第 51 卷第 3 期　　中国科学技术大学学报　　Vol. 51, No. 3
2021 年 3 月　　JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA　　Mar. 2021
Received: 2021-03-08; Revised: 2021-03-28　　doi:10.52396/JUST-2021-0053

Subgroup analysis for multi-response regression

WU Jie[1,2], ZHOU Jia[1,2*], ZHENG Zemin[1,2*]
1. International Institute of Finance, University of Science and Technology of China, Hefei 230601, China;
2. School of Management, University of Science and Technology of China, Hefei 230026, China
* Corresponding author. E-mail: tszhjia@mail.ustc.edu.cn; zhengzm@ustc.edu.cn

Abstract: Correctly identifying the subgroups in a heterogeneous population has gained increasing popularity in modern big data applications since studying the heterogeneous effect can eliminate the impact of individual differences and make the estimation results more accurate. Despite the fast growing literature, most existing methods mainly focus on the heterogeneous univariate regression and how to precisely identify subgroups in face of multiple responses remains unclear. Here, we develop a new methodology for heterogeneous multi-response regression via a concave pairwise fusion approach, which estimates the coefficient matrix and identifies the subgroup structure jointly. Besides, we provide theoretical guarantees for the proposed methodology by establishing the estimation consistency. Our numerical studies demonstrate the effectiveness of the proposed method.

Keywords: multi-response regression; subgroup analysis; concave penalties; ADMM algorithm
CLC number: O29　　**Document code**: A
2010 Mathematics Subject Classification: Primary 62H20; Secondary 62J07

1 Introduction

Rapid developments in technology have brought in massive data sets in various fields ranging from health, genomics and molecular biology, among others. In these applications, identifying subgroups from a heterogeneous population is of great importance since it reveals domain knowledge behind the data. For example, experimental studies have shown that the relative efficacy of antiretroviral drugs for treating human immunodefciency virus infection sometimes depends on baseline viral load and CD4 count[1]. Similarly, in the study of clinical trials, it has been increasingly recognized that some subgroup of patients can benefit from a treatment or suffer from an adverse effect much more than the others[2]. Therefore, recovering the heterogeneous effects has gained increasing popularity in data analyses.

A popular method for analyzing data from a heterogeneous population is to view data as coming from a mixture of subgroups, where their own sets of parameter values should be estimated using finite mixture model analysis[2-4]. However, these mixture model-based approaches need to specify the underlying distribution of data and the number of mixed components in the population, which are often difficult to satisfy in real applications. By contrast, several other methods consider the problem of exploring homogeneous effects of the covariates by assuming that the true coefficients are divided into a few clusters with common values. For example, Guo et al.[5] proposed using a pairwise L_1 fusion penalty for identifying variables in the context of Gaussian model-based cluster analysis. Chi and Lange[6] proposed to multiply nonnegative weights to the L_1 norms to reduce the bias. The major advantage of such methods is that they can detect and identify heterogeneous subgroups without knowledge of a priori classification. But the L_1 penalty generates large biases in the estimates in each iteration of the algorithm, thus leading to incorrect conclusions. Furthermore, Wang et al.[7] applied a two-stage multiple change point detection method to determine the subgroups and estimated the regression parameters. Similarly, Li et al.[8] proposed an estimation procedure combining the likelihood method and the change point detection with the binary segmentation algorithm. Despite a well estimation accuracy of these two methods, it is unclear how to verify the theoretical properties.

To address these issues, Ma and Huang[9] proposes a new method in which the heterogeneity can be modeled through subject-specifc intercepts in regression and can be implemented via a concave pairwise fusion

Citation: WU Jie, ZHOU Jia, ZHENG Zemin. Subgroup analysis for multi-response regression. J. Univ. Sci. Tech. China, 2021, 51(3): 216-227.

penalized least squares without a priori knowledge of classification, thus more desirable for identifying subgroups. But this method is only applicable to heterogeneous univariate regression, and how to deal with heterogeneous multi-response regression is still unknown, which restricts the efficiency of the method. Although there exist many literature focusing on multi-response regression[10-15], which represent the dependency between the multiple outcomes and the same set of predictors, these method are no longer valid in face of the problem related to the data from a heterogeneous population.

In this article, we develop a new methodology for heterogeneous multi-response regression, which estimates the coefficient matrix and identifies the subgroup structure jointly. The proposed estimator utilizes the idea proposed By Ma and Huang[9], that is, applying the concave pairwise fusion penalized least squares, but extends the univariate heterogeneous regression to a multiple one. The major contributions of this paper are twofold. First, we develop a new method for heterogeneous multi-response regression, which automatically divides the observations into subgroups without a prior knowledge of classification. Second, we provide theoretical guarantees for the proposed method by establishing estimation consistency and derive the convergence properties of the algorithm.

The remainder of this paper is organized as follows. Section 2 presents the model setting and our new methodology. Theoretical properties of the proposed method are established in Section 3. We provide numerical studies in Section 4. Section 5 concludes with extensions and possible future work. The proofs and additional technical details are provided in the Appendix.

2 Subgroup analysis for multi-response regression via concave pairwise fusion

2.1 Model setting

Consider the following heterogeneous multi-response regression model

$$Y = XB + C + E \tag{1}$$

where $Y=(y_1,\cdots,y_n)^{\mathrm{T}}$ is an $n\times q$ response matrix, $X=(x_1,\cdots,x_n)^{\mathrm{T}}$ is an $n\times p$ covariate matrix, B is a $p\times q$ coefficient matrix, $C=(c_1,\cdots,c_n)^{\mathrm{T}}$ is an $n\times q$ intercept matrix and $E=(e_1,\cdots,e_n)^{\mathrm{T}}$ is an $n\times q$ random error matrix. One interesting application of the model is precision medicine where the responses vector could be several phenotypes associated with some disease and predictors are a set of observed characters such as gender, sex, age and so on. After adjusting for the effects of the covariates, the distribution of the response is still heterogeneous. It means that the heterogeneity can be the result of unobserved latent factors which can be modeled through the subject-specifc intercept vectors, similar as that in Ref. [9]. Specifically, the heterogeneous structure can be modeled as follows

$$C=\begin{bmatrix} c_1^{\mathrm{T}} \\ c_2^{\mathrm{T}} \\ \cdots \\ c_n^{\mathrm{T}} \end{bmatrix}=\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1K} \\ w_{21} & w_{22} & \cdots & w_{2K} \\ \vdots & & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nK} \end{bmatrix}\begin{bmatrix} \alpha_1^{\mathrm{T}} \\ \alpha_2^{\mathrm{T}} \\ \cdots \\ \alpha_K^{\mathrm{T}} \end{bmatrix}=W\alpha,$$

where $\alpha=(\alpha_1,\cdots,\alpha_K)^{\mathrm{T}}$ is a $K\times q$ matrix and α_k is the common vector for the c_i's in the kth subgroup, and W is an $n\times K$ (latent) indicator matrix in which $w_{ik}=1$ for c_i in kth group and $w_{ik}=0$ otherwise. Substituting this into model (1) yields the final multi-response heterogeneous model:

$$Y = XB + W\alpha + E \tag{2}$$

Let $\psi=(\psi_1,\cdots,\psi_K)$ be a partition of $\{1,\cdots,n\}$. Without loss of generality, we assume that $Y=(y_1,\cdots,y_n)^{\mathrm{T}}$ are from K different groups with $K\geqslant 1$ and the data from the same group have the same intercept vector α_i, i. e. $c_i=\alpha_k$ for all $i\in\psi_k$.

Although we adopt a similar estimation idea proposed in Ref. [9], how to apply the method under multi-response regression model here is a nontrivial task even if it is conceptually straightforward. Our goal is to correctly estimate K, identify the subgroups of outcomes and accurately estimate the regression coefficient matrix B.

2.2 Algorithm

Consider that the heterogeneous treatment effects are characterized by the intercept matrix, to estimate the unknown parameter matrix B and intercept matrix C jointly, a direct idea is to solve the following penalized least square problem

$$Q_n(C,B)=\frac{1}{2}\sum_{i=1}^{n}\|y_i-x_iB-c_i\|_2^2+\sum_{i<j}p_r(\|c_i-c_j\|_2,\lambda) \tag{3}$$

where $p_r(\cdot,\lambda)$ is a penalty function indexed by the tuning parameter $\lambda\geqslant 0$, indicating the amount of penalization. By minimizing the objective function, the penalty function $p_\lambda(.)$ shrinks some $\|c_i-c_j\|_2$ to 0, thus automatically dividing the samples into different subgroups. Given $\lambda>0$, the estimates of the coefficient matrices are defined as

$$(\widehat{C}(\lambda),\widehat{B}(\lambda))=\arg\min_{C,B}Q_n(C,B,\lambda) \tag{4}$$

For ease of presentation, we set $(\widehat{C},\widehat{B})\equiv(\widehat{C}(\lambda),\widehat{B}(\lambda))$. Let $\{\widehat{\alpha}_1,\cdots,\widehat{\alpha}_{\widehat{K}}\}$ be the distinct values of $\widehat{c}_i$ and $\widehat{\mathscr{A}}_k=\{i:\widehat{c}_i=\widehat{c}_k,1\leqslant i\leqslant n\}$, $1\leqslant k\leqslant\widehat{K}$. Then an estimated partition of $\{1,\cdots,n\}$ can be $\{\widehat{\mathscr{A}}_1,\cdots,\widehat{\mathscr{A}}_{\widehat{K}}\}$.

