Autocorrelação espacial

Oswaldo Santos Baquero

VPS-FMVZ-USP

2017

Sumário

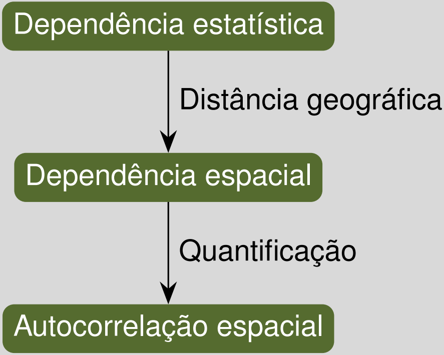
Sumário

Dependência estatística

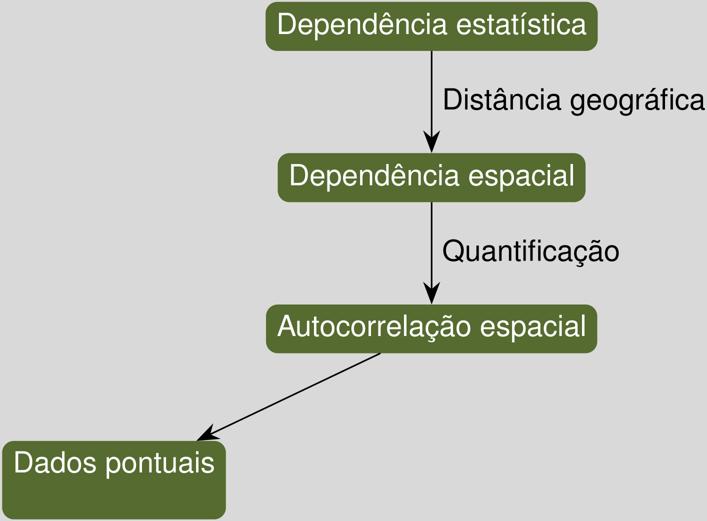
Sumário

$$\text{Dependência estatística} \xrightarrow{\text{Distância geográfica}} \text{Dependência espacial}$$

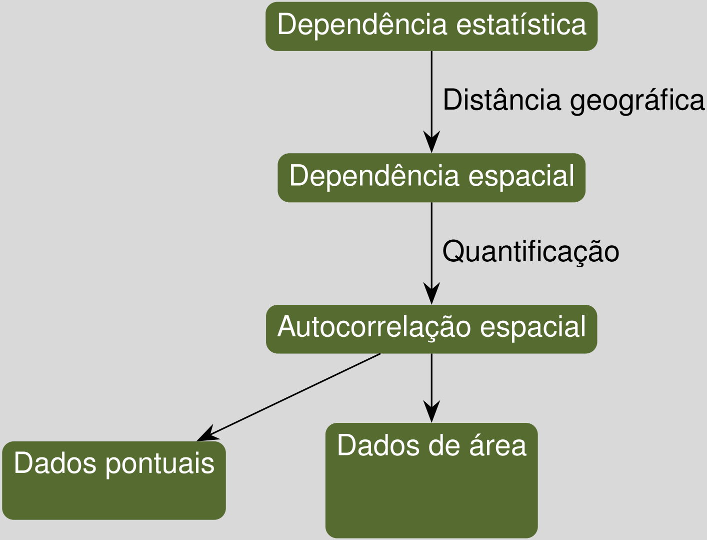
Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial

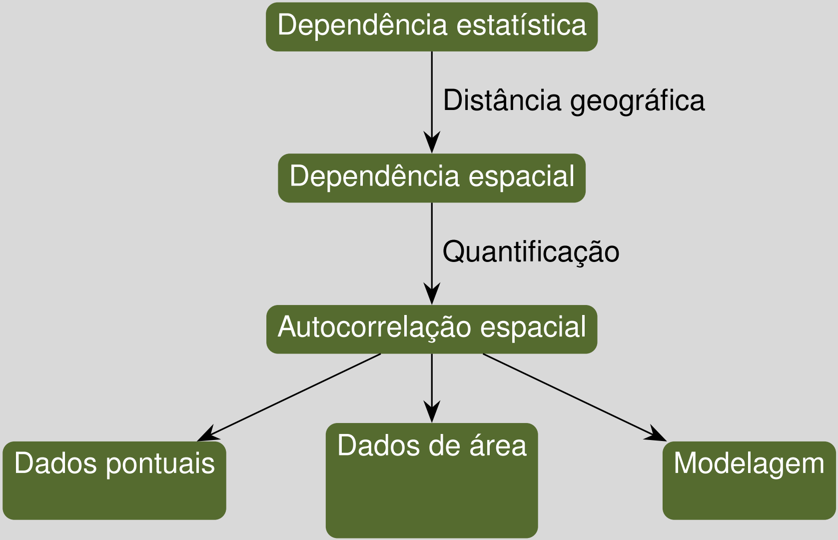
Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais

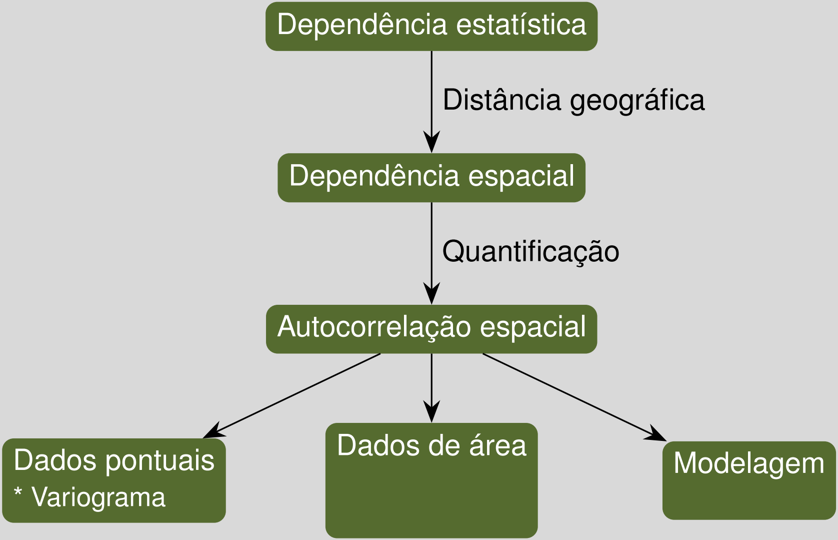
Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais
Dados de área

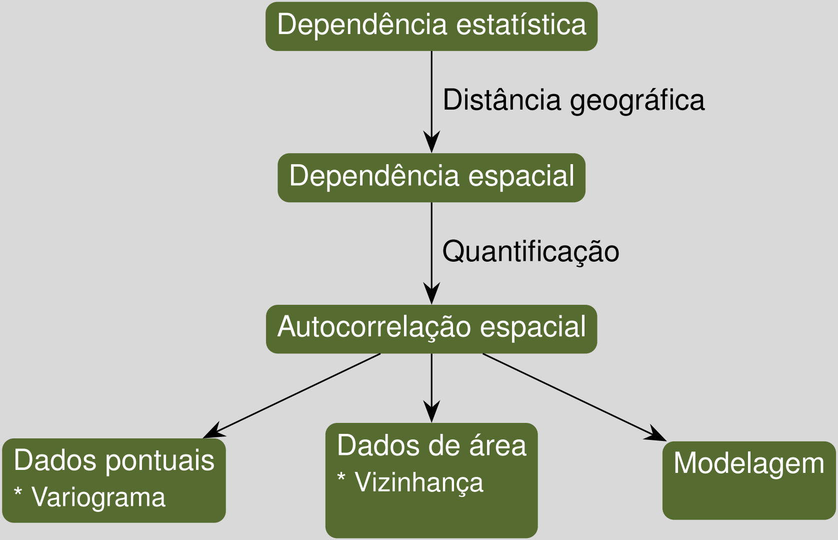
Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais
Dados de área
Modelagem

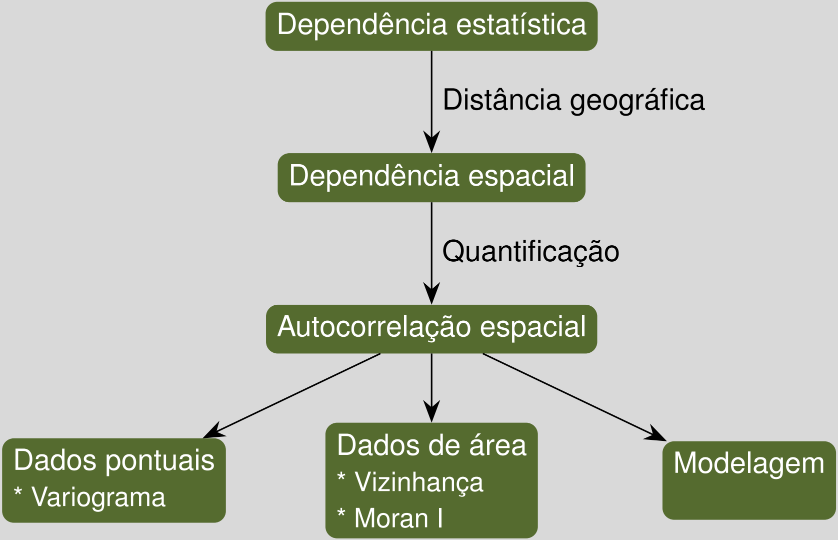
Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais
* Variograma
Dados de área
Modelagem

