

## Directional conditionally autoregressive models

Minjung Kyung<sup>a,1</sup>

<sup>a</sup>Department of Statistics, Duksung Women's University

(Received April 26, 2016; Revised July 3, 2016; Accepted July 3, 2016)

---

### Abstract

To analyze lattice or areal data, a conditionally autoregressive (CAR) model has been widely used in the field of spatial analysis. The spatial neighborhoods within CAR model are generally formed using only inter-distance or boundaries between regions. Kyung and Ghosh (2010) proposed a new class of models to accommodate spatial variations that may depend on directions. The proposed model, a directional conditionally autoregressive (DCAR) model, generalized the usual CAR model by accounting for spatial anisotropy. Properties of maximum likelihood estimators of a Gaussian DCAR are discussed. The method is illustrated using a data set of median property prices across Greater Glasgow, Scotland, in 2008.

Keywords: anisotropy, lattice data, spatial analysis, conditionally autoregressive models, maximum likelihood estimation

---

### 1. 서론

격자 또는 지역자료 (Cressie, 1993)로 알려진 지역에 대한 어떤 집합체 자료나 평균자료들이 다양한 분야의 연구에서 관측되고 있으며 공간통계 방법을 사용한 분석이 사용되고 있다. 지리학적 지역 집합이 주어졌을 때, 가까운 지역에서 수집된 관측치들은 거리상 먼 지역 자료의 관측치에 비해 서로 비슷한 특성을 띠는 경향을 보이게 된다. 지리학에서 이러한 특성은 ‘토블러의 첫 번째 법칙’으로 알려져 있다 (Miller, 2004). 통계학적인 관점에서 볼 때 이러한 특성은 거리상 가까운 지역들로부터 수집된 관측치들 사이의 자기상관관계가 먼 거리의 관측치들과의 자기상관관계보다 높기 때문에 나타나는 결과로 보여진다. 따라서 격자나 또는 불규칙한 공간으로부터 관측된 공간과정은 공간적 자기회귀 모형으로 분석된다.

관측된 격자 또는 불규칙한 공간의 자료들이 분포되어있는 공간  $\mathbf{S}$ 를 고려한다. 이 지역은 소지역들,  $S_1, \dots, S_n$ 으로 나뉘게 되고  $\mathbf{S} = \cup_{i=1}^n S_i$ ,  $S_i \cap S_j = \emptyset$ 으로 나타낼 수 있다.  $Y_i = Y(S_i)$ 를 지역  $S_i$ 로부터 수집된 반응치(자료)로 정의한다. 일반적으로 주어진 공간에서 우리는 집합체 자료에 대해 다음과 같은 일반화 선형모형을 고려할 수 있다.

$$E[\mathbf{Y}|\mathbf{Z}] = g(\mathbf{Z}) \quad \text{and} \quad \mathbf{Z} = \boldsymbol{\mu} + \boldsymbol{\eta}, \quad (1.1)$$

---

This work was supported by the Duksung Women's University research grants (3000002407).

<sup>1</sup>Department of Statistics, Duksung Women's University, 33, 144-gil, Samyang-ro, Dobong-Gu, Seoul 01369, Korea. E-mail: mkyung@duksung.ac.kr

여기서  $g(\cdot)$ 은 연결함수이고,  $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_n) = (\mathbf{Y}(S_1), \dots, \mathbf{Y}(S_n))$ 는 반응벡터이며  $\mathbf{Z} = (Z_1, \dots, Z_n) = (Z(S_1), \dots, Z(S_n))$ 는 공간 잠재 과정이다. 그리고  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n) = (\mu(S_1), \dots, \mu(S_n))$ 는 연구 공간의 추세와 경향성을 설명하는 고정효과인 대규모 변화이며,  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n) = (\eta(S_1), \dots, \eta(S_n))$ 는 평균이  $\mathbf{0}$ 이고 분산이  $\boldsymbol{\Sigma}$ 인 지역적 자기상관성을 내포한 공간적 임의 효과인 소규모 변화이다.

각 소지역들  $S_i$ 에 정의되는 대규모 변화값  $\mu_i$ 는 모수  $\beta$ 와 설명변수들의 모수적 또는 비모수적 회귀함수로 모형화 한다 (van der Linde 등, 1995). 반면 공간적 임의 효과  $\eta_i$ 에 대하여 적절한 모형을 정의하는 것은  $\eta_i$ 들의 공간적 상관성 때문에 쉬운 과정은 아니다. 이러한 공간적 공분산의 추정은 적당한 모수 형태의 선택에 따라 달라질 수 있다. 즉,  $n \times n$  공분산 행렬  $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\omega})$ 은 유한차원의 모수  $\boldsymbol{\omega}$ 의 결정적 함수로 가정한다. 그러므로 어떠한 표본 수  $n$ 과 데이터로부터 추정하는 모든 수용 가능한 모수  $\boldsymbol{\omega}$ 에 대해서도 양정치 행렬의 특성을 만족해야 한다.

지역자료 분석에 사용하는 공간적 공분산행렬을 위해 두가지의 다른 방법론을 고려할 수 있다. 우선 적절한 집합체적 공간통계적인 모형을 직접적으로 사용하여  $\boldsymbol{\Sigma}$  행렬의 원소들의 결정적 함수를 구체화하는 것이다. 반대로 조건부 공간자기회귀모형은 공분산행렬의 역행렬  $\boldsymbol{\Sigma}^{-1}(\boldsymbol{\omega})$ 의 원소들의 결정적 함수를 사용한다 (Besag, 1974). Griffith와 Csillag (1993), Rue와 Tjelmeland (2002), Hrafinkelsson과 Cressie (2003) 등이 앞서 제시한 두 방법론의 연결 가능성을 설명하였다. Song 등 (2008)은 스펙트럼 밀도들을 사용하여 이러한 가우시안 공간통계적인 모형들이 가우시안 마르코프 임의장(Gaussian Markov random fields; GMRFs)으로 표현될 수 있는 것을 증명하였다. 그렇지만 지금까지 다양한 연구에서 사용된 GMRFs의 대부분은 지역적 자료의 특징 중 하나인 이방성(anisotropy)을 고려하지 않았다. Lindgren 등 (2011)은 근사 확률적 약수렴 방법을 선형 확률적 편미분방정식에 사용하여 Matérn 계급에서 기저함수 표현으로 가우시안 임의장과 가우시안 마르코프 임의장의 명확한 연결함수를 제시할 수 있음을 증명하였다. 그리고 다양한 공간적 설정에서 생긴 역공분산 행렬이나 정밀도 행렬에 Matérn 형태의 가우시안 모형의 구성을 제시하기 위해 유한 원소 방법들이 사용된다.

실제로 통계방법론을 사용하는 사람들은 변수들간의 관계를 설명하는데 선형성과 분류모형들을 사용하고, 선형모형과 처리 효과의 가설들을 검정하고 의미있는 대비관계 설정등을 활용한다 (Schabenberger와 Gotway, 2005). 그러나 공간적 선형모형은 소지역들간의 상관성을 가정하고 각 소지역  $S_i$ 가 근접이웃지역들에 어떠한 영향을 받는지 설명한다 (Cliff와 Ord, 1981). 그러므로 우리는 이러한 지역집합체 자료들을 분석하기 위해 일반화된 선형 혼합 모형을 고려한다. 이러한 모형에서 잠재적 공간과정  $Z_i$ 는 임의 효과로 적용되며, 이러한 공간적 임의 효과 모형을 위해 조건부 자기회귀(conditionally autoregressive; CAR) 모형 (Besag, 1974, 1975; Cressie와 Chan, 1989)과 동시적 자기회귀(simultaneously autoregressive; SAR) 모형이 널리 사용되어왔다.

