

## SELEÇÃO DE CARACTERES FENOTÍPICOS PARA DIFERENCIAÇÃO DE GENÓTIPOS DE SOJA, PELA ANÁLISE DE COMPONENTES PRINCIPAIS<sup>1/</sup>

Carlos Sigueyuki Sedyama<sup>2/</sup>  
Tocio Sedyama<sup>2/</sup>  
Ney Sussumu Sakiyama<sup>2/</sup>

### 1. INTRODUÇÃO

A coleta de dados experimentais torna-se tarefa difícil quando é grande o número de caracteres da planta a ser avaliado. Daí a necessidade de selecionar, arbitrariamente ou tecnicamente, número reduzido de caracteres, para viabilizar seu estudo.

Se o interesse, num experimento, for a seleção de caracteres que determinem ou expliquem outro qualquer (por exemplo a produção de grãos), a análise de triilha (3, 5, 6) é adequada. Porém, se o interesse foi identificar as variáveis que proporcionam maiores diferenças fenotípicas entre os genótipos, pode ser utilizada a análise de componentes principais (1, 2, 4).

Com o presente trabalho os objetivos foram verificar a importância de 12 caracteres fenotípicos na diferenciação de 16 genótipos de soja, pela análise de componentes principais, e analisar a eficiência dessa técnica na escolha de caracteres para experimentos semelhantes.

### 2. MATERIAL E MÉTODOS

Dois experimentos com soja (*Glycine max* (L.) Merrill) foram instalados em Viçosa, MG — experimento I, em 26/11/1982, e experimento II, em 24/11/1983 — ambos no delineamento inteiramente casualizado, com 16 genótipos e oito e nove repetições, respectivamente. A unidade experimental foi constituída de uma única planta, com espaçamento de 1,0 metro entre plantas.

---

<sup>1/</sup> Aceito para publicação em 23-02-1989.

<sup>2/</sup> Departamento de Fitotecnia da UFV. 36570 Viçosa, MG.

Os caracteres avaliados no campo foram: altura da planta, altura de inserção da primeira vagem, número de dias da germinação à floração, número de dias da germinação à maturação, número de dias da floração à maturação, diâmetro da copa, peso de sementes, número de ramos, número de nós, número de vagens, número de sementes por vagem e peso de cem sementes.

A importância relativa de cada um desses 12 caracteres na diferenciação fenotípica dos 16 genótipos de soja foi avaliada através do coeficiente de correlação entre cada caráter e o primeiro componente principal ( $Y_1$ ). Os resultados obtidos nos experimentos I e II foram comparados, para verificar a eficiência da utilização do primeiro componente principal na escolha de caracteres mais representativos da diferenciação fenotípica dos 16 genótipos.

O primeiro componente principal ( $Y_1$ ), dado pela combinação linear de  $p$  caracteres padronizados, cuja variância  $V(Y_1)$  é máxima, foi expresso do seguinte modo (1, 2, 4):

$$Y_1 = e_1' Z = e_{11} Z_1 + e_{21} Z_2 + \dots + e_{p1} Z_p,$$

com  $\text{Cov}(Z) = P = R$

e  $e_1' e_1 = 1$ ,

em que

$Y_1$  representa o primeiro componente principal;

$Z$  representa o vetor das  $p$  variáveis padronizadas;

$e_1'$  representa a transposta do autovetor associado ao maior autovalor ( $\lambda_1$ ) da matriz  $R$ ; e

$R$  representa a matriz de correlação das variáveis  $Z_i$  ( $i = 1, 2, \dots, p$ ).

O vetor de autovalores ( $\lambda$ ) da matriz  $R$  foi obtido do seguinte modo:  $|R - \lambda| = 0$ .

O autovalor ( $e_1$ ) foi obtido por meio da expressão 
$$e_1 = \frac{v_i}{\sqrt{v_1' v_i}},$$

satisfazendo a condição de unicidade,  $e_1 = 1$ ,

sendo  $|R - \lambda_1| v_1 = 0$ .

