NAG Toolbox for MATLAB

g02cf

1 Purpose

g02cf reorders the elements in two vectors (typically vectors of means and standard deviations), and the rows and columns in two matrices (typically either matrices of sums of squares and cross-products of deviations from means and Pearson product-moment correlation coefficients, or matrices of sums of squares and cross-products about zero and correlation-like coefficients).

2 Syntax

```
[xbar, std, ssp, r, ifail] = g02cf(korder, xbar, std, ssp, r, 'n', n)
```

3 Description

Input to the function consists of:

(a) A list of the order in which the n variables are to be arranged on exit:

$$i_1, i_2, i_3, \ldots, i_n.$$

(b) A vector of means:

$$(\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n).$$

(c) A vector of standard deviations:

$$(s_1, s_2, s_3, \ldots, s_n).$$

(d) A matrix of sums of squares and cross-products of deviations from means:

$$\begin{pmatrix} S_{11} & S_{12} & S_{13} & . & . & . & S_{1n} \\ S_{21} & S_{22} & & & & & . \\ S_{31} & & & & & & . \\ . & & & & & & . \\ . & & & & & & . \\ . & & & & & & . \\ S_{n1} & S_{n2} & . & & . & . & S_{nn} \end{pmatrix}.$$

(e) A matrix of correlation coefficients:

$$\begin{pmatrix} R_{11} & R_{12} & R_{13} & . & . & . & R_{1n} \\ R_{21} & R_{22} & & & & & . \\ R_{31} & & & & & & . \\ . & & & & & & . \\ . & & & & & & . \\ . & & & & & & . \\ R_{n1} & R_{n2} & . & & . & . & R_{nn} \end{pmatrix}.$$

On exit from the function, these same vectors and matrices are reordered, in the manner specified, and contain the following information:

(i) The vector of means:

$$\left(\bar{x}_{i_1}, \bar{x}_{i_2}, \bar{x}_{i_3}, \ldots, \bar{x}_{i_n}\right).$$

(ii) The vector of standard deviations:

$$\left(s_{i_1}, s_{i_2}, s_{i_3}, \ldots s_{i_n}\right).$$

(iii) The matrix of sums of squares and cross-products of deviations from means:

$$\begin{pmatrix} S_{i_1i_1} & S_{i_1i_2} & S_{i_1i_3} & . & . & . & S_{i_1i_n} \\ S_{i_2i_1} & S_{i_2i_2} & & & & & . \\ S_{i_3i_1} & & & & & & . \\ . & & & & & & . \\ . & & & & & & . \\ . & & & & & & . \\ S_{i_ni_1} & S_{i_ni_2} & . & & . & . & S_{i_ni_n} \end{pmatrix}.$$

(iv) The matrix of correlation coefficients:

$$\begin{pmatrix} R_{i_1i_1} & R_{i_1i_2} & R_{i_1i_3} & . & . & . & R_{i_1i_n} \\ R_{i_2i_1} & R_{i_2i_2} & & & & & . \\ R_{i_3i_1} & & & & & & . \\ . & & & & & & . \\ . & & & & & & . \\ . & & & & & & . \\ R_{i_ni_1} & R_{i_ni_2} & . & & . & . & R_{i_ni_n} \end{pmatrix}.$$

Note: for sums of squares of cross-products of deviations about zero and correlation-like coefficients S_{ij} and R_{ij} should be replaced by $\tilde{S}_{ij}$ and $\tilde{R}_{ij}$ in the description of the input and output above.

4 References

None.

5 Parameters

5.1 Compulsory Input Parameters

1: **korder**(**n**) – **int32 array**

korder(i) must be set to the number of the original variable which is to be the ith variable in the re-arranged data, for $i = 1, 2, \ldots, n$.

Constraint: $1 \leq$ **korder**(i) $\leq$ **n**, for $i = 1, 2, \ldots, n$.

