

ANÁLISE DE VARIÂNCIA MULTIVARIADA PARA OS CRUZAMENTOS DIALÉLICOS¹

CARLOS ALBERTO DA SILVA LEDO²
DANIEL FURTADO FERREIRA³
MAGNO ANTÔNIO PATTO RAMALHO⁴

RESUMO – Com este trabalho objetivou-se apresentar a análise de variância multivariada para os cruzamentos dialélicos, com o intuito de fornecer meios mais eficientes para a seleção de genótipos superiores. Entre as metodologias mais comumente utilizadas para análise dialélica, cita-se a proposta por Gardner e Eberhart (1966), por meio da qual são estimados os efeitos dos genitores e da heterose entre seus híbridos. Essa metodologia, no entanto, é proposta apenas para o caso univariado. O fato é que, para a obtenção de populações superiores, os melhoristas necessitam avaliar vários caracteres para melhor inferir sobre a superioridade rela-

tiva dos mesmos. Para isso, foi adotado um modelo estendido para um número $k > 1$ de variáveis. Inicialmente, foram obtidos, pelo método dos quadrados mínimos, as expressões explícitas dos estimadores dos efeitos genéticos desse modelo e de suas respectivas variâncias. Posteriormente, foram obtidas as matrizes de soma de quadrados e produtos relacionados. Finalmente, foram apresentados os testes multivariados para as hipóteses de interesse. A análise de variância multivariada para os cruzamentos dialélicos pode ser realizada para estimar a heterose em várias características simultaneamente.

TERMOS PARA INDEXAÇÃO: Análise dialélica, seleção de múltiplas características, heterose, modelo de Gardner e Eberhart.

MULTIVARIATE ANALYSIS OF VARIANCE FOR DIALLEL CROSSES

ABSTRACT – Multivariate analysis of variance (MANOVA) was presented to provide a more efficient way to select superior genotypes from diallel crosses. Gardner and Eberhart (1966) model is one of the most commonly used methodologies for diallel analysis, in which parents (varieties) and heterosis effects are estimated, despite been presented by the authors as an univariate methodology. On the other hand, to get better populations and broaden the inference, plant

breeders use to evaluate many traits. An extension of this model to $k > 1$ traits was adopted. Explicit expressions to least squares estimates of genetic effects and their variances were worked out. Sums of squares and products matrices were also given. Multivariate tests for the hypotheses of interest were presented. MANOVA on diallel crosses can be accomplished to estimate simultaneously the heterosis in several traits.

INDEX TERMS: Diallel analysis, multiple traits selection, heterosis, Gardner and Eberhart model.

INTRODUÇÃO

No melhoramento de plantas, a hibridação das linhagens e populações representa um dos pontos fundamentais, pois possibilita a combinação de alelos favoráveis que estão em genitores diferentes, com o intuito de produzir novas cultivares adaptadas (RAMALHO et

al., 1993). Uma das dificuldades encontradas pelo melhorista se refere à escolha de materiais promissores para serem usados como genitores num programa de hibridação. Essa escolha, quando bem feita, propicia a obtenção de populações segregantes, com potencial de fornecer progênes superiores, traduzindo-se em maior eficiência do programa de melhoramento. A identifica-

-
1. Parte da tese apresentada à UNIVERSIDADE FEDERAL DE LAVRAS/UFLA, Caixa Postal 37 – 37200-000 – Lavras, MG, pelo primeiro autor, para obtenção do grau de Doutor, na área de Genética e Melhoramento de Plantas.
 2. Engenheiro Agrônomo, D.Sc., Pesquisador da Embrapa Mandioca e Fruticultura, Caixa Postal 07 – 44380-000, Cruz das Almas, BA. led@cnpmf.embrapa.br.
 3. Engenheiro Agrônomo, D.Sc., Professor do Departamento de Ciências Exatas da UFLA. Bolsista do CNPq.
 4. Engenheiro Agrônomo, D.Sc., Professor do Departamento de Biologia da UFLA. Bolsista do CNPq.

ção de metodologias que auxiliam nessa escolha eficiente tem recebido uma maior atenção dos pesquisadores (BAENZIGER e PETERSON, 1991; BERNARDO, 1996; ABREU, 1997; SOUZA 1997; CHARCOSSET et al., 1998).

