

階層ベイズモデルを用いた不動産データの復元*

Imputing missing real estate data using hierarchical Bayesian model *

瀬谷創**・堤盛人***

By・Hajime SEYA**・Morito TSUTSUMI***

1. はじめに

近年、GIS、GPS、リモートセンシング等の発達によって、不動産の価格分析のための様々な属性情報が容易に入手できる環境が整いつつある。しかしながら依然として、欠損データ (missing data) の問題は、実務においては悩ましい課題の一つである。これに対して Knight *et al.* (1998) は、ベイズ統計学に基づくマルコフ連鎖モンテカルロ (以下、MCMC) 法を用いて、欠損データを効果的に復元する方法を提案している。しかしながら、Knight *et al.* (1998) は、不動産データの重要な側面である、空間的依存性 (spatial dependence) が考慮されていないという点で問題がある。筆者らは、Knight *et al.* (1998) の方法を拡張し、データ間の空間的な依存関係を考慮しながら欠損データを復元するアプローチを開発してきた (堤他, 2007 など)。しかしながら、堤他 (2007) は、MCMC 法と他のパラメータ推定手法との実証比較を大きな目的としていたため、モデルの評価を通常の決定係数で行うなど、必ずしもベイズ統計学の枠組みでの一貫性を持っていない。したがって本研究は、堤他 (2007) のアプローチをベイズ統計学的な意味で精緻化し、データ間の空間依存関係を考慮した不動産データの復元のための一つの方法論を提示することを目的としている。

2. 不動産データと空間的依存性

(1) 空間計量経済学・空間統計学

不動産の価格・賃料評価のためのデータ分析においては、データ間の空間的な依存性を考慮することが不可欠である。不動産評価に対して空間モデリングを活用する研究は、例えば、古谷 (2004)、井上他 (2009) など近年大きな発展を見せている。空間相関の考慮の方法は、空間計量経済学と空間統計学で違いがあるが、本研究のように観測値以外におけるデータの予測 (内挿) を行う場合は、後者のアプローチのほうが実用的であるため、本

研究では後者を用いることとする。なお、紙面の都合上、空間計量経済学と空間統計学の関係性や最新の研究成果については、次の文献に譲ることとする (関係性: Griffith, 2007; 堤・瀬谷, 2010; 空間計量経済学: Anselin, 2010; 空間統計学: Gelfand *et al.*, 2010)。

(2) 空間統計モデルによる空間相関のモデル化

本研究では、次のような線形回帰型の空間統計モデルを対象とする。

$$y \sim N(\mathbf{X}\boldsymbol{\beta}, \tau^2 \mathbf{I} + \sigma^2 \mathbf{H}(\phi)). \quad (1)$$

y は $n \times 1$ の目的変数ベクトル、 \mathbf{X} は $n \times k$ の説明変数行列、 $\boldsymbol{\beta}$ は $k \times 1$ のパラメータベクトル、 \mathbf{I} は $n \times n$ の単位行列であり、 τ^2 、 σ^2 、 ϕ はそれぞれナゲット、パーシャル・シル、レンジと呼ばれるパラメータである。 $N(\cdot, \cdot)$ は正規分布を示す。 \mathbf{H} は、 $n \times n$ の相関行列であり、その i, j 要素を距離 d_{ij} に依存する関数で与える (空間過程の等方性を仮定)。本研究では、相関関数に指数型を仮定し、 $\exp(d_{ij}/\phi)$ で与えることとする。

