Prenet ペナルティによる因子分析の単純構造の推定

九州大学　マス・フォア・インダストリ研究所　廣瀬慧

因子分析モデルは，観測される変数間の相関構造から，観測されない潜在因子を推定するモデルで，心理学，経済学，マーケティングなどに応用されている．いま，p 次元確率変数ベクトル $\boldsymbol{X} = (X_1, \ldots, X_p)^T$ が，平均ベクトル $\mathbf{0}$，分散共分散行列 $\boldsymbol{\Sigma}$ を持つ多変量正規分布に従うとする．因子分析モデルでは，共分散行列が $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}^T + \boldsymbol{\Psi}$ と表される．ただし，$\boldsymbol{\Lambda}$ は $p \times m$ 因子負荷行列とし，$\boldsymbol{\Psi} = \mathrm{diag}(\psi_1, \ldots, \psi_p)$ とする．$\boldsymbol{\Psi}$ の対角成分は独自分散と呼ばれる．

いま，多変量正規分布 $N(\mathbf{0}, \boldsymbol{\Lambda}\boldsymbol{\Lambda}^T + \boldsymbol{\Psi})$ に従う N 個の標本 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N$ が観測されたとする．$\boldsymbol{S}$ を標本分散共分散行列とする．ここで，パラメータを罰則付き損失関数 $\ell_\rho(\boldsymbol{\Lambda}, \boldsymbol{\Psi})$ の最小化により推定する．

$$\ell_\rho(\boldsymbol{\Lambda}, \boldsymbol{\Psi}) = \frac{1}{2}\left\{\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{S}) - \log|\boldsymbol{\Sigma}^{-1}\boldsymbol{S}|\right\} + \rho P(\boldsymbol{\Lambda}).$$

ただし，第一項目は負の対数尤度関数に対応し，第二項目はペナルティ項とする．$\rho > 0$ はペナルティの大きさを表す調整パラメータである．

これまで，ペナルティ項 $P(\boldsymbol{\Lambda})$ として，Lasso

$$P(\boldsymbol{\Lambda}) = \sum_{i=1}^{p}\sum_{j=1}^{m}|\lambda_{ij}|$$

などの，スパース推定ができる関数が用いられてきた．パラメータをいくつか 0 と推定する Lasso ペナルティは，推定された因子が何を意味するかを解釈しやすいというメリットがある．しかしながら，因子の解釈を行うことは，因子負荷量の要素間のコントラストを明確にさせることに対応し，パラメータを「正確に」0 と推定することは必ずしも本質的ではない．その意味で，Lasso ペナルティを使うよりも，因子の解釈を目的として古くから使われてきた因子回転に基づくペナルティを用いたほうが自然である．実は，罰則付き最尤法は，因子回転の一般化であることが示されている (Hirose et al., 2015) ため，原理的には，どのような因子回転基準も罰則付き推定法へと拡張することができる．

本発表では，次の *prenet* (*pr*oduct *e*lastic *net*) ペナルティを提案する．

$$P(\boldsymbol{\Lambda}) = \sum_{i=1}^{p}\sum_{j=1}^{m-1}\sum_{k>j}\left\{\gamma|\lambda_{ij}||\lambda_{ik}| + \frac{1}{2}(1-\gamma)\lambda_{ij}^2\lambda_{ik}^2\right\}.$$

ただし，$\gamma \in [0, 1]$ は調整パラメータとする．Prenet penalty の最大の特徴は，因子負荷行列の各行の要素のペアに対する**積**に基づいて構成されているということである．

Prenet ペナルティは，様々な興味深い性質を持つ．まず，$\gamma = 0$ の時は，因子回転で使われるコーティミン基準に対応する．そのため，Prenet ペナルティはコーティミン基準の一般化であることがわかる．さらに，$\gamma > 0$ のとき，Lasso と同様，パラメータの一部を正確に 0 と推定できる．また，$\rho \to \infty$ のとき，因子負荷行列の各行の要素のうち，少なくとも $m-1$ 個の成分が 0 と推定される．言い換えれば，因子負荷行列の各行の要素のうち，非ゼロ要素が 1 つ存在する可能性がある．これは，完全単純構造と呼ばれる，因子負荷行列の望ましい構造である．さらに，このとき，k 平均法の，変数に関するクラスタリングの一般化に対応している．実際，因子負荷行列の各列に対し，非ゼロ要素を持つ変数群が同じクラスターに属する．

References

Hirose, K. and Yamamoto, M. (2015). Sparse estimation via nonconcave penalized likelihood in factor analysis model. Statistics and Computing, 25(5), 863–875.