Theor. Appl. Genet. 62, 273–276 (1982)

Factor Analysis in Clusterbean (*Cyamopsis tetragonoloba* (L.) Taub)

G. V. S. Rao and R. S. Paroda
Central Arid Zone Research Institute, Jodhpur (India)

Summary. The pattern of diversity in forty genotypes of clusterbean (*Cyamopsis tetragonoloba* (L.) Taub) was analysed using the centroid method of factor analysis based on nine characters. Three factors were found to adequately account for most of the intercorrelations in the four environments. Branches, clusters, pods and seed yield/plant could be grouped as productivity factors; days to flowering, maturity and plant height as growth factors; seeds/pod and gum content as economic factor. Interestingly, the constitution of the factors were the same in all the environments. The loadings on the economic factors were relatively unstable, whereas the loadings on productivity and growth factors were stable. It is concluded that productivity and growth factors play a pivotal role towards diversity in clusterbean. The results of factor analysis provides confirmatory evidence of diversity in clusterbean, which was earlier studied by using generalized distance, and thus proves the adequacy of the centroid method in biological investigations.

Key words: Factor analysis – Diversity – Clusterbean

Introduction

Factor analysis, a branch of multi-variate analysis, helps elucidate the number and nature of causative influences, thus aiding the selection of better genotypes. Cattel (1965) pointed out that the utility of factor analysis is not only important at exploratory stages of research but it is also important at later stages where the simultaneous action of several factors influencing a variable can be critically analysed. According to Maxwell (1961), and Rao (1964), the main advantage of factor analysis is that it requires smaller numbers of hypothetical variates to explain the covariance matrix compared to the principal component analysis wherein the hypothetical variates required are equal in number to the original variates. Thus, the factor analysis approach is of great importance in studies on biolocical populations where basic knowledge of causal influences is not known.

With these points in view, a study was undertaken in clusterbean, commonly known as guar, to obtain the factors responsible for differentiation among the genotypes. This study was considered important as clusterbean is gaining importance as an industrial crop due to a high gum (glacto-manon) content in its seed endosperm.

Materials and Methods

Forty genotypes of clusterbean (*Cyamopsis tetragonoloba* (L.) Taub.) were collected from different clusterbean growing regions in the states of Gujarat, Haryana and Rajasthan. These were then grown under four different environmental conditions at the Central Research Farms of the Central Arid Zone Research Institute, Jodhpur and Bikaner; the Regional Research Farm of the University of Udaipur at Durgapura, Jaipur; and the Experimental Farm of the Haryana Agricultural University, Hissar, in the monsoon season of 1978. A randomized block design with four replicates was used. Observations on nine characters related to fitness: (1) days to flowering; (2) plant height; (3) branches/plant; (4) clusters/plant; (5) pods/plant; (6) days to maturity; (7) number of seeds/pod; (8) seed yield/plant and (9) gum content: were taken on 5 random plants from each replication.

The centroid method explained by Holzinger and Harman (1941) was followed in the extraction of the factor loadings. As suggested by Cattell (1965) the highest correlation coefficient in each array was taken as an estimate of the array communality. The first centroid factor showed high loadings in most of the items but the second and third factors gave positive weights on some items and negative on others. These factors were, therefore, rotated in order to obtain new interpretable factors leading to the same correlation matrix with the same degree of accuracy. The procedure of orthogonal rotation (Fruchter 1967) was employed wherein two factors are taken at a time and new loadings are calculated after locating the new axes based on maximising the zero factor loadings and

0040-5752/82/0062/0273/$ 1.00

Table 1. Centroid factor matrix after orthogonal rotation for 9 characters in clusterbean at the first and second environments