가우스 조건부 자기회귀 모형은 일반화된 혼합효과 모형에서 임의 효과로 사용된다 (Breslow와 Clayton, 1993). 왜냐하면, 가우스 CAR 모형은 상당히 일반적인 규칙조건에서 낮은 차원의 조건부 가우스 분포가 공간적 CAR 과정의 결합 가우스 분포를 이룬다는 특징을 가지고 있기 때문이다. 그러므로 추정방법으로 최대우도법과 베이저안 방법이 쉽게 적용될 수 있다. 그렇지만 CAR 모형의 주요 제한 요건 중 하나는 이웃 지역들이 몇몇 거리 측정 방법들로 정의되고 방향의 효과는 완전히 무시된다는 점이다. 최근들어 연구 지역안에서 다른 부분의 지역에 서로 다른 CAR 모형을 적용시키는 방법이 사용되고 있다. 예를들어 Reich 등 (2007)은 치주질환 분석법으로 새로운 모형을 제시하면서 각각의 턱에 별도의 CAR 모형을 적용하였다. White와 Ghosh (2008)는 이웃지역의 효과를 설명하기 위해 CAR 구조 내에서 확률적 모수를 사용하였다. 그럼에도 불구하고, 내제된 공간과정이 이방성의 특성을 갖고 있다면 다른 방향의 이웃간 자기상관성의 크기가 다를 수 있다. 이러한 모형적 제한점을 고려하여 Kyung과 Ghosh (2010)는 고유 이방성을 모형화하여 기존의 CAR 과정을 일반화 시킨 방향적 조건부 자기회

귀(directional conditionally autoregressive) 모형을 제시하였다. 본 논문에서는 공간 임의 효과  $\eta_i$ 에 적용할 유연한 모형인 DCAR 모형에 대해서 살펴보고, 실제 데이터에 적용하여 활용성을 연구한다.

본 논문의 구성은 다음과 같다. 2절에서는 방향적 자기회귀 모형에 대해서 설명하고, 추정방법으로 최대우도 추정법을 적용하며, 최대우도 추정량의 특성에 대해서 살펴본다. 3절에서는 데이터 분석을 통해 방향적 자기회귀 모형을 실제 데이터에 적용하며, 4절에서는 요약과 결론으로 끝을 맺는다.

## 2. 방향적 자기회귀 모형(DCAR)

이 절에서는 식 (1.1)에서 정의한 내제된 공간과정  $Z_i$ 의 공간적 고유 이방성을 모형화하여 기존의 CAR 과정을 일반화한 방향성 자기회귀 모형을 살펴보고, 최대우도 추정법과 최대우도 추정량의 특성에 대해서 논의한다. 이해와 표기의 편의를 위해 소지역  $S_i$ 는 이차원 공간에 정의되었다 가정한다. 즉,  $S_i \in \mathbb{R}^2, \forall i$ . 그러나 이 논문의 모형과 관련된 통계적 추론 방법은 고차원 데이터에 적용으로의 확장이 용이하다. 소지역  $S_i$ 의 중심의 좌표를  $\mathbf{s}_i = (s_{1i}, s_{2i})$ 라 할 때,  $s_{1i}$ 는 수평좌표 ( $x$ -좌표)이고  $s_{2i}$ 는 수직좌표 ( $y$ -좌표)이다. 각 소지역  $S_i$ 와  $S_j$  ( $i \neq j$ ) 사이의 각도를  $\alpha_{ij}$ 라 하고, 다음과 같이 정의한다.

$$\alpha_{ij} = \alpha(S_i, S_j) = \begin{cases} \left| \tan^{-1} \left( \frac{s_{2j} - s_{2i}}{s_{1j} - s_{1i}} \right) \right|, & \text{if } s_{2j} - s_{2i} \geq 0, \\ - \left( \pi - \left| \tan^{-1} \left( \frac{s_{2j} - s_{2i}}{s_{1j} - s_{1i}} \right) \right| \right), & \text{if } s_{2j} - s_{2i} < 0. \end{cases}$$

이웃지역의 방향을 설정할때 각 소지역의 중점을 사용한다.  $\mathcal{N}_i$ 를 거리측정의 한 형태를 바탕으로 정의한 소지역  $S_i$ 의 이웃지역의 인덱스의 집합이라 하자 ( $i = 1, \dots, n$ ). 일반적으로 각각의 소지역들의 이웃지역들은 파별(clique)을 구성해야만 한다. 파별은 어떤 지역들의 집합으로 하나 혹은 그 이상의 지역으로 각각 구성되어 있으며 그 집합 내에서 모든 지역은 각각 다른 지역들과 이웃관계에 있다 (Besag, 1974). 그러므로 방향성을 가진 이웃지역의 인덱스의 집합을 정의할 때 파별이 형성되는 조건을 만족해야 한다.

각 소지역에 대해서 파별이 형성되는 방향성 이웃지역의 인덱스 집합을 잘 정의하는 것은 Hammersley-Clifford 정리 (Besag, 1974, 197–198)를 활용할 수 있는 공간과정의 확장임을 쉽게 증명할 수 있다. 그리고 전체 조건부 분포의 집합을 사용하여 공간과정의 유한 차원 결합분포를 쉽게 유도할 수 있다. 그러므로 방향성을 고려한  $\mathcal{N}_i$ 의 두 부분집합을 다음과 같이 정의한다.

$$\begin{aligned} \mathcal{N}_{i1}^* &= \left\{ j : j \in \mathcal{N}_i \text{ and } \left( 0 \leq \alpha_{ij} < \frac{\pi}{2} \text{ or } \pi \leq \alpha_{ij} < \frac{3}{2}\pi \right) \right\}, \\ \mathcal{N}_{i2}^* &= \left\{ j : j \in \mathcal{N}_i \text{ and } \left( \frac{\pi}{2} \leq \alpha_{ij} < \pi \text{ or } \frac{3}{2}\pi \leq \alpha_{ij} < 2\pi \right) \right\}. \end{aligned} \quad (2.1)$$

각  $\mathcal{N}_{i1}^*$ 와  $\mathcal{N}_{i2}^*$ 는 파별을 형성하며  $\mathcal{N}_i = \mathcal{N}_{i1}^* \cup \mathcal{N}_{i2}^*$ 가 되는 특성이 있다.  $\mathcal{N}_{i1}^*$ 은 북동남서 방향의 이웃지역의 인덱스 집합이고,  $\mathcal{N}_{i2}^*$ 는 북서남동 방향의 이웃지역의 인덱스 집합이다. 이러한 이웃지역의 인덱스 집합의 부분집합은 파별을 형성하는 더 많은 부분집합으로 나눌 수 있는데 본 연구에서는 식 (2.1)에 정의된 두 부분집합을 사용한다. 그리고 소지역  $S_i$ 의 중점이 주어지지 않거나 중점을 정의하기 불가능한 경우가 있다. 예를들어 이웃지역이 중점의 거리가 아닌 근접성으로 정의된 경우이다. 이 경우 방향성을 지닌 소그룹으로 인덱스 집합을 분리하는 것이 쉽지 않다. 한가지 제안할 수 있는 방법은 근접성의 방향을 보고 직관적으로 정의하는 것이다. 그러나 본 연구에서는 각 소그룹의 이웃지역 인덱스 집합을 방향성을 지닌 부분집합으로 분리할 수 있다고 가정한다.

