

道路の速度をクリギングによって予測する方法を提案している。これらは距離の近い道路は似たような走行速度になることを仮定して、平面上で道路の位置と道路間の距離に基づいた通常クリギングによって推定を行っている。

クリギングは予測すべき対象を含む2次元もしくは3次元空間、あるいは時間を加えた時空間で適用されることがほとんどであったが、本研究では新しい試みとして平面上に位置する出発地と目的地の座標によって作られる4次元ユークリッド空間でクリギングによる旅行時間予測を行う。

2. 普遍クリギングによる旅行時間の予測

本節では主として Cressie[2]と間瀬、武田[12]に依拠しつつ、普遍クリギングによる旅行時間予測値の導出方法について述べる。

対象地域を平面領域と仮定する。この地域の出発地・目的地ペアの座標によって構成される4次元ユークリッド空間の部分集合を D とする。任意の出発地・目的地ペア $\mathbf{s} \in D$ について、2地点間の旅行時間 $z(\mathbf{s})$ が、

$$z(\mathbf{s}) = \beta f(\mathbf{s}) + \delta(\mathbf{s}) \quad (1)$$

という構造を持つと仮定する。ただし $f(\mathbf{s})$ は平面上の2地点間の距離、 β はパラメータである。2地点間の距離は必ずしも直線距離でなくてもよい。 $E(\cdot)$ を平均、 $\text{Cov}(\cdot, \cdot)$ を共分散、 \mathbf{h} を任意の4次元空間ベクトル、コバリオグラムを $C(\mathbf{h})$ として、 $\delta(\mathbf{s})$ は次の性質を満たすと仮定する。

$$E(\delta(\mathbf{s})) = 0, \quad (2)$$

$$\text{Cov}(\delta(\mathbf{s} + \mathbf{h}), \delta(\mathbf{s})) = C(\mathbf{h}). \quad (3)$$

$E(\delta(\mathbf{s})) = \text{定数}$ 、および(3)が成り立つとき、 $\delta(\mathbf{s})$ は二次定常性を満たすという。これは任意の $\mathbf{s} \in D$ について $\delta(\mathbf{s})$ の平均が一定であり、相対的位置が \mathbf{h} 離れたときの $\delta(\mathbf{s})$ と $\delta(\mathbf{s} + \mathbf{h})$ の共分散がコバリオグラムと呼ばれる関数 $C(\mathbf{h})$ で記述できることを意味する。

コバリオグラムの関数型にはいくつかの定型が提案されている。本研究では、コバリオグラムが等方的(距離 $\|\mathbf{h}\|$ のみに依存する)であることを仮定し、指数型と呼ばれる次の関数形を使用する。

$$C(\mathbf{h}) = \theta_1 e^{-\|\mathbf{h}\|/\theta_2}. \quad (4)$$

ただし θ_1, θ_2 はパラメータである。 $\theta_1 \geq 0, \theta_2 > 0$ とする。既往研究(例えば文献[5][9])では spherical 型と呼ばれるコバリオグラム型が使用されることが多い。 spherical 型コバリオグラムは、ランダムな点が分布

する3次元空間における球に含まれる点の数から導かれる[4]。この導出過程を尊重して本研究では使用を避けた。一方で指数型は距離が増えるほど共分散は大きくなり、その増え方はしだいに緩やかになることを表現しているに過ぎない。

