

综合选择指数的决策分析*

王丽波¹, 孙洪罡², 刘璐¹, 郭满才¹, 袁志发¹

(1 西北农林科技大学 生命科学学院, 陕西 杨凌 712100;

2 中国农业大学 理学院, 北京 100083)

[摘要] 建立了综合选择指数的通径分析化模型, 即综合选择指数的相关遗传力分解模型, 提出了综合选择指数的决策分析方法, 从而使综合选择指数的路径信息最优化。提出了确定主选性状、辅助性状和限制性性状的方法, 便于育种者根据个体的性状表现进行选择。

[关键词] 综合选择指数; 通径分析; 相关遗传力; 决策分析

[中图分类号] S11⁺ 4; O212

[文献标识码] A

[文章编号] 1671-9387(2005)10-0094-03

Hazel 于 1943 年提出了综合选择指数的概念, 从而在实际中实现了对个体综合选择指数值的选择^[1]。1988 年袁志发^[2]提出了综合选择指数相关遗传进展分解原理, 实现了复合育种值相关遗传进展的剖分, 但仍未解决选择的最优路径问题。2001 年袁志发^[3]又提出了通径分析的决策分析, 以解决通径分析中的最优路径问题。本研究在前人研究基础上, 建立了综合选择指数的通径分析化模型, 进而提出综合选择指数的决策分析方法, 即综合选择指数的最优路径分析, 给出了确定主选性状、辅助性状和限制性性状的方法, 从而使育种者能从预选个体的性状表现进行直接选择。

1 综合选择指数的通径分析化模型

选择性状为 $x_1, x_2, \dots, x_m$, 记为 $X = \{x_1, x_2, \dots, x_m\}^T \sim N_m(\mu, \Sigma_p)$; X 的育种值为 $g = \{g_1, g_2, \dots, g_m\}^T \sim N_m(\mu, \Sigma_g)$; X 的环境离差值为 $e = \{e_1, e_2, \dots, e_m\}^T \sim N_m(0, \Sigma_e)$, $X = g + e$ 。设 g 与 e 独立, 则有 $\Sigma_p = \Sigma_g + \Sigma_e$, 其中 Σ_p , Σ_g 和 Σ_e 分别为 X 的表型协方差阵、遗传协方差阵和环境协方差阵。

设列向量 $a = (a_1, a_2, \dots, a_m)^T$, 已知的经济权重列向量 $w = (w_1, w_2, \dots, w_m)^T$ 。Hazel 的综合选择指数和复合育种值分别为:

$$I = a_1 x_1 + a_2 x_2 + \dots + a_m x_m = a^T X \quad (1)$$

$$H = w_1 g_1 + w_2 g_2 + \dots + w_m g_m = w^T g \quad (2)$$

I 的表型方差、遗传方差分别为:

$$\sigma_{I_p}^2 = a^T \Sigma_p a, \sigma_{I_g}^2 = a^T \Sigma_g a \quad (3)$$

H 的方差为:

$$\sigma_H^2 = w^T \Sigma_g w \quad (4)$$

I 与 H 的协方差为:

$$\text{Cov}(I, H) = a^T \Sigma_g w = w^T \Sigma_g a \quad (5)$$

I 与 H 的相关系数为:

$$r_{IH} = \frac{a^T \Sigma_g w}{\sqrt{(a^T \Sigma_p a)(w^T \Sigma_g w)}} \quad (6)$$

Hazel 确定 a 的原则为 $r_{IH} = \max$ 。故可化为如下条件极值问题:

$$\begin{cases} a^T \Sigma_g w = \max \\ a^T \Sigma_p a = 1 \end{cases} \quad (7)$$

应用拉格朗日乘数法作辅助函数

$$f(a) = a^T \Sigma_g w - \frac{\lambda}{2} (a^T \Sigma_p a - 1) \quad (8)$$

令 $\frac{df}{da} = 0$, 若只考虑 a 中各分量的比值, 则 a 可由下列线性方程组确定:

$$\Sigma_p a = \Sigma_g w \quad (9)$$

式中, $\Sigma_p = (\sigma_{p_{ij}})_{m \times m}$; $\Sigma_g = (\sigma_{g_{ij}})_{m \times m}$

由于 g 与 e 独立, 故:

$$\text{Cov}(x_i, H) = \text{Cov}(g_i + e_i, \sum_{j=1}^m w_j g_j) = \sum_{j=1}^m w_j \sigma_{g_{ij}} \quad (10)$$

所以式(9)可写为:

$$\Sigma_p a = \text{Cov}(X, H) \quad (11)$$

* [收稿日期] 2005-01-24

[作者简介] 王丽波(1980-), 女, 黑龙江大庆人, 在读硕士, 主要从事农业应用数学研究。

[通讯作者] 袁志发(1938-), 男, 陕西渭南人, 教授, 博士生导师, 主要从事农业应用数学研究。

式中, $\text{Cov}(X, H) = (\text{Cov}(x_1, H) \text{ Cov}(x_2, H) \dots \text{Cov}(x_m, H))^T$ (12)