Entre as técnicas que auxiliam na escolha de genitores, os cruzamentos dialélicos têm sido largamente utilizados por melhoristas. Isso porque possibilitam a obtenção de informações com base no comportamento "per se" de um grupo de genitores e, principalmente, por considerar sua capacidade de combinação ao formar híbridos. Entre as metodologias de escolha de genitores mais comumente utilizadas, citam-se as propostas por Griffing (1956) e por Gardner e Eberhart (1966). Para o estudo mais detalhado da heterose, a metodologia desenvolvida por Gardner e Eberhart (1966) tem sido mais utilizada por permitir o desdobramento desse efeito em heterose média, heterose do genitor e heterose específica. No entanto, essa metodologia é proposta apenas para o caso univariado.

O fato é que, para a obtenção de materiais genéticos superiores, os melhoristas necessitam avaliar vários caracteres para melhor inferir sobre a superioridade relativa dos mesmos. Na aplicação de técnicas biométricas, o que se utiliza normalmente é a análise univariada, sendo as análises combinadas, em geral, restritas a procedimentos bivariados. A análise dessas variáveis isoladamente poderá não ser suficiente para modelar o fenômeno, pois não consideram as correlações existentes entre elas. Dessa forma, a utilização da teoria de análise multivariada permite combinar as múltiplas informações contidas na unidade experimental, de modo que facilite a execução da seleção com base na combinação de variáveis, possibilitando discriminar as populações mais promissoras, principalmente no contexto genético.

A utilização de técnicas de análise multivariada tem sido pouco aproveitada na análise de cruzamentos dialélicos. Martinez Garza (1983) discutiu a técnica de análise de variância multivariada para os modelos 2 e 4 de Griffing (1956). Sakaguti (1994) apresentou, com base no modelo proposto por Griffing (1956), a técnica de análise de variância multivariada para dialelos completos não-balanceados em cruzamentos com coelhos. Atualmente, contando com a maior disponibilidade dos recursos na área da informática, as pesquisas podem explorar melhor novas técnicas de análise (Cruz, 1990).

Conduziu-se este trabalho com o objetivo de desenvolver as expressões para a análise de variância multivariada do modelo de cruzamentos dialélicos de Gardner e Eberhart (1966).

METODOLOGIA

Foram considerados um cruzamento dialélico com p genitores, suas $p(p-1)/2$ combinações híbridas e a avaliação de k variáveis. Seja $Y_{ij}^{(t)}$ o valor médio observado em um genitor, quando $i = j$, ou em uma combinação híbrida, quando $i \neq j$ e $j > 1$, na t -ésima variável, sendo $i, j = 1, 2, \dots, p$ e $t = 1, 2, \dots, k$.

Para a obtenção da extensão multivariada do método de Gardner e Eberhart (1966), foram empregados os seguintes modelos estatísticos:

$$(1) Y_{ij}^{(t)} = m^{(t)} + \bar{e}_{ij}^{(t)}$$

$$(2) Y_{ij}^{(t)} = m^{(t)} + \frac{1}{2}(v_i^{(t)} + v_j^{(t)}) + \bar{e}_{ij}^{(t)}$$

$$(3) Y_{ij}^{(t)} = m^{(t)} + \frac{1}{2}(v_i^{(t)} + v_j^{(t)}) + \theta \bar{h}^{(t)} + \bar{e}_{ij}^{(t)}$$

$$(4) Y_{ij}^{(t)} = m^{(t)} + \frac{1}{2}(v_i^{(t)} + v_j^{(t)}) + \theta(\bar{h}^{(t)} + h_i^{(t)} + h_j^{(t)}) + \bar{e}_{ij}^{(t)}$$

$$(5) Y_{ij}^{(t)} = m^{(t)} + \frac{1}{2}(v_i^{(t)} + v_j^{(t)}) + \theta(\bar{h}^{(t)} + h_i^{(t)} + h_j^{(t)} + s_{ij}^{(t)}) + \bar{e}_{ij}^{(t)}$$

em que:

- $m^{(t)}$: média geral, para a variável t ;
- $v_i^{(t)}$ e $v_j^{(t)}$: efeito dos genitores i e j , para a variável t ;
- $\bar{h}^{(t)}$: efeito da heterose média, para a variável t ;
- $h_i^{(t)}$ e $h_j^{(t)}$: efeito heterótico dos genitores i e j , para variável t ;
- $s_{ij}^{(t)}$: efeito da heterose específica entre os genitores i e j , para a variável t ;
- $\bar{e}_{ij}^{(t)}$: erro experimental médio, em que os $\bar{e}_{ij}^{(t)}$ têm distribuição multinormal k -dimensional, com vetor nulo de médias e uma matriz de variâncias e covariâncias Σ ;

$\theta = 0$, quando $i = j$ e $\theta = 1$, quando $i \neq j$.

