MODELO DE SUPERFÍCIE DE RESPOSTA COM ALGUMAS VARIÁVEIS AUXILIARES ADICIONAIS[1]

AUGUSTO RAMALHO DE MORAIS[2] e VIVALDO FRANCISCO DA CRUZ[3]

RESUMO - Estudou-se um modelo de regressão polinomial quadrático, através da metodologia de superfície de resposta, considerando-se algumas variáveis auxiliares adicionadas ao modelo, com vistas à obtenção de fórmulas que permitam avaliar o comportamento da variável dependente ante a inclusão das variáveis adicionais. Os estimadores dos parâmetros (tratamentos e regressão) foram obtidos através do método dos quadrados mínimos. Desenvolveu-se a seqüência de operações para a realização da análise estatística, considerando-se o modelo adaptado a um esquema fatorial completo para três fatores e três níveis eqüidistantes. Foram determinados ainda: os estimadores das variâncias e covariâncias dos parâmetros e a análise da variância.

Termos para indexação: regressão polinomial múltipla.

RESPONSE SURFACE MODEL WITH SOME ADDITIONAL AUXILIARY VARIABLES

ABSTRACT - A quadratic polynomial model was studied, using response surface method, with some additional variables included in the model, in order to obtain equations to evaluate the dependent variable behavior after the inclusion of the additional variables. The parameter estimators (treatment and regression) were obtained using the least squares method. An algorithm was developed to perform the statistical analysis, considering the model adapted to a complete factorial design with three factors and three levels. In addition, estimators of variance and covariance of the parameters and variance analysis were determined.

Index term: multiple polynomial regression.

INTRODUÇÃO

A metodologia de superfície de resposta é comumente utilizada com a finalidade de determinar as melhores condições ante a variação de fatores aplicados a determinado fenômeno. Sua aplicação deu-se inicialmente na indústria química, tendo sido seus fundamentos formalizados por Box & Wilson (1951). No campo agronômico, seu uso concentrou-se no estudo do rendimento de cultivares como efeito de níveis de nutrientes apliados ao solo, incluindo-se outros fatores, como: densidade de plantio e níveis de doenças (Gomez et al. 1978).

A pesquisa tem-se preocupado em buscar solução racional de utilização de fertilizantes pelos agricultores, buscando, também, os modos mais eficientes para planejar a procura dessas soluções, a partir de delineamentos experimentais que utilizem todas as combinações dos fatores, ou parte delas.

A utilização de modelos matemáticos para expressar a resposta aos nutrientes, sob a forma de superfície, teve um grande progresso a partir da introdução dos fatoriais fracionados (Finney, 1945) e dos delineamentos compostos centrais e compostos rotacionais (Box & Wilson 1951, Box & Hunter 1957).

Segundo Myers (1971), as pressuposições básicas para o estudo de superfícies de resposta são: a função existe, mas é desconhecida, sendo as variáveis envolvidas quantitativas e contínuas; a função pode ser aproximada dentro de uma região de interesse, e as variáveis independentes são controladas pelo experimento e medidas sem erro.

Cochran & Cox (1957) consideraram que o método mais informativo para analisar os resultados de um ensaio fatorial depende da natureza dos fatores: se estes representam variáveis

[1] Aceito para publicação em 2 de agosto de 1988.
Parte da dissertação apresentada pelo primeiro autor à ESALQ/USP para obtenção do título de Mestre em Agronomia. Área de concentração: Estatística e Experimentação Agronômica.

[2] Eng. - Agr., M.Sc., EMBRAPA/Centro Nacional de Pesquisa de Milho e Sorgo (CNPMS), Caixa Postal 151, CEP 35700 Sete Lagoas, MG.

[3] Eng. - Agr., Dr., Prof.-Assist., ESALQ/USP, Dep. Matemática e Estatística, Caixa Postal 09, CEP 13400 Piracicaba, SP.

quantitativas, a análise mais informativa consiste em relacionar as respostas como função dessas variáveis.

Box & Draper (1959) discutiram várias razões na escolha de um delineamento para o estudo de superfícies de resposta; incluem simples interesse na superfície, estimação de parâmetros e discriminação entre modelos.

Vários trabalhos são relatados na literatura a respeito das aplicações da superfície de resposta no campo agronômico. Entre outros, Campos (1967), Vieira (1970) e Costa (1977) verificaram que as estimativas dos parâmetros obtidas nas superfícies ajustadas são pouco precisas, com intervalos de confiança bastante amplos, dificultando, desse modo, a recomendação de fórmulas e predição de rendimentos. Por outro lado, uma análise, considerando variáveis auxiliares que estão linearmente relacionadas à variável dependente, pode contribuir para aumentar a precisão experimental e, conseqüentemente, obter estimativas mais eficientes, como ocorre nas análises de covariâncias (Cochran 1957).

Desse modo, o objetivo do presente trabalho foi apresentar um método de análise de um modelo de regressão polinomial quadrático, adaptado a um esquema fatorial completo para três fatores com três níveis eqüidistantes, considerando-se algumas variáveis auxiliares adicionais, através da metodologia de superfície de resposta.

