

공간통계분석을 이용한 지가의 입지값 측정에 관한 연구

이지영* · 황철수**

The Measurements of Locational Effects in Land Price Prediction with the Spatial Statistical Analysis

Ji-Young Lee* · Chul-Sue Hwang**

요 약

본 연구에서는 GIS의 공간통계분석을 활용하여 지가 연구에 일반적으로 활용되고 있는 특성가격모형에서 입지적 특성이 갖는 영향력을 계량적으로 설명하기 위한 분석방법을 제시하였다. 여기에는 GIS 공간분석방법 가운데 중첩과 내삽 기능을 이용한 공간자료의 처리 과정이 포함되었다. 사례연구를 위해 동대문구 회기동의 1421개 개별지가에서 54개 표준지들을 추출하여 표준지의 중심좌표를 구하고, 이 벡터 자료점들과 공간적 관련성에 기초하여 조사되지 않은 지점의 지가 예측값을 확률적으로 평가할 수 있는 크리깅 분석방법을 적용하였다. 특히 이러한 분석 과정에서 변동도를 통해 분석한 공간적 자기상관관계는 공간 의존성의 형성과정을 추정할 때 장점이 있음을 밝혔다.

주요어 : 지가, 공간통계, 크리깅, 변동도, 공간적 자기상관관계

ABSTRACT : The purpose of this paper is to quantitatively measure the effect of location in evaluating the land value through the implementation of GIS coupled with spatial statistical analysis. We assumed that the hedonic price model, which was commonly used in modelling the land value, could not explain the spatial factor effectively. In order to add the spatial factor, the analysis of the spatial autocorrelation was used. The present project used 54 standard land price samples from 1421 parcel land values and applied Kriging to predict stochastically the unsampled values on the basis of spatial autocorrelation between location of

* (주) 지도시스템

** 경희대학교 지리학과 조교수(Assistant Professor, Department of Geography, Kyung Hee University)

vector data. This study confirms that the spatial variogram analysis has an advantage of predicting spatial dependence process and revealing the positive premium and the negative penalty on location factor objectively.

Keywords: land price, spatial statistics, Kriging, variogram, spatial autocorrelation

1. 서 론

토지의 공간가치는 지형, 교통, 인구, 주변 토지이용조건 등 토지가 가지는 자연적·사회적·경제적 특성에 의해 결정이 된다. 미시적 측면에서 지역별 여러 조건에 따른 지역 특성의 차이가 반영되어 지가가 형성되기 때문에, 이를 측정하는 것은 용이하지 않다. 이처럼 토지가격은 수요요인과 공급요인 등의 지가형성요인에 의해 영향을 받게 되고, 특히 토지의 특성 중 고정성으로 인해 그 토지가 속한 지역적 요인에 따라 유동적으로 변화한다. 이러한 지역적 요인에 따른 지가형성은 토지의 물리적 특성뿐만이 아니라 입지적 특성에 의해 크게 영향을 받게 된다. 지가 평가사들은 입지가 평가에 있어 가장 중요한 영향을 미친다는 공간적 중요성의 인식에도 불구하고 거시적인 공간구조에 따른 분석을 중심으로 해왔다. 이는 입지의 복잡한 역할과 공간적 상호작용으로 인한 객관적이고 체계적인 측정이 어렵기 때문이다. 이러한 문제의식은 공간의 적절한 형성과정이라는 공간적 상호작용(spatial interaction)에 대한 인식과 공간적 의존관계(spatial dependence)를 고려한 공간 분석기능의 필요성을 부각시킨다.

본 연구에서는 GIS의 공간통계분석을 활

용하여 특성가격모형(hedonic price model)에서 입지적 특성의 영향력을 설명하기 위한 대안적 분석방법을 제시한다. 일반적으로 특성가격모형은 토지의 물리적 특성, 즉 구조적 특성과 관련이 있는 개별지가에 대한 지형적 변수와 입지적 특성에 기초하여 지가를 설명하는 것이 목적이다. 이 방법은 다중회귀분석을 사용하여 속성들의 분포를 계산하는 것이지만 입지적 특성을 통합하기가 어렵기 때문에 GIS의 공간통계분석으로 이모형을 개선하고자 한다.

2. 연구 방법 및 연구 지역

2.1 분석 시스템 구축 및 구성

GIS 내에 공간통계분석 도구를 마련하기 위한 구체적인 방법으로, S-Plus for ArcView GIS의 모듈접근법을 이용한 S+API(application programming interface)를 적용한다. 이를 통해 특정한 함수를 작성하고 S+Spatial module(MathSoft)를 통해 공간적 자기상관관계를 분석하여 ArcView와 역동적 연계방안을 구축하도록 한다. 위와 같은 구축환경을 기반으로 토지 특성 차이에 따라 설명되는 변수들의 수학적 관련성을 형성하기 위해 계량적 분석방법 중 특성가격모형을 선택한다. 이 후 물리

적 특성만을 고려한 즉, 입지적 특성변수를 생략한 모형을 구축한다. GIS의 공간통계분석기법들에 의해 초기 모형을 분석한 후 입지적 특성을 측정한다. 54개 표준지들의 입지적 특성값들 사이의 공간적 자기상관관계를 고려하여 조사되지 않은 지역의 입지적 특성값을 확률적으로 평가할 수 있는 크리깅 방법을 이용한다. 이러한 크리깅 방법을 통해 연구지역의 입지 특성의 경향을 분석할 수 있다.