방향적 관계성을 지닌 이웃지역의 인덱스의 집합인  $\mathcal{N}_{i1}^*$ 와  $\mathcal{N}_{i2}^*$ 을 기반으로 방향성 무게 행렬(directional weight matrices)인  $\mathbf{W}^{(1)} = ((w_{ij}^{(1)}))$ 와  $\mathbf{W}^{(2)} = ((w_{ij}^{(2)}))$ 를 정의한다. 예를들어 만일 소지역  $S_j$ 에 대하여  $j \in \mathcal{N}_{i1}^*$ 이면  $w_{ij}^{(1)} = 1$ 이고  $j \in \mathcal{N}_{i2}^*$ 이며  $w_{ij}^{(2)} = 1$ 로 방향 근접 행렬을 정의한다. 그러면  $\mathbf{W} = \mathbf{W}^{(1)} + \mathbf{W}^{(2)}$ 가 성립되고 이것은 일반적인 CAR 모형에서 거리를 기준으로 정의한 기존의 근접 행렬이 된다.

연구 공간의 추세와 경향성을 설명하는 고정효과인 대규모 변화를 모형화하기 위하여 어떤 일반성의 손실없이 정준적 일반화된 선형 모형  $\mu_i = \mathbf{x}_i^T \boldsymbol{\beta}$ 을 가정한다. 여기서  $\mathbf{x}_i$ 는 각 소지역  $S_i$ 에서 관측된 예측 변수들의 벡터이고,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_q)^T$ 는 회귀계수의 벡터이다. 평활 스플라인과 다항식이 사용된 비모수적 회귀 함수도 정준적 선형모형으로 표현될 수 있다 (Wahba, 1977; van der Linde 등, 1995). 식 (1.1)에서 정의한 모형으로부터

$$E[\mathbf{Z}] = \mathbf{X}\boldsymbol{\beta} \quad \text{and} \quad \text{Var}[\mathbf{Z}] = \boldsymbol{\Sigma}(\boldsymbol{\omega}) \quad (2.2)$$

가 성립함을 알 수있고, 여기서  $\boldsymbol{\omega}$ 는 공간적 자기상관 모수와 다른 분산 성분들의 벡터를 나타낸다. 식의 결과를 함께, 식 (1.1)의 모형은 이산응답자료를 위한 일반화된 선형 모형의 틀에서도 사용가능하다 (Schabenberger와 Gotway, 2005, p.353). 본 연구에서는 분산행렬  $\boldsymbol{\Sigma}_s(\boldsymbol{\omega})$ 에 대해 이방성을 고려한 모형을 살펴본다.

$\delta_1$ 과  $\delta_2$ 를 각각  $\mathcal{N}_{i1}^*$ 와  $\mathcal{N}_{i2}^*$ 에 대응하는 방향적 공간 효과로 정의한다.  $j \neq i$ 인 모든  $Z_j$ 가 주어졌을 때, 공간과정  $Z_i$ 의 조건부 분포에 대해 처음 두 적분을 사용하여 평균과 분산을 정의하면

$$\begin{aligned} E[Z_i | Z_j = z_j, j \neq i, \mathbf{x}_i] &= \mathbf{x}_i^T \boldsymbol{\beta} + \sum_{k=1}^2 \delta_k \sum_{j=1}^n w_{ij}^{(k)} (z_j - \mathbf{x}_j^T \boldsymbol{\beta}) \\ \text{Var}[Z_i | Z_j = z_j, j \neq i, \mathbf{x}_i] &= \frac{\sigma^2}{m_i} \end{aligned} \quad (2.3)$$

이고  $k = 1, 2$ 에 대하여  $w_{ij}^{(k)} \geq 0$ 이고  $w_{ii}^{(k)} = 0$ 이며  $m_i = \sum_{j=1}^n w_{ij}$ 이다.

본 연구의 표기법으로 일반적인  $\text{CAR}(\rho, \sigma^2)$ 를 적용한 공간과정  $Z_i$ 의 평균과 분산을 표기하면

$$\begin{aligned} E[Z_i | Z_j = z_j, j \neq i, \mathbf{x}_i] &= \mathbf{x}_i^T \boldsymbol{\beta} + \rho \sum_{j=1}^n w_{ij} (z_j - \mathbf{x}_j^T \boldsymbol{\beta}), \\ \text{Var}[Z_i | Z_j = z_j, j \neq i, \mathbf{x}_i] &= \frac{\sigma^2}{m_i} \end{aligned}$$

와 같고  $w_{ij} = w_{ij}^{(1)} + w_{ij}^{(2)}$ 이다. 그러므로  $\delta_1 = \delta_2 = \rho$ 일때  $\text{DCAR}(\delta_1, \delta_2, \sigma^2)$ 는  $\text{CAR}(\rho, \sigma^2)$ 로 축소되어, 일반적인 CAR 모형이 DCAR 모형에 중첩됨을 알 수 있다.

완전 조건부 집단이 주어진 것을 근거로 한 결합 분포를 이끌어내기 위해 Brook의 보조정리 (Brook, 1964)를 이용한다. 이 정리는 양의 조건을 만족시키는 완전 조건부로부터 유일한 결합분포를 구조적으로 다시 되돌릴 수 있게 해준다 (Besag, 1974; Besag와 Kooperberg, 1995). DCAR 모형에서 식 (2.1)에서 정의된  $\mathcal{N}_{i1}^*$ 와  $\mathcal{N}_{i2}^*$ 는 각각 파벌을 형성하므로 Hammersley-Clifford 정리에 근거하여, DCAR 모형의 내재된 공간과정  $Z_i$ 은 존재하며 유효한 마르코프 랜덤 필드를 구성한다. 그러므로, 각각의 완전 조건부 분포를 가우스 분포로 가정하면, DCAR 과정  $Z_i$ 의 정확한 결합 분포를 쉽게 유도할 수 있다.

## 2.1. 가우스 DCAR 모형

내재된 공간과정  $Z_i$ 의 모형화를 위하여 가우스 조건부 자기회귀 모형이 널리 사용되어 왔다. 이 절에서

는, Brook의 보조정리를 이용하여 이끌어낸 CAR의 확장모형인 방향적 조건부 자기회귀 과정의 결합분포를 살펴본다.

