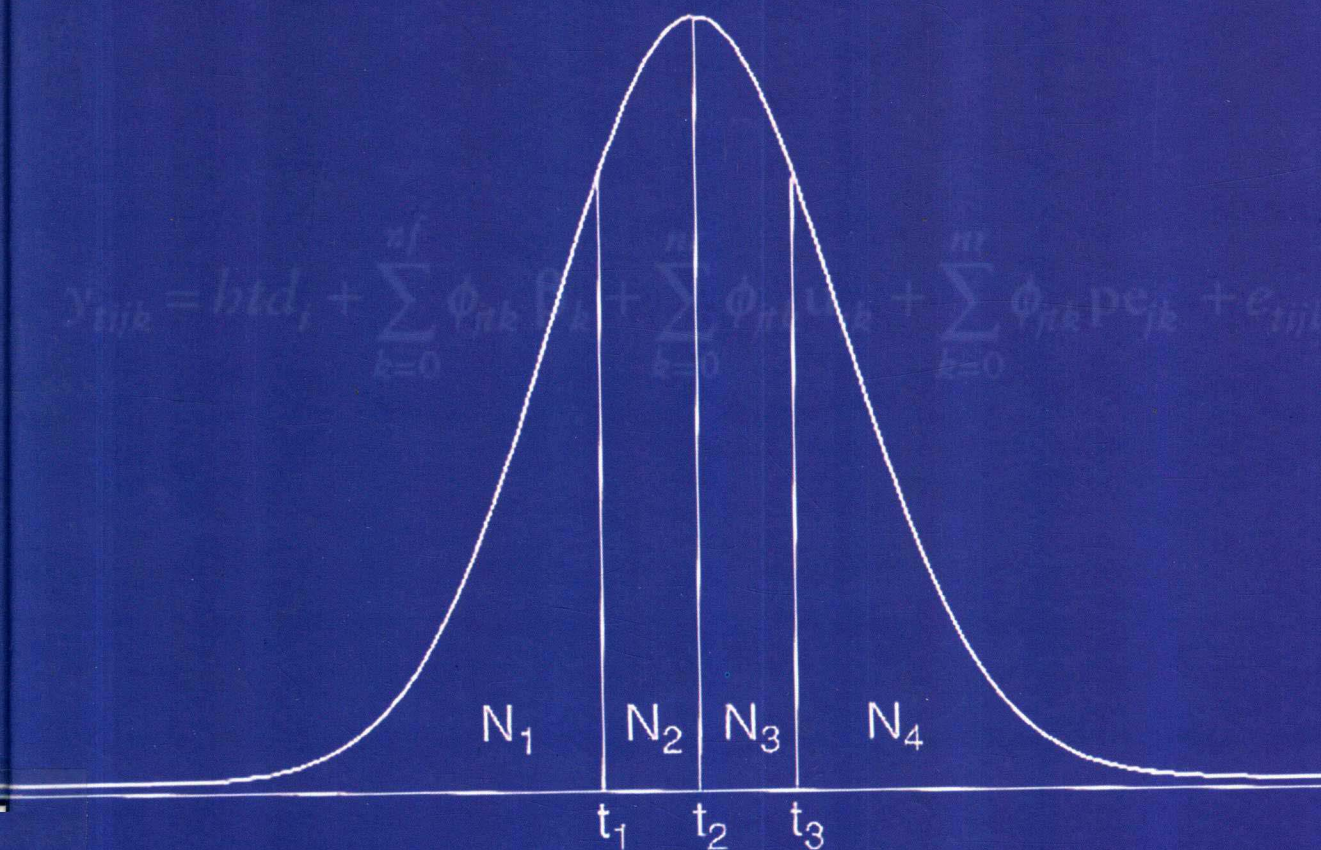


线性模型在动物育种值 预测中的应用

(第3版)

