

## 의사우도법을 이용한 공간 종속 모형의 추정 \*

이윤동<sup>1)</sup> 최혜미<sup>2)</sup>

### 요약

본 연구에서는 베리오그램 추정을 통한 공간 종속성 추정방법에 있어서 의사우도 사용 방법을 설명하고, 모의실험을 통하여 전통적으로 사용되는 다른 방법들과 그 특성을 비교하고자 한다. 의사우도를 이용한 공간 종속 추정방법들은 그 통계적 성질이 우수할 뿐만 아니라, 전통적인 방법들에서 요구되어지는 관측치가 갖는 래그(lag)들을 미리 지정된 래그로 그룹화하는 과정이 필요 없어서 활용상의 우수성도 함께 가지고 있다. 또한, 이 방법에 대한 로버스트 방법을 개발하고 그 특성을 알아 보고자 한다.

주요용어: 의사우도법, 베리오그램, 공간 종속성, 로버스트 방법.

### 1. 서론

지리통계적 공간 자료의 해석에서 사용되는 베리오그램  $2\gamma(h)$ 은 주로 공간 종속성을 모형화하거나 측정하는데 언급되는 기본적인 양으로 약안정형(weak stationarity) 또는 기본안정형(intrinsic stationary) 공간 확률과정  $Z(s)$ 에 대하여

$$2\gamma(h) = E(Z(s+h) - Z(s))^2$$

과 같이 공간 확률과정의 관측 위치에 따른 이차 적률적 차이로 정의된 값이다. 또한, 지리통계적 공간 자료의 예측에 사용되는 크리깅 방법은, 베리오그램을 그 기저함수로 사용하는 스플라인 방법으로 해석되어진다. 그러므로, 베리오그램의 추정은 지리통계적 공간 자료 해석과 예측에 필수적인 단계이다. 크리깅 방법에 사용될 베리오그램은 특정한 수리적 조건들을 필요로 하는데, 이러한 조건들을 만족시키기 위해서 보통 베리오그램은 모형 모수  $\beta$ 를 통하여  $2\gamma(h; \beta)$ 와 같이 모수적으로 모형화하고, 이 모수를 추정하여 베리오그램을 얻게 된다.

전통적으로 사용되는 베리오그램 추정방법들은 다음과 같은 두 단계에 의하여 이루어진다. 미리 선택된 래그  $h_i, i = 1, \dots, K$ 들에 대하여 관측된 공간 자료로부터 해당 베리오그램의 추정값  $2\hat{\gamma} = (2\hat{\gamma}(h_1), \dots, 2\hat{\gamma}(h_K))$ 를 비모수적 방법으로 얻는다. 다음은 모수적 적합 단계로 모수적 모형에서 얻어지는 베리오그램의 값  $2\gamma(\beta) = (2\gamma(h_1; \beta), \dots, 2\gamma(h_K; \beta))$ 과 추정된 값  $2\hat{\gamma}$  사이의 정의된 거리  $Dist(2\hat{\gamma}, 2\gamma(\beta))$ 를 최소화 하는  $\beta$ 를 찾는 방법으로 이루어진다.

\* 이 논문은 2003년도 한국학술진흥재단의 지원(KRF-2003-C00030)에 의하여 연구되었음.

1) (143-701) 서울시 광진구 회양동 1, 건국대학교 상경대학 응용통계학과, 조교수

E-mail: poisson@dreamwiz.com

2) (151-742) 서울시 관악구 신림동 산56-1, 서울대학교 자연과학대학 통계학과, 박사후 연구원

E-mail: hyemichoi@yahoo.com

첫번째 단계에서  $2\hat{\gamma}(\mathbf{h}_i)$ 를 얻는 대표적인 방법으로는 Matheron의 방법이 있다. 먼저, 관측치  $\{Z(\mathbf{s}_i) | \mathbf{s}_i \in D\}$ 들이 있을 때 그 관측지점  $\mathbf{s}_i, \mathbf{s}_j \in D$ 에 대하여, 래그 (lag)라 불리는 관측 지점 좌표들 사이의 차이  $\mathbf{h}_{ij} = \mathbf{s}_i - \mathbf{s}_j$  들을 미리 지정된 래그  $\mathbf{h}_1, \dots, \mathbf{h}_K$ 를 기준으로  $\mathbf{h}_{ij} \simeq \mathbf{h}_k$ ,  $1 \leq k \leq K$ 가 되도록  $\mathbf{h}_{ij}$ 를  $K$ 개의 그룹으로 그룹화하고, 해당 그룹 내에 속하는  $(\mathbf{s}_i, \mathbf{s}_j)$  쌍들에 대응하는  $(Z(\mathbf{s}_i) - Z(\mathbf{s}_j))^2$ 들의 평균을 구하는 방식을 사용한다. 즉,  $N(\mathbf{h}_k)$ 를  $\mathbf{h}_{ij} \simeq \mathbf{h}_k$ 인 그룹에 속하는  $(\mathbf{s}_i, \mathbf{s}_j)$  쌍들의 집합이라고 할 때,

$$2\hat{\gamma}(\mathbf{h}_k) = \frac{1}{|N(\mathbf{h}_k)|} \sum_{N(\mathbf{h}_k)} (Z(\mathbf{s}_i) - Z(\mathbf{s}_j))^2$$

를 구하게 된다. 모수적 추정법에 있어서는  $\sum_k (2\hat{\gamma}(\mathbf{h}_k) - 2\gamma(\mathbf{h}_k; \beta))^2$ 를 최소화 하는 최소자승(OLS) 방법이 많이 사용된다. 차분자승항  $(Z(\mathbf{s}_i) - Z(\mathbf{s}_j))^2$ 를 이용하는 Matheron의 방법이 가지는 자료의 변화에 대한 민감성을 해결하기 위한 대안으로 Cressie와 Hawkins(1980)는 차분자승항 대신에 차분제곱근  $|Z(\mathbf{s}_i) - Z(\mathbf{s}_j)|^{1/2}$ 을 쓰는

$$2\tilde{\gamma}(\mathbf{h}_k) = \frac{\left\{ \frac{1}{|N(\mathbf{h}_k)|} \sum_{N(\mathbf{h}_k)} |Z(\mathbf{s}_i) - Z(\mathbf{s}_j)|^{1/2} \right\}^4}{0.457 + 0.494 |N(\mathbf{h}_k)|}$$

와 같은 통계량의 사용을 제안하였다.

한편, 모수적 적합 단계에서 OLS가 갖는 효율성을 높이기 위해서 Cressie(1985)는

$$\sum_k |N(\mathbf{h}_k)| (\hat{\gamma}(\mathbf{h}_k) / \gamma(\mathbf{h}_k; \beta) - 1)^2$$

과 같이 정의된 거리함수를 최소화 하는 가중 최소자승법(WLS)을 제안하였다.

