



UNIVERSIDADE FEDERAL DE JUIZ DE FORA - *CAMPUS* Governador Valadares

ECONOMETRIA ESPACIAL

Prof. Vinícius de Azevedo Couto Firme

Econometria Espacial:

Capítulo 2 – Dados Espaciais



UNIVERSIDADE FEDERAL DE JUIZ DE FORA - *CAMPUS* Governador Valadares

ECONOMETRIA ESPACIAL

Prof. Vinícius de Azevedo Couto Firme

Estrutura da Apresentação:

1. Introdução
2. Fonte de Dados Espaciais
3. Processo Estocástico Espacial
4. Dados Espaciais e Inferência Estatística
5. Problemas Especiais com dados Espaciais



1. Introdução

- Os dados espaciais precisam informar não apenas o “quanto” um atributo variou, mas também “onde” esta variação ocorreu.
- Este tipo de dado apresenta algumas peculiaridades:
 - 1) São dados georreferenciados: a posição relativa e absoluta dos dados no mapa são importantes.
 - 2) Há multidirecionalidade de efeitos: os dados de uma região interagem com as demais em todas as direções e vice-versa, gerando endogeneidade.
 - 3) Há multidimensionalidade espacial: a dependência espacial pode variar conforme a direção especificada.



2. Fonte de Dados Espaciais

- A oferta de dados espaciais avançou juntamente com o desenvolvimento de satélites (mapas via GPS) e dos *softwares* com Sistemas de Informação Geográficas – SIG.

- Exemplos de *Softwares* Espaciais:
IPEAGEO – <http://www.ipea.gov.br/ipeageo/index.html>
GEODA/GEODA Space - <https://spatial.uchicago.edu/software>

- Algumas bases de dados espaciais disponíveis:
 - a) Socioeconômicos: <http://www.ipeadata.gov.br/>
 - b) Diversos: <http://dados.gov.br/>
 - c) Saúde: <http://tabnet.datasus.gov.br/>
 - d) Exportação/Importação: <http://comexstat.mdic.gov.br/pt/home>
 - e) Econômicos (Países): <http://www.imf.org/external/ns/cs.aspx?id=28>

- *Shapes* Espaciais:
Brasil - <http://www.usp.br/nereus/?dados=brasil>
<https://mapas.ibge.gov.br/bases-e-referenciais/bases-cartograficas/malhas-digitais>
Mundo - <http://www.mapcruzin.com/>



3. Processo Estocástico Espacial

- Os dados espaciais representam uma única realização de um processo estocástico do tipo espacial (ex: efeito do gasto em educação na renda municipal de 2010).
- Precisamos garantir que o Processo Gerador destes Dados (PGD ou PGM, caso gere mapas) seja estável. Em outras palavras, que o PGM seja representativo para o caso de outros mapas (i.e.: efeito da educação na renda deve ser válido para outros mapas).

Nota: Conceito análogo à estacionariedade das séries de tempo.

- De modo geral, um processo estocástico espacial pode ser descrito como:

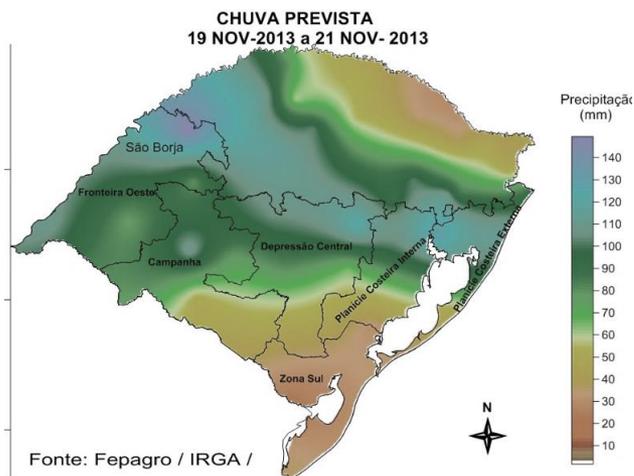
$$\{y_i: i \in D\} \tag{1}$$

Onde y_i é a variável aleatória de interesse associada à unidade espacial i , que pertence a um subconjunto fixo e finito (D) contido no espaço euclidiano de dimensão d (R^d), onde R representa um conjunto de números reais.

