

多変量モデルの頑健な事後分析

東京大学経済学部

入江 薫

一変量回帰モデルのパラメータの事後分布を外れ値から保護することを目的として、事後分布の頑健性が数学的に定義されている。過去 10 年間のうちに頑健性達成のための簡易な十分条件の研究が進み、位置尺度モデル (Desgagné, 2015)、回帰モデル (Gagnon et al., 2020)、階層モデル (Hamura et al., 2022)、ポアソン・負の二項回帰モデル (Hamura et al., 2025+) の頑健な事後分析の理論が明らかになっている。これらの研究はすべて、誤差分布を裾厚にするという古典的なロバスト統計のモデリングに基づくものだが、裾の厚さが t 分布よりも厚い（密度関数が対数項を含み、対数正規変動である）誤差分布でないと事後頑健性を達成できないという点が特徴である。

本研究ではこれまでの研究を多変量へと拡張し、多変量回帰モデルにおける事後頑健性の十分条件を与える。具体的には、 p 次元の観測値 $\mathbf{y}_i$ に対して、 $\mathbf{y}_i \sim N(\mathbf{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma})$ というモデルを考える。このクラスは共分散行列 $\boldsymbol{\Sigma}$ の推定に関心があるグラフィカルモデルを含む。いま、 $\mathbf{y}_i$ の成分（の一部）に外れ値の存在が疑われるときに、元のモデルをいかに変更すれば事後頑健性を達成できるかという問いに対して、我々は以下のモデルを提案する：

$$\mathbf{y}_i \sim N(\mathbf{X}_i\boldsymbol{\beta}, \mathbf{T}_i\boldsymbol{\Sigma}\mathbf{T}_i), \quad \mathbf{T}_i = \text{diag}(t_{i1}, \dots, t_{ip}).$$

ここで、 t_{ik} には独立に超裾厚な分布を仮定し、そのサポートは $(-\infty, -1) \cup (1, \infty)$ とする。各成分 y_{ik} の分散は $t_{ik}^2 \Sigma_{kk}$ であるから、外れ値に対して t_{ik}^2 が大きくなり誤差分散を増加させ、外れ値を誤差項で説明するという点是不変である。特徴的であるのは t_{ik} が負になりうるという点である。これにより、 y_{ik} と $y_{ik'}$ の相関は

$$\text{sgn}(t_{ik}t_{ik'}) \frac{\Sigma_{kk'}}{\Sigma_{kk}\Sigma_{k'k'}},$$

となり、外れ値が存在する（ $t_{ik} \neq 1$ または $t_{ik'} \neq 1$ の）時には相関の符号が変わり得ることが分かる。逆にいえば、その他のモデルでは外れ値の存在によって相関の絶対値が減少したり、あるいは外れ値が生じるときも通常と同じ相関を用いることになり不合理である。本講演では、このような共分散構造が事後頑健性の達成にとって本質的である点について議論する。また、類似のモデル、特に $\mathbf{T}_i$ に負値を許すモデルの研究はほとんど見当たらないことについて、参考文献を挙げつつ議論する。

本研究は羽村靖之氏（京都大学）および菅澤翔之助氏（慶應義塾大学）との共同研究である。

- Desgagné, A. (2015). “Robustness to outliers in location–scale parameter model using log-regularly varying distributions.” *The Annals of Statistics*, 43(4).
- Gagnon, P., Desgagné, A., & Bédard, M. (2020). “A new Bayesian approach to robustness against outliers in linear regression.” *Bayesian Analysis*, 15(2), 389-414.
- Hamura, Y., Irie, K., & Sugawara, S. (2022). “Log-regularly varying scale mixture of normals for robust regression.” *Computational Statistics & Data Analysis*, 173, 107517.
- Hamura, Y., Irie, K., & Sugawara, S. (2025+). “Robust hierarchical modeling of counts under zero-inflation and outliers: Theoretical robustness and efficient computation.” *Journal of the American Statistical Association*, published online.
- Hamura, Y., Irie, K., and Sugawara, S. (2025), “Outlier-robust Bayesian multivariate analysis with correlation-intact sandwich mixture,” arXiv:2508.18004.