However, the penalty function in the above optimization problem (3) is not separable in c_i. We then reformulate the problem as follows by resorting to $\delta_{ij}=c_i-c_j$,

$$L_0\{C, B, \delta\} = \frac{1}{2}\sum_{i=1}^{n} \| y_i - x_i B - c_i \|_2^2 + \sum_{i<j} p_r(\|\delta_{ij}\|_2, \lambda),$$

$$\text{s.t. } c_i - c_j - \delta_{ij} = 0 \tag{5}$$

where $\delta = \{\delta_{ij}, i<j\}^{\mathrm{T}}$. Applying the augmented Lagrangian method, we can obtain the estimates of the parameter matrix by minimizing the following problem:

$$L(C, B, \delta, v) = L_0\{C, B, \delta\} + \sum_{i<j}\langle v_{ij}, (c_i - c_j - \delta_{ij})\rangle + \frac{\upsilon}{2}\| c_i - c_j - \delta_{ij}\|_2^2 \tag{6}$$

where $\langle a, b\rangle = a^{\mathrm{T}}b$ is the inner product of two matrices a and b with the same dimensions, the dual matrix $v=\{v_{ij}^{\mathrm{T}}, i<j\}^{\mathrm{T}}$ is Lagrange multiplier matrix and υ is a penalty parameter. Then, with observed data sets (X, Y), we use an alternating direction method of multipliers (ADMM) to compute the estimates of (C, B, δ, v). At the mth iteration, we solve for the minimizer of $L(C, B, \delta, v)$ iteratively using the following three steps:

$$(C^{m+1}, B^{m+1}) = \arg\min_{C,B} L(C, B, \delta^m, v^m) \tag{7}$$

$$\delta^{m+1} = \arg\min_{\delta} L(C^{m+1}, B^{m+1}, \delta, v^m) \tag{8}$$

$$v_{ij}^{m+1} = v_{ij}^m + \upsilon(c_i^{m+1} - c_j^{m+1} - \delta_{ij}^{m+1}) \tag{9}$$

Since δ^m and v^m are fixed in the first step, the problem (7) can be simplified as

$$f(C, B) = \frac{1}{2}\sum_{i=1}^{n}\| y_i - x_i B - C \|_2^2 + \sum_{i<j}\langle v_{ij}, (c_i - c_j - \delta_{ij})\rangle + \frac{\upsilon}{2}\sum_{i<j}\| c_i - c_j - \delta_{ij}\|_2^2 + C_n,$$

where C_n is a constant independent of (C,B). Through some algebra, we rewrite $f(C,B)$ as

$$f(C, B) = \frac{1}{2}\| Y - XB - C\|_F^2 + \langle V, (AC - \delta)\rangle + \frac{\upsilon}{2}\| AC - \delta\|_F^2 + C \tag{10}$$

where $A=\{(e_i-e_j), i<j\}^{\mathrm{T}}$ with e_i being the ith unit $n\times 1$ vector whose ith element is 1 and the remaining ones are 0. Thus the updates C^{m+1} and Z^{m+1} are given by

$$C^{m+1} = (I_n + \upsilon A^{\mathrm{T}}A)^{-1}[(I_n - Q_X)Y + \upsilon A^{\mathrm{T}}(\delta^m - \frac{1}{\upsilon}V^m)] \tag{11}$$

$$B^{m+1} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}(Y - C^{m+1}) \tag{12}$$

where $Q_X = X(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}$ denotes the orthogonal projection matrix onto the range of X and I_n denotes the identity matrix. In the second step, we discard the terms independent of δ and minimize the problem (8) as follows

$$\sum_{i<j}\langle v_{ij}^{\mathrm{T}}, (c_i - c_j - \delta_{ij})\rangle + \frac{\upsilon}{2}\| c_i - c_j - \delta_{ij}\|_2^2 + p_r(\|\delta_{ij}\|_2, \lambda),$$

with respect to δ. Corresponding to the penalty p_r, this is a groupwise thresholding operator. Let $S(z, t) = (1-t/\|z\|_2)_+ z$ be the groupwise soft thresholding operator with $(x)_+=x$ if $x>0$ and $=0$, otherwise. And define $\zeta_{ij}^m = (c_i^{m+1} - c_j^{m+1}) + \upsilon^{-1}v_{ij}^m$. Then the solution path for the matrix δ is

$$\delta_{ij}^{m+1} = S(\zeta^m, \lambda/\upsilon) \tag{13}$$

for the L_1 penalty[16] or

$$\delta_{ij}^{m+1} = \begin{cases} \dfrac{S(\zeta_{ij}^m, \lambda/\upsilon)}{1-(\gamma\upsilon)} \text{ if } \|\zeta_{ij}^m\|_F \leqslant \gamma\lambda, \\ \zeta_{ij}^m, \text{ if } \|\zeta_{ij}^m\|_F > \gamma\lambda \end{cases} \tag{14}$$

for the MCP[17] with $\gamma>1/\upsilon$ or

$$\delta_{ij}^{m+1} = \begin{cases} S(\zeta_{ij}^m, \lambda/\upsilon) \text{ if } \|\zeta_{ij}^m\|_F \leqslant \lambda + \lambda/\upsilon, \\ \dfrac{S(\zeta_{ij}^m, \gamma\lambda/((\gamma-1)\upsilon))}{1-1/((\gamma-1)\upsilon))} \\ \quad \text{if } \lambda + \lambda/\upsilon \leqslant \|\zeta_{ij}^m\|_F \leqslant \gamma\lambda, \\ \zeta_{ij}^m \text{ if } \|\zeta_{ij}^m\|_F > \gamma\lambda \end{cases} \tag{15}$$

for the SCAD penalty[18] with $\gamma>1/\upsilon+1$. At last, the update of v_{ij} is given in (9). We terminate the algorithm when the stopping criterion met. To be specific, we stop the algorithm once $r^{m+1}=AC^{m+1}-\delta^{m+1}$ is close to zero such that $\|r^{m+1}\|_F<a$ for some small value a.

The efficiency of subgroups identification via our method can be seen from a simple simulation example summarized in Figure 1, where the data are generated similarly as in Section 4 except that each row of C^* is generated i. i. d. from three different vectors α_1, α_2, α_3 with equal probabilities. That is,

$$P(c_i^* = \alpha_1) = P(c_i^* = \alpha_2) = P(c_i^* = \alpha_3) = 1/3$$

with $\alpha_1=(2,\cdots,2)$, $\alpha_2=(0,\cdots,0)$ and $\alpha_3=(-2,\cdots,-2)$. It is clear that the estimators using concave penalties MCP and SCAD can accurately identify the subgroups while convex penalty Lasso merge to one value quickly due to the overshrinkage of the L_1 penalty.

3 Theoretical properties

In this section, we study the theoretical properties of the proposed estimator. Specifically, we provide sufficient conditions under which there exists a local minimizer of the objective function equal to the oracle least squares estimator with a priori knowledge of the true groups with high probability. We also derive the lower bound of the minimum difference of the coefficients between subgroups to estimate the subgroup-specific treatment effects.

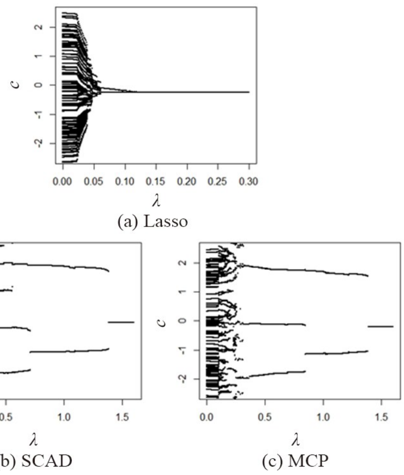

Figure 1. Solution paths for $\widehat{c}_{11}, \widehat{c}_{21}, \cdots, \widehat{c}_{n1}$ against λ for data in Section 2.