공간과정  $Z_j$   $j \neq i$ 가 주어졌을 때,  $Z_i$ 의 완전 조건부 분포가 다음과 같다고 가정한다.

$$Z_i | Z_j = z_j, \quad j \neq i, \quad \mathbf{x}_i \sim N \left( \mathbf{x}_i^T \boldsymbol{\beta} + \sum_{k=1}^2 \delta_k \sum_{j=1}^n w_{ij}^{(k)} (z_j - \mathbf{x}_j^T \boldsymbol{\beta}), \frac{\sigma^2}{m_i} \right), \quad (2.4)$$

여기서  $w_{ij}^{(k)}$ ,  $k = 1, 2$ 는 방향성 가중치이다. 내재된 공간 DCAR 과정은 마르코프 랜덤 필드임은 쉽게 증명할 수 있다. 그러므로 Brook의 보조정리와 Hammersley-Clifford 정리에 근거하여 유한차원 결합분포는 다음과 같은 다변량 가우스 분포가 된다.

$$\mathbf{Z} \sim N_n \left( \mathbf{X}\boldsymbol{\beta}, \sigma^2 \left( \mathbf{I} - \delta_1 \mathbf{W}^{(1)} - \delta_2 \mathbf{W}^{(2)} \right)^{-1} \mathbf{D} \right),$$

여기서  $\mathbf{Z} = (Z_1, \dots, Z_n)^T$ 이고  $\mathbf{D} = \text{diag}(1/m_1, \dots, 1/m_n)$ 이다. 쉬운 표기를 위해 DCAR 과정의 분산공분산 행렬을  $\boldsymbol{\Sigma}_Z^* \equiv \sigma^2 (\mathbf{I} - \delta_1 \mathbf{W}^{(1)} - \delta_2 \mathbf{W}^{(2)})^{-1} \mathbf{D}$ 와 같이 표기한다.

적합한 가우스 모형이 되기 위하여, 공분산 행렬  $\boldsymbol{\Sigma}_Z^*$ 은 양정치성을 만족해야 한다. 이분적 방향성 가중치 행렬 대신 표준화된 방향성 근접행렬  $\tilde{\mathbf{W}}^{(k)} = (\tilde{w}_{ij}^{(k)} = w_{ij}^{(k)}/m_i)$ ,  $k = 1, 2$ 를 사용하면, 유한차원 결합분포의 분산은

$$\boldsymbol{\Sigma}_Z \equiv \sigma^2 \left( \mathbf{I} - \delta_1 \tilde{\mathbf{W}}^{(1)} - \delta_2 \tilde{\mathbf{W}}^{(2)} \right)^{-1} \mathbf{D} \quad (2.5)$$

이다. Kyung과 Ghosh (2010)는 식 (2.5)에서 정의한 분산공분산행렬  $\boldsymbol{\Sigma}_Z$  대칭성은 만족하는 것을 설명하였고,  $\max(|\delta_1|, |\delta_2|) < 1$ 일때 행렬의 성질을 이용하여 양정치성을 만족함을 증명하였다. 그러므로 방향적 조건부 자기회귀 과정의 결합분포로 다음을 사용한다.

$$\mathbf{Z} \sim N_n \left( \mathbf{X}\boldsymbol{\beta}, \sigma^2 \left( \mathbf{I} - \delta_1 \tilde{\mathbf{W}}^{(1)} - \delta_2 \tilde{\mathbf{W}}^{(2)} \right)^{-1} \mathbf{D} \right). \quad (2.6)$$

다음으로 관측된 반응값과 설명변수를 바탕으로 DCAR 모형의 모수를 추정하는 방법을 살펴본다. 베이지안 추정법도 가능하지만 본 연구에서는 최대우도 추정법을 살펴보기로 한다.

## 2.2. 최대우도 추정법

이 절에서는 내재된 공간과정  $Z_i$ 의 가우스 DCAR 모형을 중심으로 최대우도 추정법에 근거한 모수 추정법과 불확실성에 근거한 관련측도에 대해 살펴본다.

자기회귀 모형의 경우, 우도는 공간상관 모수의 복잡한 비선형적 함수가 그러므로 가우스 DCAR 모형의 로그우도함수는 공분산행렬에 포함된  $\delta_1$ 과  $\delta_2$ 의 복잡한 비선형적 함수가 된다. 표기의 편의를 위하여  $\boldsymbol{\delta} = (\delta_1, \delta_2)$ 와  $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma^2, \boldsymbol{\delta}^T)$ 로 정의한다. 내재된 공간 DCAR 과정  $Z_i$ 의 결합 다변량 가우스 분포 (2.6)로 부터 로그우도 함수는 다음과 같은 형태를 갖는다.

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2} \ln |\mathbf{A}^*(\boldsymbol{\delta})^{-1} \mathbf{D}| - \frac{1}{2\sigma^2} (\mathbf{Z} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{D}^{-1} \mathbf{A}^*(\boldsymbol{\delta}) (\mathbf{Z} - \mathbf{X}\boldsymbol{\beta}), \quad (2.7)$$

여기서  $\mathbf{A}^*(\boldsymbol{\delta}) = \mathbf{I} - \delta_1 \tilde{\mathbf{W}}^{(1)} - \delta_2 \tilde{\mathbf{W}}^{(2)}$ 이고  $\mathbf{D} = \text{diag}(1/m_1, \dots, 1/m_n)$ 이다.  $\boldsymbol{\delta}$ 가 주어졌을때,  $\boldsymbol{\beta}$ 와  $\sigma^2$ 의 프로파일 최대우도 추정량은

$$\begin{aligned} \widehat{\boldsymbol{\beta}}(\boldsymbol{\delta}) &= \left( \mathbf{X}^T \mathbf{D}^{-1} \mathbf{A}^*(\boldsymbol{\delta}) \mathbf{X} \right)^{-1} \mathbf{X}^T \mathbf{D}^{-1} \mathbf{A}^*(\boldsymbol{\delta}) \mathbf{Z}, \\ \widehat{\sigma^2}(\boldsymbol{\delta}) &= n^{-1} \left( \mathbf{Z} - \mathbf{X}\widehat{\boldsymbol{\beta}}(\boldsymbol{\delta}) \right)^T \mathbf{D}^{-1} \mathbf{A}^*(\boldsymbol{\delta}) \left( \mathbf{Z} - \mathbf{X}\widehat{\boldsymbol{\beta}}(\boldsymbol{\delta}) \right) \end{aligned}$$

가 된다. 프로파일 최대우도 추정량을 다시 식 (2.7)에 대입하면  $\delta$ 의 최대우도 추정량은 프로파일 로그 우도를 최대화하여 구할 수 있다.  $\delta$ 의 프로파일 로그 우도는

$$\ell^*(\delta) = -\frac{n}{2} \ln(\widehat{\sigma^2}(\delta)) + \frac{1}{2} \ln |\mathbf{A}^*(\delta)| \quad (2.8)$$

이고 각각의  $\hat{\delta}_1$ 과  $\hat{\delta}_2$ 는 수치적 최대화과정을 통해서 구할 수 있다.

