

베이지안 추론방법을 이용한 소지역 주택매매가격지수 추정*

An Estimation of Housing Sales Price Index using Bayesian Approach to Small Area Model

우남교** · 권범준***

Woo, Nam Kyo · Kwon, Bum Jun

■ Abstract ■

Real estate price is measured by the area, location, accessibility and so on. Therefore even if the area is very small, these factors are effective to make difference of real estate price by areas. Recently the demand for small area estimation has greatly increased since that is important to formulate policies and programs, to the allocation of government funds and to regional planning. Generally, small area models are considered that include area-specific random effects (called small area effect). In this study, we consider a semiparametric specification of the Fay-Herriot model, which allows non linearities in the relationship between Y and the auxiliary variables X. We can estimate housing sales price index in seoul area with land price index as covariate variable from this model. The plan of this paper is as follows. In Chapter 2, we introduce small area model using Bayesian approach. In Chapter 3, we provide the result of the numerical study. Then we give our concluding remarks in Chapter 4.

Key Words : KHPI, Fay-Harriot model, Small area estimation, Bayesian, Semiparametric

* 본 논문은 2014년도 제1회 부동산통계 활용 및 발전을 위한 대학(원)생 논문 공모전 수상작(우수상)을 수정·보완한 논문임.

** 한국감정원 KAB부동산연구원 연구개발실 시장분석연구부 연구원 (주저자, k04614@kab.co.kr)

*** 한국감정원 KAB부동산연구원 부동산통계센터 주택통계부 대리 (공동저자, k25798@kab.co.kr)

1. 서 론

1. 연구의 배경

부동산은 우리나라 경제에서 차지하는 비중이 크다. 통계청(2013)에 따르면 우리나라 가계 자산 중 부동산 등의 실물자산이 차지하는 비중이 73.3%로 금융자산 26.7%의 3배에 달하고 있다. 이러한 상황 속에서 부동산 시장의 관심도가 점점 높아지고 있다. 높아지는 관심에 따라 최근 부동산 가격은 꾸준히 상승하고 있다. 따라서 정부 정책 또한 부동산 거래 활성화와 더불어 서민주거안정에 초점을 맞추고 있으며, 소득수준이 점점 높아짐에 따라 거주를 목적으로 인간의 생활 터전으로써의 부동산에서 투자를 목적으로 수익을 창출하는 부동산까지 다양한 가치로 변하고 있다. 저금리가 이어지면서 수익을 목적으로 한 부동산의 투자가 늘어나면서 부동산 관련 통계가 더욱 주목 받고 있다. 현재 다양한 부동산 관련 통계자료가 쏟아지고 있는 가운데 한국감정원에서 발표하고 있는 주택가격통계인 전국주택가격동향조사(KHPI), 토지가격통계인 전국지가변동률조사 등 가격 통계에 대한 국민의 관심도가 커지고 있다.

지리적 위치가 고정되어 있는 부동산은 입지 속성을 가지고 접근성에 따라서 용도가 결정되기 때문에 입지와 접근성에 대한 선호도를 유발하게 되고 그에 따른 가격을 형성하게 된다. 이러한 부동산의 속성 때문에 지역에 따라 가격의 차이가 발생하게 되고 이는 충분히 작은 소지역(small area)내에서도 그 영향을 미치게 된다. 앞서 말한 소지역이란 지리

학적인 지역을 지칭하며 이러한 소지역내에서는 동일한 지리적, 사회적 특성을 지니고 있다. 예를 들어 동일한 소득수준이 구분되어 지는 읍면동이나 명확히 특성이 구분되어 지는 시군구 등 동일한 소지역내 비슷한 속성을 가지고 있는 특징이 있다.

현재 주택가격통계인 전국 주택가격동향조사와 토지가격통계인 전국지가변동률조사는 시군구를 층화변수로 표본을 추출하여 조사되고 있어 다른 국가통계에 비교하여 상대적으로 소지역 단위로 발표되고 있다. 하지만 최근 부동산 시장은 과거보다 더욱 작은 단위(읍면동 혹은 아파트 단지 단위)에서 가격 및 동향의 차이가 발생하여 통계 공표 단위를 시군구에서 읍면동으로 확대되어야 한다는 목소리가 높아지고 있으나, 이를 위한 표본확대는 시간과 비용 등의 제약에 따른 효율성 저하로 어려움이 존재한다. 그러므로 이와 같은 소지역이 가진 속성을 직접 추정하는데 있어서 충분한 표본을 가지고 있지 않는 경우가 많아 추정값의 신뢰성과 정확성을 보장할 수 없다. 그렇기 때문에 소지역의 특성에 따라 추정하는 연구가 끊임없이 이루어지고 있고, 최근 몇 년간 이러한 소지역 추정 관련된 수요는 점점 늘어나고 있다. 소지역 추정치들은 국가 정책을 결정하고 수행하는데 활용되면서 중요성이 더욱 부각되고 있다.

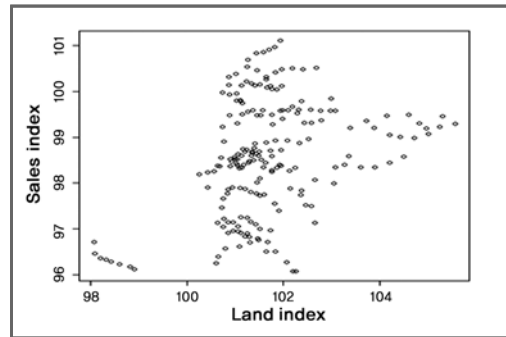
모형을 이용한 소지역 추정은 이용 가능한 변수가 타당한 정보를 제공하지 못하는 지역에 대해 믿을만한 추정치를 제공 가능하게 한다. 이러한 소지역 모형은 소지역 효과(small area effect)를 포함하고 있다. 소지역 효과는 보통

정규분포변수로 가정되고, 보조변수와 소지역 특성 효과는 소지역에 추정과 함께 선형적으로 이어져 있다. 이와 관련된 자세한 사항에 대해서는 Fay, Nelson and Litow(1993), Datta, Ghosh, Nangia and Natarajan(1996), Fay (1987), Ghosh, Nangia and Kim(1996) 등에서 확인 할 수 있다.

Fay-Herriot(1979) 모형은 가장 잘 알려진 소지역 추정 모형으로 설계모형(Design model)과 회귀분석 모형(regression model)의 조합으로 다른 지역의 특성을 이용(borrow strength)한다. Fay-Herriot 모형은 직접 조사로 추정된 추정값은 공변량의 선형함수를 가정한다. 하지만, 이 가정을 만족하지 않을 때는 소지역 모수 추정값이 편향될 수 있다. 이렇게 종속변수와 설명변수 간 비선형성을 가지는 경우, P-스플라인(Penalized-splines)을 이용한 준모수적 접근방법(Eilers and Marx, 1996)으로 해결할 수 있다.

〈그림 1〉은 서울특별시의 시군구별 지수로 주택매매가격지수와 지가지수를 그린 산점도이다. 보이는 것과 같이 뚜렷한 선형성을 확인할 수 없다. 따라서 본 연구에는 선형성을 가정하는 모수적인 방법과 비선형성을 가정하는 준모수적 방법을 적용하여 모델링을 한다. 본문 2장에서는 전통적인 소지역 모형의 정의와 준모수적 접근 방법을 소개하고 베이지안 추론을 이용하여 추정하는 방법을 제안하였다. 3장에서는 실제로 소개된 모형을 이용하여 주택매매가격지수와 지가지수를 적합 시킨 수치를 제공한다. 이때 모형이 비교 측정방법을 활용하여 적합한 모형을 찾아내고 실제 표본조사

에서 추정값의 상대표준오차(coefficient of validation)와 모형의 표준오차(Standard error)의 비교를 통하여 모형의 적합성을 확인해본다. 마지막 4장에서는 결론을 내리고 모형의 활용 방안과 시사점 등을 모색하겠다.



