

生态梯度轴 (EGA) 用于林木生态遗传的研究

——PCA方法估算EGAPC_r

① 顾万春 王维胜 李斌 郭文英

(中国林科院林业研究所, 北京, 100091)

摘 要

n 个群体的 p 维环境因子主成分分析(PCA), $Z = U'X$, 恢复原始数据线性表达成 $\hat{X} = UZ$ 。依据 λ 值, 取前 $1-r$ 个主成分, $1 \leq r \leq p$, 将 p 维环境因子降维成1元向量, 建立前 r 个主成分的生态梯度轴(EGA (PC_r)), 估算误差为 $d_{(j)}$ 。对白榆全分布区20个群体的样点6维环境因子PCA求算的EGA(PC_r), 取 $r = 1, 2, 3$, 建立3个生态梯度轴, 累计贡献率依次为63.8%, 84.3%和96.0%, 都能较好的代表环境因子。EGA(PC_r)揭示群体7个性状的梯度变异中得到验证, 并能用于判别种群变异模式。

关键词: 主成分, 生态梯度, 生态梯度轴, 梯度变异。

S718.46
S718.45
林木生态遗传 EGAPC_r

森林生态遗传学是研究广域生态条件造成种以下群体的变异, 即种群的梯度变异。Langlee (1936)最早发现梯度变异, 后人将梯度变异解释为最适性状的产物, 是基因的可变选择值使基因频率沿平行的生态梯度作有规律变化的群体变异^[1]。生态梯度是研究群体梯度变异的重要依据, 也是生态遗传学和林木遗传改良工作者十分关心的研究内容。但将多元生态因子用1元数值表达尚未见报道^[2]。

在前人工作基础上^[3,5], 利用多元统计技术^[4,5], 研究多元生态环境因子降维成1元生态梯度向量, 将多维空间抽象轴转换成1元数据轴, 使每个样点的生态环境因子成为1个数值。从而更简化、更直观的研究群体的梯度变异, 同时便于梯度变异的评价和预测。文中, 生态梯度轴(Ecological gradient axes)英文缩写EGA。用主成分分析(PCA)估算生态梯度轴, 英文缩写EGA(PC_r)。通过白榆种源20个样本分析为例^[6], 用于生态梯度及梯度变异研究, 进一步验证白榆群体连续梯度的种群模式。其结果还与作者以往研究的生态梯度值(GEI)相比较^[7]。

一、试验材料和方法

1. 试验材料 作为林木群体的种源(产地), 满足“足够大、同质、适当”的样本要求^[2,3,8], 是研究生态梯度和梯度变异的理想资料。为此, 选用全国第1批白榆种源试验的20个种源试验林结果^[6], 用于研究群体变异的性状包括花期、果熟期、展叶期、种子千粒重以及9年生树高、胸径、单株材积的的育种值等7个性状(表1)。

2. 生态环境因子选择 生态环境的变异导致种群性状复合体的平行变异, 主导环境因子对梯度变异的影响十分重要, 国内外已有大量报道^[1,2,6-9]。本研究在逐步回归分析基础上, 筛选确定的主导的生态环境因子与前人研究的多数结果一致, 即: 影响广域生态变异主

导因子是年均气温、年降水量和生长期(或无霜期);而决定综合生态因子的地理坐标是纬度,经度和海拔高度。因此,将以上6个地理生态因子用于EGA的研究。

3. 统计分析方法 研究植物群落排序和种群梯度变异研究的方法,常用极点排序

表1 白榆20个群体(样点)的地理生态因子和7个性状统计表

Table 1 The statistic table of the environmental factors and seven traits among twenty populations of American elm