이러한 대안적 연구들에도 불구하고 지금까지 알려진 베리오그램 추정 방법들은 크게 성공적이지 못하였다. 기본적으로 특정 래그를 미리 지정하고 이를 기준으로 관측치들에서 얻어지는 래그들을 그룹화 하여 추정을 수행하는 까닭으로 여러 가지 이론적 논의와 응용에 있어서 복잡성과 임의성을 배제 할 수 없다. 기존의 방법들이 기본적으로 미리 지정된 래그를 사용했던 이유는 원래의 확률과정이 정규분포를 따른다고 가정하더라도 비정규 분포를 따르게 되는 차분자승항이나 차분제곱근항들을 지정된 래그들을 기준으로 평균을 구함으로써 중심극한 정리에 의해 보다 정규분포에 가까운 형태를 갖도록 하고, 그렇게 구해진 값들을 정규분포의 로그우도 함수와 평균 추정에 있어서 동등한 역할을 하는 최소자승 방법이나 일반화 최소자승(GLS) 방법에 근거하여 거리함수를 정하고 그에 따라 모형모수를 추정하고자하는 데 그 목적이 있었다고 할 수 있다.

그러나, 근래에 일부 비정규분포들의 경우에도 로그우도 함수를 이용하는 경우와 동등한 추론을 가능하게 하는 의사우도(quasi-likelihood) 추정방법이, 일반화 선형모형에서 개발되어, 그 적용법들에 대한 연구들이 많이 진척되어 왔다. 공간 통계 분야에서도 Albert와 McShane(1995)은 일반화 추정 방정식 이론을 도입하여 공간 자료의 해석을 시도하고 있고, Ecker와 Gelfand(1997)는 베이즈인 입장에서 의사우도를 사용하여 공간 종속성을 추정하기 위한 연구를 시도한 바 있다.

본 연구에서는 Wedderburn(1974)이 제안한 의사우도 방법을 사용하여, 차분자승항들에 대하여 최소자승법에서와 같이 미리 래그를 지정하고 평균을 구하는 과정 없이 베리오그램 모형의 모수를

추정하는 방법을 설명하고, 그 효율성을 다른 방법들과 비교 하고자 한다. 제안된 방법은 래그 설정의 임의성을 배제 할 수 있을 뿐만 아니라, 평균화 과정에서의 효율성 상실을 방지할 수 있기 때문에 기존 방법에 비하여 훨씬 효율적인 추정이 가능하게 된다. 의사우도 추정방법은 기본적으로 공간 확률과정의 분포에 대한 정규성이 가정될 때 차분자승항이  $2\gamma(\mathbf{h})\chi^2(1)$ 인 분포를 갖는다는 점에 근거한다. 따라서, 제안된 의사우도를 이용한 공간 종속 추정방법이 비정규분포를 따르는 확률과정에서도 좋은 통계적 성질을 가지는지 고려해 보아야 한다. 3절에서 비정규분포를 따르는 공간 확률과정의 경우에도 로버스트한 의사우도 추정법에 대하여 살펴보고자 한다. 로버스트 의사우도 추정법은 최근 Cantoni와 Ronchetti(2001)에 의하여 심도있게 연구되었고, Adimari와 Ventura(2001)은 이 방법에 대한 적용 연구를 수행하였다. 또한, 국내에서는 최근 Cha 등(2003)에 의하여 ARCH모형 추정에 대한 적용 연구가 수행된 바 있고, So와 Shin(2001)은 단위근 검정에서의 로버스트 추정법에 대한 연구를 수행한 바 있다.

## 2. 공간 종속 모수의 의사우도 추정법

공간 자료  $Z(s_i), s_i \in D, i = 1, \dots, n$ 가 정규분포를 따르고  $2\gamma(\mathbf{h}; \beta)$ 와 같은 베리오그램으로 정의되는 종속성을 갖고, 평균  $m$ 이 미지인 약안정형 공간 확률과정  $Z(s)$ 로부터 관측된 값이라고 할 때, 그 차분자승항  $(Z(s_i) - Z(s_j))^2$ 들은  $2\gamma(s_i - s_j)\chi^2(1)$ 인 분포로 근사된다 (Cressie, 1993). 물론, 차분자승항들 사이에는 원 자료가 얻어진 확률과정에서 유전받은 종속성이 살아있게 되나, Cressie(1985)가 WLS를 제안할 때 사용한 바와 같이 차분자승항들 사이의 종속성을 무시하고 나면, 베리오그램 모형  $2\gamma(\mathbf{h}; \beta)$ 의 모수  $\beta$ 는 다음과 같이 정의되는 차분자승항  $Y_{ij} = (Z(s_i) - Z(s_j))^2$ 의 의사우도 함수를 최대화 하는  $\hat{\beta}$ 로 추정할 수 있다.

$$Q(\beta) = \sum_{ij} \int_{y_{ij}}^{\mu_{ij}} \frac{y_{ij} - t}{2t^2} dt.$$

여기서  $\mu_{ij}$ 는  $2\gamma(s_i - s_j; \beta)$ 를 나타낸다.

의사우도를 추정 모수  $\beta$ 에 대하여 미분하여

$$U(\beta) = \sum_{ij} \frac{\partial \mu_{ij}}{\partial \beta} \frac{y_{ij} - \mu_{ij}}{2\mu_{ij}^2}$$

라고 하면 의사우도로부터 얻어지는 추정 방정식은  $U(\beta) = \mathbf{0}$ 로 주어진다. 이 추정 방정식의 근은 다음과 같은 반복적 과정(McCullagh와 Nelder, 1989)을 통하여 얻어진다.

$$\beta^{(new)} = \beta^{(old)} - [E\dot{U}(\beta^{(old)})]^{-1}U(\beta^{(old)}).$$

여기서  $\dot{U}$ 는  $U$ 의  $\beta$ 에 대한 편미분 함수이다. 아래에서 임의의 함수  $f$ 의  $\beta$ 에 대한 편미분은  $\dot{f}$ 로 나타낸다. 또한, 모든 가능한 차분자승항들로 정의되는 추정방정식은 근을 구하기 위한 계산 부담이 너무 커서, 관측점  $s_i$ 들 사이의 거리가 주어진 상수값  $\lambda$ 이하인 관측치들의 차분자승항들로만 정의된 추정방정식을 고려 한다. 추정에 사용되는 차분자승항의 개수를  $n_\lambda$ 로 표시하기로 하자.  $n_\lambda = O(n)$ 인 성질을 가짐을 쉽게 알 수 있다.

추정방정식  $U(\beta) = \mathbf{0}$ 의 근  $\hat{\beta}$ 은 자료의 개수  $n$ 이 무한대로 커짐에 따라

$$\sqrt{n_\lambda}(\hat{\beta} - \beta) = \text{Normal}(\mathbf{0}, n_\lambda \mathbf{i}^{-1}(\beta)) + O_p(n_\lambda^{-1/2})$$

임을 McCullagh(1983)로부터 얻을 수 있다. 여기서  $\mu$ 은  $\mu_{ij}$ 로 이루어진 벡터라 하고,  $V$ 는  $2\mu_{ij}^2$ 를 대각원소로 갖는 대각 행렬,  $R$ 은 차분자승항들 사이의 상관행렬이라고 할 때,

$$\mathbf{i}(\beta) = (\hat{\mu}^T \mathbf{V}^{-1} \hat{\mu})(\hat{\mu}^T \mathbf{V}^{-1/2} \mathbf{R} \mathbf{V}^{-1/2} \hat{\mu})^{-1}(\hat{\mu}^T \mathbf{V}^{-1} \hat{\mu})$$

인 관계로 표현 된다. 일반적으로 공간 통계학에서 점근성에 대한 논의는 자료가 어떤 형태로 증가하는가에 따라 보다 복잡한 고려들이 필요하다. 위의 서술은 가장 기본적인 형태의 공간 통계적 점근성에 근거한 것이다. 위의 식에 대한 구체적인 증명이나, 점근성에 대한 보다 다양한 논의는 Lahiri 등(2002)에서 찾을 수 있다.

