Ministério da Agricultura e do Abastecimento

Embrapa

Pesquisa em andamento

Número 34 3p. 100 exemplares dez./1999 ISSN 1517-4921

UM ALGORITMO PARA ANÁLISE ESPACIAL DE EXPERIMENTOS DE CAMPO

Lúcio José Vivaldi[1]

A análise clássica de experimentos casualizados é entendida como análise de variância de experimentos onde os tratamentos são alocados aleatoriamente às parcelas, respeitando-se a estrutura de parcelas homogêneas fornecida pelas condições experimentais. Ocorre que, em experimentos de campo, podem existir tendências ao longo das parcelas, advindas das próprias características do solo: umidade, fertilidade, acidentes não detectáveis como por exemplo, cupim. Nesses casos, a análise de variância não é suficiente para retirar a heterogeneidade gerada por esses fatores espaciais, e como conseqüência , o erro experimental é maior.

Para resolver esse problema, surgiram os modelos de análise espacial de experimentos que levam em conta:

1) As condições do modelo clássico (casualização);

2) As variações espaciais.

Os modelos espaciais de análise são mais conhecidos como modelos de análise de vizinhanças, desde os trabalhos de Bartlett (1978) e de Wilkinson at al. (1984). Entre os vários modelos de análise, o primeiro a ser aplicado é o de Besag & Kempton (1986), denominado de modelo das primeiras diferenças e apropriado para conjuntos de parcelas que estão em disposição longitudinal. Considerando um delineamento em blocos ao acaso, com os blocos em forma longitudinal e v tratamentos, tem-se que

$$\mathbf{Y} = \mathbf{1}\mu + \mathbf{Bb} + \mathbf{T}\tau + \mathbf{e} \qquad (1)$$

onde $\mathbf{Y}$ é o vetor das observações, μ é a média geral, **1** é o vetor de um ´s, **B** é a matriz de incidência de blocos, **b** é o vetor dos efeitos de blocos, **T** é a matriz de incidência de tratamentos, τ é o vetor dos efeitos de tratamentos e **e** é o vetor dos erros que têm média zero, variância σ_2^2 e são independentes. No modelo proposto por Besag & Kempton (1986), existem variáveis $x_1, x_2, \ldots, x_v$ específicas para as parcelas 1, 2,..., v de cada bloco, com as seguintes propriedades:

(a) $x_i - x_{i+1}$, são variáveis não correlacionadas, para todo i;

(b) $E\,(x_i - x_{i+1}) = 0$

(c) $Var(x_i - x_{i+1}) = \sigma_1^2$

Essas variáveis representam a fertilidade do solo em cada parcela. Com elas, o modelo (1) torna-se:

$$\mathbf{Y} = \mathbf{1}\mu + \mathbf{Bb} + \mathbf{T}\tau + \mathbf{e} + \mathbf{x}, \qquad (2)$$

onde $\mathbf{x}$ é um vetor que comporta as variáveis x ´s, na ordem das parcelas.

[1] Pesquisador da Embrapa Cerrados.

Uma forma de analisar o modelo (2), levando em conta as propriedades (a), (b) e (c), é considerar um processo de primeiras diferenças, $y_i - y_{i+1}$ aplicado a (2), ou seja, a diferença entre duas parcelas vizinhas. Assim fazendo, obtém-se o novo modelo:

$$\mathbf{Z} = \Delta\mathbf{Y} = \Delta\mathbf{T}\tau + \Delta\mathbf{e} + \Delta\mathbf{x}, \qquad (3)$$

sendo Δ a matriz das primeiras diferenças. Agora o modelo tem as seguintes propriedades:

$$\text{(a)}\, E(\mathbf{Z}) = \Delta\mathbf{T}\tau = \mathbf{F}\tau \qquad (4)$$

$$\text{(b)}\, Var(\mathbf{Z}) = \mathbf{I}\sigma_1^2 + \Delta\Delta'\sigma_2^2 = \mathbf{V} \qquad (5)$$

