Communications for Statistical Applications and Methods
2024, Vol. 31, No. 2, 179–189
https://doi.org/10.29220/CSAM.2024.31.2.179
Print ISSN 2287-7843 / Online ISSN 2383-4757

MBRDR: R-package for response dimension reduction in multivariate regression

Heesung Ahn[a], Jae Keun Yoo[1,a]

[a]Department of Statistics, Ewha Womans University, Korea

Abstract

In multivariate regression with a high-dimensional response $\mathbf{Y} \in \mathbb{R}^r$ and a relatively low-dimensional predictor $\mathbf{X} \in \mathbb{R}^p$ (where $r \geq 2$), the statistical analysis of such data presents significant challenges due to the exponential increase in the number of parameters as the dimension of the response grows. Most existing dimension reduction techniques primarily focus on reducing the dimension of the predictors ($\mathbf{X}$), not the dimension of the response variable ($\mathbf{Y}$). Yoo and Cook (2008) introduced a response dimension reduction method that preserves information about the conditional mean $E(\mathbf{Y} \mid \mathbf{X})$. Building upon this foundational work, Yoo (2018) proposed two semi-parametric methods, principal response reduction (PRR) and principal fitted response reduction (PFRR), then expanded these methods to unstructured principal fitted response reduction (UPFRR) (Yoo, 2019). This paper reviews these four response dimension reduction methodologies mentioned above. In addition, it introduces the implementation of the `mbrdr` package in R. The `mbrdr` is a unique tool in the `R` community, as it is specifically designed for response dimension reduction, setting it apart from existing dimension reduction packages that focus solely on predictors.

Keywords: multivariate regression, nonparametric reduction, prinicipal response reduction, principal fitted response reduction, unstructured principal response reduction, R-package

1. Introduction

Sufficient dimension reduction (SDR) in regression replaces p-dimensional predictor $\mathbf{X}$ to lower dimensional linear projection $\mathbf{M}^\mathsf{T}\mathbf{X}$ without loss of information on $\mathbf{Y}|\mathbf{X}$. It can be expressed as $\mathbf{Y} \perp\!\!\!\perp \mathbf{X}|\mathbf{M}^\mathsf{T}\mathbf{X}$. For all $\mathbf{M} \in \mathbb{R}^{p\times d}$, $\mathcal{S}(\mathbf{M})$ is called a sufficient dimension reduction (SDR) subspace, where $\mathcal{S}(\mathbf{M})$ is the subspace spanned by columns of $\mathbf{M}$. The intersection of all possible dimension reduction subspace is called the central subspace $\mathcal{S}_{\mathbf{Y}|\mathbf{X}}$, which is the target of SDR. Sliced inverse regression (SIR) (Li, 1991) and sliced average variance estimation (SAVE) (Cook and Weisberg, 1991) are ones of the most popular SDR methods. SIR uses inverse mean $E(\mathbf{X}|\mathbf{Y})$ and SAVE uses inverse covariance $\text{cov}(\mathbf{X}|\mathbf{Y})$ to estimate central subspace $\mathcal{S}_{\mathbf{Y}|\mathbf{X}}$.

Multivariate regression of multi-dimensional $\mathbf{Y} \in \mathbb{R}^r|\mathbf{X} \in \mathbb{R}^p$, $r \geq 2$ has been common in many fields including repeated measures, longitudinal studies, time series data, functional data analysis. However, analysis of such data is often challenging. It is because dimension of response is high,

For Jae Keun Yoo and Heesung Ahn, this work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Korean Ministry of Education (RS-2023-00240564 and RS-2023-00217022).

[1]Corresponding author: Department of Statistics, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Korea. E-mail: peter.yoo@ewha.ac.kr

Published 31 March 2024 / journal homepage: http://csam.or.kr

while the dimension of the predictors is relatively low. Then, the number of parameters in the analysis exponentially increases as the dimension of the responses grows. Reducing the dimension of responses still capturing the information of regression would be helpful to handle such data. Unfortunately, most dimension reduction methodologies including SIR and SAVE focused on reducing the dimension of the predictors $\mathbf{X}$, not response $\mathbf{Y}$.

Yoo and Cook (2008) proposed a response dimension reduction methodology in a multivariate regression, without loss of information on the conditional mean $E(\mathbf{Y}|\mathbf{X})$. They defined two types of response dimension reduction, called linear and conditional response reduction, and provided a non-parametric methodology to estimate the linear response reduction.

Yoo (2013) investigated theoretical relation between linear and conditional response reduction under the envelope model setting (Cook *et al.*, 2010) and opened a possibility of a model-based response dimension reduction. Following this seminal work, Yoo (2018) proposed two semi-parametric response dimension reduction approaches, called principal response reduction (PRR) and principal fitted response reduction (PFRR). Yoo (2008) showed that these semi-parametric approaches outperform non-parametric method from Yoo and Cook (2008) through numerical studies and real data example. Yoo (2019) developed another model based approach called unstructured prinicpal fitted response reduction (UPFRR), which do not assume the structure of the covariance matrix in PFRR from Yoo (2018).

For the response dimension reduction for multivariate regression, the `mbrdr` package is recently developed in R. It can implement the Yoo-Cook method and three versions of model-based response dimension reduction mentioned above. The `mbrdr` package can be found in R-CRAN (https://cran.r-project.org/web/packages/mbrdr/index.html). The `mbrdr` package is the very first package in the response dimension reduction in the sufficient dimension reduction context. The existing dr package is used only for reducing the dimension of the predictors, not responses. So, the `mbrdr` is unique, and it makes the `mbrdr` valuable in the R community.