MATERIAL E MÉTODOS

A metodologia de superfície de resposta foi aplicada a um ensaio fatorial completo 3^3, com níveis eqüidistantes de três quaisquer fatores A, B e C (por ex. N, P_2O_5 e K_2O), levando-se em consideração algumas variáveis auxiliares adicionais, usando-se o seguinte modelo matemático:

$$y_{ijk} = a_{00} + a_{10}x_1 + a_{20}x_2 + a_{30}x_3 + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + a_{12}x_1x_2$$

$$a_{13}x_1x_3 + a_{23}x_2x_3 + b_1t_1 + \dots + b_Ht_H + e_{ijk}$$

onde:
x_1 representa os níveis de N (fator A);
x_2 representa os níveis de P_2O_5 (fator B);
x_3 representa os níveis de K_2O_5 (fator C);
a_{00} representa a média geral;
$a_{10}, a_{20}, \dots, a_{23}$ são os coeficientes dos parâmetros;
t_h representa o valor da h-ésima variável auxiliar adicional ($h = 1, 2, \dots, H$);
b_h é o coeficiente de regressão associado à h-ésima variável auxiliar adicional;
e_{ijk} é o erro experimental associado a y_{ijk}, tal que $e \sim N(0, \sigma^2)$.

Para facilidade de cálculos, os valores dos níveis de nutrientes e das variáveis auxiliares adicionais foram transformados, respectivamente, para:

$x_i = (d_i - \bar{d})/q$ e $t_h = w_h - \bar{w}_h$

onde:
d_i são os valores dos níveis de nutrientes, ($i = 1, 2, 3$), associados às variáveis X_1, X_2 e X_3;
$\bar{d}$ média desses valores;
q diferença entre os níveis;
w_h são os valores da h-ésima variável auxiliar adicional;
$\bar{w}_h$ média desses valores.

Os parâmetros correspondentes aos efeitos quadráticos puros não são independentes da média; desse modo, fez-se nova parametrização, e o modelo ficou da forma:

$$y = a'_{00} + a_{10}x_1 + a_{20}x_2 + a_{30}x_3 + a_{11}[x_1^2 - (\Sigma x_1^2)/n] + a_{22}[x_2^2 - (\Sigma x_2^2)/n] +$$

$$a_{33}[x_3^2 -)\Sigma x_3^2)/n] + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + b_1t_1 + \ldots + b_Ht_H + e$$

sendo: n é o número de níveis;

$$a'_{00} = a_{00} + a_{11}(\Sigma x_1^2)/n + a_{22}(\Sigma x_2^2)/n + a_{33}(\Sigma x_2^2)/n$$

Em forma matricial, o modelo adotado, é:

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \ ,$$

onde: $\underset{\sim}{Y}$ é o vetor das observações, X é a matriz dos coeficientes dos parâmetros, $\underset{\sim}{\beta}$ é o vetor dos parâmetros e $\underset{\sim}{\varepsilon}$ é o vetor dos erros experimentais, sendo que $\underset{\sim}{\varepsilon} \sim N(\phi, \sigma^2 I)$.

O sistema de equações normais obtido através do método dos quadrados mínimos é dado por $X'X\underset{\sim}{\beta} = X'Y$ e, pode ser parcionado de modo conveniente, em

$$\begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix} \begin{bmatrix} \hat{\underset{\sim}{\tau}} \\ \hat{\underset{\sim}{b}} \end{bmatrix} = \begin{bmatrix} X_1'\underset{\sim}{Y} \\ X_2'\underset{\sim}{Y} \end{bmatrix}$$

onde:
X_1 é a matriz dos valores assumidos por x_1, x_2 e x_3, de dimensão N x p;
X_2 é a matriz dos valores das variáveis auxiliares adicionais, de dimensão N x H;
$\hat{\underset{\sim}{\tau}}$ vetor dos efeitos de tratamentos, de dimensão p x 1;
$\hat{\underset{\sim}{b}}$ vetor dos coeficientes de regressão associados as variáveis auxiliares adicionais, de dimensão H x 1;
N é o número total de observações e p o número de parâmetros de tratamentos.

O sistema de equações acima pode ser representado por:

$$\begin{bmatrix} R & T_x \\ T_x' & L \end{bmatrix} \begin{bmatrix} \hat{\underset{\sim}{\tau}} \\ \hat{\underset{\sim}{b}} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{T} \\ \underset{\sim}{B} \end{bmatrix}$$

onde:
R matriz diagonal dos coeficientes dos tratamentos (modelo original), de dimensão p x p;
T_x matriz dos totais de tratamentos relativos às variáveis auxiliares adicionais, de dimensão p x H;
L matriz das somas dos quadrados e dos produtos das variáveis auxiliares adicionais, de dimensão H x H;
$\underset{\sim}{T}$ vetor dos totais de tratamentos da variável dependente, de dimensão p x 1;
$\underset{\sim}{B}$ vetor das somas de produtos da variável dependente com as variáveis auxiliares adicionais, de dimensão H x 1;

sendo:
R = diagonal r[27 18 18 18 6 6 6 12 12 12]

