Package 'MVLM'

October 12, 2022

Type Package

Title Multivariate Linear Model with Analytic p-Values

Version 0.1.4

Date 2017-03-26

Author Daniel B. McArtor (dmcartor@nd.edu) [aut, cre]

Maintainer Daniel B. McArtor <dmcartor@nd.edu>

Description Allows a user to conduct multivariate multiple regression using analytic p-values rather than classic approximate F-tests.

License GPL (>= 2)

LazyData TRUE

Depends R (>= 3.0.0),

Imports CompQuadForm, parallel

Suggests knitr, rmarkdown

VignetteBuilder knitr

URL http://github.com/dmcartor/MVLM

BugReports http://github.com/dmcartor/MVLM/issues

RoxygenNote 5.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2017-03-26 16:11:40 UTC

R topics documented:

MVLM-package *Multivariate Linear Model with Analytic p-values*

Description

The MVLM package is used to fit linear models with a multivariate outcome. It utilizes the asymptotic null distribution of the multivariate linear model test statistic to compute p-values (McArtor et al., under review). It therefore alleviates the need to use approximate p-values based Wilks Lambda, Pillai's Trace, the Hotelling-Lawley Trace, and Roy's Greatest Root.

Usage

To access this package's tutorial, type the following line into the console:

```
vignette("mvlm-vignette")
```

There is one primary function that comprises this package: `vignette('mvlm-vignette')` There is one primary functions that comprise this package: `mvlm`, which regresses a multivariate outcome onto a set of predictors. Standard functions like `summary`, `fitted`, `residuals`, and `predict` can be called on a `mvlm` output object.

References

Davies, R. B. (1980). The Distribution of a Linear Combination of chi-square Random Variables. Journal of the Royal Statistical Society. Series C (Applied Statistics), 29(3), 323-333.

Duchesne, P., & De Micheaux, P.L. (2010). Computing the distribution of quadratic forms: Further comparisons between the Liu-Tang-Zhang approximation and exact methods. Computational Statistics and Data Analysis, 54(4), 858-862.

McArtor, D. B., Grasman, R. P. P. P., Lubke, G. H., & Bergeman, C. S. (under review). The asymptotic null distribution of the multivariate linear model test statistic. Manuscript submitted for publication.

Examples

```
data(mvlmdata)
Y <- as.matrix(Y.mvlm)
mvlm.res <- mvlm(Y ~ Cont + Cat + Ord, data = X.mvlm)
summary(mvlm.res)
```

`fitted.mvlm` *Extract mvlm Fitted Values*

Description

`fitted` method for class `mvlm`.

Usage

```
## S3 method for class 'mvlm'
fitted(object, ...)
```

Arguments

`object`	Output from `mvlm`
`...`	Further arguments passed to or from other methods.

Value

A data frame of fitted values with the same dimension as the outcome data passed to `mvlm`

Author(s)

Daniel B. McArtor (dmcartor@nd.edu) [aut, cre]

Examples

```
data(mvlmdata)
Y <- as.matrix(Y.mvlm)
mvlm.res <- mvlm(Y ~ Cont + Cat + Ord, data = X.mvlm)
Y.hat <- fitted(mvlm.res)
```

`mvlm` *Conduct multivariate multiple regression and MANOVA with analytic p-values*

Description

`mvlm` is used to fit linear models with a multivariate outcome. It uses the asymptotic null distribution of the multivariate linear model test statistic to compute p-values (McArtor et al., under review). It therefore alleviates the need to use approximate p-values based Wilks' Lambda, Pillai's Trace, the Hotelling-Lawley Trace, and Roy's Greatest Root.

Usage

```
mvlm(formula, data, n.cores = 1, start.acc = 1e-20,
  contr.factor = "contr.sum", contr.ordered = "contr.poly")
```

Arguments

formula	An object of class formula where the outcome (e.g. the Y in the following formula: Y ~ x1 + x2) is a n x q matrix, where q is the number of outcome variables being regressed onto the set of predictors included in the formula.
data	Mandatory data.frame containing all of the predictors passed to formula.
n.cores	Number of cores to use in parallelization through the parallel pacakge.
start.acc	Starting accuracy of the Davies (1980) algorithm implemented in the davies function in the CompQuadForm package (Duchesne & De Micheaux, 2010) that mvlm uses to compute multivariate linear model p-values.
contr.factor	The type of contrasts used to test unordered categorical variables that have type factor. Must be a string taking one of the following values: ("contr.sum", "contr.treatment", "contr.helmert").
contr.ordered	The type of contrasts used to test ordered categorical variables that have type ordered. Must be a string taking one of the following values: ("contr.poly", "contr.sum", "contr.treatment","contr.helmert").