以上の事柄がクリギング適用に必要な仮定である。

さて、 n 組の出発地・目的地ペア間の旅行時間観測データが得られているとしよう。第 i 番目のデータの出発地・目的地ベクトルを $\mathbf{s}_i \in D$ 、旅行時間を $z(\mathbf{s}_i)$ とする ($i=1, \dots, n$)。クリギングによる任意の \mathbf{s} に関する旅行時間予測値 $\hat{Z}(\mathbf{s})$ は観測データ $z(\mathbf{s}_i)$ の重み付き平均

$$\hat{Z}(\mathbf{s}) = \sum_{i=1}^n \lambda_i z(\mathbf{s}_i) \quad (5)$$

によって得られる。ここで重みパラメータ λ_i は $\sum_{i=1}^n \lambda_i = 1$ を満たすとする(ただし λ_i は正値とは限らない)。パラメータベクトル $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)'$ は平均二乗予測誤差

$$\begin{aligned} \sigma^2(\mathbf{s}) &= E((z(\mathbf{s}) - \hat{Z}(\mathbf{s}))^2) \\ &= C(\mathbf{0}) + \boldsymbol{\lambda}' \boldsymbol{\Sigma} \boldsymbol{\lambda} - 2\boldsymbol{\lambda}' \mathbf{c} \end{aligned} \quad (6)$$

を最小にする値として求められる。ここで

$\boldsymbol{\Sigma} = \{C(\mathbf{s}_i - \mathbf{s}_j)\}_{ij}$ 、 $\mathbf{c} = (C(\mathbf{s} - \mathbf{s}_1), \dots, C(\mathbf{s} - \mathbf{s}_n))'$ と置いた。(6)にラグランジュ乗数 m を導入して、最小化すべき目的関数を

$$\begin{aligned} \phi(\boldsymbol{\lambda}, m) &= \\ &= C(\mathbf{0}) + \boldsymbol{\lambda}' \boldsymbol{\Sigma} \boldsymbol{\lambda} - 2\boldsymbol{\lambda}' \mathbf{c} - 2m \left\{ \sum_{i=1}^n \lambda_i f(\mathbf{s}_i) - f(\mathbf{s}) \right\} \end{aligned} \quad (7)$$

とする。 $\phi(\boldsymbol{\lambda}, m)$ を $\boldsymbol{\lambda}, m$ に関して偏微分して得られる正規方程式を解いたならば、

$$\boldsymbol{\lambda} = \{\mathbf{c} + X(X' \boldsymbol{\Sigma}^{-1} X)^{-1}(f(\mathbf{s}) - X' \boldsymbol{\Sigma}^{-1} \mathbf{c})\}' \boldsymbol{\Sigma}^{-1}, \quad (8)$$

$$m = -(f(\mathbf{s}) - X' \boldsymbol{\Sigma}^{-1} \mathbf{c})'(X' \boldsymbol{\Sigma}^{-1} X)^{-1} \quad (9)$$

を得る。ただし $X = \{f(\mathbf{s}_1), \dots, f(\mathbf{s}_n)\}'$ とする。 $\mathbf{z} = \{z(\mathbf{s}_1), \dots, z(\mathbf{s}_n)\}'$ として(8)を(5)に代入すると予測値 $\hat{Z}(\mathbf{s})$ の計算式を得る。

$$\begin{aligned} \hat{Z}(\mathbf{s}) &= \sum_{i=1}^n \lambda_i z(\mathbf{s}_i) = \boldsymbol{\lambda} \mathbf{z} \\ &= \hat{\beta}_{gls} f(\mathbf{s}) + \mathbf{c}' \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \hat{\beta}_{gls} X). \end{aligned} \quad (10)$$

ただし

$$\hat{\beta}_{gls} = (X' \boldsymbol{\Sigma}^{-1} X)^{-1} X' \boldsymbol{\Sigma}^{-1} \mathbf{z} \quad (11)$$

は β の一般化最小二乗推定量である。予測値の計算に必要な(1)のパラメータ β と(4)のパラメータ θ_1, θ_2 は制限付き最尤推定法によって推定する[2][10]。詳しくは3.2節で述べる。

