Package 'MultiwayRegression'

May 31, 2019

Type Package

Title Perform Tensor-on-Tensor Regression

Version 1.2

Date 2019-05-28

Author Eric F. Lock

Maintainer Eric F. Lock <elock@umn.edu>

Description Functions to predict one multi-way array (i.e., a tensor) from another multi-way array, using a low-rank CANDECOMP/PARAFAC (CP) factorization and a ridge (L_2) penalty [Lock, EF (2018) <doi:10.1080/10618600.2017.1401544>]. Also includes functions to sample from the Bayesian posterior of a tensor-on-tensor model.

License GPL-3

Imports MASS

Depends R(>= 2.10.0)

NeedsCompilation no

Repository CRAN

Date/Publication 2019-05-31 16:10:03 UTC

R topics documented:

`MultiwayRegression-package` *Perform tensor-on-tensor regression*

Description

Functions to predict one multi-way array (i.e., a tensor) from another multi-way array, using a low-rank CANDECOMP/PARAFAC (CP) factorization and a ridge (L_2) penalty. Also includes functions to sample from the Bayesian posterior of a tensor-on-tensor model.

Details

Package:	MultiwayRegression-package
Type:	Package
Version:	1.2
Date:	2019-05-28
License:	GPL-3

Author(s)

Eric F. Lock

Maintainer: Eric F. Lock <elock@umn.edu>

References

Lock, E. F. (2018). Tensor-on-tensor regression. Journal of Computational and Graphical Statistics, 27 (3): 638-647, 2018.

Examples

```
data(SimData) ##loads simulated X: 100 x 15 x 20 and Y: 100 x 5 x 10
Results <- rrr(X,Y,R=2)  ##Fit rank 2 model with no regularization
Y_pred <- ctprod(X,Results$B,2)  ##Array of fitted values
```

`ctprod` *Compute the contracted tensor product between two multiway arrays.*

Description

Computes the contracted tensor product between two multiway arrays.

Usage

```
ctprod(A,B,K)
```

Arguments

A	An array of dimension P_1 x ... x P_L x R_1 x ... x R_K.
B	An array of dimension R_1 x ... x R_K x Q_1 x ... x Q_M.
K	A positive integer, giving the number of modes to collapse.

Value

An array C of dimension P_1 x ... x P_L x Q_1 x ... x Q_M, given by the contracted tensor product of A and B.

Author(s)

Eric F. Lock

---

rrr *Penalized reduced rank regression for tensors*

Description

Fits a linear model to estimate one multi-way array from another, under the restriction that the coefficient array has given PARAFAC rank. By default, estimates are chosen to minimize a least-squares objective; an optional penalty term allows for L_2 regularization of the coefficient array.

Usage

```
rrr(X,Y,R=1,lambda=0,annealIter=0,convThresh=10^(-5), seed=0)
```

Arguments

X	A predictor array of dimension N x P_1 x ... x P_L.
Y	An outcome array of dimension N x Q_1 x ... X Q_M.
R	Assumed rank of the P_1 x ... x P_L x Q_1 x ... x Q_M coefficient array.
lambda	Ridge (L_2) penalty parameter for the coefficient array.
annealIter	Number of tempering iterations to improve initialization
convThresh	Converge threshold for the absolute difference in the objective function between two iterations
seed	Random seed for generation of initial values.

Value

U	List of length L. U[[l]]: P_l x R gives the coefficient basis for the l'th mode of X.
V	List of length M. V[[m]]: Q_m x R gives the coefficient basis for the m'th mode of Y.
B	Coefficient array of dimension P_1 x ... x P_L x Q_1 x ... x Q_M. Given by the CP factorization defined by U and V.
sse	Vector giving the sum of squared residuals at each iteration.
sseR	Vector giving the value of the objective (sse+penalty) at each iteration.

Author(s)

Eric F. Lock

References

Lock, E. F. (2018). Tensor-on-tensor regression. Journal of Computational and Graphical Statistics, 27 (3): 638-647, 2018.

Examples

```
data(SimData) ##loads simulated X: 100 x 15 x 20 and Y: 100 x 5 x 10
Results <- rrr(X,Y,R=2)  ##Fit rank 2 model with no regularization
Y_pred <- ctprod(X,Results$B,2)  ##Array of fitted values
```

---

rrrBayes *Bayesian inference for reduced rank regression*

Description

Performs Bayesian inference for a linear model to estimate one multi-way array from another, under the restriction that the coefficient array has given PARAFAC rank.

Usage

```
rrrBayes(X,Y,Inits,X.new,R=1,lambda=0,Samples=1000, thin=1,seed=0)
```

Arguments

X	A predictor array of dimension N x P_1 x ... x P_L for the training data.
Y	An outcome array of dimension N x Q_1 x ... x Q_M for the training data.
Inits	Initial values. Inits$U gives a list of length L where Inits$U[[l]]: P_l x R gives the coefficient basis for the l'th mode of X. Inits$V gives a list of length M where Inits$V[[m]]: Q_m x R gives the coefficient basis for the m'th mode of Y. Can be the output of rrr(...).

X.new	Predictor array of dimension M x P_1 x ... x P_L. Each row gives the entries for a new P_1 x ... x P_L predictor observation in vectorized form.
R	Assumed rank of the P_1 x ... x P_L x Q_1 x ... x Q_M coefficient array.
lambda	Ridge (L_2) penalty parameter for the coefficient array, inversely proportional to the variance of the coefficients under a Gaussian prior.
Samples	Length of the MCMC sampling chain.
thin	Thinning value, for thin=j, only every j'th observation in the MCMC chain is saved.
seed	Random seed for generation of initial values.

Value

An array of dimension (Samples/thin) x M x Q_1 x ... x Q_M, giving (Samples/thin) samples from the posterior predictive of the outcome array predicted by Xmat.new.

Author(s)

Eric F. Lock

References

Lock, E. F. (2018). Tensor-on-tensor regression. Journal of Computational and Graphical Statistics, 27 (3): 638-647, 2018.

SimData *Simulated multi-way data for prediction*

Description

Simulated multi-way data for prediction.

Format

- X: predictor array of dimension 100 x 15 x 20
- Y: outcome array of dimension 100 x 5 x 10

X *Simulated multi-way data for prediction*

Description

Simulated multi-way data for prediction.

Format

- X: predictor array of dimension 100 x 15 x 20
- Y: outcome array of dimension 100 x 5 x 10

Y *Simulated multi-way data for prediction*

Description

Simulated multi-way data for prediction.

Format

- X: predictor array of dimension 100 x 15 x 20
- Y: outcome array of dimension 100 x 5 x 10

Index