### 3. 로버스트 의사우도 추정법

앞 절에서 설명한 의사우도 추정법은 공간 확률과정에 대한 정규성 가정이 매우 중요한 역할을 한다. 이 정규성 가정이 어긋나지 않는 경우에는 앞 절에서 제안된 의사우도 추정법은 다음 절의 모의실험에서 확인되는 바와 같이 어느 정도 만족스러운 통계적 성질을 갖는다. 그러나, 일반적으로 통계학의 응용에서 가정되는 분포의 정규성은 엄격한 수리적 가정이라고 보기 어렵고, 따라서 정규성 가정의 이탈에서 오는 통계량의 성질을 살펴보는 것은 매우 중요한 문제라고 할 수 있다. 최근 Cantoni와 Ronchetti(2001)은 Mallow(1975)가 제안한 로버스트 추정법에 대한 의사우도를 정의하는 연구를 하였다. 위의 연구에서는 종속변수  $y_i$ ,  $i = 1, \dots, n$ 가 독립변수  $x_i$ 에 의해서 결정되는 평균  $\mu_i$ 를 갖고, 분산이 산포모수  $\phi$ 와 분산함수  $v_0(\mu)$ 에 대하여  $v(\mu) = \phi v_0(\mu)$ 로 주어질 때 그 의사우도  $Q^R(\beta) = \sum_{i=1}^n Q_i^R$ 를,

$$Q_i^R = \int^{\mu_i} \psi_c\left(\frac{y_i - t}{\sqrt{v(t)}}\right) \frac{w(\mathbf{x}_i)}{\sqrt{v(t)}} dt - \frac{1}{n} \sum_i \int^{\mu_i} E\left\{\psi_c\left(\frac{y_i - t}{\sqrt{v(t)}}\right)\right\} \frac{w(\mathbf{x}_i)}{\sqrt{v(t)}} dt$$

으로 정의하고 있다. 여기서  $\psi_c(r) = \min(c, \max(-c, r))$ ,  $c > 0$ 이다. Huber함수  $\psi_c(\cdot)$ 에서 상수  $c$ 의 선택은 바라는 추정치의 로버스트 정도와 효율성에 대한 희망 정도에 따라 임의로 선택되어진다.  $w(\mathbf{x})$ 는  $\mathbf{x}$ 에 의하여 결정되는 가중값으로 과대 영향치(influential data)의 영향력을 보정해 주기 위해서 고려되는 값이다. 보정항  $w(\mathbf{x}_i) = 1$ 인 경우는 의사우도 추정법에 대한 Huber적 방법이라고 할 것이다.

일반적인 선형모형 적합의 경우, 독립변수들이 큰 leverage를 갖게 되는 자료들에 대하여는 과대한 영향력을 감소시켜주기 위해서 작은 보정치를 주는 방식으로  $w(\cdot)$ 를 결정하지만, 관측위치들의 차이인 래그들을 (또는 래그에 대한 어떤 함수값을) 설명변수로 사용하여 공간종속성을 추정하고자 하는 경우에는 일반적 사항과는 달리 별도의 고려사항이 필요하게 된다. 기존의 공간종속성 추정 방법들에서는 모형 적합에 사용하게 될 래그를 선택하는데 있어서 매우 임의적 방법을 사용하고 있고, 선택된 래그들에 대한 가중값 조정은 Cressie가 제안한 WLS방법과 같이 추정될 모형에

따라 달라지는 방식을 사용하고 있다. 보통의 선형 모형에서 사용되는 설명변수들의 leverage를 기준으로 가중치를 고려하는 방식을 공간종속성 추정문제에 통합하여 적용하고자 하는 것은, 또 다른 좋은 연구 주제가 될 수 있을 것으로 보인다. 이에 대하여는 더 많은 연구와 논의가 필요하리라 보이므로, 본 연구에서는 가중함수의 사용을 고려하지 않는 경우, 즉 함수  $w(\cdot)$ 가 항등적으로 1인 경우만을 고려하기로 한다 (선택되어지지 않은 래그들에 대하여는  $w(\cdot)$ 가 0인 것으로 해석됨). 앞서 정의된 로버스트 의사우도를 공간 종속성 추정 문제로 한정하여 다시 써보면

$$Q_{ij}^R = \int^{\mu_{ij}} \psi_c\left(\frac{y_{ij}-t}{\sqrt{2t}}\right) \frac{w(\mathbf{h}_{ij})}{\sqrt{2t}} dt - \frac{1}{n_\lambda} \sum_{i,j} \int^{\mu_{ij}} E\left\{\psi_c\left(\frac{y_{ij}-t}{\sqrt{2t}}\right)\right\} \frac{w(\mathbf{h}_{ij})}{\sqrt{2t}} dt$$

와 같고

$$Q^R(\beta) = \sum_{ij} Q_{ij}^R$$

으로 정의 하기로 한다.

의사우도  $Q^R(\beta)$ 로부터 얻어지는 추정방정식에 대하여 논의 하기에 앞서, 다음 기호들을 정의 하기로 하자.  $\mathbf{y}$ 는  $y_{ij}$ 들로 이루어진 크기  $n_\lambda$ 인 열벡터를 나타내고, 행렬  $W$ 는  $w(\mathbf{h}_{ij}) = 1$ 를 대각원소로 갖는 대각행렬이라 하자. 즉,  $W$ 는 여기서는 단위행렬을 나타낸다. Wand(2002) 에서 사용된 원소별 벡터 연산법을 적용하여  $\mathbf{r} = (\mathbf{y} - \boldsymbol{\mu})/(\sqrt{2}\boldsymbol{\mu})$ 인 열벡터를 정의하자. 따라서, 의사우도  $Q^R$ 로부터 얻어지는 추정방정식은

$$U_R(\beta) = \dot{\boldsymbol{\mu}}^T W V^{-1/2} \psi_c(\mathbf{r}) - n_\lambda \boldsymbol{\alpha} \tag{3.1}$$

에 대하여,  $U_R(\beta) = \mathbf{0}$ 으로 주어진다. 여기서 Fisher의 일치성 조건을 만족하기 위해서는  $E U_R = \mathbf{0}$ 이어야 하므로  $\boldsymbol{\alpha}$ 는  $\dot{\boldsymbol{\mu}}^T W V^{-1/2} E \psi_c(\mathbf{r})/n_\lambda$ 인 값을 갖는다. 추정방정식  $U_R(\beta) = \mathbf{0}$ 의 근  $\hat{\beta}^R$ 은 M-추정량에 대한 일반적인 성질로부터,  $\sqrt{n_\lambda}(\hat{\beta}^R - \beta)$ 이 점근적으로 평균이  $\mathbf{0}$ 이고 분산이

$$\lim_{n \rightarrow \infty} n_\lambda (E \dot{U}_R)^{-1} E[U_R U_R^T] (E \dot{U}_R)^{-1}$$

인 정규분포를 따름을 알 수 있다. 여기에서  $E \dot{U}_R$ 과  $E[U_R U_R^T]$ 를 구체적으로 살펴 보면 다음과 같다.