3. Processo Estocástico Espacial

- O subconjunto D pode ser representado no espaço de 3 formas:

Contínua
(Dados Geoestatísticos)



Discreta
(*Lattice Data*)



Forma de Pontos
(*Point Pattern Data*)



Nota: A análise contínua utiliza a técnica *Kriging* para prever valores de y em lugares desconhecidos.



4. Dados Espaciais e a Inferência Estatística

- Para que um mapa, gerado por uma variável aleatória, seja representativo, é necessário que o PGM apresente as seguintes características:
 - a) Média Constante e Finita: $E(y_i) = \mu$
 - b) Variância Constante e Finita: $\text{Var}(y_i) = \sigma_y^2$
 - c) $\text{Cov}(y_i, y_j) = \sigma_y^2 c(d_{ij}, \phi)$

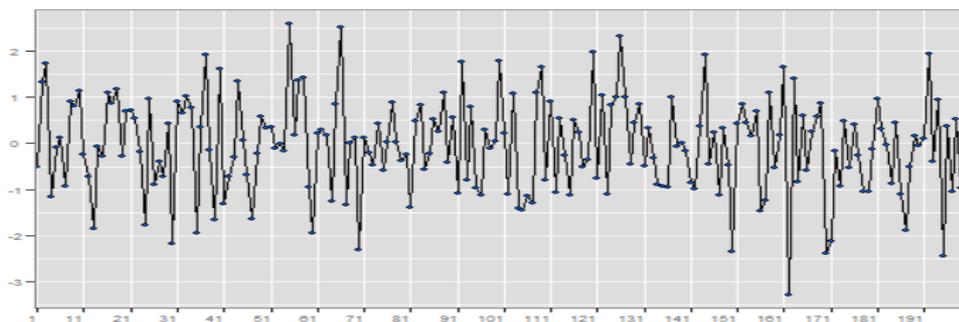
Os pressupostos “a” e “b” são semelhantes aos da Estacionariedade. Contudo, o pressuposto “c” indica que a covariância entre y_i e y_j deve depender da distância entre i e j (d_{ij}) e da direção entre eles, medida pelo coeficiente de angulação ϕ .

Como a covariância pode mudar conforme a direção (mesmo mantendo a distância), costuma-se assumir que há isotropia. Neste caso:

$$c.1) \text{Cov}(y_i, y_j) = \sigma_y^2 c(d_{ij})$$

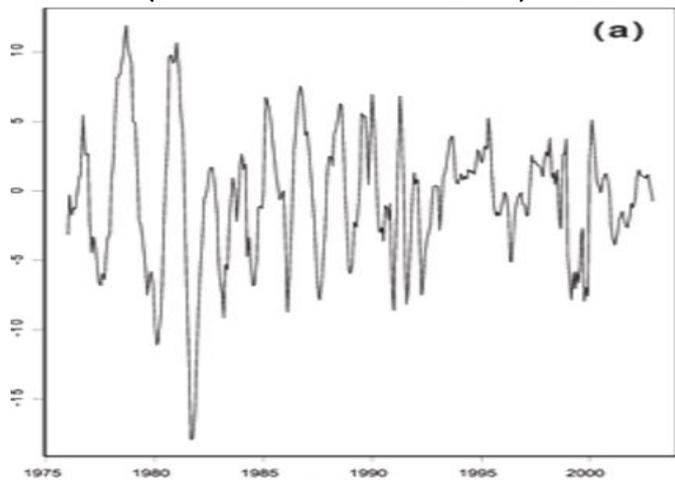
4. Dados Espaciais e a Inferência Estatística

