

1. はじめに

最近では、空間的な相関性を有する地盤物性値の分布を、地盤統計学(geostatistics)¹⁾における確率場の概念を用いてモデル化することが一般的となってきた。互いに相関性を有する2つ以上の地盤物性値の空間分布推定において、対象とする物性値のデータが少数しか得られない場合に、他の物性値を補助データとするCokriging手法が有効である。著者らは、杭基礎の基盤面レベル推定において、基盤面が地形の起伏に類似することが多いことに着目して、地表面レベルを補助データとしたCokrigingを提案、推定精度向上に有効であることを示した²⁾。

KrigingやCokrigingを用いて空間分布を推定するためには、2段階の推定問題を解く必要がある。一つはKrigingやCokrigingの前提条件となる確率場モデルの同定であり、もう一つは求められた確率場モデルから標本値を条件としたKrigingやCokrigingなどによる空間分布の推定である。ここで後者の空間分布推定は、標本値を条件とした条件付き確率場を定めているに過ぎず、その推定精度は前提となる確率場モデルの善し悪しに依存する。すなわち前者の確率場モデルの同定が、空間分布推定の精度にとって大きな要因であり、限られた標本データから確率場のパラメータを精度良く推定することは非常に重要である。しかしながら実際の適用例の多くは、この問題をあまり議論せずに、一般的に用いられているセミバリオグラムを最小二乗法により関数近似するという古典的な手法³⁾で確率場モデルを推定しているものがほとんどである。

Cokrigingによる空間分布推定の場合、多変量の確率場としてモデル化することになり、各々の物性値のセミバリオグラムの他に各物性値間の相互セミバリオグラムの定義が必要である。相互セミバリオグラムは、

$$\gamma_{12}(h) = \frac{1}{2} E\{[z_1(\mathbf{x}+h) - z_1(\mathbf{x})][z_2(\mathbf{x}+h) - z_2(\mathbf{x})]\} \quad (1)$$

と表され、この式から明らかなように、相互セミバリオグラムを積率法により求めるためには、各物性値の標本データが、同じ位置で得られている必要がある。すなわち、ある位置 $\mathbf{x}$ で1つでも標本データが得られていない物性値があれば、その位置のデータは推定に用いることができないのである。このため実際問題では標本データが特に限られ、セミバリオグラムの推定精度が悪く、計算自体が不安定となる。このことがCokriging手法があまり広く用いられていない理由の一つである。

杭基礎地盤面レベルと地表面レベルのように、一般に相関性を有する2つ以上の地盤物性値は、一方は限られたデータ数であるのに対して、もう一方のデータは数多く得られることがある。このため統計量を推定する際にも、これら全ての情報を用いることができれば、モデルの推定精度が良くなり、Cokrigingによる空間分布推定の精度も向上する。このような考えを取り入れたこれまでの研究では、各変量相互の相関特性の定式化として、同じ位置に標本を必要としない擬似クロスバリオグラム(pseudo-cross-variogram)を用いた研究^{3a)}や一般化相互共分散(generalized cross-covariance)を導入した研究^{3b)}などがあるが、いずれも積率法で求めたセミバリオグラムに関数形を当てはめて最小二乗法を行うものである。

データが不規則に分布していても比較的安定した計算が行え、またAICによりモデル適合評価が容易な推定法が、最尤法である。確率場モデル推定に最尤法を用いた研究⁴⁾はいくつか存在し、最近ではモデル推定における最尤法の優位性を数値実験により示した研究⁵⁾などもあるが、いずれも1変量の確率場を扱ったものであり、多変量問題への適用例は見あたらない。

そこで本研究では、多変量確率場モデルの推定に最尤法を導入し、必ずしも同じ位置に全てのデータが存在しない問題に一般化した定式化を提案し、数値実験により推定精度の検証を行った。

2. 最尤法による2変量確率場モデルの推定

ここでは2変量 z_1, z_2 の確率場として、一般的な次式を考える。

$$\begin{Bmatrix} Z_1(\mathbf{x}) \\ Z_2(\mathbf{x}) \end{Bmatrix} = \begin{Bmatrix} \mu_1(\mathbf{x}) \\ \mu_2(\mathbf{x}) \end{Bmatrix} + \begin{Bmatrix} \varepsilon_1(\mathbf{x}) \\ \varepsilon_2(\mathbf{x}) \end{Bmatrix} \quad (3)$$

ここで $\mu_1(\mathbf{x}), \mu_2(\mathbf{x})$ はトレンド成分で $\mathbf{x}$ の多項式 $\mathbf{X}$ を基にした関数 $\mathbf{X}\mathbf{b}$ であり、 $\varepsilon_1(\mathbf{x}), \varepsilon_2(\mathbf{x})$ はランダム成分で、平均値0、自己共分散、相互共分散が、それぞれ次式のように存在すると仮定する。

$$\mathbf{K}_{11}(h) = E[\varepsilon_1(\mathbf{x}+h) \cdot \varepsilon_1(\mathbf{x})^T]$$

$$\mathbf{K}_{22}(h) = E[\varepsilon_2(\mathbf{x}+h) \cdot \varepsilon_2(\mathbf{x})^T]$$

$$\mathbf{C}_{21}(h) = E[\varepsilon_2(\mathbf{x}+h) \cdot \varepsilon_1(\mathbf{x})^T]$$