The organization of the paper is as follows. Section 2 provides reviews for four response dimension reduction methodologies mentioned above, which are the Yoo-Cook method, PRR, PFRR, and UPFRR. The implementation of `mbrdr` is introduced in Section 3. Section 4 summarizes the work.

We will use the following notations throughout the rest of the paper. A p-dimensional random variable $\mathbf{X}$ will be denoted as $\mathbf{X} \in \mathbb{R}^p$. So, $\mathbf{X} \in \mathbb{R}^p$ means a random variable, although there is no specific mention. For $\mathbf{X} \in \mathbb{R}^p$ and $\mathbf{Y} \in \mathbb{R}^r$, we define that $\mathbf{\Sigma}_x$ and $\mathbf{\Sigma}_y$ are the covariance matrix of $\mathbf{X}$ and $\mathbf{Y}$, respectively. Additionally, it is assumed that $\mathbf{\Sigma}_x$ and $\mathbf{\Sigma}_y$ are positive-definite.

2. Collection of response dimension methodologies in `mbrdr`

2.1. Yoo-Cook method

For a multivariate regression of $\mathbf{Y} \in \mathbb{R}^r|\mathbf{X} \in \mathbb{R}^p$, it is supposed that there exists a $r \times q$ matrix $\mathbf{L}$ to have the smallest rank among all matrices to satisfy the following relation for $E(\mathbf{Y}|\mathbf{X})$:

$$E(\mathbf{Y} \mid \mathbf{X}) = E\left\{\mathbf{P}^{\mathsf{T}}_{\mathbf{L}(\mathbf{\Sigma}_y)}\mathbf{Y} \mid \mathbf{X}\right\}, \tag{2.1}$$

where $q \leq r$ and $\mathbf{P}_{\mathbf{L}(\mathbf{\Sigma}_y)} = \mathbf{L}(\mathbf{L}^{\mathsf{T}}\mathbf{\Sigma}_y\mathbf{L})^{-1}\mathbf{L}^{\mathsf{T}}\mathbf{\Sigma}_y$ is an orthogonal projection operator relative to the inner product $\langle\omega_1, \omega_2\rangle_{\mathbf{\Sigma}_y} = \omega_1^{\mathsf{T}}\mathbf{\Sigma}_y\omega_2$.

Equation (2.1) indicates that the predictors $\mathbf{X}$ have effects to the components of the conditional mean $E(\mathbf{Y}|\mathbf{X})$ only through $P_{\mathbf{L}(\mathbf{\Sigma}_y)}$. So, the q-dimensional linearly transformed $\mathbf{P}^{\mathsf{T}}_{\mathbf{\Sigma}_y}\mathbf{Y}$ can successfully replace the original response $\mathbf{Y}$ without loss of information on $E(\mathbf{Y}|\mathbf{X})$. In Yoo and Cook (2008), this response dimension reduction is called as a linear response reduction.

Next, suppose that there exists a $k \times k$ matrix $\mathbf{K}$ satisfying the following equivalences:

$$E(\mathbf{Y} \mid \mathbf{X}) = E\{E(\mathbf{Y} \mid \mathbf{X}, \mathbf{K}^{\mathrm{T}}\mathbf{Y}) \mid \mathbf{X}\} = E\{E(\mathbf{Y} \mid \mathbf{K}^{\mathrm{T}}\mathbf{Y}) \mid \mathbf{X}\} = E\{g(\mathbf{K}^{\mathrm{T}}\mathbf{Y}) \mid \mathbf{X}\}, \tag{2.2}$$

where $k \geq r$, $\mathbf{K} \neq \mathbf{I}_r$ and $g(\cdot)$ is an unknown function.

By the last equivalence, another dimension reduction of $\mathbf{Y}$ can be done, if $k < r$, and this response reduction is called a conditional response reduction. In Yoo and Cook (2008), the column spaces of $\mathbf{L}$ and $\mathbf{K}$ are called a response dimension reduction subspace.

Yoo and Cook (2008) prove that $\mathcal{S}(\mathbf{K}) \subseteq \mathcal{S}(\mathbf{L})$ for $\mathbf{L}$ and $\mathbf{K}$ in equations (2.1) and (2.2), respectivey, and that $\mathcal{S}(\mathbf{K}) = \mathcal{S}(\mathbf{L})$ under the following condition: **A1.** $E(\mathbf{Y}|\mathbf{K}^{\mathrm{T}}\mathbf{Y} = a)$ is linear in a. The condition holds, if $\mathbf{Y}$ is elliptically distributed. If condition **A1** is not satisfied, $\mathbf{Y}$ is usually power-transformed for the normality. Under condition A1, the quantity $\mathbf{\Sigma}_y^{-1}\,\mathrm{cov}(\mathbf{Y}, \mathbf{X})\mathbf{\Sigma}_x^{-1}$ is proposed to estimate $\mathbf{L}$ and $\mathbf{K}$ in Yoo and Cook (2008).