제약조건  $\max(|\delta_1|, |\delta_2|) < 1$ 을 만족하면서 로그우도함수를 최대화시키는 최대우도 추정량을 구하기 위하여 “L-BFGS-B” (Byrd 등, 1995)를 적용한다. 정보행렬을 추정하기 위하여 헤시안 행렬을 추출하여 사용할 수 있다. 즉, 모수들  $\theta$ 의 표준오차를 구하기 위하여 관측된 피셔 정보량을 사용할 수 있는 것이다. Kyung과 Ghosh (2010)는 Cressie (1993)의 과정을 적용하여 가우스 DCAR 모형 모수들의 피셔 정보행렬을 다음과 같이 구하였다.

$$\mathbf{I}(\theta) = \begin{bmatrix} \sigma^{-2} \mathbf{X}^T \mathbf{D}^{-1} \mathbf{A}^*(\delta) \mathbf{X} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}^T & \frac{n}{2} \sigma^{-4} & \frac{1}{2} \sigma^{-2} \text{tr}(\mathbf{G}_1) & \frac{1}{2} \sigma^{-2} \text{tr}(\mathbf{G}_2) \\ \mathbf{0}^T & \frac{1}{2} \sigma^{-2} \text{tr}(\mathbf{G}_1) & \frac{1}{2} \nu_1 & \frac{1}{2} \text{tr}(\mathbf{G}_1 \mathbf{G}_2) \\ \mathbf{0}^T & \frac{1}{2} \sigma^{-2} \text{tr}(\mathbf{G}_2) & \frac{1}{2} \text{tr}(\mathbf{G}_1 \mathbf{G}_2) & \frac{1}{2} \nu_2 \end{bmatrix}, \quad (2.9)$$

여기서  $\mathbf{G}_k = \tilde{\mathbf{W}}^{(k)} \mathbf{A}^*(\delta)^{-1}$ 이고  $\nu_k = \text{tr}(\mathbf{G}_k \mathbf{G}_k) = \text{tr}(\mathbf{G}_k^2)$ ,  $k = 1, 2$ 이다. 그러므로 추정된 모수들의 점근분산을 구하기 위하여 식 (2.9)에 주어진 피셔 정보행렬을 추정한다.

Sweeting (1980)은 중속된 관측치들을 바탕으로 최대우도추정량의 약일치성과 균일한 점근적 정규성의 일반적인 결과는 연구하였다. 이러한 Sweeting의 결과를 사용하여, Mardia와 Marshall (1984)는 가우스 CAR 모형 모수의 최대우도 추정량의 약일치성과 균일한 점근적 정규성의 충분조건을 증명하였다. Kyung과 Ghosh (2010)은 이러한 결과들을 바탕으로 이웃지역의 수가 표본크기  $n$ 의 증가속도보다 빠르지 않으면, 가우스 DCAR 모형 모수  $\theta$ 의 최대우도 추정량  $\hat{\theta}$ 은 점근적 정규성을 보이며과 약일치성을 만족하는 것을 증명하였다.

### 3. 자료분석

이 절에서는 공간 DCAR 모형을 스코틀랜드의 그레이터 글래스고우(Grater Glasgow)의 평균 부동산 가격자료에 적용하여 분석한다. 본 자료는 Scottish Neighborhood Statistics(SNS) 데이터베이스(<http://www.sns.gov.uk/>)에서 구할 수 있고, 지도를 그릴 수 있는 지형다각자료와 좌표가 포함되어 있는 자료를  $\mathcal{R}$ 의 CARBayes 패키지에서 쉽게 구할 수 있다.

연구지역은 그레이터 글래스고우와 클라이드(Clyde)의 건강관련 구분지역(health board)으로 271개의 중간지역(intermediate geographies; IG)으로 나뉘어져 있다. 이러한 IG은 영역넓이의 중앙값이 124헥타르이고 인구밀도의 중앙값이 4,239명인 소지역이다. 관측된 변수로는 영국 파운드로 관측된 주택가격(in thousands)과 10,000명당 범죄율(crime rate), 방의 갯수(number of rooms)의 중앙값, 부동산 판매 비율(%)(property sales), 마켓까지의 운전시간(분)(drive time to a shop), 그리고 각 IG에 가장 일반적인 부동산의 형태(type)이다. 반응변수는 주택가격(house price)이고 type변수는 명목형 변수로 ‘flat’, ‘terraced’, ‘semi-detached’와 ‘detached’의 네수준으로 표시된다. 주택가격은 50,000에서 372,800 영국 파운드로 관측되어 변화량이 크다는 것을 알 수 있다. 범죄율은 인구 10,000명당 각 지역에서 관측된 총 범죄 수를 기록한 것으로 82에서 1,994의 범위안에 있다. 그러나, 글래스고우의 도시 중심부의 범죄율은 인구 10,000명당 8,009건으로 이상치라 간주되어 데이터 분석에서 제외한다.

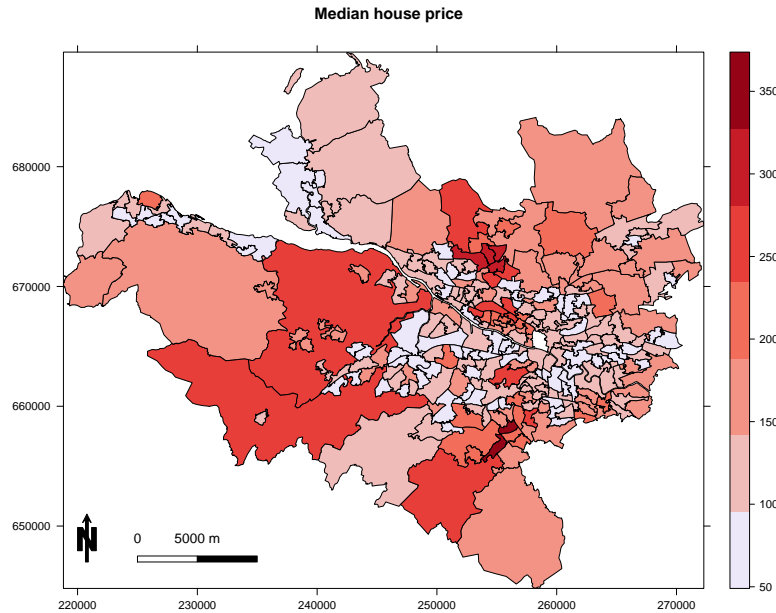


Figure 3.1. Median property prices across Greater Glasgow (in thousands), Scotland, in 2008.

스코틀랜드의 그레이터 글래스고우의 부동산 가격의 중앙값을 7개의 구간으로 나누어 그린 분포도가 Figure 3.1에 있다. Figure 3.1을 살펴보면 다른 주변지역들과의 관계가 없이 높은 부동산 가격을 보이는 몇몇 지역을 발견할 수 있다. 그림에서 어두운 진한색을 보이는 부분인데 남동쪽으로 흐르는 얇은 하얀 선으로 표시된 클라이드 강의 북쪽 지역과 중심부 지역의 두 부분으로 크게 나눌 수 있다.