$E(U_R) = \mathbf{0}$ 이므로  $E \dot{U}_R + E \dot{U}_R = \mathbf{0}$ 이고,  $(\partial/\partial \boldsymbol{\mu})(\log E) = \mathbf{r}^T V^{-1/2}$ 인 관계를 적용하면  $E \dot{U}_R = E[U_R(\log E)] = E[U_R \mathbf{r}^T] V^{-1/2} \boldsymbol{\mu}$ 이다. 따라서,  $E \dot{U}_R = -E(U_R \mathbf{r}^T) V^{-1/2} \boldsymbol{\mu}$ 임을 알 수 있다. 그런데,  $E \mathbf{r}^T = \mathbf{0}$ 이므로,

$$E \dot{U}_R = -\dot{\boldsymbol{\mu}}^T W V^{-1/2} E[\psi_c(\mathbf{r}) \mathbf{r}^T] V^{-1/2} \boldsymbol{\mu}$$

가 된다. 한편,  $\Sigma_\psi \equiv Cov(\psi_c(\mathbf{r}))$ 라 하면 (3.1)에서 정의된  $U_R$ 은

$$\begin{aligned} E[U_R U_R^T] &= \dot{\boldsymbol{\mu}}^T W V^{-1/2} E[\psi_c(\mathbf{r}) \psi_c(\mathbf{r})^T] V^{-1/2} W \boldsymbol{\mu} - n_\lambda^2 \boldsymbol{\alpha} \boldsymbol{\alpha}^T \\ &= \dot{\boldsymbol{\mu}}^T W V^{-1/2} \Sigma_\psi V^{-1/2} W \boldsymbol{\mu} \end{aligned} \tag{3.2}$$

임을 알 수 있다. 그런데, Cantoni와 Ronchetti(2001)의 1024쪽과 1028쪽에 있는 (3.2)와 관련된 식이 잘못 유도되었음을 언급하고자 한다.  $E[U_R U_R^T]/n_\lambda$ 에 대응하는 값  $Q(\psi_c, F)$ 을 얻는 과정에서,

$\mathbf{g}_i = \psi_c(r_i)w(\mathbf{x}_i)V^{-1/2}(\boldsymbol{\mu}_i)\boldsymbol{\mu}_i$ 에 대하여

$$E\{\sum_i(\mathbf{g}_i - \boldsymbol{\alpha})\sum_j(\mathbf{g}_j - \boldsymbol{\alpha})^T\} = \sum_i E(\mathbf{g}_i - \boldsymbol{\alpha})(\mathbf{g}_i - \boldsymbol{\alpha})^T$$

라고 가정하는 잘못을 범하고 있다. 자료의 독립성 유무에 관계없이  $\sum_i E(\mathbf{g}_i - \boldsymbol{\alpha}) = \mathbf{0}$ 인 조건으로부터 얻어진  $\boldsymbol{\alpha}$ 는  $\sum_{i \neq j} E(\mathbf{g}_i - \boldsymbol{\alpha})(\mathbf{g}_j - \boldsymbol{\alpha})^T = \mathbf{0}$ 을 만족시키지 못한다.

추정방정식  $U_R(\boldsymbol{\beta}) = \mathbf{0}$ 의 근은 앞 절의 일반 의사우도 경우와 같이 Fisher Scoring 근사법에 따라 기대 Hessian 행렬  $E\dot{U}_R$ 를 이용하여 Newton-Raphson 방법으로 구하는 것이 일반적이다. 그러나, 행렬  $E[\psi_c(\mathbf{r})\mathbf{r}^T]$ 의 모든 원소값들을 계산하려면 계산량이 너무 커지므로, 대각 원소만을 이용하여  $E\dot{U}_R$ 을

$$\dot{E}\dot{U}_R = -\dot{\boldsymbol{\mu}}^T W V^{-1/2} \text{diag}(E[\psi_c(\mathbf{r}) \circ \mathbf{r}]) V^{-1/2} W \dot{\boldsymbol{\mu}}$$

로 근사시켜서

$$\boldsymbol{\beta}^{(new)} = \boldsymbol{\beta}^{(old)} - [\dot{E}\dot{U}_R(\boldsymbol{\beta}^{(old)})]^{-1} U_R(\boldsymbol{\beta}^{(old)})$$

인 방법으로  $\hat{\boldsymbol{\beta}}^R$  계산했다. 연산자  $\circ$ 는 두 벡터의 원소별 곱을 의미한다. 이 방법은 차분자승항들 사이의 상관행렬  $E[\mathbf{r}\mathbf{r}^T]$ 를 이용하여,  $E[\psi_c(\mathbf{r})\mathbf{r}^T]$ 를

$$\text{diag}(E[\psi_c(\mathbf{r}) \circ \mathbf{r}])^{1/2} E[\mathbf{r}\mathbf{r}^T] \text{diag}(E[\psi_c(\mathbf{r}) \circ \mathbf{r}])^{1/2}$$

로 근사시키고, 상관행렬  $E[\mathbf{r}\mathbf{r}^T]$ 을 WLS에서와 마찬가지로 대각원소들만을 고려하여 단위행렬로 대체한 것과 동일한 근사법이다. 이는 다른 입장에서 Jacobi의 반복법을 적용한 것으로 볼 수도 있다(Young, 1971). 또한,  $n_\lambda$ 가 크지 않은 경우는 계산상의 부담 없이, Cressie(1993)의 96쪽에 주어진 차분자승항들 사이의 상관계수를 구하는 방법에 따라 행렬  $E[\mathbf{r}\mathbf{r}^T]$ 의 값을 직접 계산하여 대입하는 방법도 가능하다.

차분자승항 사이에 상관성이 존재하여  $E[\psi_c(\mathbf{r})\mathbf{r}^T]$ 가 대각행렬이 아니고, 행렬  $W$ 가 단위 행렬이 아니라고 가정하는 경우에는 앞 절에서 논의된 보통의 의사우도에서의 경우와는 다르게,  $E\dot{U}_R$ 가 비대칭성을 갖게됨에 주목할 필요가 있다. 이러한 비대칭성은  $Q_{ij}^R$ 이  $U_R$ 의 적분 형태로 표현되는 경우 그 값이 적분 경로에 무관하지 않게 됨을 의미한다. McCullagh와 Nelder(1989, p334)는 유의미한 의사우도가 되기 위해서는, 이러한 적분 과정이 적분 경로에 무관해야 한다고 언급하고 있다. Cantoni와 Ronchetti(2001)가 정의한 Mallows class에 대한 의사우도를 종속성이 있는 자료의 경우로 확장할 때 발생하는, 이러한 문제점을 해결하기 위해서 별도의 연구가 필요하다.