No modelo reduzido (1), admite-se apenas o efeito da média geral; no modelo reduzido (2), admite-se que não exista heterose nos cruzamentos entre os genitores i e j ; no modelo reduzido (3), admite-se a existência de um mesmo efeito heterótico para to-

dos os cruzamentos; no modelo reduzido (4), admite-se que a heterose não é a mesma para todos os cruzamentos, e cada genitor apresenta um efeito heterótico próprio, no modelo completo (5), admitem-se as mesmas pressuposições do modelo reduzido (4) mais um efeito adicional, resultante da heterose específica de cada cruzamento entre os genitores i e j .

Para a estimação dos efeitos da constante ($m^{(t)}$), dos genitores ($v_i^{(t)}$ e $v_j^{(t)}$) e das heteroses ($\bar{h}^{(t)}$, $h_i^{(t)}$, $h_j^{(t)}$ e $s_{ij}^{(t)}$) e de suas respectivas variâncias e somas de quadrados e produtos, utilizou-se o método dos quadrados mínimos. Assim, obtiveram-se as soluções com base nas equações normais $X'X\hat{\beta} = X'Y$, derivadas do modelo linear $Y = X\beta + \epsilon$, em que Y é matriz de médias da tabela dialélica, X é matriz com os coeficientes relacionados aos parâmetros do modelo, β é matriz de parâmetros do modelo e ϵ é a matriz de erros.

Uma vez que a matriz X não é de posto coluna completo e pela necessidade de tornar certas funções paramétricas estimáveis, foram adotadas as seguintes restrições paramétricas:

$$(1) \sum_i g_i^{(t)} = \sum_j g_j^{(t)} = 0$$

$$(2) \sum_i h_i^{(t)} = \sum_j h_j^{(t)} = 0$$

$$(3) \sum_i s_{ij}^{(t)} = \sum_j s_{ij}^{(t)} = 0$$

Os estimadores dos efeitos da constante ($m^{(t)}$), dos genitores ($v_i^{(t)}$ e $v_j^{(t)}$) e das heteroses ($\bar{h}^{(t)}$, $h_i^{(t)}$, $h_j^{(t)}$ e $s_{ij}^{(t)}$) foram obtidos por meio da solução do sistema de equações normais dada por $X'X_R\hat{\beta} = X'Y$ para cada modelo considerado.

Admitindo-se o modelo completo (5) como fixo e que se tenha interesse em testar hipóteses de funções lineares algumas estimáveis sob restrição paramétrica, foram obtidas as variâncias dos efeitos e dos contrastes entre efeitos.

As matrizes de somas de quadrados e produtos (SQP), associadas aos parâmetros de cada modelo considerado, foram dadas por:

SQP do modelo reduzido (1) = SQP da constante = $R(m^{(t)}) = \hat{\beta}'X'Y_{(1)} = C$

SQP do modelo reduzido (2) = $R(m^{(t)}, v_i^{(t)}, v_j^{(t)}) = \hat{\beta}'X'Y_{(2)}$

SQP do modelo reduzido (3) = $R(m^{(t)}, v_i^{(t)}, v_j^{(t)}, \bar{h}^{(t)}) = \hat{\beta}'X'Y_{(3)}$

SQP do modelo reduzido (4) = $R(m^{(t)}, v_i^{(t)}, v_j^{(t)}, \bar{h}^{(t)}, h_i^{(t)}, h_j^{(t)}) = \hat{\beta}'X'Y_{(4)}$

SQP do modelo completo (5) = $R(m^{(t)}, v_i^{(t)}, v_j^{(t)}, \bar{h}^{(t)}, h_i^{(t)}, h_j^{(t)}, s_{ij}^{(t)}) = \hat{\beta}'X'Y_{(5)}$

As matrizes de somas de quadrados e produtos para os efeitos do genitor, da heterose, da heterose média, da heterose do genitor e da heterose específica foram obtidas por meio de:

SQP (genitor) = $\hat{\beta}'X'Y_{(2)} - \hat{\beta}'X'Y_{(1)} = R(v_i^{(t)}, v_j^{(t)}/m^{(t)})$

