

种猪选育系统工程若干问题研究

李茂青 易英 吴天

(厦门大学自动化系 361005)

摘要 本文就种猪选育系统工程中的最优选配问题和饲料配方问题展开讨论,给出了系统建模,计算机仿真,数据库等技术的综合应用的研究结果。

关键词 种猪选育 系统工程 仿真技术 选配 饲料配方

Study on Some Problems about Breeding System Engineering of Piglet

Li Maoqing Yi Ying Wu Tian

(Xiamen University 361005)

Abstract In this paper, the optimal breeding problem and the filling a prescription for feeding pig are studied, and by use of system engineering method some researchful results also are given.

Keywords breeding system engineering; optimal breeding; fill a prescription

1 引言

种猪选育是一项复杂的系统工程,它不仅与猪种本身的遗传基因有密切关系,而且与周围环境,猪舍条件,饲养管理,饲料配方,饲喂方法等紧密相关。传统的选育方法大多基于遗传统计原理,结合生产实际,进行多世代的选育。由于猪的生长繁殖需要较长时间,使得研究周期长,人力物力消耗大。法国的“大约克”种猪就花了三十多年时间才选育成功。

由厦门市系统工程学会和福建天马种猪场联合承担的厦门市科委立项《“大约克”天马种猪选育系统工程研究与实施》项目,从另一角度对种猪选育展开了研究。该项目在传统的选育方法基础上首先确定了选育目标,制定了种猪选育系统工程的总体规划,进而结合应用了系统建模,计算机仿真,数据库及综合评价等技术。自90年开始,就总体规划中涉及到的饲料配方问题,选配选种问题,饲养管理等问题进行了研究与实施。本文重点就最优选配问题和饲料配方问题展开讨论。

2 选配问题

种猪选配是加快优良种猪选育进度的重要一环,制定和采用最优的选配方案,是确保在最短的时间内选育出高质种猪的关键所在。其主导思想如下:

根据数量遗传学和统计学原理,利用计算机对福建省天马种猪场1983~1993的种猪生产繁殖记录和系谱记录进行整理分析,计算出种猪主要繁殖性状——产仔数,初生窝重,20日窝重,断奶窝重,断奶头数的群体遗传参数——遗传力,重复力,相关系数,近交系数和各数量性状间的相关关系。根据以上计算的群体遗传参数和种猪个体的生产表型值,计算出每头种猪各繁殖性状的育种值,经分类整理后建立相应的样

本库。样本库中除含有能通过上述计算求得育种值的那些种猪记录外,还需包含那些尚无生产繁殖经历的后备种猪的记录。这些后备种猪的育种值通过以下两类模型来求得:

模型 B 能根据父母的育种值估算儿子育种值

模型 C 能根据父母的育种值估算女儿育种值

另外构造配种预测仿真模型:

模型 A 能根据母猪、公猪育种值预测其交配后的生产值。

利用 A 类模型对样本库中的所有公猪、母猪进行虚拟组合交配运算,计算其所有可能的交配的仿真结果,再根据结果,决定最优的选配策略。具体有三个步骤。

1) 建立基础数据库

为了计算种猪繁殖性状的遗传力、重复力,各性状间的相关系数,种猪个体的育种值,特建立三个基础数据库。

① 母猪生产记录库

主要字段为:母猪耳号,生产胎次,分娩日期,与配公猪耳号,公猪品种,产仔数,初生窝重,20日窝重,断奶窝重,断奶头数,转栏数

② 种猪三代系谱记录库

主要字段为:种猪耳号,种猪品种,种猪性别,出生日期,父亲耳号,母亲耳号,祖父耳号,祖母耳号,外祖父耳号,外祖母耳号,备注(注明品系)

③ 现行生产种猪库

主要字段为:种猪耳号,种猪品种,产仔数育种值,初生窝重育种值,20日窝重育种值,断奶窝重育种值,断奶头数育种值,已生产胎次或交配次数,亲缘关系链(仅对母猪)

