

Guerra na Geoestatística: uma disputa pela menor variância entre os interpoladores lineares

Lidiane Maria Ferraz Rosa ¹

Gérson Rodrigues dos Santos ²

Adriana Maria Rocha Trancoso Santos ³

Nilcilene das Graças Medeiros ^{4 5}

1 Introdução

Na arte de utilizar alguns dados observados para conhecer os não observados, sabidamente os interpoladores lineares da Geoestatística apresentam características ótimas quanto à eficácia. Em outras palavras, é comum encontrar a seguinte expressão: “a krigagem prediz sem viés e com variância mínima”. Diversos trabalhos já foram publicados na literatura especializada com o intuito de provar ou reprovar tal afirmação, sendo que, às vezes, dificuldades foram encontradas na comparação utilizando interpoladores que não apresentaram os mesmos mecanismos de mensuração das incertezas.

Santos (2010) apresenta uma interessante visão geométrica dos preditores lineares da Geoestatística. Mas, desde Journel (1977) que o interesse pelos melhores preditores geoestatísticos tem sido notado. Santos, Oliveira, Louzada & Santos (2011) mostraram que entre a krigagem linear simples e a krigagem linear universal a krigagem simples é mais precisa quanto à variância de predição.

Neste trabalho, o que está em pauta é a utilização de simulações estocásticas computacionais, utilizando o Software R, para uma disputa pela menor variância de predição. Nessa guerra declarada entre os interpoladores lineares da Geoestatística, verificou-se que a krigagem linear simples apresenta essa superioridade, mesmo que o tamanho do grid amostral esteja reduzido, seguida da krigagem universal.

¹DET - UFV. E-mail: lidiane.rosa@ufv.br

²DET - UFV. E-mail: geron.santos@ufv.br

³DEC - UFV. E-mail: adriana.trancoso@ufv.br

⁴DEC - UFV. E-mail: nilcilene.medeiros@ufv.br

⁵Os autores agradecem à Capes, Fapemig e CNPq pelo apoio financeiro.

2 Material e Métodos

Devido a dificuldade de encontrar conjuntos de dados com as características apropriadas com o objetivo deste trabalho, adotou-se a simulação estocástica computacional de 8 conjuntos de dados georeferenciados com 144, 121, 100, 81, 64, 49, 25 e 16 pontos amostrais, todos com tendência nos dados em relação às coordenadas, visando também a robustez do interpolador.

Para cada tamanho de amostra, simulações foram feitas até que um comportamento de estrutura de dependência espacial fosse percebido. Por exemplo, não era suficiente simular um processo gaussiano com parâmetros claramente estimáveis no primeiro variograma, a estrutura de dependência espacial deveria ainda existir após a remoção da tendência dos dados.

A krigagem linear simples é dada por

$$\hat{Y}_{KLS}(\mathbf{x}_0) = \lambda_0 + \sum_{i=1}^n \lambda_i Y(\mathbf{x}_i). \quad (1)$$

A krigagem linear universal é dada por

$$\hat{Y}_{KLU}(\mathbf{x}_0) = \sum_{j=1}^n \lambda_j Y(\mathbf{x}_j), \quad (2)$$

com $Y(\mathbf{x}_j) = \left[\sum_{l=1}^L a_l f_l(\mathbf{x}_j) \right] + \varepsilon'(\mathbf{x}_j) + \varepsilon''$.

A variância de krigagem para ambos os interpoladores é dada por

$$\sigma^2(\mathbf{x}_0) = \|Y(\mathbf{x}_0) - \hat{Y}_{Krig}(\mathbf{x}_0)\|^2. \quad (3)$$

Em seguida, toda a análise geoestatística foi realizada, utilizando o software livre R (R Core Team, 2013).

3 Resultados e Discussão

Diante das simulações estocásticas, onde processos gaussianos e processos desconhecidos, de médias conhecidas, acrescidos de tendência em relação às coordenadas, cujo modelo é dado por

$$m(\mathbf{x}_i) = f(x, y) = 100 - 2x + 0,3y$$

conforme Tabela 1, pode-se chegar aos resultados esperados pelo trabalho.

Tabela 1: Resultados das simulações que possibilitam a comparação entre os principais preditores lineares geoestatísticos.

	Dist. F(.) conhecida				Dist. F(.) desconhecida	
	KLS		KLU		KLS	KLU
<i>grid</i>	dim	$\sigma_{KLS}^2(\mathbf{x}_0)$	dim	$\sigma_{KLU}^2(\mathbf{x}_0)$	$\sigma_{KLS}^2(\mathbf{x}_0)$	$\sigma_{KLU}^2(\mathbf{x}_0)$
144	145	0,1427	141	0,3369	0,1955	0,3606
121	122	0,1047	118	0,3229	0,4639	0,6723
100	101	0,1836	97	0,3945	0,1841	0,4086
81	82	0,0941	78	0,3610	0,7670	2,2942
64	65	0,2077	61	0,4496	0,2143	1,9756
49	50	0,2708	46	0,5944	0,4076	1,3983
25	26	0,3566	22	0,6137	0,2343	1,4474
16	17	0,4586	13	0,8022	0,4398	0,7452

Analisando os resultados apresentados na Tabela 1, pode-se verificar que as diferenças entre as variâncias de krigagens foram razoavelmente pequenas, o que pode ser explicado pela parcimônia do modelo de tendência, pois, quanto maior o número de parâmetros do modelo de tendência maior será a variância de krigagem.

Outrossim, como as simulações, em cada um dos grandes grupos, foram repetidas para tamanhos diferentes de amostras (144, 121, 100, 81, 64, 49, 25 e 16 pontos amostrais), percebe-se um crescimento acentuado na variância de krigagem ao reduzir do tamanho amostral, o que ratifica a influência do número de informações em uma predição (ou estimação) estatística.

4 Conclusões

Nessa guerra declarada entre os interpoladores lineares da Geoestatística, verificou-se que a krigagem linear simples apresenta essa superioridade, mesmo que o tamanho do grid amostral esteja reduzido, seguida da krigagem universal, tomando como parâmetro de decisão a variância de krigagem da autovalidação.

Referências

- [1] Journel, A. G. (1977). Kriging in Terms of Projections. **Mathematical Geology**, New York, vol. 9, n. 6. p. 563-586.
- [2] R Core Team (2013). R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. Disponível em: <http://www.R-project.org/>. Acesso em: 8 de dezembro de 2013.
- [3] Santos, G. R. **Hierarquização geométrica dos preditores geoestatísticos**. Tese (Doutorado em Estatística e Experimentação Agropecuária) Setor de Ciências Exatas, Universidade Federal de Lavras, Lavras. 2010. 200 p.
- [4] Santos, G. R., Oliveira, M. S., Louzada, J. M., Santos, A. M. R. T. Krigagem Simples versus Krigagem Universal: qual o preditor mais preciso? **Revista Energia na Agricultura**, Botucatu, vol. 26, n. 2, 2011, p. 49-55.