Package ‘softImpute’

February 20, 2015

Type Package

Title matrix completion via iterative soft-thresholded svd

Version 1.0

Date 2013-03-30

Author Trevor Hastie <hastie@stanford.edu> and Rahul Mazumder
<hastie@stanford.edu>

Maintainer Trevor Hastie <hastie@stanford.edu>

Description Iterative methods for matrix completion that use
nuclear-norm regularization. There are two main approaches. The
one approach uses iterative soft-thresholded svds to impute the
missing values. The second approach uses alternating least
squares. Both have an `EM' flavor, in that at each iteration
the matrix is completed with the current estimate. For large
matrices there is a special sparse-matrix class named
``Incomplete'' that efficiently handles all computations. The
package includes procedures for centering and scaling rows,
columns or both.

License GPL-2

LazyLoad yes

Depends Matrix, methods

NeedsCompilation yes

Repository CRAN

Date/Publication 2013-04-03 07:17:36

R topics documented:

   biScale ................................................................. 2
   complete .............................................................. 4
   Incomplete ............................................................. 5
   Incomplete-class ....................................................... 6
   lambda0 ................................................................. 7
biScale

standardize a matrix to have optionally row means zero and variances one, and/or column means zero and variances one.

Description

A function for standardizing a matrix in a symmetric fashion. Generalizes the scale function in R. Works with matrices with NAs, matrices of class "Incomplete", and matrix in "sparseMatrix" format.

Usage

biScale(x, maxit = 20, thresh = 1e-09, row.center = TRUE, row.scale = TRUE, col.center = TRUE, col.scale = TRUE, trace = FALSE)

Arguments

x
maxit
thresh
row.center
row.scale
col.center
col.scale
trace

Details

This function computes a transformation

\[ X_{ij} - \alpha_i - \beta_j \]

\[ / \gamma_i \tau_j \]

to transform the matrix X. It uses an iterative algorithm based on "method-of-moments". At each step, all but one of the parameter vectors is fixed, and the remaining vector is computed to solve
the required condition. Although in general this is not guaranteed to converge, it mostly does, and quite rapidly. When there are convergence problems, remove some of the required constraints. When any of the row/column centers or scales are provided, they are used rather than estimated in the above model.

Value

A matrix like \( x \) is returned, with attributes:

- **biScale:row**: A list with elements "center" and "scale" (the \( \alpha \) and \( \gamma \) above. If no centering was done, the center component will be a vector of zeros. Likewise, of no row scaling was done, the scale component will be a vector of ones.

- **biScale:column**: Same details as **biScale:row**

For matrices with missing values, the constraints apply to the non-missing entries. If \( x \) is of class "sparseMatrix", then the sparsity is maintained, and an object of class "SparseplusLowRank" is returned, such that the low-rank part does the centering.

Note

This function will be described in detail in a forthcoming paper.

Author(s)

Trevor Hastie, with help from Andreas Buja and Steven Boyd

Maintainer: Trevor Hastie <hastie@stanford.edu>

See Also

softImpute, Incomplete, lambda0, impute, complete, and class "SparseplusLowRank"

Examples

```r
set.seed(101)
n=200
p=100
j=50
np=n*p
missfrac=0.3
x=matrix(rnorm(n*j),n,j)*matrix(rnorm(j*p),j,p)*matrix(rnorm(np),n,p)/5
xc=biScale(x)
ix=seq(np)
imiss=sample(ix,np*missfrac,replace=FALSE)
xna=x
xna[imiss]=NA
xna=biScale(xna, row.scale=FALSE, trace=TRUE)
xnaC=as(xna, "Incomplete")
xnaCb=biScale(xnaC)
nz=trunc(np*.3)
inz=sample(seq(np), nz, replace=FALSE)
i=rown(x)[inz]
j=col(x)[inz]
```
**Description**

These functions produce predictions from the low-rank solution of `softImpute`.