[英] Raphael A. Mrode 编著
于向春 张 豪 主译



 中国农业出版社

线性模型在动物育种值 预测中的应用

(第3版)

[英] Raphael A. Mrode 编著

于向春 张 豪 主译

中国农业出版社

图书在版编目 (CIP) 数据

线性模型在动物育种值预测中的应用: 第3版 /
(英) 莫罗德·拉斐尔 (Raphael A. Mrode) 编著
于向春, 张豪主译. —北京: 中国农业出版社, 2016. 12
ISBN 978-7-109-22168-0

I. ①线… II. ①莫… ②于… ③张… III. ①线性模
型-应用-畜禽育种-研究 IV. ①S813.2

中国版本图书馆 CIP 数据核字 (2016) 第 231079 号

Linear Models for the Prediction of Animal Breeding Values, 3rd Edition
By Raphael A. Mrode
© R. A. Mrode, 2014

本书简体中文版由 Raphael A. Mrode 和 CAB International 授权中国农业出版社独家出版发行。本书内容的任何部分, 事先未经出版者书面许可, 不得以任何方式或手段复制或刊载。

北京市版权局著作权合同登记号: 图字 01-2015-7499 号

中国农业出版社出版

(北京市朝阳区麦子店街 18 号楼)

(邮政编码 100125)

责任编辑 张艳晶

北京通州皇家印刷厂印刷 新华书店北京发行所发行
2016 年 12 月第 1 版 2016 年 12 月北京第 1 次印刷

开本: 787mm×1092mm 1/16 印张: 20.25

字数: 450 千字

定价: 88.00 元

(凡本版图书出现印刷、装订错误, 请向出版社发行部调换)

译 者 名 单

主 译	于向春	海南大学
	张 豪	华南农业大学
副主译	侯冠彧	中国热带农业科学院
	张燕军	内蒙古农业大学
	苏 蕊	内蒙古农业大学
	王瑞军	内蒙古农业大学
	王志英	内蒙古农业大学
	李加琪	华南农业大学
	张 哲	华南农业大学
主 审	张立岭	海南大学

序

动物育种学是应用遗传学理论指导育种实践的有关科学知识体系，该学科体系近些年来发展迅速，这得益于数量遗传学和分子遗传学理论、方法的逐步发展以及计算能力的提高。其中，线性模型理论应用于动物个体的育种值预测，使得对个体的遗传评估更加准确，极大地提高了育种效率。

R. A. Mrode 教授所著《Linear Models for the Prediction of Animal Breeding Values》，是近年来在英美等国家动物育种学科体系中最重要参考书之一。今由于向春、张豪等译者把这本经典著作编译成中文版本《线性模型在动物育种值预测中的应用》。著作综述了数量遗传学和基因组学的最新研究进展，对动物育种值预测和育种规划工作影响深远。著作内容还包括生存分析模型、社会互动模型、公畜-外祖父模型，以及 SNP 标记用于基因组育种值计算。此外，书中还含有选择指数制定、随机回归模型用于纵向数据分析、矩阵代数等知识。该书各部分内容均有实例进行例证，可以很好地平衡复杂的育种理论与实践应用之间的差距，适合动物遗传育种专业的研究生和研究者学习和参考。

对动物个体的重要经济性状如生长速度、产奶量和产绒量等进行遗传评定，是动物育种研究者的必备技能。衷心希望该著作对于读者理解动物育种的新理论和新方法有所裨益，在条件具备的情况下，积极应用这些方法和技术，大幅度提高家畜产品的数量和品质，使之造福于民。