〈그림 1〉 서울특별시 주택매매지수와 지가지수

II. 소지역 모형과 계층적 베이지안 추론

1. 소지역(Small Area)

1) 소지역 추정방법

일반적으로 소지역 추정법은 크게 직접 추정법, 간접 추정법, 모형기반 추정법으로 나눌 수 있다. 먼저 직접 추정법은 소지역의 표본 조사를 실시하여 직접 추정하는 방식이다. 소지역의 크기에 따라 표본의 개수가 늘어나는 경향이 있기 때문에 소지역을 추정하기에 충분한 표본을 추출한다면 전체지역을 생각할 때 조사에 대한 비용이나 자원 등의 면에서 비효율적이다. 반대로 전체지역을 고려하여 표본을 추출한다면 소지역에 배정되는 표본의 개수가 적어 정도 높은 추정을 할 수 없다.

두 번째 방법은 간접 추정법이다. 간접추정법은 다시 인구통계적 방법, 합성 추정법(Synthetic Estimator), 복합추정법(Composite Estimator) 등으로 나눌 수 있다. 인구통계적 방법은 가령 5년마다 하는 센서스에서 중간 해당년도의 인구를 추정하기 위해 사용하는 방법으로 출생수, 사망수 등을 활용하여 인구를 추정하는 방법이다. 합성 추정법은 큰 지역을 기준으로 표본설계를 하였을 때 큰 지역과 소지역의 구조적 특성이 같다는 조건하에서 추정 하는 방법이고 복합 추정법은 직접추정한 불안정한 추정량을 합성추정량과 가중평균을 이용하여 추정하는 방법이다.

세 번째 방법은 모형 기반 추정법이다. 이 추정법은 소지역 내 포함하고 있는 공변량으로 지역 수준 모형을 적합시켜 추정하는 방식으로 다음과 같은 장점이 있다. 첫째, 소지역 간의 복잡한 오차구조를 내포하고 있기 때문에 소지역 간의 변동을 반영하여 소지역 추정의 정도를 높일 수 있다. 둘째, 표본 자료로부터 모형의 유용성을 확인할 수 있고, 범주형 자료 및 시계열 자료 등 대부분의 자료형태를 모형화 하여 추론 할 수 있다.

본 연구에서는 모형 기반 추정법인 계층적 베이즈(HB) 방법을 이용하여 비교적 추론의 정확도가 높고 복잡한 유형의 문제들도 해결할 수 있는 MCMC(Monte Carlo Markov Chain) 방법을 활용하였다.

2) 소지역 모형

일반적으로 관심모수인 $\theta = (\theta_1, \dots, \theta_m)^T$ 는 소지역의 평균으로 가정하며 여기서 m 은 소

지역의 개수이다. $y = (y_1, \dots, y_m)^T$ 는 직접 조사되어진 추정치로 i 번째 소지역과 관련 있는 반응변수이다.

$x_i = (1, x_{i1}, \dots, x_{ip})^T$ 는 y_i 와 관련 있는 공변량 벡터로 가정한다. 그러면 지역효과 u_i 가 포함된 소지역 모형을 다음과 같이 표현할 수 있다.

$$Y_i = \theta_i + e_i \quad \text{식(1)}$$

$$\theta_i = X_i^T \beta + u_i \quad \text{식(2)}$$

여기에 β 는 회귀 계수 벡터이고, 이때 u_i 는 지역 랜덤 효과를 나타내며 평균이 0이고 분산이 σ_u^2 인 정규분포를 따르고, e_i 는 모형 오차항으로 평균이 0이고 분산이 σ_e^2 인 정규분포를 따른다고 가정한다. e_i 와 u_i 는 독립임을 가정한다.

Fay and Herriot(1979)은 식(1)과 식(2)를 선형 혼합 하여 식(3)과 같은 소지역 모형으로 표현 할 수 있다.

$$Y_i = X_i^T \beta + u_i + e_i \quad \text{식(3)}$$

이 모형의 경우 추정값 Y_i 와 공변량 X_i 가 선형관계가 있어야 활용할 수 있다.

2. 준모수적 소지역 모형

위의 Fay and Herriot(1979)에 의해 제안된 모형은 추정값과 공변량 사이에 선형관계를 만족하지 않을 경우 편향(biased)이 발생하게 된다. 따라서 설명변수와 종속변수가 비선형인

경우에는 P-스플라인(Penalized-splines)을 이용하여 모형의 준모수적 접근방법(Eilers and Marx, 1996)을 사용한다.

기본적인 준모수적 모형은 식(4)와 같다.

$$Y_{ij} = f(x_{ij}) + b_i + u_{ij} + e_{ij} \quad \text{식(4)}$$

이때 $f(x_{ij})$ 는 공변량들의 함수로 구체화되지 않은 함수이다. 설명변수와 반응변수 간에 뚜렷한 관계가 없는 경우에 활용한다. b_i 는 지역 랜덤 효과로 $b_i \sim N(0, \sigma_b^2)$ 를 가정하고, u_{ij} 와 e_{ij} 는 각각 i 번째 지역과 j 번째 시간 상호작용 효과와 오차항을 나타내며 $u_{ij} \sim N(0, \psi_j^2)$, $e_{ij} \sim N(0, \sigma_{ij}^2)$ 를 가정한다.

1) 절단된 다항식 기저함수

P스플라인에서 절단된 다항식 기저함수(Truncated Polynomial Basis Function; TPBF)를 이용하여 $f(x_{ij})$ 를 근사시켜 준모수적 모형을 식(5)와 같이 표현 할 수 있다.

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1 x_{ij} + \dots + \beta_p x_{ij}^p \quad \text{식(5)} \\ &+ \sum_{k=1}^K \gamma_k (x_{ij} - \tau_k)^p + b_i + u_{ij} + e_{ij} \\ &= X_{ij}^T \boldsymbol{\beta} + Z_{ij}^T \boldsymbol{\gamma} + b_i + u_{ij} + e_{ij} \\ &= \theta_{ij} + e_{ij} \end{aligned}$$

이때 $X_{ij} = (1, x_{ij}, \dots, x_{ij}^p)^T$ 와, $Z_{ij} = \{(x_{ij} - \tau_1)_+, \dots, (x_{ij} - \tau_K)_+\}^T$ 는 공변량 이고 이때 τ_k 는 중심계수로 절단된 다항식 기저함수의 정의이다. 또한, $\boldsymbol{\beta} = (\beta_0, \dots, \beta_p)^T$ 는 회귀계수

벡터, $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_K)^T$ 는 스플라인 계수 벡터이며 $i=1, \dots, m$, $j=1, \dots, t$ 이다. 반응변수와 설명변수들이 측정된 지역의 수를 m , 측정된 시간의 수를 t 라 한다. 위에서 p 차수의 스플라인 모델은 어떤 구체화되지 않은 평활화(smoothing) 함수를 활용하여 적절하게 근사 할 수 있다. 일반적으로, 선형 또는 이차식 모형이 적합한 커브에서 적절한 평활화 효과를 나타내어 좋은 결과를 제공한다.

