UM OLHAR CUIDADOSO SOBRE O MODELO ESPACIAL DE DECAIMENTO EXPONENCIAL

Erica Castilho Rodrigues[1,3], **Renato Assunção**[2,3]

Resumo: Um dos pontos cruciais em modelos espaciais para dados de área está na forma de modelar a estrutura de dependência. Modelos como, por exemplo, aqueles propostos por [2] e [4] têm sido utilizados em diversos tipos de aplicação. Alguns desses modelos, porém, apresentam aspectos não intuitivos, como aqueles apontados por [5]. O nosso objetivo nesse trabalho é identificar e analisar as causas de certos de certos resultados não razoáveis do modelo proposto por [4].

Palavras-chave: *Matriz Exponencial, Modelos Espaciais, Estrutura de Covariâncias.*

1 Introdução

Quando estamos lidando com dados de área temos, em geral, uma região $\mathbf{R}$ particionada em n áreas disjuntas $A_1, A_2, \ldots, A_n$. Tais que $\cup_{i=1}^{n} A_i = \mathbf{R}$. Os dados observados em cada uma dessas áreas são, tipicamente, somas ou médias de variáveis observadas em cada unidade. A fim de introduzir a dependência espacial entre elas, precisamos definir uma estrutura de vizinhança de acordo com a forma como essas áreas estão dispostas em toda a região. Uma vez definida a estrutura de vizinhança, podemos utilizá-la para definirmos modelos que reflitam a dependência espacial dos dados. Uma abordagem muito recorrente se baseia nas ideias dos modelos autoregressivos de séries temporais. Dois modelos muito populares, que incluem esse tipo de estrutura, são os modelos CAR (*conditional autoregressive*) e SAR (*simultaneous autoregressive*), que foram propostos originalmente por [2] e [6], respectivamente.

Devido ao formato de sua distribuição conjunta, a forma mais adequada de se estimar o modelo SAR é usando máxima verossimilhança. Porém, mesmo usando essa abordagem, o processo de estimação ainda não é simples, apesar da estrutra de covariância não apresentar grandes complexidades. Diante desse problema, [4] apresentam um modelo com objetivo de resolver problemas numéricos e tornar o processo de estimação mais eficiente. Eles definem a matriz de covariância de uma forma um pouco diferente. Essa nova matriz de covariância apresenta certas propriedades que simplificam bastante o processo de maximização. Apesar de resolver esse problema, o modelo proposto apresenta alguns aspectos não intuitivos. O ponto principal é que utilizando-se essa abordagem podemos ter correlações marginais e parciais entre pares de áreas que possuem sinais opostos.

O nosso objetivo então é entender melhor o modelo e apontar alguns de seus aspectos não intuitivos, tentando sempre explicitar qual o motivo do comportamento observado. O artigo está organizado da seguinte maneira. Na Seção 2 apresentamos o modelo proposto por [4]. Nas Seção 3 apontamos aspectos não intuitivos do modelo para o caso do *lattice* regular. Na Seção 4 consideramos o mesmo mapa utilizado pelos autores e os valores dos parâmetros por ele estimados. Na seção 5 apresentamos as conclusões e discussões sobre o trabalho.

[1]DEEst-Universidade Federal de Ouro Preto, ericaa_casti@yahoo.com.br

[2]DCC-Universidade Federal de Minas Gerais, assuncao@dcc.ufmg.br

[3]Agradecimento à FAPEMIG pelo apoio financeiro.

2 Definição do Modelo

A matriz de covariância do modelo SAR é dada por

$$\sigma^2(\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}-\mathbf{S})\,.$$

onde $\mathbf{S} = \rho\mathbf{D}$. Considerando-se a matriz $\mathbf{D}$ padronizada por linhas, temos que ρ está definido no intervalo $(-1, 1)$. Sob essas condições, é possível mostrar que a matriz $(\mathbf{I}-\rho\mathbf{D})^{-1}$ pode ser expressa como

$$(\mathbf{I}-\rho\mathbf{D})^{-1} = \mathbf{I} + \rho\mathbf{D} + \rho^2\mathbf{D}^2 + \rho^3\mathbf{D}^3 \ldots \tag{1}$$

Notamos, portanto, a partir de (1), que a influência de vizinhanças mais longínguas cai geometricamente nesse modelo.

