 综合性状的分析

采用“GENTESTW”分析区域试验综合性状的品种表现,必须安排在单一性状分析后立即进行,这样可以直接利用各性状的方差分量计算结果及生成的逆矩阵。

在本示例中,采用“GENTESTW C”分析了“testdata.txt”的区试数据。当运行分析软件时,显示器上会出现一些提示,使用者可根据提示输入键盘回应。由于是分析综合性状,须输入每一个性状的加权值。在本示例中 3 个性状的加权系数分别是: 0.2, 0.4, 0.4。其它提示和键盘输入均与单一性状分析的类似。

因为分析单一性状时电脑已经完成了矩阵的求逆,所以综合性状分析的运行速度很快。

综合性状区试分析结果将自动存入以“.COV”为后缀的文本文件。现以一个中文输出的结果为例,说明各项分析的排列格式。运行“GENTESTW C”分析原始数据“testdata.txt”,结果自动存放在“testdata.cov”文本文件中。该文件可以直接打印,也可以调入文字编辑软件或数据库软件重新编辑。

与单一性状的结果文件类似，第一部分列出了一些描述性的信息，第二部分列出所采用的统计分析方法的参考文献。另外列出了所采用的加权系数：

<W1>: 0.2, <W2>: 0.4, <W3>: 0.4,

和品种线性对比系数，

+<1> +<2> +<3> -<4> -<5> -<6>

表明该试验比较了品种(1, 2, 3)与亲本(4, 5, 6)的平均数。

综合性状区试分析的主体部分是各性状的方差分量以及成对性状的协方差分量估算结果：

性状 单铃重 的方差分析结果

Estimate of Var(Y) =	0.0654411
Estimate of Var(L) =	0.291648
Estimate of Var(YL) =	0.146901
Estimate of Var(GY) =	0.00458665
Estimate of Var(GL) =	0.0109432
Estimate of Var(e) =	0.113743

性状 单铃重 与 衣分% 的协方差分析结果

Estimate of Cov(Y) =	6.75109
Estimate of Cov(L) =	-22.5311
Estimate of Cov(YL) =	0.388549
Estimate of Cov(GY) =	-0.722389
Estimate of Cov(GL) =	0.414833
Estimate of Cov(e) =	2.01136

在方差和协方差分析结果以后，列出了各品种综合性状的的加权平均数及其显著性测验结果(小写字母为5%显著水准测验，大学字母为1%极显著水准测验)：

位次	品种代号	平均数	显著性测验
(10)	V 7	98.0729	a A
(9)	V 1	99.3565	ab AB
(8)	V10	100.0000	b AB
(7)	V 6	101.9190	b AB
(6)	V 8	102.5645	b B
(5)	V 5	103.1982	b B
(4)	V 4	103.3885	b B
(3)	V 9	104.0072	b B
(2)	V 2	104.2753	b B
(1)	V 3	104.5228	b B

对于每一项线性对比测验，列出了对比的二组品种代号。对于本示例，比较了(V1, V2, V3)与(V4, V5, V6)的综合性状的加权平均数间的差异，结果如下：

线性对比	C-值	标准误<S.E>	u-值
Contrast 1	-1.053509	120.792450	0.008722

区试分析还列出了品种综合性状稳定性测验的分析结果。在简要的提示后，列出了各个品种的稳定性参数的估计量、标准误和95%置信区间。以下是第一个品种综合性状的估算结果：

截距 a= 15.724	标准误 S.E.= 6.157	置信区间从 3.656 到 27.793
斜率 b= 0.836	标准误 S.E.= 0.063	置信区间从 0.713 到 0.959
相关 r= 0.954	标准误 S.E.= 0.023	置信区间从 0.909 到 0.999