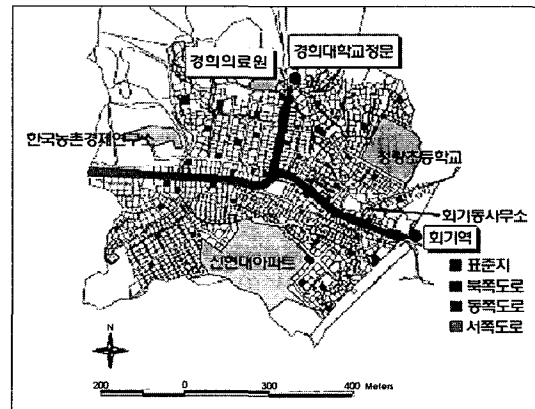
2.2 연구지역의 입지분석

본 연구에서는 서울시 동대문구 회기동이라는 지역 영향권내에서 미시적 입지(location)특성의 차이에 의해 발생하는 개별필지간의 입지적 특성을 설명한다. 이는 개별적인 요인을 중심으로 분석한 것이다. 입지는 접근성(accessibility)과 인근 지역의 질(neighborhood quality)로 크게 2가지로 분류하고 있다(McCluskey and Deddis, 2000). 접근성은 도심접근성, 공공편의시설 접근성, 유해시설과의 접근성 등이 포함되고 인근지역의 질에는 그 지역의 인구밀도, 지리적 위치, 소득수준, 지역의 지명도, 소음, 대기의 질, 수질, 일조권, 주변환경 등이 포함된다.

그러나 이러한 입지 여건들을 체계적이고 객관적으로 측정할 자료는 미비하다. 따라서 본 연구에서는 공시지가제도¹⁾에서 산정한 표준지 토지특성조사자료를 기준으로 하여 입지 변수를 정의하고 특성을

분석한다. 토지특성 조사자료는 지가형성에 가장 많은 영향을 미치는 것으로 생각되는 항목을 선정하여 기재한 것으로 조사방법은 먼저, 토지 대장, 임야 대장, 종합 토지세 과세대장, 국토이용계획 등 공적규제 사항을 도면과 나타내고 현지 조사를 통하여 실제 토지용도, 지형지세, 도로조건, 편익시설과 유해시설의 접근성, 농지의 경우 비옥도, 경지정리의 유무, 주변환경 등을 조사하여 정리하게 된다.

이와 같은 토지특성 조사자료를 기반으로 회기동 개별필지들의 토지특성을 조사한 결과 54개의 대표 표준지를 가지고 있었다. 주요 특징으로는 경희대학교, 경희의료원, 경희중고등학교, 한국농촌경제연구원, 회기시장, 회기역이라는 공공서비스 편의시설 등이 위치하고 있다. 특히 경희대학교와 경희의료원, 회기역의 입지로 유동 인구수가 많기 때문에 지가에 미치는



[그림 1] 연구지역(동대문구 회기동) 필지와 주요위치

1) 표준지공시지가는 전국의 조사대상 토지 중에서 일정수의 표준지를 선정하고 그 가격을 조사·평가하여 공시한다. 표준지는 토지이용상황이나 주변 환경, 지리적 위치 등 자연적, 사회적 조건이 일반적으로 유사하다고 인정되는 일단의 토지 중 대표성, 중용성, 안정성, 확정성의 기본 원칙에 의거하여 감정평가사가 선정한다. 표준지 가격은 개별토지 가격산정의 기준이 되므로 보다 적정한 가격이 산정되도록 하기 위해 전문 감정평가사에게 의뢰하여 조사·평가하고 있다.

영향력이 토지의 물리적 특성보다는 입지적 특성에 의해 공간적 차별화가 나타나게 된다. 또한 경희대학교 정문에서 회기역으로 가는 도로 주변 필지와 홍릉쪽으로 가는 도로 주변 필지 가격을 비교했을 때, 토지의 물리적 특성이 유사함에도 불구하고 차등적인 분포를 보이게 된다.

3. 지가의 탐색적 공간자료 분석

자료의 경험적 분포 패턴이나 상관관계를 파악하기 위해 일련의 탐색적 분석 기법들을 이용하여 회기동 지가의 공간적 분포 특성을 살펴본다. 이는 자료의 사전 조사적 의미 뿐 아니라 보다 복잡한 예측 모형의 유효성을 평가하기 위한 중요한 수단이 된다.

3.1 지가지료의 선택 · 처리

토지특성조사표²⁾에 제시된 특성으로는 시, 구, 표준지 번호, 지목, 면적, 용도지역, 용도지구, 계획시설, 계획시설저축율, 농지, 비옥도, 경지정리, 임야, 토지이용상황, 지형의 고저, 형상, 방위, 도로 접면, 유해시설 접근성(철도·고속도로 등과의 거리, 폐기물처리·수질오염방지시설 등과의 거리), 기타(결정지가, 산정지가, 전년 지가) 등 20개 특성이 있다. 따라서 이 조사표에 제시된 특성들과 결정지가에 대한 선형관계를 함수로 하여 분석한다.

한편 회기동은 1927개의 필지로 나누어

져 있으며, 이 중 학교와 병원, 도로로 이용되고 있는 필지는 제외하여 1421개의 필지만을 분석대상으로 한다. 이 중 표준지는 면적, 지목, 지리적 위치, 이용상황, 용도지역, 주위환경, 도로 접면 조건, 형상 지세, 유해시설과의 접근성, 지역적 특성 등을 토지 특성으로 하여 구청 공무원에 의해 산정된 개별공시지가와는 달리 전문가인 감정평가사와 건설교통부에 의해 산정된 것이다. 따라서 54개의 표준지를 표본추출하여 입지적 특성값을 분석하고 이를 기반으로 하여 주변 개별필지들의 입지값을 예측한다.

3.2 특성가격함수를 이용한 입지적 특성값 정의

특성가격함수를 기초로 하여 지가결정요인의 지가영향력을 설명할 때 토지특성조사표에 제시된 특성들만을 고려하여 분석한다. 지가를 분석하는데 있어 특성가격모형은 한 필지의 지가와 그 필지가 가진 토지 특성간에는 일정한 함수관계가 있다는 개념을 기초로 하고 있기 때문에 각 특성이 지가 형성에 미친 영향력을 분석해낼 수 있다.

그러나 입지성 요인의 객관적·체계적인 지표가 없기 때문에 표준지를 기준으로 한 특성가격함수에서 입지특성을 생략한 모형에서 회귀잔차들을 입지값으로 가정하여 공간적 자기상관관계를 측정하고 이를 기반으로 크리깅 내삽을 한다.