Details

Importantly, the outcome of formula must be a matrix, and the object passed to data must be a data frame containing all of the variables that are named as predictors in formula.

The conditional effects of variables of type factor or ordered in data are computed based on the type of contrasts specified by contr.factor and contr.ordered. If data contains an (ordered or unordered) factor with k levels, a k-1 degree of freedom test will be conducted corresponding to that factor and the specified contrast structure. If, instead, the user wants to assess k-1 separate single DF tests that comprise this omnibus effect (similar to the approach taken by lm), then the appropriate model matrix should be formed in advance and passed to mvlm directly in the data parameter. See the package vigentte for an example by calling vignette('mvlm-vignette').

Value

An object with nine elements and a summary function. Calling summary(mvlm.res) produces a data frame comprised of:

Statistic	Value of the corresponding test statistic.
Numer DF	Numerator degrees of freedom for each test statistic.
Pseudo R2	Size of the corresponding (omnibus or conditional) effect on the multivariate outcome. Note that the intercept term does not have an estimated effect size.
p-value	The p-value for each (omnibus or conditional) effect.

In addition to the information in the three columns comprising summary(mvlm.res), the mvlm.res object also contains:

p.prec	A data.frame reporting the precision of each p-value. These are the maximum error bound of the p-values reported by the davies function in CompQuadForm.
y.rsq	A matrix containing in its first row the overall variance explained by the model for variable comprising Y (columns). The remaining rows list the variance of each outcome that is explained by the conditional effect of each predictor.
beta.hat	Estimated regression coefficients.
adj.n	Adjusted sample size used to determine whether or not the asmptotic properties of the model are likely to hold. See McArtor et al. (under review) for more detail.
data	Original input data and the model.matrix used to fit the model.
formula	The formula passed to mvlm.

Note that the printed output of summary(res) will truncate p-values to the smallest trustworthy values, but the object returned by summary(mvlm.res) will contain the p-values as computed. If the error bound of the Davies algorithm is larger than the p-value, the only conclusion that can be drawn with certainty is that the p-value is smaller than (or equal to) the error bound.

Author(s)

Daniel B. McArtor (dmcartor@nd.edu) [aut, cre]

References

Davies, R. B. (1980). The Distribution of a Linear Combination of chi-square Random Variables. Journal of the Royal Statistical Society. Series C (Applied Statistics), 29(3), 323-333.

Duchesne, P., & De Micheaux, P.L. (2010). Computing the distribution of quadratic forms: Further comparisons between the Liu-Tang-Zhang approximation and exact methods. Computational Statistics and Data Analysis, 54(4), 858-862.

McArtor, D. B., Grasman, R. P. P. P., Lubke, G. H., & Bergeman, C. S. (under review). A new approach to conducting linear model hypothesis tests with a multivariate outcome.

Examples

```
data(mvlmdata)

Y <- as.matrix(Y.mvlm)

# Main effects model
mvlm.res <- mvlm(Y ~ Cont + Cat + Ord, data = X.mvlm)
summary(mvlm.res)

# Include two-way interactions
mvlm.res.int <- mvlm(Y ~ .^2, data = X.mvlm)
summary(mvlm.res.int)
```

predict.mvlm *mvlm Model Predictions*

Description

predict method for class mvlm. To predict using new data, the predictor data frame passed to newdata must have the same number of columns as the data used to fit the model, and the names of each variable must match the names of the original data.

Usage

```
## S3 method for class 'mvlm'
predict(object, newdata, ...)
```

Arguments

object	Output from mvlm
newdata	Data frame of observations on the predictors used to fit the model.
...	Further arguments passed to or from other methods.

Value

A data frame of predicted values

Author(s)

Daniel B. McArtor (dmcartor@nd.edu) [aut, cre]

Examples

```
data(mvlmdata)
Y.train <- as.matrix(Y.mvlm[1:150,])
X.train <- X.mvlm[1:150,]

mvlm.res <- mvlm(Y.train ~ ., data = X.train)

X.test <- X.mvlm[151:200,]
Y.predict <- predict(mvlm.res, newdata = X.test)
```

print.mvlm *Print mvlm Object*

Description

print method for class mvlm

Usage

```
## S3 method for class 'mvlm'
print(x, ...)
```

Arguments

x	Output from mvlm
...	Further arguments passed to or from other methods.