$$L = \begin{bmatrix} \Sigma t_1^2 & \Sigma t_1 t_2 & \dots & \Sigma t_1 t_H \\ \Sigma t_2 t_1 & \Sigma t_2^2 & \dots & \Sigma t_2 t_H \\ \dots & \dots & \dots & \dots \\ \Sigma t_H t_1 & \Sigma t_H t_2 & \dots & \Sigma t_H^2 \end{bmatrix} \qquad \underset{\sim}{B} = \begin{bmatrix} \Sigma y t_1 \\ \Sigma y t_2 \\ \dots \\ \Sigma y t_H \end{bmatrix}$$

$$T_X = \begin{bmatrix} 0 & \dots & 0 \\ t_{12..} - t_{10..} & \dots & t_{H2..} - t_{H0..} \\ t_{1.2.} - t_{1.0.} & \dots & t_{H.2.} - t_{H.0.} \\ t_{1..2} - t_{1..0} & \dots & t_{H..2} - t_{H..0} \\ (t_{10..} - 2t_{11..} + t_{12..})/3 & \dots & (t_{H0..} - 2t_{H1..} + t_{H2..})/3 \\ (t_{1.0.} - 2t_{1.1.} + t_{1.2.})/3 & \dots & (t_{H.0.} - 2t_{H.1.} + t_{H.2.})/3 \\ (t_{1..0} - 2t_{1..1} + t_{1..2})/3 & \dots & (t_{H..0} - 2t_{H..1} + t_{H..2})/3 \\ t_{100.} + t_{122.} - t_{102.} - t_{120.} & \dots & t_{H00.} + t_{H22.} - t_{H02.} - t_{H20.} \\ t_{10.0} + t_{12.2} - t_{10.2} - t_{12.0} & \dots & t_{H0.0} + t_{H2.2} - t_{H0.2} - t_{H2.0} \\ t_{1.00} + t_{1.22} - t_{1.02} - t_{1.20} & \dots & t_{H.00} + t_{H.22} - t_{H.02} - t_{H.20} \end{bmatrix}$$

$$\underset{\sim}{T} = \begin{bmatrix} \Sigma Y_{ijk} \\ Y_{2..} - Y_{0..} \\ Y_{.2.} - Y_{.0.} \\ Y_{..2} - Y_{..0} \\ (Y_{0..} - 2Y_{1..} + Y_{2..})/3 \\ (Y_{.0.} - 2Y_{.1.} + Y_{.2.})/3 \\ (Y_{..0} - 2Y_{..1} + Y_{..2})/3 \\ Y_{00.} + Y_{22.} + Y_{02.} - Y_{20.} \\ Y_{0.0} + Y_{2.2} - Y_{0.2} - Y_{2.0} \\ Y_{.00} + Y_{.22} - Y_{.02} - Y_{.20} \end{bmatrix} \quad \underset{\sim}{\hat{\tau}} = \begin{bmatrix} \hat{a}_{00} \\ \hat{a}_{10} \\ \hat{a}_{20} \\ \hat{a}_{30} \\ \hat{a}_{11} \\ \hat{a}_{22} \\ \hat{a}_{33} \\ \hat{a}_{12} \\ \hat{a}_{13} \\ \hat{a}_{23} \end{bmatrix} \qquad \underset{\sim}{\hat{b}} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \vdots \\ \hat{b}_H \end{bmatrix}$$

Para facilidade de notação, os índices relativos aos tratamentos foram suprimidos em algumas matrizes, na indicação de somatórios. r é o número de repetições; $t_{h.oo}$ soma das parcelas que possuem o nível o dos fatores B e C simultaneamente da h-ésima variável auxiliar adicional; $Y_{o..}$ soma das parcelas que possuem o nível o do primeiro fator (A) da variável dependente; os demais, por analogia.

RESULTADOS

Considerando-se que as matrizes R e L são não-singulares, o vetor dos estimadores dos efeitos de tratamentos ajustados para a regressão linear é dado por:

$$\hat{\underset{\sim}{\tau}} = R^{-1}\underset{\sim}{T} - R^{-1}T_{X}\underset{\sim}{\hat{b}}$$

sendo que, no caso específico considerado, os estimadores são:

$$\hat{a}_{00} = \frac{1}{27}\Sigma Y_{ijk}$$

$$\hat{a}_{10} = \frac{Y_{2..} - Y_{0..}}{18} - \frac{\hat{b}_1(t_{12..} - t_{10..}) + \hat{b}_2(t_{22..} - t_{20..}) + \ldots + \hat{b}_H(t_{H2..} - t_{H0..})}{18}$$

$$\hat{a}_{20} = \frac{Y_{.2.} - Y_{.0.}}{18} - \frac{\hat{b}_1(t_{1.2.} - t_{1.0.}) + \hat{b}_2(t_{2.2.} - t_{2.0.}) + \ldots + \hat{b}_H(t_{H.2.} - t_{H.0.})}{18}$$