参考文献

- 朱军. 1992. Mixed model approaches for estimating genetic variances and covariances. *生物数学学报* 7(1): 1-11.
- 朱军. 1993a. 作物杂种后代基因型值和杂种优势的预测方法. *生物数学学报*, 8(1):32-44.
- 朱军. 1993b. Mixed model approaches for estimating covariances between two traits with unequal design matrices. *生物数学学报*, 8(3) 24-30.
- 朱军. 1994. 广义遗传模型与数量遗传分析新方法. *浙江农业大学学报*, 20(6)551-559.
- 朱军. 1996. 包括基因型×环境互作效应的种子遗传模型及其分析方法. *遗传学报*, 23(1):56-68.
- 朱军、段家龙. 1994. Genetic models and analysis methods for sex-linked and maternal gene effects. *生物数学学报*,9(4):1-9.
- 朱军、赖鸣冈、许馥华. 1993. 作物品种区域试验非平衡资料的分析方法: 综合性状的分析. *浙江农业大学学报*, (1993) 19(3):241-247.
- 朱军、许馥华, 1994. 禾谷类作物胚乳品质性状的遗传模型及其分析方法. *作物学报*, 20(3): 264-270.
- 朱军、许馥华、赖鸣冈. 1993. 作物品种区域试验非平衡资料的分析方法: 单一性状的分析. *浙江农业大学学报*, (1993), 19(1):7-13.
- 朱军等. 1993. 作物品种间杂种优势遗传分析的新方法. *遗传学报*, 20(3): 262-271.
- 段家龙、朱军. 1994. 家蚕茧质性状的基因效应分析. *生物数学学报*, 9(4): 10-17.
- Cockerham, 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.
- Cockerham, C. C. 1980. Random and fixed effects in plant genetics.

- Theor. Appl. Genet.**, 56: 119-131.
- Cockerham, C. C. and Weir, B. S. 1977. Quadratic analyses of reciprocal crosses. **Biometrics**, 33: 187-203.
- 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.
- East, E. M. 1915. Studies on size inheritance in *Nicotiana*. **Genetics** 1: 164-176.
- Eisen, E. J. *et al.* 1966. Sex-linked and maternal effects in the diallel cross. **Aust. J. Biol. Sci.**, 19: 1061-1071.
- Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. **Trans. Roy. Soc. Edinb.**, 52: 399-433.
- Fisher, R. A. 1925. *Statistical Methods for Research Workers*. 1st ed. Oliver & Boyd. Edinburgh and London.
- Griffing, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. **Aust. J. Biol. Sci.**, 9: 463-493.
- Hallauer, A. R. and Miranda, J. B. 1981. *Quantitative Genetics in Maize Breeding*. Iowa State University Press, Ames, Iowa.
- Hartley, H. D. and Rao, J. N. K. 1967. Maximum-likelihood estimation for the mixed analysis of variance model. **Biometrika**, 54: 93-108.
- Hayman, B. I. 1954. The theory and analysis of diallel crosses. **Genetics**, 39: 789-809.
- Henderson, C. R. 1963. Selection index and expected genetic advance. In: Hanson, W. D., Robinson, H. F. eds. **Statistical Genetics and Plant Breeding**. Natl. Acad. Sci. -N R C 982: 141-163.
- Henderson, C. R. 1965. Estimation of variance and covariance components. **Biometrics**, 9: 226-252.
- Johannsen W. 1909. *Elemente der exakten Erblchkeitslehre*. Gustav

- Fischer, Jena.
- Miller, R. G. 1974. The jackknife: a review. **Biometrika**, 61: 1-15.
- Patterson, H. D. and Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. **Biometrika**, 58: 545-554.
- Rao, C. R. 1971. Estimation of variance and covariance components: MINQUE theory. **J. Multivar. Anal.**, 1: 257-275.
- 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: North-Holland.
- Searle S. R. 1968. Another look at Henderson's methods of estimating variance components. **Biometrics**, 24: 749-778.
- Searle S. R. 1970. Large sample variances of maximum likelihood estimators of variance components using unbalanced data. **Biometrics**, 26: 505-524.
- Zhu, J. 1989. Estimation of genetic variance components in the general mixed model. Ph.D. Dissertation. North Carolina State Univ. Raleigh, N.C.
- Zhu, J. 1995. Analysis of conditional effects and variance components in developmental genetics. **Genetics**, 141(4):1633-1639.
- Zhu, J. and Weir, B. S. 1994a. Analysis of cytoplasmic and maternal effects: I. a genetic model for diploid plant seeds and animals. **Theor. Appl. Genet.**, 89:153-159.
- Zhu, J. and Weir, B. S. 1994b. Analysis of cytoplasmic and maternal effects: II. genetic models for triploid endosperms. **Theor. Appl. Genet.**, 89:160-166.
- Zhu, J. and Weir, B. S. 1996. Diallel analysis for sex-linked and maternal effects. **Theor. Appl. Genet.**, 92:1-9.