予測値 $\hat{Z}(\mathbf{s})$ の95%信頼区間を得るには(8)を(6)に代入すればよい。すなわち

$$\begin{aligned} \sigma^2(\mathbf{s}) = & \\ & C(\mathbf{0}) - \mathbf{c}' \Sigma^{-1} \mathbf{c} \\ & + (\mathbf{f}(\mathbf{s}) - \mathbf{X}' \Sigma^{-1} \mathbf{c})' (\mathbf{X} \Sigma^{-1} \mathbf{X})^{-1} (\mathbf{f}(\mathbf{s}) - \mathbf{X}' \Sigma^{-1} \mathbf{c}) \end{aligned} \quad (12)$$

から95%信頼区間 $[\bar{Z}(\mathbf{s}) - 1.96\sigma(\mathbf{s}), \bar{Z}(\mathbf{s}) + 1.96\sigma(\mathbf{s})]$ が求められる。ただしこの信頼区間はコバリオグラムが既知であるとの条件下で導かれたものである。本研究のようにコバリオグラムが未知の場合、 $\sigma^2(\mathbf{s})$ はより大きくなり、したがって信頼区間の範囲も広くなる性質があることに留意しておかねばならない[2][5]。

3. 適用例：千葉県内の自動車旅行時間予測

3.1 使用データ

千葉県を対象地域として任意の2地点間の最短旅行時間予測を行う。図2の点は2007年7月現在千葉県にある市役所(36カ所)を示す。北部は人口が多く道路密度も高いが、南部は平地が少なく人口は海岸地域に集中している。したがって南部地域では道路と市役所は海の近くに分布しており、半島をよぎる道路は比較的少ない。

図2の36市を出発地・目的地とするすべてのペアを取り出す。これらを $\mathbf{s}_i \in D (i=1, \dots, 1260 (=36 \times 35))$ とし、それらの最短旅行時間を観測データ $z(\mathbf{s}_i)$ とする。対象地域は平面ではないから座標と距離の扱いに少々注意を払う必要がある。市役所の緯度経度は国土数値情報ダウンロードサービス[6]と2万5千分1地形図から取得した。 \mathbf{s}_i の出発地から目的地までの距離 $f(\mathbf{s}_i)$ は、ユークリッド距離ではなく、地球を半径6370kmの球体と仮定して経度と緯度から計算される大圏距離を使用する。大圏距離とは球面上にある2点間の球面上の最短距離である。 \mathbf{s}_i の出発地と目的地の経度緯度を $(\lambda_{i1}, \phi_{i1})$ および $(\lambda_{i2}, \phi_{i2})$ とすると大圏距離 $f(\mathbf{s}_i)$ km は

$$\begin{aligned} f(\mathbf{s}_i) = & 6370 \times \arccos(\sin \phi_{i1} \cdot \sin \phi_{i2} \\ & + \cos \phi_{i1} \cdot \cos \phi_{i2} \cdot \cos(\lambda_{i1} - \lambda_{i2})) \end{aligned}$$

で与えられる[7]。4次元ユークリッド空間の部分集合Dの要素すなわち出発地と目的地の座標の組は、経度緯度を平面直角座標系(第IX系)に変換したものを使う。

最短旅行時間 $z(\mathbf{s}_i)$ はナビタイムの車ルートサービス(有料道路優先)によって取得した[11]。ナビタイム車ルートサービスとはインターネットで出発地から目的地までの自動車による最短旅行時間と経路を検索