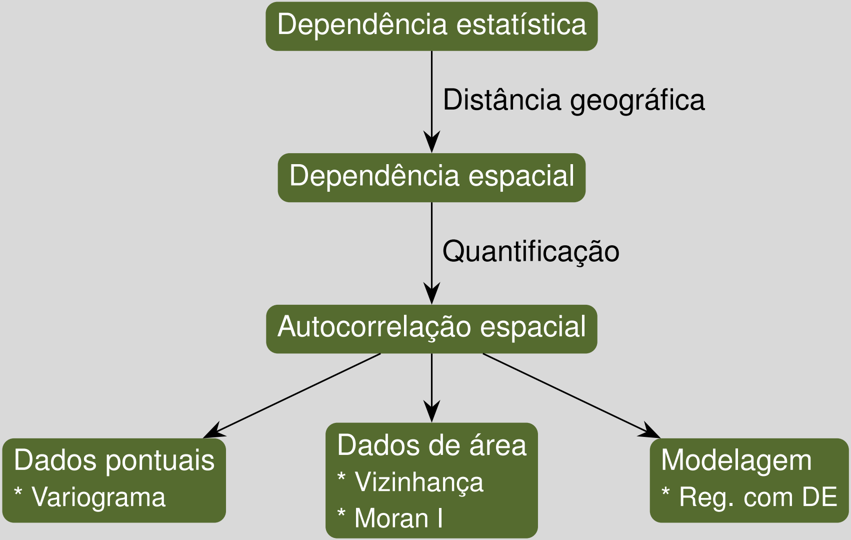
Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais
* Variograma
Dados de área
* Vizinhança
Modelagem

Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais
* Variograma
Dados de área
* Vizinhança
* Moran I
Modelagem

Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais
* Variograma
Dados de área
* Vizinhança
* Moran I
Modelagem
* Reg. com DE

Por que se preocupar com a dependência espacial?

- Métodos para dados independentes

Por que se preocupar com a dependência espacial?

- Métodos para dados independentes $\Rightarrow$ análises de dados com dependência espacial

Por que se preocupar com a dependência espacial?

- Métodos para dados independentes $\Rightarrow$ análises de dados com dependência espacial $\Rightarrow$ conclusões erradas.

Por que se preocupar com a dependência espacial?

- Métodos para dados independentes $\Rightarrow$ análises de dados com dependência espacial $\Rightarrow$ conclusões erradas.
- A predição de desfechos auxilia as ações de vigilância e a acurácia das predições pode melhorar ao levar em conta a dependência espacial.

Dependência espacial

Dependência espacial é um tipo de relação caracterizada pela *dependência estatística* de uma coleção de variáveis aleatórias geolocalizadas.

Dependência espacial

Dependência espacial é um tipo de relação caracterizada pela *dependência estatística* de uma coleção de variáveis aleatórias geolocalizadas.

A dependência estatística entre dois eventos significa que a ocorrência de um evento afeta a probabilidade de ocorrência do outro.

Dependência espacial

Dependência espacial é um tipo de relação caracterizada pela *dependência estatística* de uma coleção de variáveis aleatórias geolocalizadas.

A dependência estatística entre dois eventos significa que a ocorrência de um evento afeta a probabilidade de ocorrência do outro.

$$P(B) \neq P(B|A)$$

Dependência espacial

Dependência espacial é um tipo de relação caracterizada pela *dependência estatística* de uma coleção de variáveis aleatórias geolocalizadas.

A dependência estatística entre dois eventos significa que a ocorrência de um evento afeta a probabilidade de ocorrência do outro.

$$P(B) \neq P(B|A)$$

Nesses casos, tanto a covariância como a correlação entre A e B são diferentes de zero.

Dependência espacial

- A dependência estatística é espacial quando é expressada em função da distância entre os eventos.

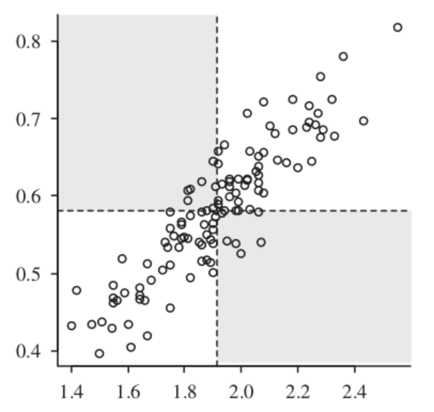
Dependência espacial

- A dependência estatística é espacial quando é expressada em função da distância entre os eventos.
- As funções da distância entre eventos são definidas formalmente em vários modelos espaciais.

Dependência espacial

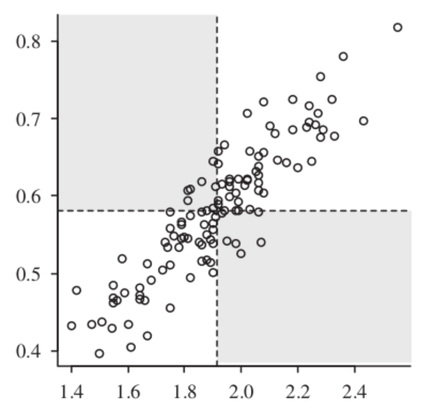
- A dependência estatística é espacial quando é expressada em função da distância entre os eventos.
- As funções da distância entre eventos são definidas formalmente em vários modelos espaciais.

Correlação



Matthiopoulos, 2011.

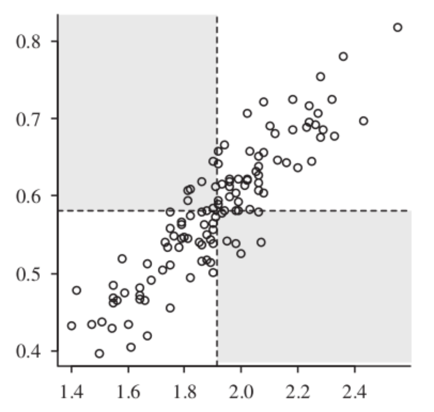
Correlação



Matthiopoulos, 2011.

- O padrão da figura indica uma possível relação linear.

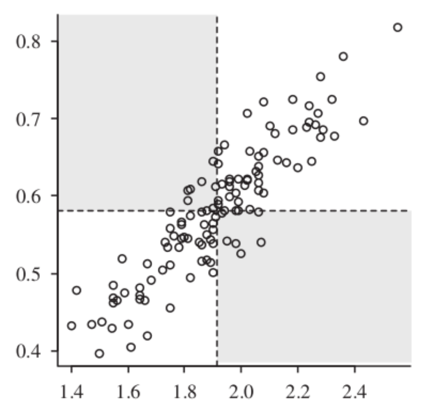
Correlação



Matthiopoulos, 2011.

- O padrão da figura indica uma possível relação linear.
- Mas podemos quantificar a força dessa relação?

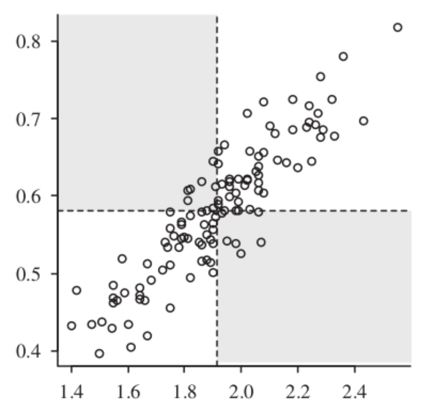
Correlação



Matthiopoulos, 2011.

- O padrão da figura indica uma possível relação linear.
- Mas podemos quantificar a força dessa relação?
- Podemos sim, medindo a correlação.

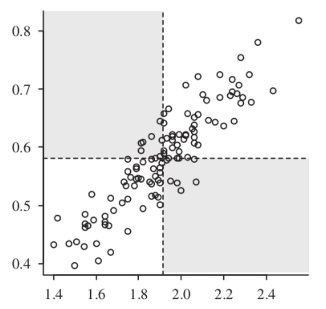
Correlação



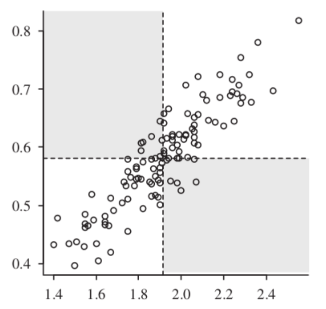
Matthiopoulos, 2011.

- O padrão da figura indica uma possível relação linear.
- Mas podemos quantificar a força dessa relação?
- Podemos sim, medindo a correlação.
- A estatística de correlação mais usada é o coeficiente de correlação de Pearson.

Correlação: o caso do coeficiente de Pearson

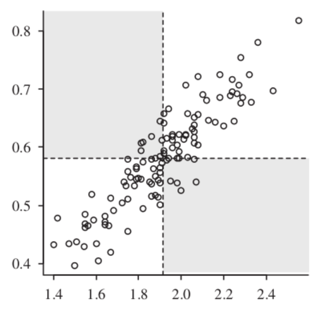


Correlação: o caso do coeficiente de Pearson



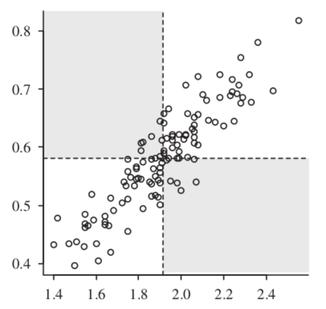
- Variáveis aleatórias: X e Y.

Correlação: o caso do coeficiente de Pearson



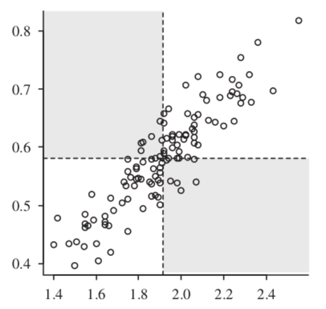
- Variáveis aleatórias: X e Y.
- Médias: $\bar{x}$ e $\bar{y}$.

Correlação: o caso do coeficiente de Pearson



- Variáveis aleatórias: X e Y.
- Médias: $\bar{x}$ e $\bar{y}$.
- Desvios padrão: s_x e s_y.