今、パラメータをベイズ推定するにあたり、パラメータ集合 (ベクトル) $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2, \tau^2, \phi)$ に関する事前分布は独立、すなわち $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\sigma^2)\pi(\tau^2)\pi(\phi)$ を満たすとしよう。ここで、ベイズの定理 $\pi(\boldsymbol{\theta} | y) \propto f(y | \boldsymbol{\theta})\pi(\boldsymbol{\theta})$ より、尤度関数と事前確率密度関数を合成することで各パラメータの事後確率密度関数を求めることができる。しかしながら、(1) 式に関する尤度の計算では、 $(\tau^2 \mathbf{I} + \sigma^2 \mathbf{H})^{-1}$ という逆行列の計算が必要なため、特に市販の計算機では大規模データへの適用が難しい。そこで、次のような階層型モデルを考える (Banerjee *et al.*, 2004)。

$$y | \boldsymbol{\beta}, \mathbf{W}, \tau^2 \sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{W}, \tau^2 \mathbf{I}), \mathbf{W} | \sigma^2, \phi \sim N(\mathbf{0}, \sigma^2 \mathbf{H}(\phi)). \quad (2)$$

$\mathbf{0}$ はその要素を 0 とする $n \times 1$ のベクトルである。ベイズの定理より、事後確率密度関数 $\pi(\boldsymbol{\beta}, \sigma^2, \tau^2, \phi | y)$ は、事前分布と (2) 式から与えられる尤度 f の積に比例する。

$$f(y | \boldsymbol{\beta}, \mathbf{W}, \tau^2)\pi(\boldsymbol{\beta})\pi(\tau^2)f(\mathbf{W} | \sigma^2, \phi)\pi(\sigma^2)\pi(\phi). \quad (3)$$

ここで、各パラメータの事前分布を、 $\boldsymbol{\beta} \sim N(\mathbf{c}, \mathbf{T})$ 、 $\sigma^2 \sim IG(a_\sigma, b_\sigma)$ 、 $\tau^2 \sim IG(a_\tau, b_\tau)$ 、 $\phi \sim Ga(a_\phi, b_\phi)$ と設定する。ただし、 $IG(\cdot, \cdot)$ は逆ガンマ分布、 $Ga(\cdot, \cdot)$ はガンマ

*キーワード: 不動産賃料、データ復元、空間統計学、階層ベイズモデル

**正会員 独立行政法人国立環境研究所

(つくば市小野川16-1 seya.hajime@nies.go.jp)

***正会員 博(工) 筑波大学大学院システム情報工学研究科

(つくば市天王台 1-1-1 tsutsumi@sk.tsukuba.ac.jp)

分布を示す。(3) 式より、各パラメータに関する条件付事後分布 (full conditional) が、次のように求められる。

$$\boldsymbol{\beta} | \mathbf{y}, \mathbf{W}, \mathbf{X}, \tau^2 \sim N(\mathbf{D}_\beta \mathbf{d}_\beta, \mathbf{D}_\beta).$$

$$\mathbf{W} | \tau^2, \sigma^2, \phi, \mathbf{y}, \mathbf{X}, \boldsymbol{\beta} \sim N(\mathbf{D}_W \mathbf{d}_W, \mathbf{D}_W).$$

$$\tau^2 | \mathbf{y}, \mathbf{X}, \boldsymbol{\beta}, \mathbf{W} \sim IG\left(a_\tau + \frac{n}{2}, b_\tau + \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{W})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{W})\right).$$

$$\sigma^2 | \mathbf{W}, \phi \sim IG\left(a_\sigma + \frac{n}{2}, b_\sigma + \frac{1}{2}\mathbf{W}\mathbf{H}^{-1}(\phi)\mathbf{W}\right).$$

$$\pi(\phi | \mathbf{W}, \sigma^2) \propto \pi(\phi) \times \exp\left(-\frac{1}{2\sigma^2}\mathbf{W}\mathbf{H}^{-1}(\phi)\mathbf{W}\right). \quad (4)$$

$$\text{ただし、 } \mathbf{D}_\beta = \left(\frac{1}{\tau^2}\mathbf{X}\mathbf{X}' + \mathbf{T}^{-1}\right)^{-1}, \quad \mathbf{d}_\beta = \frac{1}{\tau^2}\mathbf{X}'(\mathbf{y} - \mathbf{W}) + \mathbf{T}^{-1}\mathbf{c},$$

$$\mathbf{D}_W = \left(\frac{1}{\tau^2}\mathbf{I} + \frac{1}{\sigma^2}\mathbf{H}^{-1}(\phi)\right)^{-1}, \quad \mathbf{d}_W = \frac{1}{\tau^2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