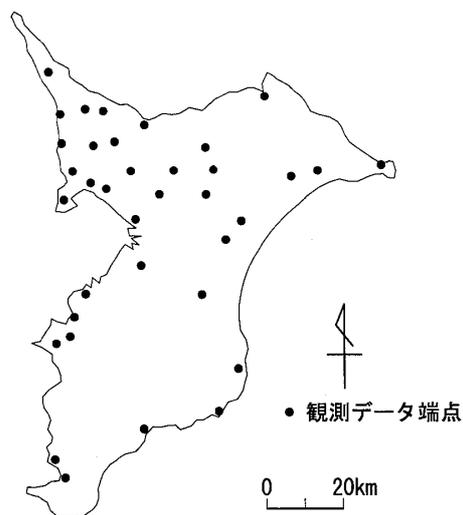


図2 千葉県の36市分布図

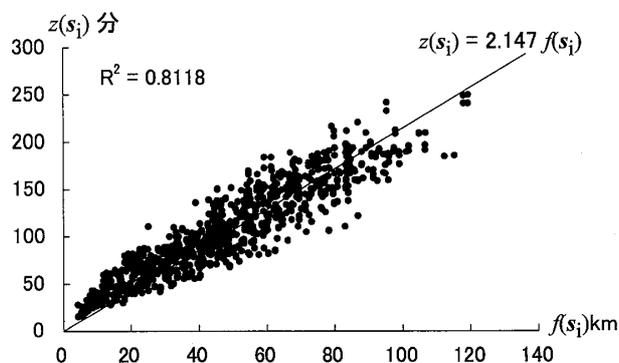


図3 大圏距離と最短旅行時間の散布図

できるサービスである。道路の設定速度の詳細は不明であるが、結果の出力から等級・幅員・上り下り別に設定されていると推測される。観測データの大圏距離と最短旅行時間の散布図を図3に示す。図3において通常の最小二乗法によって大圏距離を説明変数とする切片が0の回帰直線を求めたところ回帰係数は2.147であった。仮にこの回帰直線によって最短旅行時間を予測したならば、予測区間の95%信頼区間は、 $f(\mathbf{s})$ の大きさにもよるが $2.147f(\mathbf{s})$ を中心とした約±40分の範囲である。

3.2 コバリオグラムの推定

クリギングによって予測値を求めるためには、(1)のパラメータ β と(4)のコバリオグラム $C(\mathbf{h})$ のパラメータ θ_1, θ_2 を導出しなければならない。 $\delta(\mathbf{s})$ が正規分布に従うと仮定する。まず制限付き最尤推定法によって θ_1, θ_2 を推定し([2][10])、次に(11)によって得られる β_{gis} を β の推定値とする。

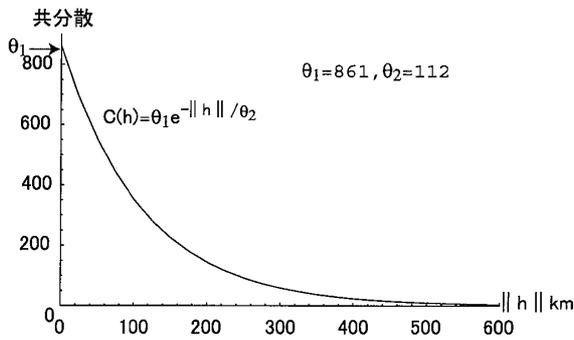


図4 推定されたコバリオグラムの図示

θ_1, θ_2 を推定するための対数尤度関数は以下のとおりである。

$$L(\theta_1, \theta_2) = -\frac{n-1}{2} \log 2\pi + \frac{1}{2} \log |X'X| - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \log |X'\Sigma^{-1}X| - \frac{1}{2} Z'\Pi Z. \quad (13)$$

ただし、 $\Pi = \Sigma^{-1} - \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$ とする。得られたパラメータ推定値は $\theta_1=861, \theta_2=112, \beta_{GIS}=1.894$ である。これらを当てはめたコバリオグラムを図4に示す。

3.3 任意の2地点間の最短旅行時間予測

千葉県内の任意の2地点間の最短旅行時間を予測する。千葉県内の3538個の郵便番号から2つを1組として100組をランダムに選び出し出発地・目的地ペアのサンプル $t_i \in D (i=1, \dots, 100)$ とする。サンプルの郵便番号に対応する住所間を移動する最短旅行時間 $z(t_i)$ をナビタイムによって求める。 $z(t_i)$ とクリギング予測値 $\hat{Z}(t_i)$ の散布図は図5のとおりである。図6は比較のため縦軸を最短旅行時間 = $2.147 \times$ 大圏距離による予測値に置き換えたものである。両図を比較すればクリギングによって予測の精度が上がっていることが見て取れる。決定係数は図6の0.8138から図5の0.9354へ上昇した。(12)によって予測値の95%信頼区間を求めると、信頼区間の幅の半分 $1.96\sigma(t_i)$ の分布は図7に示すとおりである。その平均は16.1分であり、大圏距離だけで予測した場合の約40分と比べて信頼区間の幅はかなり短くなる。