试验号 No	种源名称 Provenance	种源坐标空间 Location of Population			产地气候 Climatic factors			母树物候(月·日) Phytnology (Month/Date)				9年生种源育种值(2) Breeding Values of provenances at age of 9 years		
		北纬(度) Latitude (N)	东经(度) Longitude (E)	海拔高(m) Alt. m	年均温(°C) Annual mean temperature	无霜期(d) Frost-free days	年降水(mm) Annual precipitation	花 期 Flowering	果 熟 期 Fruiting	展 叶 期 Leaf emergence	种子千粒重(g) Thousand seeds weight(g)	树高(m) Height	胸径(cm) DBH	单株材积 (×10 ⁻³ m ³) Volume/trec
1	嫩江 Nenjiang	49.3	125.2	222	3.1	131	378	4.26	5.27	5.13	7.0	4.88	6.69	6.8504
2	肇东 Zhaodong	46.0	126	150	3.5	149	530	4.23	5.26	5.10	6.9	5.06	6.21	6.1151
3	哈尔滨 Haerbin	45.8	126.8	172	3.6	148	577	4.22	5.23	5.8	6.4	6.61	6.76	6.0498
4	玛纳斯 Manasi	44.3	86.2	470	7.8	161	206	4.7	5.2	4.24	5.9	4.89	6.05	6.6202
6	乌鲁木齐 Wulumuqi	43.8	87.5	680	5.6	153	270	4.12	5.5	5.3	6.0	5.23	6.04	5.9911
6	承德 Chengde	41.0	118.0	690	9.0	150	600	4.26	5.11	5.4	6.5	5.06	6.15	6.0094
7	武清 Wuqing	39.4	117.0	9	12.3	210	555	3.26	5.1	4.13	6.9	5.66	6.16	12.2620
6	武清东州 Dongzhuo of Wuqing	39.1	117.0	10	12.3	210	555	3.26	5.2	4.14	7.6	6.07	6.20	12.8168
9	徐水 xushui	39.0	116.7	180	12.6	212	580	3.29	5.3	4.14	7.2	6.19	6.46	13.9111
10	夏津 Xiajing	36.9	116.0	33	13.5	216	663	4.10	4.26	4.8	7.4	6.64	6.48	14.9930
11	魏县 Weixian	36.3	114.9	60	13.6	212	650	4.1	4.26	4.13	7.9	6.29	6.08	12.2836
12	固原 Guyuan	36.1	108.3	1753	6.1	140	420	4.25	5.14	5.2	6.7	5.86	7.01	6.7334
13	金乡 Jin Xiang	35.0	116.3	37	13.4	218	721	3.27	4.26	4.7	8.3	6.66	6.64	15.8110
14	睢县 Sulxian	34.5	115.1	58	14.1	217	640	3.21	4.22	4.10	8.2	6.49	6.48	14.6543
15	睢县榆厢铺 Yuxiang Pu of sui xian	34.4	115.0	58	14.1	218	640	3.22	4.20	4.10	8.4	6.69	6.75	16.0831
16	宿迁 Suqian	33.9	118.3	18	14.2	220	1060	3.26	4.20	4.8	8.5	6.30	6.40	13.9682
17	南阳 Nanyang	33.0	112.6	129.8	15.2	225	870	3.25	4.18	3.27	7.4	6.66	6.79	16.6430
16	霍丘 Huoqiu	32.3	116.3	68	15.5	233	1055	3.21	4.18	3.26	8.0	6.22	6.49	16.4358
19	海宁 Haining	30.6	120.7	30	15.8	225	1250	3.23	4.22	3.26	8.4	6.02	7.93	11.8670
20	临安 Linan	30.2	119.8	168	15.5	224	1300	3.26	4.23	3.27	6.4	6.61	6.73	15.6183

(PO)、相互平均(RA)和主成分分析(PCA)等数学方法,近年来又开始运用位置向量排序(PVO)和Gauch排序等技术^[2,8]。PCA是适用于探讨多元问题的常用方法,在群体遗传、生态遗传和林木育种等研究中被广泛运用,它在结构程序中将生态集合体或性状集合体的总变异有效地分配在新轴中,达到综合排序或分类的目的。本研究是在PCA常用结果的基础上,利用主成分(PC)性质,恢复数据(损失信息少)具有最优性,达到降维目的^[6],并对 n 个样点用前 $1-r$ 个主成分的线性组合计算出1元估算值。