パラメータの推定には、MCMC法を用いる。(4)式から分かるように、レンジパラメータ ϕ 以外は、条件付事後分布が標準的な分布となっており、容易に乱数を発生できる(ギブス・サンプラー)。しかしながら、 ϕ に関する事後分布は標準的な形式となっておらず、MH-アルゴリズムや、スライスサンプラーなどを用いる必要がある。本研究では、前者の酔歩連鎖(random walk chain)を用いることとする。まず、 ϕ を次式から発生する(ただし、 r はシミュレーションのステップ番号)。

$$\phi^* = \phi^{(r-1)} + \varepsilon_r, \quad \varepsilon_r \sim N(0, \zeta^2). \quad (5)$$

候補 ϕ^* の採択確率は、次式により与えられる。

$$\alpha(\phi^{(r-1)}, \phi^*) = \min\left(\frac{\pi(\phi^* | \mathbf{W}, \sigma^2)}{\pi(\phi^{(r-1)} | \mathbf{W}, \sigma^2)}, 1\right). \quad (6)$$

(0,1)上の一様乱数 u を発生させて、

$$\phi^{(r)} = \begin{cases} \phi^*, & u \leq \alpha(\phi^{(r-1)}, \phi^*) \\ \phi^{(r-1)}, & u > \alpha(\phi^{(r-1)}, \phi^*) \end{cases}. \quad (7)$$

とする。この方法では、採択確率が25%程度になるように試行錯誤的に ζ^2 を見つける必要がある(Roberts *et al.*, 1997)。任意地点におけるデータ値の予測は、 $\boldsymbol{\psi} = (\boldsymbol{\theta}, \mathbf{W})$ としたとき、次式により行うことが可能である(Bayesian kriging (以下、BK), Banerjee *et al.*, 2004)。

$$\begin{aligned} \pi(\mathbf{y}_0 | \mathbf{y}, \mathbf{X}_0) &= \int \pi(\mathbf{y}_0 | \mathbf{y}, \boldsymbol{\psi}, \mathbf{X}_0) \pi(\boldsymbol{\psi} | \mathbf{y}) d\boldsymbol{\psi} \\ &\approx \frac{1}{R} \sum_{r=1}^R \pi(\mathbf{y}_0 | \mathbf{y}, \boldsymbol{\psi}^{(r)}, \mathbf{X}_0). \end{aligned} \quad (8)$$

ただし、 R はMCMC法におけるサンプル数、 \mathbf{y}_0 は $m \times 1$ の予測値ベクトル、 \mathbf{X}_0 は $m \times k$ の予測地点における説明変数行列である。BKは、推定されたパラメータの点推定値をプラグインする古典的なクリギングと違って、パラメータの不確実性を予測に取り組みることが可能である。

3. 不動産データの復元

(1) Knight *et al.* (1998) の方法

本節では、Knight *et al.* (1998) で提案されたデータ復元方法を簡単に紹介する。

まず、次のような線形回帰モデルを想定する。

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i, \quad (i=1, \dots, n). \quad (9)$$

ε_i は、独立に $N(0, \xi^2)$ に従うとする。まず、欠損データがないと想定すると、尤度は次式で与えられる。

$$L(\boldsymbol{\omega}, \mathbf{y}) = (\xi^2)^{-n/2} \exp[-\sum (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i})^2 / 2\xi^2]. \quad (10)$$