3.4 観測データ数と信頼区間の関係

予測値の信頼区間を小さくするためには計算に使用する観測データ数はなるべく多い方がよいが、計算時間や手間を考慮すれば観測データをむやみに増やすことはできない。

3.1節で用いた1260個の観測データからランダム

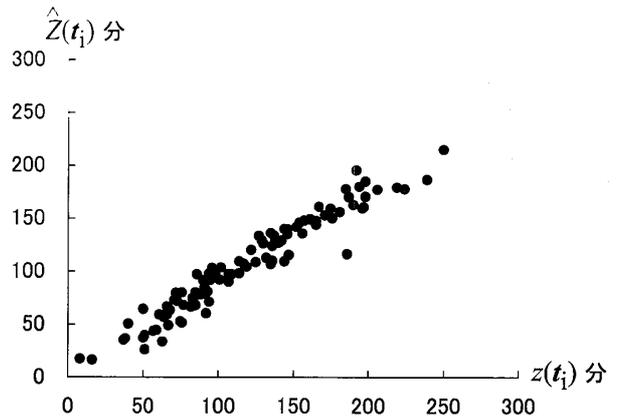


図5 サンプルの最短旅行時間とクリギング予測値の散布図

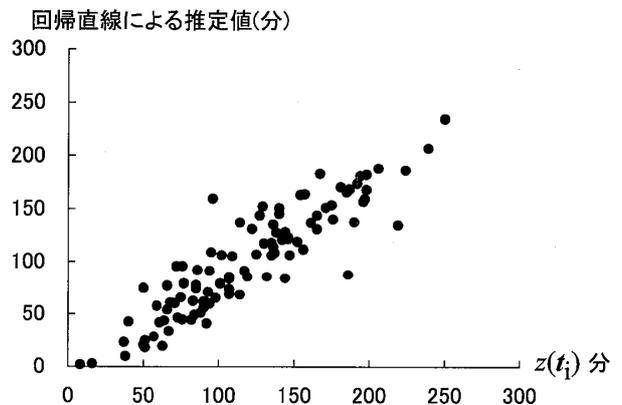


図6 サンプルの最短旅行時間と回帰直線による予測値の散布図

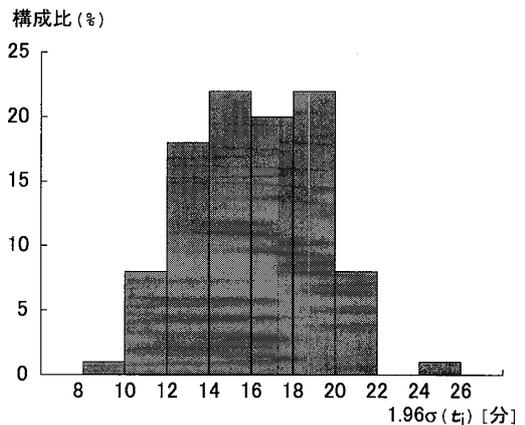
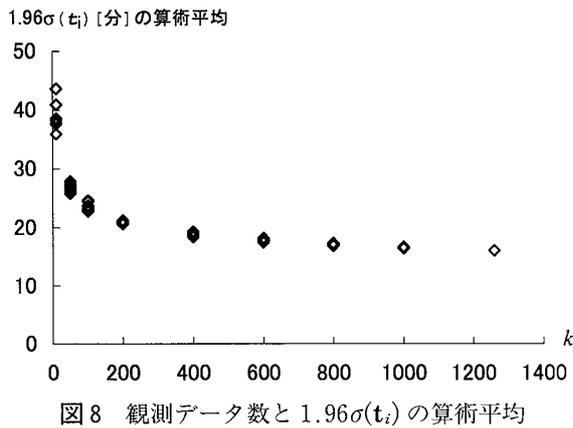