Variables	Common factor coefficients						Communality (h^2)			
	Factor I		Factor II		Factor III		E I		E II	
	E I	E II	E I	E II	E I	E II	Calculated	Original	Calculated	Original
1.	0.41	0.05	0.71	0.36	0.36	0.65	0.80	0.78	0.55	0.55
2.	0.15	0.04	0.70	0.54	0.31	0.61	0.61	0.61	0.67	0.67
3.	0.30	0.57	-0.28	0.25	-0.23	0.48	0.22	0.25	0.62	0.63
4.	0.69	0.87	0.10	0.20	0.10	-0.17	0.50	0.49	0.83	0.82
5.	0.65	0.75	0.20	0.20	0.25	0.46	0.53	0.54	0.81	0.82
6.	0.25	-0.10	0.75	0.25	0.28	0.61	0.70	0.71	0.44	0.44
7.	0.05	-0.15	0.30	0.70	0.43	0.41	0.28	0.28	0.68	0.68
8.	0.92	0.60	0.25	0.50	0.18	0.42	0.94	0.93	0.79	0.79
9.	0.14	0.08	0.33	0.70	0.40	0.53	0.29	0.28	0.78	0.79
Total							4.87	4.87	6.17	6.19
Contribution of factor	2.11	2.05	1.95	1.86	0.81	2.26				
% of total original communality	43.32	33.12	40.04	30.05	16.63	36.51	100		99.68	

Table 2. Centroid factor matrix after orthogonal rotation for 9 characters in clusterbean at the third and fourth environments

Variables	Common factor coefficients						Communality (h^2)			
	Factor I		Factor II		Factor III		E III		E IV	
	E III	E IV	E III	E IV	E III	E IV	Calculated	Original	Calculated	Original
1.	0.72	0.50	0.25	0.38	0.44	0.40	0.77	0.78	0.55	0.57
2.	0.53	0.53	0.30	0.10	0.45	0.51	0.57	0.58	0.55	0.57
3.	0.33	0.32	0.48	0.56	0.37	0.54	0.48	0.48	0.71	0.70
4.	0.30	0.02	0.63	0.68	0.45	0.28	0.69	0.69	0.54	0.54
5.	0.25	0.00	0.62	0.84	0.38	0.21	0.59	0.60	0.75	0.75
6.	0.77	0.58	0.10	0.35	0.14	0.45	0.62	0.62	0.66	0.67
7.	0.20	0.26	0.20	0.47	0.69	0.76	0.56	0.56	0.87	0.87
8.	0.35	-0.05	0.70	0.84	0.10	0.24	0.62	0.62	0.77	0.77
9.	0.15	0.20	0.40	0.18	0.47	0.51	0.40	0.43	0.33	0.34
Total							5.30	5.36	5.73	5.78
Contribution of factor	1.84	1.08	1.86	2.72	1.60	1.93				
% of total original communality	34.33	18.68	34.70	47.06	29.85	33.39	98.88		99.13	

minimising negative loadings. The final factor loadings of the nine items together with the calculated and original communalities (h^2) are presented in Tables 1 and 2.

Results

Three factors were found to adequately to explain the correlation matrices in this material since the factor matrices multiplied by their inverses were essentially equal to the original correlation matrices. The coefficients of the residual matrix were negligible in magnitude after the first three factors had been eliminated. It was significant that the same three factors were obtained in all four environments (Tables 3 and 4) indicating the proper choice of the characters and their constant

Table 3. Effect of factors on nine variables in clusterbean at the first and second environments

Variables	Environment I			Environment II		
		Loading	Factor		Loading	Factor
Branches/plant	3	0.30	I	3	0.57	I
Clusters/plant	4	0.69		4	0.87	
Pods/plant	5	0.65		5	0.75	
Seed yield/plant	8	0.92		8	0.60	
Days to flowering	1	0.71	II	1	0.65	III
Plant height	2	0.70		2	0.61	
Days to maturity	6	0.75		6	0.61	
No. of seeds/pod	7	0.43	III	7	0.70	II
Gum content	9	0.40		9	0.70	

Table 4. Effect of factors on nine variables in clusterbean at the third and fourth environments