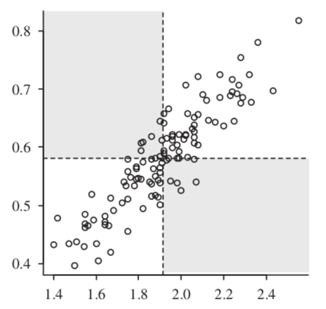
Correlação: o caso do coeficiente de Pearson



- Variáveis aleatórias: X e Y.
- Médias: $\bar{x}$ e $\bar{y}$.
- Desvios padrão: s_x e s_y.

- As diferenças $x_i - \bar{x}$ e $y_i - \bar{y}$, indicam se x_i e y_i estão acima ou abaixo da respectiva média.

Correlação: o caso do coeficiente de Pearson



- Variáveis aleatórias: X e Y.
- Médias: $\bar{x}$ e $\bar{y}$.
- Desvios padrão: s_x e s_y.

- As diferenças $x_i - \bar{x}$ e $y_i - \bar{y}$, indicam se x_i e y_i estão acima ou abaixo da respectiva média.
- O produto dessas diferenças indica o sentido da relação.

Correlação: o caso do coeficiente de Pearson

- A média dos produtos das diferenças de todos os pares de pontos diz se a relação é predominantemente positiva ou negativa.

Correlação: o caso do coeficiente de Pearson

- A média dos produtos das diferenças de todos os pares de pontos diz se a relação é predominantemente positiva ou negativa.
- Essa média nada mais é do que a covariância!

$$cov(x, y) = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$

Correlação: o caso do coeficiente de Pearson

- A média dos produtos das diferenças de todos os pares de pontos diz se a relação é predominantemente positiva ou negativa.
- Essa média nada mais é do que a covariância!

$$cov(x, y) = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$

- Ao limitar a covariância a valores entre -1 e 1, surge o coeficiente de correlação de Pearson.

Correlação: o caso do coeficiente de Pearson

- A média dos produtos das diferenças de todos os pares de pontos diz se a relação é predominantemente positiva ou negativa.
- Essa média nada mais é do que a covariância!

$$cov(x, y) = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$

- Ao limitar a covariância a valores entre -1 e 1, surge o coeficiente de correlação de Pearson.

$$r = \frac{cov(x, y)}{s_x s_y}$$

Correlação: o caso do coeficiente de Pearson

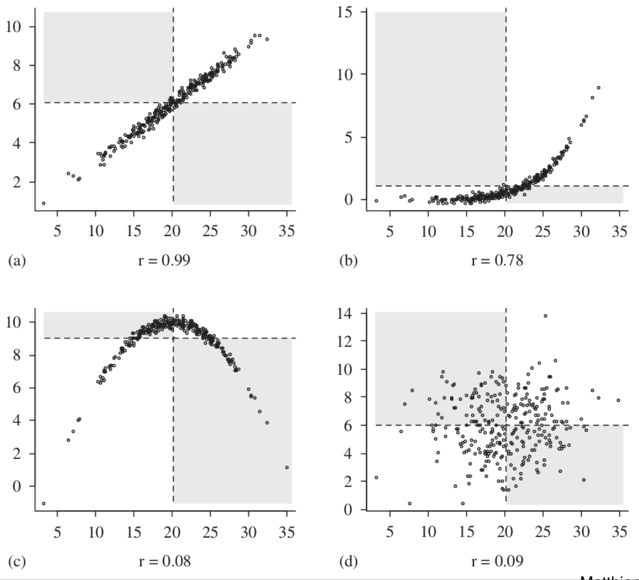
- A média dos produtos das diferenças de todos os pares de pontos diz se a relação é predominantemente positiva ou negativa.
- Essa média nada mais é do que a covariância!

$$cov(x, y) = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$

- Ao limitar a covariância a valores entre -1 e 1, surge o coeficiente de correlação de Pearson.

$$r = \frac{cov(x, y)}{s_x s_y} = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2} \sqrt{\sum_i^n (y_i - \bar{y})^2}}$$

Correlação



Matthiopoulos, 2011.

Autocorrelação

- A correlação que acabamos de ver quantifica a relação linear entre duas variáveis X e Y, com base nos pares:

$$\begin{gathered}(x_1, y_1)\\(x_2, y_2)\\\vdots\\(x_n, y_n)\end{gathered}$$

Autocorrelação

- A correlação que acabamos de ver quantifica a relação linear entre duas variáveis X e Y, com base nos pares:

$$\begin{array}{c}(x_1, y_1)\\(x_2, y_2)\\\vdots\\(x_n, y_n)\end{array}$$

- A AC quantifica a relação linear de uma variável com ela mesma.

Autocorrelação

- A correlação que acabamos de ver quantifica a relação linear entre duas variáveis X e Y, com base nos pares:

$$\begin{gathered}(x_1, y_1)\\(x_2, y_2)\\\vdots\\(x_n, y_n)\end{gathered}$$

- A AC quantifica a relação linear de uma variável com ela mesma.

- Na AC temporal, os pares são dados por $(x_1, x_{1+l}), (x_2, x_{2+l}), ...,$ (x_n, x_{n+l}), onde l é o *lag* o desfasagem temporal.

Autocorrelação

- A correlação que acabamos de ver quantifica a relação linear entre duas variáveis X e Y, com base nos pares:

$$\begin{array}{c}(x_1, y_1)\\(x_2, y_2)\\\vdots\\(x_n, y_n)\end{array}$$

- A AC quantifica a relação linear de uma variável com ela mesma.

- Na AC temporal, os pares são dados por $(x_1, x_{1+l}), (x_2, x_{2+l}), ...,$ (x_n, x_{n+l}), onde l é o *lag* o desfasagem temporal.
- Assim, para $l = 1$, os pares seriam:

$$\begin{array}{c}x_1\\(x_1, x_2)\\(x_2, x_3)\\\vdots\\(x_{n-1}, x_n)\\x_n\end{array}$$

Autocorrelação

- A AC temporal envolve uma dimensão só, o tempo.

Autocorrelação

- A AC temporal envolve uma dimensão só, o tempo.
- A ACE envolve duas dimensões (ex. latitude e longitude), mas a ideia continua sendo a quantificação da relação linear de uma variável (geolocalizada) com ela mesma.

Autocorrelação

- A AC temporal envolve uma dimensão só, o tempo.
- A ACE envolve duas dimensões (ex. latitude e longitude), mas a ideia continua sendo a quantificação da relação linear de uma variável (geolocalizada) com ela mesma.
- Existem várias medidas de correlação para dados espaciais representados por pontos e por áreas.

Autocorrelação e análises espaciais

Quantificação da ACE

- É em si uma análise espacial que fornece informação sobre a natureza dos dados.

Autocorrelação e análises espaciais

Quantificação da ACE

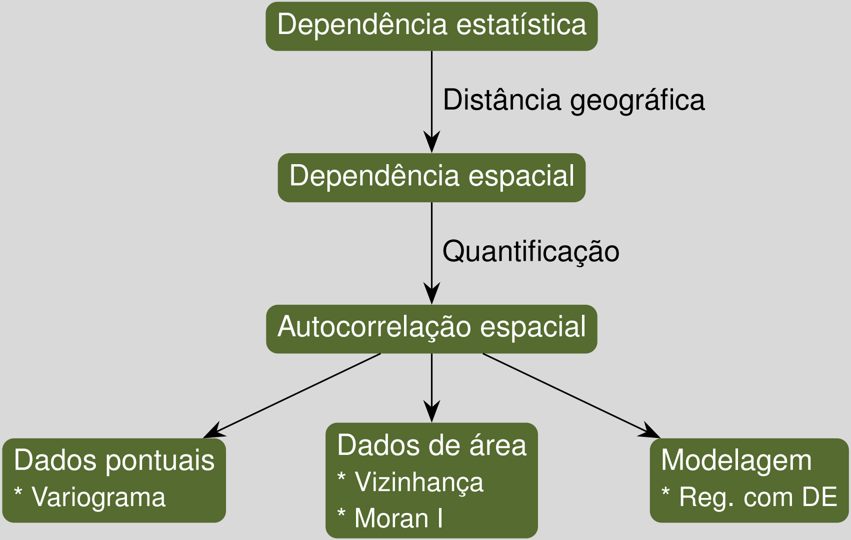
- É em si uma análise espacial que fornece informação sobre a natureza dos dados.
- Serve para determinar se os dados podem ser analisados mediante métodos que pressupõem independência estatística.

Autocorrelação e análises espaciais

Quantificação da ACE

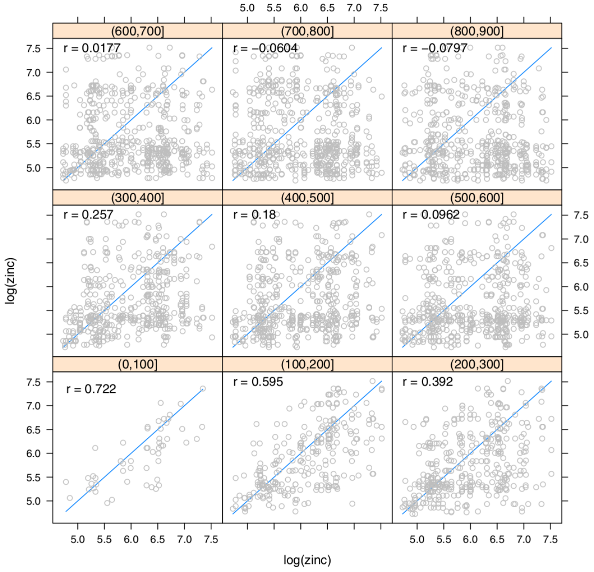
- É em si uma análise espacial que fornece informação sobre a natureza dos dados.
- Serve para determinar se os dados podem ser analisados mediante métodos que pressupõem independência estatística.
- Pode ser incorporada na modelagem para prescindir do pressuposto da independência estatística.

Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais
* Variograma
Dados de área
* Vizinhança
* Moran I
Modelagem
* Reg. com DE

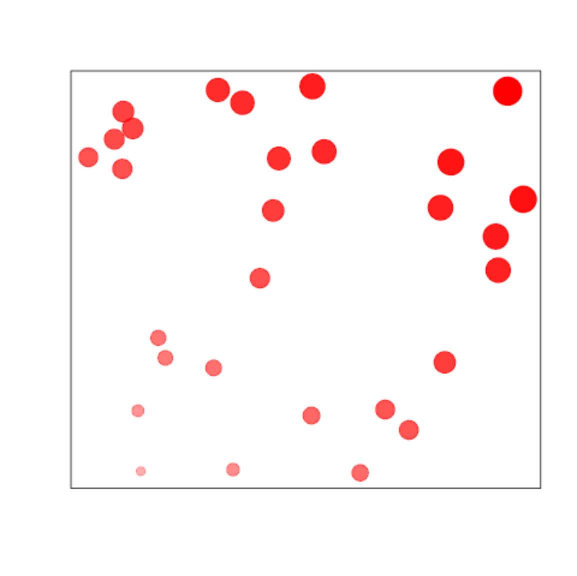
Desfasagem espacial

Para cada intervalo de desfasagem, a correlação é dada pelos pares separados por uma distância contida no intervalo.

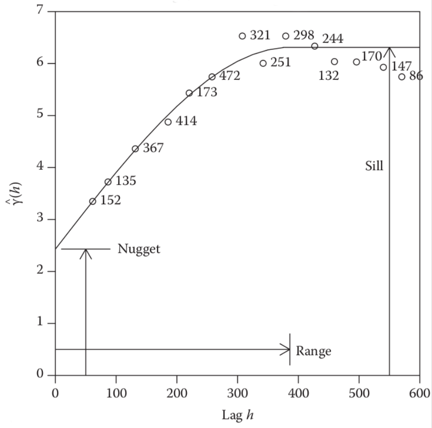


Bivand, 2013.

Variograma

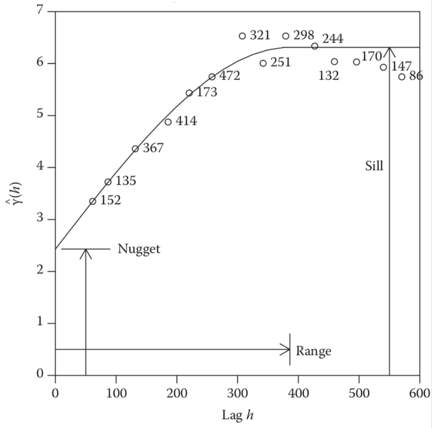


Variograma



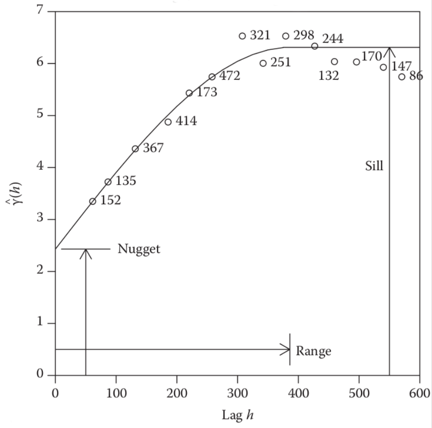
7
6
5
4
3
2
1
0
0
100
200
300
400
500
600
Lag h
$\hat{\gamma}(h)$
321
298
244
251
170
132
147
86
472
173
414
367
135
152
Sill
Nugget
Range

Variograma



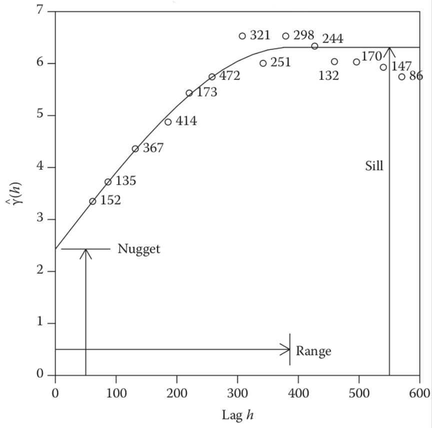
- Até o alcance, a variância aumenta com o aumento de *h* (quantos mais distantes os pontos, maior a diferença do valor das suas observações).

Variograma



- Até o alcance, a variância aumenta com o aumento de *h* (quantos mais distantes os pontos, maior a diferença do valor das suas observações).
- Após o alcance, a variância estabiliza-se (não há mais dependência espacial).

Variograma



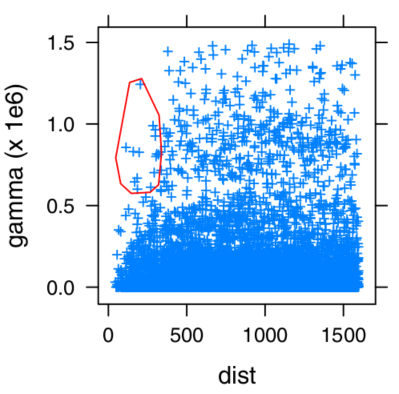
- Até o alcance, a variância aumenta com o aumento de h (quantos mais distantes os pontos, maior a diferença do valor das suas observações).
- Após o alcance, a variância estabiliza-se (não há mais dependência espacial).
- Sem erro amostral nem variabilidade de pequena escala (nugget = 0), o valor de observações com a mesma posição seria o mesmo.

Nuvem do variograma

- Distância entre um par contra a diferença quadrática dos seus valores, para todos os pares possíveis.

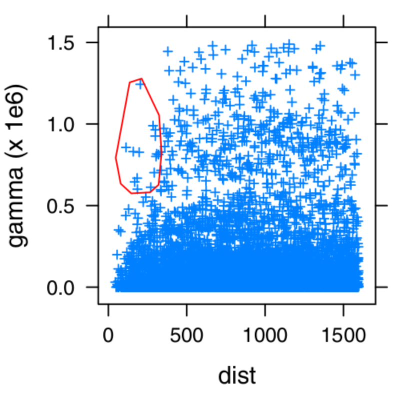
Nuvem do variograma

- Distância entre um par contra a diferença quadrática dos seus valores, para todos os pares possíveis.



Nuvem do variograma

- Distância entre um par contra a diferença quadrática dos seus valores, para todos os pares possíveis.

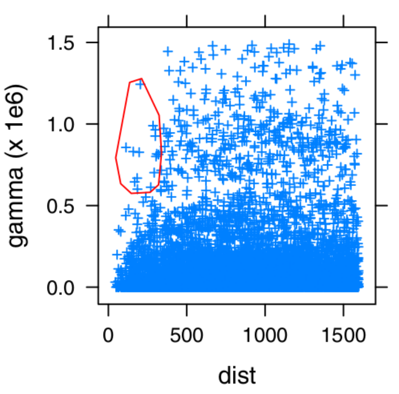


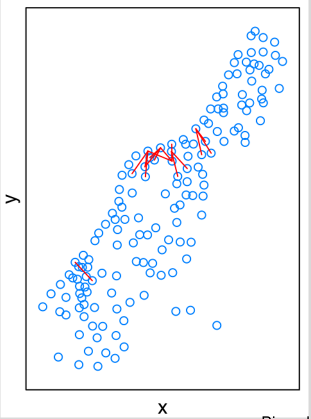
Valores extremos indicam gradientes marcados e o mapeamento dos mesmos dá uma ideia da “clusterização”.

Nuvem do variograma

- Distância entre um par contra a diferença quadrática dos seus valores, para todos os pares possíveis.

Valores extremos indicam gradientes marcados e o mapeamento dos mesmos dá uma ideia da "clusterização".

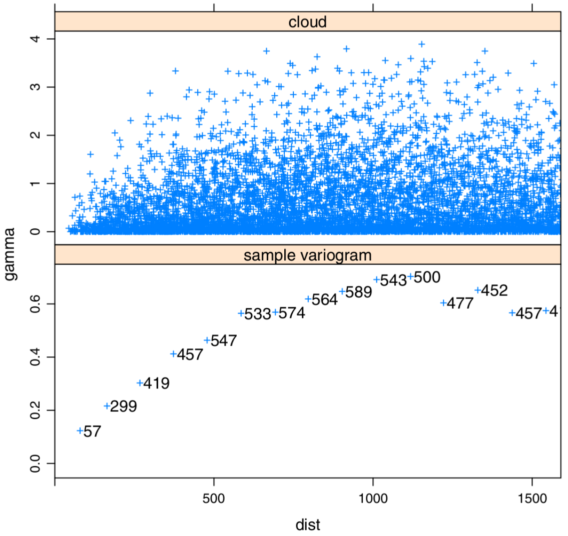




Bivand, 2013.

Variograma amostral

- Os pares são agrupados em intervalos para calcular o variograma.



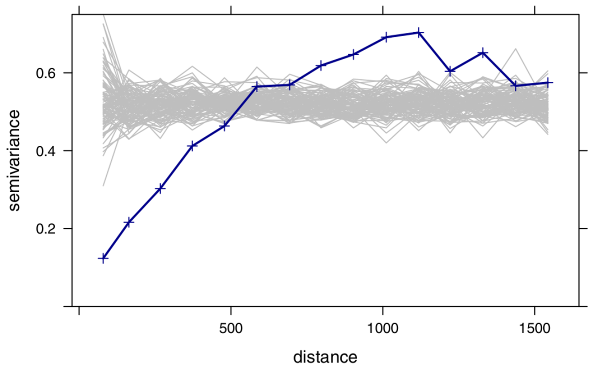
Bivand, 2013.