对于式(11)表示的方程两边分别除以 $\sigma_{p1} \sigma_H$, $\sigma_{p2} \sigma_H, \dots, \sigma_{pm} \sigma_H$, 并令 $c_j = \sigma_{pj} a_j / \sigma_H$, 则式(11)变为:

$$R_p c = R_{XH} \quad (13)$$

其中 $R_p = (r_{pij})_{m \times m}$ 为 $x_1, x_2, \dots, x_m$ 的表型相关阵, $c = (c_1, c_2, \dots, c_m)^T$ 为途径系数, $R_{XH} = (r_{x_1H}, r_{x_2H}, \dots, r_{x_mH})^T$ 为 X 与 H 的相关阵。

式(13)即是综合选择指数的途径分析化模型, 即育种目标与性状间的相关系数的分解原理。该模型将选择性状 x_i 与复合育种值 H 的相关系数剖分为 x_i 对 H 的直接贡献 c_i 与 x_i 通过 $x_j (j \neq i)$ 表型相关对 H 的间接贡献 $r_{pij} c_j$ 之和。

该模型不同于以往的途径分析化模型, 以往的途径分析化模型两边都是表型相关, 或者都是遗传相关, 笔者认为那样的途径分析化模型应用在选择指数法上是不贴切的, 不符合选择的本质, 因为运用选择指数法选择的本质是通过选择表型让复合育种值获得最大的进展。这里给出的综合选择指数的途径分析化模型是混合相关系数模型, 左边是表型相关, 右边是表型与遗传相关, 可以认为这是一种新的途径分析化模型。事实上由于 g 与 e 独立, 故有:

$$r_{x_iH} = \frac{\text{Cov}(x_i, H)}{\sigma_{x_i} \sigma_H} = \frac{\text{Cov}(g_i + e_i, H)}{\sigma_{x_i} \sigma_H} = \frac{\text{Cov}(g_i, H)}{\sigma_{x_i} \sigma_H} \quad (14)$$

又由相关遗传力的定义知, x 和 y 的相关遗传力定义为

$$h_{xy} = \frac{\text{Cov}(x, y)}{\sigma_{px} \sigma_{py}} \quad (15)$$

即性状 x_i 与复合育种值 H 的相关遗传力为:

$$h_{x_iH} = \frac{\text{Cov}(x_i, H)}{\sigma_{px_i} \sigma_{pH}} = \frac{\text{Cov}(x_i, H)}{\sigma_{px_i} \sigma_H} \quad (16)$$

比较式(14)和(16)可知:

$$r_{x_iH} = h_{x_iH} \quad (17)$$

即性状 x_i 与复合育种值 H 的表型相关系数等于 x_i 与 H 的相关遗传力, 故式(13)可变为:

$$R_p c = h_{XH} \quad (18)$$

其中 $h_{XH} = (h_{x_1H}, h_{x_2H}, \dots, h_{x_mH})^T$, 为性状与复合育种值的相关遗传力矩阵。

式(18)说明综合选择指数的途径分析化模型的实质, 是被选择的性状与复合育种值的相关遗传力的分解模型。

在途径分析中, $x_1, x_2, \dots, x_m$ 对 H 的决定系数

为:

$$R^2 = \sum_{i=1}^m c_i^2 + 2 \sum_{i=1}^m c_i r_{pij} c_j = \sum_{i=1}^m R_i^2 + \sum_{i=1}^m R_{ij} \quad (19)$$

式中, $R_i^2 = c_i^2$, 为 x_i 对 H 的直接决定系数; $R_{ij} = 2c_i r_{pij} c_j$, 为相关路 $x_i \leftarrow x_j$ 对 H 的相关决定系数 (亦称间接决定系数), 显然 $R_{ij} = R_{ji}$ 。途径分析给出了各个性状 x_i 对复合育种值 H 的直接作用和间接作用, 但没有解决相关性状 $x_1, x_2, \dots, x_m$ 与复合育种值 H 的最优路径问题, 即没有给出主要利用、限制和协同哪些性状才能使复合育种值得到最大进展的方法。

2 综合选择指数的决策分析

所谓性状 x_i 对复合育种值 H 的决策系数 $R_{(i)}$, 表示 x_i 对 H 的综合决定能力, 即:

$$R_{(i)} = R_i^2 + \sum R_{ij} \quad (20)$$

式中, $R_{(i)}$ 表示性状 x_i 对复合育种值 H 的综合决定作用, 既包括 x_i 通过直接路 $x_i \leftarrow H$ 对 H 的直接作用, 也包括 x_i 通过 $m-1$ 条表型相关路 $x_i \leftarrow x_j (j \neq i)$ 对 H 的间接作用。将 $R_{(1)}, R_{(2)}, \dots, R_{(m)}$ 由大到小排序, 设决策系数中最大者为 $R_{(i)}$, 若 $R_{(i)} > 0$, 说明 x_i 对 H 的综合促进作用最大, 故 x_i 为主选性状, 其他决策系数大于 0 的性状可作为辅助性状; 如果决策系数 $R_{(k)} < 0$, 说明 x_k 对 H 起限制作用, 故 x_k 为限制性性状, 其中最小决策系数对应的是主要限制性性状。根据各 x_i 的决策系数排序就可以将选择指数具体化、形象化, 即明确了在 $x_1, x_2, \dots, x_m$ 中应依靠、限制和保持的性状, 这样就找到了使复合育种值进展最大的选择路线。