$$\hat{a}_{30} = \frac{Y_{..2} - Y_{..0}}{18} - \frac{\hat{b}_1(t_{1..2} - t_{1..0}) + \hat{b}_2(t_{2..2} - t_{2..0}) + \ldots + \hat{b}_H(t_{H..2} - t_{H..0})}{18}$$

$$\hat{a}_{11} = \frac{Y_{2..} - Y_{1..} + Y_{2..}}{18} - \frac{\hat{b}_1(t_{10..} - 2t_{11..} + t_{12..}) + \ldots + \hat{b}_H(t_{H0..} - 2t_{H1..} + t_{H2..})}{18}$$

$$\hat{a}_{22} = \frac{Y_{.0.} - Y_{.1.} + Y_{.2.}}{18} - \frac{\hat{b}_1(t_{1.0.} - 2t_{1.1.} + t_{1.2.}) + \ldots + \hat{b}_H(t_{H.0.} - 2t_{H.1.} + t_{H.2.})}{18}$$

$$\hat{a}_{33} = \frac{Y_{..0} - Y_{..1} + Y_{..2}}{18} - \frac{\hat{b}_1(t_{1..0} - 2t_{1..1} + t_{1..2}) + \ldots + \hat{b}_H(t_{H..0} - 2t_{H..1} + t_{H..2})}{18}$$

$$\hat{a}_{12} = \frac{Y_{00.} + Y_{22.} - Y_{02.} - Y_{20.}}{12} - [\hat{b}_1(t_{100.} + t_{122.} - t_{102.} - t_{120.}) + \ldots$$

$$\ldots + \hat{b}_H(t_{H00.} + t_{H22.} - t_{H02.} - t_{H20.})]/12$$

$$\hat{a}_{13} = \frac{Y_{0.0} + Y_{2.2} - Y_{0.2} - Y_{2.0}}{12} - [\hat{b}_1(t_{10.0} + t_{12.2} - t_{10.2} - t_{12.0}) + \ldots$$

$$\ldots + \hat{b}_H(t_{H0.0} + t_{H2.2} - t_{H0.2} - t_{H2.0})]/12$$

$$\hat{a}_{23} = \frac{Y_{.00} + Y_{.22} - Y_{.02} - Y_{.20}}{12} - [\hat{b}_1(t_{1.00} + t_{1.22} - t_{1.02} - t_{1.20}) + \ldots$$

$$\ldots + \hat{b}_H(t_{H.00} + t_{H.22} - t_{H.02} - t_{H.20})]/12$$

O vetor dos estimadores dos coeficientes de regressão é dado por:

$$\underset{\sim}{b} = E_1^{-} \underset{\sim}{E} \quad , \text{ sendo: } E_1 = L - T_x' R^{-1} T_x \text{ e } \underset{\sim}{E}_0 = \underset{\sim}{B} - T_x' R^{-1} \underset{\sim}{T} \text{ ou}$$

$$E_1 = \begin{bmatrix} \text{SQRes } t_1 & \text{SPRes } t_1 t_2 & \cdots & \text{SPRes } t_1 t_H \\ & \text{SQRes } t_2 & \cdots & \text{SPRes } t_2 t_H \\ & & & \cdots \\ \text{Simétrica} & & & \text{SQRes } t_H \end{bmatrix} \quad \underset{\sim}{E}_0 = \begin{bmatrix} \text{SPRes } Yt_1 \\ \text{SPRes } Yt_2 \\ \cdots \\ \text{SPRes } Yt_H \end{bmatrix}$$

ou seja, E é a matriz das somas de quadrados e de produtos do resíduo das variáveis auxiliares adicionais, e $\underset{\sim}{E}$ é o vetor das somas de produtos do resíduo da variável dependente com as variáveis auxiliares adicionais.

Análise da variância

A soma de quadrados de parâmetros é obtida por:

$$SQP = \underset{\sim}{\beta}' X'Y' = \underset{\sim}{\hat{t}}'\underset{\sim}{T} + \underset{\sim}{b}'\underset{\sim}{B} = \underset{\sim}{T}'R^{-1}\underset{\sim}{T} + \underset{\sim}{b}'\underset{\sim}{E}_0$$

onde: $\underset{\sim}{T}'\underset{\sim}{R}^{-1}\underset{\sim}{T}$ é a soma de quadrados de tratamentos e, $\underset{\sim}{b}'\underset{\sim}{E}_0$ a soma de quadrados de regressão linear.

A soma de quadrados de resíduo ajustada para regressão será:

$$SQR(aj) = \underset{\sim}{Y}'\underset{\sim}{Y} - \underset{\sim}{T}'R^{-1}\underset{\sim}{T} - \underset{\sim}{b}'\underset{\sim}{E}_0$$

As somas de quadrados dos tratamentos, ajustada para regressão linear, foram obtidas utilizando-se o método do resíduo condicional (Graybill 1961). Por exemplo, para o efeito de nitrogênio quadrático, o modelo matemático reduzido é o seguinte:

$$Y = a'_{00} + a_{10}x_1 + a_{20}x + a_{30}x_3 + a_{22}[x_2^2 - (\Sigma x_2^2)/n] + a_{33}[x_3^2 - (\Sigma x_3^2)/n] + a_{12}x_1x_2 +$$

$$+ a_{13}x_1x_3 + a_{23}x_2x_3 + b_1t_1 + \ldots + b_Ht_H + e \quad ,$$

com as mesmas especificações do modelo completo.