$\{z_i(\mathbf{x}_j), i=1, \dots, N_1\}$ および $\{z_i(\mathbf{x}_j), i=1, \dots, N_2\}$ の観測値(N_1+N_2)個から同時確率密度関数である尤度が次式で与えられる。

$$p(\mathbf{z}|\boldsymbol{\theta}) = (2\pi)^{-(N_1+N_2)/2} |\mathbf{Q}|^{1/2} \exp\left\{-\frac{1}{2}(\mathbf{z}-\mathbf{X}\mathbf{b})^T \mathbf{Q}^{-1}(\mathbf{z}-\mathbf{X}\mathbf{b})\right\} \quad (4)$$

$$\mathbf{z} = \begin{matrix} N_1 \uparrow \\ N_2 \downarrow \end{matrix} \begin{Bmatrix} z_1 \\ z_2 \end{Bmatrix}, \quad \mathbf{Q} = \begin{matrix} N_1 \leftrightarrow & \leftrightarrow N_2 \\ \begin{Bmatrix} \mathbf{K}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{K}_{22} \end{Bmatrix} \end{matrix}$$

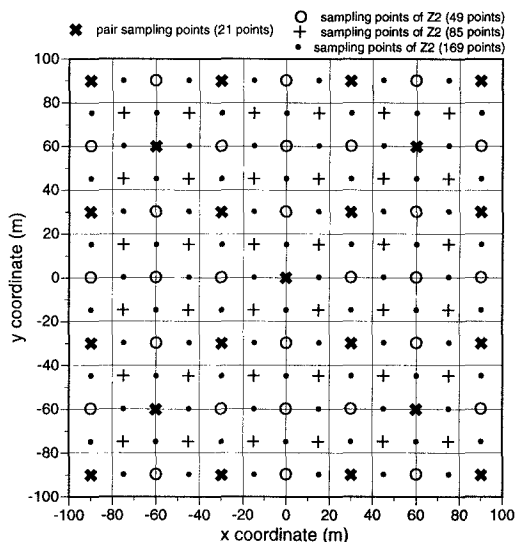


図1 数値実験モデル(データ抽出位置)

Keywords: 地盤統計学, 最尤法, 相互共分散, 確率場, Cokriging

連絡先: 〒100-0011 東京都千代田区内幸町2-2-2 富国生命ビル 清水建設(株) TEL03-3508-8101 FAX03-3508-2196

θ は推定すべきパラメータであり、次式の負の対数尤度を最小化することにより決定する。

$$L(\mathbf{z}|\theta) = \frac{(N_1 + N_2)}{2} \ln(2\pi) + \frac{1}{2} \ln|\mathbf{Q}| + \frac{1}{2} (\mathbf{z} - \mathbf{X}\mathbf{b})^T \mathbf{Q}^{-1} (\mathbf{z} - \mathbf{X}\mathbf{b}) \quad (5)$$

以上の定式化からわかるように、2つの観測値の位置および数は一致している必要はなく、全く一致しない場合でも推定が可能である。

3. 数値実験による検証

提案手法の検証を行うために数値実験を実施した。検証の目的は、 z_1, z_2 が共に観測される位置に加えて、 z_1 だけが観測される位置を設定し、 z_2 の観測値データだけを追加して推定精度の向上を示すことである。数値実験はモンテカルロシミュレーションにより、 $100\text{m} \times 100\text{m}$ 、5m間隔格子点において2変量確率場を30個作成し、各々最尤推定計算を実施した。確率場は共分散関数を、相関距離(range)20m、分散30、2変量間の相関係数0.5とした次式で示す指数関数型モデルを用いた。

$$\mathbf{K}_{11}(\mathbf{h}) = \mathbf{K}_{22}(\mathbf{h}) = 30 \cdot \exp\{-\mathbf{h}/20\}$$

$$\mathbf{C}_{12}(\mathbf{h}) = \mathbf{C}_{21}(\mathbf{h}) = 0.5 \cdot 30 \cdot \exp\{-\mathbf{h}/20\}$$

平均値は一定として、 z_1 を0、 z_2 を50とした。

図1は、推定に用いるデータ抽出位置を示すもので、*印が z_1, z_2 がペアで観測される21点である。検討ケースは、その21点だけを用いて推定するケースに加え、○印の z_2 データを加えて z_2 の標本数を計49点としたケース、+印を加えて計85点としたケース、■印を加えて計169点とした4ケース実施した。