3.1 Notation and conditions

Let $\mathscr{M}_{\psi} = \{ C \in \mathbb{R}^{n\times q} \}$: $c_i = c_j$, for $i, j \in \psi_k$, $1 \leqslant k \leqslant K$. For each $C \in \mathscr{M}_{\psi}$, It can be written as $C = W\alpha$, where $\alpha = (\alpha_1, \cdots, \alpha_K)^{\mathrm{T}}$ and α_k is a q-dimension vector of the kth subgroup for $k = 1, \cdots, K$. Simple calculation shows

$$W^{\mathrm{T}} W = \operatorname{diag}(|\psi_1|, \cdots, |\psi_K|),$$

where $|\psi_k|$ denotes the number of elements in ψ_k. Denote the minimum and maximum group sizes by $|\psi_{\min}| = \min_{1\leqslant k\leqslant K} |\psi_k|$ and $|\psi_{\max}| = \max_{1\leqslant k\leqslant K} |\psi_k|$, respectively. For any positive numbers a_n and b_n, let $a_n \gg b_n$ denote $a_n^{-1} b_n = o(1)$. For any vector $\zeta = (\zeta_1, \cdots, \zeta_s)^{\mathrm{T}} \in \mathbb{R}^s$, define the infinity norm of the vector as $\|\zeta\|_\infty = \max_{1\leqslant l\leqslant s} |\zeta_l|$. For any matrix $A = (A_{ij})_{i=1,j=1}^{s,t}$, define the infinity norm of the matrix as $\|A\|_\infty = \max_{1\leqslant l\leqslant s} \sum_{j=1}^{t} |A_{ij}|$. For any symmetric matrix $A_{s\times s}$, define its L_2 norm by $\|A\|_2 = \max_{\zeta\in R^s, \|\zeta\|_2=1} \|A\zeta\|_2$, and let $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ be the smallest and largest eigenvalues of A, respectively. Denote $U = (W, X)$, and the notation $\mathbf{1}(\cdot)$ denotes an indicator function. Finally, denote the scaled penalty function by

$$\rho(t) = \lambda^{-1} P_\gamma(t, \lambda).$$

We make the following basic assumptions.

Condition 3.1 For some constant $a > 0$, the function $\rho(t)$ is constant once $t \geqslant a\lambda$. It is symmetric, non-decreasing and concave on $[0, \infty)$. In addition, $\rho'(t)$ exists and is continuous except for a finite number values of t, $\rho'(0+) = 1$ and $\rho(0) = 0$.

Condition 3.2 The noise vector $\epsilon = (\epsilon_1, \cdots, \epsilon_n)^{\mathrm{T}}$ has sub-Gaussian tails such that $P(|a^{\mathrm{T}}\epsilon| > \|a\|_2 x) \leqslant 2\exp(-c_1 x^2)$ for any vector $a \in \mathbb{R}^n$ and $x > 0$, where $0 < c_1 < \infty$.

Condition 3.3 For any $1 \leqslant l \leqslant q$, assume $\sum_{i=1}^{n} d_{il}^2 = n$, and for any $1 \leqslant j \leqslant p$, assume $\sum_{i=1}^{n} x_{ij}^2 \mathbf{1}\{i \in \psi_k\} = |\psi_k|$. $\lambda_{\min}(U^{\mathrm{T}} U) \geqslant C_1 |\psi_{\min}|$, $\sup_i \|X_i\|_2 = C_2\sqrt{p}$ for some constants $0 < C_1 < \infty$ and $0 < C_2 < \infty$.

Conditions 3.1 puts a mild assumption on the penalties and it is obviously that the concave penalties such as MCP and SCAD satisfy it. Condition 3.2 is standard in the penalized regression in high-dimensional settings. Condition 3.3 imposes conditions on the eigenvalues of the population covariance matrix $U^{\mathrm{T}} U$. Note that $\lambda_{\min}(W^{\mathrm{T}} W) = |\psi_{\min}|$. By assuming $\lambda_{\min}(X^{\mathrm{T}} X) = Cn$, we have

$$\lambda_{\min} | (W, X)^{\mathrm{T}} (W, X) | \leqslant \min(|\psi_{\min}|, Cn).$$

The equality holds when $W^{\mathrm{T}} X = 0$. Hence, we assume $\lambda_{\min}[(W, X)^{\mathrm{T}} (W, X)] \geqslant C_1 |\psi_{\min}|$ for some constant $0 < C_1 < 1$.

Before giving the detail theorem, we also need to introduce a new definition called oracle estimators, which are not real estimators but theoretical constructions useful for stating the properties of the proposed estimator. When the true group memberships $\psi_1, \cdots, \psi_K$ are known, the oracle estimators for B and C are solved by

$$(\widehat{B}^{\mathrm{or}}, \widehat{C}^{\mathrm{or}}) = \arg\min_{C\in\mathbb{R}^{n\times q},\, B\in\mathbb{R}^{p\times q}} \frac{1}{2} \|Y - XB - C\|_F^2,$$

and correspondingly, the oracle estimators for the common intercept vector α and the coefficient matrix B are given by

$$(\widehat{B}^{\mathrm{or}}, \widehat{\alpha}^{\mathrm{or}}) = \arg\min_{\alpha\in\mathbb{R}^{K\times q},\, B\in\mathbb{R}^{p\times q}} \frac{1}{2} \|Y - XB - W\alpha\|_F^2 \tag{16}$$

Let $\alpha^0 = (\alpha_1, \cdots, \alpha_K)^{\mathrm{T}} \in \mathbb{R}^{K\times q}$, where α_k^0 is the underlying common intercept vector for the group ψ_k. Let B^0 be the underlying regression coefficient matrix. Now we are ready to show the main results.

Theorem 3.1 Suppose $|\psi_{\min}| \gg q\sqrt{(K+p)\, n\log n}$. Then under Conditions 3.1 − 3.3, we have with probability at least $1 - 2q(K+p)\, n^{-1}$,

$$\|((\widehat{\alpha}^{\mathrm{or}} - \alpha^0)^{\mathrm{T}}, (\widehat{B}^{\mathrm{or}} - B^0)^{\mathrm{T}},)^{\mathrm{T}}\|_F \leqslant \phi_n \tag{17}$$

and

$$\|\widehat{C}^{\mathrm{or}} - C^0\|_2 \leqslant \sqrt{|\psi_{\max}|}\, \phi_n, \ \sup_i \|\widehat{c}_i^{\mathrm{or}} - c_i^0\|_F \leqslant \phi_n,$$

where

$$\phi_n = C C_1^{-1} |\psi_{\min}|^{-1} q \sqrt{K+p}\sqrt{n\log n} \tag{18}$$

Theorem 3.1 establishes the estimated error bound for the oracle estimator $(\widehat{\alpha}^{\mathrm{or}}, \widehat{B}^{\mathrm{or}})$ obtained by solving least squares problem (1). Considering that $K \geqslant n|\psi_{\min}|$ and $|\psi_{\min}| \gg q\sqrt{(K+p)n\log n}$, K, p and q should satisfy $q\sqrt{K+p}\sqrt{n\log n}=o(n)$. Therefore, the upper bound ϕ_n of the estimation error can converge to 0 for sufficient large n.

Remark 3.1 For $K\leqslant 2$, let

$$b_n=\min_{i\in\psi_k,\, j\in\psi_{k'},\, k\neq k'}\| c_i^0-c_j^0\|_2=\min_{k\neq k'}\|\alpha_k^0-\alpha_{k'}{}^0\|_2$$

be the minimal difference of the common values between two groups.

Theorem 3.2 Suppose the conditions in Theorem 3.1 hold. If $b_n>a\lambda$ and $\lambda\gg\phi_n$, for some constant $a>0$, where ϕ_n is given in (18), then there exists a local minimizer $(\widehat{C}(\lambda),\widehat{B}(\lambda))$ of the objective function $Q_n(C,B;\lambda)$ given in (4) satisfying

$$P((\widehat{C}(\lambda),\widehat{B}(\lambda))=((\widehat{C}^{\rm or},\widehat{B}^{\rm or})))\to 1.$$

Theorem 3.2 describes that the oracle estimator $(\widehat{C}^{\rm or},\widehat{B}^{\rm or})$ is a local minimizer of function (4). Combined with Theorem 3.1, we conclude that the estimation error bound of $(\widehat{C},\widehat{B})$ solved by (4) can converge to 0 with sufficient large n.

Proposition 3.1 Let $r^m=AC^m-\delta^m$ and $s^m=vA^{\rm T}(\delta^m-\delta^{m+1})$ be the primal residual and the dual residual in the ADMM described above, respectively. It holds that $\lim_{m\to\infty}\|r^m\|_F^2=0$ and $\lim_{m\to\infty}\|s^m\|_F^2=0$ for the MCP and SCAD penalties.

Therefore, the proposed Algorithm always achieves a local minimum of Q_n, starting from some reasonable initial values. We suggest to give the initial matrices by solving $L_R(C,B)=\frac{1}{2}\|Y-XB-C\|_F^2$ and the solutions are $B_R=(X^{\rm T}X)^{-1}X^{\rm T}Y$ and $C_R=Y-XB_R$.

4 Simulation studies

In this section, we use simulated data to investigate the finite sample performance of the proposed method via two concave penalties, the smoothly clipped absolute deviation (SCAD), the minimax concave penalty (MCP) and one convex penalty (Lasso), in comparison with the classic method reduced rank regression (RRR). Among them, MCP, SCAD and Lasso have two tuning parameters including rank and sparse parameter, which are selected jointly by BIC. By contrast, RRR only has one tuning parameter rank, which is also tuned by BIC for fair comparison of all methods.