**Usage**

```r
complete(x, object, unscale = TRUE)
impute(object, i, j, unscale = TRUE)
```

**Arguments**

- `x`: a matrix with NAs or a matrix of class "Incomplete".
- `object`: an svd object with components `u`, `d` and `v`.
- `i`: vector of row indices for the locations to be predicted.
- `j`: vector of column indices for the locations to be predicted.
- `unscale`: if `object` has `biScale` attributes, and `unscale=TRUE`, the imputations reversed the centering and scaling on the predictions.

**Details**

`impute` returns a vector of predictions, using the reconstructed low-rank matrix representation represented by `object`. It is used by `complete`, which returns a complete matrix with all the missing values imputed.

**Value**

Either a vector of predictions or a complete matrix. WARNING: if `x` has large dimensions, the matrix returned by `complete` might be too large.

**Author(s)**

Trevor Hastie

**See Also**

`softImpute`, `biScale` and `Incomplete`
**Examples**

```r
set.seed(101)
n=200
p=100
J=50
np=n*p
missfrac=0.3
x=matrix(rnorm(n*p),n,J)+matrix(rnorm(p),n,p)/5
ix=seq(np)
imiss=sample(ix,np*missfrac,replace=FALSE)
xna=x
xna[imiss]=NA
fit1=softImpute(xna,rank=50,lambda=30)
complete(xna,fit1)
```

---

**Incomplete**

`create a matrix of class Incomplete`

**Description**

creates an object of class `Incomplete`, which inherits from class `dgCMatrix`, a specific instance of class `sparseMatrix`

**Usage**

`Incomplete(i, j, x)`

**Arguments**

- **i**: row indices
- **j**: column indices
- **x**: a vector of values

**Details**

The matrix is represented in sparse-matrix format, except the "zeros" represent missing values. Real zeros are represented explicitly as values.

**Value**

a matrix of class `Incomplete` which inherits from class `dgCMatrix`

**Author(s)**

Trevor Hastie and Rahul Mazumder

**See Also**

`softImpute`
Examples

```r
set.seed(101)
n=200
p=100
J=50
np=n*p
missfrac=0.3
x=matrix(rnorm(n*J),n,J)%*%matrix(rnorm(J*p),J,p)+matrix(rnorm(np),n,p)/5
ix=seq(np)
imiss=sample(ix,np*missfrac,replace=FALSE)
xna=x
xna[imiss]=NA
xnaC=as(xna,"Incomplete")
### here we do it a different way to demonstrate Incomplete
### In practise the observed values are stored in this matrix format.
i = row(xna)[-imiss]
j = col(xna)[-imiss]
xnaC=Incomplete(i,j,x=x[-imiss])
```

Incomplete-class

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a sparse matrix inheriting from class dgCMatrix with the NAs represented as zeros</td>
</tr>
</tbody>
</table>

Objects from the Class

Objects can be created by calls of the form `new("Incomplete", ...)` or by calling the function `Incomplete`

Slots

i: Object of class "integer" ~
p: Object of class "integer" ~
Dim: Object of class "integer" ~
Dimnames: Object of class "list" ~
x: Object of class "numeric" ~
factors: Object of class "list" ~

Extends

**Methods**

- `as.matrix` signature(x = "Incomplete"): ...
- `coerce` signature(from = "matrix", to = "Incomplete"): ...
- `complete` signature(x = "Incomplete"): ...

**Author(s)**

Trevor Hastie and Rahul Mazumder

**See Also**

`biScale`, `softImpute`, `Incomplete`, `impute`, `complete`

**Examples**

```r
showClass("Incomplete")
set.seed(101)
n=200
p=100
J=50
np=n*p
missfrac=0.3
x=matrix(rnorm(n*J),n,J)*matrix(rnorm(J*p),J,p)*matrix(rnorm(np),n,p)/5
ix=seq(np)
imiss=sample(ix,np*missfrac,replace=FALSE)
xna=x
xna[imiss]=NA
xnaC=as(xna,"Incomplete")
```

---

`\texttt{\textbf{lambda0}}`  \hspace{10cm} \textit{compute the smallest value for lambda such that softImpute(x,lambda) returns the zero solution.}

---

**Description**

this determines the "starting" lambda for a sequence of values for softImpute, and all nonzero solutions would require a smaller value for lambda.