A proporção da variância total explicada pelo primeiro componente principal

foi dada por  $\frac{\lambda_1}{p}$ ,

sendo

$\lambda_1$  o maior autovalor da matriz  $R$ , representando a variância do primeiro componente principal, e

$p$  a variância total de  $p$  componentes principais. Eis que

$$p = \lambda_1 + \lambda_2 + \dots + \lambda_p$$

A correlação entre o primeiro componente principal e a variável  $Z_i$  ( $i = 1, 2, \dots, p$ ) foi dada por

$$r_{y_1 Z_i} = e_{i1} \sqrt{\lambda_1}$$

### 3. RESULTADOS E DISCUSSÃO

Os dados relativos aos 12 caracteres fenotípicos de soja avaliados nos experimentos I e II são apresentados, resumidamente, nos Quadros 1 e 2, respectivamente.

Observou-se, em ambos os experimentos, diferença significativa entre os genótipos ( $P < 0,05$ ), pelo teste F, para cada caráter, analisado isoladamente.

No experimento I, o primeiro componente principal ( $Y_1$ ), formado pela combinação linear dos 12 caracteres iniciais, correlacionou-se com seis desses caracteres ( $P < 0,05$ ), a saber: altura da planta, diâmetro da copa, peso de sementes, número de ramos, número de nós e número de vagens, ao passo que, no experimento II, os caracteres correlacionados com o primeiro componente principal foram sete: altura de inserção da primeira vagem, número de dias da germinação à floração, número de dias da floração à maturação, peso de sementes, número de vagens, número de sementes por vagem e peso de cem sementes.

Comparando esses resultados, observou-se que somente os caracteres peso de sementes e número de vagens foram comuns aos dois experimentos, o que indica que os caracteres escolhidos no experimento I não seriam indicados para o estudo da diferenciação fenotípica dos genótipos no experimento II, e vice-versa, quando o estudo fosse efetuado apenas com o primeiro componente principal.

### 4. RESUMO E CONCLUSÕES

A análise de componentes principais foi utilizada para estudar a variabilidade fenotípica de 12 genótipos de soja.

A correlação simples entre o primeiro componente principal e o valor do caráter padronizado,  $Z_i$  ( $i=1,2,\dots, p$ ), foi usada para identificar os melhores dentre 12 caracteres, nos experimentos I e II, conduzidos em 1982 e 1983, respectivamente. Somente dois desses caracteres correlacionados foram comuns aos dois experimentos, o que indica que os caracteres selecionados em um experimento não devem ser indicados para outro, quando o estudo é efetuado apenas com o primeiro componente principal.

### 5. SUMMARY

#### (SELECTION OF PHENOTYPIC CHARACTERISTICS FOR THE DIFFERENTIATION OF SOYBEAN GENOTYPES BY MEANS OF PRINCIPAL COMPONENT ANALYSIS)

Principal component analysis was used to study the phenotypic variability among sixteen soybean varieties.

The simple correlation of the first principal component with the standardized variable  $Z_i$  ( $i=1,2,\dots, p$ ) was used to identify the best among the traits evaluated in the experiments I and II, carried out in 1982 and 1983, respectively. Only two correlated traits were the same in both experiments, meaning that the traits chosen in one experiment could be inadequate for use in another, when the study is made only with the first principal component.

