IV-207

統計モデルにおける誤差項の分散共分散行列の構造化に関する代数的考察

東京大学大学院　正会員　堤　盛人
東京大学大学院　正会員　清水英範
東京大学大学院　学生会員　井出裕史

1. はじめに

社会資本整備が地域へ及ぼす諸影響を客観的かつ定量的に把握するために様々なモデル分析が行われるが、分析に当たってはモデルを統計モデルとして記述し、実際のデータを用いてそのパラメータを推定する必要がある。誤差の確率密度関数を明示的に仮定するか否かに関わらず、通常、誤差項の１次モーメントである期待値、２次モーメントである分散共分散に何らかの仮定がおかれる。期待値については一般性を失わず０とすることができるが、分散共分散行列についてはその構造を事前に知ることができないため、多くの場合単位行列の定数倍が仮定される。しかし、その仮定が妥当でない時には、推定パラメータの信頼性が損なわれるといった問題が生じるため、統計学やその応用分野において様々な工夫がこらされてきたが、それらの相互関係は必ずしも明確に理解できるものではない。

本稿では、誤差の分散共分散行列についてのモデル化手法について、それらの相互関係を明確にすることを目的とした基本的考え方を示す。

2. 分散共分散行列の構造化

以下、表記上の簡単化のため、y を被説明変数、$x_h\ (h=1,\cdots,m)$ を説明変数、$\beta_h\ (h=0,1,\cdots,m)$ をパラメータとする線形モデルについて、誤差項を含んだ統計モデルを以下のように行列表記する。

$$\boldsymbol{y} = \boldsymbol{X\beta} + \varepsilon \qquad \cdots (1)$$

$$\boldsymbol{y} = (y_1,\cdots,y_i,\cdots,y_n)^t,\quad \boldsymbol{\beta} = (\beta_0,\beta_1,\cdots,\beta_m)^t,\quad \varepsilon = (\varepsilon_1,\cdots,\varepsilon_i,\cdots,\varepsilon_n)^t,\quad \boldsymbol{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{m1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{mn} \end{pmatrix}$$

ここで i は分析の対象とする地域の地点もしくはゾーンの番号、$\boldsymbol{\varepsilon}$ は誤差項である。最小二乗法や最尤法の適用に際しては、通常、次の仮定を満たすものと仮定される。

$$E(\varepsilon) = \boldsymbol{0} \qquad \cdots (2)$$

$$\boldsymbol{\Sigma} \equiv Var(\varepsilon) = \sigma^2 \boldsymbol{I} \qquad \sigma^2 ：定数 \qquad \cdots (3)$$

地域を対象としたモデル分析においては、$\boldsymbol{\Sigma}$ の非対角項が０にならない系列相関や、非対角項が０であっても対角要素の大きさが異なる分散不均一性が生じることが少なくない。そのような場合の対処法として、次のような誤差項の構造化が行われる。

(a) $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} = \lambda \boldsymbol{W\varepsilon} + \boldsymbol{\mu}$ $\cdots (4)$

(b) $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} = \lambda \boldsymbol{W\mu} + \boldsymbol{\mu}$ $\cdots (5)$

(c) $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} = \boldsymbol{Wv} + \boldsymbol{\mu}$ $\cdots (6)$

$\boldsymbol{\mu} = (\mu_1,\cdots,\mu_n)^t, \boldsymbol{v} = (v_1,\cdots,v_n)^t$ ：確率変数

$Var(\boldsymbol{\mu}) = \sigma_\mu^2 \boldsymbol{I}, Var(\boldsymbol{v}) = \sigma_v^2 \boldsymbol{I}, E(\boldsymbol{\mu v}^t) = 0$ $\cdots (7)$

$\lambda, \sigma_\mu^2, \sigma_v^2$ ：定数

$\boldsymbol{W} = (w_{ij}),\ w_{ij} = c_j / d_{ij}^\alpha\ (i \neq j),\quad w_{ii} = 0$ $\cdots (8)$

d_{ij} ：地点 ij 間の空間距離 d_{ij}

α ：定数　c_j ：$\sum_i \mathrm{w}_{ij} = 1$ とする規格化定数

(a)〜(c)のようなモデルは、誤差修正モデル（error correction model : ECM）と呼ばれる。これらは仮定から、誤差の分散共分散行列を直接以下のように構造化したと考えても同じである。

(a)' $\boldsymbol{\Sigma}_a = \sigma_\mu{}^2 \left[(\boldsymbol{I} - \lambda \boldsymbol{W})^t (\boldsymbol{I} - \lambda \mathbf{W}) \right]^{-1}$ $\cdots (9)$