3. 계층적 베이지안 추론

본 모형의 목적은 소지역 특성치 $\boldsymbol{\theta}$ 를 추정하는 것이고, i 번째 지역의 반응변수를 $\mathbf{y}_i = (y_{i1}, \dots, y_{it})^T$, 공변량은 $\mathbf{X}_i = (X_{i1}, \dots, X_{mt})^T$ 와 $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{mt})^T$ 라 하자. 위 식을 이용하여 다음과 같은 계층적 베이지안 모형을 정의 할 수 있다.

$$\begin{aligned} y_{ij} | \theta_{ij}, \sigma_{ij}^2 &\sim N(\theta_{ij}, \sigma_{ij}^2) \\ \theta_{ij} | \mathbf{X}_{ij}, \boldsymbol{\beta}, \mathbf{Z}_{ij}, \boldsymbol{\gamma}, b_i, \psi_j^2 &\sim N(\mathbf{X}_{ij}^T \boldsymbol{\beta} + \mathbf{Z}_{ij}^T \boldsymbol{\gamma} + b_i, \psi_j^2) \\ b_i &\stackrel{iid}{\sim} N(0, \sigma_b^2) \\ \boldsymbol{\gamma} | \sigma_\gamma^2 &\sim N(\mathbf{0}, \sigma_\gamma^2 \mathbf{I}) \\ u_{ij} &\sim N(0, \psi_j^2) \\ e_{ij} &\sim N(0, \sigma_{ij}^2) \end{aligned}$$

모수공간은 $\boldsymbol{\Omega}_i = (\boldsymbol{\theta}_i, \boldsymbol{\beta}, \boldsymbol{\gamma}, b_i, \psi_i, \sigma_b^2, \sigma_\gamma^2)$ 이고, 전체 모수 공간은 $\boldsymbol{\Omega} = \boldsymbol{\Omega}_1 \times \dots \times \boldsymbol{\Omega}_m$ 이다.

1) 우도함수

준모수적 모델의 우도함수는 식(6)과 같다.

$$L(\mathbf{y}_i, \mathbf{X}_i, \mathbf{Z}_i | \Omega_i) \propto L(\mathbf{y}_i | \boldsymbol{\beta}, \boldsymbol{\gamma}, b_i, \psi^2, \mathbf{X}_i, \mathbf{Z}_i) \quad \text{식(6)}$$

$$\times L(b_i | \sigma_b^2) L(\boldsymbol{\gamma} | \sigma_\gamma^2)$$

여기서, $L(U|a, b)$ 는 평균이 a , 분산이 b 인 정규분포의 확률밀도함수를 나타내며, $L(b_i | \sigma_b^2)$ 와 $L(\boldsymbol{\gamma} | \sigma_\gamma^2)$ 는 평균이 0이고 분산이 각각 σ_b^2 , σ_γ^2 인 정규분포를 나타낸다.

2) 사전분포(Prior distribution)

베이지안 프레임에서는 모수에 대한 사전 분포 정의가 필요하다. 사전적인 정보가 없을 경우 무정보적 사전 분포(noninformative prior)를 활용하고, 있는 경우 공액사전분포(conjugate prior)를 활용한다. 따라서 다항 계수 $\boldsymbol{\beta}$ 에 대한 사전분포로는 정보가 없는 균등분포를 정의 하였고, 각각 분산의 역수는 감

마분포를 사전분포로 활용하였다. 이러한 사전분포는 상호간에 독립을 가정한다. 즉, 사전 분포는 다음과 같다.

$$\boldsymbol{\beta} \sim \text{Uniform}(\mathbf{R}^{(p+1)})$$

$$(\psi_j^2)^{(-1)} \sim \text{Gamma}(c_j, d_j)$$

$$(\sigma_b^2)^{(-1)} \sim \text{Gamma}(c, d)$$

$$(\sigma_\gamma^2)^{(-1)} \sim \text{Gamma}(c_\gamma, d_\gamma)$$

$X \sim \text{Gamma}(a, b)$ 는 형상 모수 a , 비율 모수 b 를 가지는 감마분포를 나타내며, $f(x) \propto x^{a-1} \exp(-bx)$, $x \geq 0$ 이다.

3) 사후분포 및 추론

앞서 설정된 계층적 베이지안 모델을 이용하여 자료가 주어졌을 때 사후분포는 우도함수와 사전분포의 곱 형태인 식(7)과 같다.

$$p(\boldsymbol{\Omega} | \mathbf{y}, \mathbf{X}, \mathbf{Z}) \quad \text{식(7)}$$

$$\propto \prod_{i=1}^m L(\mathbf{y}_i, \mathbf{X}_i, \mathbf{Z}_i | \Omega_i) \pi(\boldsymbol{\beta}) \pi(\sigma_b^2) \pi(\sigma_\gamma^2) \prod_{j=1}^t \pi(\psi_j^2)$$

〈표 1〉 모수별 조건부 분포

모 수	조건부 분포
$[\theta_{ij} \boldsymbol{\beta}, \boldsymbol{\gamma}, \psi^2, b, \mathbf{X}, \mathbf{Z}]$	$N[(\frac{1}{\sigma_{ij}^2} + \frac{1}{\psi_j^2})^{-1}(\frac{y_{ij}}{\sigma_{ij}^2} + \frac{(\mathbf{X}_{ij}'\boldsymbol{\beta} + \mathbf{Z}_{ij}'\boldsymbol{\gamma} + b_i)}{\psi_j^2}), (\frac{1}{\sigma_{ij}^2} + \frac{1}{\psi_j^2})^{-1}]$
$[b_i \boldsymbol{\beta}, \boldsymbol{\gamma}, \theta, \psi^2, \sigma_b^2, \mathbf{X}, \mathbf{Z}]$	$N[(\frac{1}{\sigma_b^2} + \sum_{j=1}^t \frac{1}{\psi_j^2})^{-1}(\sum_{j=1}^t \frac{1}{\psi_j^2}(\theta_{ij} - \mathbf{X}_{ij}^T \boldsymbol{\beta} - \mathbf{Z}_{ij}^T \boldsymbol{\gamma})), (\frac{1}{\sigma_b^2} + \sum_{j=1}^t \frac{1}{\psi_j^2})^{-1}]$
$[\boldsymbol{\beta} \boldsymbol{\gamma}, \theta, b, \psi^2, \mathbf{X}, \mathbf{Z}]$	$N[(\sum_{i=1}^m \sum_{j=1}^t \frac{X_{ij} X_{ij}^T}{\psi_j^2})^{-1}(\sum_{i=1}^m \sum_{j=1}^t \frac{X_{ij}}{\psi_j^2}(\theta_{ij} - \mathbf{Z}_{ij}^T \boldsymbol{\gamma} - b_i)), (\sum_{i=1}^m \sum_{j=1}^t \frac{X_{ij} X_{ij}^T}{\psi_j^2})^{-1}]$
$[\boldsymbol{\gamma} \boldsymbol{\beta}, \theta, b, \psi^2, \sigma_\gamma^2, \mathbf{X}, \mathbf{Z}]$	$N[(\sum_{i=1}^m \sum_{j=1}^t \frac{Z_{ij} Z_{ij}^T}{\psi_j^2} + \frac{1}{\sigma_\gamma^2} \mathbf{I})^{-1}(\sum_{i=1}^m \sum_{j=1}^t \frac{Z_{ij}}{\psi_j^2}(\theta_{ij} - \mathbf{X}_{ij}^T \boldsymbol{\beta} - b_i)), (\sum_{i=1}^m \sum_{j=1}^t \frac{Z_{ij} Z_{ij}^T}{\psi_j^2} + \frac{1}{\sigma_\gamma^2} \mathbf{I})^{-1}]$
$[(\sigma_\gamma^2)^{-1} \boldsymbol{\gamma}]$	$\text{Gamma}(\frac{K}{2} + c_\gamma, \frac{1}{2} \boldsymbol{\gamma}^T \boldsymbol{\gamma} + d_\gamma)$
$[(\psi_j^2)^{-1} \boldsymbol{\beta}, \boldsymbol{\gamma}, \theta, b, \mathbf{X}, \mathbf{Z}]$	$\text{Gamma}(c_j + \frac{m}{2}, \frac{1}{2} \sum_{i=1}^m (\theta_{ij} - \mathbf{X}_{ij}^T \boldsymbol{\beta} - \mathbf{Z}_{ij}^T \boldsymbol{\gamma} - b_i)^2 + d_j)$
$[(\sigma_b^2)^{-1} b]$	$\text{Gamma}(\frac{m}{2} + c, \frac{1}{2} \sum_{i=1}^m b_i^2 + d)$