Value

p-value	Analytical p-values for the omnibus test and each predictor

Author(s)

Daniel B. McArtor (dmcartor@nd.edu) [aut, cre]

residuals.mvlm *Extract mvlm Residuals*

Description

residuals method for class mvlm.

Usage

```
## S3 method for class 'mvlm'
residuals(object, ...)
```

Arguments

object	Output from mvlm
...	Further arguments passed to or from other methods.

Value

A data frame of residuals with the same dimension as the outcome data passed to mvlm

Author(s)

Daniel B. McArtor (dmcartor@nd.edu) [aut, cre]

Examples

```
data(mvlmdata)
Y <- as.matrix(Y.mvlm)
mvlm.res <- mvlm(Y ~ Cont + Cat + Ord, data = X.mvlm)
Y.resid <- resid(mvlm.res)
```

`summary.mvlm` *Summarizing mvlm Results*

Description

`summary` method for class `mvlm`

Usage

```
## S3 method for class 'mvlm'
summary(object, ...)
```

Arguments

`object`	Output from `mvlm`
`...`	Further arguments passed to or from other methods.

Value

Calling `summary(mvlm.res)` produces a data frame comprised of:

`Statistic`	Value of the corresponding test statistic.
`Numer DF`	Numerator degrees of freedom for each test statistic.
`Pseudo R2`	Size of the corresponding (omnibus or conditional) effect on the multivariate outcome. Note that the intercept term does not have an estimated effect size.
`p-value`	The p-value for each (omnibus or conditional) effect.

In addition to the information in the three columns comprising `summary(mvlm.res)`, the `mvlm.res` object also contains:

`p.prec`	A data.frame reporting the precision of each p-value. These are the maximum error bound of the p-values reported by the `davies` function in `CompQuadForm`.
`y.rsq`	A matrix containing in its first row the overall variance explained by the model for variable comprising Y (columns). The remaining rows list the variance of each outcome that is explained by the conditional effect of each predictor.
`beta.hat`	Estimated regression coefficients.

`adj.n`	Adjusted sample size used to determine whether or not the asmptotic properties of the model are likely to hold. See McArtor et al. (under review) for more detail.
`data`	Original input data and the `model.matrix` used to fit the model.

Note that the printed output of `summary(res)` will truncate p-values to the smallest trustworthy values, but the object returned by `summary(mvlm.res)` will contain the p-values as computed. If the error bound of the Davies algorithm is larger than the p-value, the only conclusion that can be drawn with certainty is that the p-value is smaller than (or equal to) the error bound.

Author(s)

Daniel B. McArtor (dmcartor@nd.edu) [aut, cre]

References

Davies, R. B. (1980). The Distribution of a Linear Combination of chi-square Random Variables. Journal of the Royal Statistical Society. Series C (Applied Statistics), 29(3), 323-333.

Duchesne, P., & De Micheaux, P.L. (2010). Computing the distribution of quadratic forms: Further comparisons between the Liu-Tang-Zhang approximation and exact methods. Computational Statistics and Data Analysis, 54(4), 858-862.

McArtor, D. B., Grasman, R. P. P. P., Lubke, G. H., & Bergeman, C. S. (under review). A new approach to conducting linear model hypothesis tests with a multivariate outcome.

Examples

```
data(mvlmdata)

Y <- as.matrix(Y.mvlm)

# Main effects model
mvlm.res <- mvlm(Y ~ Cont + Cat + Ord, data = X.mvlm)
summary(mvlm.res)

# Include two-way interactions
mvlm.res.int <- mvlm(Y ~ .^2, data = X.mvlm)
summary(mvlm.res.int)
```

X.mvlm *Simulated predictor data to illustrate the mvlm package.*

Description

See package vignette by calling `vignette('mvlm-vignette')`.

Usage

```
X.mvlm
```

Format

An object of class `data.frame` with 200 rows and 3 columns.

`Y.mvlm` *Simulated outcome data to illustrate the mvlm package.*

Description

See package vignette by calling `vignette('mvlm-vignette')`.

Usage

```
Y.mvlm
```

Format

An object of class `matrix` with 200 rows and 5 columns.

Index