

黄山松种群拓广遗传参数的研究

童再康 范义荣

(浙江林学院, 临安 311300)

摘要 利用黄山松种源试验两个试验点的材料, 运用多元分析方法所定义的估算拓广遗传参数方法, 分别从协方差阵和相关阵出发估算了黄山松种群的拓广表型方差、拓广遗传方差、拓广遗传力和拓广遗传通径系数。结果表明使用全部特征根估算拓广遗传参数是不合适的, 95%左右的累计贡献率最为适宜。此时所估算的拓广遗传力以协方差阵为基础的是0.32~0.48, 以相关阵为基础的是0.66~0.87, 且从相关阵出发所估算的拓广遗传参数可比性强。依据初步分析结果, 对估算方法和实践意义进行了较为详细的讨论。

关键词 黄山松; 遗传分析; 遗传变异; 参数估计; 多元分析

中图分类号 S791.259

定义数量性状的遗传力、遗传方差、遗传变异系数和遗传相关系数等遗传参数以往都是从单个性状出发的^[1,2]。它们只能应用于单个性状的选择, 遗传增益估算等, 也只能从单个性状上来衡量种群的遗传分化, 而难以从各类性状上综合性地把握种群的遗传分化程度。然而, 随着生产的发展和人们需要的提高, 林木育种工作者越来越重视生长、适应性、抗性和品质等多类性状同时改良的问题, 并开展了大量的多性状综合选择工作^[1,3]。这种多性状的综合育种工作须以多性状综合遗传参数作指导, 以此来估测育种效果, 制定合理的育种策略。戴君惕等(1990)利用多元统计方法定义了若干拓广遗传参数, 并用棉花和紫花苜蓿两个实例进行了估算^[4]。本文试图对黄山松种源试验数据分别从原始数据和标准化数据出发估算拓广遗传方差、拓广遗传力和拓广遗传通径系数, 以了解黄山松种群在多维空间的遗传变异状况, 为黄山松的遗传改良提供科学依据。

1 材料与方 法

1.1 材 料

材料来自黄山松种源试验浙江临安和泰顺两个试验点1981年营造的试验林(8年生)。试验概况参见文献5。调查的性状有: 树高、胸径(或地径)、材积(材积指数)、枝数、冠幅、分枝角、结实率和虫害指数共8个, 可以归纳为生长、分枝、结实和抗性4类。

1.2 数 学 方 法

收稿日期: 1991-05-22, 1992-03-23收到修改稿

设两组变量 y_1 (m 维) 和 y_2 (q 维) 有联合协方差阵 Σ :

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

则
$$\prod_{k=1}^m \lambda_k^2 = \frac{|\Sigma_{21} \Sigma_{12}|}{|\Sigma_{11}| \cdot |\Sigma_{22}|}$$

设 $m \leq q$, 定义拓广相关系数 $\tilde{\rho}$ 为 m 个典范相关系数的几何平均值:

$$\tilde{\rho} = \left(\prod_{k=1}^m \lambda_k \right)^{\frac{1}{m}} = \left(\frac{|\Sigma_{21} \Sigma_{12}|}{|\Sigma_{11}| |\Sigma_{22}|} \right)^{1/2m}$$

如果设 $m = q$, 则定义拓广方差和拓广协方差如下:

$$\tilde{V}_1 = |\Sigma_{11}|^{1/m} = \left(\prod_{k=1}^m \lambda_k \right)^{1/m}$$

$$\tilde{V}_2 = |\Sigma_{22}|^{1/m} = \left(\prod_{k=1}^m \lambda_k^{-1} \right)^{1/m}$$

$$\widetilde{COV} = |\Sigma_{12}|^{1/m}$$

这果 $\lambda_k, \lambda_k^{-1}$ 分为 Σ_{11} 和 Σ_{22} 的第 k 个特征根, 则:

$$\tilde{\rho} = \frac{\widetilde{COV}}{(\tilde{V}_1 \tilde{V}_2)^{\frac{1}{2}}}$$

其形式同一元情况相似。在考虑 m 个数量性状时, 有 $p_j = g_j + e_j$ ($j = 1, 2, 3, \dots, m$), 假设 $\widetilde{COV}(g_j, e_j) = 0$, 则 p 和 g 向量组的聚合协方差阵为:

$$\Sigma = \begin{pmatrix} \Sigma_p & \Sigma_{p,g} \\ \Sigma_{g,p} & \Sigma_g \end{pmatrix} = \begin{pmatrix} \Sigma_p & \Sigma_g \\ \Sigma_g & \Sigma_p \end{pmatrix}$$