θ_{ij} 의 주변 사후 분포는 분석적으로 매우 복잡하기 때문에 고차원의 적분이 필요하게 되며 모수 추정의 문제는 매우 어렵다. 그러나 베이저안 추론 방법을 활용하여 다른 연관 모수들이 주어졌을 때 θ_{ij} 의 조건 분포를 활용하면 쉽고 정확한 추정이 가능하다. 위에서 설정된 사후분포를 활용하여 각 모수의 조건부 분포를 구해보면 <표 1>과 같다. 조건부 분포를 구하는 과정은 부록(Appendix)으로 남겨 둔다.

모든 분포가 각각 표준분포를 따르므로 깁스 샘플링(Gibbs sampler)을 활용하여 모수를 추정할 수 있다.

III. 데이터를 이용한 소지역 모형 적합

1. 데이터 설명

데이터분석에 사용된 자료는 한국감정원이 작성하고 있는 주택가격동향조사의 주택매매 가격지수와 전국 지가변동률 조사의 지가지수이다. 두 가지의 통계 모두 국민의 관심도가 높은 자료로 주택가격동향조사는 전국 주택 매매가격 및 전세가격의 변동과 시장동향을 조사, 분석하여 정책 수립 등에 활용되는 통계로 1986년에 국민은행에서 시작되어 현재 한국감정원이 조사, 작성 하고 있다. 주택매매 가격지수는 매월 206개 시군구를 조사, 발표하는 대규모 조사이다. 전국지가변동률 조사의 지가지수는 1975년에 시작되어 현재 한국감정원이 조사, 작성하고 있는 통계로 전국 지가의 변동 상황을 조사하여 토지정책 수립의 기

초자료로 제공된다. 지가지수 또한 매월 252개 시군구를 조사, 발표하는 대규모 조사이다.

본 분석에 사용된 자료는 2014년 1월부터 2014년 9월까지 공표된 서울특별시의 시군구별 지수를 수치분석에 활용하였다. 2014년 1월부터 8월까지의 모형적합에 활용하였고, 9월의 지가 지수를 활용하여 주택매매가격지수를 추정하는 방식으로 수치분석을 실시하였다. 두 가지 지수의 관계는 <그림 1>을 통해 확인 할 수 있다. 두 지수간에 뚜렷한 선형관계를 확인 할 수 없다. 그렇기 때문에 앞서 설명한 선형관계를 가정하는 소지역 모형과 P-스플라인을 이용한 준모수적 모형 두 가지 모형이 모두 적합 하였다. 특히 준모수적 모형에서는 노트의 개수를 0,3,5개로 조절한 모형 3가지를 고려 하였다.

2. 모형 비교척도

앞에서 세운 모형 중 어떠한 모형이 적합한 지 확인하기 위하여 추정 값의 차이를 비교하는 4가지의 비교척도(Comparative Measures)를 사용하였다. ARB척도는 i 번째 지역의 실제 조사되어진 값(c_i)과 모형을 적합시킨 추정

<표 2> 모형의 비교척도

비교방법	척도(Measures)
• Average Relative Bias(ARB)	$\frac{1}{m} \sum_{i=1}^m \frac{ c_i - e_i }{c_i}$
• Average Squared Relative Bias(ASRB)	$\frac{1}{m} \sum_{i=1}^m \frac{(c_i - e_i)^2}{c_i^2}$
• Average Absolute Bias(AAB)	$\frac{1}{m} \sum_{i=1}^m c_i - e_i $
• Average Squared Deviation(ASD)	$\frac{1}{m} \sum_{i=1}^m (c_i - e_i)^2$

값(e_i)의 차이를 실제값으로 나눈 비율의 평균이고, ASRB 척도는 ARB척도의 비율을 제공한 평균, AAB, ASD 척도는 각각 차이에 대한 절대값의 평균, 제곱의 평균이며 척도 값이 낮을 수록 적합한 모형이라 할 수 있다.

3. 수치분석

서울지역의 25개 시군구의 2014년 1월부터 9월까지의 주택 종합유형의 매매가격지수와 지가지수 자료를 활용하여 베이지안 추론을 실시하였다. 본 분석에는 4가지의 모형을 설정하였다. 모형1은 선형을 가정한 전통적인 소지역 모형, 모형2는 비선형인 준모수적 방법을 이용하여 TPBF를 기저함수로 사용하고 노트(knot)를 3개로 정의한 모형, 모형3은 TPBF를 기저함수로 사용하고 노트(knot)를 5개로 정의한 모형, 모형4는 TPBF를 기저함수로 사용하고 노트를 정의하지 않은 모형으로 각각 설정하였다. 노트는 비선형 모형에서 구분경계로 경계 결정은 자료의 분위수를 활용하여 결정하였다.

모수추정을 위한 시뮬레이션은 깁스표집기

에서 20,000번의 반복추출을 실시하여 15,000번을 burn-in 한 후, 모수를 추론하였다.

분석을 위하여 주어진 자료에 대한 변수 설정은 y_{ij} 는 i 번째 시군구 j 월의 주택 종합 매매지수 ($i=1, \dots, 25, j=1, \dots, 8$)이고, x_{ij} 는 i 번째 시군구 j 월의 지가지수 ($i=1, \dots, 25, j=1, \dots, 8$) 이다.

모형1~4의 비교를 위해 정의된 모형 비교 측정도구를 활용하여 계산한 결과가 <표 3>에 나타나 있다. 모형1은 종속변수와 설명변수의 선형성이 만족하는 경우 사용하는 모형이므로 다른 모형에 비해 높은 값이 나왔고, 모형4는 노트를 설정하지 않아 높은 값으로 나타났다. 노트를 설정한 모형2와 모형3을 비교한 결과 유사하게 나와 모형3을 이용하여 소지역 모델링을 실시하였다. 모형 3으로 적합한 사후평균들은 아래에서 확인 할 수 있다.