실제 분석에 있어 특성가격함수를 추

2) 이 토지특성조사표는 동대문구청 지적과에서 2000년도 지가를 산정한 알파(α) 데이터 중 회기동 일부만을 구축한 자료이다.

정하기 위하여 분석 대상필지의 지가를 종속변수로 19개³⁾의 지가결정요인을 독립변수로 하는 다중회귀분석을 선형 $[y=f(x)]$, 비선형 $[\log y=f(\log x)]$, 준로그 $[\log y=f(x)]$, 역준로그 $[y=f(\log x)]$ 의 4가지 함수형태에 적용하였다. 독립변수는 모든 요인(변수)을 투입하는 입력(Enter) 방법으로 한다. 그 결과 준로그함수 형태가 가장 우수하게 나타나($R^2:66\%$) 이 함수를 선택했으며, 변수들의 공선성(multi-collinearity) 검증 결과 계획시설 구분과 계획시설저축율이 0.9이상으로 나타나 계획시설저축율을 제외했다. 그것은 여러 변수들이 서로 완전한 선형관계에 있음을 의미하는 공선성 검증에서, 만약 다중공선성이 존재하면 결정계수가 크거나 통계적으로 회귀식이 의미있다고 하더라도 신뢰할 수 없는 경우가 발생하기 때문이다. 보통의 경우 공선성이 0.8이상이면 공선성이 있는 것으로 추정한다. 또한 토지특성조사에서 지목, 기타, 농지, 임야, 비옥도, 경지정리, 유해 폐기물 등은 모든 필지에 따라 같거나 해당경우가 없기 때문에 제외하여, 실제분석에 이용된 토지의 물리적 특성으로 고려한 변수들은 면적, 용도지역, 용도지구, 토지이용상황, 계획시설구분, 지형의 고저, 형상, 방위, 도로조건으로 한정한다. 이러한 변수만을 고려한 모형을 구축하여 54개 표준지공시지가의 각각의 회귀잔차값을 구하고 이를 입지적 특성값으로 가정한다.

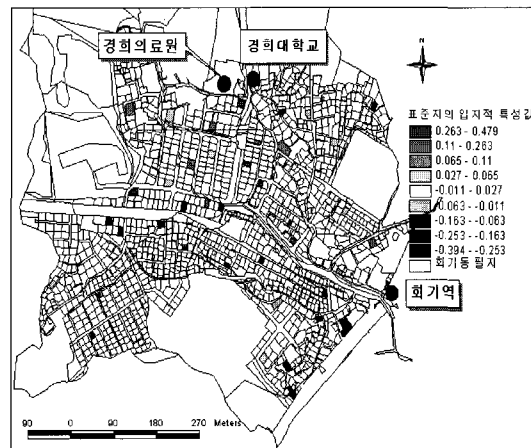
$$\log \text{지가} = 14.82249 - 0.00002458503 * \text{면적} + 0.02591855 * \text{용도지역} + 0.0238647 * \text{용도지구} - 0.07$$

$$893717 * \text{계획시설} + 0.005821861 * \text{토지이용상 황} - 0.1094059 * \text{지형고저} - 0.0107049 * \text{지형모양} - 0.006249612 * \text{지형방위} - 0.09090884 * \text{도로접면} + 44(\epsilon) \dots \dots \text{모형 ①}$$

3.3 입지적 특성의 공간 구조

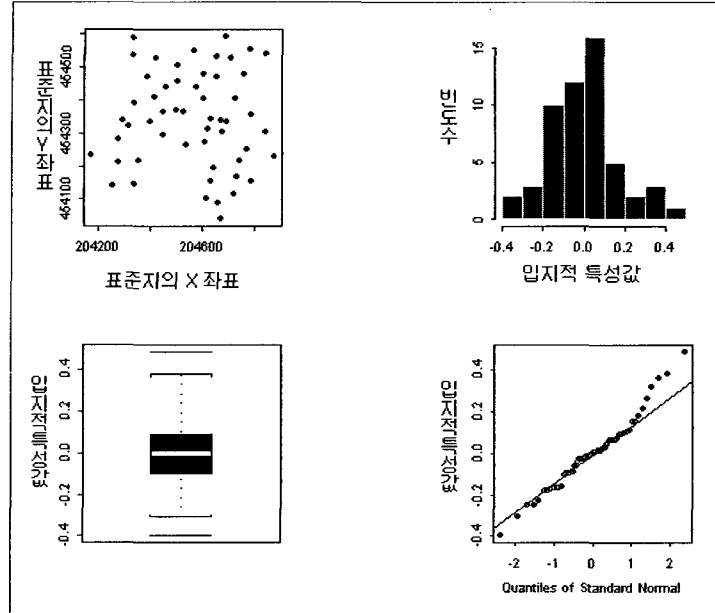
입지적 특성변수를 생략한 모형 ①에서 추정된 회귀잔차값을 입지적 특성값으로 가정하였을 때 이러한 입지적 특성값에 따른 공간구조를 살펴보고 공간적 경향을 파악한다.

먼저, 표준지의 공간적 분포와 모형 ①의 회귀잔차값을 잔차도로 나타낸 결과 경희의료원과 인접한 표준지의 잔차값이 가장 높았으며 도로와 인접한 필지들이 다른 표준지보다 상대적으로 높았다. 즉, 잔차값이 높다는 것은 물리적 특성보다는 입지성 특성에 의한 영향이 크다고 할 수 있다. 이러한 표준지의 입지적 특성값의