PCA方法的要点是:设 n 个样点(群体), $i=1,2,\dots,n$;对应有 j 个生态因子(变量), $j=1,2,\dots,p$, X 为一个 $n \times p$ 阶阵,写作 R^p ;把 X 转置成 $p \times n$ 阵写作 R^n , R^p 对偶于 R^n ,即 R^p/R^n 。 $Z=U'X$, X 为中心化阵(或用标准化阵 \tilde{X}), U 为正交阵,PC向量的线性组合可以恢复原数据矩阵的表述,而且线性组合的系数就是正交阵的 λ 的元素。因此,以上矩阵式能写成:

$$\hat{X}=UZ \quad (1)$$

取 $1-r$ 个主成分的 n 个样点的线性组合,就是 n 个样点的生态梯度轴 $EGA(PC_r)$ 。写成:

$$EGA(PC_r)=\hat{X}_{(i)}=\sum_{j=1}^r U_{ij}Z_{(j)} \quad (2)$$

式中, $i=1,2,\dots,n$; $1 \leq r \leq p$; $z_{(j)}$ 理解成对应 u 变量的主成分“权值”。利用主成分正交性及欧氏空间(R^n)的商高定理,(2)式的误差为:

$$d_{(i)}=\sum_{j=r+1}^p \lambda_j \quad (3)$$

二、结果分析

1. 环境因子的主成分分析结果

将表1中北纬、东经、海拔高、年均温、无霜期和年降水量6个地理生态因子的20个样点数据,进行PCA。分析结果(表2),前3个PC特征根依次为3.8269, 1.2300和0.7003,累计贡献率(%)依次为68.8014, 84.3016和95.9732。贡献率是特征根(λ)的率值,反映了综合6个地理生态因子的能力大小。不难看出,第1个PC就有较大的代表性,前3个PC信息损失量只有约4%。依据前3个PC将20个样点(种源)排序(图1,表3),结合从表2的累计因子负荷量(h_j^2)可以看出,第1主成分反映无霜期,年均温、纬度;第2主成分主要反映年均温,无霜期和纬度;第3主成分主要反映纬度、海拔高和无霜期。

2. 利用PCA结果估算生态梯度轴($EGA(PC_r)$)

PCA结果的主成分表示方差大小,不能直接解释某1变量意义,它反映多因子投影在多维空间的密度,第1PC是最大密度中心,离开这个复合轴,密度在所有方向上都降低。为了简化、直观地表示生态梯度,应用(2)式进行降维计算,求算生态梯度轴,估算误差并进行 F 检验。取前1个PC、前2个PC和前3个PC的计算,计算结果列于表3右边3列。 $EGA(PC_1)$, $EGA(PC_2)$ 和 $EGA(PC_3)$,相应的信息损失量分别为36.2%, 15.7%和4.0%。前3个PC的方差分析和取前1个,前2个,前3个主成分的 F 检验表明(表4),3个 $EGA(PC_r)$ 都达到足够的可靠性($\alpha=0.01$),都能够被用于生态梯度研究。

3. $EGA(PC_r)$ 的生态梯度的代表性

表2 6个地理生态因子PCA的λ, L 和 h²
Table 2 The values of λ, L and h² of six environmental factors

变 量 Variable	因子负荷量 L _{ij} Factor load			累计因子负荷量 h _i ² = ΣL _{ij} ² Accumulated factor load		
	Z ₁	Z ₂	Z ₃	前 1 个PC The first one PC	前 2 个PC The first two PC	前 3 个PC The first Three PC
X=1	-0.8744	0.4062	-0.2398	0.7646 ^③	0.9296 ^③	0.9671 ^①
X=2	0.3718	0.8190	0.3827	0.1382	0.8090	0.9589
X=3	-0.6100	-0.5517	0.5540	0.3721 ^②	0.6766 ^①	0.9834 ^②
X=4	0.9378	-0.2689	-0.1325	0.8795 ^①	0.9516 ^②	0.9692 ^③
X=5	0.9584	-0.1073	-0.2077	0.9185	0.9300	0.9732
X=6	0.8689	0.0792	0.3540	0.7550	0.7613	0.8666
特征根λ Eigenvalue	3.8281	1.2300	0.7003	(3.8079)	(5.0580)	(5.7684)
贡献率(%) Contribution rate	63.8014	20.5002	11.6716			
累计贡献率(%) Accumulated contribution rate	63.8014	84.3016	95.9732	(63.80)	(84.30)	(95.97)