是为之序！

张豪
2016.12.27

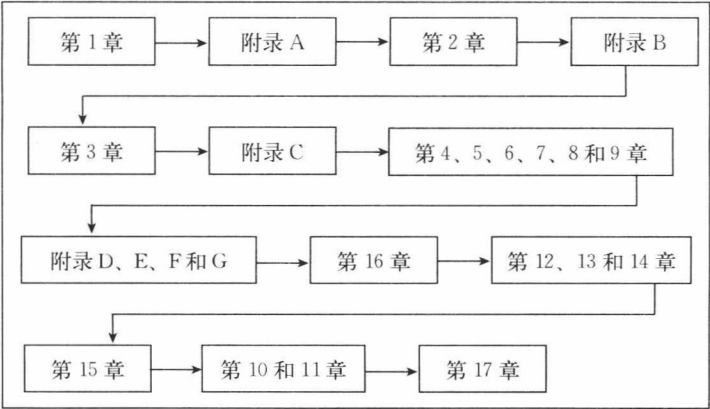
译者自绪

根据美国农业部的研究,品种选育在提高畜禽生产性能的诸多因素中作用最大,占40%左右。掌握动物遗传评定的理论和方法是每个动物育种工作者必需的技能。R. A. Mrode 博士所著《Linear Models for the Prediction of Animal Breeding Values》,中文译名《线性模型在动物育种值预测中的应用》,已经修订了两次,是目前动物遗传评定方面的权威专著。我们组织人员进行翻译,反复校对,历时两年多时间,终于得以完成。

原书共有17章,7个附录,并有参考文献和主题词索引,16开本共343页。除第13章介绍等级资料的分析方法、第14章介绍截尾数据的分析外,其他章节均是关于数量性状的分析方法。附录A介绍了线性代数的基本知识,附录B为分子血缘相关矩阵的快速算法,附录C对BLUP的混合模型方程组进行了推导,它们连同前4章为本书的基础内容。第5、6、7和第9章为本书的核心内容,介绍的方法是目前绝大多数国家和育种公司使用的方法。第11章的基因组选择也已经在发达国家的奶牛业中普遍应用。第8章介绍了如何分析群居对动物生产性能影响的遗传分析方法。第12章重点说明无近交群体非加性遗传效应的分析方法。第10章为DNA标记在动物育种值预测中的应用。第17章的线性方程求解方法,既适用于BLUP的混合模型方程组计算,也适用于方差组分分析。第15章对方差组分估计原理和方法进行了简单介绍。第16章讲述了贝叶斯理论应用于动物育种值计算的经典内容。

本书内容较全面,但并没有包括动物育种中线性模型的所有内容,如第10章只讲述了标记QTL间方差协方差矩阵的计算,只叙述了3种方法中的1种。对于一本专著,内容必然有所选择,虽然作者的选择不一定符合每个人的偏好。

线性模型可能是动物科学研究生最难学的一门课程,读者可以按照书的章节顺序阅读,也可以选择几章学习,译者建议的学习顺序为:



本书既可以当作教材，也可以当作参考书。书后的参考文献为广大研究生和从事动物遗传育种研究工作者提供了有用的文献资料，可以作为研究和学习需要重点阅读的文献。

由于译者水平所限，译著中一定存在不少不足甚至错误，欢迎读者提出意见和建议。

华南农业大学动物科学学院 张豪

2016 年 12 月于广州

译 者 的 话

关于《线性模型在动物育种值预测中的应用》(第3版)译书,其内容、特色、阅读方法和适用对象等,在李金泉教授、张豪教授和原著者 R. A. Mrode 博士的序或绪言中已有较为详细的说明和介绍,读者从中会大受裨益。

我是在李金泉教授的指导下,逐渐接触数量遗传学和动物育种学的,并已从相关教学和研究十多载。李老师引领我入门,多年来在我成长的过程中,李老师也给予了诸多鼓励和帮助,否则就不可能有本译著的问世,千言万语,无尽感激!

