

Estimadores de *Kernel*

Wagner Hugo Bonat
Henrique Dallazuana

5 de outubro de 2007

1 Metodologia

1.1 Estimador de Intensidade ("Kernel Estimation")

Uma das alternativas simples para gerar uma superfície bidimensional, com base em valores pontuais, é ajustar uma função bidimensional sobre as amostras consideradas, compondo uma superfície cujo valor será proporcional à intensidade dos valores das amostras locais. A formulação geral para este tipo de interpolação é:

$$\hat{z}_i = \frac{\sum_{j=1}^n w_{ij} z_j}{\sum_{j=1}^n w_{ij}} \quad (1)$$

onde z_i é o valor da cota de um ponto i qualquer da grade; z_j é a cota de uma amostra j vizinha do ponto i da grade; e, w_{ij} é um fator de ponderação. Existem várias opções para estes tipos de interpoladores, as opções mais comuns são:

- Vizinho mais próximo - É definida pela escolha de apenas uma amostra vizinha para cada ponto da grade, este interpolador deve ser usado quando se deseja manter os valores de cotas das amostras na grade, sem gerar valores intermediários.
- Média Simples - Este considera o valor da cota z do elemento da grade igual a média aritmética dos valores das cotas vizinhas, nesse caso considera-se o fator de ponderação $w_{ij} = \frac{1}{n}$ para qualquer amostra considerada.
- Média Ponderada - Neste caso o valor da cota de cada elemento da grade é definido pela média ponderada dos valores de cota das amostras

vizinhas, a ponderação mais usada na prática é o inverso da distância euclidiana do ponto da grade à amostra considerada ou seja:

$$w_{ij} = \frac{1}{d_{ij}^k} \quad (2)$$

onde, k é o expoente da distância, geralmente, igual a 1 ou 2 e; d_{ij} é o valor de distância da amostra j ao ponto i da grade expresso por:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3)$$

De uma forma geral as superfícies geradas por estas funções simples apresentam defeitos típicos, no caso de vizinho mais próximo e média simples tendem a produzir superfícies com variações abruptas e com embricação respectivamente, já no caso do inverso da medida do quadrado da distância, os máximos locais tendem a ser muito acentuados, formando "picos" artificiais.

Dado estes problemas faz-se necessário utilizar um refinamento destes interpoladores, um refinamento destes é a classe de estimadores descrita na literatura como *kernel estimators* ou *estimadores de densidade não paramétricos*. Esta classe de estimadores supõe que a densidade do fenômeno varia localmente de forma suave, sem "picos" nem "descontinuidades", e o objetivo é produzir uma superfície mais suave e mais representativa do fenômeno em estudo.

O *kernel estimator* possui dois parâmetros básicos são:

1. um raio de influência que define a vizinhança do ponto a ser interpolado;
2. uma função de estimação com propriedades "convenientes" de suavização do fenômeno.

De forma geral, em toda posição z_j cujo valor queremos estimar, o estimador de intensidade será calculado com base nos valores das amostras z_1, z_2, \dots, z_n contidos em um raio de tamanho τ , e da distância euclidiana d_{ij} entre a i -ésima posição e a j -ésima amostra a partir de uma função do tipo:

$$\hat{z}_i = \frac{\sum_{j=1}^n k\left(\frac{d_{ij}}{\tau}\right) z_j}{\sum_{j=1}^n k\left(\frac{d_{ij}}{\tau}\right)} \quad , \quad d_{ij} \leq \tau \quad (4)$$

Essa fórmula é uma generalização da equação (1) na qual os pesos w_{ij} serão substituídos por uma função generalizada dependente da distância, uma das possíveis funções e a que será utilizada neste trabalho é conhecida como *kernel* gaussiano:

$$k(x, y, \tau) = \frac{1}{2\pi\tau} \exp - \left[\frac{d_{ij}^2}{2\tau^2} \right] \quad (5)$$

O *kernel estimator* é uma alternativa viável a métodos mais sofisticados de interpolação, pois não requerem a parametrização da estrutura de correlação espacial (como do caso da geoestatística). As superfícies interpoladas são suaves e aproximam o fenômeno em estudo. As desvantagens desses estimadores são a forte dependência do raio de busca e a excessiva suavização da superfície que pode, em alguns casos, esconder variações locais importantes.