図7 $1.96\sigma(t_i)$ の分布

に k 個の観測データを取り出す。取り出された観測データを使用して、3.3節の100サンプル t_i についてあらためてクリギング予測値を計算し、さらに $1.96\sigma(t_i)$ の算術平均を求める。これを $k=10, 50, 100, 200, 400, 600, 800, 1000$ について各10回繰り返した結果を図8に表す。また、すべての観測デー



タを使用したとき ($k=n=1260$) の結果も示す。ただしコバリオグラムパラメータは3.2節で導出した $n=1260$ のときの推定値 ($\theta_1=861$, $\theta_2=112$, $\beta_{GIS}=1.894$) を共通して使用した。

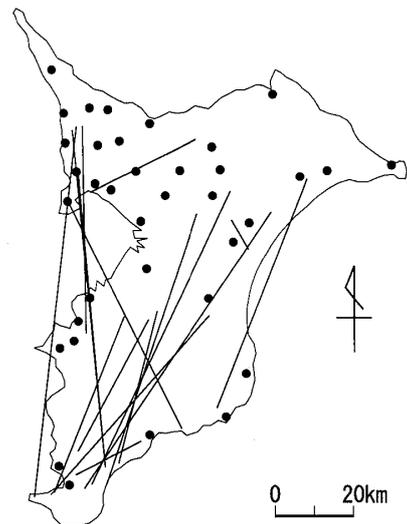
$1.96\sigma(t_i)$ の算術平均は k が大きくなるに従って小さくなるが見て取れる。 k が600を超えるあたりから低下の程度は緩やかになる。ただし、これは k に関わらず同じコバリオグラムを使用した得られた結果である。 k および取得する観測データの位置が異なれば当然ながらコバリオグラムパラメータも変化するはずであるが、ここでは簡便のためコバリオグラムを既知であるとして信頼区間の大きさを比較した。それゆえ k が小さいときの信頼区間は現実にはもっと大きくなる可能性がある。

地域が異なれば道路・山岳・河川に伴う橋の位置・湖沼などの分布も違うから、より確実なことを述べるためには他の地域で同様の実験を行う必要があるが、対象地域が都道府県程度の大きさであれば観測データは600程度使用すればよいことが推定される。

なお2節で述べたように、ここでの結果は既知でないコバリオグラムから導出されたものであるから、本当の精度はもっと悪いはずであることを付言しておく。

3.5 予測誤差の検討

3.3節の結果について予測誤差を吟味する。 $3\sigma(t_i) < |\hat{Z}(t_i) - z(t_i)|$ となる予測誤差の大きい16サンプルを取り出し、予測誤差が生じた原因について考察する。図9はこれら16サンプルの出発地と目的地を結ぶ線分を図示したものである。最短時間経路はいずれも山岳地域など道路の限られた地域を經由しているため経路は直線から大きく離れており、クリギング予測時間はナビタイムによる旅行時間よりもかなり小さい。南部地域における観測データは数が少ないうに、市役



所が海岸に近い地域に偏っていることが誤差の原因である。

これらは市役所を代表点とした観測データに基づいてクリギングを行う際に当然想定すべき誤差であった。南部の内陸地域に観測データ端点を設定しておけば、予測誤差はもっと小さくなったであろう。クリギングによって最短旅行時間を予測するためには、観測データの数だけでなくそれらの地理分布にも注意すべきことを示している。