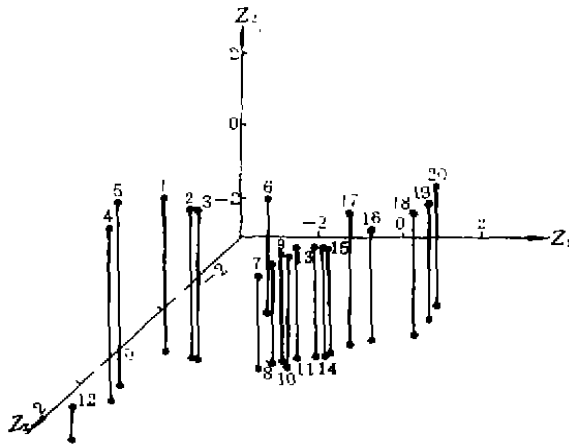


图1 6个地理生态因子20个样点的PCA的3维坐标
Fig.1 The three-dimension Coordinates of PCA from six environmental factors among twenty populations

0.99289, 0.99694和0.99914, 残差标准差为0.07446, 0.05009和0.03001。表明生态梯度轴对6个地理生态因子配合程度很高。

以上简单回归与复回归分析, 都证明了EGA(PC₁)能均衡反映多维生态梯度, 验证了PCA方法估算的生态梯度轴充分代表了综合地理生态因子在空间上的投影密度。而且, 比起PCA排序, EGA(PC₁)能直接描述生态梯度, 具有更直观更简化的优点, 有利于梯度变异的研究。

4. EGA(PC₁)用于群体梯度变异的评价

林木群体与多元生态因子之间相似性或协变性, 通过简化和凝缩成 EGA(PC₁)与性状间的简单关系, 有助于对种群变异的主要趋势和格局加以解释和鉴别。这种简化的生态梯度, 是生态遗传、群体遗传和林木遗传改良研究中渴望实现的方法之一。下边通过白榆种源

(1) 3个生态梯度轴分别与6个地理生态因子单变量相关分析结果(表5), 18个相关系数中仅1个达不到显著性水平, 除与东经、海拔高相关系数较低外, 其余相关系数都在0.7358—0.9584; 其相关性明显高于6个地理生态因子间平均相关程度。另外, 3个生态梯度轴间相关极密切, 相关系数在0.8932—0.9750。可以认定, EGA(PC₁)综合代表各个地理生态因子的属性, 是任何1个地理生态因子无法替代的。

(2) 3个 EGA(PC_i)与6个地理生态因子的复回归分析结果, 偏回归系数介于0.9961—1.0071, 复相关系数依次为

表3 前3个PC坐标与取1—3个PC的PCA(PC)

Table 3 The Coordinates of the first three PC and the principle Components values of the first to third PC

主成分 Principle component	点号: No. of point	各点主成分坐标 Component of each point			取1—3个PC的位点回归值 Regression Value of each point to the first, first two and first three		
		Z(1)	Z(2)	Z(3)	前1个PC The first one PC	前2个PC The first two PC	前3个PC The first three PC
1	1	-2.9000	2.1144	0.1395	-2.4492	-1.7311	-1.6119
2	2	-2.0287	2.0008	0.3117	-1.7117	-1.0322	-0.7658
3	3	-1.9548	2.0313	0.4700	-1.6510	-0.9612	-0.5595
4	4	-2.8608	-1.6970	-1.4971	-2.4162	-2.9925	-4.2720
5	5	-3.0766	-1.7671	-0.9045	-2.5984	-3.1985	-3.9801
6	6	-1.6034	0.0935	0.3651	-1.3542	-1.3237	-0.5843
7	7	0.3279	0.4258	-0.6833	0.2769	0.4215	-0.1667
8	8	0.3522	0.4040	-0.6705	0.2975	0.4347	-0.1384
9	9	0.3027	0.0756	-0.4304	0.2557	0.2814	-0.0865
10	10	0.7260	0.0712	-0.6338	0.6149	0.6391	0.0974
11	11	0.3244	-0.0559	-0.4557	0.6962	0.6772	0.2877
12	12	-2.8214	-2.2387	2.4260	-2.3829	-3.1431	-1.0686
13	13	1.1465	-0.0038	-0.2933	0.9682	0.9669	0.7162
14	14	1.0898	-0.2065	-0.4215	0.9197	0.8496	0.4893
15	15	1.0824	-0.2178	-0.4136	0.9142	0.8402	0.4867
16	16	1.9222	0.1209	0.2720	1.6234	1.6645	1.8969
17	17	1.6957	-0.6054	-0.0977	1.4322	1.2266	1.1431
18	18	2.2969	-0.3041	0.2045	1.9399	1.8366	2.0114
19	19	2.7685	-0.0044	0.7586	2.3382	2.3367	2.9851
20	20	2.7065	-0.2368	1.0590	2.2856	2.2052	3.1103