SQP (heterose) = $\hat{\beta}'X'Y_{(5)} - \hat{\beta}'X'Y_{(2)} = R(\bar{h}^{(t)}, h_i^{(t)}, h_j^{(t)}, s_{ij}^{(t)}/m^{(t)}, v_i^{(t)}, v_j^{(t)})$

SQP (heterose média) = $\hat{\beta}'X'Y_{(3)} - \hat{\beta}'X'Y_{(2)} = R(\bar{h}^{(t)}/m^{(t)}, v_i^{(t)}, v_j^{(t)})$

SQP (heterose do genitor) = $\hat{\beta}'X'Y_{(4)} - \hat{\beta}'X'Y_{(3)} = R(h_i^{(t)}, h_j^{(t)}/m^{(t)}, v_i^{(t)}, v_j^{(t)}, \bar{h}^{(t)})$

SQP (heterose específica) = $\hat{\beta}'X'Y_{(5)} - \hat{\beta}'X'Y_{(4)} = R(s_{ij}^{(t)}/m^{(t)}, v_i^{(t)}, v_j^{(t)}, \bar{h}^{(t)}, h_i^{(t)}, h_j^{(t)})$

O esquema da análise de variância multivariada, para o modelo proposto por Gardner e Eberhart (1966), é apresentado na Tabela 1.

Para a aplicação dos testes multivariados para hipótese de igualdade dos efeitos genéticos do modelo, são utilizados quatro critérios apresentados na Tabela 2. Esses critérios podem ser aproximados para uma distribuição de F .

TABELA 1 – Esquema da análise de variância multivariada para o modelo de análise dialéctica proposto por Gardner e Eberhart (1966).

FV	GL	Matriz de SQP
Genitor	(p-1)	SQP (genitor)
Heterose	p(p-1)/2	SQP (heterose)
Heterose média	1	SQP (heterose média)
Heterose do genitor	(p-1)	SQP (heterose do genitor)
Heterose específica	p(p-3)/2	SQP (heterose específica)
Resíduo	v ¹	SQP (resíduo) ²

¹n é o número de graus de liberdade do resíduo.

²SQP (resíduo) é a matriz de soma de quadrados e produtos do resíduo dividida pelo número de observações que deram origem às médias da tabela dialéctica, obtida em análise de variância multivariada preliminar.

TABELA 2 – Estatísticas multivariadas e suas equivalências aproximadas com a distribuição de F.

Critério	Estatística	Aproximação de F	GL de F
Wilks	$\Lambda = \frac{ E }{ H+E } = \prod \frac{1}{1+\lambda_i}$	$F = \left(\frac{1-\Lambda^{t-1}}{\Lambda^{t-1}} \right) \left(\frac{rt-2f}{pq} \right)$	$v_1 = pq$ $v_2 = rt - 2f$
Traço de Pillai	$V = \text{tr}[H(H+E)^{-1}] = \sum \frac{\lambda_i}{1+\lambda_i}$	$F = \left(\frac{V}{s-V} \right) \left(\frac{2n+s+1}{2m+s+1} \right)$	$v_1 = s(2m+s+1)$ $v_2 = 2(2n+s+1)$
Raiz máxima de Roy	$\theta = \lambda_i$	$F = \frac{\theta(v-d+q)}{d}$	$v_1 = d$ $v_2 = v - d + q$
Traço de Hotelling-Lawley	$U = \text{tr}(HE^{-1}) = \sum \lambda_i$	$F = \frac{2(sn+1)U}{s^2(2m+s+1)}$	$v_1 = s(2m+s+1)$ $v_2 = 2(sn+1)$

em que p: número de variáveis = posto(H + E); q: GL de tratamento (ou do contraste); v : GL do resíduo; s = min(p, q); r = v - (p - q + 1)/2; f = (pq - 2)/4; d = max(p, q); m = (|p - q| - 1)/2; n = (v - p - 1)/2; H: matriz de somas de quadrados e produtos da fonte de variação da hipótese nula (H₀); E: matriz de somas de quadrados e produtos do resíduo médio; e

$$t = \begin{cases} \sqrt{\frac{p^2q^2 - 4}{p^2 + q^2 - 5}}, & \text{se } (p^2 + q^2 - 5) > 0 \\ 1, & \text{caso contrário.} \end{cases}$$