即拓广表型方差:

$$\tilde{p} = |\Sigma_p|^{1/m} = \left(\prod_{k=1}^m \lambda_k \right)^{1/m}$$

拓广遗传方差:

$$\tilde{g} = |\Sigma_g|^{1/m} = \left(\prod_{k=1}^m \lambda_k^{-1} \right)^{1/m}$$

拓广遗传通径系数:

$$\begin{aligned} \tilde{p}_{(p,g)} &= |\Sigma_{p,g}, \Sigma_{g,p}|^{1/2m} / (|\Sigma_p| |\Sigma_g|)^{1/2m} \\ &= [|\Sigma_g|^2 / |\Sigma_p| \cdot |\Sigma_g|^{1/2m}] / (|\Sigma_p| |\Sigma_g|)^{1/2m} = \sqrt{\tilde{g} / \tilde{p}} \end{aligned}$$

拓广遗传力:

$$\tilde{h}^2 = \tilde{\rho}_{(p,g)}^2 = \tilde{g} / \tilde{p}$$

2 结果与分析

2.1 各数量性状的遗传参数的估算

利用两个试验点试验林材料估算各性状的遗传力、遗传方差、遗传变异系数和遗传相关系数等，结果列表 1 中。由表 1 可见黄山松主要生长性状遗传力较高，但在试验点间有一定差异。这同杉木、马尾松上所得的结果相似。分枝性状中，冠幅的遗传力较高(0.62~0.88)，而侧枝角的遗传力最低(0~0.13)，即侧枝角是 1 个很不稳定的性状，易受环境条件的影响，其在群体间遗传变异不大。从遗传变异系数看，除侧枝角和分枝数的变异较小外，其余性状都较大，可见开展黄山松遗传改良的潜力很大。从遗传相关分析知，生长、分枝性状间有很高的相关，同结实性状有中等程度相关，而同抗性性状无多大关系。可见抗虫性与生长性状受不同的基因控制，即选择生长快、抗性强的群体是可能的。

表 1 各试验点试验林遗传参数估值

Table 1 Estimated values of genetic parameters of test plantations in 2 sites

性 状	临 安 试 点			泰 顺 试 点		
	遗传方差	遗传力	遗传变异系数 (%)	遗传方差	遗传力	遗传变异系数 (%)
树 高	167.92	0.77	7.98	166.71	0.87	17.44
胸径(或地径)	0.0621	0.71	18.33	0.4078	0.90	17.84
材 积	4.34×10^{-9}	0.76	21.11	2.67×10^{-7}	0.90	48.13
分 枝 数	1.07	0.38	4.46	2.71	0.76	9.10
侧 枝 角	0	0	0	0.43	0.13	0.97
冠 幅	110.61	0.62	6.77	183.64	0.88	15.39
结果指数	57.10	0.69	48.48	17.91	0.79	103.56
虫害指数	—	—	—	0.1637	0.85	40.68

注：遗传相关系数表中略

2.2 多性状的拓广遗传参数的估算

多性状的拓广遗传参数分别从原始数据和标准化数据出发，即分别从协方差阵和相关阵出发估算。先计算各矩阵的特征根，并计算累计贡献率，结果列于表 2 和表 3 中。

表 2 协方差阵的特征根与累计贡献率

Table 2 Eigen values and cumulative contribution rate of covariance matrix

特 征 根 数		1	2	3	4	5	6	7	8
临	表 特 征 根	828.47	185.56	45.32	3.41	0.028	-7.9×10^{-8}		
	型 累 计 贡 献 率 (%)	77.95	95.41	99.68	99.99	100			
安	基 因 型 特 征 根	281.41	57.55	0.74	6.86×10^{-8}	-4.1×10^{-3}	-2.94		
	累 计 贡 献 率 (%)	82.84	99.78	100					
泰	表 特 征 根	579.77	41.96	24.01	10.86	3.48	0.24	0.05	-4.42×10^{-7}
	型 累 计 贡 献 率 (%)	87.8	94.1	97.8	99.40	99.96	100		
顺	基 因 型 特 征 根	351.58	15.65	4.42	0.69	0.004	-2.5×10^{-7}	-0.02	-0.39
	累 计 贡 献 率 (%)	94.4	98.6	99.8	99.97	100			