A forma matricial desse modelo reduzido é $\underset{\sim}{Y} = Z\underset{\sim}{\Theta} + \varepsilon$, cujo sistema de equações normais é $Z'Z\underset{\sim}{\Theta} = Z'\underset{\sim}{Y}$ ou

$$\begin{bmatrix} X_0'X_0 & X_0'X_2 \\ X_2'X_0 & X_2'X_2 \end{bmatrix} \begin{bmatrix} \hat{\underset{\sim}{\tau}}_0 \\ \hat{\underset{\sim}{b}}_0 \end{bmatrix} = \begin{bmatrix} X_0'\underset{\sim}{Y} \\ X_2'\underset{\sim}{Y} \end{bmatrix}$$

onde: X_0 matriz formada pela exclusão da coluna relativa ao efeito de nitrogênio quadrático da matriz X_1;

$\underset{\sim}{\tau}_0$ vetor $\underset{\sim}{\tau}$ sem o efeito de nitrogênio quadrático.

Representando o sistema de equações normais por:

$$\begin{bmatrix} R_0 & T_{x0} \\ T'_{x0} & L \end{bmatrix} \begin{bmatrix} \hat{\underset{\sim}{\tau}}_0 \\ \hat{\underset{\sim}{b}}_0 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{T}_0 \\ \underset{\sim}{B} \end{bmatrix}$$

onde: R_0 matriz diagonal dos coeficientes dos tratamentos na ausência de nitrogênio quadrático;

T_{x0} matriz dos totais de tratamentos relativos às variáveis adicionais na ausência de nitrogênio quadrático;

$\underset{\sim}{T}_0$ vetor dos totais de tratamentos da variável dependente na ausência de nitrogênio quadrático.

Sabendo-se que as matrizes R_0 e L são não-singulares, os estimadores dos efeitos de tratamentos ajustados para regressão linear, do modelo reduzido, são dados por:

$$\hat{\underset{\sim}{\tau}}_0 = R_0^{-1}\underset{\sim}{T}_0 - R_0^{-1}T_{x0}\hat{\underset{\sim}{b}}_0$$

ou ainda:

$$\hat{a}^*_{00} = \frac{1}{27}\Sigma Y_{ijk}$$

$$\hat{a}^*_{10} = \frac{Y_{2..} - Y_{0..}}{18} - \frac{b^*_1(t_{12..} - t_{10..}) + b^*_2(t_{22..} - t_{20..}) + \ldots + b^*_H(t_{H2..} - t_{H0..})}{18}$$

$$\hat{a}^*_{20} = \frac{Y_{.2.} - Y_{.0.}}{18} - \frac{b^*_1(t_{1.2.} - t_{1.0.}) + b^*_2(t_{2.2.} - t_{2.0.}) + \ldots + b^*_H(t_{H.2.} - t_{H.0.})}{18}$$

$$\hat{a}^*_{30} = \frac{Y_{..2} - Y_{..0}}{18} - \frac{b^*_1(t_{1..2} - t_{1..0}) + b^*_2(t_{2..2} - t_{2..0}) + \ldots + b^*_H(t_{H..2} - t_{H..0})}{18}$$

$$\hat{a}^*_{22} = \frac{Y_{.0.} - 2Y_{.1.} + Y_{.2.}}{18} - \frac{b^*_1(t_{1.0.} - t_{1.1.} + t_{1.2.}) + \ldots + b^*_H(t_{H.0.} - 2t_{H.1.} + t_{H.2.})}{18}$$

$$\hat{a}^*_{33} = \frac{Y_{..0} - 2Y_{..1} + Y_{..0}}{18} - \frac{b^*_1(t_{1..0} - t_{1..1} + t_{1..2}) + \ldots + b^*_H(t_{H..0} - 2t_{H..1} + t_{H..2})}{18}$$

$$\hat{a}^*_{12} = \frac{Y_{00\cdot} + Y_{22\cdot} - Y_{02\cdot} - Y_{20\cdot}}{12} - [b^*_1(t_{100\cdot} + t_{122\cdot} - t_{102\cdot} - t_{120\cdot}) + \ldots$$

$$\ldots + b^*_H(t_{H00\cdot} + t_{H22\cdot} - t_{H02\cdot} - t_{H20\cdot})]/12$$

$$\hat{a}^*_{13} = \frac{Y_{0\cdot 0} + Y_{2\cdot 2} - Y_{0\cdot 2} - Y_{2\cdot 0}}{12} - [b^*_1(t_{10\cdot 0} + t_{12\cdot 2} - t_{10\cdot 2} - t_{12\cdot 0}) + \ldots$$

$$\ldots + b^*_H(t_{H0\cdot 0} + t_{H2\cdot 2} - t_{H0\cdot 2} - t_{H2\cdot 0})]/12$$

$$\hat{a}^*_{23} = \frac{Y_{\cdot 00} + Y_{\cdot 22} - Y_{\cdot 02} - Y_{\cdot 20}}{12} - [b^*(t_{1\cdot 00} + t_{1\cdot 22} - t_{1\cdot 02} - t_{1\cdot 20}) + \ldots$$

$$\ldots + b^*_H(t_{H\cdot 00} + t_{H\cdot 22} - t_{H\cdot 02} - t_{H\cdot 20})]/12$$