表4 6个地理因子PCA前3个PC的方差分析

Table 4 The variance analysis of the first three PC of six environmental factors

SV	df	SS	MS	F(1)
第1 PC λ_1 First PC	$n+p-3=23$	3.8281	0.166439	22.045**
第2 PC λ_2 Second PC	$n+p-5=21$	1.2300	0.058571	7.758**
第3 PC λ_3 Third PC	$n+p-7=19$	0.7003	0.036858	4.882**
误差 Errors	$(n-4)(p-4)=32$	0.2416	0.007560	
取前1个PC, $(EGA(PC_1))$; $F_{1,1} = 5.5716^{**}$ First one PC				$(df = 23, 32)$
取前2个PC, $(EGA(PC_2))$; $F_{2,1-2} = 7.4449^{**}$ First two PC				$(df = 44, 61)$
取前3个PC, $(EGA(PC_3))$; $F_{3,1-3} = 10.4480^{**}$ First three pc				$(df = 73, 32)$

(1)**表示 $\alpha=0.01$ 水平差异显著。

**Expresses that the variation is significant at $\alpha=0.01$ level.

7个性状与 $EGA(PC_i)$ 关系的分析, 阐明 $EGA(PC_i)$ 揭示梯度变异的合理性。

(1) 生态梯度轴与群体的7个性状相关极密切, 3个 $EGA(PC_i)$ 的相关系数区间分

表5 EGA(PC)与13个变量的相关性比较

Table 5 Comparison of correlations between EGA(PC) and the thirteen variates

变量 Variable	任一变量与其他12变量相关系数均值与标准差 The mean value and standard deviation of Correlation-Coefficients of each variable to other 12 variables	3个梯度轴与变量相关系数 Correlation-coefficient of the 3 gradients of PC		
		EGA(PC ₁)	EGA(PC ₂)	EGA(PC ₃)
北纬 x_1 Latitude x_1	0.6641 ± 0.2503 (-0.0981—0.9001)	-0.8744	-0.7632	-0.7954
东经 x_2 Longitude	0.2664 ± 0.1380 (-0.0981—0.4676)	0.8718	0.5445	0.6525
海拔 x_3 Alt	0.4601 ± 0.1745 (0.1742—0.6239)	-0.6112	-0.7185	-0.4471
年均温 x_4 Annual mean temperature	0.7047 ± 0.2260 (0.1044—0.9604)	0.9378	0.8547	0.7958
无霜期 x_5 Frost-free days	0.7370 ± 0.1954 (0.2112—0.9604)	0.9584	0.9102	0.7561
年降水 x_6 Annual precipitation	0.6147 ± 0.1402 (-0.4184—0.7882)	0.8689	0.8648	0.9342
花期 x_7 Flowering	0.6725 ± 0.1633 (-0.3220—0.8445)	-0.8435	-0.8229	-0.7443
果熟期 x_8 Fruiting	0.6723 ± 0.1928 (-0.2409—0.8448)	-0.8508	-0.8031	-0.7607
展叶期 x_9 Leaf emergence	0.6051 ± 0.2151 (-0.090—0.7980)	-0.7777	-0.7090	-0.6047
千粒重 x_{10} Thousand seeds weight	0.7012 ± 0.1376 (0.4674—0.8448)	0.8814	0.8309	0.8581
9年生树高(x_{11}) Height at age of 9 years	0.6911 ± 0.2099 (0.2281—0.9315)	0.8651	0.8085	0.7329
9年生胸径(x_{12}) DBH at age of 9 years	0.7063 ± 0.1854 (0.3195—0.9578)	0.8738	0.8433	0.7388
9年生单株材积(x_{13}) Volume/tree at age of 9 years	0.7108 ± 0.1932 (0.2780—0.9753)	0.8873	0.8453	0.7454
平均值 ± σ Mean value	0.6266 ± 0.1327	0.8155 ± 0.1517	0.7976 ± 0.0928	0.7314 ± 0.1131

相关系数显著值: Significant Value of Correlation-coefficient,

$$\alpha_{0.05}(18) = 0.4438, \alpha_{0.01}(18) = 0.5614, \alpha_{0.001}(18) = 0.6787$$