我和张豪教授是本译著的共同主译,这也是我们之间的第二次合作。第一次合作是在2011年共同翻译了《动物科学生物统计导论》(第2版)一书。张豪教授严谨的态度、务实的学风和渊博的知识令我十分钦佩,并获益匪浅,在此深致谢意。同时也感谢侯冠或博士、张燕军博士等译者,没有大家的共同努力和辛勤付出,就不能及时完成书稿的翻译和校对工作。

特别感谢张文广教授在本书翻译过程中提出了很多宝贵意见。

在我多年的学习中,我不止一遍地详读过以下书籍,从中获得了大量知识并奠定了翻译本学科英文经典著作的基础。在这里按出版时间先后列出:内蒙古农牧学院主编的《家畜育种学》(第二版,1990)、张沅和张勤编著的《畜禽育种中的线性模型》(1993)、盛志廉和吴常信编著的《数量遗传学》(1995)、Falconer 和 Mackay 所著的《Introduction to Quantitative Genetics》(第4版,1996)、盛志廉和陈瑶生编著的《数量遗传学》(1999)、张立岭编写的《数量遗传学》(内部讲义,2000)、张沅主编的《家畜育种学》(2001),以及张勤著的《动物遗传育种中的计算方法》(2007)。在此谨向这些前辈们致以崇高的敬意!我是怀着崇敬的心情开始翻译这部经典著作的,在翻译过程中力求忠实于原文又能文笔流畅,惟恐文字表达和对动物遗传育种学科的

见解不及前辈们精辟和严谨。虽然已经尽己所能,但限于能力、水平和悟性等问题,文中一定存在不少缺点和错误,恳请读者批评指正。

本译著还得到下列教学和研究项目的资助,分别是:内蒙古农业大学的家畜遗传繁育系列课程国家级教学团队建设项目、家畜育种学国家级资源共享课建设项目;国家星火计划重点项目——特优儋州鸡扩繁与生态养殖技术产业化(2015GA800001);国家生猪产业技术体系(杂交配套CARS-36);海南省高等学校科学研究项目——热带农业复合应用型人才培养与实践(Hnjgzd2014-01);国家自然科学基金青年基金项目(31302011);海南省农业科技服务体系建设专项(HNXH201534),在此表示真诚感谢。

最后,感谢CABI出版社能够授权允许翻译原著,这有利于动物育种理论和方法的广泛传播和应用;诚挚感谢中国农业出版社大力支持该译著的出版!

于向春

2016年12月于海口

原 著 序

最佳线性无偏预测 (BLUP) 已被广泛用于家畜的遗传评估。自 Hender-son (1949) 首次提出该方法以来, 根据遗传评估模型的不同, 如从早期的公畜模型、公畜和外祖父模型已经发展到单性状动物模型、多性状动物模型, 以及近年来应用于纵向数据分析的随机回归模型和用于基因组数据分析的 SNP-BLUP 或 GBLUP。计算方法和计算能力的进步, 加快了该方法的发展。目前, 多数国家的家畜遗传评估体系均使用基于动物模型或随机回归模型的 BLUP 方法。

尽管取得了显著进步并且相关研究文献数量有所增加, 但目前还没有简单而直接的书籍来介绍线性模型在预测育种值中的应用。更重要的是, 在发展中国家, 由于相关杂志或文献的缺乏, 该方法应用较少。该书的出版, 平衡了理论和实际应用间的鸿沟。本书参考了多种出版物的相关文献、结合了相关领域研究者以及过去 14 年来国家评估项目的经验, 内容丰富、实用。

开始的第 1 章介绍育种值预测时的基本模型和一些假设, 以及如何使用个体本身记录进行育种值预测; 接着介绍使用选择指数结合多性状和不同个体信息来预测育种值。从第 3 章至第 5 章, 介绍在遗传评估中, BLUP 在单性状和多性状模型中的应用原理。其中也介绍了通过不同转换简化多性状模型的评估过程, 接着介绍母体性状模型和社会互动模型。第 9 章介绍用于纵向数据分析的随机回归模型。接下来的一章介绍结合遗传标记信息进行标记辅助选择。第 11 章介绍基因组选择。第 12 章介绍非加性遗传动物模型, 以及考虑显性和互作效应时, 几个快速计算相关系数矩阵的逆的算法。接着介绍阈值性状模型和生存模型。在第 15 章和第 16 章, 介绍方差组分估计的基本概念、吉布斯抽样在遗传参数估计中的应用, 以及在单性状和多性状模型中的应用。最后一章介绍混合模型方程组的求解方法, 以及数据迭代的一些公式。在本书中, 理解遗传评估的原理需要基本矩阵代数知识。对于不熟悉矩阵代数的读