2.2. Principal response reduction

We consider the following multivariate regression model with assuming $E(\mathbf{Y}) = 0$ and $E(\mathbf{X}) = 0$ without loss of generality:

$$\mathbf{Y} = \mathbf{\Gamma}\boldsymbol{\nu}_x + \boldsymbol{\varepsilon}, \tag{2.3}$$

where $\mathbf{\Gamma} \in \mathbb{R}^{r \times d}$ with $\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma} = \mathbf{I}_d$ and $d \leq r$, $\boldsymbol{\varepsilon} \sim N(0, \mathbf{\Sigma}), \mathrm{cov}(\boldsymbol{\nu}_x, \boldsymbol{\varepsilon}) = 0$ and $\boldsymbol{\nu}_x$ is a d-dimensional unknown random function of the predictors $\mathbf{X}$ with a positive definite sample covariance and $\mathbf{\Sigma}_x\boldsymbol{\nu}_x = 0$. If $\boldsymbol{\nu}_x = \mathbf{X}$, the model in (2.3) is the same as a multivariate linear regression.

A crucial assumption required for the model given in (2.3) is that $\mathcal{S}(\mathbf{\Gamma})$ is an invariant and reducing subspace of $\mathbf{\Sigma}$. This guarantees that $\mathbf{\Sigma} = \mathbf{\Gamma}\mathbf{\Omega}\mathbf{\Gamma}^{\mathrm{T}} + \mathbf{\Gamma}_0\mathbf{\Omega}_0\mathbf{\Gamma}_0^{\mathrm{T}}$, where $\mathbf{\Gamma}_0 \in \mathbb{R}^{r \times (r-d)}$ with $\mathbf{\Gamma}_0^{\mathrm{T}}\mathbf{\Gamma}_0 = \mathbf{I}_{r-d}$ and $\mathbf{\Gamma}_0^{\mathrm{T}}\mathbf{\Gamma} = 0$, $\mathbf{\Omega} = \mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{\Gamma}$ and $\mathbf{\Omega}_0 = \mathbf{\Gamma}_0^{\mathrm{T}}\mathbf{\Sigma}\mathbf{\Gamma}_0$.

According to Yoo (2018), for model (2.3), we have that $E(\mathbf{Y}|\mathbf{X}) = E(\mathbf{P}^{\mathrm{T}}_{\mathbf{\Gamma}(\mathbf{\Sigma}_y)}\mathbf{Y}|\mathbf{X})$. That is, the original response $\mathbf{Y}$ can be reduced through $\mathbf{\Gamma}$ without loss of information of $E(\mathbf{Y}|\mathbf{X})$.

Then, the unknown $\mathbf{\Gamma}$ in model (2.3) is estimated by maximizing its likelihood function, because the normality of $\boldsymbol{\varepsilon}$ is assumed. Denote $\hat{\mathbf{\Sigma}}_y$ as an unusual moment estimator of $\mathbf{\Sigma}_y$. Yoo (2018) proves that the maximum likelihood estimator (MLE) of $\mathbf{\Gamma}$ is a set of the eigenvectors corresponding to the first d largest eigenvalues of $\hat{\mathbf{\Sigma}}_y$. This dimension reduction is called principal response reduction (PRR).

2.3. Principal fitted response reduction (PFRR)

For PRR, the information of $\mathbf{X}$ is excluded in the estimation of $\mathbf{\Gamma}$. To incorporate the predictors $\mathbf{X}$, it is assumed that $\boldsymbol{\nu}_x = \boldsymbol{\psi}\mathbf{f}_x$:

$$\mathbf{Y} = \mathbf{\Gamma}\boldsymbol{\psi}\mathbf{f}_x + \boldsymbol{\varepsilon}, \tag{2.4}$$

where $\boldsymbol{\psi}$ is an unknown $d \times q$ matrix, and $\mathbf{f}_x \in \mathbb{R}^q$ is a known q dimensional vector-valued function of $\mathbf{X}$ with $\mathbf{\Sigma}_x\mathbf{f}_x = 0$. For convenience, the following notations are defined.

$\mathbb{Y}$: the $n \times r$ data matrix for the responses.

$\mathbb{X}$: the $n \times p$ data matrix for the predictors.

$\mathbb{F}$: the $q \times n$ matrix constructed by stacking $\mathbf{f}_x^{\mathrm{T}}$ and $\mathbf{P}_{\mathbb{F}} = \mathbb{F}(\mathbb{F}^{\mathrm{T}}\mathbb{F})^{-1}\mathbb{F}^{\mathrm{T}}$.

$\hat{\boldsymbol{\Sigma}}_{fit} = \mathbb{Y}^{\mathrm{T}}\mathbf{P}_{\mathbb{F}}\mathbb{Y}/n$ and $\hat{\boldsymbol{\Sigma}}_{res} = \hat{\boldsymbol{\Sigma}}_{y} - \hat{\boldsymbol{\Sigma}}_{fit}$.

As the candidates of $\mathbf{f}_x$, Yoo (2018) considers $\mathbf{X}, \mathbf{X}^2, \exp(\mathbf{X})$, their combinations and the cluster indicator of $\mathbf{X}$ constructed from the K-means clustering algorithm. If $\mathbf{f}_x = \mathbf{X}$, $\mathbf{P}_{\mathbb{F}}\mathbb{Y}$ is equal to the ordinary least squares, and hence $\hat{\boldsymbol{\Sigma}}_{fit}$ is the regression product sums of square.