Variables	Environment III			Environment IV		
		Loading	Factor		Loading	Factor
Branches/plant	3	0.48	II	3	0.56	II
Clusters/plant	4	0.63		4	0.68	
Pods/plant	5	0.62		5	0.64	
Seed yield/plant	8	0.70		8	0.84	
Days to flowering	1	0.72	I	1	0.50	I
Plant height	2	0.53		2	0.53	
Days to maturity	6	0.77		6	0.58	
No. of seeds/pod	7	0.69	III	7	0.76	III
Gum content	9	0.47		9	0.51	

importance in different environments. Days to flowering, plant height and days to maturity constituted one factor – this was considered as the growth factor. Branches/plant, clusters/plant, pods/plant and seed yield, which would determine the productivity, formed another factor, – the productivity factor. The third factor consisted of number of seeds/pod and gum content and were, therefore, designated as the economic factor.

Discussion

This study was initiated with a preliminary knowledge of the characters included. The characters chosen were important contributors to yield and fitness. This material had already been subjected to studies on divergence and the characters chosen had already been found to be appropriate and adequate (Rao 1980). Hence, the main aim in subjecting the same material to factor analysis was to find out whether this method would be able to provide any meaningful factors responsible for differentiation among the genotypes of clusterbean.

The study revealed the utility of factor analysis in providing fewer stable factors to delineate the divergent populations. The rule where $(p+k)$ should be less than $(p-k)^2$ (p = number of characters, k = number of factors) was applicable, as indicated previously by Lawley and Maxwell (1963).

Different schools of thought exist with regard to the correct method of estimating array communalities. In fact, various methods have been tried by many workers in many different disciplines in this connection. In the present investigation, the estimated communalities were adequate for drawing conclusions as the three factors together accounted for 100 per cent of the total communality.

Even with the smaller number of variables included in this study, factor analysis was potent enough to isolate the different factors responsible for differentiation. A comparison of the factor loadings from different environments revealed substantial changes in the sizes of loading. However, the composition of the variables in the factors remained essentially the same. Though the economic factor was stable with regard to its constitution, the loadings in different environments varied largely as compared to growth factor and productivity factor. Morever, the contribution of the economic factor towards original communality was less.

Therefore, it can be inferred that growth and productivity factors, and not gum content, appear to be major causative influences for diversity.

In this collection, there is a good scope for selecting genotypes for better adaptation due to the growth factor because flower is an important factor for growth and therefore, for adaptation (Murty and Arunachalam 1970).

Seed yield potential, rather than the gum content, can be improved, as the productivity factor includes in it the yield components branches, clusters and pods with high loadings rather than such economic factors as gum content. Hence, selection based on plant types having a higher number of pods, clusters and branches would enable breeders to realize desired gains in clusterbean.

Literature

Cattel, R.B. (1965): Factor analysis: An Introduction to essentials I. The purpose and underlying models. Biometrics **21,** 190–215

Cattel, R. B. (1965): Factor analysis: An Introduction to essentials II. The role of factor analysis in research. Biometrics, **21,** 405–435

Fruchter, B. (1967): An Introduction to Factor Analysis: D. Van Nosli and Co.

Holzinger, K.J.; Harman, H.H. (1941): Factor Analysis. Chicago: University of Chicago Press
Lawley, O.N.; Maxwell, A.E. (1963): Factor analysis as a Method. London: Butterworths
Maxwell, A.E. (1961): Recent trends in factor analysis. J. Roy. Statis. Soc. A. **124**, 49–59
Murty, B.R.; Arunachalam, V. (1970): Factor analysis of diversity in the genus *sorghum* Indian J. Genet. **27**, 123–135
Rao, C.R. (1964): The use of interpretation of principal component analysis in applied research. Sankhya **26A**, 329–358
Rao, G.V.S. (1980): Studies on genetic variability and association of seed characters with gum content in guar (*Cyamopsis tetragonoloba* (L.) Taub.) Ph. D. Thesis, Jodhpur: University of Jodhpur

Received January 22, 1982
Communicated by G. S. Khush

Dr. G. V. S. Rao
Central Arid Zone Research Institute
Jodhpur (India)

Prof. Dr. R. S. Paroda
Department of Plant Breeding
Haryana Agricultural University
Hissar 125004 (India)