

Dados espaciais

Patrícia de Siqueira Ramos

UNIFAL-MG, *campus* Varginha

18 de Outubro de 2017

O que são dados espaciais

- Dados espaciais: quanto varia algum fenômeno preocupando-se em saber onde ocorre tal variação
- Esse tipo de dados possui duas propriedades:
 - magnitude da variação do atributo do fenômeno em estudo (valor da variável)
 - localização geográfica do atributo (como os dados estão localizados)

Características dos dados espaciais

- georreferenciamento - posição relativa ou absoluta dos dados no mapa importa
- multidirecionalidade - dados observados em regiões interagem em todas as direções

Surgimento dos dados espaciais

- Escaneamento de mapas em papel
- Fotos aéreas digitalizadas
- Mais recentemente, sensoriamento remoto e desenvolvimento do GPS (*Global Positioning System*)
- Evolução dos SIG (Sistemas de Informações Geográficas) - programas que permitem a estocagem, organização, visualização e análise de dados espaciais
- Com o aumento da disponibilidade de dados espaciais, houve a necessidade do desenvolvimento de técnicas de análise desses dados
- Usando um programa SIG é possível associar informações de atributo e de localização, surgindo os dados espaciais

Fontes de dados espaciais

- A maioria das fontes de dados espaciais vêm de agências estatais públicas (no Brasil pelo IBGE - Instituto Brasileiro de Geografia e Estatística)
- Grande parte dos dados pode ser obtida no site do IPEA (Instituto de Pesquisa Econômica Aplicada) que compila dados de várias fontes ou no próprio IBGE (<ftp://geoftp.ibge.gov.br/>)

Fontes de dados espaciais

- Algumas fontes de dados são amostrais e outras são censitárias:
 - PNAD (Pesquisa Nacional por Amostra de Domicílios), Pesquisa Agrícola Municipal, Pesquisa Pecuária Municipal etc.
 - Censo, Censo Agropecuário, registros administrativos do mercado de trabalho formal e dos fluxos migratórios (pelo Ministério do Trabalho e Emprego) - relatório anual de informações sociais (Rais)

Como são os dados espaciais

- Numa planilha deve haver a codificação das unidades espaciais
- Cada unidade deve possuir uma identificação própria (código com valores inteiros consecutivos e não repetidos)
- Programas SIG fazem a junção da tabela com os dados sobre “quanto variou” com a localização geográfica de “onde variou” pela verificação do código da região

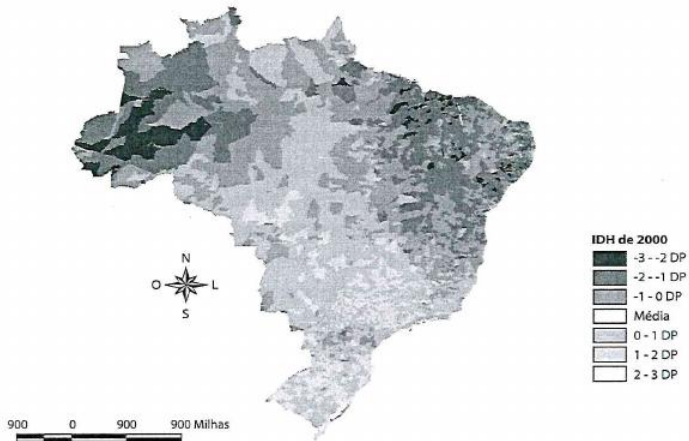
Como são os dados espaciais

- Na internet podemos ter acesso a mapas digitais nos vários formatos dos programas SIG existentes
- Há extensões .shp, .mif
- Um arquivo *shapefile* na verdade não é apenas um arquivo, ele possui 3 componentes principais:
 - .shp contém a geometria
 - .dbf contém os atributos como uma tabela relacional
 - .prj contém a projeção do mapa em texto

Processo estocástico espacial

- Modelos determinísticos \times estocásticos
- Os dados espaciais representam uma única realização de um processo estocástico do tipo espacial
- Por exemplo, um mapa com o IDH municipal para o Brasil é uma realização, dentre inúmeras possíveis, de um “processo gerador de mapas”
- Ou seja, um mapa com o IDH municipal para o Brasil é a única realização que está disponível de uma população de mapas com o IDH municipal para o Brasil que poderia ter sido realizada

Ex.: dados pontuais



IDH dos municípios brasileiros em 2000

Processo estocástico espacial

- Se a região de interesse é denotada por R e o valor de $Y(s)$ em cada local s ($s \in R$) é uma v.a., a coleção de v.a.s:

$$\{y(s) : s \in R\}$$

é um processo estocástico em R (também chamado de campo aleatório em R)

Processo estocástico espacial

- Se a região de interesse é denotada por R e o valor de $Y(s)$ em cada local s ($s \in R$) é uma v.a., a coleção de v.a.s:

$$\{y(s) : s \in R\}$$

é um processo estocástico em R (também chamado de campo aleatório em R)

- Há um grande paralelo entre processos estocásticos espaciais e temporais:

$$\{y(t) : t \in T\},$$

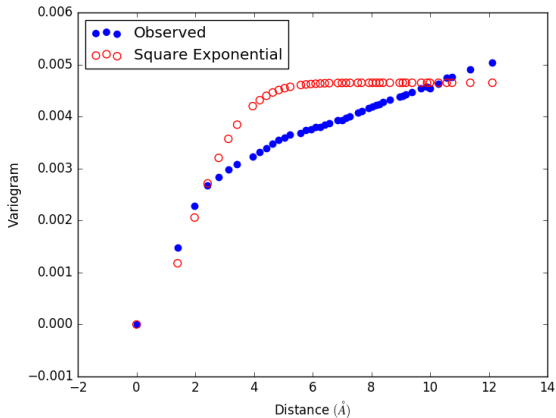
em que o conjunto T é um intervalo de tempo contínuo (e provavelmente não limitado)