We adopted a similar simulation setting as that in Ref. [9] by extending the univariate response variable to muli-response variables and generate 100 data sets from model (1) with $(n,p,K)=(100,5,2)$ and $q=3$, 6. For each data set, the rows of the design matrix X are independently and identically distributed (i. i. d.) generated from $N(0,\Sigma_x)$, where Σ_x is with diagonal elements 1 and off-diagonal elements 0.3, thus bringing in predictor correlation. Similarly, each row of the error matrix E is drawn i. i. d. from $N(0,0.5^2\Sigma_E)$, where Σ_E has the same compound symmetry structure as Σ_x. All entries in the coefficient matrix B^* are generated i. i. d. from independent uniform [0.5, 1]. Denote the true intercept matrix by $C^*=(c_1^*,\cdots,c_n^*)^{\rm T}$. Then each row of C^* is generated i. i. d. from two different vectors α_1, α_2 with equal probabilities. That is, we generate them from the distribution: $P(c_i^*=\alpha_1)=P(c_i^*=\alpha_2)=1/2$ with $\alpha_1=(2,\cdots,2)$ and $\alpha_2=(0,\cdots,0)$.

To compare the aforementioned methods, we calculate ${\rm Est}(\widehat{B})=\|B^*-\widehat{B}\|_F^2/(pq)$ and ${\rm Est}(\widehat{C})=\|C^*-\widehat{C}\|_F^2/(nq)$ for estimation and ${\rm Err}(\widehat{B},\widehat{C})=\|XB^*+C^*-X\widehat{B}-\widehat{C}\|_F^2/(nq)$ for prediction. Meanwhile, to evaluate the subgroup identification performance, we report the mean squared errors for K different subgroups, which are defined as ${\rm Est}(\widehat{\alpha}_1)=\|\alpha_1^*-\widehat{\alpha}_1\|_2^2/q$ and ${\rm Est}(\widehat{\alpha}_2)=\|\alpha_2^*-\widehat{\alpha}_2\|_2^2/q$. At last, the average of the estimated intercept vectors from all repetitions $(\widehat{K})$ and the percentage of correct identification (K-per) are also summarized in Table 1.

Table 1. The sample mean and standard deviation (S. D.) of estimators (1×10^2). $n=100$, $p=12$, $K=2$.

	Method	Est($\widehat{B}$)	Est($\widehat{C}$)	Err($\widehat{B},\widehat{C}$)	Est($\widehat{\alpha}_1$)	Est($\widehat{\alpha}_2$)	$\widehat{K}$	K-per (%)
$q=3$	RRR	0.647 (0.111)	180.000 (0.000)	177.747 (0.102)	–	–	–	–
	MCP	0.090 (0.031)	0.921 (0.256)	1.218 (0.280)	0.816 (0.292)	1.045 (0.418)	2.000 (0)	100.000
	SCAD	0.120 (0.100)	0.889 (1.446)	1.053 (1.172)	0.218 (0.142)	0.402 (0.280)	2.000(0)	100.000
	Lasso	0.381 (0.076)	99.901 (0.051)	99.115 (0.121)	–	--	1.000 (0)	0.000
$q=6$	RRR	1.296 (0.101)	176.000 (0.000)	171.307 (0.074)	–	–	–	–
	MCP	0.086 (0.022)	0.719 (0.225)	1.011 (0.236)	0.737 (0.250)	0.702 (0.287)	2.000 (0)	100.000
	SCAD	0.080 (0.018)	0.170 (0.069)	0.487 (0.091)	0.180 (0.108)	0.160 (0.093)	2.000 (0)	100.000
	Lasso	2.698 (0.234)	100.061 (0.052)	89.666 (0.120)	–	–	1.000 (0)	0.000

In view of the results in Table 1, it is clear that our proposed estimators using concave penalties MCP and SCAD enjoy higher accuracy than utilizing convex penalty Lasso or classic method RRR which does not consider heterogeneous effects. Although similar to MCP and SCAD, Lasso consider a heterogeneous effects, but it tends to over-shrink large coefficients, thus leads to biased estimates and unable to correctly recover subgroups. In view of the heterogeneous estimation results in Table 1, both MCP and SCAD can identify the subgroups while RRR and Lasso have worse performance than the MCP and SCAD penalties. Specifically, the heterogeneity related measures do not apply to Lasso and RRR since they do not recover the latent subgroup structures. From another point of view, Lasso and RRR can not recover and utilize the heterogeneous structure, which in turn lowers its estimation and prediction accuracies about coefficient matrices B^* and C^*.

Figure 2 displays the fusiongram for $\widehat{c}_{11}, \widehat{c}_{21}, \cdots, \widehat{c}_{n1}$, the first element in $\widehat{c}_i$, $i=1,\cdots,n$, against different sparse parameter λ for the data. In view of the results, it is clear that MCP and SCAD have similar solution paths as shown in Figure 1 and their estimated values of $\widehat{c}_{i1}$ converge to two different values 2 and 0, which equal to the true heterogeneous intercept values. By contrast, the L_1 penalty shrinks the value quickly and converges to one value once λ exceed a certain constant.

5 Application to the yeast cell cycle data set

In this section, we will analyze the yeast cell cycle data set originally studied in Ref. [19]. This data set consists of 524 yeast genes, the RNA transcript levels (X) of which can be regulated by transition factors (TF) within the eukaryotic cell cycle. Specifically, 21 of the TFs were experimentally confirmed related to cell cycle process. It covers approximately two cell cycle periods with measurements at 7 min intervals for 119 min with a total of 18 time points (Y). Similar to Ref. [13], we considered the multi-response regression model to estimate the association coefficient matrix between the transition factors and the cell cycle gene expression data. Then we artificially added the intercept matrix C similarly as in Section 4 to generate subgroup structure, where $C=(c_1,\cdots,c_n)^{\mathrm{T}}$. Each row of C is generated i. i. d. from from the distribution:

$$P(c_i=\alpha_1)=P(c_i=\alpha_2)=1/2$$

with $\alpha_1=(2,\cdots,2)$ and $\alpha_2=(0,\cdots,0)$. Based on the processed data set, we identify the subgroups via Lasso, SCAD and MCP. Since the underlying model is unknown, we report the surrogate estimation error

$$\mathrm{Err}(\widehat{B},\widehat{C})=\|Y-X\widehat{B}-\widehat{C}\|_F/(nq)$$

with the true covariate matrix X to evaluate the estimation accuracy of the regression coefficient matrix. The results for our proposed method are summarized in Table 2. In view of the results, MCP achieved the highest accuracy in view of the estimation errors. Both MCP and SCAD can successfully identify the true subgroups.

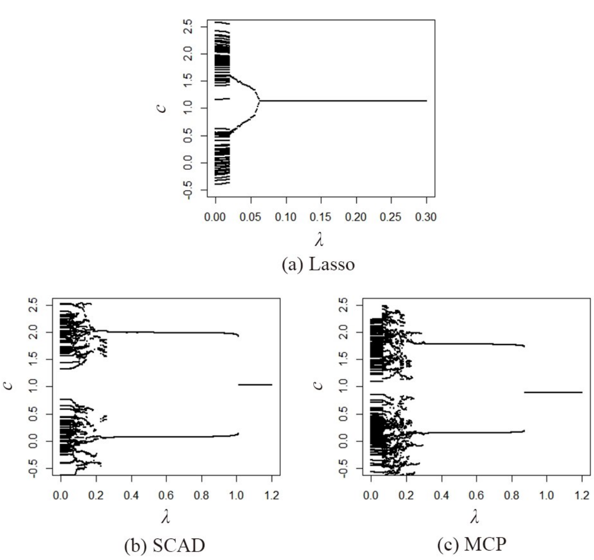

Figure 2. Solution paths for $\widehat{c}_{11}, \widehat{c}_{21}, \cdots, \widehat{c}_{n1}$ against λ for data in Section 4.

Table 2. The performance measures of estimators (1×10^2).

Method	Err($\widehat{B},\widehat{C}$)	Est($\widehat{\alpha}_1$)	Est($\widehat{\alpha}_2$)	$\widehat{K}$
Lasso	165.590	–	–	–
MCP	10.072	4.164	4.930	2.000
SCAD	12.368	6.782	8.976	2.000

6 Conclusions

In this paper, we have extended the method for heterogeneous univariate-response regression to multi-response regression, which recovers regression coefficient matrix and latent heterogeneous factors by concave pairwise fusion penalties. Numerical studies demonstrate the statistical accuracy of the proposed method. Our estimation procedure may be extended to deal with data containing censoring, measurement errors and outliers or more general model settings such as the generalized linear model, which will be interesting topics for future research.

Acknowledgments

This work is supported by National Natural Science Foundation of China (Nos. 72071187, 11671374, 71731010, and 71921001), and Fundamental Research Funds for the Central Universities (Nos. WK3470000017 and WK2040000027).

Conflict of interest

The authors declare no conflict of interest.

Author information

WU Jie is currently a PhD candidate at School of Management, University of Science and Technology of China. She received her BS degree from Anhui University of Technology in 2017. Her research mainly focuses on high dimensional variable selection, classification.

ZHOU jia (corresponding author) is currently a PhD candidate at School of Management, University of Science and Technology of China. She received her BS degree from Anhui University of Finance and Economics in 2016. Her research mainly focuses on high dimensional variable selection and statistic inference.