<표 3> 모형1~모형4 모형 비교측정 결과

모형	ARB	ASRB	AAB	ASD
모형1(FH model)	0,0107	0,0002	1,0527	1,6289
모형2(3 knot)	0,0080	0,0001	0,7884	1,0069
모형3(5 knot)	0,0081	0,0001	0,8048	0,9405
모형4(w/o knot)	0,0103	0,0001	1,0144	1,4099

<표 4> 모형3에 대한 모수의 사후평균

Parameter	사후평균	S.E.	C.I.	
			2.5%	97.5%
β_0	57,0006	0,00040	56,9999	57,0014
β_1	0,4119	0,00000	0,4119	0,4119
γ_1	0,4832	0,00005	0,4831	0,4833
γ_2	-0,8671	0,00012	-0,8673	-0,8669
γ_3	-0,4043	0,00013	-0,4045	-0,4040
γ_4	0,2272	0,00004	0,2270	0,2274
γ_5	0,4442	0,00004	0,4442	0,4443

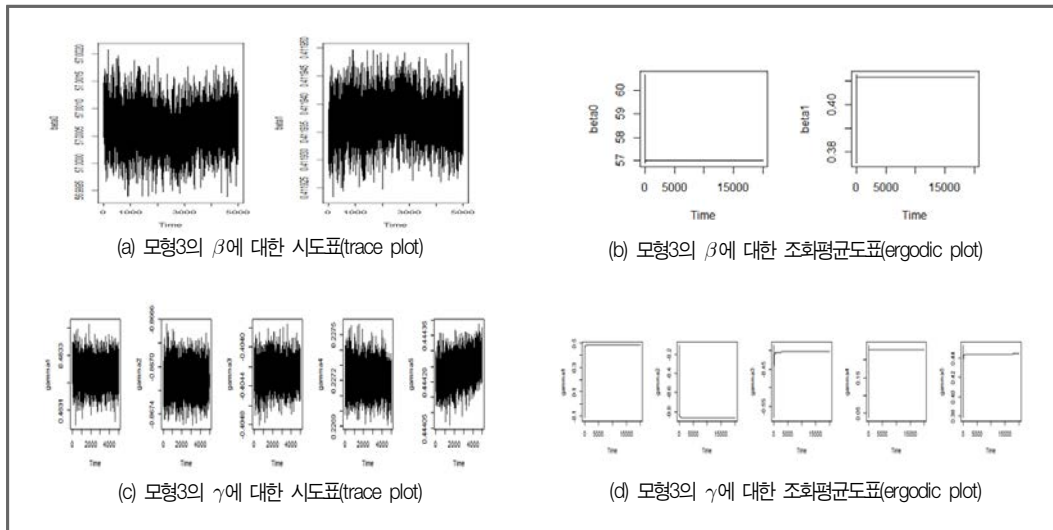
〈표 5〉 모형3에 의한 주택매매가격지수(추정치)

구분	시군구	'14. 1	2	3	4	5	6	7	8
추 정 치	종로	99.7	99.7	99.6	99.5	99.5	99.6	99.6	99.6
	중	98.9	99.2	99.2	99.3	99.2	99.1	99.1	99.1
	용산	96.7	96.7	96.8	96.8	96.9	97.0	97.1	97.1
	성동	97.7	97.8	98.0	98.2	98.2	98.2	98.2	98.2
	광진	96.6	96.7	96.7	96.8	96.9	97.0	97.0	96.9
	동대문	97.9	98.0	98.0	98.2	98.3	98.3	98.3	98.3
	중랑	100.0	100.1	100.2	100.3	100.3	100.3	100.2	100.2
	성북	97.5	97.6	97.6	97.8	97.9	97.9	98.0	98.0
	강북	98.1	98.1	98.4	98.5	98.5	98.5	98.4	98.4
	도봉	98.8	98.9	98.9	99.0	99.1	99.1	99.2	99.2
	노원	98.3	98.3	98.4	98.5	98.6	98.6	98.6	98.7
	은평	99.1	99.1	99.2	99.1	99.0	99.0	98.9	98.9
	서대문	98.1	98.2	98.3	98.5	98.5	98.5	98.4	98.3
	마포	99.4	99.5	99.4	99.3	99.3	99.3	99.2	99.3
	양천	97.3	97.4	97.4	97.4	97.3	97.3	97.3	97.3
	강서	98.4	98.4	98.3	98.3	98.2	98.2	98.2	98.3
	구로	98.9	99.0	99.2	99.2	99.2	99.1	99.0	99.0
	금천	98.7	98.8	98.9	98.9	98.8	98.7	98.7	98.7
	영등포	96.9	97.0	97.2	97.3	97.3	97.2	97.2	97.1
	동작	99.5	99.6	99.9	99.9	99.8	99.8	99.8	99.7
	관악	99.2	99.4	99.5	99.5	99.5	99.4	99.4	99.4
S.E.	서초	97.8	97.8	97.9	98.0	98.1	98.2	98.3	98.4
	강남	98.9	99.0	99.2	99.3	99.3	99.4	99.5	99.6
	송파	99.6	99.7	99.9	100.0	100.1	100.1	100.2	100.3
	강동	100.1	100.0	100.0	100.0	100.0	100.0	100.0	100.1
	종로	0.1655	0.1655	0.1655	0.1654	0.1653	0.1657	0.1657	0.1656
	중	0.0924	0.0925	0.0923	0.0922	0.0922	0.0922	0.0923	0.0923
	용산	0.1314	0.1315	0.1312	0.1315	0.1313	0.1312	0.1312	0.1313
	성동	0.1657	0.1657	0.1658	0.1657	0.1654	0.1656	0.1659	0.1657
	광진	0.1808	0.1810	0.1809	0.1810	0.1807	0.1810	0.1810	0.1808
	동대문	0.0678	0.0676	0.0676	0.0677	0.0677	0.0677	0.0677	0.0674
	중랑	0.0745	0.0746	0.0745	0.0745	0.0744	0.0745	0.0747	0.0748
	성북	0.1944	0.1945	0.1946	0.1942	0.1943	0.1946	0.1943	0.1947
	강북	0.0590	0.0586	0.0590	0.0590	0.0587	0.0587	0.0589	0.0588
	도봉	0.0697	0.0696	0.0697	0.0696	0.0697	0.0696	0.0695	0.0693
	노원	0.0777	0.0780	0.0779	0.0779	0.0779	0.0778	0.0778	0.0777
	은평	0.1673	0.1673	0.1674	0.1674	0.1671	0.1674	0.1673	0.1674
	서대문	0.1261	0.1262	0.1261	0.1261	0.1262	0.1262	0.1264	0.1261
	마포	0.1772	0.1773	0.1770	0.1770	0.1771	0.1771	0.1772	0.1772
	양천	0.1479	0.1477	0.1480	0.1474	0.1477	0.1476	0.1479	0.1478
	강서	0.1036	0.1036	0.1033	0.1036	0.1035	0.1035	0.1034	0.1033
	구로	0.1296	0.1295	0.1297	0.1294	0.1294	0.1296	0.1295	0.1294
	금천	0.0701	0.0702	0.0702	0.0705	0.0703	0.0701	0.0702	0.0702
	영등포	0.1067	0.1068	0.1067	0.1067	0.1067	0.1069	0.1066	0.1069
	동작	0.1002	0.1004	0.1007	0.1006	0.1002	0.1005	0.1005	0.1004
	관악	0.1507	0.1507	0.1505	0.1508	0.1507	0.1507	0.1511	0.1509
	서초	0.1448	0.1449	0.1448	0.1447	0.1449	0.1449	0.1448	0.1447
	강남	0.1708	0.1708	0.1708	0.1709	0.1709	0.1708	0.1708	0.1708
	송파	0.1457	0.1460	0.1460	0.1459	0.1457	0.1458	0.1460	0.1458
	강동	0.1723	0.1725	0.1724	0.1724	0.1724	0.1723	0.1723	0.1723

〈표 4〉는 모형3에 대한 사후분포의 평균과 S.E., 95% C.I 모수 추정치들이 나타나있고, 〈그림 2〉는 모형의 수렴성 진단을 위한 β 와 γ 의 시도표(Trace plot)와 조화평균 도표(Ergodic mean plot)이다. 두 가지의 모수 모두 시도표

는 일정한 패턴이 없이 정상(stationary)분포로 수렴하고, 조화평균 도표도 안정한 값을 보임에 따라 수렴하는 것을 확인 할 수 있다.