Significância do variograma amostral

- Variogramas adicionais são calculados com permutações dos valores das observações, mantendo as posições (simulações de Monte Carlo).

Significância do variograma amostral

- Variogramas adicionais são calculados com permutações dos valores das observações, mantendo as posições (simulações de Monte Carlo).
- O variograma amostral não é significativo se fica contido no envelope de simulação.



Plant, 2012.

Variogramas ajustados

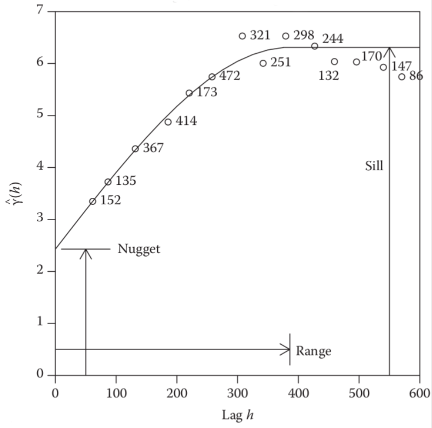
- Existem modelos paramétricos e não paramétricos para ajustar variogramas.

Variogramas ajustados

- Existem modelos paramétricos e não paramétricos para ajustar variogramas.

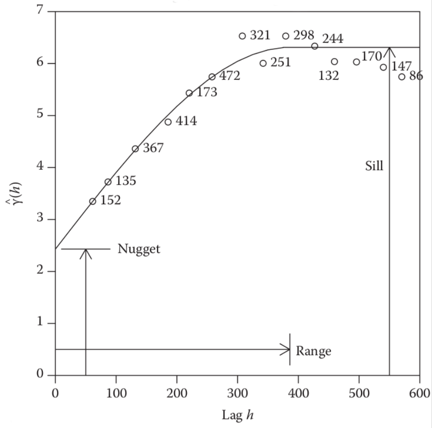
- **Nugget**
- **Exponencial**
- **Esférico**
- **Gaussiano**
- **Matérn**
- **Potência**
- Circular
- Linear
- Bessel
- Pentaesférico
- Periódico
- Hole
- Logarítmico
- Spline
- Legendre
- Erro de medição
- Intercepto

Modelo esférico com nugget



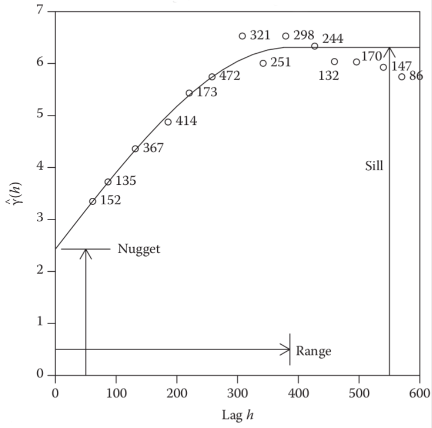
- Até o alcance, a variância aumenta com o aumento de *h* (quantos mais distantes os pontos, maior a diferença do valor das suas observações).

Modelo esférico com nugget



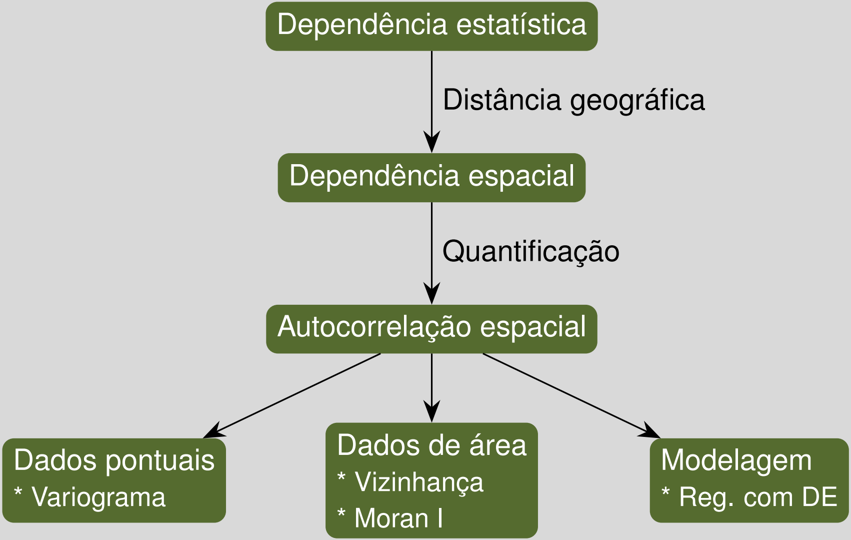
- Até o alcance, a variância aumenta com o aumento de *h* (quantos mais distantes os pontos, maior a diferença do valor das suas observações).
- Após o alcance, a variância estabiliza-se (não há mais dependência espacial).

Modelo esférico com nugget



- Até o alcance, a variância aumenta com o aumento de *h* (quantos mais distantes os pontos, maior a diferença do valor das suas observações).
- Após o alcance, a variância estabiliza-se (não há mais dependência espacial).
- Sem erro amostral nem variabilidade de pequena escala (nugget = 0), o valor de observações com a mesma posição seria o mesmo.

Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais
* Variograma
Dados de área
* Vizinhança
* Moran I
Modelagem
* Reg. com DE

Vizinhança e autocorrelação

Escolha de critérios de vizinhança

Vizinhança e autocorrelação

Escolha de critérios de vizinhança

↓

Identificação de relações de vizinhança

Vizinhança e autocorrelação

Escolha de critérios de vizinhança

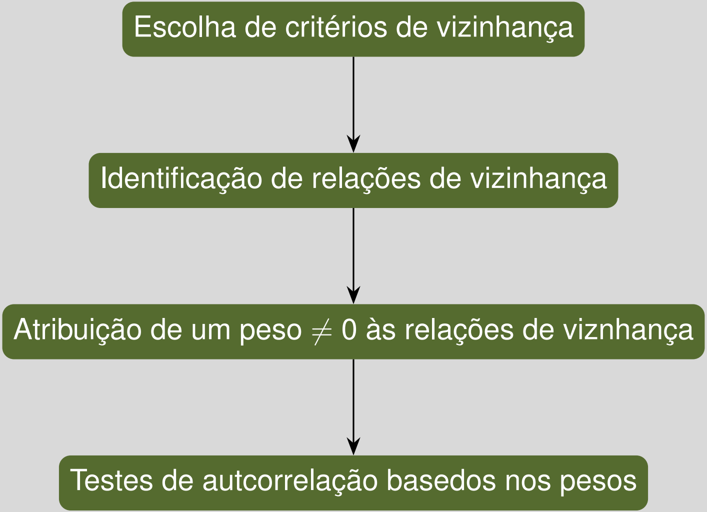
↓

Identificação de relações de vizinhança

↓

Atribuição de um peso $\neq 0$ às relações de viznhança

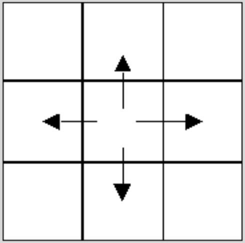
Vizinhança e autocorrelação



Vizinhança por contiguidade

Vizinhos tipo Rook

Lados em comum.



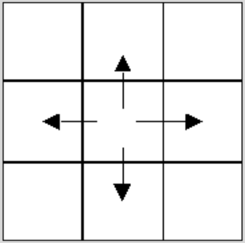
Vizinhança por contiguidade

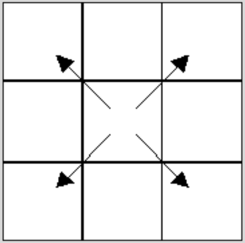
Vizinhos tipo Rook

Lados em comum.

Vizinhos tipo Bishop

Pontos em comum.





Vizinhança por contiguidade

Vizinhos tipo Rook

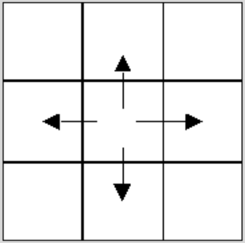
Lados em comum.

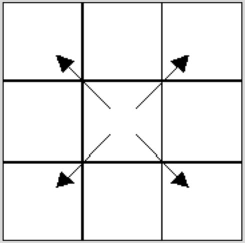
Vizinhos tipo Bishop

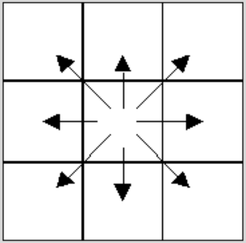
Pontos em comum.

Vizinhos tipo Queen

Lados ou pontos em comum.







Vizinhança baseada em grafos

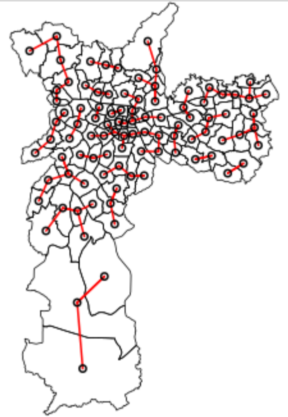
- Os polígonos podem ser representados pelos seus centroides.

Vizinhança baseada em grafos

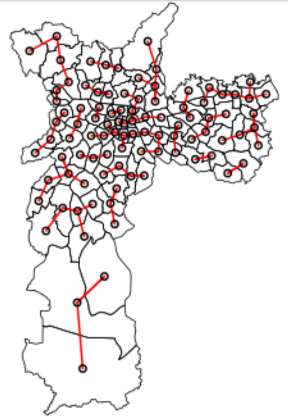
- Os polígonos podem ser representados pelos seus centroides.
- Os critérios de vizinhança entre centroides são baseados em grafos.