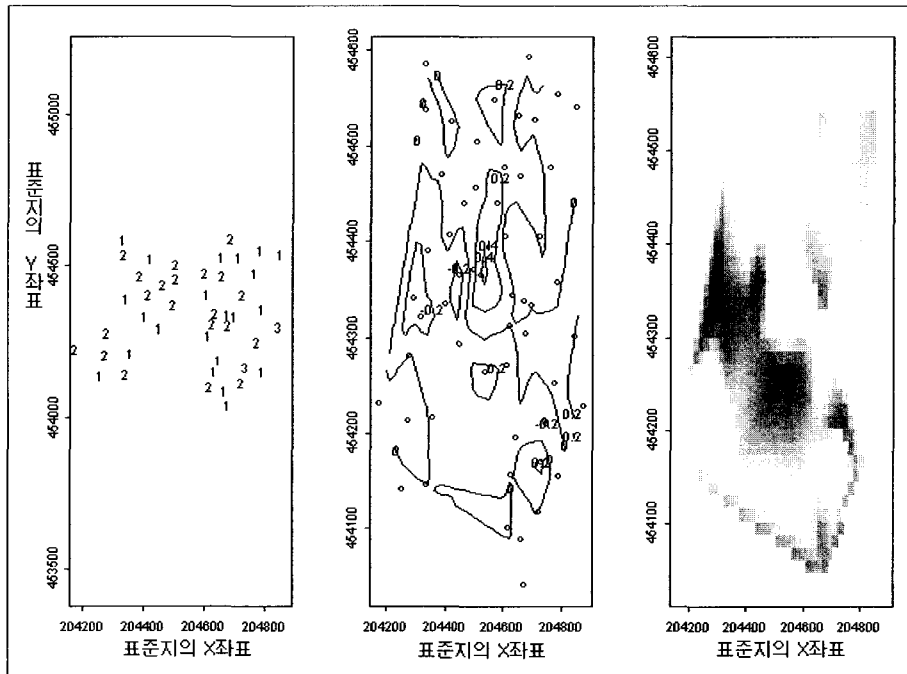


[그림 2] 입지적 특성값의 산포도

3) '시(市)'를 제외한 토지특성변수를 의미한다.



[그림 3] 표준지 통계분석



[그림 4] 입지적 특성값의 공간 경향분석

경험적 분포패턴을 파악하기 위한 가장 기본적인 통계 분석 절차라고 할 수 있는 모듈로서 히스트그램, 산포도, box plot, QQ plot, 줄기-잎-그림(stem-and-leaf diagram)과 같은 그래프와 자료의 기술적 요약 모듈이 구현될 수 있다. 또한 자료의 분포 패턴을 보다 직관적으로 파악 및 시각화하고자 할 때 이용하는 분석 기법으로서 지도를 이용한 패턴 분석은 질의 기능과 공간적 이레지점을 손쉽게 규명할 수 있는 기반을 제공하기도 한다.

표준지 자료의 기술적 통계 분석 결과는 최고값 0.479, 평균 0.686, 중위수 0.03, 최소값 -0.394로 다음과 같다. boxplot은 자료의 분포적 특징을 볼 수 있는 가장 기본적인 분석으로써 속성을 쉽게 파악하기 위한 위치적 대표치를 나타내는 중위수를 보여준다. 또한 자료의 그룹화에 따른 분포를 보여주는 히스트그램과 qq plot를 통해 정규분포로 전환해야하는지를 결정한다. 아래의 자료값의 경우 자료의 정규성 검증을 구체적으로 파악하기 위해 왜곡도를 측정하였는데 그 결과 0.42로 정규분포를 이룬다고 할 수 있다(그림 3).

다음은 공간 경향을 살펴보기 위해 각각의 점자료의 속성값을 서열척도하여 plot한 결과이다. 북쪽 도로를 중심으로 표준지 입지적 특성의 높은 값들이 집중되어 있는 것을 볼 수 있으며, 이를 이미지로 나타냈을 때 더욱 뚜렷한 모습을 확인할 수 있다(그림 4).

4. 지가의 입지값 측정

4.1 입지적 특성의 공간적 자기상관관계 분석

공간변동에 대한 모든 정보를 포함 및 요약하는 변동도를 통해 공간적 자기상관관계(또는 공간적 상관성)를 측정하고 표준지의 입지값을 기반으로 하여 인접한 개별필지의 입지성도 추정한다. 먼저, 입지값의 경험적 변동도 분포 패턴을 수행하여 거리(h)의 증가에 따른 속성 값의 상관성 변동도를 분석한 후 이를 이론적 변동도 모형에 적합시켜 크리깅의 이론적인 공분산 모형을 정의하도록 한다. h에 대한 $\gamma(h)$ 의 그래프를 변동도라 하는데 변동도는 위의 지역화 변수 이론의 내재적 가정⁴⁾에 의해 명시된 조건이 만족되면 분산은 표본자료로부터 다음과 같이 추정될 수 있다.

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n [Z(x_i) - Z(x_i + h)]^2$$

이 때 n은 거리 h에 의해 분리된 표본 점쌍(54개의 표준지 벡터 좌표)의 수이며, 표본간격 h를 래그(lag)라 한다. 이는 공간적 상호의존성을 알아보기 위해 이용하는 탐색적 분석기법으로써 입지값의 변동이 변화하는 스케일(scale)과 관련된 정보를 제공한다. 입지적 특성값의 공간적 상관성을 측정하기 위한 경험적 변동도 분석 결과는 <표 2>와 같다.

4) 차이의 고정(stationarity of difference)과 차이의 변동이라는 두가지 조건은 지역화 변수 이론의 내재적 가정에 요구되는 필요조건을 결정한다. 이것은 일단 구조적 영향이 설명되면 나머지 변동이 등질적이고 지점간의 차이는 단순한 거리함수가 된다는 것을 의미한다.