ZHENG Zemin (corresponding author) is Full Professor at Department of Management, University of Science and Technology of China (USTC). He received his BS degree from USTC in 2010 and PhD degree from University of Southern California in 2015. Afterwards, he is working in Department of Management, USTC. His research interests include high dimensional statistical inference and big data problems. He has published some high-level articles in top journals of Statistics, including *Journal of Royal Statistical Society Series B*, *The Annals of Statistics*, *Operations Research*, and *Journal of Machine Learning Research*. In addition, he has also presided over several scientific research projects, including the Youth Project of the National Natural Foundation of China (NSFC), General Projects of the NSFC.

References

[1] Zhang Z, Nie L, Soon G, et al. The use of covariates and random effects in evaluating predictive biomarkers under a potential outcome framework. Annals of Applied Statistics, 2014, 8(4): 2336–2355.

[2] Shen J, He X. Inference for subgroup analysis with a structured logistic-normal mixture model. Journal of the American Statistical Association, 2015, 110(509): 303–312.

[3] Hastie T, Tibshirani R. Discriminant analysis by Gaussian mixtures. Journal of the Royal Statistical Society Series B, 1966, 58 (1): 155–176.

[4] Wei S, Kosorok M. Latent supervised learning. Journal of the American Statistical Association, 2013, 108 (503): 957–970.

[5] Guo F J, Levina E, Michailidis G, et al. Pairwise variable selection for high-dimensional model-based clustering. Biometrics, 2010, 66(3): 793–804.

[6] Chi E C, Lange K. Splitting methods for convex clustering. Journal of Computational and Graphical Statistics, 2015, 24(4): 994–1013.

[7] Wang J, Li J, Li Y, et al. A model-based multithreshold method for subgroup identification. Statistics in Medicine, 2019, 38: 2605–2631.

[8] Li J, Yue M, Zhang, W. Subgroup identification via homogeneity pursuit for dense longitudinal/spatial data. Statistics in Medicine, 2019, 38: 3256–3271.

[9] Ma S, Huang J. A concave pairwise fusion approach to subgroup analysis. Journal of the American Statistical Association, 2017, 112(517): 410–423.

[10] Izenman A. Reduced-rank regression for the multivariate linear model. Journal of Multivariate Analysis, 1975, 5 (2): 248–264.

[11] Reinsel G, Velu R. Multivariate Reduced-Rank Regression: Theory and Applications. New York: Springer, 1998.

[12] Yuan M, Ekici A, Lu Z, et al. Dimension reduction and coefficient estimation in multivariate linear regression. Journal of the Royal Statistical Society Series B, 2007, 69 (3): 329–346.

[13] Chen L, Huang J Z. Sparse reduced-rank regression for simultaneous dimension reduction and variable selection. Journal of the American Statistical Association, 2012, 107: 1533–1545.

[14] Liu H, Wang L, Zhao T. Calibrated multivariate regression with application to neural semantic basis discovery. Journal of Machine Learning Research, 2015, 16: 1579–1606.

[15] Zheng Z, Bahadori M T, Liu Y, et al. Scalable interpretable multi-response regression via SEED. Journal of Machine Learning Research, 2019, 20: 1–34.

[16] Tibshirani R J. Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society Series B, 1996, 58(1): 267–288.

[17] Zhang C H. Nearly unbiased variable selection under minimax concave penalty. Annals of Statistics, 2010, 38 (2): 894–942.

[18] Fan J, Li R. Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American Statistical Association, 2011, 96(459): 1348–1360.

[19] Wang L, Chen G, Li H. Group SCAD regression analysis for microarray time course gene expression data. Bioinformatics, 2007, 23(12): 1486–1494.

[20] Tseng P. Convergence of a block coordinate descent method for nondifferentiable minimization. Journal of Optimization Theory and Applications, 2001, 109: 475–494.

基于多响应回归的子群分析

吴捷[1,2],周佳[1,2*],郑泽敏[1,2*]
1. 中国科学技术大学国际金融研究院,安徽合肥 230601;
2. 中国科学技术大学管理学院,安徽合肥 230026
* 通讯作者. E-mail:tszhjia@mail.ustc.edu.cn; zhengzm@ustc.edu.cn

摘要: 由于研究异质效应可以消除个体差异的影响,使估计结果更加准确,因此在现代大数据应用中,正确识别异质群体中的亚群越来越受欢迎.尽管文献增长迅速,但现有的方法大多集中在异质单响应回归上,如何在多响应问题中准确识别亚组仍不清楚.本文提出了一种新的基于凹融合的异质多响应回归方法,该方法能同时估计系数矩阵并识别子群结构.此外,通过建立估计一致性,为所提出的方法提供了理论保证.数值研究证明了该方法的有效性.

关键词: 多响应回归;子群分析;凹惩罚;ADMM 算法

Appendix

A.1 Proof of Proposition 3.1

By the definition of δ^{m+1}, we have

$$L(C^{m+1},B^{m+1},\delta^{m+1},v^m)\leqslant L(C^{m+1},B^{m+1},\delta,v^m)$$

for any δ. Define

$$f^{m+1}=\inf_{AC^{m+1}-\delta=0}\left\{\frac{1}{2}\|Y-XB^{m+1}-C^{m+1}\|_F^2+\sum_{i<j}p_r(\|\delta_{ij}\|_2,\lambda)\right\}=\inf_{AC^{m+1}-\delta=0}L(C^{m+1},B^{m+1},\delta,v^m).$$

Then

$$L(C^{m+1},B^{m+1},\delta^{m+1},v^m)\leqslant f^{m+1}.$$

Let t be an integer. Since $v^{m+1}=v^m+\upsilon(AC^{m+1}-\delta^{m+1})$, then we have

$$v^{m+t-1}=v^m+\sum_{i=1}^{t-1}(AC^{m+i}-\delta^{m+i}),$$

and thus

$$L(C^{m+t},B^{m+t},\delta^{m+t},v^{m+t-1})=\frac{1}{2}\|Y-XB^{m+t}-C^{m+t}\|_F^2+\langle v^{m+t-1},AC^{m+t}-\delta^{m+t}\rangle+$$
$$\frac{\upsilon}{2}\|AC^{m+t}-\delta^{m+t}\|_F^2+\sum_{i<j}p_r(\|\delta_{ij}{}^{m+t}\|_2,\lambda)=\frac{1}{2}\|Y-XB^{m+t}-C^{m+t}\|_F^2+\langle v^m,AC^{m+t}-\delta^{m+t}\rangle+$$
$$\upsilon\sum_{i=1}^{t-1}\langle AC^{m+i}-\delta^{m+i},AC^{m+t}-\delta^{m+t}\rangle+\frac{\upsilon}{2}\|AC^{m+t}-\delta^{m+t}\|_F^2+\sum_{i<j}p_r(\|\delta_{ij}{}^{m+t}\|_2,\lambda)\leqslant f^{m+t}.$$

Since the objective function $L(C,B,\delta,v)$ is differentiable with respect to and is convex with respect to, by applying the results in Theorem 4.1 of Ref. [20], the sequence (C^m,B^m,δ^m) has a limit point, denoted by (C^*,B^*,δ^*). Then we have

$$f^*=\lim_{m\to\infty}f^{m+1}=\lim_{m\to\infty}f^{m+t}=\inf_{AC^*-\delta=0}\left\{\frac{1}{2}\|Y-XB^*-C^*\|_F^2+\sum_{i<j}p_r(\|\delta_{ij}^*\|_2,\lambda)\right\},$$

and for all $t\geqslant 0$

$$\lim_{m\to\infty}L(C^{m+t},B^{m+t},\delta^{m+t},v^{m+t-1})=$$
$$\frac{1}{2}\|Y-XB^*-C^*\|_F^2+\sum_{i<j}p_r(\|\delta_{ij}^*\|_2,\lambda)+\lim_{m\to\infty}\langle v^m,AC^*-\delta^*\rangle+\left(t-\frac{1}{2}\right)\upsilon\|AC^*-\delta^*\|_F^2\leqslant f^*.$$

Hence, $\lim_{m\to\infty}\|r^m\|_F^2=r^*=\|AC^*-\delta^*\|_F^2=0$.

Since C^{m+1} minimizes $L(C,B^m,\delta^m,v^m)$ by definition, we have that

$$L(C,Z^m,\delta^m,v^m)/\partial C=0,$$

and moreover,

$$L(C^{m+1},B^m,\delta^m,v^m)/\partial C=XB^{m+1}+C^m-Y+A^{\mathrm{T}}V^m+\upsilon A^{\mathrm{T}}(AC^{m+1}-\delta^m)=$$

$$XB^{m+1} + C^m - Y + A^{\mathrm{T}}(V^m + \upsilon(AC^{m+1} - \delta^m)) =$$
$$XB^{m+1} + C^m - Y + A^{\mathrm{T}}(V^{m+1} - \upsilon(AC^{m+1} - \delta^{m+1}) + \upsilon(AC^{m+1} - \delta^m)) =$$
$$XB^{m+1} + C^m - Y + A^{\mathrm{T}} V^{m+1} + \upsilon A^{\mathrm{T}}(\delta^{m+1} - \delta^m).$$

Therefore,

$$s^{m+1} = \upsilon A^{\mathrm{T}}(\delta^{m+1} - \delta^m) = -(XB^{m+1} + C^m - Y) - A^{\mathrm{T}} V^{m+1}.$$

Since $\| AC^* - \delta^* \|_F^2 = \mathbf{0}$,

$$\lim_{m\to\infty} L(C, B^m, \delta^m, v^m)/\partial C = \lim_{m\to\infty} XB^{m+1} + C^m - Y + A^{\mathrm{T}} V^{m+1} = XB^* + C^* - Y + A^{\mathrm{T}} V^* = 0.$$

Therefore, $\lim_{m\to\infty} s^{m+1} = \mathbf{0}$.