**Usage**

```r
lambda0(x, lambda = 0, maxit = 100, trace.it = FALSE, thresh = 1e-05)
```
Arguments

x  
An m by n matrix. Large matrices can be in "sparseMatrix" format, as well as "SparseplusLowRank". The latter arise after centering sparse matrices, for example with biScale, as well as in applications such as softImpute. The remaining arguments only apply to matrices x in "sparseMatrix", "Incomplete", or "SparseplusLowRank" format.

lambda  
As in svd.als, using a value for lambda can speed up iterations. As long as the solution is not zero, the value returned adds back this value.

maxit  
maximum number of iterations.

trace.it  
with trace.it=TRUE, convergence progress is reported.

thresh  
convergence threshold, measured as the relative changed in the Frobenius norm between two successive estimates.

Details

It is the largest singular value for the matrix, with zeros replacing missing values. It uses svd.als with rank=2.

Value

a single number, the largest singular value

Author(s)

Trevor Hastie, Rahul Mazumder
Maintainer: Trevor Hastie <hastie@stanford.edu>

References


See Also

softImpute,Incomplete, and svd.als.

Examples

set.seed(101)
n=200
p=100
J=50
mp=m*p
missfrac=0.3
x=matrix(rnorm(n*J),n,J)*matrix(rnorm(J*p),J,p)+matrix(rnorm(np),n,p)/5
ix=seq(np)
imiss=sample(ix,np*missfrac,replace=FALSE)
fit a low-rank matrix approximation to a matrix with missing values via nuclear-norm regularization. The algorithm works like EM, filling in the missing values with the current guess, and then solving the optimization problem on the complete matrix using a soft-thresholded SVD. Special sparse-matrix classes available for very large matrices.

Usage

softImpute(x, rank.max = 2, lambda = 0, type = c("als", "svd"), thresh = 1e-05, maxit = 100, trace.it = FALSE, warm.start = NULL, final.svd = TRUE)

Arguments

x An m by n matrix with NAs. For large matrices can be of class "Incomplete", in which case the missing values are represented as pseudo zeros leading to dramatic storage reduction. x can have been centered and scaled via biScale, and this information is carried along with the solution.

rank.max This restricts the rank of the solution. If sufficiently large, and with type="svd", the solution solves the nuclear-norm convex matrix-completion problem. In this case the number of nonzero singular values returned will be less than or equal to rank.max. If smaller ranks are used, the solution is not guaranteed to solve the problem, although still results in good local minima.

lambda nuclear-norm regularization parameter. If lambda=0, the algorithm reverts to "hardImpute", for which convergence is typically slower, and to local minimum. Ideally lambda should be chosen so that the solution reached has rank slightly less than rank.max. See also lambda0() for computing the smallest lambda with a zero solution.

type two algorithms are implements, type="svd" or the default type="als". The "svd" algorithm repeatedly computes the svd of the completed matrix, and soft thresholds its singular values. Each new soft-thresholded svd is used to re-impute the missing entries. For large matrices of class "Incomplete", the svd is achieved by an efficient form of alternating orthogonal ridge regression. The "als" algorithm uses this same alternating ridge regression, but updates the imputation at each step, leading to quite substantial speedups in some cases. The "als" approach does not currently have the same theoretical convergence guarantees as the "svd" approach.

thresh convergence threshold, measured as the relative changed in the Frobenius norm between two successive estimates.

maxit maximum number of iterations.
trace.it       with trace.it=TRUE, convergence progress is reported.

warm.start   an svd object can be supplied as a warm start. This is particularly useful when
             constructing a path of solutions with decreasing values of lambda and increasing
             rank.max. The previous solution can be provided directly as a warm start for
             the next.

final.svd     only applicable to type="als". The alternating ridge-regressions do not lead
to exact zeros. With the default final.svd=TRUE, at the final iteration, a one
step unregularized iteration is performed, followed by soft-thresholding of the
singular values, leading to hard zeros.