위에서 산출된 모형을 이용하여 추정된 주택 매매가격지수의 추정치들이 〈표 5〉에 나타나



〈그림 2〉 모형3에 대한 모수의 사후평균

〈표 6〉 '14. 9월의 주택매매가격지수(참값)와 모형 추정치 비교

시군구	실제조사		모형추정		시군구	실제조사		모형추정	
	지수	CV	지수	CV		지수	CV	지수	CV
종로	99.7	0.1275	99.7	0.1658	마포	100.7	0.4014	99.4	0.1784
중	99.9	0.4590	99.1	0.0930	양천	96.7	0.5254	97.2	0.1520
용산	96.1	0.4812	96.9	0.1354	강서	97.2	0.4843	98.3	0.1053
성동	98.5	0.7742	97.8	0.1694	구로	97.6	0.4972	98.8	0.1312
광진	96.8	0.6146	96.6	0.1872	금천	98.5	0.6018	98.6	0.0712
동대문	97.9	0.5296	97.8	0.0691	영등포	98.6	0.4362	96.9	0.1100
중랑	100.1	0.4847	99.9	0.0749	동작	101.3	0.3676	99.5	0.1010
성북	98.6	0.2707	97.5	0.1997	관악	100.2	0.2101	99.2	0.1519
강북	97.0	0.3881	98.1	0.0600	서초	98.9	0.3823	98.5	0.1471
도봉	97.0	0.4571	98.7	0.0706	강남	100.1	0.7208	99.8	0.1721
노원	99.9	0.4188	98.1	0.0793	송파	99.6	0.6592	100.4	0.1452
은평	99.2	0.7241	99.0	0.1690	강동	100.6	0.4521	100.2	0.1720
서대문	98.8	0.3338	98.1	0.1285					

있다. 매월 추정치에 대한 상대표준오차(CV)가 안정적으로 나와 추정의 정도가 높아 졌음을 확인 할 수 있다. 이는 <표 6>의 '14년 9월에 실제 조사에 의해 산출된 지수에 대한 상대표준오차와 소지역 모형에 의한 추정된 지수에 대한 상대표준오차를 비교해 보면 명확히 확인 할 수 있다. 다만, 추정 값과 참값의 차이가 조금은 크게 나타난 지역은 주변지역의 영향으로 인해 공간적평활화(spatial smoothing)된 결과로 해석된다. 또한 공변량으로 사용되어진 지가지수의 경우 측정오차가 발생하기 때문에 좀 더 정확한 값으로 대체를 한다면 좀 더 좋은 결과를 얻을 수 있을 것으로 기대된다.

IV. 결 론

1. 결과 및 시사점

본 연구는 주택매매가격 지수와 지가지수를 소지역 모형에 적합 시켜 지가지수를 이용하여 주택매매가격 지수를 추정해 보았다. 앞의 결과에서 확인 할 수 있듯이 선형성을 만족하지 않는 두 가지 가격지수의 소지역 모델링을 위해서는 준모수적 방법인 P-스플라인을 활용한 노트를 설정한 TPBF 모형이 가장 적합한 모형으로 나타났다. 모형을 적합 시켜 공간적 평활화가 나타나지만 추정 값들이 실제 조사가격에 비해 안정적으로 나타나는 것으로 확인할 수 있다. 다만, 정확한 참값을 공변량으로 활용해야 하지만 측정오차가 존재하는 지가지수를 활용하였기 때문에 그 측정오차를

고려한 모형은 차후에 과제로 남겨 두었다.

가격 지수의 소지역 추정은 세부시장 정보가 필요한 부동산 시장을 분석하는데 반드시 필요하다. 이는 정부의 부동산 정책을 뒷받침해주는 역할 뿐만 아니라 부동산이 가지고 있는 속성인 입지에 따른 가격의 차이를 확인 할 수 있는 중요한 역할을 한다. 특히 현재 시군구별 가격통계가 나오지 않는 월세가격 동향조사, 상업용 임대동향조사 등은 시도별 혹은 상권별 가격통계에 의존할 수밖에 없다. 하지만 본 연구에서 알 수 있듯이 시군구별 가격통계가 생산되는 주택가격동향조사의 매매 가격지수, 전세 가격지수, 전국지가변동률조사의 지가지수 등과 같이 관련이 있는 공변량을 찾을 수 있는 경우 위의 방법과 같이 모형화하여 시군구별 지수를 추정해 낼 수 있을 것이다. 또한, 상호 관련이 있는 가격지수들의 소지역 모델링을 통하여 조사가격에 대해 상호 검정하고 특이한 지역에 대해 검증 과정을 거쳐 조사 정확도를 제고 할 수 있는 척도가 될 수 있을 것이다.

또한, 지가지수는 공표지역이 시군구보다 하위시장인 읍면동으로 확대되었다. 이는 동일한 모형을 활용하여 주택가격동향조사의 표본 확대 없이 현재 조사된 데이터만으로도 정도 높은 읍면동별 주택 매매가격지수 추정으로 보다 효과적으로 소지역의 동향을 파악할 수 있을 것으로 기대된다. 이는 조사 예산 확대 없이도 현재의 공표단위보다 하위 단위에서의 가격 동향을 확인 할 수 있을 것으로 보인다. 모든 통계에서 소지역 추정이 중요시 여겨지고 있는 시점에서 한국감정원이 생산하고 있는

가격지수인 전세가격지수, 월세가격지수, 상업용 임대가격지수 등 부동산 가격 통계가 소지역 추정을 통해 통계가 작성이 되길 기대해 본다.

참고문헌

1. Datta, G. S., Ghosh, M., Nangia, N., and Natarajan, K., 1996, *Estimation of Median Income of Four-Person Families : A Bayesian Approach*, in *Bayesian Analysis in Statistics and Econometrics : Essays in Honor of Arnold Zellener*, eds. D. A. Berry, K. M. Chaloner, and J. K. Geweke, New York : Wiley, pp. 129-140.
2. Eilers, P. H. C., and Marx, B. D. 1996, "Flexible Smoothing with B-splines and penalties (with Discussion)", *Statistical Science*, 11, 89-121.
3. Fay, R. E., 1987, *Application of Multi-variate Regression to Small Domain Estimation*, in *Small Area Statistics*, eds. R. Platek, J. N. K. Rao, C. E. Samdal, and M. P. Singh, New York : Wiley, pp.91-102.
4. Fay, R. E. and Herriot, R. A., 1979, "Estimates of Income for Small Places: An Application of James-Stein Procedure to Census Data", *Journal of the American Statistical Association*, 74, 269-277.
5. Fay, R. E., Nelson, C. T., and Litow, L., 1993, "Estimation of Median Income for 4-Person Families by State, In Indirect Estimators in Federal Programs", *Statistical Policy Working Paper 21*, Washington, DC : Statistical Policy Office, Office of Management and Budget, pp.901-917.
6. Ghosh, M., Nangia, N., and Kim, D. H., 1996, "Estimation of Median Income of Four-Person Families : A Bayesian Time Series Approach", *Journal of the American Statistical Association*, 91, 1423-1431.
7. Ghosh, M. and Rao, J. N. K., 1994, "Small Area Estimation: An Appraisal", *Statistical Science*, 9, 55-76.
8. Kim, D. H., 2005, *Introduction to the Bayesian Statistics using R and Winbugs*, Korea : Freedom Academy. (in Korean)
9. Nitzoufras, I., 2009, *Bayesian Modeling Using WinBUGS*, New York : Wiley.
10. Rao, J. N. K., 2003, *Small Area Estimation*, New York : Wiley.
11. Ruppert, D., 2002, "Selecting the Number of Knots for Penalized Splines", *Journal of Computational and Graphical Statistics*, 11, 735-757.
12. Ruppert, D., and Carroll, R., 1999, "Spatially-Adaptive Penalties for Spline Fitting", *Australian and New Zealand Journal of Statistics*, 42, 205-223.
13. Heo, S. Y., 2013, A Comparative Study in Bayesian Semiparametric Approach to Small Area Estimation, thesis of Master degree, Department of Statistics, Kyungpook National University.
14. Choi, S. M., 2009, Robust Bayesian Estimation of Small Area Means, thesis of Master degree, Department of Statistics, Kyungpook National University.
15. Schumaker, L. L., 1981, *Spline Functions : Basic Theory*, New York : Wiley.