Os estimadores dos coeficientes de regressão, para modelo reduzido, será:

$$\hat{\underset{\sim}{b}}_0 = (E^*_1)^{-1}\underset{\sim}{E}^*_0$$

onde: $E^*_1 = L - T'_{x_0} R_0^{-1} T_{x_0}$ e $\underset{\sim}{E}^*_0 = \underset{\sim}{B} - T'_{x_0} R_0^{-1} \underset{\sim}{T}_0$

ou

$$\underset{\sim}{E}^*_0 = \begin{bmatrix} \text{SPRes } t_1Y + \dfrac{(Y_{0..} - 2Y_{1..} + Y_{2..})(t_{10..} - 2t_{11..} + t_{12..})}{54} \\[2ex] \text{SPRes } t_2Y + \dfrac{(Y_{0..} - 2Y_{1..} + Y_{2..})(t_{20..} - 2t_{21..} + t_{22..})}{54} \\ \text{-------------------------} \\ \text{SPRes } t_HY + \dfrac{(Y_{0..} - 2Y_{1..} + Y_{2..})(t_{H0..} - 2t_{H1..} + t_{H2..})}{54} \end{bmatrix}$$

$$E^*_1 = \begin{bmatrix} \text{SQRes } t_1 + \dfrac{(t_{10..} - 2t_{11..} + t_{12..})^2}{54} & \text{SPRes } t_1t_2 + \dfrac{(t_{10..} - 2t_{11..} + t_{12..})(t_{10..} - 2t_{21..} + t_{22..})}{54} & \ldots & \text{SPRes } t_1t_H + \dfrac{(t_{10..} - 2t_{11..} + t_{12..})(t_{H0..} - 2t_{H1..} + t_{H2..})}{54} \\[2ex] & \text{SQRes } t_2 + \dfrac{(t_{20..} - 2t_{21..} + t_{22..})^2}{54} & \ldots & \text{SPRes } t_2t_H + \dfrac{(t_{20..} - 2t_{21..} + t_{22..})(t_{H0..} - 2t_{H1..} + t_{H2..})}{54} \\ \text{-----} & \text{-----} & \text{-----} & \text{-----} \\ \text{Simétrica} & & & \text{SQRes } t_H + \dfrac{(t_{H0..} - 2t_{H1..} + t_{H2..})^2}{54} \end{bmatrix}$$

sendo, E^*_1 a matriz $E_1 + A$, onde A é uma matriz constituída pelas somas de quadrados e de produtos das variáveis auxiliares adicionais do efeito de tratamento que se pretende ajustar (no caso, efeito de nitrogênio quadrático). E^*_0 é o vetor $\underset{\sim}{E}_0 + A^*$, onde A^* é um vetor constituído pelas somas de produtos da variável dependente com as variáveis auxiliares adicionais do efeito de tratamento que se pretende ajustar.

A soma de quadrados de parâmetros do modelo reduzido para o exemplo em questão (sem o efeito de nitrogênio quadrático) é dada por;

$$SQP(r) = \underset{\sim}{\hat{\theta}}'Z'\underset{\sim}{Y} = \underset{\sim}{\hat{T}}{}_0'T_0 + \underset{\sim}{\hat{b}}{}_0'\underset{\sim}{B} = \underset{\sim}{T}_0 R_0^{-1} \underset{\sim}{T}_0 + \underset{\sim}{\hat{b}}{}_0' \underset{\sim}{E}{}_0^*$$

onde: $\underset{\sim}{T}'_0 R^{-1}_0 \underset{\sim}{T}_0$ é a soma de quadrados de tratamentos do modelo reduzido, e, $\underset{\sim}{b}'_0 E^*_0$ a soma de quadrados de regressão linear na ausência do efeito a ser ajustado (no caso, nitrogênio quadrático).

Assim, a soma de quadrados de nitrogênio quadrático ajustada para regressão linear, é obtida pela diferença:

$$SQN''(aj) = SQP - SQP(r) = T'R^{-1}\underset{\sim}{T} + \underset{\sim}{\hat{b}}'\underset{\sim}{E}_0 - \underset{\sim}{T}{}_0'R_0^{-1}\underset{\sim}{T}_0 - \underset{\sim}{\hat{b}}{}_0'\underset{\sim}{E}{}_0^*$$

ou, simplificando

$$SQN''(aj) = SQN'' + \underset{\sim}{\hat{b}}'\underset{\sim}{E}_0 - \underset{\sim}{\hat{b}}{}_0'\underset{\sim}{E}{}_0^*$$

Para os demais efeitos, procede-se de maneira análoga.