- Relembrando Estacionariedade (séries de tempo)

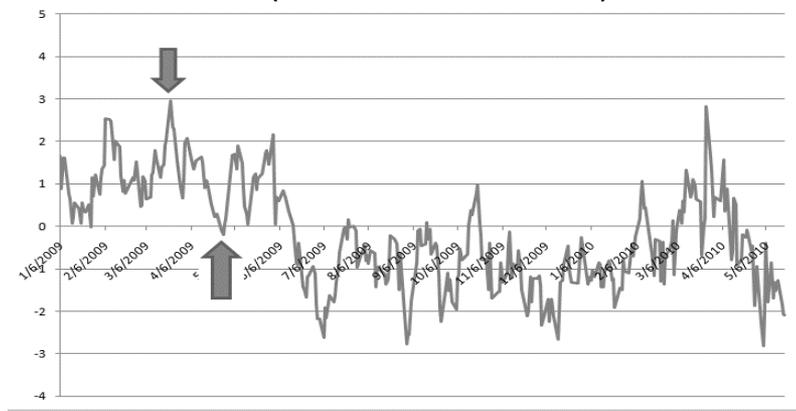


Série Estacionária

Não Estacionária
 (variância não constante)



Não Estacionária
 (média não constante)





5. Problemas Especiais com Dados Espaciais

5.1) Falácia Ecológica: Consiste em inferir sobre o comportamento individual usando dados agregados.

Ex. 1: Estudo de Robinson (1950):

- a) Correlação entre pop. Estrangeira e taxa de alfabetização c/ dados agregados (48 estados EUA) = 0.58.
- b) Correlação entre pop. Estrangeira e taxa de alfabetização c/ dados individuais (amostra dos mesmos 48 estados) = -0.11.

Ex. 2: Câncer de Esôfago é mais frequente no RS que em MG (estados).

Conclusão: Efeito chimarrão.

Problema: Será que as pessoas (indivíduos) com câncer de esôfago no RS tomam chimarrão?

Solução: não há. Sugere-se cautela ao extrapolar análises agregadas para o nível individual.



5. Problemas Especiais com Dados Espaciais

5.2) Problema da Unidade de Área Modificável (MAUP)*: diferentes formas de agregação e/ou reordenamento podem induzir à diferentes resultados.

* *Modifiable Areal Unit Problem (MAUP).*

200	100	400	10	20	10
200	300	300	50	10	5
500	400	100	60	10	20

Base Population University degree (count)

a) Efeito Escala:

Nota: Associado à autocorrelação Espacial
(envolve agregação dos dados)

5 %	20 %	3 %	6 %
25 %	3 %	2 %	8 %
12 %	3 %	20 %	9 %

a - scale effect

b) Efeito Zoneamento:

Nota: Associado à heterogeneidade Espacial
(envolve reordenação dos dados)

10 %	4 %	
7 %	16 %	6 %
8 %		

b - zoning effect

Nota: corrigir dependência espacial atenua MAUP.



5. Problemas Especiais com Dados Espaciais

5.2. Um exemplo prático de MAUP:

A				B				C			
x=5	x=4	x=5	x=4	x=4.5	x=4.5	x=4.5	x=3.8	y=4.3	y=3.5		
y=3	y=5	y=6	y=2	y=4.0	y=4.0						
x=2	x=7	x=3	x=3	x=4.5	x=3.0	x=3.5	x=3.8	y=3.0	y=3.0		
y=3	y=6	y=2	y=4	y=4.5	y=3.0						
x=2	x=5	x=3	x=5	x=3.5	x=4.0	x=3.5	x=3.8	y=3.0	y=3.0		
y=1	y=5	y=4	y=3	y=3.0	y=3.5						
x=1	x=6	x=3	x=4	x=3.5	x=3.5	4.7	4.0	4.7	2.0		
y=2	y=4	y=1	y=4	y=3.0	y=2.5						
D				E				F			
3.5	5.5	4.0	3.5	2.5	5.5	3.5	4.0	3.6	4.0		
3.0	5.5	4.0	3.0							2.5	5.0
1.5	5.5	3.0	4.5								
1.5	4.5	2.5	3.5								

