Short introduction to package MAVE with example

Hang Weiqiang(E0010758@u.nus.edu)

January 29, 2017

1 Introduction

Package MAVE provides several methods to find central space and central mean space using MAVE and OPG method which are proposed by [6, 7]. It also implements sliced inverse regression of kernel version. Formal definition of central space and central mean space can be found in [3, 4]. Sliced inverse regression was proposed by [1], and more discussion can be found in [2].

The main part of package MAVE is written in C++ based on RcppArmadillo package. If there is any problem during installation, please update your Rcpp package and install RcppArmadillo package. It may cause problems.

2 Usage

The primary function in this package is MAVE function. The input arguments are a $n \times p$ covariate matrix X, $n \times 1$ respond matrix Y and the method argument for dimension reduction. The options for the method argument are 'csopg', 'csmave','meanopg','meanmave' and 'ksir', and the default is 'csopg'. 'csopg' and 'csmave' are methods of finding CS by OPG and MAVE, 'meanopg' and 'meanmave' are methods of finding CMS by OPG and MAVE, 'ksir' is the sliced inverse regression of kernel version. Its usage is very simple. The following is a simple example to show how to use it:

```
> X <- matrix(rnorm(400),100,4)
> eps <- matrix(rnorm(100),100,1)
> Y <- as.matrix(X[,1]+X[,2]+(X[,3]+X[,4])*eps)
> library(MAVE)
> rd <- MAVE(X,Y)
```

In the example above, 'csopg' is actually used. In many cases, the result of OPG is as good as MAVE but it is much more time-efficient, hence OPG method is recommended.

The class of object created by MAVE function is mave, which is a S3 class. It contains the basis matrix of central (mean) space of different dimension. The print method is implemented for mave . By typing the name of the returned object, the output can be viewed.

```
> rd

Call:
MAVE(x = X, y = Y)

            [,1]
[1,]  0.66483671
[2,]  0.72679329
[3,]  0.03167453
[4,] -0.16958890
            [,1]        [,2]
[1,]  0.66268528 0.001818867
[2,]  0.72767421 0.110249969
[3,]  0.07025732 0.626130594
[4,] -0.16248807 0.771882190
           [,1]        [,2]       [,3]
[1,]  0.6639452 -0.06147113 -0.1857438
[2,]  0.7242846  0.29150253  0.0897172
[3,]  0.0139921  0.19976314  0.9478183
[4,] -0.1854530  0.93345716 -0.2430854
(only the first 3  central space  are displayed,
        to display the space of dimension k, call object$dir[[k]])
```

The output will display first three matrix of central (mean) space selected by MAVE function. To select best direction by cross-validation method, DIM method can be used.

```
> rd <- DIM(rd)
> rd

Call:
DIM(rd = rd)

Dimension        1         2         3         4
```

```
CV-value        0.19         0.18         0.22         0.28

The selected dimension is  2
```

The output show the corresponding cv-value for different dimensions. More additional information is stored in the object. These includes:

```
> names(rd)

[1] "ky"     "x"      "call"   "method" "dir"    "cv"
```

In order to access other information such as central (mean) space of all dimension, just type

```
> rd$dir

[[1]]
            [,1]
[1,]  0.66483671
[2,]  0.72679329
[3,]  0.03167453
[4,] -0.16958890

[[2]]
            [,1]        [,2]
[1,]  0.66268528 0.001818867
[2,]  0.72767421 0.110249969
[3,]  0.07025732 0.626130594
[4,] -0.16248807 0.771882190

[[3]]
           [,1]        [,2]       [,3]
[1,]  0.6639452 -0.06147113 -0.1857438
[2,]  0.7242846  0.29150253  0.0897172
[3,]  0.0139921  0.19976314  0.9478183
[4,] -0.1854530  0.93345716 -0.2430854

[[4]]
            [,1]       [,2]        [,3]        [,4]
[1,]  0.64941568 -0.5545730 -0.20253596 -0.47925699
[2,]  0.73516269  0.4864200 -0.07695991  0.46584178
[3,]  0.05296834  0.6530715  0.13879256 -0.74258241
[4,] -0.18705466  0.1712952 -0.96632965 -0.04330746
```

References

[1] Li, K. C. (1991). Sliced inverse regression for dimension reduction. Journal of the American Statistical Association, 86(414), 316-327.

[2] Cook, R.D. and Weisberg, S.(1991). Discussion of Li(1991). Journal of the American Statistical Association, 86, 328-332.

[3] Cook, R.D.(1998), Regression Graphics. New York: Wiley

[4] Cook, R. D., and Li, B. (2002). Dimension reduction for conditional mean in regression. Annals of Statistics, 455-474.

[5] Fan, J., and Gijbels, I. (1996). Local Polynomial Modelling and Its Applications, New York: Chapman and Hall.

[6] Xia, Y., Tong, H., Li, W. K., and Zhu, L. X. (2002). An adaptive estimation of dimension reduction space. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64(3), 363-410.

[7] Wang, H., and Xia, Y. (2008). Sliced regression for dimension reduction. Journal of the American Statistical Association, 103(482), 811-821.