A.2 Proof of Theorem 3.1

For every $C \in \mathscr{M}_\psi$, it can be written as $C = W\alpha$. Recall $U = (W, X)$. We have

$$\begin{pmatrix} \widehat{C}^{\mathrm{or}} \\ \widehat{B}^{\mathrm{or}} \end{pmatrix} = \arg\min_{\alpha\in\mathbb{R}^{K\times q},\, B\in\mathbb{R}^{p\times q}} \frac{1}{2} \| Y - XB - C \|_F^2 = \arg\min_{\alpha\in\mathbb{R}^{K\times q},\, B\in\mathbb{R}^{p\times q}} \frac{1}{2} \| Y - XB - W\alpha \|_F^2.$$

Thus

$$\begin{pmatrix} \widehat{\alpha}^{\mathrm{or}} \\ \widehat{B}^{\mathrm{or}} \end{pmatrix} = [(W, X)^{\mathrm{T}}(W, X)]^{-1}(W, X)^{\mathrm{T}} Y = (U^{\mathrm{T}} U)^{-1} U^{\mathrm{T}} Y.$$

Then

$$\begin{pmatrix} \widehat{\alpha}^{\mathrm{or}} - \alpha^0 \\ \widehat{B}^{\mathrm{or}} - B^0 \end{pmatrix} = (U^{\mathrm{T}} U)^{-1} U^{\mathrm{T}} E.$$

Hence

$$\left\| \begin{pmatrix} \widehat{\alpha}^{\mathrm{or}} - \alpha^0 \\ \widehat{B}^{\mathrm{or}} - B^0 \end{pmatrix} \right\|_F \leqslant \| (U^{\mathrm{T}} U)^{-1} \|_2 \| U^{\mathrm{T}} E \|_F \tag{A1}$$

By Condition 3.1, we have

$$\| (U^{\mathrm{T}} U)^{-1} \|_2 \leqslant C_1^{-1} | \psi_{\min} |^{-1} \tag{A2}$$

Moreover, since

$$E = [\epsilon_1, \cdots, \epsilon_q],$$

where $\epsilon_i \in \mathbb{R}^{n\times 1}$ is a vector with its elements i.i.d. sub-Gaussian variable. Since every row of the matrix E is i.i.d. vectors, each element of the vectors ϵ_i is i.i.d. sub-Gaussian variables.

$$\| U^{\mathrm{T}} E \|_F = \| U^{\mathrm{T}}(\epsilon_1, \cdots, \epsilon_q) \|_F = \sum_{i^*=1}^{q} \| U^{\mathrm{T}} \epsilon_{i^*} \|_2.$$

Additionally,

$$P(\| U^{\mathrm{T}} \epsilon_{i^*} \|_\infty > C\sqrt{n\log n}) \leqslant P(\| W^{\mathrm{T}} \epsilon_{i^*} \|_\infty > C\sqrt{n\log n}) + P(\| X^{\mathrm{T}} \epsilon_{i^*} \|_\infty > C\sqrt{n\log n}),$$

for some constant $0 < C < \infty$. Then by Condition 3.3,

$$P(\| W^{\mathrm{T}} \epsilon_{i^*} \|_\infty > C\sqrt{n\log n}) \leqslant \sum_{k=1}^{K} P(| \sum_{j\in\psi_k} \epsilon_{i^*j} | > C\sqrt{n\log n}) \leqslant$$
$$\sum_{k=1}^{K} P(| \sum_{j\in\psi_k} \epsilon_{i^*j} | > \sqrt{|\psi_k|}\, C\sqrt{\log n}) \leqslant 2K\exp(-c_1 C^2 \log n) = 2Kn^{-c_1C^2}.$$

By union bound, Condition 3.1 that $\| X_k \|_2 = \sqrt{n}$, where X_k is the kth column of X, and Condition 3.3,

$$P(\| X^{\mathrm{T}} \epsilon_{i^*} \|_\infty > C\sqrt{n\log n}) \leqslant \sum_{j=1}^{p} P(| X_j^{\mathrm{T}} \epsilon | > \sqrt{n}\, C\sqrt{\log n}) \leqslant 2p\exp(-c_1 C^2 \log n) = 2pn^{-c_1C^2}.$$

It follows that

$$P(\| U^{\mathrm{T}} \epsilon_{i^*} \|_\infty > C\sqrt{n\log n}) \leqslant 2(K+p)\, n^{-c_1C^2}.$$

Since $\| U^{\mathrm{T}} \epsilon_{i^*} \|_2 = \sqrt{K+p}\, \| U^{\mathrm{T}} \epsilon_{i^*} \|_\infty$, then

$$P(\| U^{\mathrm{T}} \epsilon_{i^*} \|_2 > C\sqrt{K+p}\sqrt{n\log n}) \leqslant 2(K+p)\, n^{-c_1C^2}.$$

Then,

$$P(\| U^{\mathrm{T}} E \|_F > Cq\sqrt{K+p}\sqrt{n\log n}) = P(\sum_{i^*=1}^{q} \| U^{\mathrm{T}} \epsilon_{i^*} \|_2 > Cq\sqrt{K+p}\sqrt{n\log n}) \leqslant$$
$$\sum_{i^*=1}^{q} P(\| U^{\mathrm{T}} \epsilon_{i^*} \|_2 > C\sqrt{K+p}\sqrt{n\log n}) \leqslant 2q(K+p)\, n^{-c_1C^2} \tag{A3}$$

Therefore, by (A1), (A2) and (A3), we have with probability at least $1-2q(K+p)n^{-c_1C^2}$,

$$\left\| \begin{pmatrix} \widehat{\alpha}^{\mathrm{or}} - \alpha^0 \\ \widehat{B}^{\mathrm{or}} - B^0 \end{pmatrix} \right\|_F \leqslant CC_1^{-1} \mid \psi_{\min} \mid^{-1} q\sqrt{K+p}\sqrt{n\log n}.$$

The result (9) in Theorem 3.1 is proved by letting $C_1 = c_1^{-1/2}$.

A.3 Proof of Theorem 3.2

In this section we show the results in Theorem 3.2. Define

$$L_n(C, B) = \frac{1}{2}\| Y - XB - C \|_F^2,\ P_n(C) = \lambda \sum_{i<j} \rho(\| c_i - c_j \|_2),$$

$$L_n^{\psi}(\alpha, B) = \frac{1}{2}\| Y - XB - W\alpha \|_F^2,\ P_n^{\psi}(\alpha) = \lambda \sum_{k<k'} \mid \psi_k \mid\mid \psi_{k'} \mid \rho(\| \alpha_i - \alpha_j \|_2),$$

and let

$$Q_n(C, B) = L_n(C,B) + P_n(C),\ Q_n^{\psi}(\alpha, B) = L_n^{\psi}(\alpha,B) + P_n^{\psi}(\alpha).$$

Let $T: \mathcal{M}_{\psi} \to \mathbb{R}^{K\times q}$ be the mapping that $T(C)$ is the $K\times q$ matrix consisting of K vectors with dimension $q\times 1$ and its kth vector component equals to the common value of c_i for $i \in \psi_k$. Let $T^*: \mathbb{R}^{n\times q} \to \mathbb{R}^{K\times q}$ be the mapping that $T^*(C) = \{ |\psi|^{-1} \sum_{i\in\psi_k} c_i,\ k=1,\cdots,K \}^{\mathrm{T}}$. Clearly, when $C \in \mathcal{M}_{\psi}$, $T(C) = T^*(C)$.