K vizinhos mais próximos

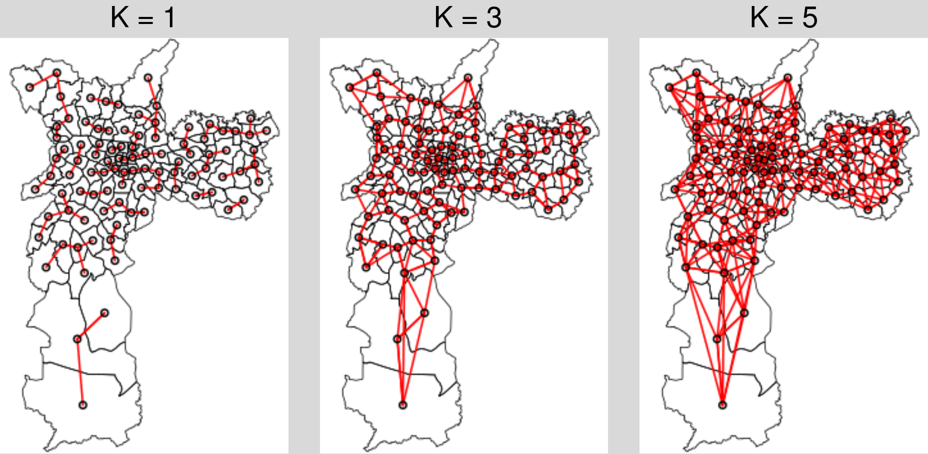
K = 1



K vizinhos mais próximos



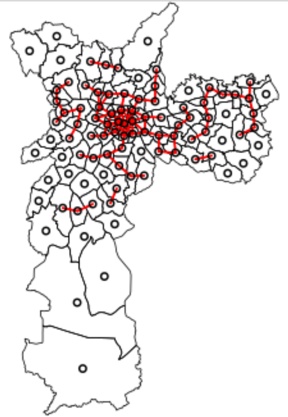
K vizinhos mais próximos



Vizinhança baseada em distância

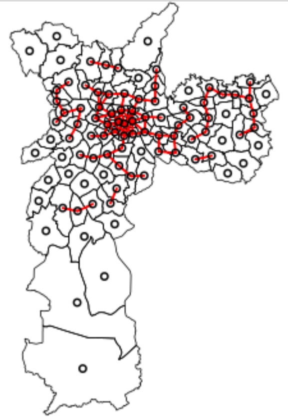
Vizinhança baseada em distância

3 km



Vizinhança baseada em distância

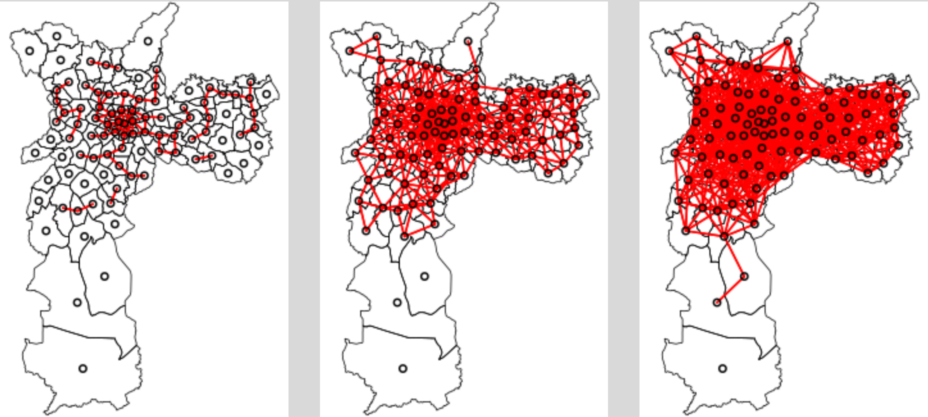
3 km



6 km

Vizinhança baseada em distância

3 km 6 km 9 km



Vizinhança de ordens maiores

- Os tipos de vizinhança que vimos até o momento são de primeira ordem.

Vizinhança de ordens maiores

- Os tipos de vizinhança que vimos até o momento são de primeira ordem.
- Na vizinhança de 2a ordem, são considerados os vizinhos dos vizinhos.

Vizinhança de ordens maiores

- Os tipos de vizinhança que vimos até o momento são de primeira ordem.
- Na vizinhança de 2a ordem, são considerados os vizinhos dos vizinhos.
- Na vizinhança de 3a ordem, os vizinhos dos vinhos dos vizinhos.

Vizinhança de ordens maiores

- Os tipos de vizinhança que vimos até o momento são de primeira ordem.
- Na vizinhança de 2a ordem, são considerados os vizinhos dos vizinhos.
- Na vizinhança de 3a ordem, os vizinhos dos vinhos dos vizinhos.
- A ideia é a mesma para vizinhos de ordens maiores.

Pesos

- As relações de vizinhança podem ser representadas em uma matriz W de pesos.

Pesos

- As relações de vizinhança podem ser representadas em uma matriz W de pesos.
- Se $w_{ij} \neq 0$, então i e j são vizinhos.

Pesos

- As relações de vizinhança podem ser representadas em uma matriz W de pesos.
- Se $w_{ij} \neq 0$, então i e j são vizinhos.

$$W = \begin{array}{cc} & \begin{array}{ccc} a & b & c \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

Estilos de pesos

W

- Pesos iguais com somatória por linha igual a 1.

Estilos de pesos

W

- Pesos iguais com somatória por linha igual a 1.
- Quanto mais vizinhos tiver uma área, menor o peso deles.

Estilos de pesos

W

- Pesos iguais com somatória por linha igual a 1.
- Quanto mais vizinhos tiver uma área, menor o peso deles.

B

- O peso para todos os vizinhos é 1.

Estilos de pesos

W

- Pesos iguais com somatória por linha igual a 1.
- Quanto mais vizinhos tiver uma área, menor o peso deles.

B

- O peso para todos os vizinhos é 1.
- Somatória por linha igual ao número de vizinhos.

Estilos de pesos

W

- Pesos iguais com somatória por linha igual a 1.
- Quanto mais vizinhos tiver uma área, menor o peso deles.

B

- O peso para todos os vizinhos é 1.
- Somatória por linha igual ao número de vizinhos.

C

Pesos iguais com somatória total igual ao número de áreas.

Estilos de pesos

W

- Pesos iguais com somatória por linha igual a 1.
- Quanto mais vizinhos tiver uma área, menor o peso deles.

B

- O peso para todos os vizinhos é 1.
- Somatória por linha igual ao número de vizinhos.

C

Pesos iguais com somatória total igual ao número de áreas.

U

Pesos iguais com somatória total igual a 1.

Estilos de pesos

S

- Os pesos variam menos que em W.

Estilos de pesos

S

- Os pesos variam menos que em W.
- Somatória por linha maior que em W e menor que em B, C e U.

Estilos de pesos

S

- Os pesos variam menos que em W.
- Somatória por linha maior que em W e menor que em B, C e U.

Inverso da distância

Pesos iguais ao inverso (ou inverso quadrático) da distância entre os vizinhos.

Escolha de critérios de vizinhança e do tipo de pesos

- Influencia os resultados dos testes e da modelagem da autocorrelação espacial.

Escolha de critérios de vizinhança e do tipo de pesos

- Influencia os resultados dos testes e da modelagem da autocorrelação espacial.
- Depende da natureza dos dados, por exemplo:
    - Se não há razões para escolher um tipo específico de pesos, o tipo binário (B), é uma escolha razoável.

Medidas

Globais

Medidas

Globais

- **Moran I**

Medidas

Globais

- **Moran I**
- Getis Ord G

Medidas

Globais

- **Moran I**
- Getis Ord G
- Geary C

Medidas

Globais

- **Moran I**
- Getis Ord G
- Geary C
- Produto cruzado geral de Mantel

Medidas

Globais

- **Moran I**
- Getis Ord G
- Geary C
- Produto cruzado geral de Mantel
- “Join-Count” (dados nominais)

Medidas

Globais

- **Moran I**
- Getis Ord G
- Geary C
- Produto cruzado geral de Mantel
- “Join-Count” (dados nominais)

Locais (LISA)

Medidas

Globais

- **Moran I**
- Getis Ord G
- Geary C
- Produto cruzado geral de Mantel
- “Join-Count” (dados nominais)

Locais (LISA)

- **Moran I**

Medidas

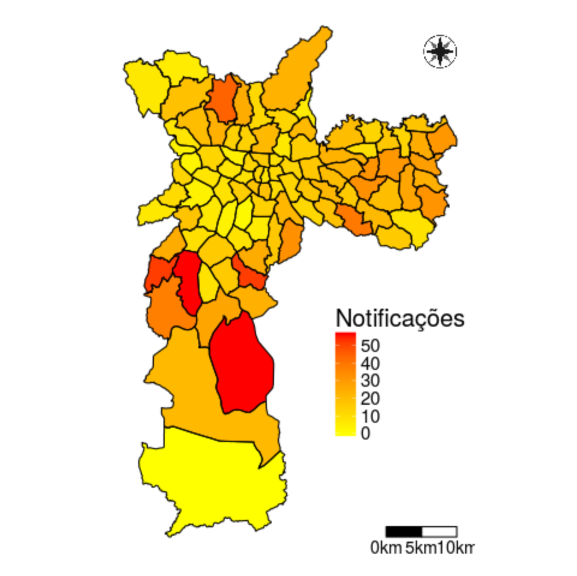
Globais

- **Moran I**
- Getis Ord G
- Geary C
- Produto cruzado geral de Mantel
- “Join-Count” (dados nominais)

Locais (LISA)

- **Moran I**
- Getis Ord G

Moran I global



Moran I global

- A esperança ($-1/(n-1)$) é negativa e tende a zero com o aumento do tamanho amostral.