Details

SoftImpute solves the following problem for a matrix $X$ with missing entries:

$$\min ||X - M||^2_o + \lambda ||M||_*.$$

Here $||\cdot||_o$ is the Frobenius norm, restricted to the entries corresponding to the non-missing
entries of $X$, and $||M||_*$ is the nuclear norm of $M$ (sum of singular values). For full details of the
"svd" algorithm are described in the reference below. The "als" algorithm will be described in a
forthcoming article. Both methods employ special sparse-matrix tricks for large matrices with
many missing values. This package creates a new sparse-matrix class "sparsepluslowrank" for
matrices of the form

$$x + ab'$$

where $x$ is sparse and $a$ and $b$ are tall skinny matrices, hence $ab'$ is low rank. Methods for efficient
left and right matrix multiplication are provided for this class. For large matrices, the function
Incomplete() can be used to build the appropriate sparse input matrix from market-format data.

Value

An svd object is returned, with components "u", "d", and "v". If the solution has zeros in "d", the
solution is truncated to rank one more than the number of zeros (so the zero is visible). If the
input matrix had been centered and scaled by biScale, the scaling details are assigned as attributes
inherited from the input matrix.

Author(s)

Trevor Hastie, Rahul Mazumder
Maintainer: Trevor Hastie <hastie@stanford.edu>

References

Learning Large Incomplete Matrices, http://www.stanford.edu/~hastie/Papers/mazumder1PaN
pdf

See Also

biScale, svd.als, Incomplete, lambda0, impute, complete
**Examples**

```r
set.seed(101)
n=200
p=100
J=50
np=n*p
missfrac=0.3
x=matrix(rnorm(n*J),n,J)%*%matrix(rnorm(J*p),J,p)+matrix(rnorm(np),n,p)/5
ix=seq(np)
imiss=sample(ix,np*missfrac,replace=FALSE)
xna=x
xna[imiss]=NA
### uses regular matrix method for matrices with NAs
fit1=softImpute(xna,rank=50,lambda=30)
### uses sparse matrix method for matrices of class "Incomplete"
xnaC=as(xna,"Incomplete")
fit2=softImpute(xnaC,rank=50,lambda=30)
### uses "svd" algorithm
fit3=softImpute(xnaC,rank=50,lambda=30,type="svd")
imp=complete(xna,fit1)
### first scale xna
xnas=biScale(xna)
fit4=softImpute(xnas,rank=50,lambda=10)
imp=complete(xna,fit4)
imp=imp[fit4,1:3,7,2:5,10])
imp=imp[fit4,1:3,7,2:5,10,unscale=FALSE]#ignore scaling and centering
```

---

**SparseplusLowRank-class**

*Class* "SparseplusLowRank"

---

**Description**

A structured matrix made up of a sparse part plus a low-rank part, all which can be stored and operated on efficiently.

**Objects from the Class**

Objects can be created by calls of the form `new("SparseplusLowRank", ...)` or by a call to `splr`

**Slots**

- `x`: Object of class "sparseMatrix"
- `a`: Object of class "matrix"
- `b`: Object of class "matrix"
Methods

%*% signature(x = "ANY", y = "SparseplusLowRank"): ...
%*% signature(x = "SparseplusLowRank", y = "ANY"): ...

as.matrix signature(x = "SparseplusLowRank"): ...
colMeans signature(x = "SparseplusLowRank"): ...
colSums signature(x = "SparseplusLowRank"): ...
dim signature(x = "SparseplusLowRank"): ...
norm signature(x = "SparseplusLowRank", type = "character"): ...
rowMeans signature(x = "SparseplusLowRank"): ...
rowSums signature(x = "SparseplusLowRank"): ...
svd.als signature(x = "SparseplusLowRank"): ...