别是: 0.7777—0.8873, 0.7090—0.8809和0.7329—0.8581(表5相关系数为绝对值), 都达到极显著水平。表明群体性状与EGA(PC₁)具有相似性能, 较好地揭示群体的连续梯度变异。

(2) EGA(PC₁)与7个性状复回归分析结果, 偏回归系数约等于1, 3个生态梯度轴的复回归系数依次是0.99453, 0.99427和0.99936, 相应的残标准差是0.09892, 0.10924和0.08485。从总体上反应了生态梯度与梯度变异的相似性。这一结论与以往研究结果一致^[6,7], 表明EGA(PC₁)可用于揭示群体的梯度变异模式。

(3) 在确定种群属于梯度变异的模式后, 可以根据n个样点EGA(PC₁)在一定显著性水平上多重检验, 初拟种源区或育种区的区划, 并与群体性状的梯度变异作吻合性验证, 其区划效果类同于典型相关分析的效果。作者曾报道^[7]用标准化复回归系数权值估算生态梯度值(EGI), 就采用LSD检验区划种源区的做法。另经分析, EGI与3个EGA(PC₁)间相关系数0.9433—0.9539。

5. EGA(PC₁)可由3维地理坐标建立预测模型

白榆20个种源是由树种全分布区抽样所得, 有关20个样点的EGA(PC₁)大体上代表该

树种群体的生态梯度。对于分布区内20个样点之外任何样点,可以进行 $EGA(PC_i)$ 预测。用3维地理坐标(或6个地理生态因子)与 $EGA(PC)$ 建立回归式,引用(2)式建立回归模型(表6)。结果看出,3个生态梯度轴分别与3维地理因子的回归有足够的可靠性,而且误差较小。

同样,在已知 $EGA(PC_i)$ 前提下能够预测估计群体性状值。下边以群体平均千粒重和9年生树高两性状(\hat{y})为例,与3个生态梯度轴(x)简单回归(表7)。

表6 3个 $EGA(PC)$ (Y) 与 3 维地理坐标 (X) 复回归分析

Table 6 The multiple regression analysis of three $EGA(PC)$ (Y) to three-dimensional geographical co-ordinates(X)

统计量 Statistic value	平均数 Mean		标准差 Standard deviation		
	X(1)	38.04500	5.34230	X(2)	114.53500
X(3)	249.29000	399.08300	Y(1)	-0.00005	1.65241
Y(2)	-0.00010	1.69404	Y(3)	-0.00053	1.84046

复回归分析结果 Result of multiple regression analysis	系数 Coefficient		回归系数 Regression Coefficient		
	Y(1)	Y(2)	Y(1)	Y(2)	Y(3)
B(0) =	7.7712	3.3582	-1.5588		
B(1) =	-0.2442	-0.2066	-0.2508		
B(2) =	-0.0171	0.0438	0.0877		
B(3) =	-0.0018	-0.0021	-0.0004		

偏相关系数 Partial Correlation-Coefficient			
X(1) =	-0.9925524	-0.9901141	-0.9729004
X(2) =	0.7080533	0.9346108	0.0944539
X(3) =	-0.09697446	-0.9787428	-0.3845015

复相关系数 Multiple Correlation-Coefficient			
R(4) =	0.9954386	0.9958678	0.9853902
Se =	0.1762758	0.1720695	0.3404399

表7 3个 $EGA(PC_i)$ 与种子千粒重、9年生树高回归的估算模型

Table 7 The models of estimating the 1000 Seeds weight and the tree height of nine years old and three $EGA(PC_i)$

项目 (Item)	$EGA(PC_1)$				$EGA(PC_2)$				$EGA(PC_3)$			
	b_0	b	r	S_e	b_0	b	r	S_e	b_0	b	r	S_e
种子千粒重 Thousand seeds weight \hat{y}_1	7.40002	0.43854	0.8814	0.40495	7.40004	0.42736	0.8809	0.41014	7.40003	0.38335	0.8581	0.44499
9年生树高 Height at age of 9 years	5.99252	0.34650	0.8738	0.34994	5.99253	0.31574	0.8433	0.4153	5.99264	0.26355	0.7329	0.47465

表7中回归系数 b , 是在一定截距下, $EGA(PC_i)$ 变动时性状值变异的倍率关系, 如 $EGA(PC_1)$ 每相差1, 树高差值0.3465m, 余类推。