- Diferença: o tempo é naturalmente ordenado (passado-presente-futuro) e o espaço não

Processo estocástico gerador de dados

- Sequência de variáveis aleatórias ordenadas segundo o critério geográfico
- Os dados espaciais constituem uma amostra de pontos de uma distribuição espacial contínua na forma de uma superfície e que pode ser modelada usando o variograma
- Variograma: recurso geoestatístico que revela a força da associação entre os pares de localidades como uma função contínua da distância que os separa

Variograma



Processo estocástico gerador de dados

- Com a Geoestatística é possível inferir informação em relação a dados não observados sobre o plano espacial de observações amostradas utilizando krigagem
- Geoestatística: ramo da Estatística que procura modelar a informação contínua de um fenômeno através do espaço.

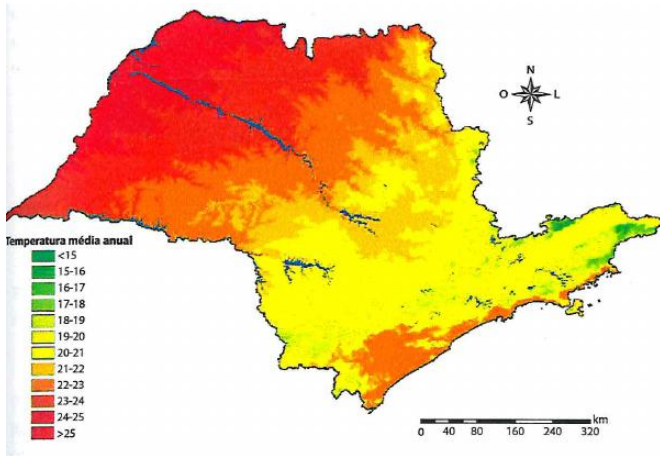
Processo estocástico gerador de dados

- Com a Geoestatística é possível inferir informação em relação a dados não observados sobre o plano espacial de observações amostradas utilizando krigagem
- Geoestatística: ramo da Estatística que procura modelar a informação contínua de um fenômeno através do espaço.
- krigagem: técnica que usa combinação linear de pesos em lugares conhecidos para prever um valor desconhecido, gerando uma superfície de previsão contínua

Processo estocástico gerador de dados

- A variação contínua ocorrendo em qualquer ponto do mapa é mais adequada para se estudar fenômenos físicos (pressão atmosférica, elevação do terreno, temperatura)
- No exemplo a seguir, as variações da temperatura média, por serem contínuas, são suaves ao longo do mapa
- Assim, em qualquer ponto do mapa é possível obter um valor de temperatura (se um instrumento de medida de alta precisão estiver disponível)

Ex.: dados contínuos

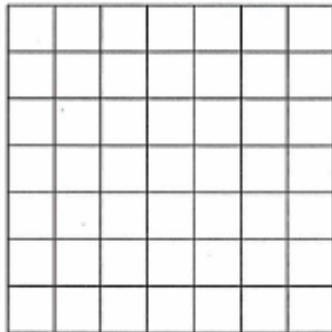


Temperatura média anual do ar no estado de SP

Dados descontínuos

- Em Ciências Sociais, geralmente, dados espaciais com variação contínua costumam não ter muito interesse
- Os dados espaciais descontínuos são agrupados em polígonos e vistos como uma única realização de um processo estocástico espacial
- Os polígonos podem ser regulares ou irregulares

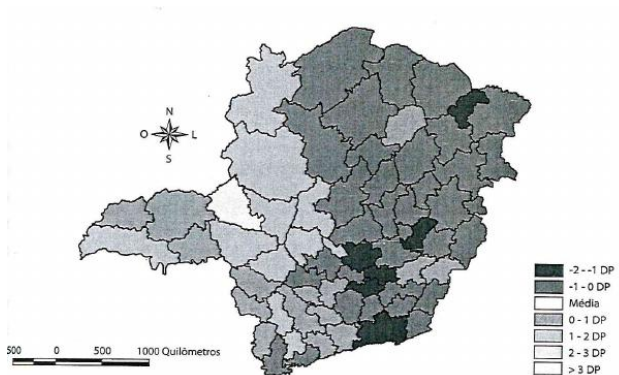
Ex.: Gride 7×7 - polígono regular



Representação de dados descontínuos

- A maioria dos dados descontínuos são representados em polígonos irregulares
- O tamanho e a forma das fronteiras são diferentes ao longo do espaço

Ex.: polígono irregular



Valor da produção agropecuária per capita em Minas Gerais (1996)
- 66 microrregiões

Representação em polígonos irregulares

- No exemplo, os valores da produção agropecuária per capita só variam quando passamos de uma microrregião para outra, cruzando as fronteiras
- A variação do fenômeno representado pelos dados é constante dentro de uma região
- A variação só ocorre quando se ultrapassa a fronteira, passando de uma microrregião para outra

Dados pontuais

- Há também a possibilidade de representar dados pontuais
- Os pontos são localidades de eventos discretos
- Dados pontuais são comuns em epidemiologia (ex.: casos de dengue num município) e criminologia (ex.: ocorrência de crimes nos bairros em uma cidade)

Ex.: dados pontuais



Amostra de cidades com TV a cabo (sem data)