QUADRO 1 - Estimativas dos desvios-padrão das médias, dos valores de F de 12 caracteres e das correlações entre o primeiro componente principal ( $Y_1$ ) e os valores padronizados do caráter  $Z_i$  ( $i=1,2,\dots,p$ ), no estudo de 16 genótipos de soja (Experimento I)

Caráter(x)	s(x)	Média	F	$Y_1$ p=12	$Y_1$ p=6
Altura da planta (cm)	14,38	55,82	8,08a	0,80*	0,79*
Alt. de inserção 1.ª vagem (cm)	2,92	9,53	4,02a	0,34ns	
Dias para floração	5,57	63,66	4,01a	-0,20ns	
Dias para maturação	2,84	142,51	1,99a	0,38ns	
Dias da floração à maturação	6,49	78,85	2,86a	0,34ns	
Diâmetro da copa (cm)	11,07	54,59	6,67a	0,81*	0,82*
Peso de sementes (g)	29,30	57,01	3,51a	0,90*	0,90*
Número de ramos	2,43	8,87	4,44a	0,83*	0,87*
Número de nós	3,62	16,75	5,02a	0,78*	0,80*
Número de vagens	100,96	219,94	2,87a	0,83*	0,89*
Número de sementes por vagem	0,43	2,13	3,17a	0,30ns	
Peso de cem sementes (g)	2,10	12,19	8,00a	0,18ns	
Proporção explicada da variância total.				0,385	0,713
Variância do primeiro componente principal ( $\lambda_1$ )				4,62	4,28

a Significativo a 1%, pelo teste de F.

\* Correlação simples entre  $Y_1$  e  $Z_i$  ( $i=1,2,\dots,p$ ) significativa, a 5%.

QUADRO 2 - Estimativas dos desvios-padrão das médias, dos valores de F de 12 caracteres e das correlações entre o primeiro componente principal ( $Y_1$ ) e os valores padronizados do caráter  $Z_i(i=1,2,\dots,p)$ , no estudo de 16 genótipos de soja (Experimento II)

Caráter (x)	s(x)	Média	F	$Y_1$ p=12	$Y_1$ p=6
Altura da planta (cm)	12,35	89,85	11,27a	0,44ns	
Alt. inserção 1ª vagem (cm)	6,08	20,07	7,86a	0,66*	-0,59*
Dias para floração	6,29	72,42	8,04a	0,75*	-0,75*
Dias para maturação	2,41	147,03	8,90a	0,49ns	
Dias da floração à maturação	5,44	74,62	4,04a	-0,65*	0,71*
Diâmetro da copa (cm)	10,82	61,77	2,90a	-0,27ns	
Peso de sementes (g)	22,70	66,59	1,75b	-0,52*	0,58*
Número de ramos	2,71	9,90	2,49a	-0,05ns	
Número de nós	3,47	22,40	6,53a	0,26ns	
Número de vagens	101,85	305,71	3,07a	-0,67*	0,71*
Número de sementes por vagem	0,02	2,20	10,88a	-0,59*	0,63*
Peso de cem sementes (g)	1,67	10,06	9,14a	0,65*	-0,63*
Proporção explicada da variância total.					
Variância do primeiro componente principal ( $\lambda_1$ )				0,290	0,439
				3,48	3,07

a, b Significativos, a 1% e 5%, respectivamente, pelo teste de F.

\* Correlação simples entre  $Y_1$  e  $Z_i(i=1,2,\dots,p)$  significativa, a 5%.

## 6. LITERATURA CITADA

1. HARRIS, R.J. *A primer of multivariate statistics*. New York, Academic Press, 1975. 332 p.
2. JOHNSON, R.A. & WICHERN, D.W. *Applied multivariate statistical analysis*. New Jersey, Prentice-Hall, 1982. 594 p.
3. KEMPTHORNE, O. *An introduction to genetic statistics*. Iowa, The Iowa State University Press, 1973. 545 p.
4. TATSUOKA, M.M. *Multivariate analysis — Techniques for educational and psychological research*. New York, John Wiley & Sons, 1971. 310 p.
5. WRIGHT, S. Correlation and causation. Part I. Method of path coefficients. *J. Agric. Res.*, 20:557-585. 1921.
6. WRIGHT, S. The method of path coefficients. *Annals of Mathematical Statistics*, 5:161-215. 1934.