The MLE of $\boldsymbol{\Gamma}$ in model (2.4) is not in a close form. The likelihood function in $\boldsymbol{\Gamma}$ is summarized as follows according to Yoo (2018):

$$L(\boldsymbol{\Gamma}, \boldsymbol{\Gamma}_0) = -\frac{n}{2}\log\left|\boldsymbol{\Gamma}_0^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_y\boldsymbol{\Gamma}_0\right| - \frac{n}{2}\log\left|\boldsymbol{\Gamma}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_{res}\boldsymbol{\Gamma}\right|.$$

Therefore, the MLE of $\boldsymbol{\Gamma}$ is clearly affected by both $\hat{\boldsymbol{\Sigma}}_y$ and $\hat{\boldsymbol{\Sigma}}_{res}$. A sequential selection algorithm among a set of all the eigenvectors of $\hat{\boldsymbol{\Sigma}}_y$, $\hat{\boldsymbol{\Sigma}}_{fit}$ and $\hat{\boldsymbol{\Sigma}}_{res}$ is adopted from Cook (2007: Section 6.2). This approach to estimate $\boldsymbol{\Gamma}$ is called principal fitted response reduction (PFRR).

2.4. Unstructured principal fitted response reduction

We consider the following model with assuming that $\boldsymbol{\varepsilon} \sim N(0, \boldsymbol{\Sigma} > 0)$ and $\mathrm{cov}(\boldsymbol{\nu}_x, \boldsymbol{\varepsilon}) = 0$:

$$\mathbf{Y} = \boldsymbol{\Gamma}\nu_x + \boldsymbol{\varepsilon}. \tag{2.5}$$

The difference between models (2.3) and (2.5) is the structure of $\boldsymbol{\Sigma}$. In model 2.5, the condition that $\boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{\Omega}\boldsymbol{\Gamma}^{\mathrm{T}} + \boldsymbol{\Gamma}_0\boldsymbol{\Omega}_0\boldsymbol{\Gamma}_0^{\mathrm{T}}$ is not required any more.

In Yoo (2019), the following relationship between $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_y$ for the invariant condition is shown that $\mathcal{S}(\boldsymbol{\Sigma}\boldsymbol{\Gamma}) \subseteq \mathcal{S}(\boldsymbol{\Gamma})$ if and only if $\mathcal{S}(\boldsymbol{\Sigma}_y\boldsymbol{\Gamma}) \subseteq \mathcal{S}(\boldsymbol{\Gamma})$. That is, the invariant condition for $\boldsymbol{\Sigma}$ is equivalent to that for $\boldsymbol{\Sigma}_y$. According to Yoo (2019), it is established that $E(\mathbf{Y}|\mathbf{X}) = E(\mathbf{P}^{\mathrm{T}}_{\Gamma(\Sigma_y)}\mathbf{Y}|\mathbf{X})$ for model (2.5), if the invariance of $\mathcal{S}(\boldsymbol{\Gamma})$ for $\boldsymbol{\Sigma}_y$ holds. So, hereafter, the invariant condition of $\boldsymbol{\Gamma}$ for $\boldsymbol{\Sigma}_y$ will be assumed in model (2.5).

To incorporate the information of $\mathbf{X}$ in the estimation of $\boldsymbol{\Gamma}$, its fitted component version is constructed as follows:

$$\mathbf{Y} = \boldsymbol{\Gamma}\boldsymbol{\psi}\mathbf{f}_x + \boldsymbol{\varepsilon}. \tag{2.6}$$

We define the following quantities:

$\mathbf{E}_d$ and $\mathcal{S}_d(\mathbf{E})$ are the first d largest eigenvectors of a matrix $\mathbf{E}_d$ and the column subspace of $\mathbf{E}_d$, respectively.

$\mathbf{B} = \hat{\boldsymbol{\Sigma}}^{-1/2}\hat{\boldsymbol{\Sigma}}_{fit}\hat{\boldsymbol{\Sigma}}^{-1/2}$, $\mathbf{B}_{res} = \hat{\boldsymbol{\Sigma}}_{res}^{-1/2}\hat{\boldsymbol{\Sigma}}_{fit}\hat{\boldsymbol{\Sigma}}_{res}^{-1/2}$, and $\mathbf{B}_y = \hat{\boldsymbol{\Sigma}}_y^{-1/2}\hat{\boldsymbol{\Sigma}}_{fit}\hat{\boldsymbol{\Sigma}}_y^{-1/2}$.

$\hat{\boldsymbol{\Lambda}} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_q)$ and $\hat{\mathbf{V}} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_q)$ are the ordered eigenvalues and corresponding eigenvectors of $\mathbf{B}_{res}$.

$\hat{\mathbf{K}}_d = \mathrm{diag}(0, \ldots, 0, \hat{\lambda}_{d+1}, \ldots, \hat{\lambda}_q)$.

Then, under model (2.6), the following results are derived in Yoo (2019):

(1) $\hat{\mathcal{S}}(\boldsymbol{\Gamma}) = \hat{\boldsymbol{\Sigma}}^{1/2}\mathcal{S}_d(\mathbf{B})$ or $\hat{\boldsymbol{\Gamma}} = \hat{\boldsymbol{\Sigma}}^{1/2}\mathbf{B}_d$.

(2) $\hat{\boldsymbol{\Sigma}} = \hat{\boldsymbol{\Sigma}}_{res} + \hat{\boldsymbol{\Sigma}}_{res}^{1/2}\hat{\mathbf{V}}\hat{\mathbf{K}}_d\hat{\mathbf{V}}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_{res}^{1/2} = \hat{\boldsymbol{\Sigma}}_{res}^{1/2}(\mathbf{I}_r + \hat{\mathbf{V}}\hat{\mathbf{K}}_d\hat{\mathbf{V}}^{\mathrm{T}})\hat{\boldsymbol{\Sigma}}_{res}^{1/2}$.

