リーマン多様体上における多変量回帰モデルの低ランク推定

電気通信大学 大学院情報理工学研究科　吉川 剛平
電気通信大学 大学院情報理工学研究科　川野 秀一

1. はじめに

多変量低ランク回帰 (Izenman, 1975) とは，多変量回帰モデルの回帰係数に低ランク性を仮定した回帰モデルであり，その低ランク性によって推定すべきパラメータの個数は削減される．多変量低ランク回帰の拡張の一つに，特異値分解を用いた共スパース因子回帰 (co-sparse factor regression)(Mishra et al., 2017) が提案されている．このモデルには低ランク制約に加えて，直交制約とスパース制約が課されており，逐次的に潜在因子とパラメータの推定を行なっている．しかしながら，逐次推定アルゴリズムは推定する潜在因子の個数が多くなると解が不安定になるということが知られている．

本報告では，解の不安定性を克服するために，共スパース因子回帰を多様体最適化に基づいて推定する手法を提案する．多様体最適化 (Absil et al., 2008) を用いることで，すべての潜在因子間の関係を考慮に入れてパラメータの同時推定を実行することが可能である．これにより真のランクを適切に推定することができ，問題点を克服することができると期待される．数値実験により，提案手法の有効性を検証する．

2. リーマン多様体上における共スパース因子回帰

目的変数 $\mathbf{y}_i = [y_{i1}, \ldots, y_{iq}]^\mathsf{T} \in \mathbb{R}^q$ と説明変数 $\mathbf{x}_i = [x_{i1}, \ldots, x_{ip}]^\mathsf{T} \in \mathbb{R}^p$ に関して，n 組のデータ $\{(\mathbf{y}_i, \mathbf{x}_i), i = 1, \ldots, n\}$ が得られたとする．本報告では，次の最小化問題を考える．

$$\min_{\substack{\mathbf{U}\in\mathrm{StG}(r,p),\\ \mathbf{D}\in\mathbb{R}^{r\times r},\\ \mathbf{V}\in\mathrm{St}(r,q)}} \frac{1}{2}\left\|\mathbf{Y} - \mathbf{X}\mathbf{U}\mathbf{D}\mathbf{V}^\mathsf{T}\right\|_F^2 + n\lambda_1 \sum_{i=1}^{p}\sum_{j=1}^{r} w_{ij}^{(u)}|u_{ij}| + n\alpha\lambda_2 \sum_{i=1}^{q}\sum_{j=1}^{r} w_{ij}^{(v)}|v_{ij}| + n\sqrt{q}(1-\alpha)\lambda_2 \sum_{j=1}^{r} w_j^{(d)}\mathbb{1}\left(\mathbf{v}_j \neq \mathbf{0}\right).$$

ここで，$\mathbf{U} \in \mathbb{R}^{p\times r}$, $\mathbf{V} \in \mathbb{R}^{q\times r}$, $\mathbf{D} = \mathrm{diag}\{d_1, \ldots, d_r\}$ で，回帰係数 $\mathbf{C} = \mathbf{U}\mathbf{D}\mathbf{V}^\mathsf{T}$ とする．回帰係数には，低ランク制約 $\mathrm{rank}(\mathbf{C}) \leq r = \min(r_x, q)$，$r_x = \mathrm{rank}(\mathbf{X})$ が課されている．また，$\lambda_1, \lambda_2 > 0$ は正則化パラメータで，第 2 項目と第 3 項目は要素毎にスパース性を誘引する正則化項である．第 4 項目の $\mathbb{1}(\cdot)$ は，指示関数であり，低ランク性を誘引する硬式作用素型の正則化項である．ここで，$w_{ij}^{(u)}, w_{ij}^{(v)}, w_{ij}^{(d)}$ は正の重みである．α は 0 から 1 の間に値をとるチューニングパラメータであり，α によって $\mathbf{V}$ に作用するスパース性と低ランク性のバランスを調整する．$\mathbf{U}, \mathbf{V}$ の実行可能領域は，次に示すシュティーフェル多様体の構造を持つ．

$$\mathrm{StG}(r,p) = \left\{\mathbf{U} \in \mathbb{R}^{p\times r} \mid \mathbf{U}^\mathsf{T}\mathbf{G}\mathbf{U} = \mathbf{I}_r,\ \mathbf{G} = \mathbf{X}^\mathsf{T}\mathbf{X}/n\right\},$$
$$\mathrm{St}(r,q) = \left\{\mathbf{V} \in \mathbb{R}^{q\times r} \mid \mathbf{V}^\mathsf{T}\mathbf{V} = \mathbf{I}_r\right\}.$$

このように，推定するパラメータの実行可能領域がリーマン多様体を形成しているとき，リーマン多様体上で直接パラメータを推定することが可能となる．推定アルゴリズムの詳細および数値実験の結果は当日紹介する．

Absil, P.-A., Mahony, R., and Sepulchre, R. (2008) *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ.
Izenman, A. J. (1975) Reduced-rank regression for the multivariate linear model. *J. Multivar. Anal.*, **5**(2), 248–264.
Mishra, A., Dey, D. K., and Chen, K. (2017) Sequential co-sparse factor regression. *J. Comput. Graph. Stat.*, **26**(4), 814–825.