부동산가격에 자연로그 변환을 적용한 변수를 반응변수  $Z_i = \ln(Y_i)$ 로 사용하고 가우스 분포를 가정한다. 각 변수들간 산점도행렬을 살펴보면 선형모형의 적용을 위하여 범죄율과 쇼핑센터까지의 운전시간에 자연로그변환을 적용하는 것이 적절하다고 판단하여 자연로그변환된 범죄율과 운전시간을 설명변수로 대신 사용한다. 산점도행렬은 생략한다. 그러므로 주어진 자료를 위해 고려하는 모형은 다음과 같다.

$$Z_i = \mathbf{X}_i\boldsymbol{\beta} + \epsilon_i, \quad i = 1, \dots, n.$$

Model 1.  $\epsilon_i \sim N(0, \sigma^2)$ : iid errors

Model 2.  $\boldsymbol{\epsilon} \sim N\left(\mathbf{0}, \sigma^2 \left(\mathbf{I} - \rho\tilde{\mathbf{W}}\right)^{-1} \mathbf{D}\right)$ : CAR errors

Model 3.  $\boldsymbol{\epsilon} \sim N\left(\mathbf{0}, \sigma^2 \left(\mathbf{I} - \delta_1\mathbf{W}^{(1)} - \delta_2\mathbf{W}^{(2)}\right)^{-1} \mathbf{D}\right)$ : DCAR errors

주어진 자료에 CAR 모형과 DCAR 모형을 적용하여 비교한다. 두 모형을 비교하기 위하여 아카이케 정보기준(Akaike information criterion; AIC), 베이지안 정보 기준(Bayesian information criterion; BIC)

$$\text{AIC} = -2\ell(\hat{\boldsymbol{\theta}}) + 2k, \quad \text{BIC} = -2\ell(\hat{\boldsymbol{\theta}}) + 2\ln(n)$$

**Table 3.1.** Maximum Likelihood estimates of parameters based on fitting CAR and DCAR models to the median property prices in Greater Glasgow

Parameter	iid		CAR		GCAR	
	Est.	Std.Err.	Est.	Std.Err.	Est.	Std.Err.
$\rho$	-	-	0.925	0.075	0.987	0.147
$\alpha$	-	-	-	-	0.744	0.135
$\sigma^2$	0.227	-	2.408	0.209	2.409	0.213
Intercept	5.275	0.277	5.035	0.273	5.038	0.273
log(crime)	-0.207	0.031	-0.190	0.030	-0.191	0.030
Rooms	0.229	0.027	0.245	0.026	0.245	0.026
Sales(%)	0.002	0.0004	0.003	0.0003	0.003	0.0003
Type	-0.062	0.012	-0.039	0.023	-0.040	0.023
log(time)	-0.079	0.025	-0.026	0.030	-0.027	0.003
AIC	-	-	25.118		24.131	
BIC	-	-	57.504		52.918	

을 고려한다.  $\ell(\boldsymbol{\theta})$ 는 식 (2.7)에서 정의된 로그우도함수값이고  $k$ 는 모수의 갯수이다. AIC와 BIC는 모형선택에 가장 많이 사용되는 통계량으로 AIC는 미니맥스 비율 최적의 특성이 있고, BIC는 참모형을 선택하는데 일치성이라는 특성이 있다. 그러므로 두 통계량을 모두 사용하여 비교한다. 그리고 평균제곱오차

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (z_i - \hat{z}_i)^2$$

를 사용한다. 예측값  $\hat{z}_i$ 으로는 최량 선형 비편향 추정량(best linear unbiased estimator)을 계산값을 사용하는데 다음과 같은 형태로 표현한다.

$$\text{Model 1. } \mathbf{X}_i \hat{\boldsymbol{\beta}}$$

$$\text{Model 2. } \mathbf{X}_i \hat{\boldsymbol{\beta}} + \hat{\rho} \sum_{j=1}^n \tilde{w}_{ij} (z_j - \mathbf{X}_j \hat{\boldsymbol{\beta}})$$

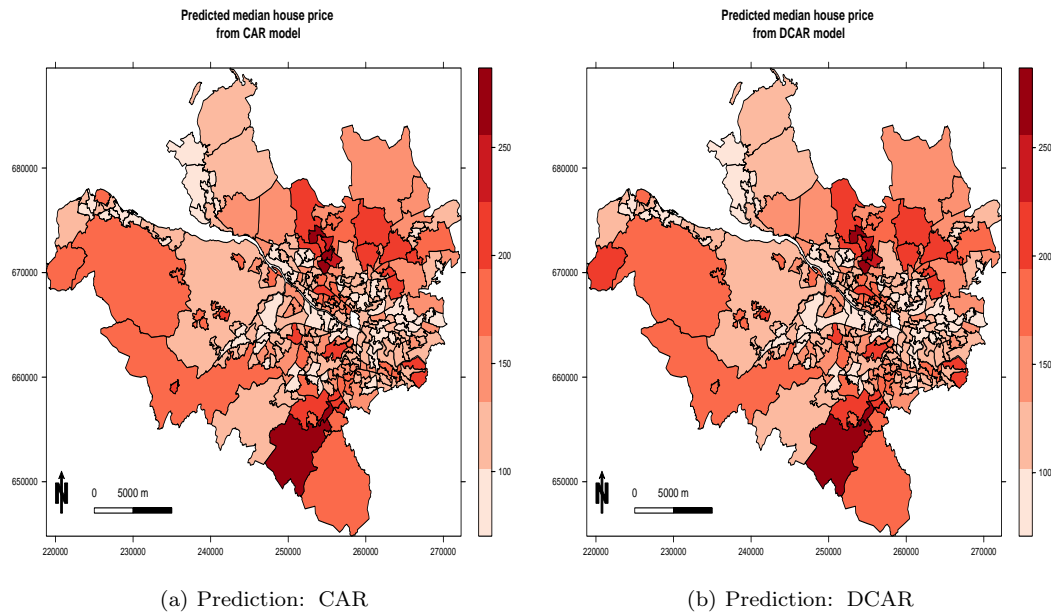
$$\text{Model 3. } \mathbf{X}_i \boldsymbol{\beta} + \hat{\delta}_1 \sum_{j=1}^n \tilde{w}_{ij}^{(1)} (z_j - \mathbf{X}_j \hat{\boldsymbol{\beta}}) + \hat{\delta}_2 \sum_{j=1}^n \tilde{w}_{ij}^{(2)} (z_j - \mathbf{X}_j \hat{\boldsymbol{\beta}})$$

선형모형, CAR 모형 그리고 DCAR 모형의 모수의 최대우도추정치와 표준오차가 Table 3.1에 있다. 이 표를 살펴보면, 모든 모형의 최대우도 추정치의 값이 매우 유사하다. 세 모형 모두에 유의한 설명 변수들은 로그변환된 10,000명당 범죄율과, 방의 갯수, 그리고 부동산 판매비율이다. 그러나 부동산의 형태의 최대우도 추정치는 선형모형의 경우 CAR 모형이나 DCAR 모형의 두 배 정도 작은 값으로 나타났다, 로그변환된 마켓까지의 운전시간(log(time))의 최대우도 추정치는 선형모형의 경우가 다른 모형의 세 배 정도 작은 값으로 나타났다. 주의할 점은 표준오차와 함께 고려했을 때, 선형모형에서는 type과 log(time)이 유의하지만 CAR 모형이나 DCAR 모형의 경우 두 변수는 유의하지 않다는 차이를 보이고 있다. 분산의 추정값을 살펴보면 독립오차의 경우 추정값이 공간자기상관 모형들의 10배정도 작게 나타난다. 이는 공간적 임의 효과를 고려하면 부동산 형태와 로그변환된 마켓까지의 운전시간은 부동산 가격에 영향을 미치지 않는다는 의미이다. 즉, 부동산 형태와 마켓까지의 운전시간은 공간적 상관성의 효과로 설명가능하다는 의미라 할 수 있다.