이제  $\boldsymbol{\alpha}$ 를 정의하기 위해서 사용한 벡터  $E\psi_c(\mathbf{r})$ , 점근 분포의 공분산을 구하기 위해서 사용한 크기  $n_\lambda \times n_\lambda$ 인 행렬  $E[\psi_c(\mathbf{r})\mathbf{r}^T]$ 와  $E[\psi_c(\mathbf{r})\psi_c(\mathbf{r})^T]$ 을 구하는 방법을 살펴보기로 하자. 일반적으로 차분자승항들 사이의 상관성이 존재하는 경우에  $n_\lambda \times n_\lambda$ 인 행렬  $E[\psi_c(\mathbf{r})\mathbf{r}^T]$ 와  $E[\psi_c(\mathbf{r})\psi_c(\mathbf{r})^T]$ 을 직접 구하는 것은 까다로운 일이다. 그러나, 큰 어려움 없이 다음과 같은 결과를 얻을 수 있다.

$$\begin{aligned} E\psi_c(\mathbf{r}) &= 1 \cdot \frac{1}{\sqrt{2}} \{aF_1(a) + b\bar{F}_1(b) - F_3(a) - \bar{F}_3(b)\}, \\ E\psi_c^2(\mathbf{r}) &= 1 \cdot [(3/2)(F_5(b) - F_5(a)) - (F_3(b) - F_3(a)) + (1/2)(F_1(b) - F_1(a)) \\ &\quad + (1/2)(a-1)^2 F_1(a) + (1/2)(b-1)^2 \bar{F}_1(b)], \end{aligned}$$

$$E[\psi_c(\mathbf{r}) \circ \mathbf{r}] = \mathbf{1} \cdot (1/2)[(a-1)(F_3(a) - F_1(a)) - (b-1)(F_3(b) - F_1(b)) \\ + 3(F_5(b) - F_5(a)) - 2(F_3(b) - F_3(a)) + (F_1(b) - F_1(a))].$$

위에서  $F_i(\cdot)$ 는  $\chi^2(i)$ 분포의 누적분포함수를 의미하고  $\bar{F}_i = 1 - F_i$ 이다. 상수  $a, b$ 는 각각  $a = \max(0, 1 - \sqrt{2}c)$ 와  $b = 1 + \sqrt{2}c$ 로 정의된다.

#### 4. 모의실험을 통한 비교

이 절에서는 앞 절에서 설명된 단순 의사우도 추정방법 (QLE)과 로버스트 의사우도 추정법 (RQLE)의 특성을 살펴보기 위하여 모의 실험을 수행한다. 먼저 공간확률과정 생성을 위해 사용되는 좌표로는, 85개의 관측점을 갖는 Wolfcamp-Aquifer 자료(Cressie, 1993)의 관측좌표들을 이용한다. 모형 모수의 상대적 크기를 조정하기 위하여 Wolfcamp-Aquifer 자료의  $x$ 좌표와  $y$ 좌표 값을 각각 1/100로 규모를 줄여서 사용한다. 베리오그램 모형으로는 단조성을 갖는 Matérn 족에서 가장 대표적인 지수형 모형(E-model)과 가우스 모형(G-model)을 설정한다. 각 모형은

$$\begin{aligned} \text{E-model: } 2\gamma(\mathbf{h}; \beta) &= 2\sigma^2[1 - \exp(-(\beta_0 + \beta_1|h_x| + \beta_2|h_y|))] \cdot I(\mathbf{h} \neq \mathbf{0}), \\ \text{G-model: } 2\gamma(\mathbf{h}; \beta) &= 2\sigma^2[1 - \exp(-(\beta_0 + \beta_1h_x^2 + \beta_2h_y^2))] \cdot I(\mathbf{h} \neq \mathbf{0}). \end{aligned}$$

과 같이 정의된다. 여기서 함수  $I(\cdot)$ 는 괄호안의 조건이 참이 되는 경우에 1을 갖고 그렇지 않은 경우에 0을 갖는 특성 함수이다.  $\mathbf{h} = (h_x, h_y)$ 이고,  $\beta_i > 0, i = 0, 1, 2$ 이다. 여기서  $2\sigma^2[1 - \exp(-\beta_0)]$ 는 조각효과(nugget effect)를 의미한다.  $\sigma^2$ 은 공간 확률과정  $Z(s)$ 의 분산이고 그 값이 알려져 있다고 가정한다( $\sigma^2 = 1$ ). 모수의 값은 지수형과 정규형 각각의 경우에 대하여  $\beta = (0.3, 0.2, 0.4)$  인 경우를 실험한다(이와 다른 모수 값들에 대하여도 실험하였으나 결과의 양태에 큰 차이가 없었음).

Matheron의 베리오그램 추정량  $2\hat{\gamma}(\mathbf{h})$ 와 OLS, WLS를 결합한 방법들을 고려하였고(이들을 각각 MOLS, MWLS라 함), 로버스트 베리오그램 추정량  $2\hat{\gamma}(\mathbf{h})$ 와 WLS를 결합한 방법(RWLS)를 비교 대상으로 선정하였다. QLE와 RQLE를 구하기 위해서는 유클리디안 거리  $d(\cdot, \cdot)$ 을 기준으로 전체 관측 지점의 차이들 중  $d(\mathbf{s}_i, \mathbf{s}_j) < 1.0$ 이 되는 771개의  $(\mathbf{s}_i, \mathbf{s}_j)$ 쌍들을 사용하였다. 공정한 비교를 위하여 MOLS, MWLS나 RWLS를 구할 때  $d(\mathbf{s}_i, \mathbf{s}_j) < 1.0$ 인 영역의 값을 그림 4.1과 같이  $d(\mathbf{s}_i, \mathbf{s}_j) = 0.25, d(\mathbf{s}_i, \mathbf{s}_j) = 0.5, d(\mathbf{s}_i, \mathbf{s}_j) = 0.75$ 인 원과,  $15^\circ, \dots, 180^\circ$ 를 지나는 축 방향선을 기준으로 48개의 그룹으로 묶어서 래그들을 그룹화하였다. MOLS, MWLS나 RWLS 경우에 이외에도 다양한 방법으로 래그 지정이 가능하지만, 미리 실험된 여러 방법들 중에서 이들의 통계적 성질이 공통적으로 우수한 예를 잡았다.

공간 자료가 갖는 분포의 특성을 비교하기 위하여, 표준정규분포, 형태모수와 척도모수가 각각 1인 감마분포를 평균이 0이 되도록 평행이동한 분포  $\text{Gamma}(1, 1) - 1$ , 자유도가 3인  $t$ -분포에 분산이 1이 되도록  $1/\sqrt{3}$ 을 곱한 분포들로부터 난수를 발생시켰다. 발생된 난수를 행렬  $A$ 를 이용하여 선형변환하였다. 여기서 행렬  $A$ 는  $AA^T$ 가 앞에서 가정된 베리오그램 모형에서 얻어지는 공분산 행렬이 되는 행렬이다. 즉, 두가지 베리오그램 모형(E-model과 G-model)과 세 가지 분포 모형(표준정규분포,  $t$ -분포, 감마분포)들을 결합한 여섯 가지 경우들에 대하여 실험하였다.