<표 2> 변동도 요약

	거리(m)	gamma	자료점수	범위
1	38.58671	0.03771959	12	0
2	56.24665	0.02321406	25	0
3	75.93158	0.02211489	44	0
4	92.72903	0.02966625	37	0
5	110.73778	0.03258860	52	0
6	129.88339	0.03356642	52	0
7	148.18000	0.03115189	59	0
8	165.75709	0.02929690	60	0
9	187.52750	0.03306128	73	0
10	202.82828	0.03608348	52	0
11	222.40364	0.02957924	69	0
12	239.86844	0.04000260	64	0
13	259.85767	0.03454708	68	0
14	277.93460	0.03538175	61	0
15	295.36052	0.02763741	59	0
16	314.43222	0.02970445	61	0
17	334.18050	0.02955043	68	0
18	353.08846	0.03505344	49	0

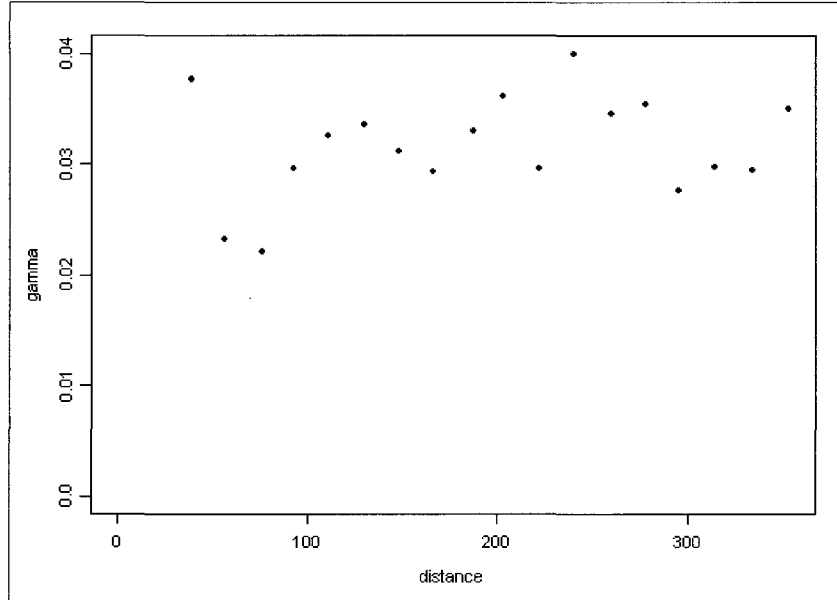
lag의 범위는 변동도에서 계산된 평균 거리에 의해 정의되고 이러한 lag 수는 전체거리⁵⁾를 고려하여 20개로 정의한다. 경험적 변동도는 표본점 수가 30개 이상을 포함하고 있는 거리 h만을 고려해야 하며 거리 신뢰성(the distance of reliability)은 최대거리의 1/2보다 작아야 한다. 이러한 분포를 그래프로 나타내면 [그림 5]와 같은 분포패턴을 보인다.

회기동 지역의 입지적 특성의 분산도는 특징적인 패턴을 나타내고 있는데, 100과 200단위의 lag에서 분산이 최대로 증가하는 것을 볼 수 있으며, 이 보다 큰 lag에서는 전반적으로 일정한 값을 유지하고 있다. 이 밖의 몇 가지 특징으로는 첫째, 큰 값의 lag에서 평탄해진다. 이 값을 실(sill)이라고 하는데, 이 값의 lag에서는 차

이의 분산이 거리에 따라 변하지 않기 때문에 자료점 사이에는 어떠한 공간적 종속성도 없다는 것을 의미한다. 둘째로 $\gamma(h)$ 의 원점에서 실까지 상승하며, 이것에서 h의 레인지(range)가 구해진다. 이 레인지는 지점사이의 차이가 공간적으로 종속적인 거리의 범위를 나타내기 때문에 분산도에 있어 매우 중요한 부분이다. 즉, 이 레인지(range)내에서는 지점사이의 거리가 가까울수록 유사할 가능성이 높다는 것을 의미한다. 이는 거리내 입지값의 공간적 분포가 매우 유사하다는 정보를 알려주고 거리 가중치를 이용하여 크리깅 내삽을 할 때 원도의 크기를 정하는 데에 필수적이다. 반면 내삽지점에서 자료점까지의 거리가 레인지보다 크다면, 그 자료점은 너무 멀리 떨어져 있기 때문에 내삽에 있어서 의미가 적다는 것을 의미한다. 마지막으로 h가 0에 접근할 때 $\gamma(h)$ 값이 양의 값을 가지는 것은 공간적으로 상호 무관한 잔차 즉, 노이즈 때문이다. ϵ^2 을 너겟분산(nugget variance)이라고 한다. 너겟분산은 측정오차의 잔차변동과 표본간격 이하의 거리에서 발생하는 공간적 변동의 합이며 분리할 수 없다. 즉, 표준지의 물리적 특성과 입지적 특성이외에도 다른 지가형성요인이 있다는 것을 의미하거나 측정오차를 의미할 수도 있다.

이와 같은 경험적 변동도는 공간 패턴을 밝히고 최적의 내삽을 하기 위한 거리 정보를 제공하게 된다. 이러한 정보를 기반으로 이론적 변동도 모형(theoretical variogram model)을 구축하게 된다. 한편, 이론적 변

5) 54개 표본점들간의 거리 행렬을 통해 구축된 거리들의 평균거리를 의미한다.



(평균 lag : 18.52553 lag수 : 20 최대거리 : 370.5105)