此库用于记录现有的生产种猪和这些种猪的生产性能状态。

2) 群体遗传参数的计算及个体育种值估算

① 遗传力的计算

遗传力是通过数量性状的表型值预测其育种值可靠性大小的比值,也可以认为是亲代与子代相似程度的指数,反映着亲代变异传递到子代的能力。此处遗传力的计算采用了盛志廉推导的混合亲缘系数公式:

遗传力

$$h^2 = r(HS) \sqrt{VA}$$

$$r(HS) = \frac{\sum df_w \sum SS_s - \sum df_s \sum SS_w}{\sum df_w \sum SS_s + \left\{ \sum (N_0 \cdot df_s - \sum df_s) \right\} \sum SS_s}$$

$$\sqrt{VA} = 0.25 \left[1 + \frac{\sum_i \sum_j^d n^2 - N}{\sum_i \sum_j^d n^2} - N \right]$$

其中 SS_w 为公猪内平方和; SS_s 为公猪间平方和; df_w 为公猪内自由度; df_s 为公猪间自由度; N_0 为公猪

子女加权平均数 = $\frac{1}{k-1} \left[\sum n_i \frac{n_i^2}{\sum n_i} \right]$; \sqrt{VA} 为混合家系亲缘系数; S 为公猪数; d 为与配母猪数。

由上公式天马大约克种猪繁殖性状的遗传力如表 1

② 重复力的计算

重复力采用公式 $t = \frac{MS_s - MS_w}{MS_s + (N_0 - 1) \cdot MS_w}$

其中 MS_s 是不同母猪之间各数量性状表型值的方差; MS_w 是不同胎次间各数量性状表型值的方差; N_0 是子女平均加权系数。

表 1

繁殖性状	产仔数	初生窝重	20日窝重	断奶窝重	断奶头数
遗传力	0.219	0.118	0.287	0.216	0.213
相关系数	0.071	0.037	0.092	0.069	0.068
亲缘系数	0.322	0.319	0.323	0.322	0.319
N_0	3.360	3.170	3.390	3.410	3.560
运算组数	44	42	43	43	45
运算记录数	75	58	79	59	83

计算出各性状重复力如表 2

表 2

繁殖性状	产仔数	初生窝重	20日窝重	断奶窝重	断奶头数
重复力	0.255	0.177	0.323	0.226	0.199
N_0	6.290	6.970	7.690	5.660	5.650
运算组数	34	41	16	35	34
运算记录数	214	286	123	198	192

③ 性状间的遗传相关和表型相关

采用运算公式如下:

$$\text{遗传相关系数 } V_A(x,y) = \frac{C_0V_b(xy)}{\sigma_g^2(x) \cdot \sigma_g^2(y)} \quad (x,y \text{ 为育种值})$$

$$\text{表型相关系数 } V_p(x,y) = \frac{C_0V(xy)}{\sigma^2(x) \cdot \sigma^2(y)}$$

其中: $C_0V_b(xy)$: 组间性状 x 和 y 育种值的协方差; $C_0V(xy)$: 组间性状 x 和 y 表型值的协方差;
 $\sigma_g^2(x)$, $\sigma_g^2(y)$: 组间性状 x (育种值, 表型值) 的方差; $\sigma^2(x)$, $\sigma^2(y)$: 组间性状 y (育种值, 表型值) 的方差。
 以公猪分组, 所求相关系数结果如下:

相关性状	产仔数 ~ 初生窝重	初生窝重 ~ 20日窝重	20日窝重 ~ 断奶窝重	断奶窝重 ~ 断奶头数	产仔数 ~ 断奶头数
遗传相关值	0.64770	0.66420	0.75730	0.81170	0.95930
表型相关值	0.30830	0.49820	0.69240	0.68370	0.32540