논문접수일 : 2016년 4월 10일

심사(수정)일 : 1차 2016년 5월 26일

게재확정일 : 2016년 5월 30일

국문초록

지리적 위치가 고정되어 있는 부동산은 입지 속성을 가지고 접근성에 따라서 용도가 결정되기 때문에 입지와 접근성에 대한 선호도를 유발하게 되고 그에 따른 가격을 형성하게 된다. 이러한 부동산의 속성 때문에 지역에 따라 가격의 차이가 발생하게 되고 이는 충분히 작은 소지역(small area)내에서도 그 영향을 미치게 되므로 소지역의 통계가 중요시 되고 있다. 이러한 소지역 추정에는 통계적 모형화를 통하여 이용 가능한 변수가 타당한 정보를 제공하지 못하는 지역에 대해 믿을만한 추정치를 제공 가능하게 한다. Fay-Herriot(1979) 모형은 가장 잘 알려진 소지역 추정 모형으로 설계모형(Design model)과 회귀 분석 모형(regression model)의 조합된 모형이며 선형성이 만족하지 않을 경우 준모수적 접근방법으로 해결 가능하다. 이 논문에서는 4가지의 후보 모형을 설정하여 계층적 베이지안 추론 방법을 활용하고, 베이지안 추정시 발생하는 수치적 어려움을 해결하기 위해 깃스 표집기를 사용하였다. 각 모형은 네가지 비교 척도를 사용하여 모형을 선택하고 추정값의 정도를 비교하였다. 이를 통해 향후 읍면동별 지가지수의 소지역 모델링을 통해 읍면동별 주택매매가격지수 뿐만 아니라 전세가격지수, 월세가격지수, 상업용 임대가격지수 등 부동산 가격과 관련된 지수들의 추정이 가능하게 될 것이다.

주제어 : 전국주택가격동향조사, Fay-Herriot 모형, 소지역 추정, 베이지안, 준모수적 방법

부 록(Appendix)

1. 깁스 표집기(Gibbs Sampler) 활용을 위한 모수의 조건부 분포

1) θ_{ij} 의 조건분포

$$\begin{aligned}\pi(\theta_{ij} | \beta, \gamma, \psi^2, \mathbf{b}, \mathbf{X}, \mathbf{Z}) &\propto \exp\left(-\frac{(y_{ij} - \theta_{ij})^2}{2\sigma_{ij}^2}\right) \exp\left(-\frac{(\theta_{ij} - X_{ij}^T \beta - Z_{ij}^T \gamma - b_i)^2}{2\psi_j^2}\right) \\ &\propto \exp\left(-\frac{(\psi_j^2 + \sigma_{ij}^2)\theta_{ij}^2 - 2(y_{ij}\psi_j^2 + (X_{ij}^T \beta + Z_{ij}^T \gamma + b_i)\sigma_{ij}^2)}{2\sigma_{ij}^2\psi_j^2}\right) \\ &\propto \exp\left(-\frac{(\theta_{ij} - \frac{y_{ij}\psi_j^2 + (X_{ij}^T \beta + Z_{ij}^T \gamma + b_i)\sigma_{ij}^2}{\sigma_{ij}^2 + \psi_j^2})^2}{2\frac{\sigma_{ij}^2\psi_j^2}{\sigma_{ij}^2 + \psi_j^2}}\right)\end{aligned}$$

이므로

$$[\theta_{ij} | \beta, \gamma, \psi^2, \mathbf{b}, \mathbf{X}, \mathbf{Z}] \sim N\left(\left(\frac{1}{\sigma_{ij}^2} + \frac{1}{\psi_j^2}\right)^{-1} \left(\frac{y_{ij}}{\sigma_{ij}^2} + \frac{(X_{ij}^T \beta + Z_{ij}^T \gamma + b_i)}{\psi_j^2}\right), \left(\frac{1}{\sigma_{ij}^2} + \frac{1}{\psi_j^2}\right)^{-1}\right]$$

2) b_i 의 조건분포

$$\begin{aligned}\pi(b_i | \beta, \gamma, \theta, \psi^2, \sigma_b^2, \mathbf{X}, \mathbf{Z}) &\propto \prod_{j=1}^t \left\{ \exp\left(-\frac{(\theta_{ij} - X_{ij}^T \beta - Z_{ij}^T \gamma - b_i)^2}{2\psi_j^2}\right) \right\} \exp\left(-\frac{b_i^2}{2\sigma_b^2}\right) \\ &\propto \exp\left\{-\frac{\frac{1}{2} \left(b_i^2 - 2 \left(\sum_{j=1}^t \frac{\theta_{ij} - X_{ij}^T \beta - Z_{ij}^T \gamma}{\psi_j^2} \right) / \left(\sum_{j=1}^t \frac{1}{\psi_j^2} + \frac{1}{\sigma_b^2} \right) b_i \right)}{\left(\sum_{j=1}^t \frac{1}{\psi_j^2} + \frac{1}{\sigma_b^2} \right)^{-1}}\right\}\end{aligned}$$

이므로

$$\begin{aligned}[b_i | \beta, \gamma, \theta, \psi^2, \sigma_b^2, \mathbf{X}, \mathbf{Z}] \\ \sim N\left[\left(\frac{1}{\sigma_b^2} + \sum_{j=1}^t \frac{1}{\psi_j^2}\right)^{-1} \left(\sum_{j=1}^t \frac{1}{\psi_j^2} (\theta_{ij} - X_{ij}^T \beta - Z_{ij}^T \gamma)\right), \left(\frac{1}{\sigma_b^2} + \sum_{j=1}^t \frac{1}{\psi_j^2}\right)^{-1}\right]\end{aligned}$$