RESULTADOS E DISCUSSÃO

As expressões dos estimadores dos efeitos da constante ($m^{(t)}$), dos genitores ($v_i^{(t)}$ e $v_j^{(t)}$) e das heterosises ($\bar{h}^{(t)}$, $h_i^{(t)}$, $h_j^{(t)}$ e $s_{ij}^{(t)}$) são dadas por:

$$(i) \hat{m}^{(t)} = \frac{1}{p} Y_G^{(t)}$$

$$(ii) \hat{v}_i^{(t)} = Y_{ii}^{(t)} - \frac{1}{p} Y_G^{(t)}$$

$$(iii) \bar{h}^{(t)} = \frac{1}{p(p-1)} (Y_{..}^{(t)} + Y_H^{(t)} - p Y_G^{(t)})$$

$$(iv) \hat{h}_i^{(t)} = \frac{1}{p-2} \left(Y_{i.}^{(t)} - \frac{p}{2} Y_{ii}^{(t)} - \frac{2}{p} Y_H^{(t)} + \frac{p-2}{2p} Y_G^{(t)} \right)$$

$$(v) \hat{s}_{ij}^{(t)} = Y_{ij}^{(t)} - \frac{1}{p-2} \left[(Y_{i.}^{(t)} - Y_{ii}^{(t)}) + (Y_{.j}^{(t)} - Y_{jj}^{(t)}) - \frac{2}{p-1} Y_H^{(t)} \right]$$

em que:

$Y_G^{(t)}$: somatório dos valores médios dos genitores, para a variável t;

$Y_H^{(t)}$: somatório dos valores médios das combinações híbridas, para a variável t;

$Y_{..}^{(t)}$: somatório dos valores médios dos genitores e das combinações híbridas, para a variável t;

$Y_{i.}^{(t)}$: somatório dos valores médios de um genitor i e de suas respectivas combinações híbridas, para a variável t;

$Y_{ii}^{(t)}$: valor médio do genitor i, para a variável t.

As expressões das variâncias dos efeitos genéticos do modelo e das variâncias entre contrastes dos efeitos são dadas por:

$$(i) \hat{V}(\hat{m}^{(t)}) = \frac{1}{p} \frac{QMR^{(t)}}{r}$$

$$(ii) \hat{V}(\hat{v}_i^{(t)}) = \frac{p-1}{p} \frac{QMR^{(t)}}{r}$$

$$(iii) \hat{V}(\hat{v}_i^{(t)} - \hat{v}_j^{(t)}) = 2 \frac{QMR^{(t)}}{r}$$

$$(iv) \hat{V}(\bar{h}^{(t)}) = \frac{p+1}{p(p-1)} \frac{QMR^{(t)}}{r}$$

$$(v) \hat{V}(\hat{h}_i^{(t)}) = \frac{(p-1)(p+2)}{4p(p-2)} \frac{QMR^{(t)}}{r}$$

$$(vi) \hat{V}(\hat{h}_i^{(t)} - \hat{h}_j^{(t)}) = \frac{p+2}{2(p-2)} \frac{QMR^{(t)}}{r}$$

$$(vii) \hat{V}(\hat{s}_{ij}^{(t)}) = \frac{p-3}{p-1} \frac{QMR^{(t)}}{r}$$

$$(viii) \hat{V}(\hat{s}_{ij}^{(t)} - \hat{s}_{ik}^{(t)}) = \frac{2(p-3)}{p-2} \frac{QMR^{(t)}}{r}$$

$$(ix) \hat{V}(\hat{s}_{ij}^{(t)} - \hat{s}_{km}^{(t)}) = \frac{2(p-4)}{p-2} \frac{QMR^{(t)}}{r}$$

em que:

$QMR^{(t)}$: quadrado médio do resíduo obtido em análise de variância univariada preliminar, para a variável t;

r: número de observações que deram origem aos valores médios da tabela dialéctica.

Observa-se que as expressões dos estimadores dos efeitos genéticos do modelo, de suas variâncias e das variâncias entre contrastes dos efeitos foram as mesmas obtidas, considerando o modelo univariado de Gardner e Eberhart (1966).

As matrizes de SQP resultantes do desenvolvimento de $R(\cdot/\cdot)$, para as diferentes reduções consideradas, apresentam dimensão $k \times k$. Na diagonal da matriz, têm-se as somas de quadrados e, fora da diagonal, têm-se as somas de produtos. Seja o elemento a_{mn} pertencente a essa matriz de SQP, em que o m representa a m-ésima linha e o n representa a n-ésima coluna, para $m, n = 1, 2, \dots, t, \dots, k$.