**Table 3.2.** Mean squared error:  $(1/n) \sum_{i=1}^n (z_i - \hat{z}_i)^2$

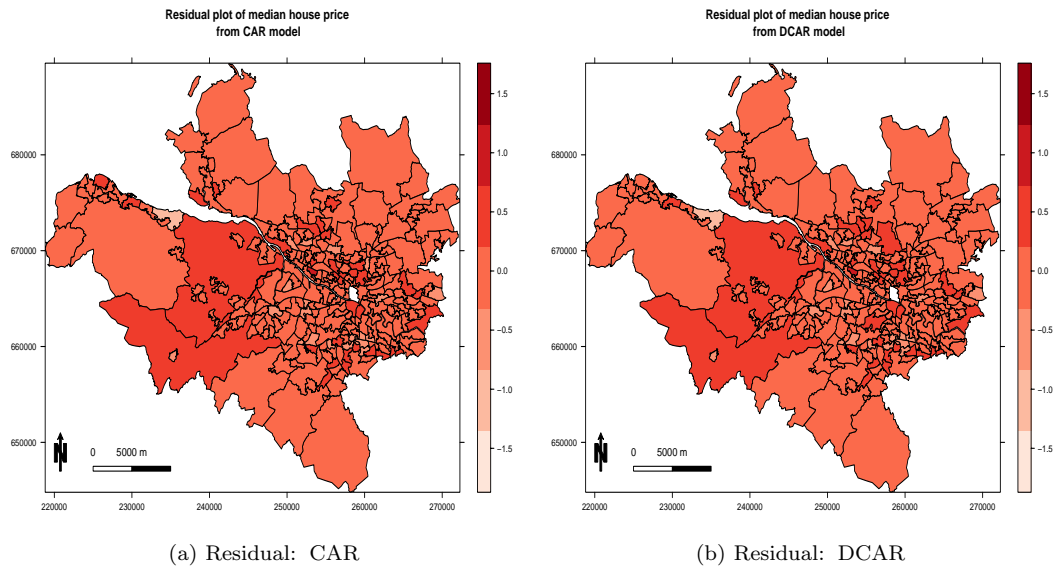
	평균제곱오차
Model 1 (iid error)	0.050
Model 2 (CAR error)	0.044
Model 3 (DCAR error)	0.040



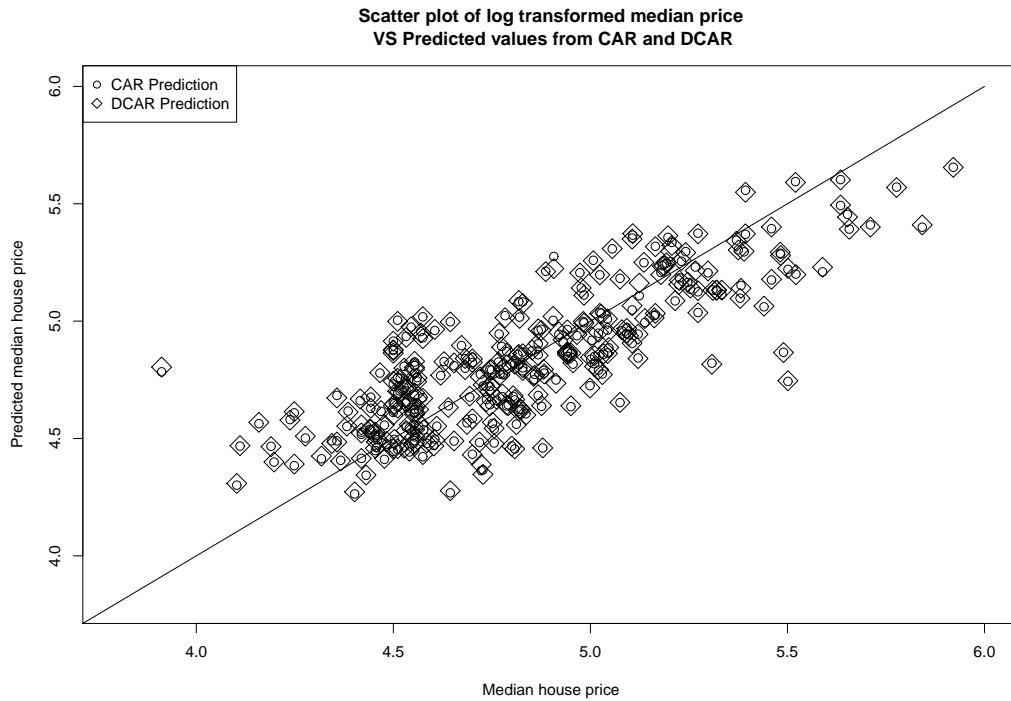
**Figure 3.2.** Predicted median property prices in Greater Glasgow.

CAR 모형의 공간효과 모수  $\rho$ 의 추정치는 0.925로 DCAR의 방향적 공간효과 모수들은  $\delta_1$ 과  $\delta_2$ 의 추정치인 0.987과 0.744의 사이의 값으로 나타난다. 방향성을 고려하지 않은 공간효과 모형에서도 유의한 양의 공간효과를 확인할 수 있으며, 방향성을 고려했을 때 북동남서 방향의 공간상관관계 (0.744)가 북서남동 방향의 공간상관관계 (0.987)보다는 수치적으로 작은 효과를 보이고 있다. AIC와 BIC를 비교하면 CAR 과정보다는 DCAR 과정이 조금은 더 그레이터 클래스고우의 부동산가격 분석 모형에 적합함을 알 수 있다. 그러나 그 차이는 무시할 만 하다. 마찬가지로 Table 3.2의 평균제곱오차를 살펴보면, 독립오차의 경우보다 공간자기상관 모형들의 경우 더 적합함을 알 수 있다. CAR 과정과 DCAR 과정은 어느 모형이 더 좋다고 평가할 수 없을만큼 평균제곱오차값이 유사하다.

Figure 3.2와 Figure 3.3는 공간자기상관을 임의효과에 포함한 모형들의 추정분포도와 잔차분포도이다. Figure 3.2를 살펴보면 (a)를 통해 공간자기상관의 효과로 거리상 가까운 지역의 값은 비슷한 경향을 나타내는 것을 볼 수 있으며, (b)를 통해 방향성을 포함한 효과의 추정값도 북동남서 방향과 북서남동 방향의 지역의 값이 비슷한 경향을 나타내고 있는 것을 알 수 있다. 공간자기상관을 모형에 포함시키면 전체적인 공간패턴은 매끄러워진다. CAR 모형과 DCAR 모형의 분포도가 비슷해보이지만 클라이드 강 주변 지역들에 작은 차이들을 발견할 수 있다. Figure 3.3의 잔차분포도를 살펴보면 공간자기상관 모형들의 잔차들이 연구지역 전체에 대한 큰 경향성은 없는 것을 관측할 수 있다. 이는 공간자기상관의 효과로 거리상 가까운 지역의 값을 비슷하게 만드는 영향으로 지역전체의 잔차에 존재하는 경향성을 약화시



**Figure 3.3.** Residuals from the predicted median property prices in Greater Glasgow by using CAR model and GCAR model.



**Figure 3.4.** Scatterplot of regional estimated prices from DCAR versus those from CAR for the median property prices in Greater Glasgow.

키기 때문이다. 그렇지만 앞의 표에서 설명한 것 처럼, DCAR 모형과 CAR 모형의 차이는 크지 않지만 클라이트 강의 북쪽 지역과 중심지역의 부동산가격을 예측하는 데는 DCAR 모형이 조금 더 좋은 모형임을 알 수 있다.