(b)' $\boldsymbol{\Sigma}_b = \sigma_\mu{}^2 \left[(\boldsymbol{I} + \lambda \boldsymbol{W})^t (\boldsymbol{I} + \lambda \mathbf{W}) \right]$ $\cdots (10)$

(c)' $\boldsymbol{\Sigma}_c = \sigma_v{}^2 \boldsymbol{W}^t \boldsymbol{W} + \sigma_\mu{}^2 \boldsymbol{I}$ $\cdots (11)$

3. 分散共分散行列の空間構造 [1),2)]

分散共分散行列 $\boldsymbol{\Sigma}$ は、$\det \boldsymbol{\Sigma} \neq 0$ に限定すると一般に正値対称行列である。（$\det \boldsymbol{\Sigma} = 0$ の場合についてはここでは扱わない。なお、本章の議論についての詳細は、参考文献(1) を参照されたい。）

キーワード：分散共分散行列、構造化、正値対称行列

連絡先　：　〒113　文京区本郷７－３－１　東京大学大学院工学系研究科　社会基盤工学専攻
TEL：03-3812-2111（代）　内線 6128　FAX：03-5689-7290

正値対称行列のなす集合を $PD(n)$ とする。

$$PD(n)=\{A\in GL(n,\boldsymbol{R})\,|\,x^tAx\geq 0,\forall x\in\boldsymbol{R}^n,$$

$$\mathbf{x}^tA\mathbf{x}=0 \text{ は } x=\boldsymbol{0} \text{ の時にのみ成立 }\} \quad \cdots(12)$$

$GL(n)$：一般線形群

$$GL(n)=\{A\in M(n,n)\,|\det A\neq 0\}^{3)} \quad \cdots(13)$$

対称行列であるから、$\boldsymbol{P}=(P^{kl})\in PD(n)$ に対し、$k\leq l$ なる成分 $\{P^{kl}\}_{k\leq l}$ のみを考えれば十分であり、自然な基底として $\{E_{kl}\}$ を導入することができる。ここで、$\{E_{kl}\}$ は、(k,l)成分と(l,k)成分が 1 でそれ以外は 0 な対称行列である。

$$\boldsymbol{P}=\sum_{1\leq l\leq n}\sum_{1\leq k\leq n}\boldsymbol{P}^{kl}\boldsymbol{E}_{kl} \quad \cdots(14)$$

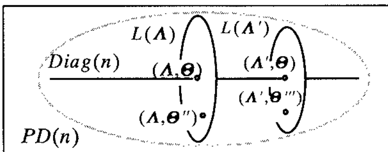
$\boldsymbol{P}\in PD(n)$ は、対角行列 $\boldsymbol{\Lambda}$ と直行行列 $\boldsymbol{\Theta}\in O(n)$（$O(n)=\left\{A\in GL(n)\,\middle|\,A^tA=\boldsymbol{I}\right\}$：直交行列群）を用いて固有値分解でき、$PD(n)$が対角行列と直交行列の組 $(\boldsymbol{\Lambda},\boldsymbol{\Theta})$ でパラメータ付けできる。分解の一意性として次の制約を付加する。

$$Diag(n)\equiv\left\{\Lambda\,\middle|\,\boldsymbol{\Lambda}=diag(\lambda_1,\cdots,\lambda_n),\lambda_1>\cdots\lambda_n>0\right\} \quad \cdots(15)$$

これにより、$PD(n)$ は $P=(\boldsymbol{\Lambda},\boldsymbol{\Theta})$ の近傍で局所座標系をなす。このとき、すべての $(\boldsymbol{\Lambda},\boldsymbol{\Theta})$ において直交するような空間 $Diag(n)$と L $(\boldsymbol{\Lambda})$ を用いて次のように分解可能である。

$$PD(n)\simeq Diag(n)\otimes L(\boldsymbol{\Lambda}) \quad \cdots(16)$$

$$L(\boldsymbol{\Lambda})\equiv\left\{\boldsymbol{\Theta}^t\boldsymbol{\Lambda}\boldsymbol{\Theta}\,\middle|\,\boldsymbol{\Theta}\in O(n)\right\}\simeq O(n) \quad \cdots(17)$$



図－1　$PD(n)$ の空間構造の概念図

4. 構造化された分散共分散行列の空間構造

次に、(a)'～(b)'のモデル化によって実現される分散共分散行列が、$PD(n)$ のなかでどのように埋め込まれているかについて考察する。

対称行列$\boldsymbol{W}^t\boldsymbol{W}$ を次のように固有値分解する。

$$\boldsymbol{W}^t\boldsymbol{W}=\boldsymbol{\Theta}_w^t\boldsymbol{\Lambda}_w\boldsymbol{\Theta}_w \quad \cdots(18)$$