Estatísticas						
	n	\bar{x}	\bar{y}	S_x^2	S_y^2	r_{xy}
A	16	3.88	3.44	2.36	2.37	0.66
B	8	3.88	3.44	0.30	0.40	0.88
C	4	3.88	3.44	0.14	0.26	0.94
D	8	3.88	3.44	1.55	1.34	0.95
E	4	3.88	3.44	1.17	0.98	0.98
F	4	4.06	3.36	0.16	0.93	0.64
F_p	4	3.88	3.44	0.18	0.48	0.80

B e D (8x8) = efeito zoneamento
 C e E (4x4) = efeito zoneamento
 C e F (4x4) = efeito zoneamento
 B e C (8x4) = efeito escala
 D e E (8x4) = efeito escala
 D e F (8x4) = efeito escala

5. Problemas Especiais com Dados Espaciais

5.3) Efeito de Beirada (*edge effect*): ocorre quando a área analisada não capta toda a dependência espacial.

Nota: pode causar viés e perda de eficiência (compromete testes de hipóteses).

Ex: Uma análise do crescimento do PIB dos municípios de MG poderá gerar resultados equivocados caso não consideremos os vizinhos de outros estados.



Soluções para *Edge Effect*:

- Ampliar área de estudo;
- Estimar via MQG atribuindo menor peso aos municípios de fronteira.
- Modelar “efeito fronteira” via *dummies* (apenas para os municípios de fronteira).

5. Problemas Especiais com Dados Espaciais

5.4) Outliers Espaciais e Pontos de Alavancagem: são observações com valores extremos e/ou que apresentam dependência espacial distinta das demais.

Problema: podem enviesar os resultados associados ao restante da amostra.

▪ Origem dos outliers espaciais:

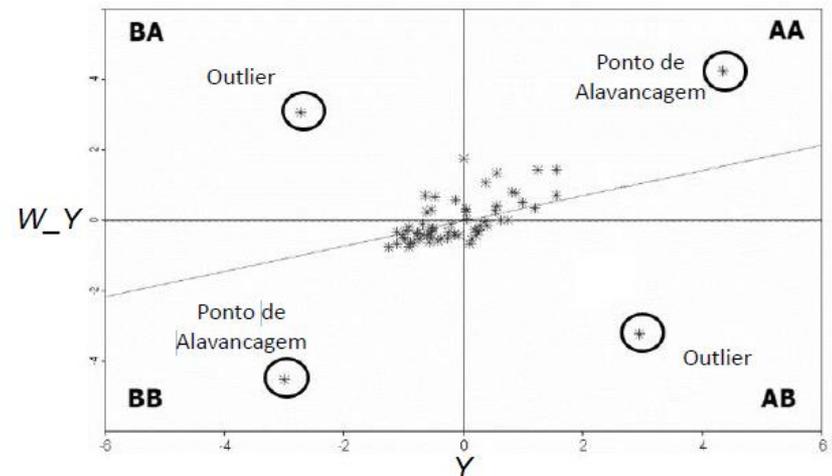
a) erros de medida ou de digitação

Solução: excluir da amostra.

b) Observações com características diferenciadas (merecem investigação!).

Solução: controlar *outliers* via *dummies*.

Outliers e Pontos de Alavancagem Espaciais



Ex. 1: Capitais podem exibir características bastante distintas das demais cidades.

Ex. 2: Condomínio fechado cercado por favelas (análise do crime).



UNIVERSIDADE FEDERAL DE JUIZ DE FORA - *CAMPUS* Governador Valadares
ECONOMETRIA ESPACIAL
Prof. Vinícius de Azevedo Couto Firme

Referência

1. ALMEIDA, E. *Econometria Espacial Aplicada*. 1ª ed. Alínea, 2012.