RELAÇÃO ENTRE FUNÇÃO DE COVARIÂNCIA E A FUNÇÃO DE SEMIVARIÂNCIA ATRAVÉS DE TÉCNICAS ESPECTRAIS

Edilenia Queiroz Pereira (PIBIC/FA/UEM), Diogo Francisco Rossoni, diogo.rossoni@gmail.com.

Universidade Estadual de Maringá / Centro de Ciências Exatas e da Terra /Maringá, PR.

Ciências Exatas e da Terra / Probabilidade e Estatística.

Palavras-chave: Densidade espectral, Estatística Espacial, Semivariância.

Resumo:

Pesquisadores da área espacial se deparam muitas vezes com grandes conjuntos de dados, normalmente coletados em uma região de grande porte. Neste contexto, a densidade espectral pode ser uma alternativa viável para o estudo de dados espacialmente correlacionados. Uma vez que, trabalhando no domínio espectral pode-se obter a semivariância a partir da função de covariância.

Introdução

A manipulação de grandes conjuntos de dados na área espacial é, de certo modo, problemática para as técnicas Geoestatísticas, seja por ter que lidar com a inversão de uma matriz de covariância excessivamente grande para se obter a função de probabilidade ou que exija um grande nível de processamento computacional (MATEU; JUAN; PORCU, 2007).
Neste contexto, a abordagem espectral pode ser uma alternativa viável para o estudo de dados espacialmente correlacionados. Uma abordagem inicial a métodos espectrais é apresentada por Cressie (1993). A eficiência desses métodos espectrais depende da geometria do domínio espacial, do método de estimação, do número de observações, e muitas outras variáveis que têm de ser levadas em consideração para se realizar uma análise correta (MATEU; JUAN; PORCU, 2007).
Este trabalho tem o objetivo apresentar a relação existente entre a função de covariância e a função de semivariâncias através de técnicas espectrais.

Materiais e métodos

No enfoque Geoestatístico, a semivariância $\{\gamma(\mathbf{h})\}$ é uma medida usualmente utilizada, a qual é definida como uma forma de medir o nível de dependência espacial entre duas variáveis. Considere dois pontos separados por uma distância $\mathbf{h}$, tem-se que a variação entre esses dois pontos é caracterizada pela função de semivariograma que é definida como:

$$\gamma(h) = \frac{1}{2} E\left\{\left[Z(x+h) - Z(x)\right]^2\right\}. \tag{1}$$

Tem-se que quando a estacionariedade de segunda ordem dos dados é estabelecida, é possível estabelecer uma relação direta entre a covariância e a semivariância. Sendo, essa relação dada por:

$$\gamma(h) = \mathrm{cov}(0) - \mathrm{cov}(h). \tag{2}$$

A covariância pode ser obtida através das funções de densidade espectrais. O periodograma é uma forma alternativa de estimar a função de densidade espectral, sendo de fundamental importância por permitir a identificação de periodicidade existente nos dados. Desta forma para um processo espacial $Z(s)$ observado em um malha regular de locais igualmente espaçados, $D = \{\mathbf{s} = (s_1, s_2) : s_1 = 0, \ldots, n_1 - 1, s_2 = 0, \ldots, n_2 - 1\}, D \subset R^2$, com $N = n_1 n_2$ pontos. Tem-se o periodograma espacial dado por:

$$I_N(\omega_0) = \frac{\delta_1 \delta_2}{(2\pi)^2 n_1 n_2} \left| \sum_{\mathbf{s} \in D} Z(\Delta \mathbf{s}) \exp(-i \Delta \mathbf{s}^t \omega) \right|^2 \tag{3}$$

em que $\Delta \mathbf{s} = \delta_1 s_1 \delta_2 s_2$ é o vetor de distância entre as observação próximas; ω é a frequência; $n_1 n_2$ é o número de pontos amostrados; δ_1 é a localização em uma direção (direção do eixo x); δ_2 é a localização em outra direção (direção do eixo y);

Função de densidade exponencial

A função de densidade espectral de um processo espacial isotrópico com uma covariância exponencial é definida como (GELFAND et al., 2010):

$$f(\omega) = \frac{1}{2} \sigma (\pi \alpha)^{-\frac{1}{2}} \exp\left(-\frac{\omega^2}{4\alpha}\right) \tag{4}$$

sendo a covariância definida como:

$$C(h) = \sigma \exp\left\{-\alpha h^2\right\}. \tag{5}$$

Os dados experimentais foram obtidos por meio de simulação no software estatístico R (R CORE TEAM, 2015). Para isso, utilizou-se o pacote *RandomFields*, o qual permite a análise e simulação de processos aleatórios espaciais. Assim, utilizou-se a função *RFsimulate* segundo o modelo de covariância exponencial com variância dois para a geração de 160.000 observações em um grid regular.