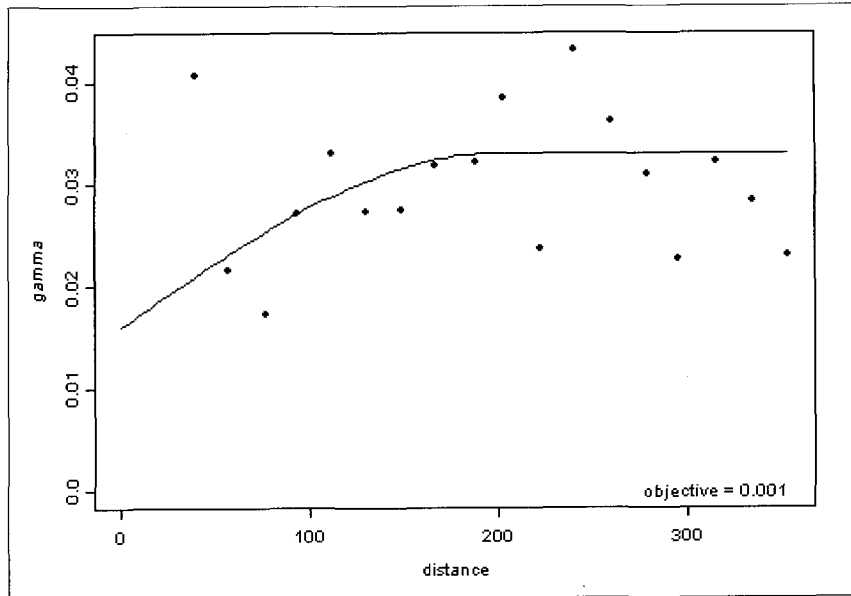
[그림 5] 입지적 특성값의 경험적 변동도 분포

동도 모형은 크게 4가지 구면(spherical)모형, 지수(exponential)모형, 가우스(gaussian) 모형, 선형(linear)모형으로 분류한다. 너겟 분산과 실이 명확하면 구면모형과 지수모형을 적용하고 변화가 매우 평탄하고 너겟분산이 공간적 의존성 확률 분산(ϵ')보다 작으면 가우스 모형을 적용한다. 선형 모형은 분산도가 실을 갖고 있지 않고 변동도의 곡선이 평탄할 때, 또는 레인지의 크기가 내삽거리보다 훨씬 클 때 사용된다. 회기동의 경험적 변동도 패턴분석 결과 너겟분산과 실이 분명하고 회기동 지역내라는 범위가 한정되어 있기 때문에 구면모형이 가장 적절하다. 경험적 변동도 패턴에 따라 계산한 너겟분산값은 0.018, 실 0.033, 레인지 198이므로 이를 기반으로 구면모형을 구축하면 아래와 같다.

$$\gamma(h) = \begin{cases} c_0 + c_1 \left\{ \frac{3h}{2a} - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right\} & \text{for } 0 < h < a \\ c_0 + c_1 & \text{for } h \geq a \end{cases}$$

a : 범위(range) c_0 : 너겟분산 $sill = c_0 + c_1$

이와 같은 변동도 추정과 모형은 내삽과 공간구조적 분석에 있어 매우 중요하다. 임의가 아닌 수학적 모형을 통해 각 거리 시차(h)내의 점들로부터 계산된 가중치를 적용한 가중치 최소제곱법(weighted least squares method)을 사용하기 때문이다. 즉 공간적 패턴 분석을 통하여 알아 낸 레인지(range)로 최적의 내삽을 하기 위한 거리정보를 얻는다. 또한 공분산 모형으로부터 예측할 분산을 최소화하여 주변의 점들로부터 조사되지 않은 점들을 추정할 수 있다.



[그림 6] 입지적 특성값의 구면 모형

4.2 입지적 특성값 추정을 위한 크리깅 내삽

크리깅은 자료점들 사이의 변화가 공간 거리에 따라 변화하는 비율을 기본으로 하는 방법이다. 내삽할 모수가 지역화 변수로 처리될 수 있다는 기본 가정에서 출발하여 공분산 모형로부터 예측할 분산을 최소화하는 선형 회귀식을 이용한다. 즉, 자료의 위치에 기초하여 가중치를 결정하는 것으로 그 절차는 IDW(inverse distance weighted interpolation) 내삽과 같으며 가중치는 변동도를 바탕으로 한 지리통계적(geostatistical) 공간분석으로부터 추출된다.

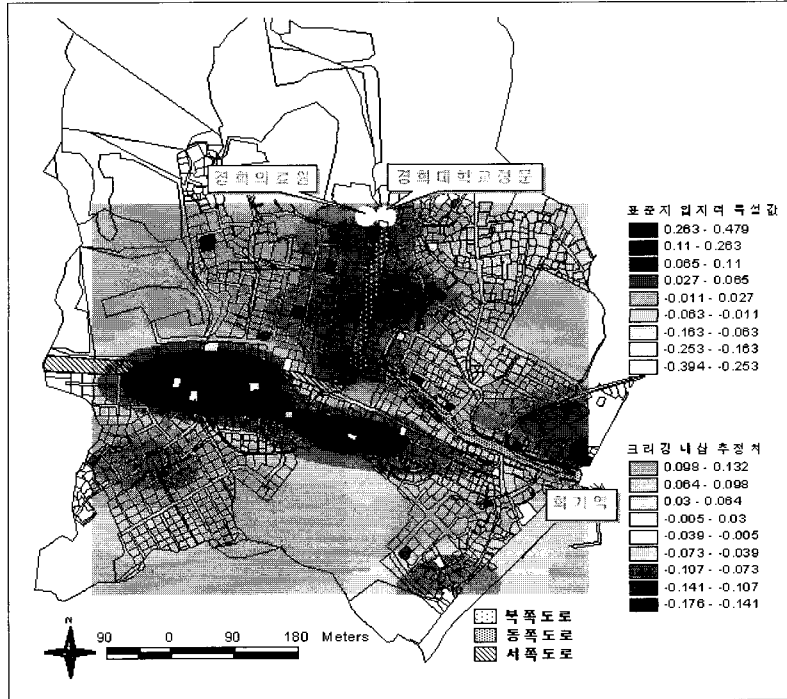
[그림 7]은 등방성 선형모형을 사용하여 표준지의 입지적 특성값을 내삽한 것이다. 연구지역의 추정된 값을 알아보면 최고값 0.132, 최저값 -0.175, 평균 -0.159,

표준편차 0.448이다. 연구지역을 그리드로 구축하여 내삽값을 바탕으로 등치선도로 변환한다.