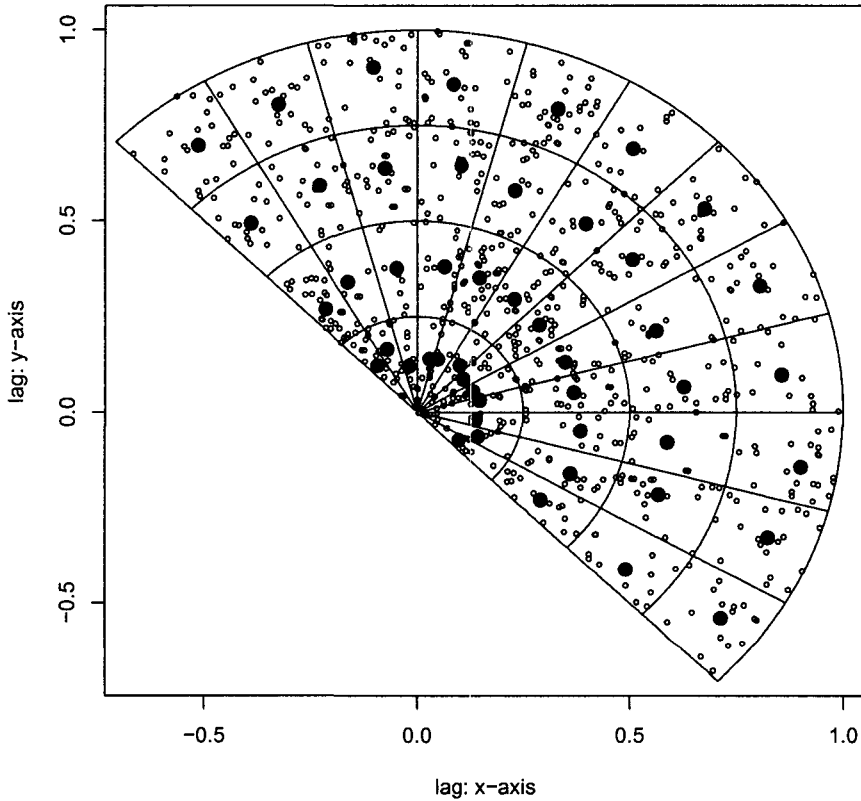


그림 4.1: 최소자승법을 위해 설정된 래그의 그룹화 방법

위와 같은 설정하에서 각각 500회 모의실험 결과가 그림 4.2와 4.3에 요약되어 있다. 상자그림에서  $b_0$ ,  $b_1$ ,  $b_2$ 는 각각  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ 의 추정치들을 의미하고, 점선들은 0.2, 0.3, 0.4를 지나는 선들이다.

그림 4.2는 E-model에서 정규분포와 t-분포를 갖는 경우에 MOLS, RWLS, QLE, RQLE를 상자그림으로 비교하였다. 그림 4.3는 G-model에서 정규분포와 감마분포를 갖는 경우의 비교 결과를 나타낸다. MWLS는 MOLS나 RWLS에 비하여 우수한 성질을 보이고 있었으나, 그 수치적 성질이 RWLS와 마찬가지로 좋지 못하여 수렴에 실패한 경우들이 많아 공정한 비교가 어려웠다. 그래서, MWLS의 실험 결과를 상자 그림으로 보여주지 않았다. 추정치들의 평균값(ave)과 평균자승오차(mse)는 표 4.1과 4.2가 보여 주고 있다.

QLE방법은 대체로 만족스러운 성질을 보였고, 특히 정규분포의 경우 좋은 성질을 갖고 있음을 볼 수 있다. RQLE는 정규분포인 경우나 비정규분포인 경우 모두 작은 값으로의 약간의 편향(bias)이



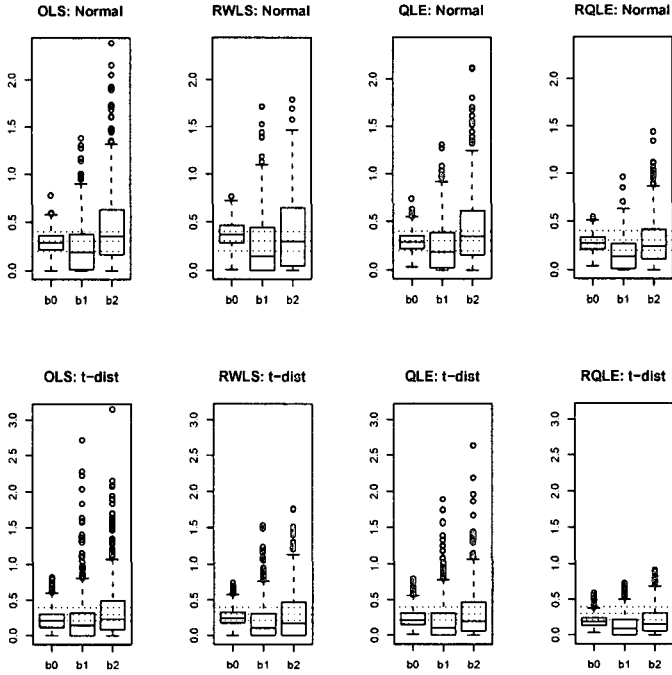


그림 4.2: 지수형 모형(E-model)에서의 비교를 위한 상자그림

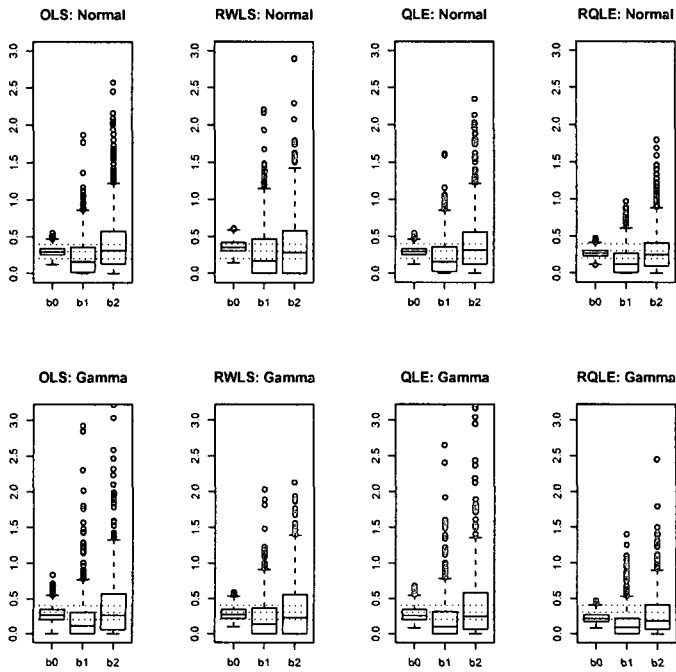


그림 4.3: 가우스 모형(G-model)에서의 비교를 위한 상자그림

표 4.1: E-model 에서의 모의실험 결과, (ave, 100×mse)