Moran I global

- A esperança ($-1/(n-1)$) é negativa e tende a zero com o aumento do tamanho amostral.
- $I > -1/(n-1) \Rightarrow$ agrupamento de valores similares.

Moran I global

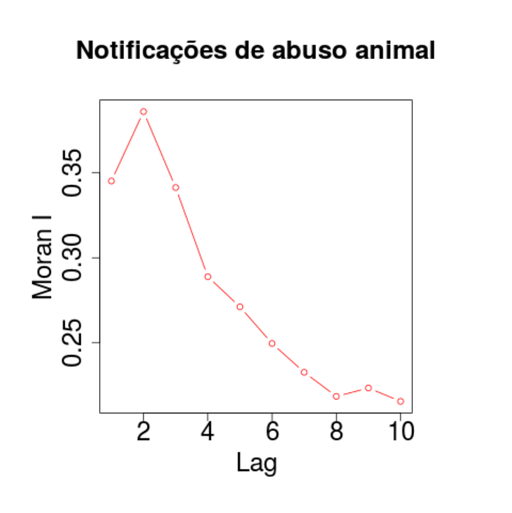
- A esperança ($-1/(n-1)$) é negativa e tende a zero com o aumento do tamanho amostral.
- $I > -1/(n-1) \Rightarrow$ agrupamento de valores similares.
- $I < -1/(n-1) \Rightarrow$ agrupamento de valores dissimilares.

Correlograma baseado no Moran I global

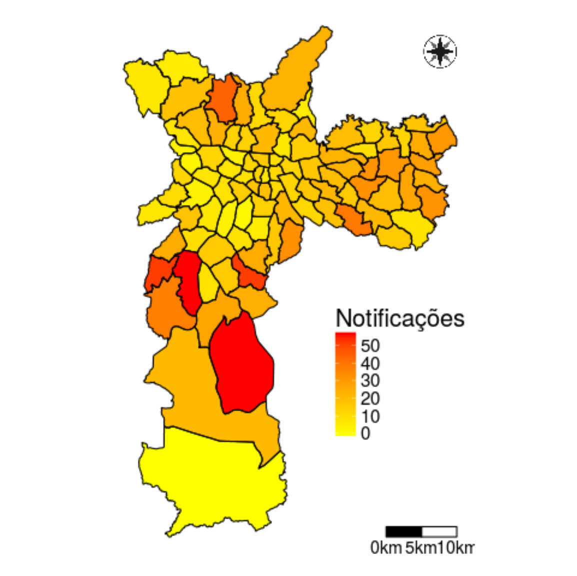
O Moran I é calculado para diferentes ordens (lags) de vizinhança.

Correlograma baseado no Moran I global

O Moran I é calculado para diferentes ordens (lags) de vizinhança.



Moran I local



Moran I local

A avaliação da autocorrelação de forma localizada serve para detectar:

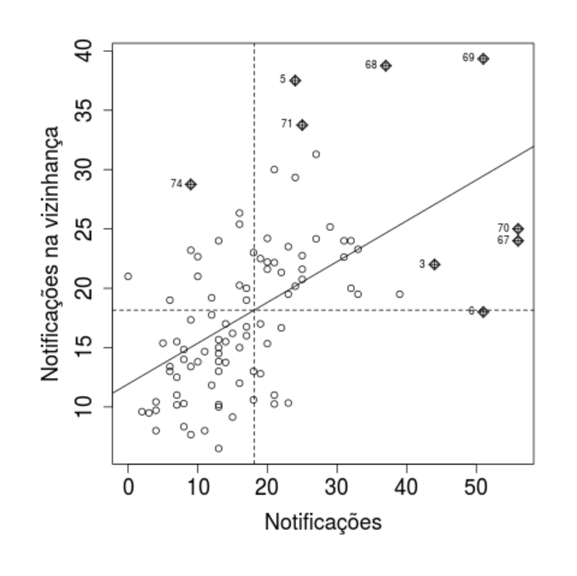
- Agrupamentos: observações com vizinhos similares.

Moran I local

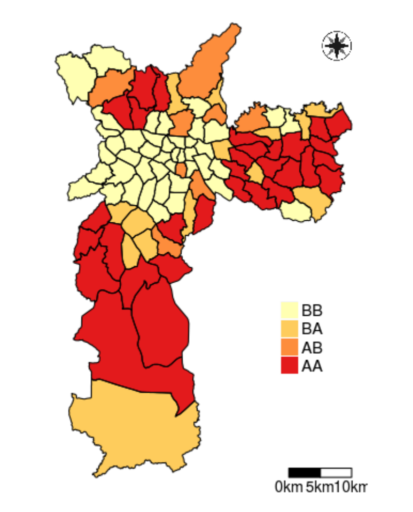
A avaliação da autocorrelação de forma localizada serve para detectar:

- Agrupamentos: observações com vizinhos similares.
- “Hotspots”: observações com vizinhos dissimilares.

Maran I scatterplot



Moran I local



Significância do Moran I

Formas de avaliação

- Analítica assumindo uma distribuição normal.

Significância do Moran I

Formas de avaliação

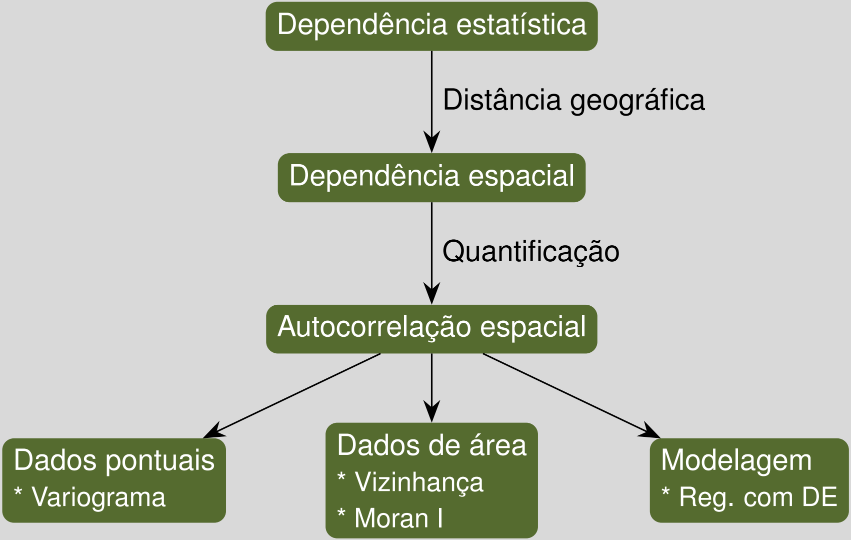
- Analítica assumindo uma distribuição normal.
- Testes de permutação.

Significância do Moran I

Formas de avaliação

- Analítica assumindo uma distribuição normal.
- Testes de permutação.

Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais
* Variograma
Dados de área
* Vizinhança
* Moran I
Modelagem
* Reg. com DE

Autocorrelação na modelagem espacial

- A modelagem espacial prescinde do pressuposto da independência espacial entre as observações.

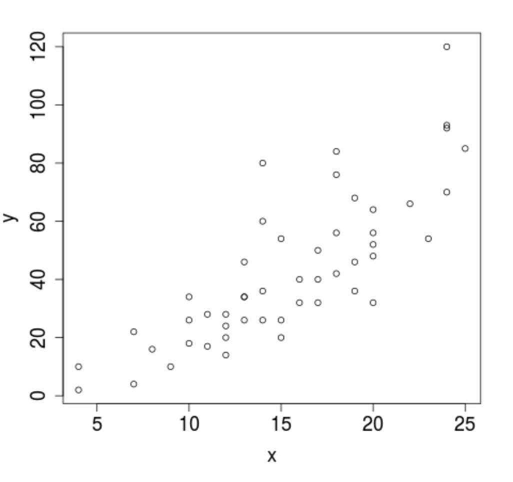
Autocorrelação na modelagem espacial

- A modelagem espacial prescinde do pressuposto da independência espacial entre as observações.
- No contexto da regressão, uma abordagem comum é usar um modelo convencional e testar a autocorrelação espacial nos resíduos.

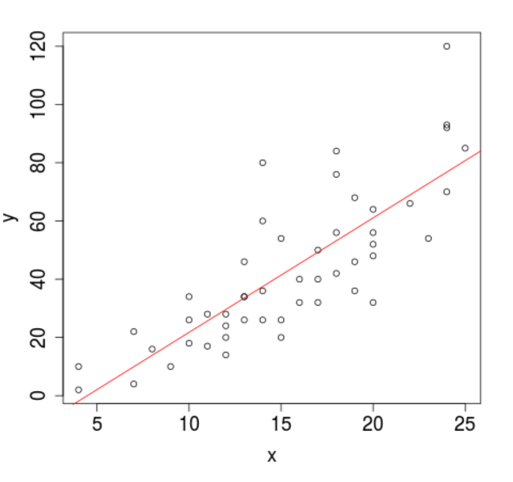
Autocorrelação na modelagem espacial

- A modelagem espacial prescinde do pressuposto da independência espacial entre as observações.
- No contexto da regressão, uma abordagem comum é usar um modelo convencional e testar a autocorrelação espacial nos resíduos.
- Se existe autocorrelação, deve ser usado um modelo que incorpore a autocorrelação espacial.