3) β 의 조건분포

$$\begin{aligned}
 \pi(\beta | \gamma, \theta, \psi^2, \mathbf{X}, \mathbf{Z}) &\propto \prod_{i=1}^m \left[\prod_{j=1}^t \left\{ \exp\left(-\frac{(\theta_{ij} - X_{ij}^T \beta - Z_{ij}^T \gamma - b_i)^2}{2\psi_j^2}\right) \right\} \right] \pi(\beta) \\
 &\propto \exp\left(-\sum_{i=1}^m \sum_{j=1}^t \frac{(X_{ij}^T \beta)^2 - 2\beta^T X_{ij}(\theta_{ij} - Z_{ij}^T \gamma - b_i)}{2\psi_j^2}\right) \\
 &\propto \exp\left[-\frac{1}{2}(\beta^T \left(\sum_{i=1}^m \sum_{j=1}^t \frac{X_{ij} X_{ij}^T}{\psi_j^2}\right) \beta - 2 \sum_{i=1}^m \sum_{j=1}^t \beta^T \frac{X_{ij}(\theta_{ij} - Z_{ij}^T \gamma - b_i)}{\psi_j^2})\right]
 \end{aligned}$$

이므로

$$\begin{aligned}
 &[\beta | \gamma, \theta, \mathbf{b}, \psi^2, \mathbf{X}, \mathbf{Z}] \\
 &\sim N\left[\left(\sum_{i=1}^m \sum_{j=1}^t \frac{X_{ij} X_{ij}^T}{\psi_j^2}\right)^{-1} \left(\sum_{i=1}^m \sum_{j=1}^t \frac{X_{ij}}{\psi_j^2} (\theta_{ij} - Z_{ij}^T \gamma - b_i)\right), \left(\sum_{i=1}^m \sum_{j=1}^t \frac{X_{ij} X_{ij}^T}{\psi_j^2}\right)^{-1}\right]
 \end{aligned}$$

4) γ 의 조건분포

$$\begin{aligned}
 \pi(\gamma | \beta, \theta, \psi^2, \sigma_\gamma^2, \mathbf{X}, \mathbf{Z}) &\propto \prod_{i=1}^m \left[\prod_{j=1}^t \left\{ \exp\left(-\frac{(\theta_{ij} - X_{ij}^T \beta - Z_{ij}^T \gamma - b_i)^2}{2\psi_j^2}\right) \right\} \right] \frac{1}{\sigma_\gamma^K} \exp\left(-\frac{\gamma^T \gamma}{2\sigma_\gamma^2}\right) \\
 &\propto \exp\left(-\frac{1}{2} \left(\sum_{i=1}^m \sum_{j=1}^t \frac{\gamma^T Z_{ij} Z_{ij}^T \gamma - 2\gamma^T Z_{ij}(\theta_{ij} - X_{ij}^T \beta - b_i)}{\psi_j^2} + \frac{\gamma^T \gamma}{\sigma_\gamma^2}\right)\right) \\
 &\propto \exp\left[-\frac{1}{2} \left\{ \gamma^T \left(\sum_{i=1}^m \sum_{j=1}^t \frac{Z_{ij} Z_{ij}^T}{\psi_j^2} + \frac{1}{\sigma_\gamma^2} \mathbf{I}\right) \gamma - 2 \sum_{i=1}^m \sum_{j=1}^t \frac{Z_{ij}(\theta_{ij} - X_{ij}^T \beta - b_i)}{\psi_j^2} \right\}\right]
 \end{aligned}$$

이므로

$$\begin{aligned}
 &[\gamma | \beta, \theta, \mathbf{b}, \psi^2, \sigma_\gamma^2, \mathbf{X}, \mathbf{Z}] \\
 &\sim N\left[\left(\sum_{i=1}^m \sum_{j=1}^t \frac{Z_{ij} Z_{ij}^T}{\psi_j^2} + \frac{1}{\sigma_\gamma^2} \mathbf{I}\right)^{-1} \left(\sum_{i=1}^m \sum_{j=1}^t \frac{Z_{ij}}{\psi_j^2} (\theta_{ij} - X_{ij}^T \beta - b_i)\right), \left(\sum_{i=1}^m \sum_{j=1}^t \frac{Z_{ij} Z_{ij}^T}{\psi_j^2} + \frac{1}{\sigma_\gamma^2} \mathbf{I}\right)^{-1}\right]
 \end{aligned}$$

5) $(\sigma_\gamma^2)^{-1}$ 의 조건분포

$$\begin{aligned}\pi((\sigma_\gamma^2)^{-1}|\boldsymbol{\gamma}) &\propto \frac{1}{\sigma_\gamma^K} \exp\left(-\frac{\boldsymbol{\gamma}^T \boldsymbol{\gamma}}{2\sigma_\gamma^2}\right) \left(\frac{1}{\sigma_b^2}\right)^{c_\gamma-1} \exp\left(-\frac{d_\gamma}{\sigma_\gamma^2}\right) \\ &\propto \left(\frac{1}{\sigma_\gamma^2}\right)^{\frac{K}{2}+c_\gamma-1} \exp\left(-\frac{\frac{1}{2}\boldsymbol{\gamma}^T \boldsymbol{\gamma} + d_\gamma}{\sigma_\gamma^2}\right)\end{aligned}$$

이므로

$$[(\sigma_\gamma^2)^{-1}|\boldsymbol{\gamma}] \sim \text{Gamma}\left(\frac{K}{2}+c_\gamma, \frac{1}{2}\boldsymbol{\gamma}^T \boldsymbol{\gamma} + d_\gamma\right)$$

6) $(\psi_j^2)^{-1}$ 의 조건분포

$$\begin{aligned}\pi((\psi_j^2)^{-1}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \mathbf{b}, \mathbf{X}, \mathbf{Z}) &\propto \prod_{i=1}^m \left\{ \frac{1}{\psi_j^2} \exp\left(-\frac{(\theta_{ij} - X_{ij}^T \boldsymbol{\beta} - Z_{ij}^T \boldsymbol{\gamma} - b_i)^2}{2\psi_j^2}\right) \right\} \left(\frac{1}{\psi_j^2}\right)^{c_j-1} \exp\left(-\frac{d_j}{\psi_j^2}\right) \\ &\propto \left(\frac{1}{\psi_j^2}\right)^{\frac{m}{2}+c_j-1} \exp\left(-\frac{\frac{1}{2}\sum_{i=1}^m (\theta_{ij} - X_{ij}^T \boldsymbol{\beta} - Z_{ij}^T \boldsymbol{\gamma} - b_i)^2 + d_j}{\psi_j^2}\right)\end{aligned}$$

이므로

$$[(\psi_j^2)^{-1}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \mathbf{b}, \mathbf{X}, \mathbf{Z}] \sim \text{Gamma}\left(c_j + \frac{m}{2}, \frac{1}{2}\sum_{i=1}^m (\theta_{ij} - X_{ij}^T \boldsymbol{\beta} - Z_{ij}^T \boldsymbol{\gamma} - b_i)^2 + d_j\right)$$

7) $(\sigma_b^2)^{-1}$ 의 조건분포

$$\begin{aligned}\pi((\sigma_b^2)^{-1}|\mathbf{b}) &\propto \prod_{i=1}^m \left\{ \frac{1}{\sigma_b^2} \exp\left(-\frac{b_i^2}{2\sigma_b^2}\right) \right\} \left(\frac{1}{\sigma_b^2}\right)^{c-1} \exp\left(-\frac{d}{\sigma_b^2}\right) \\ &\propto \left(\frac{1}{\sigma_b^2}\right)^{\frac{m}{2}+c-1} \exp\left(-\frac{\frac{1}{2}\sum_{i=1}^m b_i^2 + d}{\sigma_b^2}\right)\end{aligned}$$

이므로

$$[(\sigma_b^2)^{-1}|\mathbf{b}] \sim \text{Gamma}\left(\frac{m}{2}+c, \frac{1}{2}\sum_{i=1}^m b_i^2 + d\right)$$