Por definição, sabe-se que $D_T = E[\underset{\sim}{\hat{T}} - E(\underset{\sim}{\hat{T}})][\underset{\sim}{\hat{T}} - E(\underset{\sim}{\hat{T}})]'$ é a matriz de dispersão dos efeitos de tratamentos ajustados. Daí, obtem-se:

$D_\tau = (R^{-1} + R^{-1}T_x E_1^{-1} T'R^{-1})\sigma^2$, que é a matriz dos estimadores das variâncias e covariâncias dos efeitos de tratamentos ajustados pela regressão.

De modo análogo, tem-se que $D_b = E_1^{-1}\sigma^2$, que é a matriz dos estimadores dasvariâncias e covariâncias dos efeitos dos coeficientes de regressão.

TABELA 1. Esquema de análise da variância para os efeitos de tratamentos ajustados de acordo com a regressão linear.

Fontes de variação	GL	S.Q.
Parâmetros	p + H	$\underset{\sim}{T}'R^{-1}T + \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0$
Tratamentos (aj)	p - 1	$\underset{\sim}{T}'R^{-1}T + \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0 - \hat{a}_{00}\Sigma Y - \hat{\underset{\sim}{b}}'\underset{\sim}{B}$
Nit. Lin. (aj)	1	$SQN' + \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0 - \hat{b}'_0 - \hat{\underset{\sim}{b}}'_0\underset{\sim}{E}^*_0$
Fosf. Lin. (aj)	1	$SQP' + \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0 - \hat{\underset{\sim}{b}}_1\underset{\sim}{E}^*_1$
Pot. Lin. (aj)	1	$SQK' + \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0 - \hat{\underset{\sim}{b}}'_2\underset{\sim}{E}^*_2$
Nitrog. Quad. (aj)	1	$SQN'' + \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0 - \hat{\underset{\sim}{b}}'_3\underset{\sim}{E}^*_3$
Fosf. Quad. (aj)	1	$SQP'' + \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0 - \hat{\underset{\sim}{b}}'_4\underset{\sim}{E}^*_4$
Pot. Quad. (aj)	1	$SQK'' + \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0 - \hat{\underset{\sim}{b}}'_5\underset{\sim}{E}^*_5$
Int. N' x P' (aj)	1	$SQN' \times P' + \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0 - \hat{\underset{\sim}{b}}'_6\underset{\sim}{E}^*_6$
Int. N' x K' (aj)	1	$SQN' \times K' + \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0 - \hat{\underset{\sim}{b}}'_7\underset{\sim}{E}^*_7$
Int. P' x K' (aj)	1	$SQP' \times K' - \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0 - \hat{\underset{\sim}{b}}'_8\underset{\sim}{E}^*_8$
Resíduo (aj)	N-p-H	$Y'Y - \underset{\sim}{T}'R^{-1}\underset{\sim}{T} - \hat{\underset{\sim}{b}}'\underset{\sim}{E}_0$
Total	N	$Y'Y$

p = número de parâmetros de tratamentos (no caso p = 10).

Exemplo numérico

A título de ilustração, considerou-se um conjunto de dados fictícios de um experimento de adubação, em esquema fatorial 3^3, sem repetição. Os dados observados (Tabela 2) de produção de massa seca (y), número inicial de plantas (w_1) e pH do solo antes da aplicação dos tratamentos (w_2), foram submetidos à análise da variância (Tabela 3) de acordo com modelo proposto; foram, também, obtidas as estimativas dos parâmetros e respectivos desvios-padrão, (Tabela 4).

TABELA 2. Valores observados de produção de massa seca (y), número inicial de plantas (w_1) e pH do slo (w_2) em função dos níveis de tratamentos.

Tratamento	y	w_1	w_2	Tratamento	y	w_1	w_2	Tratamento	y	w_1	w_2	Soma y	Soma w_1	Soma w_2
000	40	55	7,1	100	53	56	5,9	200	68	54	5,7	161	165	18,7
001	140	63	6,9	101	349	64	4,4	201	342	65	4,8	831	192	16,1
002	230	65	5,0	102	216	55	6,2	202	269	58	7,3	775	178	18,5
010	41	62	7,0	110	92	54	5,6	210	77	55	6,4	210	171	19,0
011	166	63	6,1	111	326	63	5,2	211	390	64	5,2	880	190	16,5
012	150	66	5,0	112	412	57	6,8	212	300	56	7,2	862	179	19,0
020	50	57	6,9	120	62	56	8,0	220	71	54	6,8	183	167	21,7
021	153	59	6,2	121	353	64	4,9	221	326	61	6,2	832	184	17,3
022	299	60	4,8	122	255	58	8,2	222	282	58	8,0	836	176	21,0
Soma	1269	550	55,0		2178	527	55,2		2125	525	57,6	5572	1602	167,0

TABELA 3. Somas de quadrados e de produtos das variáveis produção de massa seca (y), número inicial de plantas (w_1) e pH do solo (w_2), e a análise da variância.