三、结论与讨论

1. EGA(PC_r)建立在PCA的基础上,运用 $n \times p$ 阵阶 R^* 与 $p \times n$ 阵阶 R^* 对偶的原理,取 $1-r$ 个主成分 $1 \leq r \leq p$, $\hat{x}_{(i)} = \sum_{j=1}^r u_{ij} z_{(j)}$ 成立, $i=1, 2, \dots, n, j=1, 2, \dots, p$, 计算出 n 个样点 $1-r$ 主成分投影数值,并能估算出相应的误差。进而建立了取 $1-r$ 个主成分的1维的生态梯度轴[EGA(PC_r)]。

2. 通过白榆种源试验20样点(种源)的6个地理生态因子分析,证明EGA(PC_r)能充分代表综合生态因子;并能揭示群体7个性状与生态梯度的相似性,实用于梯度变异研究。在3个EGA(PC_r)中,第2个生态梯度轴较好,它综合代表性高于第1个轴;线性关系优于第3个轴。

3. EGA(PC_r)是综合多维地理生态因子的1元向量,是特定的生态距离,对于每一个样点仅为1个数值,用于生态梯度分析和梯度变异研究具有简便性和实用性。在连续分布的树种分布区中,通过抽样研究获得的EGA(PC_r),可以估计分布区内任何1个样点的生态梯度,并能预测群体性状的梯度变异。采用常规生物统计技术,能够估算出生态梯度轴的估算值的误差,也能估算出性状梯度变异的预测值的误差。

4. 本文运用主成分分析方法由于隐含的是线性模型,其应用受到了一定限制。PCA能否普遍适用于生态遗传学研究有待进一步验证。但从本研究来看,EGA(PC_r)方法能用于广域连续分布种群的生态梯度和梯度变异研究,并能用于检验梯度变异或非梯度变异的种群变异模式。另外,对于连续分布的广域树种,种群分布中心通常在生态梯度中心附近^[1-3],本文用生态梯度轴描述生态梯度,其种群分布中心即在O点附近。在进行生态梯度研究时,对原始数集标准化后再PCA,能提高EGA(PC_r)的代表性。

参 考 文 献

- [1] Stern, K and Roche L., Genetics of Forest Ecosystems, Springer-Verlag Berlin Heideberg, New York, 1974 91—146
- [2] 梅里尔美, D.J [美] 1981, 黄瑞复等译, 生态遗传学, 科学出版社, 1981, 265—302
- [3] Whittaker, R. H, 王伯荪译, 植物群落排序, 科学出版社, 1988, 127—152, 176—228
- [4] 张尧庭等, 多元统计分析引论, 科学出版社, 1982, 302—339
- [5] 胡国定等, 多元数据分析方法——纯代数处理, 南开大学出版社, 1990, 149—181
- [6] 顾万春等, 白榆种源及家系选种研究, 林业科学, 1987, 23(4):415—424
- [7] 顾万春, 白榆生态遗传的研究, 东北林业大学出版社, 全国林木遗传育种论文汇编, 1990, 60—64
- [8] Holgate, P, Genotype frequencies in Selection of a clone, Henedity, 1964, (19):507—509
- [9] 木村资生, 集团遗传学概论, 东京培风馆, 1960, 109—113, 234—244

STUDY ON THE APPLICATION OF ECOLOGICAL GRADIENT AXES(*EGA*) IN FOREST ECOLOGICAL GENETICS

—*EGA*(*PCr*) ESTIMATED BY USING PCA METHOD

Gu Wan-Chun Wang Wei-Sheng Li Bin Guo Wen-Ying

(*Research Institute of Forestry, CAF, Beijing, 100091*)

Principal components ($Z = U'X$) of p -dimensional environmental factors in few populations were turned into a linear expression of primary data ($X = UZ$). The estimate error was named d_j . *EGA*(*PC*,)s were calculated according to principal component analysis(PCA) of 6-dimensional environmental factors among the twenty provenances of American elm. The cumulative contributions of 3 *EGA*(*PCR*)s selected according to eigenvalue were 63.8%, 84.3% and 96.0%, respectively. Most information from environmental factors was contained in these 3 *EGA*(*PC*,). *EGA*(*PC*,)method was tested well in investigated population gradient variability of 7 traits and could be used to identify the model of population gradient variation.

Key words: principal components, ecological gradient, ecological gradient axes, gradient variation.