A ideia de [4] é propor um modelo no qual essa influência de vizinhanças mais distantes cai de maneira mais rápida. Uma possibilidade seria impor um decaimento exponencial. Isso é feito substituindo, na versão original do modelo, a matriz $(\mathbf{I}-\rho\mathbf{D})^{-1}$ pela matriz

$$\mathbf{I} + \alpha\mathbf{D} + \frac{\alpha^2}{2!}\mathbf{D}^2 + \frac{\alpha^3}{3!}\mathbf{D}^3 \ldots$$

Essa matriz é denominada matriz exponencial e é denotada por $e^{\alpha\mathbf{D}}$. Portanto, no modelo proposto pelos autores, a matriz de covariâncias é definida como

$$\Sigma_\alpha = \sigma^2\left(e^{-\alpha\mathbf{D}'}e^{-\alpha\mathbf{D}}\right)$$

onde $\mathbf{D}$ é uma matriz de pesos espaciais não-negativa $n \times n$. E, tipicamente, defini-se $\mathbf{D}_{ij} > 0$ se i e j são vizinhas e $\mathbf{D}_{ij} = 0$ caso contrário.

3 Lattice Regular

Essa definição, porém, leva a alguns resultados não intuitivos. Veremos agora vários exemplos que mostram que essa definição pode induzir correlações marginais e condicionais com sinais opostos. Em outras palavras, podemos ter um par de áreas i e j cuja correlação marginal é positiva, porém quando condicionamos nas demais áreas do mapa, a correlação condicional passa a ser negativa. Esse tipo de problema já foi amplamente estudado quando estamos lidando com estrutura de dependência entre variáveis e é conhecido como "Paradoxo de Simpson" ([3]). Esse paradoxo diz que o sinal da correlação entre duas variáveis pode se modificar de acordo com os grupos que consideramos dentro da população. Ou seja, a correlação marginal, sem considerar o fator grupo, tem um sinal diferente da correlação obtida quando se condiciona no grupo ao qual o indivíduo pertence.

Esse tipo de comportamento não intuitivo para modelos espaciais foi apontado por [1] (veja página 169) que falam que para os modelos SAR e CAR *the transformation to (the covariance matrix)* Σ_Y *is very complicated and very non-linear. Positive conditional association can become negative unconditional association.* Ocorre porém, que para os modelos SAR e CAR esse tipo de comportamento é observado apenas em situações de pouco interesse prático, que são aquelas em que existe uma associação espacial negativa, ou seja, uma repulsão entre as áreas. Já para o modelo exponencial, esse comportamento é observado mesmo quando temos associação positiva entre as áreas.

Vamos considerar primeiramente uma situação simples, em que estamos em um *lattice* regular. Para esse tipo de estrutura é possível mostrar que para quaisquer pares de vizinhos de ordem ímpar, a correlação marginal será sempre positiva, porém quando condicionamos nas demais áreas do mapa essa correlação passa a ter valor negativo. Esse não é um resultado razoável para um modelo de dependência espacial, visto que uma correlação marginal positiva significa que se a área i apresenta um valor acima da média, a área j tenderá também a apresentar um valor acima do esperado. Porém, se consideramos conhecidos os valores de todas as áreas do mapa, o fato da área i ter um valor acima da média, induz a área j a ter um valor abaixo da média.

Para os vizinhos de ordem par observamos ainda outra restrição forte do modelo. Ele só permite que existam correlações marginais positivas entre vizinhos desse tipo. Ou seja, vizinhos de segunda, quarta, sexta ordem só poderão estar positivamente correlacionados quando utilizamos esse modelo, independentemente do valor de α.

Observe que, apesar do caso do *lattice* regular ser um caso muito específico de estrutura de vizinhança, ele está presente em muita situações práticas, como, por exemplo, no processamento de imagens.

4 Exemplo de LeSage e Peace

Como um terceiro exemplo no qual esse comportamento não intuitivo é observado, vamos considerar o mapa utilizado na aplicação apresentada por [4]. Os autores analisam a participação dos americanos nas eleições presidenciais no ano de 2000. Os dados são agregados a nível de condado e todo o estado do Texas é excluído da análise devido a problemas de irregularidades nas eleições. Os autores consideram dois tipos de vizinhança: adjacência e k vizinhos mais próximos. Por motivos de simplificação, consideraremos aqui apenas o critério de adjacência. O valor estimado para α na aplicação foi de -0.74 e esse será o valor utilizado aqui.