ただし、 $\boldsymbol{\omega} = (\beta_0, \beta_1, \beta_2, \xi^2)$ である。簡単のため、パラメータの事前分布が $\pi(\boldsymbol{\omega}) = \pi(\beta_0, \beta_1, \beta_2) \pi(\xi^2)$ と表現できるとし、 $\pi(\beta_0, \beta_1, \beta_2) = 1$ 、 $\xi^2 \sim IG(a_\xi, b_\xi)$ と設定しよう。すると、条件付事後分布が、次式により与えられる。

$$\beta_0 | \beta_1, \beta_2, \xi^2, \mathbf{y} \sim N\left(\frac{\sum (y_i - \beta_1 x_{1i} - \beta_2 x_{2i})}{n}, \frac{\xi^2}{n}\right),$$

$$\beta_1 | \beta_0, \beta_2, \xi^2, \mathbf{y} \sim N\left(\frac{\sum x_{1i} (y_i - \beta_0 - \beta_2 x_{2i})}{\sum x_{1i}^2}, \frac{\xi^2}{\sum x_{1i}^2}\right),$$

$$\beta_2 | \beta_0, \beta_1, \xi^2, \mathbf{y} \sim N\left(\frac{\sum x_{2i} (y_i - \beta_0 - \beta_1 x_{1i})}{\sum x_{2i}^2}, \frac{\xi^2}{\sum x_{2i}^2}\right),$$

$$\xi^2 | \beta_0, \beta_1, \beta_2, \mathbf{y} \sim IG\left(a_\xi + \frac{n}{2}, b_\xi + 0.5 \sum (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i})^2\right). \quad (11)$$

これをもとにギブス・サンプリングを行えばよい。

今、 y_1 と x_{12} の両方が欠損しているとしよう。このとき、欠損データをパラメータとみなし、パラメータ空間を次のように拡大する(data augment)。

$$\tilde{\boldsymbol{\omega}} = (\beta_0, \beta_1, \beta_2, \xi^2, y_1, x_{12}).$$

y_1 と x_{12} に関する条件付事後分布は次式で与えられる。

$$y_1 | \beta_0, \beta_1, \beta_2, \xi^2, \mathbf{y}_{-1}, x_{12} \sim N(\beta_0 + \beta_1 x_{11} + \beta_2 x_{21}, \xi^2),$$

$$x_{12} | \beta_0, \beta_1, \beta_2, \xi^2, \mathbf{y} \sim N\left(\frac{y_2 - \beta_0 - \beta_2 x_{22}}{\beta_1}, \frac{\xi^2}{\beta_1^2}\right). \quad (12)$$

ただし、 \mathbf{y}_{-1} は y_1 を除いた y_i の観測値ベクトルを示す。(12)式をギブス・サンプラーに追加すれば、欠損データの復元とパラメータの同定が同時に可能となる。

(2) 空間相関を考慮したデータ復元方法

空間相関を考慮した場合は、(4)、(5)の条件付き分布からの乱数発生に、前節の方法を組み合わせればよい。今、(2)式において、地域 p の目的変数 y_p と、地域 q の l 番目の属性変数 x_{lq} が欠損しているとし、パラメータ空間を $\tilde{\boldsymbol{\psi}} = (\boldsymbol{\psi}, y_p, x_{lq})$ と拡大しよう。このとき、 y_p と、 x_{lq} の条件付事後分布は、次式により与えられる。

$$y_p | \boldsymbol{\psi}, \mathbf{y}_{-p}, x_{lq} \sim N\left(\beta_0 + \sum_h \beta_h x_{hp} + w_p, \tau^2\right),$$

$$x_{lq} | \boldsymbol{\psi}, \mathbf{y} \sim N\left(\frac{(y_q - w_q - \beta_0 - \sum_{h \neq l} \beta_h x_{hq})}{\beta_l}, \frac{\tau^2}{\beta_l^2}\right). \quad (13)$$

空間相関を考慮した欠損値復元方法としては、他に

MCMC 法ベースの LeSage and Pace (2004)、E-M アルゴリズムベースの Griffith (2009) や堤・村上 (2009) 等が提案されている。しかしこれらはいずれも空間計量経済学のモデルに基づいたものであり、直接任意点の補間 (内挿) に用いることはできない。また、本手法は説明変数の欠損にも対応しているという点で利点を有している。