(3) $L^d_{UPFRR} = (-n/2)\log|\hat{\boldsymbol{\Sigma}}_{res}| + (n/2)\sum_{i=d+1}^{q}\log(1 + \hat{\lambda}_i)$.

(4) $\hat{\mathcal{S}}(\boldsymbol{\Gamma}) = \hat{\boldsymbol{\Sigma}}^{1/2}\mathcal{S}_d(\mathbf{B}) = \hat{\boldsymbol{\Sigma}}_{res}^{1/2}\mathcal{S}_d(\mathbf{B}_{res}) = \hat{\boldsymbol{\Sigma}}_y^{1/2}\mathcal{S}_d(\mathbf{B}_y)$.

The response reduction through model (2.6) will be called unstructured principal fitted response reduction (UPFRR).

3. Illustration of `mbrdr` package

3.1. Outline of `mbrdr` package

R-package `mbrdr` can realize four response dimension reduction methodologies discussed in Section 2. The arguments are as follows.

```
mdrdr(formula, method="upfrr", data, subset, na.action=na.fail, weights).
```

The main function `mbrdr` creates "mbrdr" class and four subclasses depending on the values of `mbrdr`. The values of `method` and its resulting subclasses are as follows. The default is "upfrr".

`method = "yc"`: Yoo-Cook method/ "yc" subclass

`method = "prr"`: principal response reduction / "prr" subclass

`method = "pfrr"`: principal fitted response reduction / "pfrr" subclass

`method = "upfrr"`: unstructured principal fitted response reduction / "upfrr" subclass.

The function `mbrdr` provides eigenvectors for response dimension reduction and test statistics about decision of its dimension. For the estimation of the true dimension d, cumulative sum of eigenvalues is provided in all four methods. These methods follow a sequence of hypothesis test, $H_0 : d_y = m$ vs. $H_1 : d_y = r$ (Rao, 1965). Starting from $m = 0$, if H_0 is rejected, we increment m by 1 and test again. $\hat{d}_y$ is determined to be m at the first time H_0 is not rejected. The test statistic of Yoo-Cook method is provided in Yoo and Cook (2008: Section 3.3). Unlike Yoo-Cook method, prr, pfrr, and upfrr use maximum likelihood estimator to estimate $\mathbf{\Gamma}$. For the test statistic, these three methods adopt likelihood ratio test (LRT) to decide the optimal dimension for reduction. For pfrr and upfrr, d can be estimated by χ^2 statistic, which are $\chi^2_{q(r-m)}$ and $\chi^2_{(q-m)(r-m)}$ respectively. However, χ^2 statistics cannot be applied to prr. It is because $\mathbf{\Omega}$ is not estimable, then the covariance matrix $\mathbf{\Sigma}$ cannot be estimated too. On the other hand, for example, in pfrr, $\mathbf{\Omega}$ and $\mathbf{\Omega}_0$ can be estimated with $\hat{\mathbf{\Gamma}}^{\mathrm{T}}\hat{\mathbf{\Sigma}}_{res}\hat{\mathbf{\Gamma}}^{\mathrm{T}}$ and $\hat{\mathbf{\Gamma}}_0^{\mathrm{T}}\hat{\mathbf{\Sigma}}_{\mathrm{y}}\hat{\mathbf{\Gamma}}_0^{\mathrm{T}}$, respectively. For this reason, `mbrdr` package provide the result of χ^2-test for dimension only for `method = "pfrr"` and `method = "upfrr"`.

To fit `method = "pfrr"` and `method = "upfrr"` in `mbrdr` function, users should select $\mathbf{f}_x$ via `fx.choice` option. The option `fx.choice` is implemented through a function `choose.fx`. The function `choose.fx` returns $n \times q$ matix of $\mathbf{f}_x$. q depends on the choice of $\mathbf{f}_x$. When user run the function `mbrdr` with fx.choice specified, the function `choose.fx` is automatically implemented. In the function `choose.fx`, the predictor $\mathbf{X}$ are normalized to have zero sample means and a sample correlation matrix, instead of a sample covariance matrix. The option `fx.choice` has the following four values.

`fx.choice = 1`: $\mathbf{f}_x = \mathbf{X}$.

`fx.choice = 2`: $\mathbf{f}_x = (\mathbf{X}, \mathbf{X}^2)$.

`fx.choice = 3`: $\mathbf{f}_x = (\mathbf{X}, \exp(\mathbf{X}))$.

`fx.choice = 4`: $(c - 1)$ dummy variables for c-cluster indicators constructed through K-means clustering $\mathbf{X}$.

For `fx.choice = 4`, `nclust` should be specified in `mbrdr` function. The command `set.seed(0)` will be used for producing the same clustering results in the following examples. The value of

Table 1: The result for `mbrdr` in upfrr

(a) The eigenvector for the dimension reduction

	Dir1	Dir2	Dir3
A4	0.3336	−0.376512	0.8292
B4	−0.3784	0.001862	0.398
A6	0.6515	−0.489923	0.3486
B6	−0.5666	−0.786264	−0.1789

(b) The χ^2-test result for the dimension

	Stat	df	p-value
0D vs ≥ 1D	144.013	36	0.0000
1D vs ≥ 2D	30.087	24	0.1819
2D vs ≥ 3D	7.071	14	0.9319
3D vs ≥ 4D	2.253	6	0.8950