		Normal		t-dist'n		Gamma	
		ave	100×mse	ave	100×mse	ave	100×mse
MOLS	$\beta_0$	0.285	1.25	0.224	3.36	0.269	2.30
	$\beta_1$	0.247	6.84	0.252	13.75	0.233	10.30
	$\beta_2$	0.445	15.84	0.357	18.19	0.450	23.04
MWLS	$\beta_0$	0.353	1.85	0.283	2.07	0.331	2.13
	$\beta_1$	0.255	8.65	0.214	9.58	0.209	7.83
	$\beta_2$	0.394	13.01	0.248	11.34	0.333	12.89
RWLS	$\beta_0$	0.373	2.39	0.258	2.11	0.289	1.84
	$\beta_1$	0.271	11.59	0.216	8.85	0.235	9.71
	$\beta_2$	0.401	14.36	0.300	13.55	0.375	16.04
QLE	$\beta_0$	0.292	0.99	0.243	2.36	0.278	1.76
	$\beta_1$	0.242	6.03	0.219	9.75	0.220	8.82
	$\beta_2$	0.422	13.15	0.321	15.35	0.423	18.28
RQLE	$\beta_0$	0.275	0.79	0.192	2.02	0.223	1.33
	$\beta_1$	0.172	2.91	0.135	2.80	0.138	2.81
	$\beta_2$	0.290	7.10	0.208	7.59	0.258	7.51

표 4.2: G-model 에서의 모의실험 결과, (ave, 100×mse)

		Normal		t-dist'n		Gamma	
		ave	100×mse	ave	100×mse	ave	100×mse
MOLS	$\beta_0$	0.298	0.49	0.233	1.74	0.283	1.42
	$\beta_1$	0.236	7.14	0.224	13.30	0.236	13.80
	$\beta_2$	0.432	18.90	0.381	39.74	0.421	29.29
MWLS	$\beta_0$	0.346	0.79	0.280	1.06	0.326	1.33
	$\beta_1$	0.246	8.03	0.212	11.34	0.207	8.68
	$\beta_2$	0.415	18.78	0.301	15.03	0.315	15.80
RWLS	$\beta_0$	0.364	1.14	0.254	1.10	0.281	0.94
	$\beta_1$	0.317	16.83	0.229	11.33	0.256	10.92
	$\beta_2$	0.393	19.44	0.315	14.64	0.384	21.22
QLE	$\beta_0$	0.296	0.46	0.238	1.40	0.283	1.30
	$\beta_1$	0.234	6.66	0.216	14.16	0.238	13.85
	$\beta_2$	0.431	18.51	0.360	31.58	0.405	24.55
RQLE	$\beta_0$	0.270	0.45	0.230	1.43	0.223	1.14
	$\beta_1$	0.174	3.56	0.150	3.51	0.157	4.47
	$\beta_2$	0.317	9.94	0.280	10.32	0.280	10.70

보이기는 하나, 다른 추정방법들에 비하여 mse가 매우 작은 것을 확인 할 수 있다. RWLS는 정규 분포인 경우들에 비하여 비정규분포인 경우에 상대적으로 좋은 성질을 보이고 있다. 특히, 주목할 만한 점은 MWLS나 RWLS의 경우 추정치를 구하는 수치적 과정에서 수렴하지 않거나 제한값 (constraint value)으로 수렴하는 경우가 자주 발생하였으나(이런 경우는 비교에서 제외함), QLE 나 RQLE 모두 수치적 과정에서 수렴에 실패하거나 제한 값을 갖는 경우가 거의 발생하지 않았다.

QLE나 RQLE가 정규분포인 경우에 반복 제한 횟수 50회 이내에 수렴 값을 찾지 못한 경우는 2% 미만이었다.

### 5. 지형 자료에서의 종속성 추정

이 절에서는 Venables과 Ripley(2002)와 Ribeiro와 Diggle(2003)에 각각 topo, elevation이라는 이름으로 실린 52개의 관측치를 갖는 Davis(1973)의 어떤 지역의 지형학적 고도를 측정한 자료를 이용하여, 앞서 살펴본 QLE와 RQLE를 기존의 방법들과 비교하여 보기로 한다.

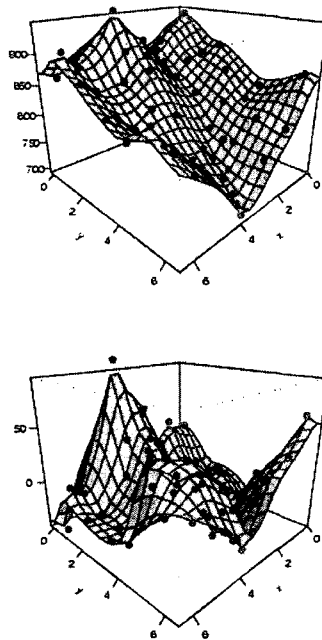


그림 5.1: 지형 자료의 관측위치와 관측값, 원자료 (상) 상관분석자료 (하)

그림 5.1은 관측 위치들과 그에 해당하는 관측 값들을 나타내고 있다. y 축 방향으로 뚜렷한 선형성과 x축 방향으로 V자형의 경향이 보이고 있다. Peixoto(1990)가 언급한 translation invariant 성질을 만족시킬 수 있도록 이차 이하의 전체 monomial들로 독립변수를 구성하는 회귀 모형을 적합하여, 이러한 경향을 제거하였다. 회귀 모형 적합 후에 남은 잔차들이 그림 5.1에 나타나 있다. 경향성이 제거된 자료의 공간 상관을 분산이 1이 되도록 표준화 하였다. 래그들을 분석한 결과 래그의 최대 크기는 대략 8정도였다. 표준화된 자료들에 대한 베리오그램 클라우드를 통해서 E-model이나 G-model이 적절함을 확인 하였다. 베리오그램 클라우드에 대한 상자그림(그림 5.2)을 바탕으로 래그의 크기가 2.5이하가 되는 398개의 관측 위치 쌍  $(s_i, s_j)$ 을 골랐다. 앞서와 비슷한 방법으로 선택된 래그들의 공간을 반경 기준으로 3등분하고, 각 기준으로 15°씩 12등분하여 36개의 그룹으로 만

들어 MOLS, MWLS, RWLS를 구하였다.