Resultados e Discussão

Por meio do software estatístico R e do pacote *fractaldim,* utilizando os dados simulados, foi estimado o periodograma para 150 distâncias h. A distribuição empírica do periodograma espacial é apresentada na Figura 1.
Para ajustar o modelo exponencial teórico – equação (4) – ao periodograma espacial, foi utilizada a função *nls* do software R. A Tabela 1 apresenta as estimativas dos parâmetros do modelo exponencial e a Soma Quadrática do Resíduo (SQR).

Tabela 1: Estimativas dos parâmetros da função de densidade exponencial.

σ	α	SQR
52.78	0.51	227.1

Com essas estimativas tem-se que a função de densidade espectral estimada de um processo espacial isotrópico com uma covariância exponencial é:

$$\hat{f}(\omega) = \frac{1}{2} 0.51(\pi 52.78)^{-\frac{1}{2}} \exp\left(-\frac{\omega^2}{4(52.78)}\right) \tag{6}$$

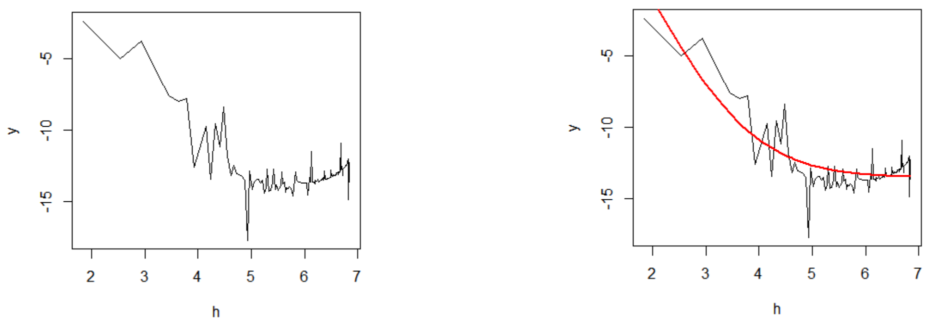

Figura 1: Distribuição empírica dos dados e modelo exponencial ajustado.

Sendo a covariância estimada definida como:

$$\hat{C}(h) = 52.78 \exp\left\{-0.51 h^2\right\}. \tag{7}$$

Conhecendo a covariância pode-se agora encontrar o semivariograma estimado pela equação (2). A Figura 2 apresenta a covariância estimada pela equação (7) e sua respectiva semivariância.

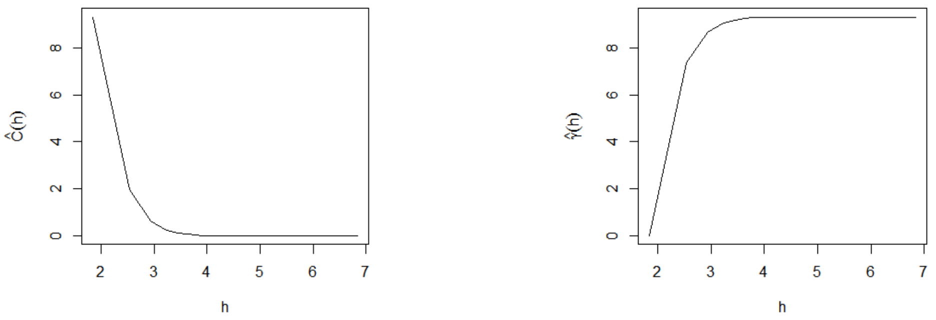

Figura 2: Covariância e semivariograma estimado.

Conclusões

Através do periodograma espacial foi possível ajustar uma função de densidade espectral. A partir da função de densidade espectral foi possível estimar a covariância e a semivariância.

Agradecimentos

A Fundação Araucária pelo apoio financeiro.

Referências

CRESSIE, N. **Statistics for spatial data**. 2. ed. New York: John Wiley, 1993.

GELFAND, A. et al. **Handbook os Spatial Statistics**. London: CRC Press, 2010.

MATEU, J; JUAN, P; PORCU, E. Geostatistical Analysis Through Spectral Techniques: Some Words of Caution. **Communications in Statistics - Simulation and Computation**, v. 36, n. 5, p. 1035–1051, 2007.

R CORE TEAM. **R: A language and environment for statistical computing.** Vienna, Austria R Foundation for Statistical Computing, 2015. Disponível em: <http://www.r-project.org/.>. Acesso em: 10 jun. 2015.