4. おわりに

本研究では最短旅行時間の予測にクリギングの応用を試みた。さまざまな制約から適用例における観測データにはナビタイムデータを使用した。千葉県を対象とした事例計算において信頼区間は予測値を中心とする ± 16.1 分の範囲であったが、これは未知コバリオグラムによるものであることとナビタイムデータには現実の走行時間データに想定される計測誤差が含まれていないことを割り引いておく必要がある。

最短旅行時間の推定にクリギングを適用する利点の一つは、時間帯や曜日によって道路の速度が大きく変化する地域を通過する旅行時間の推定が容易にできることである。当該地域のさまざまな時間帯および曜日の最短旅行時間観測データを十分に取得することができれば、クリギング計算に使用する観測データを出発する時間帯に対応して交換することによって精度の高い最短旅行時間推定が可能となる。

さらなる予測精度向上のためには、道路ネットワークデータの利用を検討すべきであろう。例えば出発地

から目的地までの距離 $f(s)$ として使用した大圏距離を、最短距離経路の道路距離に取り替えれば、もっと信頼区間を小さくできるかもしれない。

一方でクリギングによる予測では旅行時間に対応した経路が分からないという大きな欠点がある。さしあたっては最短距離経路などで代用するにせよ、予測旅行時間から経路を推定する方法を考えるなど対策を講じなければならない。

本研究の地図作製には MANDARA および Illustrator を使用した。計算は主として Mathematica を用いた。

謝辞 本研究は科学研究費補助金「都市内・都市間交通網と公共施設配置に関する数理的ならびに実証的研究」(基盤研究(A), 課題番号 17201037, 研究代表者 南山大学伏見正則教授) の補助を受けた。本稿をまとめるにあたり、東京大学大学院工学系研究科の井上亮先生、筑波大学大学院システム情報工学研究科の堤盛人先生、ならびに本誌査読者から貴重なご示唆をいただいた。ここに謝意を表します。

参考文献

- [1] H. Braxmeier, V. Schmidt and E. Spodarev: "Kriged road-traffic maps," In: J. Pilz (ed.) *Interfacing Geostatistics, GIS and Spatial Data Bases, Proceedings of the International Conference, StatGIS 2003*, Springer-Series: *Advances in Spatial Science* (2005), 39-52.
- [2] N. Cressie: *Statistics for Spatial Data. Revised Edition*, John Wiley & Sons, 29-170, 1993.
- [3] J. Du and L. Aultman-Hall: "Using Spatial Analysis to Estimate Link Travel Times on Local Roads," *Transportation Research Board 2006 Annual Meeting CD-ROM*, 06-0676 (2006).
- [4] H. Wackernagel 原著:『地球統計学』, 森北出版, 2003.
- [5] 井上亮, 木越尚之, 清水英範:「時空間クリギングの地価推定への適用可能性の検討」, 『地理情報システム学会講演論文集』, 14 (2005), 39-42.
- [6] 国土交通省: 国土数値情報ダウンロードサービス, <http://nlftp.mlit.go.jp/ksj>
- [7] 国土交通省国土地理院 監修:『数値地図ユーザーズガイド (第2版補訂版)』, 日本地図センター, 1998.
- [8] 杉浦淳徳, 島川陽一:「Kriging 法を用いた土地価格の推定と空間分析」, 『高等専門学校教育と研究: 日本高専学会誌』, Vol. 10, No. 4 (2005), 42-49.
- [9] 瀬谷創, 堤盛人, 井上亮, 石田東生, 岡本直久:「covariogram に移動平均モデルを用いた時空間 Kriging」, 応用地域学会, 12月9日~10日, 2006, 広島.
- [10] 丹後俊郎:『統計モデル入門』, 朝倉書店, 2000.
- [11] ナビタイムジャパン:ナビタイム車ルート検索, <http://www.navitime.co.jp/drive/>
- [12] 間瀬茂, 武田純:『空間データモデリング』, 共立出版, 135-151, 1999.