3 应用举例及其分析

在鸡的选择育种中, Kempthorne 等用 5 个性状制订了鸡的选择指数, 其中 4 个选择指数是: x_1 表示成年体重; x_2 表示卵重; x_3 表示初产日龄; x_4 表示至 72 周的产卵数的 $1/3^{[1]}$ 。其表型方差阵 Σ_p , 遗传协方差阵 Σ_g 分别为:

$$\Sigma_p = \begin{bmatrix} 34 & 6.775\ 04 & 0 & -2.198\ 81 \\ 6.775\ 04 & 21.6 & 0 & -1.752\ 56 \\ 0 & 0 & 13 & -10.877\ 09 \\ -2.198\ 8 & -1.752\ 6 & -10.877\ 09 & -10.877\ 09 \end{bmatrix}$$

$$\Sigma_g = \begin{bmatrix} 15.30 & 3.855\ 03 & 0 & 1.978\ 25 \\ 3.855\ 03 & 10.80 & 0 & 1.662\ 60 \\ 0 & 0 & 5.20 & -3.845\ 67 \\ 1.978\ 25 & 1.662\ 60 & -3.845\ 67 & 11.376 \end{bmatrix}$$

各性状的经济权重 $\omega = (-2.50, 7.20, 0, 0.0805)^T$ 。综合选择指数的通径分析和决策分析结果见表1。

表1 综合选择指数的通径分析与决策分析结果

Table 1 Path analysis and decision analysis on comprehensive selection index

性状 Character	通径系数 Path coefficient	表型相关系数 Phenotypic related coefficient	决策系数 Decision coefficient
x_1	$c_1 = -0.0517$	$r_{p12} = 0.25$ $r_{p13} = 0$ $r_{p14} = -0.05$	$R_{(1)} = -0.06534$
x_2	$c_2 = 0.4455$	$r_{p23} = 0$ $r_{p24} = -0.05$	$R_{(2)} = 0.1710$
x_3	$c_3 = -0.1140$	$r_{p34} = -0.40$	$R_{(3)} = 0.0458$
x_4	$c_4 = 0.3596$		$R_{(4)} = 0.0640$

由表1可知, $R_{(2)} > R_{(4)} > R_{(3)} > R_{(1)}$, 其中 $R_{(2)}$ 最大且为正, 说明卵重在4个相关性状的关系网中对复合育种值的综合促进作用最大, 故卵重为主选性状; $R_{(4)}$ 和 $R_{(3)}$ 排在 $R_{(2)}$ 之后也为正, 说明至72周的产卵数的1/3和初产日龄对复合育种值也起促进作用, 但没有卵重的促进作用大, 故为辅助选择性状; $R_{(1)}$ 最小且为负, 说明成年体重是主要限制性状, 选择时应加以限制。上述结论与实际育种结果基本上相符合。

4 结 论

本研究在Hazel的综合选择指数基础上, 建立了综合选择指数的通径分析化模型, 其本质是综合

选择指数的相关遗传力分解模型。在此基础上给出了综合选择指数的决策分析模型, 即综合选择指数的最优路径分析, 各性状的决策系数表示其对复合育种值的综合决定作用。在实际应用中算出各性状的决策系数后, 可将其从大到小排序, 决策系数最大且为正值的, 对应的是主选性状; 其他决策系数为正的性状可以作为辅助性状; 决策系数为负的对的是限制性状, 其中决策系数最小的, 对应的是主要限制性状。这样可使综合选择指数更直观化、形象化, 便于进行实际育种。另外这种决策分析法可以推广到约束选择指数、最宜选择指数等其他类型的选择指数分析中。

[参考文献]

- [1] 吴仲贤 统计遗传学[M] 北京: 科学出版社, 1979
- [2] 袁志发 选择指数与相关遗传进展的分解原理[J] 西北农业大学学报, 1988, 16(4): 31- 34
- [3] 袁志发 决策系数- 通径分析中的决策指标[J] 西北农林科技大学学报(自然科学版), 2001, 29(5): 131- 133

Decision analysis on comprehensive selection index

WANG Li-bo¹, SUN Hong-gang², LIU Lu¹, GUO Man-cai¹, YUAN Zhi-fa¹

(1 College of Life Sciences, Northwest A & F University, Yangling, Shaanxi 712100, China;

2 College of Sciences, China Agricultural University, Beijing 100083, China)

Abstract: This paper establishes the path analytical model of comprehensive selection index, that is analyzing model on correlated heritability of comprehensive selection index. Decision analysis on comprehensive selection index is advanced to optimize its path information. The paper also shows ways to ascertain main selective, assistant and restrained trait, which allow breeders to select by individual phenotypes

Key words: comprehensive selection index; path analysis; correlated heritability; decision analysis