相关性状	产仔数 ~ 20日窝重	初生窝重 ~ 断奶窝重	产仔数 ~ 断奶窝重	20日窝重 ~ 断奶头数	初生窝重 ~ 断奶头数
遗传相关值	0.88470	0.75860	0.23630	0.72500	0.82990
表型相关值	0.28820	0.48550	0.23230	0.73470	0.57070

运算中公猪的分组数为: 27组, 参与运算总数据量为 459个。

④ 个体育种值的估算

由于猪是多胎动物,其个体育种值的估算采用多次度量值的均值进行计算,选用公式:

$$A = P + (P_k - P) \cdot \frac{k}{1 + (k-1)t} h^2 \quad (1)$$

其中 A 为个体某性状的育种值; P 是某性状的群体均值; P_k 是个体某性状 k 次度量均值; t 为重复力; h^2 为某种性状的遗传力。

3) 形成最优选配方案

选配是指如何选择具有较好的育种值的公猪和母猪进行交配以生产出兼具父母良好性状的后代。因此如能事先预测出交配后的结果再依此结果而有计划地制定选配方案,必将极大地缩短育种周期。

① 建模

显然,后代的繁殖性状与亲代之间不可否认地存在着一种因果关系,本课题,基于前述已求交配公猪,母猪个体性状育种值及生产值之样本库,采用回归分析法建立生产值预测模型:(简称 A 类模型,使用样本点 308 条)

	A 类预测模型	标准差	平均值
产仔数	$4.667 \times F_{1+} - 5.566 \times M_{1-} - 92.406$	1.220	10.4
初生窝重:	$6.643 \times F_{2+} - 7.910 \times M_{2-} - 184.85$	1.89	14.3
20日窝重:	$2.177 \times F_{3+} - 3.965 \times M_{3-} - 232.173$	5.339	46.6
断奶窝重:	$2.603 \times F_{4+} - 4.303 \times M_{4-} - 1017.85$	24.317	182.3
断奶头数:	$2.60 \times F_{5+} - 3.539 \times M_{5-} - 48.621$	1.091	9.8

用作选配的种猪中还有一类尚无生产记录的后备公猪和后备母猪,此时无法用(1)式估算其育种值。同理可求得这类后备种猪育种值的预测模型(依据其父母亲育种值)

后备公猪育种值预测模型(简称 B 类模型,使用样本点 20 个)

	B 类预测模型	标准差	平均值
产仔数育种值	$-0.543 \times F_{1-} - 0.014 \times M_{1-} - 15.546$	0.064	10.29
初生窝重育种值:	$-0.015 \times F_{2-} - 0.034 \times M_{2-} - 14.342$	0.068	14.13
20日窝重育种值:	$-0.028 \times F_{3+} - 0.001 \times M_{3+} - 46.732$	0.155	46.81
断奶窝重育种值:	$-0.502 \times F_{4-} - 0.214 \times M_{4-} - 303.903$	1.453	183.4
断奶头数育种值:	$-0.178 \times F_{5+} - 0.002 \times M_{5+} - 11.217$	0.055	9.83

后备母猪育种值预测模型(简称 C 类模型,使用样本点 81 个)

	C 类预测模型	标准差	平均值
产仔数育种值	$-0.015 \times F_{1+} - 0.103 \times M_{1+} - 9.192$	0.202	10.32
初生窝重育种值:	$-0.166 \times F_{2+} - 0.281 \times M_{2+} - 12.166$	0.156	14.37
20日窝重育种值:	$0.047 \times F_{3-} - 0.080 \times M_{3+} - 47.134$	1.009	45.94
断奶窝重育种值:	$0.469 \times F_{4-} - 0.021 \times M_{4+} - 94.922$	5.905	182.9
断奶头数育种值:	$0.246 \times F_{5-} - 0.077 \times M_{5+} - 7.782$	0.286	9.82