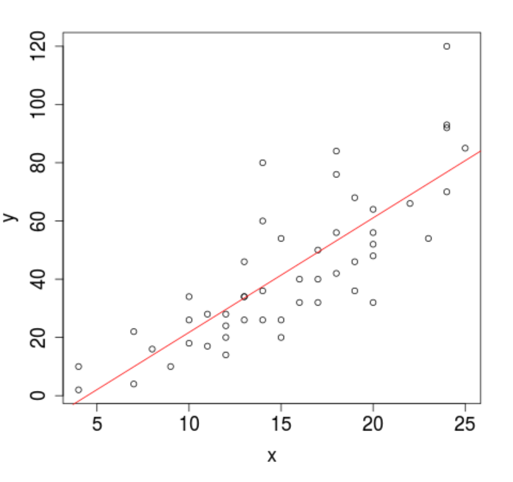
Regressões simples



Regressões simples

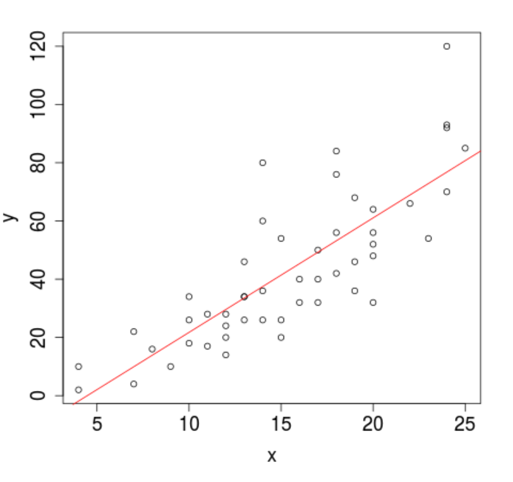


Regressões simples



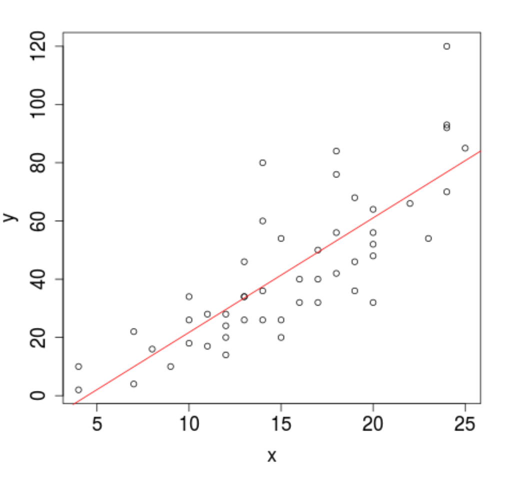
$$y_i = \beta_0$$

Regressões simples



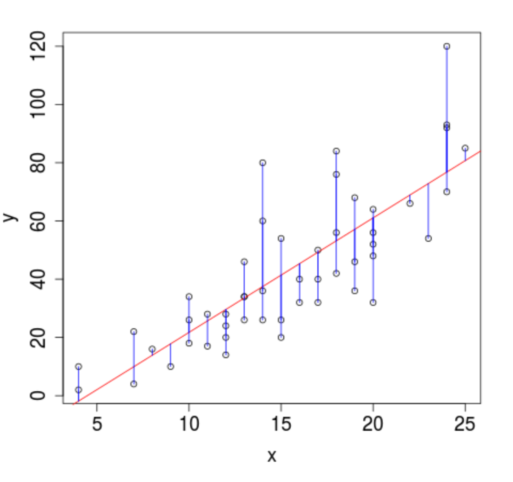
$$y_i = \beta_0 + \beta_1 x_i$$

Regressões simples



$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

Regressões simples



$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

Regressões simples espaciais

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

- Autorregressões simultâneas (SAR).

Regressões simples espaciais

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

- Autorregressões simultâneas (SAR).
- Autorregressões condicionais (CAR).

Regressões simples espaciais

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

- Autorregressões simultâneas (SAR).
- Autorregressões condicionais (CAR).
- Regressões espacias Bayesianas.

Regressões sem dependência

Regressão convencional: $y_i = \beta_0 + \beta_1 x_i + e_i$

Regressões sem dependência

Regressão convencional: $y_i = \beta_0 + \beta_1 x_i + e_i$
Pressupõe-se que as observações são independentes e a estrutura da covariância é dada por:

Obs	1	2	3	4	5	6	7	8	9	10
1	σ^2	0	0	0	0	0	0	0	0	0
2	0	σ^2	0	0	0	0	0	0	0	0
3	0	0	σ^2	0	0	0	0	0	0	0
4	0	0	0	σ^2	0	0	0	0	0	0
5	0	0	0	0	σ^2	0	0	0	0	0
6	0	0	0	0	0	σ^2	0	0	0	0
7	0	0	0	0	0	0	σ^2	0	0	0
8	0	0	0	0	0	0	0	σ^2	0	0
9	0	0	0	0	0	0	0	0	σ^2	0
10	0	0	0	0	0	0	0	0	0	σ^2

Regressões com dependência

Regressão multinível: $y_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij}$

Regressões com dependência

Regressão multinível: $y_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij}; \quad \beta_{0j} = b_0 + v_{0j}$

Regressões com dependência

Regressão multinível: $y_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij}; \quad \beta_{0j} = b_0 + v_{0j}$
As observações de um mesmo grupo podem ser dependentes e a estrutura da covariância é dada por:

Grupo		1	1	1	2	2	2	2	3	3	3
	Obs	1	2	3	1	2	3	4	1	2	3
1	1	Ω	Ψ	Ψ	0	0	0	0	0	0	0
1	2	Ψ	Ω	Ψ	0	0	0	0	0	0	0
1	3	Ψ	Ψ	Ω	0	0	0	0	0	0	0
2	1	0	0	0	Ω	Ψ	Ψ	Ψ	0	0	0
2	2	0	0	0	Ψ	Ω	Ψ	Ψ	0	0	0
2	3	0	0	0	Ψ	Ψ	Ω	Ψ	0	0	0
2	4	0	0	0	Ψ	Ψ	Ψ	Ω	0	0	0
3	1	0	0	0	0	0	0	0	Ω	Ψ	Ψ
3	2	0	0	0	0	0	0	0	Ψ	Ω	Ψ
3	3	0	0	0	0	0	0	0	Ψ	Ψ	Ω

Regressões com dependência espacial explícita

Autorregressão simultânea (SAR)

$$y_i = \beta_0 + \beta_1 x_1 + e_i$$

Regressões com dependência espacial explícita

Autorregressão simultânea (SAR)

$$y_i = \beta_0 + \beta_1 x_1 + e_i$$

$$e_i = \sum_{j=1}^{m} b_{ij} e_j + \epsilon_i$$

Regressões com dependência espacial explícita

Autorregressão simultânea (SAR)

$$y_i = \beta_0 + \beta_1 x_1 + e_i$$

$$e_i = \sum_{j=1}^{m} b_{ij} e_j + \epsilon_i$$

Autorregressão condicionada (CAR)

$$y_i = \beta_0 + \beta_1 x_1 + e_i$$

Regressões com dependência espacial explícita

Autorregressão simultânea (SAR)

$$y_i = \beta_0 + \beta_1 x_1 + e_i$$

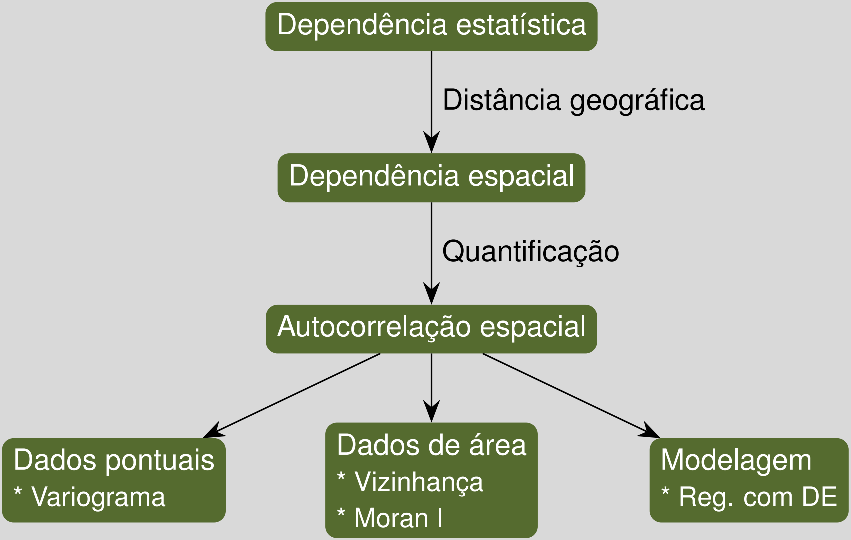
$$e_i = \sum_{j=1}^{m} b_{ij} e_j + \epsilon_i$$

Autorregressão condicionada (CAR)

$$y_i = \beta_0 + \beta_1 x_1 + e_i$$

$$e_i | e_{j \sim i} \sim N\left(\sum_{j \sim i} \frac{c_{ij} e_j}{\sum_{j \sim i} c_{ij}}, \frac{\sigma^2_{e_i}}{\sum_{j \sim i} c_{ij}} \right) + \epsilon_i$$

Sumário


Dependência estatística
Distância geográfica
Dependência espacial
Quantificação
Autocorrelação espacial
Dados pontuais
* Variograma
Dados de área
* Vizinhança
* Moran I
Modelagem
* Reg. com DE

Leituras sugeridas

- Bivand, R.; Pebesma, E; Gómez-Rubio, V. (2008). Applied spatial data analysis with R. 1st ed. Springer.
- Bivand, R.; Pebesma, E; Gómez-Rubio, V. (2012). Applied spatial data analysis with R. 2nd ed. Springer.
- Martins-Melo, F. R., Lima, M. D. S., Alencar, C. H., Ramos Jr, A. N., Carvalho, F. H. C., Machado, M. M. T., & Heukelbach, J. (2014). Temporal trends and spatial distribution of unsafe abortion in Brazil, 1996-2012. Revista de saude publica, 48(3), 508-520.
- Matthiopoulos, J. (2011). How to be a quantitative ecologist: the'A to R'of green mathematics and statistics. John Wiley & Sons.
- Plant, R. E. (2012). Spatial data analysis in ecology and agriculture using R. cRc Press.