推定するパラメータは、 z_1, z_2 各々の平均および分散、相関距離および相関係数の6個である。このうち空間特性を表す上で、最も重要なパラメータである相関距離の推定値30個について、ヒストグラムで表したものが図2である。ペアの観測点21点のデータで推定した結果は、非常にばらつきが大きいのにに対して、 z_2 のデータを追加したケースでは追加数に応じて、ばらつきが小さくなっている。推定値の平均も17.9mから19.3mと真値(20m)に近くなり、精度が向上している。表1は各パラメータの推定結果を平均と標準偏差で示したものである。この表から追加するデータ z_2 に関連するパラメータ(z_2 の平均、分散)は、相関距離と同じく追加数に応じて、推定値のばらつきが小さくなり平均も真値に近づいて推定精度が向上しているのがわかる。またそれ以外のパラメータ(z_1 の平均、分散、相関係数)についても、若干ではあるがばらつきが小さくなり、推定精度が向上している。

4. おわりに

これまでCokrigingのための多変量確率場の推定には、各々の変量がそれぞれ同じ位置で観測されたデータしか用いることができなかった。このため実問題ではデータ数が限られ、精度の良い推定ができないという問題点があった。本研究では最尤法を用いて、各変量が任意の位置で観測されたデータを用いても推定可能な、2変量確率場推定法の定式化を提案し、数値実験によりその推定精度の検証を行った。その結果、これまでは利用できなかった一方の変量しか得られないデータを取り込むことで、推定精度を向上させることができた。

参考文献

- 1) Matheron, G.: Principles of geostatistics., Economic Geology, Vol.58, pp.1246-1266, 1963.
- 2) 本多・鈴木・上田・近藤: 地形情報を用いた基礎地盤面のモデル化と推定, 土木学会論文集, No.561/III-38, pp.63-74, 1997.
- 3) Clark, I., et al.: MUCK-A Novel Approach to Co-Kriging. in B. E. Buxton (Ed.), Proc. of the Conference on Geostatistical, Sensitivity, and Uncertainty: Methods for Ground-Water Flow and Radionuclide Transport Modeling: Batelle Press, Columbus, p. 473-494, 1989.
- 4) Myers, D. E.: Pseudo-cross variograms, positive-definiteness, and cokriging, Math. Geol., Vol.23, pp.805-816, 1991.
- 5) Künsch, H. R., Parpriz, A., and Bassi, F.: Generalized cross-covariances and their estimation, Math. Geol., Vol.29, pp.779-799, 1997.
- 6) Kitanidis, P. K.: Statistical estimation of polynomial generalized covariance functions and hydrologic applications, Water Resources Res., Vol.19, No.2, pp.909-921, 1983.
- 7) Pardo-Igúzquiza, E.: Maximum likelihood estimation of spatial covariance parameters, Math. Geol., Vol.30, No.1, pp.95-108, 1998.

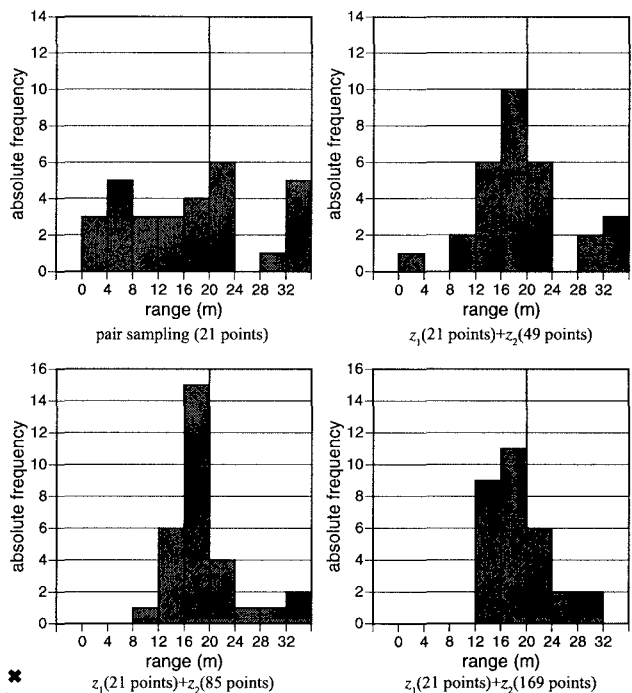


図2 推定結果のヒストグラム（相関距離）

表1 推定結果の一覧表（30個の平均と標準偏差）

	variance of z1	variance of z2	correlation coefficient	range	mean of z1	mean of z2
pair sampling (21points)	29.25	26.95	0.44	17.91	-0.25	49.66
	11.60	8.03	0.16	11.70	1.44	1.39
z1(21points)+z2(49points)	29.37	28.93	0.46	19.03	-0.27	49.75
	10.23	7.31	0.15	7.37	1.45	1.32
z1(21points)+z2(85points)	29.34	29.94	0.46	19.15	-0.28	49.70
	10.05	6.03	0.15	5.22	1.46	1.33
z1(21points)+z2(169points)	29.63	29.89	0.46	19.26	-0.30	49.68
	10.71	5.26	0.15	4.03	1.43	1.22

upper: mean of 30 estimates / lower: std.dev. of 30 estimates