4. 東京 23 区の賃料データを用いた実証分析

(1) データの概要

本稿では、アットホーム株式会社より提供を受けた、平成 18 年 1 月から 5 月にかけて取引情報として登録された東京 23 区の賃料データの一部を実証分析に用いる (堤他 (2007) で用いたのと同じもの)。採用した説明変数は表 1 のとおりである。交通利便性を示す指標としては、各物件から東京・大手町・新宿の各駅までの所要時間のうち最短時間を用いることとする。データには最寄り駅までのバス・徒歩による時間等も含まれているのでこれを利用する。ただし、鉄道所要時間については、別途、(株)ヴァル研究所の「駅すばあと」(2006 年 10 月版) により算出した。

(2) 分析の概要

本稿では、(1) のデータセットからランダムにサンプリングした 150 件の物件を分析に用いる。この 150 件を完全データ (以下、CD) と定義する。ここで、150 件からランダムに $v = [10, 40, 70] \%$ を抽出し、抽出された物件は、賃料又はいずれか一つの説明変数 (連続変数のみ) が欠損していると仮定する。具体的には、抽出された物件について $[\ln(\text{apartment rent value}), \text{train}, \text{walk}, \text{bus}, \text{floor area}(\log), \text{age}(\text{sqr})]$ のいずれかを欠損させ、その後 MCMC 法を用いて完全データの復元を行う。そして、復元された完全データをもとに、50 点の検証用データ (以下、VD) に対する賃料予測を行い、RMSE によって予測精度の評価を行う。図 1 に、CD 及び VD の位置を示す。この CD と VD は、他のパラメータ推定手法との比較を試みた堤他 (2007) と同一のものであるため、結果の比較も可能である。

なお、本研究では Matlab7.5.0 を用いてコードを作成したが、MCMC 法のプログラミングは複雑であり自作コードは誤りを含みやすいため、Geweke (2004) の Joint distribution test によって自作したコードに問題がないことを確認している。また、周辺尤度の計算には Gelfand and Dey (1994) の方法を用いる。

パラメータの初期値としては、 $c = 0$ 、 $T = 1.0 \times 10^{12} \times I$ 、 $a_\sigma = b_\sigma = a_\tau = b_\tau = 5.0 \times 10^{-3}$ 、 $\phi = 10$ 、 $a_\phi = 30$ 、 $b_\phi = 10$ を与え、発生した 11,000 個のサンプルから burn-in 1,000 個を除いた 10,000 個を推論に利用している。

表 1 説明変数の内容

表中の表記	変数の内容
constant	定数
train	最寄り鉄道駅から東京/大手町/新宿までの最短時間(分)
BoT, Tokyo	都営鉄道ダミー
Odakyu	小田急線ダミー
walk	最寄り鉄道駅までの徒歩による所要時間(分)
bus	最寄りバス停から最寄り鉄道駅までの時間(分)
floor area (log)	専有面積 (m ²) の平方根
age (sqr)	築年数 (年) の平方根
reinforced concrete	鉄筋コンクリート造 (該当:1 該当しない:0)
nos. of rooms	部屋の数
one-room type	ワンルームタイプ (該当:1 該当しない:0)
1K-type	1Kタイプ (該当:1 該当しない:0)
parking lot	駐車場の有無 (有:1 無:0)
self-locking	オートロック (該当:1 該当しない:0)



図 1 パラメータ推定用 150 物件 (灰色)、検証用 50 物件 (黒色) の位置

(3) 分析結果と考察

表 2 に、欠損なし (完全データ) のケースにおける、空間相関を考慮していないケース ("Non-spatial") と考慮したケース ("Spatial") に関するパラメータ推定結果を示す。対数周辺尤度から判断して、実データのへの当てはまりは空間相関を考慮したモデルで大きく改善されていることが分かる。ベイズ推論では、信用区間 (C.I.) に 0 が含まれているかという観点からパラメータの有意性の検定を行う。トレンドパラメータの符号は直観と整合するものとなっているものの、95%信用区間で見ると判断して、いくつかの変数 (例えば専有面積) は統計的に有意とはなっていないことが分かる。しかしながらこれらの変数を残しても、以下の議論に本質的な影響を与えないことから、筆者らの過去の研究 (例えば、堤他, 2007) と予測結果を比較できるという点を重要視し、あえてそのまま用いることとする。