`fx.choice` must be one of 1, 2, 3 and 4. If users want to choose user-own f_x other than `fx.choice`, the option `fx` should be used. If a matrix is given in `fx`, it surpasses any values in `fx.choice`. The default values for `fx.choice`, `nclust` and `fx` are 1, 5 and NULL, respectively. The candidate function f_x can be determined by examining scatter plot matrices between $\mathbf{Y}$ and $\mathbf{X}$. If a linear trend is observed between $\mathbf{Y}$ and $\mathbf{X}$, $f_x = \mathbf{X}$ is recommended. In the case of a non-linear trend, considerations may include $f_x = (\mathbf{X}, \mathbf{X}^2), \exp(\mathbf{X}), (\mathbf{X}, \exp(\mathbf{X}))$. If there is no patterns between $\mathbf{Y}$ and $\mathbf{X}$, cluster indicator from K-means clustering algorithm can be used as f_x. According to numerical studies, $f_x = \mathbf{X}$ is considered a suitable default choice. For illustration of how to use `mbrdr` function, the data set named `mps` is included in the `mbrdr` package. In the following section, we apply the package to the `mps` data.

3.2. Real data example 1: Minneapolis elementary school data

To illustrate the outputs of `mbrdr`, Minneapolis elementary school data in 1972 is adopted. The data set named `mps` is available in the `mbrdr`. The data set is made of 63 observations with 4 responses and 11 predictors. The responses are A4, B4, which are percentage of 4^{th} graders scoring above/below average on a standard 4^{th} grade vocabulary test in 1972, and A6, B6, which are percentage of 6^{th} graders scoring above/below average on a standard 6^{th} grade comprehension test in 1972. 9 predictors will be used in the model: (1) the percentage of children receiving aid to families with dependent children (AFDC), (2) the average percentage of children in attendance during the year (Attend), (3) the percentage of children in the school not living with both parents (B), (4) the number of children enrolled in the school (Enrol), (5) the percentage of adults in the school area who have completed high school (HS), (6) the percent minority children in the area (Minority), (7) the percentage of children who started in a school, but did not finish there (Mobility), (8) the percentage of persons in the school area who are above the federal poverty levels (Poverty), (9) pupil-teacher ratio (PTR). This data set would be suitable for response dimension reduction since the dimension of response is relatively high considering the observation size and the dimension of predictor.

The function `mbrdr` with default options; `method = "upfrr"`, `fx.choice = 1` is as follows. It uses predictors $\mathbf{X}$ in a normalized form.

```
library(mbrdr)
attach(mps)
X<-cbind(AFDC,Attend,B,Enrol,HS,Minority,Mobility,
Poverty,PTR)
Y<-cbind(A4, B4, A6, B6)
rdr0<-mbrdr(Y~X)
summary(rdr0).
```

The results of the function `mbrdr` is in the Table 1. The χ^2-test determines that $\hat{d}_y = 1$ with p-value 0 in a significance level $\alpha = 0.05$. Then, users can extract the eigenvector to reduce the dimension of $\mathbf{Y}$ by the command `summary(rdr0)$ evectors[,1]`. If interpreting the response reduction result

Table 2: The result for `mbrdr` in non-normalized upfrr

(a) The eigenvector for the dimension reduction

	Dir1	Dir2	Dir3
A4	0.3082	−0.34713	0.8836
B4	−0.3578	−0.05721	0.1658
A6	0.6721	−0.52751	−0.4205
B6	−0.5704	−0.77328	−0.1221

(b) The χ^2-test result for the dimension

	Stat	df	p-value
0D vs ≥ 1D	142.730	36	0.0000
1D vs ≥ 2D	29.038	24	0.2187
2D vs ≥ 3D	6.186	14	0.9616
3D vs ≥ 4D	1.645	6	0.9493

Table 3: The result for `mbrdr` in pfrr (fx.choice = 1)

(a) The eigenvector for the dimension reduction

	Dir1	Dir2	Dir3
A4	−0.4310	−0.2592	0.6927
B4	0.3335	−0.6947	−0.4140
A6	−0.6166	−0.5494	−0.1943
B6	0.5681	−0.3852	0.5577

(b) The χ^2-test result for the dimension

	Stat	df	p-value
0D vs ≥ 1D	144.013	36	0.0000
1D vs ≥ 2D	38.489	27	0.0704
2D vs ≥ 3D	23.617	18	0.1680
3D vs ≥ 4D	8.985	9	0.4386

Table 4: The result for `mbrdr` in pfrr (fx.choice = 4, nclust = 6)

(a) The eigenvector for the dimension reduction

	Dir1	Dir2	Dir3
A4	−0.2650	0.4786	0.5218
B4	−0.6901	−0.3662	0.4824
A6	−0.5577	0.5896	−0.5377
B6	−0.3774	−0.5378	−0.4538

(b) The χ^2-test result for the dimension

	Stat	df	p-value
0D vs ≥ 1D	22.217	20	0.3288
1D vs ≥ 2D	14.344	15	0.4996
2D vs ≥ 3D	6.670	10	0.7562
3D vs ≥ 4D	2.117	5	0.8327

through the eigenvectors in the Table 1(a), the first direction indicates a linear combination of 4 responses, placing more weight on A6, B6 than others. The second one rarely represents B4. The third one has strong focus on A4 and does not incorporate B6.