表3 相关阵的特征根与累计贡献率

Table 3 Eigen values and cumulative contribution rate of correlation matrix

特征根数		1	2	3	4	5	6	7	8
临 型	特征根	4.37	1.00	0.28	0.22	0.11	0.25		
	累计贡献率 (%)	72.70	89.47	94.09	97.80	99.58	100		
安 型	特征根	5.13	1.03	0.15	8.0×10^{-5}	-0.13	-0.18		
	累计贡献率 (%)	81.24	97.59	99.99	100				
泰 型	特征根	4.55	1.18	1.00	0.76	0.34	0.09	0.05	0.02
	累计贡献率 (%)	56.90	71.70	84.20	93.70	98.00	99.10	99.70	100
顺 型	特征根	5.70	1.57	0.78	0.11	0.02	-7.0×10^{-3}	-0.02	-0.15
	累计贡献率 (%)	69.70	88.90	98.40	99.70	100			

由表2知, λ_1 和 λ_2 的累计贡献率达95%左右, 且基因型较表型更为集中, 其所对应的第1, 2典范变量几乎包括各单性状的信息。同样, 在表3中, $\lambda_1, \lambda_2, \lambda_3$ 和 λ_4 的累计贡献率达95%。同协方差阵相比, 相关阵的信息量稍分散些, 但两者的第1典范变量集中反映树高、胸径等生长量性状。依据累计贡献率大小分别估算2种矩阵的各种拓广遗传参数, 结果见表4。

表4 不同组合特征根的拓广遗传参数估值

Table 4 Estimated values of generalized genetic parameter of each combination of eigen values

组 合	协 方 阵					相 关 阵					
	累计贡献率 (%)	\tilde{p}	\tilde{g}	\tilde{h}^2	$\tilde{\rho}$	累计贡献率 (%)	\tilde{p}	\tilde{g}	\tilde{h}^2	$\tilde{\rho}$	
临	λ_1	77.95~82.84	828.47	281.41	0.34	0.58	72.77~81.24	4.36	5.12	1.17	1.08
	$\lambda_1\lambda_2$	95.41~99.78	392.08	127.26	0.32	0.57	89.47~97.59	2.09	2.30	1.10	1.05
安	$\lambda_1\lambda_2\lambda_3$	99.68~100	190.99	22.90	0.12	0.35	94.09~99.99	1.07	0.93	0.87	0.90
	$\lambda_1\lambda_2\lambda_3\lambda_4$	99.99~100	69.84	0.17	0.002	0.005	97.80~100	0.72	0.14	0.19	0.44
泰	λ_1	87.80~94.40	579.77	351.58	0.61	0.78	56.90~69.70	4.55	5.70	1.25	1.12
	$\lambda_1\lambda_2$	94.10~98.60	155.97	74.18	0.48	0.69	71.70~88.90	2.32	2.99	1.19	1.14
顺	$\lambda_1\lambda_2\lambda_3$	97.80~99.80	83.59	28.97	0.35	0.59	84.20~98.40	1.75	1.91	1.09	1.05
	$\lambda_1\lambda_2\lambda_3\lambda_4$	99.40~99.97	50.19	11.38	0.23	0.48	93.40~99.70	1.42	0.94	0.66	0.81
	$\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5$	99.96~100	29.43	2.32	0.08	0.28	98.00~100	1.07	0.43	0.40	0.63

由表4知, 随累计贡献率增加, 拓广遗传、表型方差均显著下降。尤其当累计贡献率超过95%以后, 各拓广遗传参数值下降更显著, 以至失真。当使用全部 λ 时, 其值有所回升。可见, 在估算拓广遗传参数时, 累计贡献率是一个关键的因子, 且有一定的最适值。由表中看, 不管表型还是基因型或者协方差阵还是相关阵, 95%左右的累计贡献率对本试验是最适的。从两种矩阵比较看, 相近的累计贡献率时, 以相关阵估算的拓广遗传力高, 接近于单性状的

情形, 而且各参数估值(特别是方差)在试点间有可比性。从各试点看, 其从相关阵出发的拓广遗传力与单性状中多数性状的遗传力相似。两试点间的差异由立地和性状组合上的差异综合引起。

3 讨论

3.1 从多元总体出发, 利用典型相关与主成分分析进行多元遗传分析, 使传统的单元的遗传分析向多元延伸, 估算各种拓广遗传参数, 其意义可以从几何的角度来理解。在 n 维空间中, 协方差阵的行列式值 $|\Sigma|$ 是以 Σ 各特征根 $(\lambda_1, \dots, \lambda_n)$ 为棱长的 n 维超棱柱体的体积, 以此来作为该多元总体的测度^[4]。同样 $|\Sigma|^{1/n}$ 则是这个多元总体每一维向的平均变异度, 那么 $\tilde{\rho} = \frac{|\Sigma_1|^{1/n}}{|\Sigma_2|^{1/n}}$

