

1. はじめに

本研究は2地点間の最短旅行時間予測にクリギングを適用する。図1のように相互の距離が近い平面上の3点 P_A, P_B, P_C ならびに P_X, P_Y, P_Z があるとす。 P_C から P_Z まで自動車で移動するときの最短旅行時間を予測したい。 P_A から P_X までおよび P_B から P_Y までの最短旅行時間データがあるならば、最短旅行時間が直線距離におおよそ比例することに加えて、これら4地点と P_C, P_Z 相互の距離が推測の手がかりとなる。

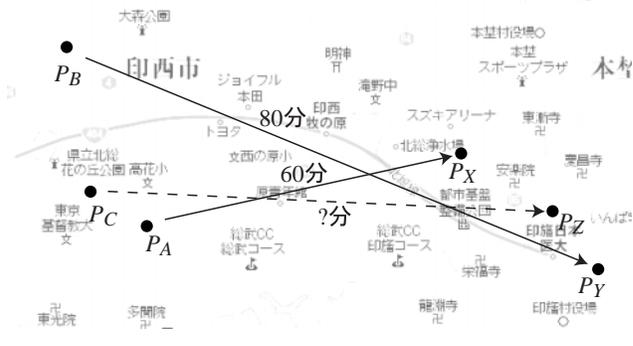


図 1: 未知の旅行時間の推測

クリギングとはこのような地理データに基づいた推測を統計学の枠組みで展開した手法である。距離が近いほど似た値を示す性質をもつ空間データに対して、既知観測データの空間分布に基づいて未知データを推定することができる。クリギングはもともと地下の鉱床の空間分布を推定する鉱床学的手法として発展し、近年では空間分布を取り扱う分野で広く応用されている [1]。都市地域分野では地価の予測や細街路の自動車速度の推定などの既往研究がある [1][3][5][7]。

本研究では、出発地と目的地の位置を表すそれぞれの座標によって構成される4次元空間においてクリギングの一手法である普遍クリギングを適用する。 $P_A, P_X, P_B, P_Y, P_C, P_Z$ それぞれの最短旅行時間経路は、経験的に P_A, P_B, P_C および P_X, P_Y, P_Z が相互に近いほど経路の重複部分が増える。迂回の程度や移動速度も類似し、従って両経路の旅行時間は似かよった値になる。このことは、平面上の2地点間の移動を出発地と目的地の位置座標で構成される4次元空間における点で表すならば、点間の距離が近いほど似た旅行時間になると言い換えられる。それゆえ、この4次元空間でクリギングの利用による最短旅行時間予測が可能であると言える。

2地点間の最短旅行時間を予測するためには、道路網ネットワークデータにダイクストラ法などの経路探索を適用する方法が一般的である。精度の良い予測を行うためには各道路に付与される速度が重要である。

速度は道路の等級や幅員あるいは通過交通量に応じて定められることが多いが、機械的に定めると都市部などで実態にそぐわない予測値が出ることもある。これを避けるためには詳細な速度調査に基づく設定が必要になる。

本研究で述べるクリギングを利用した最短旅行時間予測は、道路ごとに速度を設定した道路網データの代わりに、複数の地点間の最短旅行時間観測データを使用する。従って予測精度を高めるためには信頼性の高い観測データを数多く用意しなければならない。もし観測データに最短とは言えない旅行時間データが多数含まれているならば、予測値は最短旅行時間として信用できないからである。

先に少し触れたようにクリギングの適用範囲は幅広いが、旅行時間の予測に適用した既往研究に限ればそれほど多いとは言えない。Braxmeier, H., Schmidt, V. and Spodarev, E. [1] および Du, J. and Aultman-Hall, L. [3] は、GPSによって計測された道路の自動車走行速度を観測データとして使用し、データが取得できない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。ただしこれらは4次元空間でクリギングを適用した研究ではない。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、平面上で推定が行われている。

2. 普遍クリギングによる旅行時間の予測

本節では Cressie [2] と間瀬, 武田 [11] 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

みうらひでとし 明海大学不動産学部
〒 279-8550 千葉県浦安市明海 1-2
hmiura@meikai.ac.jp

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 $s \in D$ について，2地点間の旅行時間 $z(s)$ が，

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

という構造を持つと仮定する．ただし $f(s)$ 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， $\delta(s)$ は次の性質を満たすと仮定する．