者来说,可参见附录 A。理解该书也需要读者具备一定的数量遗传学知识。

文中有许多例子用来说明遗传评定模型的应用,解释了根据混合模型方程组如何求解随机效应和固定效应,同时也向读者论证了不同来源信息经适当加权得到不同模型的个体育种值。

尽管作者已付出诸多努力确保该书内容的准确,但倘若读者发现文中出现了错误,敬请告之。

Robin Thompson 教授编写了方差组分估计一章,尽管工作繁忙,他还是阅读了该书的大部分文稿,对此表示感谢。Ben Hayes 博士、Ricardo Pong-Wong 博士和 John A. Woolliams 教授修改了基因组选择一章,感谢他们提出了有价值的建议。Gabor Meszaros 博士和 Sue Brotherstone 博士在百忙中修改了生存分析一章,在此表示感谢。感谢 Denny Cruz 教授和 Victor Olori 博士修订了降低多性状动物模型维数的方法等章节。非常感谢 W. G. Hill 教授和 G. Swanson 先生通读了该书稿,他们的评论和建议使得本书更加完善。Martin Lidauer 博士和 Ismo Strandén 博士阅读了本书的部分章节,在此表示感谢。感谢 Sebastian Mucha 博士帮助汇图和制表。另外, Mike Coffey 博士以及 C. Smith 和 B. W. Kennedy 两位教授的写作经验对本书非常有益。我向 R. L. Quaas 教授表达感谢,感谢他允许使用尚未出版的关于近交系数计算的相关内容。感谢澳大利亚新英格兰大学的动物遗传育种系,允许我使用他们编写的 BLUP 手册用于本书第 2 章的编写,以及依阿华州立大学的 Fernando R. L. 教授在 2010 年夏季课程的授课内容。对我的妻子 Doris 帮助打印书稿表达诚挚的谢意。特别感谢在我抽空准备书稿时,我的同事 Kevwe、Joshua 和 Esther 给予的帮助,以及所有鼓励关心我的朋友们!

R. A. Mrode
苏格兰农学院
爱丁堡

缩 写 词

- ADG (average daily gain): 平均日增重
ARHS (adjusted right-hand side): 校正右手项
BFAT (backfat thickness): 背膘厚度
BLUE (best linear unbiased estimator): 最佳线性无偏估计量
BLUP (best linear unbiased prediction): 最佳线性无偏预测
BV (breeding value): 育种值
BW (birth weight): 出生重
CF (covariance function): 协方差函数
CR (correlated response): 相关选择反应
DBV (direct breeding value): 直接育种值
DGV (direct genomic breeding values): 直接基因组育种值
DIM (days in milk): 母牛测定日
DRB (deregressed breeding values): 逆回归育种值
DPS (durable performance sum): 持续生产力之和
DYD (daughter yield deviation): 女儿产奶量离差
EBV (estimated breeding value): 估计育种值
EDC (effective daughter contributions): 有效女儿贡献
EM (expectation maximization): 期望最大值
ETA (estimated transmitting ability): 估计传递力
FA (factor analysis): 因子分析
GEBV (genomic breeding values): 基因组育种值
GLS (generalized least squares): 广义最小二乘
GR (growth rate): 生长速度
HTD (herd-test-day): 场测定日
HYS (herd - year - season): 场-年-季
IBD (identical by descent): 同源相同
IGE (indirect genetic effects): 间接遗传效应
INET (index net): 指数网
LD (linkage disequilibrium): 连锁不平衡
LGR (lean growth rate): 瘦肉生长速度
LP (lean percent): 瘦肉率
LPL (length of productive life): 生产寿命

