

Utility: nagdmc_scale

Purpose

nagdmc_scale scales continuous data values to have zero mean and unit variance.

Declaration

```
#include <nagdmc.h>

void nagdmc_scale(long rec1, long nvar, long nrec, long dblk, double data[],
                  long nxvar, long xvar[], int *info);
```

Parameters

- 1: **rec1** – long *Input*
On entry: the index in the data of the first data record used in the analysis.
Constraint: **rec1** ≥ 0 .
- 2: **nvar** – long *Input*
On entry: the number of variables in the data.
Constraint: **nvar** ≥ 1 .
- 3: **nrec** – long *Input*
On entry: the number of consecutive records, beginning at **rec1**, used in the analysis.
Constraint: **nrec** > 1 .
- 4: **dblk** – long *Input*
On entry: the total number of records in the data block.
Constraint: **dblk** $\geq \text{rec1} + \text{nrec}$.
- 5: **data**[**dblk** * **nvar**] – double *Input/Output*
On entry: the (clean) data values for the j th variable (for $j = 0, 1, \dots, \text{nvar} - 1$) are stored in **data**[$i * \text{nvar} + j$], for $i = 0, 1, \dots, \text{dblk} - 1$.
On exit: the scaled data values.
- 6: **nxvar** – long *Input*
On entry: the number of variables in the analysis. If **nxvar** = 0, all variables in the data are used in the analysis.
Constraint: $0 \leq \text{nxvar} \leq \text{nvar}$.
- 7: **xvar**[**nxvar**] – long *Input*
On entry: the indices indicating the position in **data** in which the variables are stored. If **nxvar** = 0 then **xvar** must be 0, and the indices of variables are given by $j = 0, 1, \dots, \text{nvar} - 1$.
Constraints: if **nxvar** > 0 , $0 \leq \text{xvar}[i] < \text{nvar}$, for $i = 0, 1, \dots, \text{nxvar} - 1$; otherwise **xvar** must be 0.
- 8: **info** – int * *Output*
On exit: **info** gives information on the success of the function call:
 - 0: the function successfully completed its task.
 - i ; $i = 1, 2, 3, 4, 6, 7$: the specification of the i th formal parameter was incorrect.
 - 99: the function failed to allocate enough memory.

Notation

nrec number of data records n .
data set of data records X and, on output, the scaled data Z .
nxvar determines the number of variables in X to scale, p .

Description

Let X be a set of n data records x_i , for $i = 1, 2, \dots, n$, and x_{ij} be the value of the j th variable for the i th data record, for $j = 1, 2, \dots, p$. Furthermore, let μ_j be the mean value of the j th variable in X and σ_j denote the standard deviation about this mean value.

The scaled data set Z has the elements:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}, \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, p.$$

The elements in column j of Z have zero mean and unit variance.

References and Further Reading

None.

See Also

None.