When users want to use the predictors without normalization, users should specify the `fx` in the function `mbrdr` . In the following code, `fx0` means the original predictor $\mathbf{X}$. The function `update(object,...)` can be used to update the previous call and re-fit a model.

```
#non-normalized fx
fx0<-as.matrix(mps[,c(5:7, 9:14)])
rdr.upfrr.fx0<-update(rdr0, fx = fx0)
summary(rdr.upfrr.fx0).
```

The test statistics, p-value and the eigenvectors in Table 2 are similar to the normalized version given in Table 1. The test result is also same for $\hat{d}_y = 1$ with p-value 0.

```
#fitting pfrr, fx.choice = 1
set.seed(0)
rdr.pfrr1<-update(rdr0, method = "pfrr", fx.choice = 1)
summary(rdr.pfrr1).
```

For `method = "pfrr"`, users should specify `fx.choice`. Additionally, if `fx.choice = 4`, `nclust` should be identified. In Table 3, $\hat{d}_y$ returns to be 1 because the test for $d = 0$ vs $d = 1$ is accepted in 95 % significance level and the test for $d = 1$ vs $d = 2$ is rejected in the same level. On the other hand, in Table 4, which implement `method = "pfrr", fx.choice = 4, nclust = 6`, the χ^2-test rejects every hypothesis of dimension for reduction d. It indicates that the response dimension reduction is not proper for the $\mathbf{f}_x$ specified. For `nclust` = 8, 9 and 10, the χ^2-test about $d = 0$ vs $d = 1$ returns to be significant, and the dimension is determined to be 1 in $\alpha = 0.05$.

Table 5: The result for `mbrdr` in Yoo-Cook method

(a) The eigenvector for the dimension reduction

	Dir1	Dir2	Dir3
A4	−0.5511	0.6679	0.4138
B4	−0.5971	−0.6192	0.4102
A6	0.4311	−0.2800	0.6328
B6	0.3924	0.3035	0.5099

(b) The cumulative sum of eigenvalues

	Stat
H0: $d = 0$	0.2076
H0: $d = 1$	0.0331
H0: $d = 2$	0.004

Table 6: The result for `mbrdr` in prr

(a) The eigenvector for the dimension reduction

	Dir1	Dir2	Dir3
A4	0.4786	0.6617	0.2650
B4	−0.3662	−0.3895	0.6901
A6	0.5896	−0.2356	0.5577
B6	−0.5378	0.5957	0.3774

(b) The cumulative sum of eigenvalues

	Stat
H0: $d = 0$	39187
H0: $d = 1$	7936
H0: $d = 2$	2455

```
#fitting pfrr, fx.choice = 4, nclust = 6
set.seed(0)
rdr.pfrr4.6<-update(rdr0,method = "pfrr",fx.choice = 4,nclust = 6)
summary(rdr.pfrr4.6).
#fitting yoo-cook
rdr.yc<-update(rdr0, method = "yc")
summary(rdr.yc).

#fitting prr
rdr.prr<-update(rdr0, method = "prr")
summary(rdr.prr).
```

For `method = "yc"` and `method = "prr"`, χ^2-test is not provided. Instead, the cumulative sum of the eigenvalues, corresponding to the eigenvectors in Table 5, 6, is presented in the return value `stats`. Table 5, 6 shows the cumulative sum of eigenvalues for Yoo-Cook method and prr, respectively. The dimension d is determined by `min(which(summary( rdr.yc)$ stat/sum(rdr.yc$ evalues)>0.95))`. Although cumulative sums of eigenvalues in Table 5, 6(b) looks completely different, cumulative proportions are similar. Cumulative proportion is calculated by `summary(rdr.yc) $stat/sum(rdr.yc $evalues)`, which is part of command choosing d. For Yoo-Cook method, cumulative proportions are 63.01, 10.05, 1.22 for $d = 0, 1, 2$, respectively. For prr, cumulative proportions are 63, 12.76, 3.94 for $d = 0, 1, 2$, respectively. In 95% confidence level, $\hat{d}$ for both Yoo-Cook method and prr is determined to be 1.

3.3. Real data example 2: Epilepsy data

Another real data analysis will be performed using Epilepsy data from Thall and Vail (1990). The data set is about a clinical treatment of 59 epilepsy patients. The patients were randomly devided into 2 groups, one was provided with the anti-epileptic drug Progabide, and the other received a placebo. The number of seizures in the last 2 weeks had been reported 4 times, total of 8 weeks. Four 2-weeks seizure counts are response variables, while the treatment of placebo or progabide, the baseline seizure count, and age are predictors. For treatment variables, placebo is coded as 0 and progabide is coded as 1. We will use logarithm scale for baseline seizure count and age. For this data set, response

Table 7: The result for mbrdr in upfrr

(a) The eigenvector for the dimension reduction

	Dir1	Dir2	Dir3
y1	0.3307	0.78189	−0.4145
y2	0.4992	0.61812	−0.5210
y3	0.5108	0.07975	−0.1132
y4	0.6169	0.01504	0.7375

(b) The χ^2-test result for the dimension

	Stat	df	p-value
0D vs ≥ 1D	42.970	12	0.0000
1D vs ≥ 2D	4.427	5	0.6191
2D vs ≥ 3D	1.344	2	0.5106
3D vs ≥ 4D	0.000	0	1.0000