式中 F_i 为第 i 性状父亲的育种值; M_i 为第 i 性状母亲的育种值。

② 完善生产种猪数据库的数据

在初始的生产种猪数据库中,只有种猪耳号、性别等信息。此时,借助于系谱数据库及预测模型将种猪各性状的育种值求出,填入生产种猪数据库中。

根据种猪是否有繁殖生产历史,分别采用 1)或借助于系谱数据库采用 B C 类模型,求出每头种猪各性状的育种值,填入生产种猪数据库中,使其数据完善。

③ 最优选配方案的生成

将每头公猪与所有在亲缘关系链中没有此公猪耳号的母猪虚拟交配,借助 A 类预测模型,仿真出生产值。再按单性状仿真结果进行排序,分别求得关于产仔数,初生窝重,20日窝重,断奶窝重和断奶头数的最优选配方案。在实际选育工作中,还要对多个性状综合选优,这时需先构造综合选择指数公式,而后代入各交配对的各性状仿真结果。再对综合选择指数进行排序,得最优选配方案(相对多性状综合选优而言)。

本课题采用偏回归系数计算综合选择指数法,求得综合指数计算公式为

$$H = 0.08883F_{1+} + 0.00356F_{2+} + 0.14738F_{3+} + 0.044753F_{4+} + 0.15848F_{5+}$$

式中 F_i 为性状 i 的仿真结果。

3 饲料配方问题

饲料配方问题长期以来多采用线性规划模型去处理,该类模型虽然在实际应用中取得了不少成效,但由于地区不同,饲料的营养成分含量及各种相关因素的变化,特别存在如下两个问题难以解决。

① 饲料的适口性问题

求得的配方虽然在营养标准方面可达到一定的标准并在单价成本上也能得到最优的结果,但此类配方往往在适口性方面差,动物不爱吃,这好比人类不能光吃满足各种营养标准的营养液(可能营养价值也高,也省钱),因此,其配方实际上难适用。

② 为解决①,常常在饲料的某些原料上加上重量约束,以确保适口性和某些别的因素,但这往往又造成线性规划问题无解。

本文将线性规划中的硬约束同时转化为软约束后,采用的是多目标规划模型。

假设所需的营养有 N 个指标,可供选择的原料有 M 种,问题为:求一组原料 $x_i (i = 1, 2, \dots, M)$ 极小化目标函数:

$$\min S = g_1(d_{N+1}^+, d_{N+1}^-, \dots, d_{N+k}^+, d_{N+k}^-, w) + g_2(d_{N+k+1}^+, d_{N+k+1}^-, w) + g_3(d_1^+, d_2^+, d_1^-, d_2^-, w) + g_4(d_3^+, d_4^+, d_3^-, d_4^-, w) + g_5(d_5^+, \dots, d_N^+, d_5^-, \dots, d_N^-, w) + g_6(d_{N+k+2}^+, d_{N+k+2}^-, w)$$

满足约束条件:

$$\text{原料营养成分的约束: } \sum_{j=1}^M a_{ij}x_j + d_i^- - d_i^+ = b_i \quad i = 1, 2, \dots, N$$

$$\text{原料使用量的约束: } x_{u_j} + d_{N+1}^- - d_{N+1}^+ = p_j \quad j = 1, 2, \dots, k \quad u_j \in \{1, 2, \dots, M\}$$

$$\text{原料配比总和的约束: } \sum_{j=1}^M x_j + d_{N+k+1}^- - d_{N+k+1}^+ = E$$

$$\text{价格函数的约束: } \sum_{j=1}^M c_j x_j + d_{N+k+2}^- - d_{N+k+2}^+ = F$$

式中: x_j 为第 j 种原料的使用量 ($j = 1, 2, \dots, M$);

a_{ij} 为第 j 种单位原料含有第 i 种营养成分的量 ($j = 1, 2, \dots, M; i = 1, 2, \dots, N$);

c_j 为第 j 种原料的单位价格 ($j = 1, 2, \dots, M$); b_i 为第 i 种营养成分的目标值 ($i = 1, 2, \dots, N$);

p_j 为第 j 种原料的极限使用量 ($j = 1, 2, \dots, k$); E 为所有原料配比总和的值; F 为配方价格的预定值;