则是两个多元总体重叠度的测度, 可见 \tilde{h}^2 是多元总体表型平均变异中遗传平均变异的份量。

3.2 多元总体中各元的测量单位影响该总体平均变异度值, 杨德等(1982)在多性状的典范性状研究中认为原始数据的单位影响典范性状值, 应考虑标准化数据^[7]。在典范分析、主成分分析中数据经 $(x_i - \bar{x})/s$ 标准化, 即从相关阵 (R) , 而不是从协方差阵 (Σ) 出发, 这样可以消除测量单位的影响^[6]。当使用标准化数据时, 表型协方差阵 (Σ_p) 就成了表型相关阵 (R_p) , 而以育种值为基础的遗传协方差阵 (Σ_g) 就成了遗传相关阵 (R_g) 。杨德等(1982, 1983)认为典范性状不能从 R_g 出发, 而应从相关遗传力阵 (H_g) 出发^[7,8], 即应用表型的标准差标准化遗传方差(或协方差)。然而仅考虑 Σ_g 的标准化时, 应以遗传标准差进行标准化。从黄山松种群的估算数据看, 从标准化的数据出发, \tilde{p} , \tilde{g} 可比性强, \tilde{h}^2 值高。

3.3 多元总体的总变异度是以 Σ 的各特征根为棱长的 n 维超棱柱体的体积来度量的, Σ 是一个正定阵, 具 $(\lambda_1 > \lambda_2 > \dots > \lambda_n \geq 0)$, 然而实际 Σ_g 或 R_g 计算中最小特征根接近于零, 甚至有稍小于 0 的负数, 因 Σ_g, R_g 是从田间试验数据方差分析计算的结果, 常常会使 Σ_g, R_g 为非正定阵。从几何学知, n 维超棱柱体体积极易受最小棱长度的影响。黄山松种群拓广遗传力的估值中明显地反映出这种不稳定, 而从主成分(或典型相关)分析看, 总体的信息量主要反映在几个大的特征根所对应的几个典范变量或分量上, 即最小的 r 个特征根存在与否不会影响 n 维总体的信息量, 因此要降到 $n-r$ 维, 由此需要确定一个降维的标准。从黄山松数据的估值看, 95% 左右的累计贡献率是合适的, 过高、过低均造成不稳定。

3.4 拓广遗传参数的估算有助于了解多个数量性状的平均遗传变异程度。这种多元遗传分析对全面地、概括性地了解物种的变异, 制定合理的育种策略具有参考意义。从黄山松种群的拓广遗传力、拓广遗传方差看, 其种内的变异中遗传部分平均占的比例较大, 从所研究的性状看, 从生长、抗性性状上进行选择潜力很大。

参 考 文 献

- 1 马育华. 植物数量遗传基础. 南京: 江苏科学技术出版社, 1980, 280~375
- 2 王明麻. 林木育种学概况. 北京: 中国林业出版社, 1989, 191~210
- 3 Zobel B. *Applied to Forest Tree Improvement*. New York, John Wiley & Sons, 1984
- 4 戴君惕等. 遗传学报, 1990, 17(3): 161~167
- 5 范义荣. 林业科技通讯, 1990, (8): 11~13

- 6 唐守正. 多元统计分析方法. 北京, 中国林业出版社, 1986, 20~36, 90~101
- 7 杨 德等. 遗传学报, 1982, 9(3): 188~195
- 8 戴君扬. 遗传学报, 1983, 10(5): 375~383

Tong Zaikang (Zhejiang Forestry College, Lin'an 311300, PRC), Fan Yirong.

Study on Generalized Genetic Parameters of *Pinus taiwanensis* Population. I.
Zhejiang For. Coll., 1993, 10(1): 43~48

Abstract: With the methods defined generalized genetic parameter, generalized phenotypic variance (\tilde{p}), genetic variance (\tilde{g}), heritability (\tilde{h}^2) and genetic correlation coefficient ($\tilde{\rho}$) were estimated based on covariance matrix and correlation matrix respectively, using data of provenance test plantations of *Pinus taiwanensis* in 2 sites, Zhejiang Province. The results showed that it was unsuitable that generalized genetic parameters was estimated based on all eigenvalues, and about 95% of cumulative contribution rate was the most suitable. Generalized heritability estimated are 0.32~0.48 and 0.66~0.87 based on covariance matrix and correlation matrix respectively. And comparability of the generalized genetic parameters were better based on correlation matrix. The detailed discussion was made about estimating methods and practical application based on the results of primary analysis.

Key words: *Pinus taiwanensis* Hayata; genetic analysis; genetical variance; parameter estimations; multivariate analysis