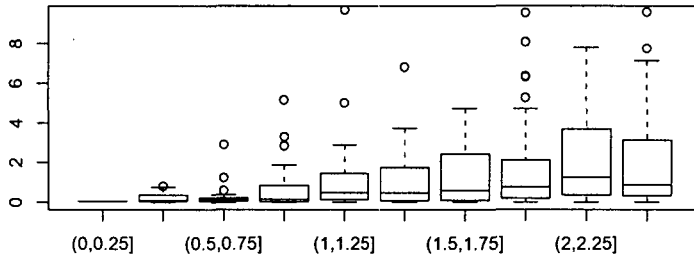


그림 5.2: 지형 자료에서의 베리오그램 클라우드에 대한 상자 그림

이 자료는 서로 다른 위치에서 동일한 관측값을 갖는 경우가 발생하여 그 차분자승항이 0이 되어, deviance 계산을 위해 로그 계산을 필요로 하는 QLE를 계산할 때 어려움이 발생하였다. 이를 보정해주기 위하여 0에 매우 가까운 다른 값으로 대체하여 QLE를 구하였다. 그 대체 효과를 보기 위하여 대체 값에 따른 민감도를 검사하였으나, 추정량에 미치는 효과는 무시할 수 있는 정도였다.

표 5.1은 각 추정치 모두 동일하게  $\beta_0$ 에 대하여는 매우 작은 값을 보여주고 있다. 이는 앞서 베리오그램 클라우드에 대한 상자 그림에서 보여지는 것과 일관된 결과라고 생각된다. RWLS는 다른 추정치들에 비하여 약간 큰 반면, RQLE는 다소 작은 값을 가짐이 확인되었다. MOLS와 QLE는 매우 비슷한 값을 가짐을 볼 수 있다.

표 5.1: 지형 자료에 대한 공간 종속성

		$\beta_0$	$\beta_1$	$\beta_2$
E-model	MOLS	0.000	0.869	0.457
	MWLS	0.000	2.009	0.522
	RWLS	0.000	2.145	0.379
	QLE	0.000	0.904	0.426
	RQLE	0.000	0.624	0.410
G-model	MOLS	0.000	0.781	0.439
	MWLS	0.000	1.059	0.648
	RWLS	0.007	1.089	0.491
	QLE	0.000	0.728	0.430
	RQLE	0.021	0.438	0.265

## 6. 맺음말

본 연구에서는 공간 자료의 종속성 추정을 위한 의사우도 추정방법(QLE)을 제안하고, 그 효율

성을 기존의 방법들과 모의실험을 통하여 비교하여 보았다. 또한, QLE 방법의 로버스트 변형으로 Cantoni와 Ronchetti(2001)방법에 근거한 로버스트 의사우도 추정법 RQLE 방법을 개발하여 비교하였다. 모의실험에서 QLE와 RQLE 방법은 다른 전통적인 방법들에 비하여 매우 우수한 특성을 보였다. QLE와 RQLE의 우수성은 본 논문에 열거된 경우 외에도 연구 과정 중에 실험된 다양한 종속 모형과 다른 모수 값들에 대하여 공통적으로 나타남을 확인할 수 있었다. 뿐만 아니라, QLE와 RQLE 방법은 전통적인 방법들에 비하여 래그들의 그룹화를 위하여 미리 특정 래그를 지정해야 하는데서 오는 임의성을 배제할 수 있어서 현실적인 사용상의 장점을 가지고 있다.

### 감사의 글

본 논문을 심사해주신 편집위원과 두분의 심사위원께 깊은 감사를 드립니다.

### 참고문헌

- Adimari, G and Ventura, L. (2001). Robust inference for generalized linear models with application to logistic regression, *Statistics and Probability Letters*, **51**, 413-419.
- Albert, P. and McShane, M. (1995). A generalized estimating equations approach for spatially correlated binary data: applications to the analysis of neuroimaging data, *Biometrics*, **51**, 627-638.
- Cantoni, E. and Ronchetti, E. (2001). Robust inference for generalized linear models, *Journal of the American Statistical Association*, **96**, 1022-1030.
- Cha, K., Kim, S. and Lee, S. (2003). Robust estimation using quasi-score estimating functions for nonlinear time series models, *Journal of the Korean Statistical Society*, **32**, 385-400.
- Cressie, N. (1985). Fitting variogram models by weighted least squares, *Journal of the International Association for Mathematical Geology*, **17**, 563-586.
- Cressie, N. (1993). *Statistics for Spatial Data*, revised edition. Wiley.
- Cressie, N and Hawkins, D. M. (1980). Robust estimation of the variogram, I, *Journal of the International Association for Mathematical Geology*, **12**, 115-125.
- Davis, J. C. (1973). *Statistics and Data Analysis in Geology*, Wiley.
- Ecker, M. and Gelfand, A. (1997). Bayesian variogram modeling for an isopropic spatial process, *Journal of Agricultural, Biological and Environmental Statistics*, **2**, 347-369.
- Lahiri, S., Lee, Y. D. and Cressie, N. (2002). On asymptotic distribution and asymptotic efficiency of least squares estimators of spatial variogram parameters, *Journal of Statistical Planning and Inference*, **103**, 65-85.
- Mallows, C. L. (1975). On some topics in robustness, *Technical Memorandum*, Bell Telephone Laboratorie, Murray Hill.
- McCullagh, P. (1983). Quasi-likelihood functions, *The Annals of Statistics*, **11**, 59-67.
- McCullagh, P. and Nelder, J. (1989). *Generalized Linear Model*, 2nd edition, Chapman and Hall.
- Peixoto, J. L. (1990). A property of well-formulated polynomial regression models, *The American Statistician*, **44**, 26-30.
- Ribeiro, P. and Diggle, P. (2003). GeoR package manual at R contributed package site, <http://www.cran.r-project.org>.
- So, B. S. and Shin, D. W. (2001). An invariant sign test for random walks based on recursive median adjustment, *Journal of Econometrics*, **102**, 197-229.

- Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*, Fourth edition Springer.
- Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models and Gauss Newton method, *Biometrika*, **61**, 439–447.
- Wand, M. P. (2002). Vector differential calculus in statistics, *The American Statisticians*, **56**, 55–62
- Young, D. M. (1971). *Iterative Solution of Large Linear System*, Academic Press.

[ 2004년 2월 접수, 2004년 5월 채택 ]

# Estimation of Spatial Dependence by Quasi-likelihood Method \*

Yoondong Lee <sup>1)</sup> Hyemi Choi <sup>2)</sup>

## ABSTRACT

In this paper, we suggest quasi-likelihood estimation (QLE) method and its robust version in estimating spatial dependence modelled through variogram used for spatial data modelling. We compare the statistical characteristics of the estimators with other popular least squares estimators of parameters for variogram model by simulation study. The QLE method for estimating spatial dependence has the advantages that it does not need the concept of lags commonly required for least squares estimation methods as well as its statistical superiority. The QLE method also shows the statistical superiority to the other methods for the tested Gaussian and non-Gaussian spatial processes.

*Keywords:* Quasi-Likelihood method; Variogram; Spatial dependence; Robust method.

---

\* This work was supported by the Korea Research Foundation Grant KRF-2003-C00030.

1) Assistant professor, Dept. of Applied Statistics, Konkuk University, 1 Hwayang-dong  
Gwangjin-gu, Seoul, 143-701

E-mail: poisson@dreamwiz.com

2) Post-doctoral researcher, Dept. of Statistics, Seoul National University, San 56-1 Sillim-dong  
Kwanak-gu, Seoul, 151-742

E-mail: hyemichoi@yahoo.com