Journal of the Korean Statistical Society 42 (2013) 143–148

Contents lists available at SciVerse ScienceDirect

Journal of the Korean Statistical Society

journal homepage: www.elsevier.com/locate/jkss

A theoretical view of the envelope model for multivariate linear regression as response dimension reduction

Jae Keun Yoo *

Department of Statistics, Ewha Womans University, Seoul 120-750, Republic of Korea

ARTICLE INFO

Article history:
Received 1 August 2011
Accepted 12 March 2012
Available online 5 April 2012

AMS 2000 subject classifications:
primary 62G08
secondary 62H05

Keywords:
Envelope model
Multivariate linear regression
Reducing subspaces
Response dimension reduction

ABSTRACT

The envelope model recently developed for the classical multivariate linear regression have potential gain in efficiency in estimating unknown parameters over usual maximum likelihood estimation. In this paper, we theoretically investigate the envelope model as dimension reduction for response variables and connect them to existing methods.

1. Introduction

When the interest is placed on changes of multi-dimensional responses in distribution as predictors vary, multivariate linear regression should be one of the popular statistical tools.

The classical multivariate linear regression of $\mathbf{Y} \in \mathbb{R}^r|\mathbf{X} \in \mathbb{R}^p$ with $r \geq 2$ is as follows:

$$\mathbf{Y}|\mathbf{X} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^r$ is an intercept vector, $\boldsymbol{\beta} \in \mathbb{R}^{r \times p}$ is an unknown coefficient matrix, the error vector $\boldsymbol{\varepsilon} \in \mathbb{R}^r \sim MN(0, \boldsymbol{\Sigma} \geq 0) \perp\!\!\!\perp \mathbf{X}$. A notation '$\perp\!\!\!\perp$' indicates statistical independence, and MN stands for the multivariate normal distribution. In addition, it is assumed that $\boldsymbol{\Sigma} > 0$ throughout the rest of the paper.

When dimensions of $\mathbf{Y}$ and $\mathbf{X}$ are high, the maximum likelihood estimation (MLE) for the parameters, especially the regression coefficient matrix $\boldsymbol{\beta}$, may not be efficient. This is problematic, when prediction of responses for a new observation of $\mathbf{X}$ is a main issue in the regression study.

As an alternative in such case, Cook, Li, and Chiaromonte (2010, CLC) proposed the envelope model under some conditions, which force a connection between the conditional mean $E(\mathbf{Y}|\mathbf{X})$ and the conditional covariance, $\text{cov}(\mathbf{Y}|\mathbf{X}) = \boldsymbol{\Sigma}$. By constructing the minimal reducing subspace of $\boldsymbol{\Sigma}$, which is fully informative to $E(\mathbf{Y}|\mathbf{X})$, dimensions of the parameter in model (1) are reduced, and it leads more efficient MLE than the usual MLE. This model is called the *envelope model*. Another interpretation of the envelope model is to partition the response subspace into the reducing subspace and its complement. Since the former subspace is fully informative to $E(\mathbf{Y}|\mathbf{X})$, the projection of response variables onto the reducing subspace can be thought as dimension reduction of responses.

* Tel.: +82 2 3277 6717.
E-mail address: peter.yoo@ewha.ac.kr.

1226-3192/$ – see front matter
doi:10.1016/j.jkss.2012.03.004

The main purpose of this paper is to provide a theoretical view of the envelope model for multivariate regression as response dimension reduction. Since, under the envelope model, lower-dimensional linearly transformed responses given $\mathbf{X}$ alone are informative to $E(\mathbf{Y}|\mathbf{X})$, the linear combination of $\mathbf{Y}$ can be considered as dimension-reduced responses. For this, we theoretically investigate the envelope model in the context of response dimension reduction developed by Yoo and Cook (2008, YC).

Response dimension reduction becomes important in various fields of study. Analysis of repeated measures, longitudinal data, functional data, curve or time series data is often difficult due to high dimensionality of $\mathbf{Y}$, although the dimension of $\mathbf{X}$ is relatively low. The study of such data would be facilitated if we could find a low dimensional linear transform of $\mathbf{Y}$ that adequately describes the regression relationship. For example, Leurgans, Moyeed, and Silverman (1993) applied the canonical correlation analysis to functional data in order to apply smoothing in a suitable way by reducing the dimensions. More recently, Li, Aragon, Shedden, and Agnan (2003) have reduced the dimension of multivariate response to adequately analyze China climate data. The data includes 12 dimensional responses.

The organization of the paper is as follows. Section 2 is devoted to reviewing both the envelope model for multivariate linear regression and response dimension reduction in YC. Section 3 contains a new interpretation of the envelope model as response dimension reduction in the context of YC. In Section 4, we summarize our work.

For notational conveniences, define $\boldsymbol{\Sigma}_{\mathbf{U}} = \text{cov}(\mathbf{U})$ for a random vector $\mathbf{U} \in \mathbb{R}^u$, and $\mathcal{S}(\mathbf{B})$ stands for a subspace spanned by the columns of $\mathbf{B} \in \mathbb{R}^{r\times p}$. And, for a subspace $\mathcal{S}$ of $\mathbb{R}^r$, $\mathcal{S}^{\perp}$ stands for its orthogonal complement. The original paper of the envelope model by CLC can be acquired from http://www.stat.umn.edu/~dennis/RecentArticles/CLC.pdf.