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

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

$E(\delta(s)) = \text{定数}$ ，および (3) が成り立つとき， $\delta(s)$ は二次定常性を満たすという．これは任意の $s \in D$ について $\delta(s)$ の平均が一定であり，相対的位置が \mathbf{h} 離れたときの $\delta(s)$ と $\delta(s + \mathbf{h})$ 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．

$$C(\mathbf{h}) = \begin{cases} \theta_0 + \theta_1, & \|\mathbf{h}\| = 0, \\ \theta_1 e^{-\|\mathbf{h}\|/\theta_2}, & \|\mathbf{h}\| > 0. \end{cases} \quad (4)$$

ただし $\theta_0, \theta_1, \theta_2$ はパラメータである． $\theta_0 \geq 0, \theta_1 \geq 0, \theta_2 > 0$ とする．既往研究 (例えば [5][8]) では spherical 型と呼ばれる型が使用されることが多いが，これは対象空間の次元が3を超えると適用できない．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 $s_i \in D$ ，旅行時間を $z(s_i)$ とする ($i = 1, \dots, n$)．クリギングによる任意の s に関する旅行時間予測値 $\hat{Z}(s)$ は観測データ $z(s_i)$ の重み付き平均

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

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 は正值とは限らない)．パラメータベクトル $\lambda = (\lambda_1, \dots, \lambda_n)'$ は平均二乗予測誤差

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

を最小にする値として求められる．ここで

$\Sigma = \{C(s_i - s_j)\}_{ij}$ ， $c = (C(s - s_1), \dots, C(s - s_n))'$ 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

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

とする． $\phi(\lambda, m)$ を λ, 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，

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

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

を得る．ただし $X = \{f(s_1), \dots, f(s_n)\}'$ とする． $z = \{z(s_1), \dots, z(s_n)\}'$ として (8) を (5) に代入すると予測値 $\hat{Z}(s)$ の計算式を得る．

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

ただし

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

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 $\theta_0, \theta_1, \theta_2$ は制限付き最尤推定法によって推定する [2][9]．詳しくは3.2節で述べる．

予測値 $\hat{Z}(s)$ 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

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

から 95 %信頼区間 $[\hat{Z}(s) - 1.96\sigma(s), \hat{Z}(s) + 1.96\sigma(s)]$ が求められる。

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ヶ所) 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。従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・目的地とする全てのペアを取り出す。これらを $s_i (i = 1, \dots, 1260 (= 36 \times 35))$ とし、それらの最短旅行時間を観測データ $z(s_i)$ 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。 s_i の出発地と目的地の合わせて 4 つの座標値は、緯度と経度を平面直角座標系 (第 I X 系) に変換したものを使用する。出発地から目的地までの距離 $f(s_i)$ は、ユークリッド距離ではなく、地球を 6370km の球体と仮定して経度と緯度から計算される大円距離を使用する。

最短旅行時間 $z(s_i)$ 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 [10]。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できるサービスである。道路の設定速度は不明であるが、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。観測データの大円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円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、 $f(s)$ の大きさにもよるが $2.147f(s)$ を中心とした約 ± 40 分の範囲である。

3.2. セミ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 $\theta_0, \theta_1, \theta_2$ を導出しなければならない。 $\delta(s)$ が正規分