ϕ は、1km 程度と、やや狭い値に推定された (指数型の場合、空間相関が 0.05 となる直観的なレンジ (effective range) は ϕ の約 3 倍)。しかしながら、図 2 に示す通り、サンプリングはバランス良く行われており、

表2 パラメータの推定結果

Method	Non spatial		Spatial	
	Coef.	C.I.95%	Coef.	C.I.95%
constant	9.64	[9.18,10.1]	9.52	[9.14,9.91]
train	-0.715	[-0.87,-0.563]	-0.512	[-0.703,-0.311]
BoT, Tokyo *	-0.112	[-0.194,-0.0306]	-0.0969	[-0.17,-0.0264]
Odakyu *	-0.372	[-0.615,-0.133]	-0.446	[-0.648,-0.254]
walk	-0.0317	[-0.0548,-0.0078]	-0.0351	[-0.053,-0.0174]
bus	-0.0347	[-0.0531,-0.0163]	-0.0321	[-0.0502,-0.015]
floor area (log)	0.0439	[-0.0405,0.128]	0.0115	[-0.0521,0.0769]
age (sq)	-0.0153	[-0.0701,0.0386]	-0.0150	[-0.0591,0.0311]
rc *	0.267	[0.116,0.421]	0.229	[0.0846,0.374]
nos. of rooms	0.0833	[0.0172,0.149]	0.0629	[0.0127,0.111]
one-room type *	0.0069	[-0.308,0.331]	-0.0286	[-0.337,0.276]
1K-type *	-0.0374	[-0.116,0.0412]	-0.0414	[-0.11,0.0272]
parking lot *	0.0433	[-0.0748,0.161]	0.0432	[-0.0436,0.129]
self-locking *	0.0607	[-0.0426,0.164]	0.0621	[-0.0116,0.138]
nugget	0.0355	[0.0286, 0.0438]	0.0041	[0.0008,0.0088]
partial-sill			0.0273	[0.0183,0.0369]
range			1.09	[0.815,1.390]
log marginal likelihood	37.6		204	

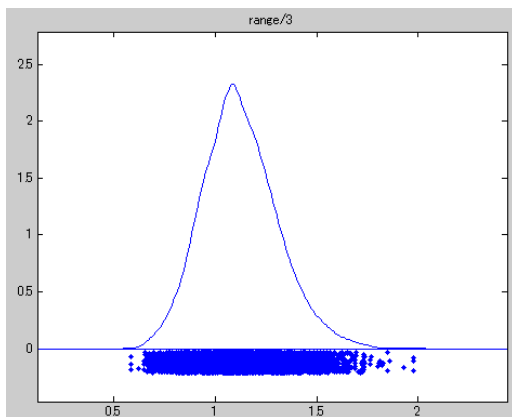


図2 レンジパラメータの事後分布

表3 予測精度の比較

Method	Non-spatial				Spatial				
	0%	10%	40%	70%	0%	10%	40%	70%	
RMSE	MCMC	0.772	0.781	0.768	0.745	0.609	0.607	0.641	0.653
	Discard	-----	0.757	0.847	0.825	-----	0.579	0.682	0.780