Author(s)

Trevor Hastie and Rahul Mazumder

See Also

softImpute.splr

Examples

showClass("SparseplusLowRank")
x=matrix(sample(c(0,1),15,replace=TRUE),5,3)
x=as(x, "sparseMatrix")
a=matrix(rnorm(10),5,2)
b=matrix(rnorm(6),3,2)
new("SparseplusLowRank", x=x, a=a, b=b)
splr(x,a,b)

splr

create a SparseplusLowRank object

Description

create an object of class SparseplusLowRank which can be efficiently stored and for which efficient linear algebra operations are possible.

Usage

splr(x, a = NULL, b = NULL)
svd.als

Arguments

x sparse matrix with dimension say m x n
a matrix with m rows and number of columns r less than \( \min(\text{dim}(x)) \)
b matrix with n rows and number of columns r less than \( \min(\text{dim}(x)) \)

Value

an object of S4 class \( \text{SparseplusLowRank} \) is returned with slots x, a and b

Author(s)

Trevor Hastie

See Also

\( \text{SparseplusLowRank-class}, \text{softImpute} \)

Examples

```r
x = matrix(sample(c(3, 0), 15, replace=TRUE), 5, 3)
x = as(x, "sparseMatrix")
a = matrix(rnorm(10), 5, 2)
b = matrix(rnorm(6), 3, 2)
new("SparseplusLowRank", x=x, a=a, b=b)
 splr(x, a, b)
```

svd.als compute a low rank soft-thresholded svd by alternating orthogonal ridge regression

Description

fit a low-rank svd to a complete matrix by alternating orthogonal ridge regression. Special sparse-matrix classes available for very large matrices, including "SparseplusLowRank" versions for row and column centered sparse matrices.

Usage

```r
svd.als(x, rank.max = 2, lambda = 0, thresh = 1e-05, maxit = 100, trace.it = FALSE, warm.start = NULL,

Arguments

x An m by n matrix. Large matrices can be in "sparseMatrix" format, as well as "SparseplusLowRank". The latter arise after centering sparse matrices, for example with biScale, as well as in applications such as softImpute.
rank.max The maximum rank for the solution. This is also the dimension of the left and right matrices of orthogonal singular vectors.
lambda The regularization parameter. \(\lambda = 0\) corresponds to an accelerated version of the orthogonal QR-algorithm. With \(\lambda > 0\) the algorithm amounts to alternating orthogonal ridge regression.

thresh convergence threshold, measured as the relative change in the Frobenius norm between two successive estimates.

maxit maximum number of iterations.

trace.it with \texttt{trace.it=TRUE}, convergence progress is reported.

warm.start an svd object can be supplied as a warm start. If the solution requested has higher rank than the warm start, the additional subspace is initialized with random Gaussians (and then orthogonalized wrt the rest).

final.svd Although in theory, this algorithm converges to the solution to a nuclear-norm regularized low-rank matrix approximation problem, with potentially some singular values equal to zero, in practice only near-zeros are achieved. This final step does one more iteration with \(\lambda = 0\), followed by soft-thresholding.

Details

This algorithm solves the problem

\[
\min \| X - M \|_F^2 + \lambda \| M \|_*
\]

subject to \(\text{rank}(M) \leq r\), where \(\| M \|_* \) is the nuclear norm of \(M\) (sum of singular values). It achieves this by solving the related problem

\[
\min \| X - AB' \|_F^2 + \lambda/2(\| A \|_F^2 + \| B \|_F^2)
\]

subject to \(\text{rank}(A) = \text{rank}(B) \leq r\). The solution is a rank-restricted, soft-thresholded SVD of \(X\).

Value

An svd object is returned, with components "u", "d", and "v".

u an \(m\) by \(\text{rank.max}\) matrix with the left orthogonal singular vectors

d a vector of length \(\text{rank.max}\) of soft-thresholded singular values

v an \(n\) by \(\text{rank.max}\) matrix with the right orthogonal singular vectors

Author(s)

Trevor Hastie, Rahul Mazumder
Maintainer: Trevor Hastie <hastie@stanford.edu>

References


svd.als

See Also

biScale, softImpute, Incomplete, lambda0, impute.complete

Examples