Desta forma,

$$Var(y_i - y_j) = \sigma_1^2 \mid i - j \mid + 2\sigma_2^2 \qquad (6)$$

Isto é, a variância da diferença entre duas parcelas depende também da distância entre elas. A expressão (5) pode ser escrita da forma:

$$\frac{1}{\sigma_1^2}\mathbf{V} = \mathbf{I} + \alpha\Delta\Delta' \qquad (7)$$

onde $\alpha = \frac{\sigma_2^2}{\sigma_1^2}$. O novo parâmetro, α (alfa), é a razão entre o erro experimental e o erro devido à tendência entre as parcelas dentro do bloco; se o erro experimental é desprezível e σ_1^2 é grande, alfa tende para zero, e no caso inverso, alfa tende para o infinito. Espera-se que alfa não atinja os extremos e sim um valor que congregue as duas fontes de variação. A análise de um experimento com o modelo (3) proporciona menor erro padrão da diferença entre dois tratamentos.

Admitindo-se que os dados seguem a distribuição normal, aplica-se método da máxima verossimilhança para se estimar alfa. Um algoritmo que estima os parâmetros envolvidos em (3) possui os seguintes passos:

1) Estimar o vetor $\hat{\tau}_\alpha$ de efeitos de tratamentos, usando quadrados mínimos e um valor inicial para alfa (que pode ser zero);
2) Obter $\hat{\mathbf{e}} = \mathbf{Z} - \mathbf{F}\hat{\tau}_\alpha$ (erros estimados);
3) Obter o valor de alfa que minimiza $L = \log(\det(\mathbf{V})) + (n-1)\log(\hat{\mathbf{e}}'\mathbf{V}^{-1}\hat{\mathbf{e}})$ (alfa faz parte de $\mathbf{V}$);
4) Retorne ao passo (1) e recalcule $\hat{\tau}_\alpha$ O processo termina segundo uma regra de parada, quando $\hat{\tau}_\alpha$ não mais se modifica.

Com os resultados, calcula-se a matriz de variâncias e covariâncias de $\hat{\tau}_\alpha$, dada por:

$$Var\,(\hat{\tau}_\alpha) = (\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\sigma_1^2 \qquad (8)$$

$$\mathbf{V} = \mathbf{I} + \hat{\alpha}\Delta\Delta', \hat{\alpha} = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2}, \; \hat{\sigma}_1^2 = \frac{\hat{\mathbf{e}}'\mathbf{V}^{-1}\hat{\mathbf{e}}}{n - v - 2} \qquad (9)$$

Esse algoritimo foi implementado numa primeira versão particularizada (para cada experimento) em linguagem SAS (1997). Pela sua aplicação em experimentos da Embrapa Cerrados, observou-se que o processo converge lentamente em alguns casos e não converge em outros, o que implica a busca de melhor algoritmo. Para ilustrar seu uso e aplicação apresentam-se os resultados da análise de um experimento com soja (BSB- FTCe, subprojeto 04.0.94.321-25, ano 1997/1998), com 27 variedades por bloco:

	Análise de variância	Análise espacial
Erro exp.	464768	309696
Erro de tendência	-----	65892
Média das variâncias	232388	207227
Eficiência	-----	10%

A eficiência da análise espacial foi calculada, tendo como base a média das variâncias de todas as possíveis comparações entre dois tratamentos, nas duas análises.

REFERÊNCIAS BIBLIOGRÁFICAS

BARTLETT, M.S. Nearest-neighbour models in the analysis of field experiments. **Journal of the Royal Statistical Society, Serie B,** London, v.40, p.147-174, 1978.

BESAG, J.; KEMPTON, R. Statistical analysis of experiments using neighbouring plots. **Biometrics,** Washington, v.42, p.231-251, 1986.

SAS INSTITUTE INC. (Cary, NC). **Release 6.12.** Cary,1997.

WILKINSON, G.N.; ECKERT, S.R.; HANCOCK, T.W.; MAYO, O. Nearest-neighbour (NN) analysis of field experiments. **Journal of the Royal Statistical Society, Serie B,** London, v.45, p.151-211, 1983.

Embrapa

Empresa Brasileira de Pesquisa Agropecuária
Embrapa Cerrados
Ministério da Agricultura e do Abastecimento
BR 020, km 18, Rodovia Brasília/Fortaleza, Caixa Postal 08223
CEP 73301-970, Planaltina, DF
Telefone: (61) 388-9898 FAX: (61) 388-9879