2. Review: envelope model and response dimension reduction

2.1. Envelopes

Envelopes are subspaces and, especially, are originated from the concepts of invariant and reducing subspaces. So, first, we start with an invariant subspace. A subspace $\mathcal{S}$ of $\mathbb{R}^r$ is an invariant subspace of $\mathbf{M} \in \mathbb{R}^{r\times r}$, if $\mathbf{M}\mathcal{S} \subseteq \mathcal{S}$. Moreover, if $\mathbf{M}\mathcal{S}^{\perp} \subseteq \mathcal{S}^{\perp}$, $\mathcal{S}$ is a reducing subspace of $\mathbf{M}$.

Now we define $\mathbf{M}$-envelopes as follows. Let $\mathbf{M} \in \mathbb{S}^{r\times r}$ and let $\mathcal{S} \subseteq \mathcal{S}(\mathbf{M})$, where $\mathbb{S}$ stands for symmetric matrices. The $\mathbf{M}$-envelope of $\mathcal{S}$, notationally $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$, is the intersection of all reducing subspaces of $\mathbf{M}$ that contain $\mathcal{S}$. By the definition, $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$ is the minimal and unique reducing subspace among all possible ones. For more details regarding invariant and reducing subspaces and $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$, readers can refer Section 2 of CLC.

To develop an envelope model under (1), we consider $\mathbf{M}$ as $\boldsymbol{\Sigma}$, which is the conditional covariance of $\text{cov}(\mathbf{Y}|\mathbf{X})$ or the covariance of the random error vector $\boldsymbol{\varepsilon}$ in (1). Also, as a choice of $\mathcal{S}$, we consider $\mathcal{B} = \mathcal{S}(\boldsymbol{\beta})$. Letting $d = \dim(\mathcal{B})$ and $u = \dim\{\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})\}$, it is assumed that $0 < d \leq u \leq r$ throughout the rest of the paper. By following the definition of $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$, $\boldsymbol{\Sigma}$ should be partitioned along with $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$ and $\mathcal{E}^{\perp}_{\boldsymbol{\Sigma}}(\mathcal{B})$ in the envelope model.

2.2. Envelope model for multivariate linear regression

For the classical multivariate linear regression in (1), we connect $\boldsymbol{\Sigma}$ and $\mathcal{B}$ through $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$ by assuming the existence of $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$. It should be again noted that $\mathcal{B} \subseteq \mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$ and $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$ reduces $\boldsymbol{\Sigma}$. We will denote that a $r \times u$ matrix $\boldsymbol{\Gamma}$ is a semi-orthogonal basis matrix of $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$ throughout the rest of the paper.

Then we can state that (1) $\boldsymbol{\beta}$ does not have full-column rank, if $u < r$; (2) if so, parts of $\boldsymbol{\beta}$ are fully informative to regression; (3) $\boldsymbol{\Gamma}$, that is $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$, can fully explain $\boldsymbol{\beta}$, because $\boldsymbol{\beta} = \boldsymbol{\Gamma}\boldsymbol{\nu}$; (4) it implies that the MLE of $\boldsymbol{\beta}$ can be obtained through lower-dimensional matrix $\boldsymbol{\Gamma}$. Along with the statements above, the following results are easily derived under the existence of $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$ in (1).

R1. $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$ with $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2 = 0$ and $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B}) = \mathcal{S}(\boldsymbol{\Sigma}_1)$.

R2. Model (1) can be re-written as follows:

$$\mathbf{Y}|\mathbf{X} = \boldsymbol{\alpha} + \boldsymbol{\Gamma}\boldsymbol{\nu}\mathbf{X} + \boldsymbol{\varepsilon}. \tag{2}$$

R3. $\boldsymbol{\Sigma}_1 = \boldsymbol{\Gamma}\boldsymbol{\Omega}\boldsymbol{\Gamma}^{\mathrm{T}}$ and $\boldsymbol{\Sigma}_2 = \boldsymbol{\Gamma}_0\boldsymbol{\Omega}_0\boldsymbol{\Gamma}_0^{\mathrm{T}}$, where a $r \times (r-u)$ matrix $\boldsymbol{\Gamma}_0$ is the orthogonal complement of $\boldsymbol{\Gamma}$, $\boldsymbol{\Omega} = \boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\Gamma}$, and $\boldsymbol{\Omega}_0 = \boldsymbol{\Gamma}_0^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\Gamma}_0$.

R4. $\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{Y} \perp\!\!\!\perp \boldsymbol{\Gamma}_0^{\mathrm{T}}\mathbf{Y}|\mathbf{X}$.

R5. $\boldsymbol{\Gamma}_0^{\mathrm{T}}\mathbf{Y} \perp\!\!\!\perp \mathbf{X}$ and $\boldsymbol{\Gamma}_0^{\mathrm{T}}\mathbf{Y} \perp\!\!\!\perp \mathbf{X}|\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{Y}$.

Results R1–R3 directly come from properties of envelopes. Result R4 can be proved by R3 under (1), and it rules out any possibility that $\boldsymbol{\Gamma}_0^{\mathrm{T}}\mathbf{X}$ contributes to the regression. The relation of $\boldsymbol{\beta} = \boldsymbol{\Gamma}\boldsymbol{\nu}$ and R4 directly implies R5.

Now we consider model (2) as an alternative of (1), and the model in (2) is called the *envelope model* for multivariate linear regression. To have insight about how efficient model (2) can be, we compare the total number of parameters for both. In model (2), it should be $r + pu + u(r-u) + u(u+1)/2 + (r-u)(r-u+1)/2 = r + pu + r(r+1)/2$, while model (1) has $r + pr + r(r+1)/2$ parameters. The difference between the two is $p(r-u)$, and with high dimensional p and relatively small u to r, the difference clearly gets bigger.