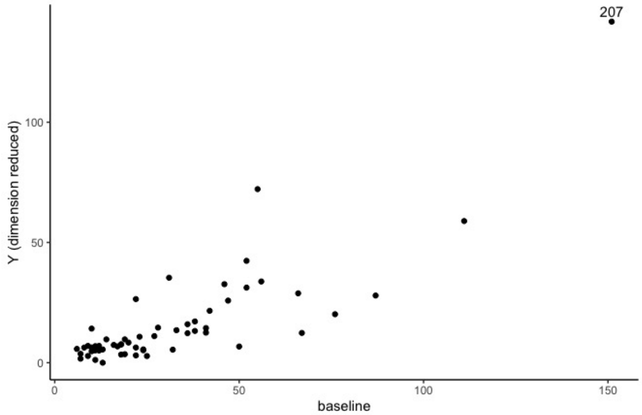

Figure 1: *Scatterplot between dimension-reduced response and baseline.*

Table 8: The result for mbrdr in upfrr after deleting an outlier

(a) The eigenvector for the dimension reduction

	Dir1	Dir2	Dir3
y1	0.5365	0.7467	−0.2838
y2	0.2959	−0.5954	−0.4050
y3	0.4172	−0.2229	−0.4960
y4	0.6712	−0.1959	0.7137

(b) The χ^2-test result for the dimension

	Stat	df	p-value
0D vs ≥ 1D	63.813	12	0.0000
1D vs ≥ 2D	6.519	6	0.3676
2D vs ≥ 3D	0.969	2	0.6160
3D vs ≥ 4D	0.000	0	1.0000

dimension reduction would be suitable since the number of response is bigger than that of predictor. In this section 3.3, only the method UPFRR will proceed.

```
#fitting upfrr
Y_ep<-cbind(epil$y1, epil$y2, epil$y3, epil$y4)
X_ep<-cbind(epil$trt, log(epil$base), log(epil$age))
rdr_ep0<-mbrdr(Y_ep~X_ep, data = epil)
summary(rdr_ep0).
```

The χ^2-test for UPFRR determines that $\hat{d}_y = 1$ with p-value 0 in a significance level $\alpha = 0.05$. In Figure 1, Scatter plot between dimension-reduced response and the predictor baseline suggests that patient # 207 is an outlier, who have much higher baseline counts than other patients. After deleting this observation, choice of dimension $\hat{d}_y$ is still 1 but the eigenvectors are changed significantly. The result after eliminating the outlier is in Table 8. The dimension-reduced response can be obtained by command `Y_ep %*% rdr_ep $ evectors[,1]`.

```
epil<-epil0[-which(epil0$id==207),]
Y_ep<-cbind(epil$y1, epil$y2, epil$y3, epil$y4)
X_ep<-cbind(epil$trt, log(epil$base), log(epil$age))

#upfrr after eliminating outlier
rdr_ep<-mbrdr(Y_ep~X_ep, data = epil)
summary(rdr_ep)
Y_ep%*%rdr_ep$evectors[,1]
```

4. Discussion

In multivariate regression of $\mathbf{Y} \in \mathbb{R}^r | \mathbf{X} \in \mathbb{R}^p$, $r \geq 2$, multi-dimensional response $\mathbf{Y}$ makes the analysis difficult because of the curse of dimension. Response dimension reduction would help solve such difficulty if it still contains the original information of $E(\mathbf{Y}|\mathbf{X})$. Recently, many response dimension reduction methodologies have been developed. Section 2 provides researches on four response dimension reduction methodologies, which are Yoo-Cook method, principal response reduction (PRR), principal fitted response reduction (PFRR), unstructured principal fitted response reduction (UPFRR).

In Section 3.2, the R-package `mbrdr` has been described by using the data set `mps` which is available in the package. Another real data analysis has been conducted with Epilepsy data set for clear understanding of the package in Section 3.3. The package `mbrdr` can implement four response dimension methods mentioned in Section 2. It provides the test for the optimal dimension $\hat{d}$ and the eigenvector for response dimension reduction. The package `mbrdr` enables users to fit these response dimension reduction methods for multivariate regression.

Acknowledgements

For Jae Keun Yoo and Heesung Ahn, this work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Korean Ministry of Education (RS-2023-00240564 and RS-2023-00217022).

References

Cook RD (2007). Fisher lecture: Dimension reduction in regression, *Statistical Science*, **22**, 1–26.

Cook RD, Li B, and Chiaromonte F (2010). Envelope models for parsimonious and efficient multivariate linear regression, *Statistica Sinica*, **20**, 927–960.

Cook RD and Weisberg S (1991). Sliced inverse regression for dimension reduction: Comment, *Journal of the American Statistical Association*, **86**, 328–332.

Li KC (1991). Sliced inverse regression for dimension reduction, *Journal of the American Statistical Association*, **86**, 316–327.

Rao CR (1965). *Linear Statistical Inference and Its Applications*, Wiley, New York.

Thall PF and Vail SC (1990). Some covariance models for longitudinal count data with overdispersion, *Biometrics*, **46**, 657–671.

Yoo JK and Cook RD (2008). Response dimension reduction for the conditional mean in multivariate regression, *Computational Statistics & Data Analysis*, **53**, 334–343.

Yoo JK (2013). A theoretical view of the envelope model for multivariate linear regression as response dimension reduction, *Journal of the Korean Statistical Society*, **42**, 143–148.

Yoo JK (2018). Response dimension reduction: Model-based approach, *Journal of Theoretical and*

Applied Statistics, **52**, 409–425.
Yoo JK (2019). Unstructured principal fitted response reduction in multivariate regression, *Journal of the Korean Statistical Society*, **48**, 561–567.

Received January 4, 2024; Revised February 3, 2024; Accepted February 3, 2024