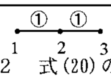
この時、(c)'によってモデル化される分散共分散行列は、次式のように表現可能である。

$$\boldsymbol{\Sigma}_c=\sigma_v\boldsymbol{\Theta}_w^t\boldsymbol{\Lambda}_w\boldsymbol{\Theta}_w+\sigma_\mu\boldsymbol{I}$$

$$=\boldsymbol{\Theta}_w^t(\sigma_v\boldsymbol{\Lambda}_w+\sigma_\mu\boldsymbol{I})\boldsymbol{\Theta}_w \quad \cdots(19)$$

このように、(c)'の構造化は分散共分散行列の集合 $PD(n)$ の中で、対角行列と直交行列の組 $(\boldsymbol{\Lambda},\boldsymbol{\Theta})$ のうち対角行列 $\boldsymbol{\Lambda}$ のみを、σ_v,σ_μ をパラメータとして変化させていることに他ならない。

一方、(a)'及び(b)'については、σ_μ をパラメータとして、対角行列 $\boldsymbol{\Lambda}$ だけでなく直交行列 $\boldsymbol{\Theta}$ も変化させる。これについて、次のような行列 $\boldsymbol{W}$ を用いて簡単な数値実験を行った（$\sigma_\mu=1$に固定）。



図－2　式(20)の行列に対応する地点の位置関係

$$\mathbf{W}=\begin{bmatrix}0 & 2/3 & 1/3\\ 1/2 & 0 & 1/2\\ 1/3 & 2/3 & 0\end{bmatrix} \quad \cdots(20)$$

結果について、次式の固有値分解の形で示す。

$$\boldsymbol{\Sigma}=\boldsymbol{\Theta}^t\boldsymbol{\Lambda}\boldsymbol{\Theta}$$

(i) $\lambda=0.10$（上段：(a)'　下段：(b)'）

$$\begin{bmatrix}1.02 & 0.12 & 0.08\\ & 1.19 & 0.12\\ Sym & & 1.02\end{bmatrix}=\begin{bmatrix}0.55 & 0.71 & 0.44\\ 0.63 & 0.00 & -0.78\\ 0.55 & -0.71 & 0.44\end{bmatrix}^t\begin{bmatrix}1.24 & & 0\\ & 0.94 & \\ 0 & & 0.88\end{bmatrix}\boldsymbol{\Theta}_{a1}$$

$$\begin{bmatrix}1.00 & 0.12 & 0.07\\ & 1.01 & 0.12\\ & & 1.00\end{bmatrix}=\begin{bmatrix}0.55 & 0.71 & 0.44\\ 0.63 & 0.00 & -0.78\\ 0.55 & -0.71 & 0.44\end{bmatrix}^t\begin{bmatrix}1.24 & & \\ & 0.93 & \\ & & 0.87\end{bmatrix}\boldsymbol{\Theta}_{b1}$$

(ii) $\lambda=0.50$（上段：(a)'　下段：(b)'）

$$\begin{bmatrix}1.74 & 1.10 & 1.01\\ & 1.84 & 1.12\\ & & 1.74\end{bmatrix}=\begin{bmatrix}0.56 & 0.71 & 0.43\\ 0.61 & 0.00 & -0.79\\ 0.56 & -0.71 & 0.46\end{bmatrix}^t\begin{bmatrix}4.04 & & \\ & 0.56 & \\ & & 0.73\end{bmatrix}\boldsymbol{\Theta}_{a2}$$

$$\begin{bmatrix}1.10 & 064 & 0.40\\ & 1.22 & 0.64\\ & & 1.10\end{bmatrix}=\begin{bmatrix}0.53 & 0.71 & 0.46\\ 0.65 & 0.00 & -0.76\\ 0.54 & -0.71 & 0.46\end{bmatrix}^t\begin{bmatrix}2.27 & & \\ & 0.69 & \\ & & 0.44\end{bmatrix}\boldsymbol{\Theta}_{b2}$$

本ケースでは、上に示したように対角行列の変化に比べて直交行列の変化が緩やかである。同様の傾向は、いろいろな3×3 の行列 $\boldsymbol{W}$ についても認められた。無論、$Diag(n)$と L $(\boldsymbol{\Lambda})$に導入されるべき自然な計量は、単純なユークリッド空間におけるそれとは異なるため [1]、これについてはより詳細な検討を続ける必要がある。

【参考文献】

(1) Ohara, A. and Kitamori, T. (1992) : Distance on Gramians and Their Applications, Recent Advances in Systems, Control, Networks and Signal Processing, Kimura, H. and Kodama, S. (eds.), Vol. I, pp.449-454, Mita Press.

(2) 前地直記 (1996)：対称行列空間の情報幾何とその応用, 東京大学大学院工学系研究科計数工学科専攻修士論文

(3) 横田一郎(1990)：古典型単純リー群, 現代数学社