By calculation, for every $C \in \mathcal{M}_{\psi}$, we have $P_n(C) = P_n^{\psi}(T(C))$ and for every $\alpha \in \mathbb{R}^{K\times q}$, we have $P_n(T^{-1}(\alpha)) = P_n^{\psi}(\alpha)$. Hence

$$Q_n(C, B) = Q_n^{\psi}(T(C), B),\ Q_n^{\psi}(\alpha, B) = Q_n(T^{-1}(\alpha), B) \qquad \text{(A4)}$$

Consider the neighborhood of (C^0, B^0):

$$\Theta = \{ C \in \mathbb{R}^{n\times q}, B \in \mathbb{R}^{p\times q} : \| ((C - C^0)^{\mathrm{T}}, (B - B^0)^{\mathrm{T}})^{\mathrm{T}} \|_F \leqslant \phi_n \}.$$

By Theorem 3.1, there exists an event E_1 in which

$$\| ((C - C^0)^{\mathrm{T}}, (B - B^0)^{\mathrm{T}})^{\mathrm{T}} \|_F \leqslant \phi_n,$$

and $P(E_1^C) \leqslant 2q(K+p)n^{-1}$. Hence $(\widehat{C}^{\mathrm{or}}, \widehat{B}^{\mathrm{or}}) \in \Theta$ in E_1. For any $C \in \mathbb{R}^{n\times q}$, let $C^* = T^{-1}(T^*(C))$. We show that $(\widehat{C}^{\mathrm{or}}, \widehat{Z}^{\mathrm{or}})$ is a strictly local minimizer of the objective function (6) with probability approaching 1 through the following two steps.

(Ⅰ) In the event E_1, $Q_n(C^*, B) > Q_n(\widehat{C}^{\mathrm{or}}, \widehat{B}^{\mathrm{or}})$ for any $(C^{\mathrm{T}}, B^{\mathrm{T}})^{\mathrm{T}} \in \Theta$ and $((C^*)^{\mathrm{T}}, (B)^{\mathrm{T}})^{\mathrm{T}} \neq ((\widehat{C}^{\mathrm{or}})^{\mathrm{T}}, (\widehat{B}^{\mathrm{or}})^{\mathrm{T}})^{\mathrm{T}}$.

(Ⅱ) There is an event E_2 such that $P(E_2^C) \leqslant 2n^{-1}$. In $E_1 \cap E_2$, there is a neighborhood of $((\widehat{C}^{\mathrm{or}})^{\mathrm{T}}, (\widehat{B}^{\mathrm{or}})^{\mathrm{T}})^{\mathrm{T}}$, denoted by Θ_n such that $Q_n(C, B) \geqslant Q_n(C^*, B)$ for any $((C^*)^{\mathrm{T}}, (B)^{\mathrm{T}})^{\mathrm{T}} \in \Theta \cap \Theta_n$ for sufficiently large n.

Therefore, by the results in (Ⅰ) and (Ⅱ), we have $Q_n(C, B) \geqslant Q_n(\widehat{C}^{\mathrm{or}}, \widehat{B}^{\mathrm{or}})$ for any $((C)^{\mathrm{T}}, (B)^{\mathrm{T}})^{\mathrm{T}} \in \Theta \cap \Theta_n$ and $((C^*)^{\mathrm{T}}, (B)^{\mathrm{T}})^{\mathrm{T}} \neq ((\widehat{C}^{\mathrm{or}})^{\mathrm{T}}, (\widehat{Z}^{\mathrm{or}})^{\mathrm{T}})^{\mathrm{T}}$ in $E_1 \cap E_2$, so that $((\widehat{C}^{\mathrm{or}})^{\mathrm{T}}, \widehat{B}^{\mathrm{or}})^{\mathrm{T}})^{\mathrm{T}}$ is a strict local minimizer of $Q_n(C, B)$ (A4) over the event $E_1 \cap E_2$ with $P(E_1 \cap E_2) \geqslant 1-2q(K+p+1)n^{-1}$ for sufficiently large n.

In the following we prove the result in (Ⅰ). We first show $P_n^{\psi}(T^*(C)) = C_n$ for any $C \in \Theta$, where C_n is a constant which does not depend on C. Let $T^*(C) = \alpha = (\alpha_1^{\mathrm{T}}, \cdots, \alpha_K^{\mathrm{T}})^{\mathrm{T}}$. It suffices to show that $\| \alpha_k - \alpha_{k'} \|_2 > a\lambda$ for all k and k'. Then by Condition 3.1, $\rho(\| \alpha_k - \alpha_{k'} \|_2)$ is a constant, and as a result $P_n^{\psi}(T^*(C))$ is a constant. Since

$$\| \alpha_k - \alpha_{k'} \|_2 \geqslant \| \alpha_k^0 - \alpha_{k'}^0 \|_2 - \sup_k \| \alpha_k - \alpha_k^0 \|_2,$$

and

$$\sup_k \| \alpha_k - \alpha_k^0 \|_2^2 = \sup_k \| \mid \psi_k \mid^{-1} \sum_{i\in\psi_k} c_i - \alpha_k^0 \|_2^2 = \sup_k \| \mid \psi_k \mid^{-1} \sum_{i\in\psi_k} (c_i - c_i^0) \|_2^2 =$$

$$\sup_k \mid \psi_k \mid^{-2} \| \sum_{i\in\psi_k} (c_i - c_i^0) \|_2^2 \leqslant \sup_k \mid \psi_k \mid^{-1} \sum_{i\in\psi_k} \| (c_i - c_i^0) \|_2^2 \leqslant \sup_i \| (c_i - c_i^0) \|_2^2 \leqslant \phi_n^2 \qquad \text{(A5)}$$

then for all k and k'.

$$\| \alpha_k - \alpha_{k'} \|_2 \geqslant \| \alpha_k^0 - \alpha_{k'}^0 \|_2 - \sup_k \| \alpha_k - \alpha_k^0 \|_2 \geqslant b_n - a\,\phi_n > a\,\lambda,$$

where the last inequality follows from the assumption that $b_n > a\lambda \gg \phi_n$. Therefore, we have $P_n^{\psi}(T^*(C)) = C_n$, and hence $Q_n^{\psi}(T^*(C), B) = L_n^{\psi}(T^*(C), B) + C_n$ for all $(C^{\mathrm{T}}, B^{\mathrm{T}})^{\mathrm{T}} \in \Theta$. Since $((\widehat{C}^{\mathrm{or}})^{\mathrm{T}}, (\widehat{B}^{\mathrm{or}})^{\mathrm{T}})^{\mathrm{T}}$ is the unique global minimizer of $L_n^{\psi}(T^*(\alpha), B)$, then $L_n^{\psi}(T^*(C), B) > L_n(\widehat{\alpha}^{\mathrm{or}}, \widehat{B}^{\mathrm{or}})$ for all $(T^*(C)^{\mathrm{T}}, B^{\mathrm{T}})^{\mathrm{T}} \neq ((\widehat{\alpha}^{\mathrm{or}})^{\mathrm{T}}, (\widehat{B}^{\mathrm{or}})^{\mathrm{T}})^{\mathrm{T}}$ and hence $Q_n^{\psi}(T^*(C), B) > Q_n^{\psi}(\widehat{\alpha}^{\mathrm{or}}, \widehat{B}^{\mathrm{or}})$ for all $T^*(C) \neq \widehat{\alpha}^{\mathrm{or}}$. By (A4), we have $Q_n^{\psi}(\widehat{\alpha}^{\mathrm{or}}, \widehat{B}^{\mathrm{or}}) = Q_n(\widehat{C}^{\mathrm{or}}, \widehat{B}^{\mathrm{or}})$ and

$$Q_n^{\psi}(T^*(C), B) = Q_n(T^{-1}(T^*(C)), B) = Q_n(C^*, B).$$

Therefore, $Q_n(C^*, B) \leqslant Q_n(\widehat{C}^{\mathrm{or}}, \widehat{B}^{\mathrm{or}})$ for all $C^* \neq \widehat{C}^{\mathrm{or}}$, and the result in (i) is proved.

Next we prove the result in (Ⅱ). For a positive sequence t_n, let $\Theta_n = \{c_i: \sup_i \| c_i - \widehat{c}_i^{\mathrm{or}} \|_2\}$. For $(C^{\mathrm{T}}, B^{\mathrm{T}})^{\mathrm{T}} \in \Theta \cap \Theta_n$, by Taylor's expansion, we have

$$Q_n(C, B) - Q_n(C^*, B) = \Gamma_1 + \Gamma_2,$$

where

$$\Gamma_1 = -(Y - (I_n, X)((C^m)^{\mathrm{T}}, (B)^{\mathrm{T}})^{\mathrm{T}}) \cdot (C - C^*),$$

$$\Gamma_2 = \sum_{i=1}^{n} \frac{\partial P_n(C^m)}{\partial c_i^{\mathrm{T}}} \cdot (c_i - c_i^*).$$

Here "·" denotes a dot product and $C^m = \alpha C + (1-\alpha)C^*$ for some constant $\alpha \in (0,1)$. Moreover,

$$\begin{aligned}\Gamma_2 = {} & \lambda \sum_{j>i} \rho'(\| c_i^m - c_j^m \|_2) \| c_i^m - c_j^m \|_2^{-1} (c_i^m - c_j^m)^{\mathrm{T}} (c_i - c_i^*) + \\ & \lambda \sum_{j<i} \rho'(\| c_i^m - c_j^m \|_2) \| c_i^m - c_j^m \|_2^{-1} (c_i^m - c_j^m)^{\mathrm{T}} (c_i - c_i^*) = \\ & \lambda \sum_{j>i} \rho'(\| c_i^m - c_j^m \|_2) \| c_i^m - c_j^m \|_2^{-1} (c_i^m - c_j^m)^{\mathrm{T}} (c_i - c_i^*) + \\ & \lambda \sum_{i<j} \rho'(\| c_j^m - c_i^m \|_2) \| c_j^m - c_i^m \|_2^{-1} (c_j^m - c_i^m)^{\mathrm{T}} (c_j - c_j^*) = \\ & \lambda \sum_{j>i} \rho'(\| c_i^m - c_j^m \|_2) \| c_i^m - c_j^m \|_2^{-1} (c_i^m - c_j^m)^{\mathrm{T}} \{(c_i - c_i^*) - (c_j - c_j^*)\}\end{aligned} \tag{A6}$$

