Input

Input

Input

Input

Principal Component Analysis: nagdmc_pca_score

Purpose

nagdmc_pca_score computes principal component scores for a data record following a principal component analysis (PCA).

Declaration

```
#include <nagdmc.h>
void nagdmc_pca_score(long nvar, long nscores, long std, double data[],
                     double xbar[], double s[], double loadings[],
                     double results[], double scores[], int *info);
```

Parameters

1: nvar - long Input On entry: the number of variables in the data. Constraint: nvar > 1. 2: nscores - long Input

On entry: the number of scores to be computed.

Constraint: 0 < nscores < nvar.

std - int3:

On entry: indicates how the scores are to be standardised. If std = 1, the scores are standardised to have unit variance; otherwise std must be set equal to zero, and the scores are standardised to have variance equal to the corresponding eigenvalue.

Constraint: $\mathbf{std} \in \{0, 1\}$.

4: data[nvar] - double

On entry: the data record for which to compute PCA scores.

xbar[nvar] - double 5:

> On entry: the vector of mean values of the variables for the data used by nagdmc_pca to compute the PCA.

s[nvar] - double 6:

On entry: the vector of standard deviations or scaling factors. If pcatype was set to 2 in **nagdmc_pca**, the standard deviations must be provided in s; else if **pcatype** was set to 3, s must contain the user-supplied standardisations; otherwise \mathbf{s} is not referenced and must be set to 0. Constraints: if referenced, $\mathbf{s}[i] > 0.0$, for $i = 0, 1, \dots, \mathbf{nvar} - 1$.

```
7:
     loadings[nvar*nvar] - double
                                                                                                Input
     On entry: the loadings from the principal component analysis as returned by nagdmc_pca.
```

8:	${f results}[6*nvar]-{f double}$	Input
	On entry: the variance decomposition results as returned by nagdmc_pca .	
9:	scores[nscores] - double	Output

On exit: the computed PCA scores for the data record, data.

10:info - int * Output

On exit: info gives information on the success of the function call:

0: the function successfully completed its task.

i; i = i = 1, 2, 3, 6: the specification of the *i*th formal parameter was incorrect.

99: the function failed to allocate enough memory.

Notation

data	the data record, x .
loadings	the array of PCA loadings, a .
nscores	the number of principal component scores to compute, l .
nvar	the number of variables in the data, p .
scores	the PCA scores for x, f_k , for $k = 1, 2, \ldots, l$.

Description

Given an array a of loadings following a PCA, the kth principal component score of a data record x is given by:

$$f_k = \sum_{j=1}^p a_{kj} x_j, \quad k = 1, 2, \dots, l,$$

where p is the number of variables in the data, and $l \leq p$ is the number of scores to calculate. Scores can be standardised so that the variance of the scores for each principal component is 1.0 or equal to the corresponding eigenvalue.

References and Further Reading

Chatfield C and Collins A J (1980) Introduction to Multivariate Analysis Chapman and Hall.

Krzanowski W J (1990) Principles of Multivariate Analysis Oxford University Press.

See Also

computes a PCA on data values. nagdmc_pca the example calling program. pca_ex.c