Online Updating for Multivariate Regression Model via Modified Cholesky decomposition

JooChul Lee, HeeSeung Kim

1 Introduction

Online updating is one of the important statistical methods as big data arrives in streams. Standard statistics tool to analyze is challenging for handling big data efficiently. Schifano et al.(2016) developed iterative estimating method without data storage requirements for the linear models and estimating equations. In this project, we will consider the multivariate regression model for the online updating. Pourahmadi (1999) studied the maximum likelihood estimators of a generalised linear model for the covariance matrix. In the researcher's study, mean and covariance of response variables are estimated by using covariates. To estimates parameters corresponding to the mean and covariance, the Newton-Raphson algorithm is used. For the online updating in the multivariate regression model, two estimating equations will be considered; one is for the parameters of the mean, and the other is for the parameters of the covariance. We need to consider two estimating equations at the same time to get cumulative estimators. Two steps should be proposed to handle two equations. For the first step, we fix the parameters of the covariance and then estimated the parameters of the mean. Then, the estimated parameters of the mean are fixed to estimate the parameters of the covariance at the second step.

2 Model and Method

Consider the i.i.d sample $\{\mathbf{Y}_i, \mathbf{X}_i\}$ for $i = 1, 2, \ldots, n$ where $\mathbf{Y}_i$, $\mathbf{X}_i$ is m, p and dimensional vectors, respectively. The multivariate regression model is

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i$$

where $\boldsymbol{\beta}$ is the p dimensional vector and $\boldsymbol{\varepsilon}_i \sim N(\boldsymbol{\mu}, \Sigma)$ and $\mathbf{Cov}(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_j)$ for $i \neq j$

2.1 MLE via Modified Cholesky decomposition

By the modified Cholesky decomposition, Σ^{-1} can be expressed by the following

$$\Sigma^{-1}(\boldsymbol{\phi}, \boldsymbol{\sigma}) = C(\boldsymbol{\phi})' D(\boldsymbol{\sigma}) C(\boldsymbol{\phi})$$

where

$$\begin{aligned} -2L(\boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}) &= n \log |\Sigma| + \sum_{i=1}^{n} (y_i - X_i \beta)' \Sigma^{-1} (y_i - X_i \beta) \\ &= n \sum_{i=1}^{p} log \sigma_t^2 + \sum_{i=1}^{n} (r_i - \mathbf{Z}(i)\gamma)' D^{-1} (r_i - \mathbf{Z}(i)\gamma) \end{aligned}$$

where $\mathbf{Z}(i) = (z(i,1), \ldots, z(i,p))'$, $z(i,t) = \sum_{j=1}^{t-1} r_{ij} z_{tj}$

For the score function and Fisher information,

$$U_1(\boldsymbol{\beta}, \boldsymbol{\phi}) = \sum_{i=1}^{n} \mathbf{X}_i' \Sigma^{-1} \boldsymbol{r}_i, \quad \mathbf{I}_{11} = \sum_{i=1}^{n} \mathbf{X}_i' \Sigma^{-1} \mathbf{X}_i$$

$$U_2(\boldsymbol{\beta}, \boldsymbol{\phi}) = \sum_{i=1}^{n} \mathbf{Z}'(i) D^{-1} (\boldsymbol{r}_i - \mathbf{Z}(i)\boldsymbol{\phi}), \quad \mathbf{I}_{22} = nW$$

where $W = \sum_{i=1}^{n} \sigma_i^{-2} W_t$, $W_t = \sum_{j=1}^{t-1} \sum_{k=1}^{t-1} \sigma_{jk} z_{tk} z_{tl}'$, and $\boldsymbol{r}_i = \mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta} = (r_{1i}, \ldots, r_{mi})'$. $\mathbf{Z}(i) = (z(i,1), \ldots, z(i,m))'$ is the $m \times q$ matrix where $z(i,t) = \sum_{j=1}^{t-1} r_{ij} z_{tj}$. Given σ_i^2, $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\phi}})$ are obtained using Fisher scoring algorithm.

2.2 Online Updating Method

Let $\{(\mathbf{Y}_{ki}, \mathbf{X}_{ki}, \mathbf{Z}_{ki}, i = 1, \ldots, n_k\}$ be the k^{th} subset. The score functions for subset k are

$$U_{1,k}(\boldsymbol{\beta}, \boldsymbol{\phi}) = \sum_{i=1}^{n_k} \Psi_1(\mathbf{X}_{ki}, \mathbf{Z}_{ki}, \boldsymbol{\beta}, \boldsymbol{\phi})$$

$$U_{2,k}(\boldsymbol{\beta}, \boldsymbol{\phi}) = \sum_{i=1}^{n_k} \Psi_2(\mathbf{X}_{ki}, \mathbf{Z}_{ki}, \boldsymbol{\beta}, \boldsymbol{\phi})$$

A Taylor Expansion of $-U_{1,k}(\boldsymbol{\beta}, \boldsymbol{\phi}), U_{2,k}(\boldsymbol{\beta}, \boldsymbol{\phi})$ at $\hat{\boldsymbol{\beta}}_{n_k,k}, \hat{\boldsymbol{\phi}}_{n_k,k}$ respectively is given by

$$-U_{1,k}(\boldsymbol{\beta}, \boldsymbol{\phi}) = \mathbf{A}_{1,k}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{n_k,k}) + \mathbf{R}_1$$

$$-U_{2,k}(\boldsymbol{\beta}, \boldsymbol{\phi}) = \mathbf{A}_{2,k}(\boldsymbol{\phi} - \hat{\boldsymbol{\phi}}_{n_k,k}) + \mathbf{R}_2$$

The cumulative estimators are

$$\hat{\boldsymbol{\beta}}_K = (\sum_{k=1}^{K} \mathbf{A}_{1,k}(\phi))^{-1} \sum_{k=1}^{K} \mathbf{A}_{1,k} \hat{\boldsymbol{\beta}}_{n_k,k} \tag{1}$$

$$\hat{\boldsymbol{\phi}}_K = (\sum_{k=1}^{K} \mathbf{A}_{2,k}(\beta))^{-1} \sum_{k=1}^{K} \mathbf{A}_{2,k} \hat{\boldsymbol{\phi}}_{n_k,k} \tag{2}$$

2.3 Algorithm

In this project, we propose an algorithm to find $(\hat{\boldsymbol{\beta}}_K, \hat{\boldsymbol{\phi}}_K)$

Step 1. At k^{th} subset, $(\hat{\boldsymbol{\beta}}_{n_k,k}, \hat{\boldsymbol{\phi}}_{n_k,k})$ are obtained by using Newton-Raphson algorithm.

Step 2. At fixed $\boldsymbol{\phi}$, Obtain $\hat{\boldsymbol{\beta}}_K$ based on (1). $\hat{\boldsymbol{\phi}}_{K-1}$ is used as the initial value.

Step 3. At fixed $\boldsymbol{\beta}$, Obtain $\hat{\boldsymbol{\phi}}_K$ based on (2).

Repeat Step 2 - 3, until convergence

3 Numerical experiments

3.1 Simulation

3.2 Real data analysis