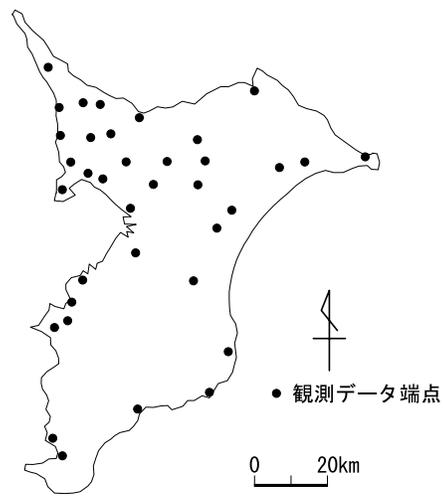


図 2: 千葉県の 36 市分布図

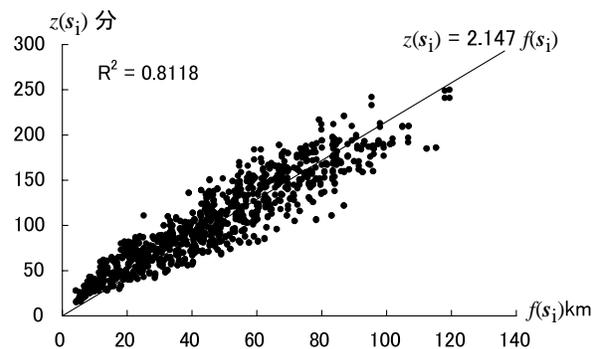


図 3: 大円距離と最短旅行時間の散布図

布に従うと仮定する。まず制限付き最尤推定法によって $\theta_0, \theta_1, \theta_2$ を推定し ([2][9])、次に (11) によって得られる β_{gls} を β の推定値とする。

$\theta_0, \theta_1, \theta_2$ を推定するための対数尤度関数は以下のとおりである。

$$L(\theta_0, \theta_1, \theta_2) = -\frac{n-1}{2} \log 2\pi + \frac{1}{2} \log |\mathbf{X}'\mathbf{X}| - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \log |\mathbf{X}'\Sigma^{-1}\mathbf{X}| - \frac{1}{2} \mathbf{Z}'\Pi\mathbf{Z}. \quad (13)$$

ただし、 $\Pi = \Sigma^{-1} - \Sigma^{-1}\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}$ とする。

得られたパラメータ推定値は $\theta_0 = 0, \theta_1 = 861, \theta_2 = 112$ $\beta_{gls} = 1.894$ 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

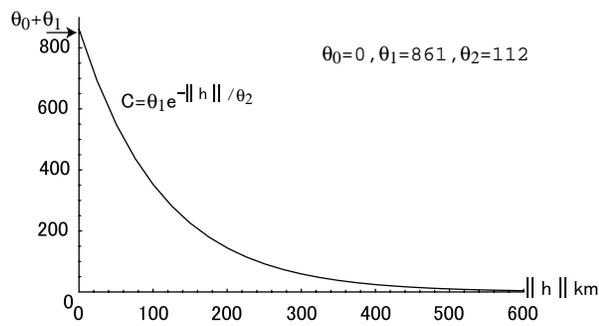


図 4: 推定されたコバリオグラムの図示

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 $t_i (i = 1, \dots, 100)$ 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 × 大円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、大円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．

なお (5) から明らかなように、観測データである 36 市役所間の最短旅行時間 $z(s_i)$ はクリギングによる最短旅行時間予測値 $\hat{Z}(s_i)$ と完全に一致する．

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

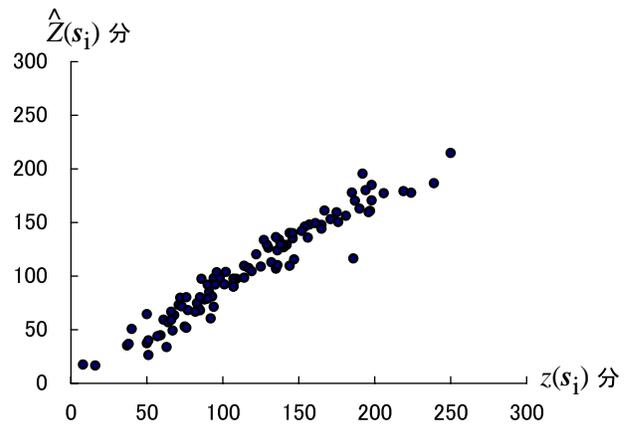


図 5: サンプルの最短旅行時間とクリギング予測値の散布図

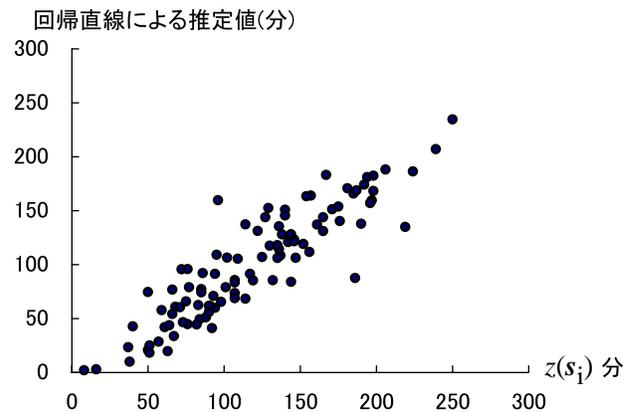


図 6: サンプルの最短旅行時間と回帰直線による予測値の散布図

値 ($\theta_0 = 0, \theta_1 = 861, \theta_2 = 112, \beta_{gls} = 1.894$) 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．