Figure 3.4를 보면 DCAR 공간모형의 예측치( $\square$ )가 CAR 공간모형의 예측치( $\circ$ )보다 직선식에 조금 더 가까운 것을 볼 수 있다. 이것은 그레이터 글래스고우의 부동산가격에 대하여 추정된 DCAR 추정치와 CAR 추정치의 비율에 대한 산점도로, 예측값이 관측된 값과 유사할 수록 기울기 1인 직선식에 더 가깝다는 의미를 갖는다. DCAR 예측치와 CAR 예측치 모두 직선식에 가깝지만, DCAR 공간 모형의 예측치가 조금 더 가깝다는 것으로, DCAR 모형이 추정된 공간자기상관을 바탕으로 연구지역의 추정값들을 더 안정화 시킨다고 할 수 있다.

#### 4. 결론

방향적 조건부 자기회귀 모형은 고유 이방성을 모형화하여 기존의 CAR 과정을 일반화 시킨 모형이다. 즉, 모든 방향으로 일정한 공간적 자기상관 무게를 부여하는 CAR모형의 한계에 대하여 강한 방향적 직교 공간적 자기상관성을 공간 임의의 효과에 포함하여 일반화 시킨 것이 DCAR 모형이다.

스코틀랜드 그레이터 글래스고우의 로그변환된 부동산 가격에 가우스 분포를 가정하며 공간 임의의 효과를 포함하여 분석하고 비교하였다. 공간자기상관의 효과로 거리상 가까운 지역의 값을 비슷하게 만드는 영향으로 지역전체의 잔차에 존재하는 경향성을 약화시키기 때문에 공간자기상관 모형들의 잔차들이 연구지역 전체에 대한 큰 경향성은 없는 것을 관측할 수 있다. 그러나 방향성을 고려했을 때 북동남서 방향의 공간상관관계 (0.744)가 북서남동 방향의 공간상관관계 (0.987)보다는 수치적으로 작은 효과를 보이고 있고, 모형비교에서도 DCAR 모형이 주어진 자료에 조금 더 적합함을 알 수 있다.

2절에서 정의한 이웃지역의 인덱스 집합의 부분집합은 파벌을 형성하는 더 많은 부분집합으로 나눌 수 있다. 그러나 이러한 방향성에 대한 부분집합의 갯수를 늘리면 더 많은 수의 모수를 추정해야하는 문제가 발생하며, 적당한 수의 부분집합의 갯수를 구하는 방법을 앞으로의 연구주제로 고려한다.

#### References

- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems (with discussion), *Journal of the Royal Statistical Society, Series B*, **36**, 192–236.
- Besag, J. (1975). Spatial analysis of non-lattice data, *The Statistician*, **24**, 179–195.
- Besag, J. and Kooperberg, C. (1995). On Conditional and Intrinsic Autoregression, *Biometrika*, **82**, 733–746.
- Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association*, **88**, 9–25.
- Brook, D. (1964). On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems, *Biometrika*, **51**, 481–483.
- Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization, *SIAM Journal on Scientific Computing*, **16**, 1190–1208.
- Cliff, A. D. and Ord, J. K. (1981). *Spatial Processes: Models & Applications*, Pion Limited.
- Cressie, N. (1993). *Statistics for Spatial Data*, John Wiley & Sons, Inc.
- Cressie, N. and Chan, N. H. (1989). Spatial modeling of regional variables, *Journal of the American Statistical Association*, **84**, 393–401.
- Griffith, D. A. and Csillag, F. (1993). Exploring relationships between semi-variogram and spatial autoregressive models, *Papers in Regional Science*, **72**, 283–295.
- Hrafnkelsson, B. and Cressie, N. (2003). Hierarchical modeling of count data with application to nuclear fall-out, *Journal of Environmental and Ecological Statistics*, **10**, 179–200.

- Kyung, M. and Ghosh, S. K. (2010). Maximum likelihood estimation for directional conditionally autoregressive models, *Journal of Statistical Planning and Inference*, **140**, 3160–3179.
- Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The SPDE approach, *Journal of the Royal Statistical Society, Series B*, **73**, 423–498.
- Mardia, K. V. and Marshall, R. J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression, *Biometrika*, **71**, 135–146.
- Miller, H. J. (2004). Tobler's first law and spatial analysis, *Annals of the Association of American Geographers*, **94**, 284–295.
- Reich, B. J., Hodges, J. S., and Carlin, B. P. (2007). Spatial analyses of periodontal data using conditionally autoregressive priors having two classes of neighbor relations, *Journal of the American Statistical Association*, **102**, 44–55.
- Rue, H. and Tjelmeland, H. (2002). Fitting Gaussian Markov random fields to Gaussian fields, *Scandinavian Journal of Statistics*, **29**, 31–49.
- Schabenberger, O. and Gotway, C. A. (2005). *Statistical Methods for Spatial Data Analysis*, Chapman & Hall/CRC.
- Song, H. R., Fuentes, M. and Ghosh, S. (2008). A comparative study of Gaussian geostatistical models and Gaussian Markov random field models, *Journal of Multivariate Analysis*, **99**, 1681–1697.
- Sweeting, T. J. (1980). Uniform asymptotic normality of the maximum likelihood estimator, *Annals of Statistics*, **8**, 1375–1381.
- van der Linde, A., Witzko, K.-H., and Jockel, K.-H. (1995). Spatial-temporal analysis of mortality using spline, *Biometrics*, **4**, 1352–1360.
- Wahba, G. (1977). Practical approximate solutions to linear operator equations when the data are noisy, *SIAM Journal on Numerical Analysis*, **14**, 651–667.
- White, G. and Ghosh, S. K. (2008). A stochastic neighborhood conditional auto-regressive model for spatial data, *Computational Statistics and Data Analysis*, **53**, 3033–3046.

# 방향성을 고려한 공간적 조건부 자기회귀 모형

경민정<sup>a,1</sup>

<sup>a</sup>덕성여자대학교 정보통계학과

(2016년 4월 26일 접수, 2016년 7월 3일 수정, 2016년 7월 3일 채택)

---

## 요약

공간통계 방법 중 지역에 대한 어떤 집합체 자료나 평균자료들을 분석하는데 일반적으로 공간적 자기회귀(conditionally autoregressive) 모형을 사용한다. 공간적 자기회귀 모형에 정의되는 공간적 이웃 소지역들은 중점의 거리나 근접성으로 정의된다. Kyung과 Ghosh (2010)는 방향에 따라서 이웃간 자기상관성의 크기가 다른 공간적 확장 모형을 제시하였다. 제안된 방향적 조건부 자기회귀(directional conditionally autoregressive) 모형은 고유 이방성을 모형화하여 기존의 CAR과정을 일반화한다. 제시한 방향적 조건부 자기회귀모형의 최대우도 추정량의 특성에 대해 설명하였고, 스코틀랜드 그레이터 글래스고우의 로그변환된 부동산 가격에 적용하여 조건부 자기회귀모형과 비교하였다.

주요용어: 이방성, 공간자료분석, 조건부 자기회귀모형, 최대우도추정법

---

이 논문은 덕성여자대학교 교내연구비 지원을 받아 수행되었습니다 (3000002407).

<sup>1</sup>(01369) 서울시 도봉구 삼양로 144길 33, 덕성여자대학교 정보통계학과. E-mail: mkyung@duksung.ac.kr