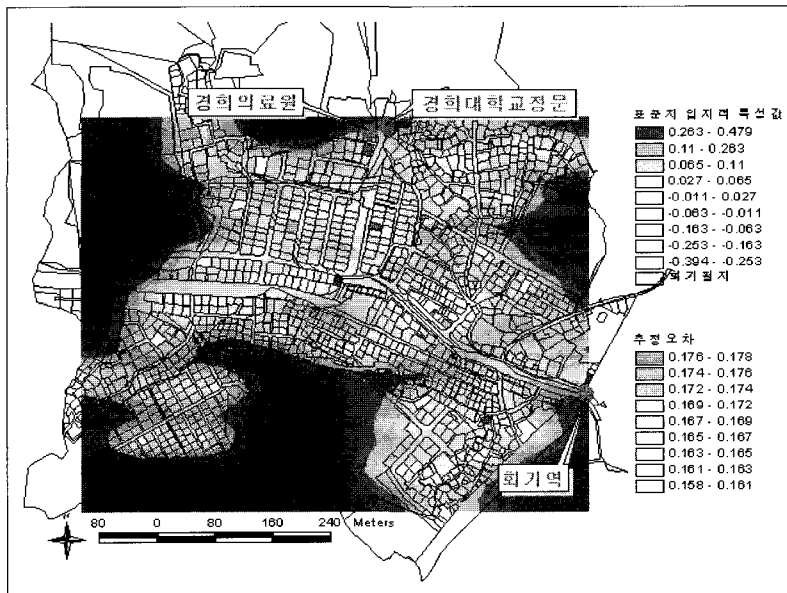
입지적 특성값이 높은 지역은 도로를 경계로 경희대학교와 의료원이 위치한 북쪽 필지들이 남쪽 필지들보다 높게 추정되었으며 동쪽 도로, 즉 회기역 방향의 필지들도 매우 높은 값을 갖는다. [그림 8]은 내삽값에 대한 크리깅 분산 즉 추정오차를 나타낸 것으로써 내삽의 신뢰성을 검증할 수 있게 한다.

이처럼 경희대학교, 의료원, 회기역과의 접근성 요인에 의해 입지적 특성값이 거리에 따라 차등적인 모습을 보인다. 도로와 인접한 필지들의 입지적 특성값이 일반적으로 높게 정의될 수 있으나 분석 결과 서쪽으로 주변의 입지적 특성값은 연구지역에서 가장 낮게 추정이 되었다.

공간통계분석을 이용한 지가의 입지값 측정에 관한 연구



[그림 7] 회기동 필지와 내삽추정치 중첩



[그림 8] 입지 특성값의 추정오차

또한 서쪽 도로와 동쪽 도로를 중심으로 한 필지들은 유사한 토지특성에도 불구하고 입지적 특성값의 상대적 차이가 크다. 이는 경희대학교, 의료원, 회기역과의 접근성에 따른 관계를 설명해 주고 있는 것이다.

이와 같은 분석에서처럼 크리깅은 통계이론과 많은 계산량에 의존하고 있는 진보된 기법이다. 원리상 크리깅은 이상적인 내삽 방법이며, 추정의 단계에서 국소적 추정치를 개선할 수 있는 지역의 크기와 형태를 파악할 수 있다. 부가적으로 내삽에 수반되는 오차의 추정치도 산출한다. 특히 국지적 변화가 존재할 때 크리깅(ordinary kriging) 내삽이 유용하다. 따라서 개별필지의 미시적인 입지 분석에 있어 큰 장점을 가지고 있다.

다음은 크리깅 내삽으로 추정된 값들을 입지적 변수로 가정하여 특성가격모형을 구축하고 모형들의 설명력을 비교한다.

4.3 공간적 특성가격의 모형

크리깅 내삽을 통해 추정된 입지적 특성값을 하나의 설명변수로 하여 모든 변수를 투입하는 입력방법으로 공간적 특성가격모형(spatial hedonic price model)을 구축한다.

우선 1421개의 필지내에 포함된 입지적 특성값만을 추출한다.⁶⁾ 즉 연구지역의 필지내에만 포함된 내삽 벡터좌표들을 추출한 결과 525개였다. 따라서 특성가격모형 구축은 이 525개 벡터자료의 내삽값을 기반으로 이루어진다. 이것은 지가에 대한 정확한 수치를 추정하는 것이 아니라 지가에 영향을 미치는 입지값의 수준을 위

치에 따라 추정하기 위한 것이다.

특성가격모형은 3가지로 구축한다. 첫째, 토지특성조사표에 제시된 모든 변수들 총 20개 중 13개 변수를 고려한 모형을 구축한다. 전체면적, 용도지역 1·2, 용도지구, 계획시설 구분, 토지이용상황, 지형의 고저·형상·방위, 도로접면, 유해철도와의 접근성, 표준지번호를 변수로 하여 모형 ②를 구축한다. 총 20개 변수 중 계획시설 구분과 계획시설저축율이 공선성이 있는 것으로 검증되어 계획시설저축율을 제외하고 지목, 농지, 임야, 비옥도, 경지정리, 유해폐기물과의 접근성, 기타 특성들은 모든 필지에 따라 같거나 해당 경우가 없기 때문에 제외하였다. 둘째 모형 ③은 물리적 특성만을 고려한 즉, 전체면적, 용도지역, 용도지구, 계획시설구분, 토지이용상황, 지형의 고저·형상·방위, 도로접면을 설명변수로 하여 구축한다. 마지막으로 물리적 특성변수와 추정치 즉 입지적 특성변수를 설명변수로 하여 모형 ④를 구축한다. 구축된 모형들은 아래와 같다.

$$\log \text{지가} = 3879388 + 326.5914 * \text{전체면적} + 48646.13 * \text{용도지역1} - 36655.96 * \text{용도지역2} + 33846.83 * \text{용도지구} - 98366.03 * \text{계획시설구분} - 392.32 * \text{토지이용상황} - 76056.3 * \text{고저} - 24603.86 * \text{형상} - 47833.6 * \text{방위} - 124172.2 * \text{도로접면} + 6026.855 * \text{유해철도접근성} - 1445.023 * \text{표준지번호} + \epsilon' \dots \dots \text{모형 ②}$$

$$\log \text{지가} = 2395682 - 1.472894 * \text{전체면적} + 39359.228 * \text{용도지역} + 32743.53 * \text{용도지구} - 63237.16 * \text{계획시설구분} + 557.0008 * \text{토지이용상황} - 69617.49 * \text{고저} - 26191.04 * \text{형상} - 48680.45 * \text{방위} - 123894.3 * \text{도로접면} + \epsilon' \dots \dots \text{모형 ③}$$