Os elementos da diagonal (a_{tt}) e os demais elementos (a_{mn}) da matriz de SQP para efeito do genitor são dados por:

$$a_{tt} = \frac{1}{p+2} \left[\sum_i (Y_{ii}^{(t)} + Y_{i.}^{(t)})^2 - \frac{4}{p} Y_{..}^{2(t)} \right]$$

e

$$a_{mn} = \frac{1}{p+2} \left[\sum_i (Y_{ii}^{(m)} + Y_{i.}^{(m)})(Y_{ii}^{(n)} + Y_{i.}^{(n)}) - \frac{4}{p} Y_{..}^{(m)} Y_{..}^{(n)} \right]$$

A matriz de SQP para o efeito da heterose foi dada por:

$$a_{tt} = \sum_i \sum_{i \leq j} Y_{ij}^{2(t)} - \frac{1}{p+2} \sum_i (Y_{ii}^{(t)} + Y_{i\cdot}^{(t)})^2 + \frac{2}{(p+1)(p+2)} Y_{..}^{2(t)}$$

e

$$a_{mn} = \sum_i \sum_{i \leq j} Y_{ij}^{(m)} Y_{ij}^{(n)} - \frac{1}{p+2} \sum_i (Y_{ii}^{(m)} + Y_{i\cdot}^{(m)})(Y_{ii}^{(n)} + Y_{i\cdot}^{(n)}) + \frac{2}{(p+1)(p+2)} Y_{..}^{2(m)} Y_{..}^{2(n)}$$

Para a matriz de SQP, para o efeito da heterose média, obtiveram-se:

$$a_{tt} = \frac{1}{p} Y_G^{2(t)} + \frac{2}{p(p-1)} Y_H^{2(t)} - \frac{2}{p(p+1)} Y_{..}^{2(t)}$$

e

$$a_{mn} = \frac{1}{p} Y_G^{(m)} Y_G^{(n)} + \frac{2}{p(p-1)} Y_H^{(m)} Y_H^{(n)} - \frac{2}{p(p+1)} Y_{..}^{(m)} Y_{..}^{(n)}$$

O efeito da heterose do genitor é dado por:

$$a_{tt} = \sum_i Y_{ii}^{2(t)} - \frac{1}{p} Y_G^{2(t)} + \frac{1}{p-2} \left[\sum_i (Y_{i\cdot}^{(t)} - Y_{ii}^{(t)})^2 - \frac{4}{p} Y_H^{2(t)} \right] - \frac{1}{p+2} \left[\sum_i (Y_{ii}^{(t)} + Y_{i\cdot}^{(t)})^2 - \frac{4}{p} Y_{..}^{2(t)} \right]$$

e

$$a_{mn} = \sum_i Y_{ii}^{(m)} Y_{ii}^{(n)} - \frac{1}{p} Y_G^{(m)} Y_G^{(n)} + \frac{1}{p-2} \left[\sum_i (Y_{i\cdot}^{(m)} - Y_{ii}^{(m)})(Y_{i\cdot}^{(n)} - Y_{ii}^{(n)}) - \frac{4}{p} Y_H^{(m)} Y_H^{(n)} \right] - \frac{1}{p+2} \left[\sum_i (Y_{ii}^{(m)} + Y_{i\cdot}^{(m)})(Y_{ii}^{(n)} + Y_{i\cdot}^{(n)}) - \frac{4}{p} Y_{..}^{(m)} Y_{..}^{(n)} \right]$$

Para o efeito da heterose específica, têm-se:

$$a_{tt} = \sum_i \sum_{i < j} Y_{ij}^{2(t)} - \frac{1}{p-2} \sum_i (Y_{i\cdot}^{(t)} - Y_{ii}^{(t)})^2 + \frac{2}{(p-1)(p-2)} Y_H^{2(t)}$$

e

$$a_{mn} = \sum_i \sum_{i < j} Y_{ij}^{(m)} Y_{ij}^{(n)} - \frac{1}{p-2} \sum_i (Y_{i\cdot}^{(m)} - Y_{ii}^{(m)})(Y_{i\cdot}^{(n)} - Y_{ii}^{(n)}) + \frac{2}{(p-1)(p-2)} Y_H^{(m)} Y_H^{(n)}$$

Observa-se que as expressões das somas de quadrados dos efeitos genéticos do modelo no caso univariado são equivalentes aos elementos da diagonal da matriz de somas de quadrados e produtos para o caso multivariado, considerando uma determinada variável t . Para os demais elementos fora da diagonal, tem-se a soma de produtos entre duas variáveis consideradas.