LS (life span): 预期寿命

LSE (least squares equations): 最小二乘方程组

MACE (multi-trait across-country evaluation): 多性状、不同国家的遗传评定

MAS (marker-assisted selection): 标记辅助选择

MBLUP (multivariate best linear unbiased prediction): 多性状最佳线性无偏预测

MCMC (Markov chain Monte Carlo): 马可夫链蒙特卡罗方法

MGD (maternal granddams): 外祖母

MGS (maternal grandsire): 外祖父

ML (marker locus): 标记位点

MME (mixed model equations): 混合模型方程组

MQTL (quantitative trait locus linked to a marker): 标记 QTL

MS (muscle score): 肌肉评分

PC (progeny contribution): 后代贡献

PCG (preconditioned conjugate gradient): 先验共轭梯度

PEC (prediction error covariance): 预测误差协方差

PEV (prediction error variance): 预测误差方差

PIN (production index): 生产指数

PLI (profitable life index): 利润指数

PPA (probable producing ability): 潜在生产力

PTA (predicted transmitting ability): 预测传递力

PWG (post-weaning gain): 断奶后增重

PYD (progeny yield deviation): 后代产奶量差异

QTL (quantitative trait locus): 数量性状座位

RAM (reduced animal model): 简约动物模型

REML [restricted (or residual) maximum likelihood]: 约束最大似然法

RHS (right-hand side): 右手项

RP (residual polygenic): 剩余多基因

RR (random regression): 随机回归

RRM (random regression model): 随机回归模型

RRS (risk ratios): 风险率

SBV (associative breeding value): 间接育种值

SCC (somatic cell count): 体细胞数

SEP (standard error of prediction): 预测标准误

SNP (single nucleotide polymorphism): 单核苷酸多态

SSR (simple sequence repeats): 简单序列重复

TBV (total breeding value): 总育种值

WWG (pre-weaning gain): 断奶前增重

XFA (extended factor analysis): 扩展因子分析

YSP (year - season - parity): 年-季-胎次

目 录

序

译者自绪

译者的话

原著序

缩写词

1 利用不同信息来源进行遗传评定	1
1.1 引言	1
1.2 基本模型	1
1.3 根据个体本身信息预测育种值	2
1.3.1 单一记录	2
1.3.2 重复记录	3
1.4 根据后裔记录估计育种值	5
1.5 通过系谱预测育种值	7
1.6 根据相关性状预测育种值	8
1.7 选择指数	9
1.7.1 选择指数的精确性	10
1.7.2 利用不同来源信息计算相关性状选择指数的例子	11
1.7.3 综合育种值的预测	14
1.7.4 使用预测遗传价值计算总经济指数	15
1.7.5 约束选择指数	16
1.7.6 结合了育种值和遗传标记信息的选择指数	17
2 亲属间的遗传协方差	19
2.1 简介	19
2.2 分子亲缘相关矩阵	19
2.3 亲缘相关矩阵的分解	20
2.4 计算亲缘相关矩阵的逆	22
2.4.1 不考虑近交时, 计算分子亲缘相关矩阵的逆	22
2.4.2 考虑近交时, 计算分子亲缘相关矩阵的逆	24