d_i^+, d_i^- 为第 i 个营养成分指标的正负偏离变量 ($i = 1, 2, \dots, N$);

d_{N+j}^+, d_{N+j}^- 为第 j 种原料使用量的正负偏离变量 ($j = 1, 2, \dots, k; j \in \{1, 2, \dots, M\}$);

d_{N+k+1}^+, d_{N+k+1}^- 为原料配比(配方比例)总和的正负偏离变量;

d_{N+k+2}^+, d_{N+k+2}^- 为配方价格的正负偏离变量

g_1 为 k 种原料使用量的正负偏离变量的目标函数; g_2 为原料配比总和的正负偏离变量的目标函数; g_3 为消化能和粗蛋白正负偏离变量的目标函数; g_4 为钙和磷的正负偏离变量的目标函数; g_5 为其它营养指标的正负偏离变量的目标函数, 如赖氨酸、苏氨酸、粗纤维等; g_6 为配方价格的正负偏离变量的目标函数。

针对断奶至 30 公斤体重阶段的猪的饲料问题, 应用多目标规划模型(略), 按国家营养标准及当地可用原料: 玉米、大麦、小麦麸、大豆饼、鱼粉、槐叶粉、骨粉、石粉的营养含量进行处理后, 得出一配方(略), 其营养成份:

消化能 3201.97 兆卡/吨 可消化蛋白质 115609.46 克/吨
粗蛋白质 150494.46 克/吨 赖氨酸 7252.449 克/吨
钙 4453.29 克/吨 磷 3998.16 克/吨
纤维素 344855.86 克/吨

此配方于 1993 年在福建天马种猪场进行应用对照试验。共有 4 个配方分组进行试验比较, 另三个为该场过去常用的不同配方, 经试验, 结果如下:

试验结果

(单位: 头, 天, 公斤, 元)

项 组	头数	天数	始重	末重	总增量	日增量	总用料	料肉比	每公斤 增重成本
试验配方组	20	12	22.7	30.2	150.0	0.63	309.2	2.06	2.13
参考配方 1 组	20	12	20.0	26.8	95.5	0.40	247.5	2.59	2.51
参考配方 2 组	10	12	21.3	26.2	49.0	0.41	134.8	2.75	2.49
参考配方 3 组	10	12	21.4	26.2	47.5	0.40	131.8	2.77	2.32

显然, 采用多目标规划得出的试验配方在试验中, 日增量、料肉比、及每公斤增重成本等几个重要参数都优于其它参照组。

4 结束语

本文从系统工程的角度出发探讨了种猪选育中的若干问题, 特别引进了建模、计算机仿真、数据库等多种技术, 力求在传统的方法和观念上有所突破。该研究结果已在福建天马种猪场投入使用, 并已取得较为显著的经济效益和社会效益。但由于历史数据的条件限制, 在模型上尚未能考虑季节的影响和生产胎次的影响因素。在今后逐渐完善数据的过程中将进一步优化模型。

参考文献

- 1 吴仲贤著. 统计遗传学. 科学出版社, 1979.
- 2 高仁之编著. 数量遗传学. 四川大学出版社, 1990.
- 3 裴新澍编著. 数理遗传与育种. 上海科技出版社, 1987.
- 4 陈瑶生. 遗传参数估计原理初探——亲本间亲缘相关时的参数估计方法. 遗传学报, 1990, 17(2): 116 - 129.
- 5 周士谔, 李子先等. 数量遗传学中一种新的求综合性状的方法. 遗传学报, 1989, 16(4): 267- 275.
- 6 J. A. 斯普里特, G. C. 范斯蒂恩斯特著. 计算机辅助建模和仿真. 科学出版社, 1991.
- 7 Sang M. Lee. 决策分析的目标规划. 清华大学出版社, 1986.
- 8 王殿选, 李至时. 多目标决策的对比系数法. 系统工程理论与实践, 1988, 18(4).