6) GeolProcessing Wizard extension를 이용한 것으로 한 필지내 입지적 특성값은 벡터좌표들의 평균값으로 대표한다.

$$\log \text{지가} = 2510781 - 0.7067437 * \text{전체면적} + 26523.54 * \text{용도지역} + 32002.7 * \text{용도지구} + 288993.39 * \text{계획시설구분} - 1701.17 * \text{토지이용상황} - 75119.97 * \text{고차} - 26952.61 * \text{형상} - 35586.59 * \text{방위} - 130509.9 * \text{도로접면} + 3540182 * \text{입지적 특성값} + \epsilon' \dots \dots \text{모형 ④}$$

각각의 특성가격모형의 적합도를 검증하기 위해 R^2 를 비교하였다. 즉 모형 ② 67%, 모형 ③ 66%, 모형 ④ 77% 등으로 선형모형에서 설명변수들이 종속변수인 지가의 변이도를 설명한 비율이다. 모형 ②와 모형 ③의 설명계수는 큰 차이를 보이지 않고 있는데 이는 물리적 변수로 선택한 토지의 특성들이 지가에 많은 영향을 주고 있다는 것을 의미한다. 특히 추정된 입지적 특성값을 하나의 설명변수로서 추가한 모형 ④는 모형 ②보다 R^2 의 값이 10%이상 높은 것으로 나타났다. 이는 잔차값의 공간적 상관성을 고려한 설명변수의 영향 때문이다.

5. 결 론

지가형성에 있어 입지가 가장 중요한 영향을 미친다는 중요성에도 불구하고 입지의 복잡한 역할과 공간적 상호작용으로 인한 객관적이고 체계적인 측정에 어려움이 있었다. 이러한 문제의식은 공간적 상호작용에 대한 인식과 공간적 의존관계를 고려한 공간분석기능의 필요성을 부각시킨다. 이에 본 연구에서는 GIS의 공간통계분석을 활용하여 입지적 특성값을 추정하도록 한다. 이러한 입지적 추정은 중첩이나 내삽과 같은 자료의 처리 과정을 통

해 수행될 수 있는데, 이 접근법은 지리통계학 자료를 처리할 때 많이 이용된다. 즉, 표본 추출된 54개 표준지의 벡터 자료점들과의 공간적 관련성에 기초하여 조사되지 않은 지역의 예측값을 확률적으로 평가할 수 있는 크리깅 분석방법을 이용한 것이다.

이러한 분석과정에서 탐색적 자료 분석 기법(ESDA)인 변동도 분석을 통해 구축된 정보는 경향면 내삽 모형을 구축할 때 입지 측정을 위한 대안적 접근 방법을 제시하고 있다. 즉 크리깅 내삽과정에서 확인할 수 있는 공간적 자기상관관계는 공간 의존성의 형성과정을 추정할 때 이점을 준다. 이러한 이점을 통해서 우리는 입지에 대한 긍정적 영향력(premium)과 부정적 영향력(penalty)을 고려할 수 있게 하여 지가평가에 적용이 가능할 수 있다. 또한 이러한 입지 추정은 측정된 입지적 특성을 중심으로 입지 특성이 유사한 인근지역의 경계범위 설정이나 개별필지의 지가 산정시 입지의 예측력을 높일 수 있다. 실제로 다양한 변동도 그래픽과 크리깅, 역거리 가중치(IDW) 기술뿐만 아니라 경향분석, 공간적 패턴 분석, 공간적 자기상관관계 분석 등과 같은 분석들은 시장의 변화성을 설명하기 위해 고려해야만 하는 근린지역의 요인들을 발견하는데 기여하게 될 것이다.

참 고 문 헌

Anselin, 1998, "Exploratory Spatial Data Analysis in a Geocomputational Environ-

- ment", in Longley, P.L., Brooks, S.M., McDonnell, R., and Macmillan, B. (eds.), *Geocomputation: A Primer*, John Wiley & Sons Ltd., pp. 77-94.
- Anselin, 1998, "GIS Research Infrastructure for Spatial Analysis of Real Estate Markets", *Journal of Housing Research* 9(1), pp. 113-133.
- Can, A., 1996, "Weight Matrices and Spatial Autocorrelation Statistics using a Topological Vector Data Model", *International Journal of GIS* Vol. 10(8), pp. 1009-1017.
- Can, A., 1998, "GIS and Spatial Analysis of Housing and Mortgage Markets", *Journal of Housing Research* Vol. 9(1), pp. 61-86.
- Castle III, H., 1998, *GIS In Real Estate: Integrating, Analyzing and Presenting Locational Information*, Appraisal Institute.
- Dubin, R.A., 1992, "Spatial autocorrelation and neighborhood quality", *Regional Science and Urban Economics*, Vol. 22, pp. 433-452.
- Dubin, R.A., 1998a, "Predicting House Prices Using Multiple Listings Data", *Journal of Real Estate Finance and Economics* Vol. 17(1), pp. 35-59.
- François, D.R., and Thériault, M., 1992, "Integrating GIS to Hedonic Price Modeling: An Application to the Quebec Region", *Property tax Journal* Vol. 11(1), pp. 50-51.
- François, D.R., 1999, "Sorting out access and Neighbourhood Factors In Hedonic Price Modelling", *Journal of Property Investment & Finance* Vol. 18(3), pp. 291-295.
- Gatrell, A.C., and Bailey, T.C., 1995, *Interactive spatial data analysis*, Longman Scientific & Technical, pp. 143-244.
- Mccluskey, W.J., and Deddis, W.G., 2000, "The Application of Surface Generated Interpolation Models for The Prediction of Residential Property Values", *Journal of Property Investment and Finance* Vol. 18(2), pp. 162-176.
- Rodriguez, M., Sirmans, C.F., and Marks, A.P., 1995, "Using GIS to improve real estate analysis", *The Journal of Real Estate Research* Vol. 10(2), pp. 163.
- Ward, R.D., Weaver, J.R., and German, J.C., 1999, "Improving CAMA Models Using GIS/Response Surface Analysis Location Factors", *Assessment Journal* Vol. 6(1), pp. 30-38.