

従属型一般化線形モデルにおけるモデル評価

九州大学 江口 翔一

Bayes アプローチにおける相対モデル記述評価においては、データの対数周辺尤度の漸近挙動に基づいて最適モデルを決定する。すなわち、対数周辺尤度の主要項を使いやすい形で近似することができれば記述的モデル評価規準として用いることができる。より一般の疑似対数尤度を対象として、漸近混合正規性やモデル誤特定の場合まで視野に入れたモデル設定で古典的な Bayes 情報量規準 BIC の拡張を行うことは理論上重要である。本発表では特に、線形回帰モデルを拡張したモデルである一般化線形モデルに関して、従属データの場合まで統一的に扱える正則条件を与える。本結果により、Lv and Liu (2014) の結果が大幅に拡張される。

パラメータ θ の事前分布を π 、疑似対数尤度を \mathbb{H}_n とする。適当な条件の下、対数周辺疑似尤度の確率展開が導出できる [2]:

$$\log \left(\int_{\Theta} \exp\{\mathbb{H}_n(\theta)\} \pi(\theta) d\theta \right) = \mathbb{H}_n(\hat{\theta}_n) - \frac{1}{2} \log \left| -\partial_{\theta}^2 \mathbb{H}_n(\hat{\theta}_n) \right| + O_p(1). \quad (1)$$

この発散項に着目することにより、Bayes モデル (π, \mathbb{H}_n) に対する疑似 Bayes 型情報量規準 (Quasi BIC; QBIC) を

$$\text{QBIC} := -2\mathbb{H}_n(\hat{\theta}_n) + \log \left| -\partial_{\theta}^2 \mathbb{H}_n(\hat{\theta}_n) \right|$$

という形で与えることができる。候補モデルごとに QBIC を計算し、それが最小となるモデルを最適なモデルとして選択する。

$(Y_j, X_j)_{j=1}^n$ を観測データ、 $b(\cdot)$ を既知の凸関数とすると、一般化線形モデルにおける疑似対数尤度を以下のように表すことができる:

$$\mathbb{H}_n(\theta) = \sum_{j=1}^n (Y_j X_j' \theta - b(X_j' \theta)).$$

この疑似対数尤度に対して (1) が成立するための条件を考察する。ここでは、 \mathbf{X}_n が従属データ、各 Y_j は X_j にのみ依存するものとし、 $Y_1|X_1, \dots, Y_n|X_n$ が独立である状況を扱う。特に、 \mathbf{X}_n が non-random で Y_1, \dots, Y_n が独立である場合その条件は簡単に与えることができる。このとき、適当な条件の下、QBIC の第二項は $\log \left| \sum_{j=1}^n \partial_{\theta} b(X_j' \hat{\theta}_n) X_j X_j' \right|$ となる。

参考文献

- [1] Lv, J. and Liu, J. S. (2014), Model selection principles in misspecified models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, **76**, 141–167.
- [2] Eguchi, S. and Masuda, H. (2015), Quasi-Bayesian comparison for LAQ models. MI Preprint Series.