```r
# create a matrix and run the algorithm
set.seed(101)
n=100
p=50
J=25
np=n*p
x=matrix(rnorm(n*J), n,J)*matrix(rnorm(J*p), J,p)*matrix(rnorm(np),n,p)/5
fit=svd.als(x, rank=25, lambda=50)
fit$d

pmax(svd(x)$d-50, 0)

# now create a sparse matrix and do the same
nnz=trunc(np*.3)
inz=sample(seq(np), nnz, replace=FALSE)
i=row(x)[inz]
j=col(x)[inz]
x=rnorm(nnz)
xS=sparseMatrix(x=x, i=i, j=j)
fit2=svd.als(xS, rank=20, lambda=7)
fit2$d

pmax(svd(as.matrix(xS))$d-7, 0)
```
Index

*Topic array
  biScale, 2
  complete, 4
  Incomplete, 5
  lambda0, 7
  softImpute, 9
  svd.als, 13
*Topic classes
  Incomplete-class, 6
  SparseplusLowRank-class, 11
  splr, 12
*Topic models
  biScale, 2
  complete, 4
  Incomplete, 5
  lambda0, 7
  softImpute, 9
  splr, 12
  svd.als, 13
*Topic multivariate
  biScale, 2
  complete, 4
  Incomplete, 5
  lambda0, 7
  softImpute, 9
  splr, 12
  svd.als, 13
  %*,ANY,SparseplusLowRank-method
    (SparseplusLowRank-class), 11
  %*,SparseplusLowRank,ANY-method
    (SparseplusLowRank-class), 11
  as.matrix,Incomplete-method
    (Incomplete-class), 6
  as.matrix,SparseplusLowRank-method
    (SparseplusLowRank-class), 11
biScale, 2
coerce,matrix,Incomplete-method
  (Incomplete-class), 5
coerce,sparseMatrix,Incomplete-method
  (Incomplete-class), 6
colMeans,SparseplusLowRank-method
  (SparseplusLowRank-class), 11
colSums,SparseplusLowRank-method
  (SparseplusLowRank-class), 11
complete, 4
complete,Incomplete-method(complete), 4
complete,matrix-method(complete), 4
compMatrix, 6
cSparseMatrix, 6
dgCMatrix, 6
dim,SparseplusLowRank-method
  (SparseplusLowRank-class), 11
dMatrix, 6
dsparseMatrix, 6
genralMatrix, 6
impute(complete), 4
Incomplete, 5
Incomplete-class, 6
lambda0, 7
lambda0,Incomplete-method(lambda0), 7
lambda0,sparseMatrix-method(lambda0), 7
lambda0,SparseplusLowRank-method
  (lambda0), 7
Matrix, 6
norm,SparseplusLowRank,character-method
  (SparseplusLowRank-class), 11
rowMeans,SparseplusLowRank-method
  (SparseplusLowRank-class), 11
rowSums,SparseplusLowRank-method
  (SparseplusLowRank-class), 11
softImpute, 9
sparseMatrix, 6
SparseplusLowRank-class, 11
splr, 12
svd.als, 13
svd.als,sparseMatrix-method(svd.als), 13
svd.als,SparseplusLowRank-method
 (svd.als), 13