2.5	计算父亲-外祖父模型中亲缘相关矩阵的逆	26
2.6	例子	27
3	育种值的最佳线性无偏预测：有一个随机效应的单性状模型	29
3.1	引言	29
3.2	理论背景简介	29
3.3	单个个体的评估模型（动物模型）	31
3.3.1	构建混合模型方程组	32
3.3.2	后裔产量的离差	36
3.3.3	评估的精确性	38
3.4	公畜模型	39
	例子	39
3.5	缩约动物模型	41
3.5.1	定义模型	42
3.5.2	例子	43
3.5.3	其他算法	46
3.6	有分组的动物模型	46
	例子	48
4	育种值的最佳线性无偏预测：有随机环境效应的动物模型	53
4.1	引言	53
4.2	重复力模型	53
4.2.1	定义模型	53
4.2.2	例子	54
4.2.3	计算女儿产量的离差	57
4.3	有共同环境效应的动物模型	58
4.3.1	模型定义	58
4.3.2	例子	58
5	育种值的最佳线性无偏预测：多性状动物模型	61
5.1	引言	61
5.2	设计矩阵相同，没有缺失记录	61
5.2.1	模型定义	61
5.2.2	例子	63
5.2.3	多性状动物模型遗传评定的剖分	64
5.2.4	多性状动物模型评定的精确性	66
5.2.5	在多性状动物模型中计算女儿产量的离差	67
5.3	设计矩阵相同，有缺失记录	68

例子	68
5.4 设计矩阵不相等	70
5.4.1 例子	70
5.4.2 在多性状动物模型中计算女儿产量的离差	72
5.5 不存在环境协方差时的多性状动物模型	73
5.5.1 不同亲属的记录性状不同	73
5.5.2 多性状、不同国家奶牛的遗传评定 (MACE)	75
6 降低多性状动物模型维数的方法	83
6.1 引言	83
6.2 典型转换	83
6.2.1 模型	84
6.2.2 例子	84
6.3 乔列斯基转换	86
6.3.1 计算转换矩阵和定义模型	86
6.3.2 例子	86
6.4 因子分析和主成分分析	88
6.4.1 因子分析	89
6.4.2 主成分分析	92
6.4.3 降秩 PC 模型分析	93
7 母体性状模型: 动物模型和缩约动物模型	95
7.1 引言	95
7.2 单个母体性状的动物模型	95
例子	96
7.3 有母体遗传效应的缩约动物模型	100
例子	101
7.4 父亲-外祖父模型	104
8 群居互作模型	105
8.1 简介	105
8.2 有群居互作效应的动物模型	107
例子	108
8.3 互作模型估计方程的剖分	110
8.4 使用相关误差结构 (correlated error structure) 进行分析	112
9 纵向数据的分析	113
9.1 引言	113

9.2	固定回归模型	113
例子		114
9.3	随机回归模型	119
9.3.1	例子	120
9.3.2	在随机回归模型中个体效应解的剖分	123
9.3.3	计算女儿产量的离差	126
9.3.4	估计育种值的可靠性	126
9.3.5	利用样条函数的随机回归模型	127
9.3.6	母体性状的随机回归模型	129
9.4	协方差函数	129
降阶协方差函数的拟合		131
9.5	随机回归模型与协方差函数的等价性	135
10	遗传标记在育种值预测中的使用	136
10.1	引言	136
10.2	定义含有标记信息的模型	136
10.3	计算 MQTL 效应的协方差矩阵 (G_v)	137
例子		138
10.4	G_v 的其他算法	139
10.5	计算 G_v 的逆	141
10.6	当含有标记信息时, 预测育种值	144
例子		145
10.7	直接预测 MQTL 的加性遗传值	146
例子		147
10.8	预测总的加性遗传值	148
例子		149
10.9	当两个标记间有一个 QTL 时的分析	150
10.9.1	基本模型	150
10.9.2	计算协方差矩阵 G	151
10.9.3	例子	153
11	基因组育种值的计算以及基因组选择	156
11.1	引言	156
11.2	一般线性模型	157
11.3	基因型的编码和标准化	157
11.4	把 SNP 效应当作固定效应的模型	157
11.5	计算 SNP 效应的混合线性模型	161
11.5.1	SNP - BLUP 模型	161