When $i,j \in \phi_k$, $c_i^* = c_j^*$, and $c_i^m - c_j^m = \alpha(c_i - c_j)$. Thus,

$$\begin{aligned}\Gamma_2 = {} & \lambda \sum_{k=1}^{K} \sum_{i,j\in\psi_k, i<j} \rho'(\| c_i^m - c_j^m \|_2) \| c_i^m - c_j^m \|_2^{-1} (c_i^m - c_j^m)^{\mathrm{T}} (c_i - c_j) + \\ & \lambda \sum_{k<k'} \sum_{i\in\psi_k, j\in\psi_{k'}} \rho'(\| c_i^m - c_j^m \|_2) \| c_i^m - c_j^m \|_2^{-1} (c_i^m - c_j^m)^{\mathrm{T}} \{(c_i - c_i^*) - (c_j - c_j^*)\}.\end{aligned}$$

Moreover,

$$\sup_i \| c_i^* - c_i^0 \|_2^2 = \sup_k \| \alpha_k - \alpha_k^0 \|_2^2 \leqslant \phi_n^2 \tag{A7}$$

where the last inequality follows from (A5). Since c_i^m is between c_i and c_i^*,

$$\sup_i \| c_i^m - c_i^0 \|_2 = \alpha \sup_i \| c_i - c_i^0 \|_2 + (1-\alpha) \sup_i \| c_i^* - c_i^0 \|_2 \leqslant \alpha\phi_n + (1-\alpha)\phi_n = \phi_n \tag{A8}$$

Hence for $k \neq k'$, $i \in \psi_k$, $j' \in \psi_{k'}$,

$$\| c_i^m - c_j^m \|_2 \geqslant \min_{i\in\psi_k, j'\in\psi_{k'}} \| c_i^0 - c_j^0 \|_2 - 2\max_i \| c_i^m - c_i^0 \|_2 \geqslant b_n - 2\phi_n > a\lambda,$$

and thus $\rho'(\| c_i^m - c_j^m \|_2) = 0$. Therefore,

$$\begin{aligned}\Gamma_2 = {} & \lambda \sum_{k=1}^{K} \sum_{i,j\in\psi_k, i<j} \rho'(\| c_i^m - c_j^m \|_2) \| c_i^m - c_j^m \|_2^{-1} (c_i^m - c_j^m)^{\mathrm{T}} (c_i - c_j) = \\ & \lambda \sum_{k=1}^{K} \sum_{i,j\in\psi_k, i<j} \rho'(\| c_i^m - c_j^m \|_2) \| c_i - c_j \|_2,\end{aligned}$$

where the last step follows from $c_i^m - c_j^m = \alpha(c_i - c_j)$. Furthermore, by the same reasoning as (A5), we have

$$\sup_i \| c_i^* - \widehat{c}_i^{\mathrm{or}} \|_2 = \sup_k \| \alpha_k - \widehat{\alpha}_k^{\mathrm{or}} \|_2 \leqslant \sup_i \| c - \widehat{c}_i^{\mathrm{or}} \|_2.$$

Then

$$\begin{aligned}& \sup_i \| c_i^m - c_j^m \|_2 \leqslant 2 \sup_i \| c_i^m - c_i^* \|_2 \leqslant 2 \sup_i \| c_i - c_i^* \|_2 \leqslant \\ & 2(\sup_i \| c_i - \widehat{c}_i^{\mathrm{or}} \|_2 + \| c_i^* - c_i^* \|_2) \leqslant 4 \sup_i \| c_i - \widehat{c}_i^{\mathrm{or}} \|_2 \leqslant 4t_n.\end{aligned}$$

Hence, $\rho'(\| c_i^m - c_j^m \|_2) \geqslant \rho'(4t_n)$ by concavity of $\rho(\cdot)$. As a result,

$$\Gamma_2 \geqslant \sum_{k=1}^{K} \sum_{i,j\in\psi_k, i<j} \lambda \rho'(4t_n) \| c_i - c_j \|_F \tag{A9}$$

Let

$$Q = (Q_1^{\mathrm{T}}, \cdots, Q_n^{\mathrm{T}})^{\mathrm{T}} = (Y - XB - C^m)^{\mathrm{T}}.$$

Then

$$\Gamma_1 = -Q^{\mathrm{T}} \cdot (C - C^*) = \sum_{k=1}^{K} \sum_{i,j\in\psi_k} \frac{Q_i^{\mathrm{T}} \cdot (c_i - c_j)}{|\psi_k|} =$$

$$-\sum_{k=1}^{K}\sum_{i,j\in\psi_k}\frac{Q_i(c_i-c_j)^{\mathrm{T}}}{2|\psi_k|}-\sum_{k=1}^{K}\sum_{i,j\in\psi_k}\frac{Q_i(c_i-c_j)^{\mathrm{T}}}{2|\psi_k|}=$$
$$-\sum_{k=1}^{K}\sum_{i,j\in\psi_k}\frac{(Q_i-Q_j)(c_i-c_j)^{\mathrm{T}}}{2|\psi_k|}=$$
$$-\sum_{k=1}^{K}\sum_{i,j\in\psi_k,\ i<j}\frac{(Q_i-Q_j)(c_i-c_j)^{\mathrm{T}}}{|\psi_k|} \quad (A10)$$

Moreover,

$$Q_i = y_i - x_i B - c_i^m = E_i + x_i(B^0 - B) + (c_i^0 - c_i^m),$$

and then

$$\sup_i \| Q_i \|_2 \leqslant \sup_i \| E_i + x_i(B^0 - B) + (c_i^0 - c_i^m) \|_2.$$

By Condition 3.3 that $\sup_i \| X_i \|_2 \leqslant C_2\sqrt{p}$, (A8) that $\sup_i \| c_i^m - c_i^0 \|_2 \leqslant \phi_n$ and $\| B^0 - B \|_F \leqslant \phi_n$, we have

$$\sup_i \| Q_i \|_2 \leqslant \| E_i \|_2 + C_2\sqrt{p}\phi_n + \phi_n,$$

By Condition 3.2

$$P(\| E_i \|_2^2 > 2c_1^{-1} q\log n) \leqslant P(\sum_{i=1}^{q} e_{ij}^2 \geqslant 2c_1^{-1} q\log n) =$$
$$\sum_{i=1}^{q} P(e_{ij}^2 \geqslant 2c_1^{-1}\log n) = \sum_{i=1}^{q} P(| e_{ij} | \geqslant \sqrt{2c_1^{-1}\log n}) \leqslant 2qn^{-2}.$$

Thus there is an event E_2 such that $P(E_2{}^C) \leqslant 2qn^{-2}$, and over the event E_2,

$$\sup_i \| Q_i \|_2 \leqslant \sqrt{2c_1^{-1}}\sqrt{q\log n} + C_2\sqrt{p}\phi_n + \phi_n,$$

Then

$$\left| \frac{(Q_i - Q_j)(c_i - c_j)^{\mathrm{T}}}{|\psi_k|} \right| \leqslant |\psi_{\min}|^{-1} 2 \sup_i Q_i(c_i - c_j)^{\mathrm{T}} \leqslant |\psi_{\min}|^{-1} 2 \sup_i \| Q_i \|_2 \| c_i - c_j \|_2 \leqslant$$
$$2 |\psi_{\min}|^{-1}(\sqrt{2c_1^{-1}}\sqrt{q\log n} + C_2\sqrt{p}\phi_n + \phi_n) \| c_i - c_j \|_2.$$

Therefore, the above results together with (A9) and (A10) yield that

$$Q_n(C, B) - Q_n(C^*, B) \geqslant \sum_{k=1}^{K}\sum_{i,j\in\psi_k,\ i<j} \{\lambda \rho'(4t_n) - 2|\psi_{\min}|^{-1}(\sqrt{2c_1^{-1} q\log n} + C_2\sqrt{p}\phi_n + \phi_n)\} \| c_i - c_j \|_2.$$

Let $t_n = o(1)$, then $\rho'(4t_n) \to 1$. Since $\lambda \gg \phi_n$, $p = o(n)$, and $|\psi_{\min}|^{-1} p = o(1)$, then $\lambda \gg |\psi_{\min}|^{-1}\sqrt{q\log n}$, $\lambda \gg |\psi_{\min}|^{-1}\sqrt{p}\phi_n$ and $\lambda \gg |\psi_{\min}|^{-1} p\,\phi_n$. Therefore, $Q_n(C, B) - Q_n(C^*, B) > 0$ for sufficiently large n, so that the result in (Ⅱ) is proved.