この結果から、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 程度使用すればよく、それを超えて観測データを用意しても予測信頼区間は大幅に小さくすることはできないことが分かる．

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

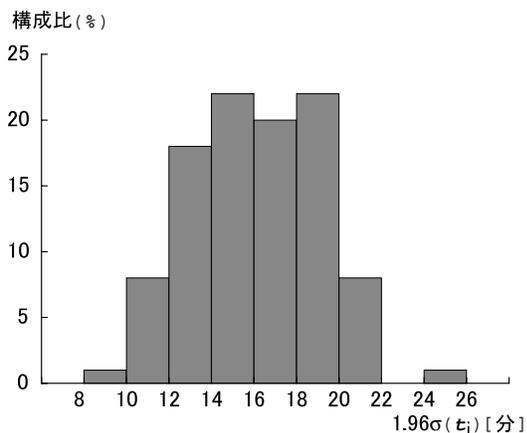


図 7: $1.96\sigma(t_i)$ の分布

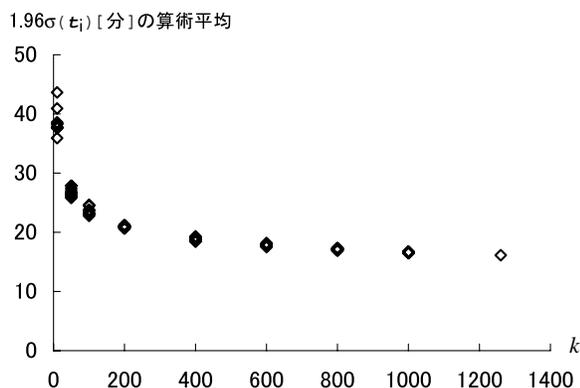


図 8: 観測データ数と $1.96\sigma(t_i)$ の算術平均

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えに，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

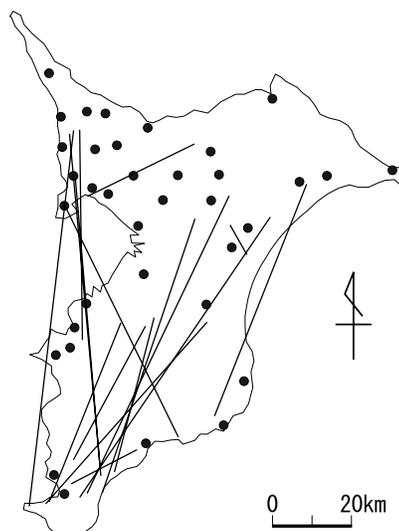


図 9: 予測誤差の大きいサンプル

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，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．

予測精度向上のために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 $f(s)$ として使用した大円距離を，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) 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と筑波大学大学院システム情報工学研究科の堤盛人先生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] Braxmeier, H., Schmidt, V. and Spodarev, E. :
"Kriged road-traffic maps" ,*In: J. Pilz (ed.) Interfacing Geostatistics, GIS and Spatial Data Bases, Proceedings of the International Conference, Stat-GIS 2003, Springer-Series: Advances in Spatial Science*(2005), 39-52.
- [2] Cressie, N. : *Statistics for Spatial Data. Revised Edition*, John Wiley & Sons, 29-170, 1993.
- [3] Du, J. and Aultman-Hall, L. : "Using Spatial Analysis to Estimate Link Travel Times on Local Roads" , *Transportation Research Board 2006 Annual Meeting CD-ROM* , 06-0676(2006).
- [4] Wackernagel, H 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] 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 」 『高等専門学校の教育と研究 : 日本高専学会誌』 , Vol.10, No.4(2005) , 42-49 .
- [8] 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.
- [9] 丹後俊郎 : 『統計モデル入門』 , 朝倉書店, 2000 .
- [10] 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/>
- [11] 間瀬 茂 , 武田 純 : 『空間データモデリング』 , 共立出版 , 135-151 , 1999 .