A extensão multivariada para o modelo de Gardner e Eberhart (1966) foi facilmente obtida pela derivação do modelo linear $Y = X\beta + \varepsilon$, impondo-se restrições convenientes. Geneticamente, além de proporcionarem considerável simplificação, as restrições proporcionam estimadores cujo significado biológico tem sido de grande interesse para os geneticistas

(CRUZ e REGAZZI, 1997). Segundo Camussi et al. (1985), o método de análise multivariada pode ser facilmente aplicado para uma larga variedade de modelos genéticos, ressaltando-se apenas os casos em que os efeitos ambientais não estão suficientemente controlados para assegurar a homocedasticidade dos dados.

Uma vantagem da extensão multivariada, quando comparada com a metodologia univariada tradicional, é a possibilidade de se estimar as matrizes de variâncias e covariâncias dos efeitos genéticos do modelo. Dessa forma, podem-se estimar as correlações fenotípicas e genotípicas entre esses efeitos, podendo-se, conseqüentemente, utilizar tais informações para a orientação de programas de melhoramento.

De maneira geral, as informações fornecidas pela modelagem univariada são contempladas pela extensão multivariada, sem, no entanto, levar em consideração um nível de significância conjunto e o aproveitamento das correlações existentes entre as variáveis (DEMÉTRIO, 1985). Para utilização da informação da análise de variância multivariada, o melhorista pode utilizar uma ampla gama de metodologias multivariadas para complementar os resultados obtidos e propiciar subsídios para utilização prática das informações resultantes. Dentre as várias possibilidades, destacam-se as variáveis canônicas, os componentes principais, a análise de fatores e os índices de seleção de Smith e Hazel (CRUZ e REGAZZI, 1997).

Essas metodologias multivariadas podem ser aplicadas a cada efeito genético do modelo (genitor, heterose média, heterose do genitor, heterose específica), tanto nas estimativas fenotípicas como nas matrizes de variâncias e covariâncias genéticas correspondentes. Para essa última situação, são necessárias as obtenções das esperanças das matrizes de variâncias e covariâncias de cada efeito genético do modelo submetido à análise de variância. Essas expressões, infelizmente, não estão disponíveis na literatura e podem ser alvo de trabalhos futuros.

Para ilustrar como aplicar essas técnicas e considerando que o objetivo do melhorista tenha foco na heterose específica, é considerada a representação da matriz de variância e covariância dessa fonte de variação por S_{HE} . Para se construir um índice com o qual o melhorista possa selecionar combinações híbridas mais promissoras, pode-se pensar em utilizar a técnica dos componentes principais. Assim, busca-se uma combinação linear dos efeitos genéticos $S_{ij}^{(t)}$ de um determinado híbrido ij com relação as $t = 1, 2, \dots, k$ variáveis.

Essa combinação linear deve maximizar a variabilidade total contida em S_{HE} . Para isso, é necessário resolver o sistema de equações homogêneo dado por:

$$(S_{HE} - \lambda_t I) \underline{e}_t = \underline{0}$$

em que λ_t e \underline{e}_t são os pares de autovalores e autovetores de S_{HE} , com $t = 1, 2, 3, \dots, k$.

Uma melhor alternativa para esse método seria aplicar a técnica das variáveis canônicas. A vantagem adicional dessa técnica em relação aos componentes principais é incorporar a informação residual para gerar a combinação linear almejada. O sistema de equações homogêneo formado é dado por:

$$(S_{HE} - \lambda_t S) \underline{e}_t = \underline{0}$$

em que S é a estimativa da matriz de variâncias e covariâncias do resíduo e os demais termos são definidos como anteriormente. As combinações lineares dos efeitos genotípicos $S_{ij}^{(t)}$, para i e j fixados, são obtidos pelos elementos de \underline{e}_t , cuja variabilidade captada do sistema é dada pela magnitude de λ_t , com $t = 1, 2, \dots, k$.

Uma terceira alternativa a essas metodologias é a técnica de análise de fatores. Essa é uma técnica refinada que poderia trazer informações adicionais da importância de cada variável para os fatores retidos na explicação de S_{HE} , por meio das comunalidades.

Finalmente, se as estimativas dos componentes genéticos da heterose específica (S_{GHE}) estiverem disponíveis, é possível estipular um índice de seleção baseado na teoria do índice clássico de Smith e Hazel (CRUZ e REGAZZI, 1997) por:

$$\underline{b} = S_{HE}^{-1} S_{GHE} \underline{a}$$

em que \underline{a} é o vetor de pesos econômicos e \underline{b} é o vetor dos coeficientes do índice de seleção que estipula a combinação linear dos efeitos genéticos de $S_{ij}^{(t)}$.