Fontes de variação	G.L.	Somas de quadrados, de produtos e valores do teste F: yy	F	w_1w_1	w_2w_2	yw_1	yw_2	w_1w_2	yy(aj.)	F(aj.)
Total	26	402818,2956		3243,6720	30,3141	5984,6667	-1128,6482	-62,13333		
Tratamento	(9)	344658,6111	11,19**	280,5278	18,9533	5268,2222	-626,3694	-45,6360	299755,0402	13,90**
A'	1	40707,5556	11,90**	34,7222	0,3756	-1188,8889	123,6444	-3,6111	50274,6324	20,98**
B'	1	392,0000	< 1,00	3,5556	2,4939	-37,3333	31,2667	-2,9778	6585,7030	2,75
C'	1	204586,7222	59,80**	50,0000	0,0200	3198,3333	-63,9667	-1,0000	195516,2771	81,57**
A''	1	17137,8519	5,01*	8,1667	0,0896	-374,1111	-39,1926	0,8556	13684,7826	5,71**
B''	1	1557,4074	< 1,00	6,0000	0,3424	96,6667	-23,0926	-1,4333	179,3916	< 1,00
C''	1	78814,2407	23,04**	170.6667	6,0669	3667,5556	-691,4870	-32,1778	19557,5907	8,16*
A' x B'	1	705,3333	< 1,00	0,7500	1,5408	-23,0000	-32,9667	1,0750	705,3502	< 1,00
A' x C'	1	630,7500	< 1,00	5,3333	8,0033	-58,0000	71,0500	-6,5333	13227,7100	5,52*
B' x C'	1	126,7500	< 1,00	1,3333	0,0208	-13,0000	-1,6250	0,1667	23,6026	< 1,00
Resíduo	17	58159,6845		2963,1442	11,3608	716,4445	-502,2788	-16,4973	35953,0958	
QM. resíduo		3421,1579							–	
Regressão	2	22206,5913							11103,2957	4,63*
QM. resíduo (aj.)	15	–							2396,8729	
CV (%)		28,34							23,72	

TABELA 4. Estimativas dos parâmetros para modelos sem e com as variáveis auxiliares adicionais e respectivos desvios-padrão.

Parâmetro	Estimativa ± desvio-padrão	Estimativa ajustada ± desvio-padrão
a_{00}	206,3704 ± 11,2565	206,3704 ± 9,4219
a_{10}	47,5557 ± 13,7864	53,9365 ± 11,7769
a_{20}	4,6667 ± 13,7864	21,1236 ± 12,7435
a_{30}	106,6111 ± 13,7864	105,1445 ± 11,6417
a_{11}	-53,4444 ± 23,8787	-48,0349 ± 20,1023
a_{22}	-16,2111 ± 23,8787	-5,5523 ± 20,2945
a_{33}	-114,6111 ± 23,8787	-70,1710 ± 25,0008
a_{12}	-7,6667 ± 16,8848	8,1792 ± 15,0769
a_{13}	7,2500 ± 16,8848	43,3584 ± 18,4572
a_{23}	3,2500 ± 16,8848	1,4061 ± 14,1503
b_1		-0,00439 ± 0,9030
b_2		-44,2180 ± 14,5841
r^2	0,8556	0,9107

REFERÊNCIAS

BOX, G.E.P. & DRAPER, N.R. A basis for the selection of a response surface design. **J. Am. Stat. Assoc.**, Washington, **54**:622-54, 1959.

BOX, G.E.P. & HUNTER, J.S. Multifactor experimental designs for exploring response surfaces. **Ann. Math. Stat.**, Hayward, **28**(1):195-241, 1957.

BOX, G.E.P. & WILSON, K.B. On the experimental attainment of optimum conditions. **J.R. Stat. Soc. B.**, Londres, **13**:1-45, 1951.

CAMPOS, H. **Aspectos da aplicação das superfícies de resposta a ensaios fatoriais 3³ de adubação.** Piracicaba, ESALQ/USP, 1967. 82p. Tese Livre-Docência.

COCHRAN, W.G. Analysis of covariance: its nature and uses. **Biometrics**, Raleigh, **13**:261-81, 1957.

COCHRAN, W.G. & COX, G.M. **Experimental designs.** 2. ed. New York, John Wiley, 1957. 611p.

COSTA, R.A. da. **Funções de produção ajustadas a ensaios fatoriais 3³ de adubação de arroz.** Piracicaba, ESALQ/USP, 1977. 80p. Tese Mestrado.

FINNEY, D.J. The fractional replication of factorial arrangements. **Ann. Eugenics**, **12**:291-301, 1945.

GOMEZ, J.H.; MENDEZ, I.R.; MORALES, A.C.; O'REILLY, F.T. Aplicaciones Agronómicas de la metodología de superficie de respuesta. **Agrociência**, Chapingo, **32**:125-35, 1978.

GRAYBILL, F.A. **An introduction to linear statistical models.** New York, McGraw-Hill, 1961. 463p.

MYERS, R.H. **Response surface methodology.** Boston, Allyn & Bacon, 1971. 246p.

VIEIRA, S. **Aspectos das funções de produção ajustadas aos ensaios fatoriais 3³ de adubação.** Piracicaba, ESALQ/USP, 1970. 160p. Tese Doutoramento.