Observamos nesse caso que todos vizinhos de primeira ordem possuem correlação marginal positiva. Isso já era esperado pois para $\alpha < 0$ todos os termos da matriz de covariância são positivos. Esse mesmo comportamento é o observado para a correlação parcial, ou seja, a correlação parcial entre vizinhos de primeira ordem também é positiva, assim como a marginal. Isso se repete para quaisquer pares de vizinhos de ordem ímpar. Por outro lado, quando consideramos os vizinhos de segunda ordem existe a troca de sinal entre as correlações marginais e parciais. Esse comportamento se repete, em geral, para vizinhos de ordem par.

Notamos então que as correlações para esse mapa se comportam da mesma maneira que o gráfico regular. Analisando-se outros mapas, com configurações completamente distintas, para esse valor de α, o mesmo comportamento foi observado. É possível mostrar que, de fato, se α não assume valores grandes, esse tipo de comportamento será observado independente da disposição espacial das áreas.

Pode-se mostrar ainda que para determinados tipos de mapa e determinadas estruturas de vizinhança existe a troca de sinal mesmo para vizinhos de primeira ordem.

4.1 Como a correlação parcial varia com α

Iremos verificar agora como a correlação parcial entre as áreas se comporta se variarmos o valor de α. Pode-se mostrar que à medida que ρ aumenta, α diminui. Tendo isso em vista, seria razoável que as correlações marginais e condicionais tivessem um comportamento decrescente com relação ao parâmetro α. Veremos aqui que para o caso das correlações condicionais esse fato não é sempre verdade.

Considerando novamente o mesmo mapa utilizado por [4]. Percebe-se que para vizinhos de primeira ordem a correlação condicional decresce com o valor de α. E, como já explicado anteriormente, esse é um comportamento razoável, pois significa que à medida que a associação espacial entre as áreas fica mais forte, a correlação condicional entre determinados pares de áreas irá crescer.

Por outro lado, para os vizinhos de segunda ordem, a correlação condicional cresce com o valor de α. Percebe-se que nesse caso, à medida que o valor de α diminui, associação espacial fica mais forte, a correlação parcial entre vizinhos de segunda ordem aumenta, porém ela tem sinal negativo. Esse não é um comportamento razoável, pois o que seria esperado era que à medida que a associação espacial aumentasse a correlação parcial entre as áreas deveria aumentar, porém no sentido positivo e não negativo, como ocorre. É possível mostrar que esse tipo de comportamento generaliza-se para qualquer tipo de estrutura espacial, desde que o valor de α não seja muito grande.

5 Conclusão

Mostramos nesse trabalho que o modelo proposto por [4], apesar de apresentar vantagens, principalmente no que se refere à eficiência computacional, apresenta alguns aspectos pouco intuitivos. Esse tipo de comportamento pode ser observado para outros modelo amplamente utilizados, como por exemplo o modelos CAR. Porém no caso desse último modelo isso só ocorre quando o parâmetro de dependência espacial tem sinal negativo, ou seja, um caso com pouca aplicabilidade em problemas reais. Já para o modelo de [4], esse comportamento é observado mesmo em situações de grande interesse prático. Portanto, a aplicação desse modelo deve ser feita com cautela, visto que os resultados apresentados podem estar refletindo um tipo de dependência espacial que não possui menor sentido prático.

Referências

[1] Banerjee, B., Carling, B., Gelfand, A., Hierarchical Modeling and Analysis for Spatial Data, *Chapman and Hall/CRC*, 2003.

[2] Bessag, J., Spatial Interaction and the Statistical Analysis of Lattice Systems, *Journal of the Royal Statistical Society. Series B (Methodological)*, **36**, 192-236, 1974.

[3] Blyth, C. R., On Simpson's paradox and the sure-thing principle, *Journal of the American Statistical Association*, **67**, 364-366, 1972.

[4] Lesage, J. P., Pace, K., A matrix exponential spatial specification, *Journal of Econometrics*, **140**, 190-214, 2007.

[5] Wall, M. M., A close look at the spatial structure implied by the CAR and SAR models, *Journal of Statistical Planning and Inference*, **121**, 311-324, 2004.

[6] Whitle, P., On stationary process in the plane, *Biometrika*, **41**, 434-449, 1954.

[7] Bussab, W. O., *Hierarchical dichotomous partitions in cluster analysis*. 1976. 236 p., PhD Thesis, University of London, Londres, 1976.

[8] Bussab, W. O., Barroso, L. P., Knott, M., Best Linear Unbiased Predictor With Incomplete Data, *Communications in Statistics. Theory and Methods*, **21**, 121-129, 1998.

[9] Bussab, W. O., Morettin, P. A., *Estatística Básica*. 6^a ed. São Paulo: Saraiva, 2009.