Geweke (1992) の収束診断における p 値では、全パラメータについて事後分布への収束が示唆されている（紙面の都合上結果は省略）。

表3は、予測結果の RMSE を示したものである。実務では、欠損データを除外してパラメータの推定を行うことが非常に多いが、その方法では、例えば40%欠損の場合、すなわち、90物件のみのデータを用いると、RMSEが0.847 (Non-spatial)、0.682 (Spatial) となり、両者ともに完全データの場合と比較して予測精度が大きく低下した。これに対し、Knight *et al.* (1998)、あるいは本研究で用いた手法を用いれば、例え賃料または一部の説明変数の7割に欠損が含まれていたとしても、それを復元しながら、完全データの場合とほぼ同等かやや劣る程度の精度で予測が行えることが示唆された。また、空間的な相関を考慮することで、予測精度はさらに大きく改善された。無論この結果は、たかだか一つのCDとVDに対する結果であり、観測・予測地点の配置によっ

て結果は変わる可能性がある。しかしながら、予測精度の高さ、MCMC サンプラーに組み込むだけというプログラミングの容易さ等を考慮すると、実務的にも非常に有用な方法に成りうるものといえよう。

参考文献

- 井上亮・清水英範・吉田雄太郎・李勇鶴 (2009). 「時空間クリギングによる東京23区・全用途地域を対象とした公示地価の分布と変遷の視覚化」, 『GIS—理論と応用』, 17 (1), 13–24.
- 堤盛人・瀬谷創 (2010). 「便益計測への空間ヘドニック・アプローチの適用」, 『土木学会論文集』, 66 (2), (登載決定).
- 堤盛人・村上大輔 (2009). 「市町村合併による統計データの集計単位変更に対する方策の模索：空間計量経済モデルを用いた分析への対処法」, 応用地域学会, 第23回研究発表大会, 山形大学, 12月12日–13日, 2009.
- 堤盛人・吉田靖・瀬谷創・川口有一郎 (2007). 「MCMC法によるデータ欠損問題と空間的相関を考慮した不動産賃料予測モデル：東京23区における賃貸マンション市場の実証分析」, 『ジャレフ・ジャーナル (不動産ファイナンス・不動産経済学研究)』, 3, 1–13.
- 古谷和之 (2004). 「ベイズ地理的加重回帰モデルの地価モデル推定への適用」, 『都市計画論文集』, 39 (3), 787–792.
- Anselin, L. (2010). Thirty years of spatial econometrics, *Papers in Regional Science*, 89 (1), 3–25.
- Banerjee, S., Carlin, B.P. and Gelfand, A.E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*, Chapman & Hall/CRC.
- Gelfand, A.E. and Dey, D. (1994). Bayesian model choice: asymptotic and exact calculations, *Journal Royal Statistical Society B*, 56, 501–514.
- Gelfand, A.E., Diggle, P.J., Fuentes, M. and Guttorp, P. (2010). *Handbook of Spatial Statistics*, Chapman & Hall/CRC.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments, In Bernardo, J.M., Berger, J.O., Dawid, A.P. and Smith, A.F.M, Editors, *Bayesian Statistics 4*, 169–193, Oxford University Press.
- Geweke, J. (2004). Getting it right: Joint distribution tests of posterior simulators, *Journal of the American Statistical Association*, 99, 799–804.
- Griffith, D.A. (2007). An equation by any other name is still the same: on spatial econometrics and spatial statistics, *The Annals of Regional Science*, 41, 209–227.
- Griffith, D.A. (2010). Spatial filtering and missing georeferenced data imputation: A comparison of the Getis and Griffith methods, In Anselin, L. and Rey, S.J. Editors, *Perspectives on Spatial Data Analysis*, 227–234, Springer.
- Knight, J. R., Sirmans, C. F., Gelfand, A. E. and Ghosh, S. K. (1998). Analyzing real estate data problems using the Gibbs sampler, *Real Estate Economics*, 26 (3), 469–492.
- LeSage, J.P. and Pace, R.K. (2004). Models for Spatially Dependent Missing Data, *The Journal of Real Estate Finance and Economics*, 29 (2), 233–254.
- Roberts, G.O., Gelman, A. and Gilks, W.R. (1997). Weak convergence and optimal scaling of random walk metropolis algorithms, *Annals of Applied Probability*, 7, 110–120.