Todas essas técnicas podem ser utilizadas em complementação à análise de variância multivariada e se constituem nas vantagens das abordagens multivariadas em relação às técnicas univariadas. No entanto, a utilização de algumas dessas metodologias só é possível com a obtenção dos componentes de variância para esse modelo, o que não foi encontrado na literatura. Entretanto, algumas das análises, complementares à análise de variância multivariada, podem ser prontamente realizadas. É conveniente salientar que, se o objetivo do melhorista for centrado no efeito do genitor, ou de hete-

rose média ou de heterose do genitor, as técnicas descritas anteriormente também se aplicam. Para isso, basta substituir S_{HE} pelas correspondentes matrizes de variâncias e covariâncias de interesse.

CONCLUSÕES

Expressões para somas de quadrados e produtos, estimadores dos efeitos genéticos, variâncias dos efeitos e dos contrastes entre efeitos e testes de hipóteses foram obtidos.

A análise de variância multivariada para os cruzamentos dialélicos pode ser realizada para estimar a heterose em várias características, simultaneamente.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABREU, A. F. B. **Predição do potencial genético de populações segregantes do feijoeiro utilizando genitores inter-raciais**. 1997. 79 f. Tese (Doutorado em Genética e Melhoramento de Plantas) – Universidade Federal de Lavras, Lavras, 1997.
- BAENZIGER, P. S.; PETERSON, E. J. Genetic variation: its origin and use for breeding self-pollinated species. In: STALKER, H. T.; MULTRIPHY, J. P. **Plant breeding in the 1990's**. North Carolina: North Carolina State University, 1991. p. 69-100.
- BERNARDO, R. Best linear unbiased prediction of maize single-cross performance. **Crop Science**, Madison, v. 36, n. 1, p. 50-56, Jan./Feb. 1996.
- CAMUSSI, A.; OTTAVIANO, E.; CALINSKY, T.; KACZMAREK, Z. Genetic distance based on quantitative traits. **Genetics**, Austin, v. 111, n. 4, p. 945-962, 1985.
- CHARCOSSET, A.; BONISSEAU, B.; TOUCHEBEUF, O.; BURSTIN, J.; DUBREUIL, P.; BARRIERE, Y.; GALLAIS, A.; DENIS, J. B. Prediction of maize hybrid silage performance using marker data: comparison of several models for specific combining ability. **Crop Science**, Madison, v. 38, n. 1, p. 38-44, Jan./Feb. 1998.
- CRUZ, C. D. **Aplicação de algumas técnicas multivariadas no melhoramento de plantas**. 1990. 188 f. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 1990.
- CRUZ, C. D.; REGAZZI, A. J. **Modelos biométricos aplicados ao melhoramento genético**. 2. ed. Viçosa: UFV, 1997. 390 p.
- DEMÉTRIO, C. G. B. **Análise multidimensional para dados de cana-de-açúcar**. 1985. 144 f. Tese (Doutorado em Genética e Melhoramento) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 1985.
- GARDNER, C. O.; EBERHART, S. A. Analysis and interpretation of the variety cross diallel and related population. **Biometrics**, Washington, v. 22, n. 3, p. 439-452, Sept. 1966.
- GRIFFING, B. Concept of general and specific combining ability in relation to diallel crossing systems. **Australian Journal of Biological Sciences**, Melbourne, v. 9, p. 463-493, 1956.
- MARTINEZ GARZA, A. **Diseños y analysis de experimentos de cruza dialélicas**. Chapingo: Centro de Estadística y Cálculo, 1983. 252 p.
- RAMALHO, M. A. P.; SANTOS, J. B. S.; ZIMMERMANN, M. J. O. **Genética quantitativa em plantas autógamas: aplicações ao melhoramento do feijoeiro**. Goiânia: UFG, 1993. 271 p.
- SAKAGUTI, E. S. **Utilização de técnicas de análise multivariadas na avaliação de cruzamentos dialélicos em coelhos**. 1994. 172 f. Dissertação (Mestrado em Zootecnia) – Universidade Federal de Viçosa, Viçosa, 1994.
- SOUZA, E. A. **Alternativas experimentais na avaliação de progênies em programas de melhoramento genético vegetal**. 1997. 122 f. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 1997.