2: **xbar**(**n**) – **double array**

xbar(i) must be set to the mean of variable i, for $i = 1, 2, \ldots, n$.

3: **std**(**n**) – **double array**

std(i) must be set to the standard deviation of variable i, for $i = 1, 2, \ldots, n$.

4: **ssp**(**ldssp**,**n**) – **double array**

ldssp, the first dimension of the array, must be at least **n**.

ssp(i,j) must be set to the sum of cross-products of deviations from means S_{ij} (or about zero $\tilde{S}_{ij}$) for variables i and j, for $i,j = 1, 2, \ldots, n$.

5: **r**(**ldr**,**n**) – **double array**

ldr, the first dimension of the array, must be at least **n**.

r(i,j) must be set to the Pearson product-moment correlation coefficient R_{ij} (or the correlation-like coefficient $\tilde{R}_{ij}$) for variables i and j, for $i,j = 1, 2, \ldots, n$.

5.2 Optional Input Parameters

1: **n – int32 scalar**

Default: The dimension of the arrays **xbar**, **std**, **ssp**, **r**. (An error is raised if these dimensions are not equal.)

n, the number of variables in the input data.

Constraint: $\mathbf{n} \geq 2$.

5.3 Input Parameters Omitted from the MATLAB Interface

ldssp, ldr, kwork

5.4 Output Parameters

1: **xbar(n) – double array**

xbar(i) contains the mean of variable k where $k = \mathbf{korder}(i)$, for $i = 1, 2, \ldots, n$.

2: **std(n) – double array**

std(i) contains the standard deviation of variable k where $k = \mathbf{korder}(i)$, for $i = 1, 2, \ldots, n$.

3: **ssp(ldssp,n) – double array**

ssp(i,j) contains the sum of cross-products of deviations from means S_{kl} (or about zero $\tilde{S}_{kl}$) for variables k and l, where $k = \mathbf{korder}(i)$, and $l = \mathbf{korder}(j)$, $i,j = 1, 2, \ldots, n$.

4: **r(ldr,n) – double array**

r(i,j) contains the Pearson product-moment correlation coefficient R_{kl} (or the correlation-like coefficient $\tilde{R}_{kl}$) for variables k and l, where $k = \mathbf{korder}(i)$ and $l = \mathbf{korder}(j)$, for $i,j = 1, 2, \ldots, n$.

5: **ifail – int32 scalar**

0 unless the function detects an error (see Section 6).

6 Error Indicators and Warnings

Errors or warnings detected by the function:

ifail $= 1$

On entry, $\mathbf{n} < 2$.

ifail $= 2$

On entry, **ldssp** $<$ **n**,
or **ldr** $<$ **n**.

ifail $= 3$

On entry, $\mathbf{korder}(i) < 1$,
or $\mathbf{korder}(i) > \mathbf{n}$ for some $i = 1, 2, \ldots, n$.

ifail $= 4$

On entry, there is not a one-to-one correspondence between the old variables and the new variables; at least one of the original variables is not included in the new set, and consequently at least one other variable has been included more than once.

7 Accuracy

Not applicable.

8 Further Comments

The time taken by g02cf depends on n and the amount of re-arrangement involved.

The function is intended primarily for use when a set of variables is to be reordered for use in a regression, and is described accordingly. There is however no reason why the function should not also be used to reorder vectors and matrices which contain any other non-statistical information; the matrices need not be symmetric.

The function may be used either with sums of squares and cross-products of deviations from means and Pearson product-moment correlation coefficients in connection with a regression involving a constant, or with sums of squares and cross-products about zero and correlation-like coefficients in connection with a regression with no constant.

9 Example

```
korder = [int32(1);
     int32(3);
     int32(2)];
xbar = [5.4;
     5.8;
     2.8];
std = [4.98;
     5.0695;
     1.924];
ssp = [99.2, -57.6, 6.4;
     -57.6, 102.8, -29.2;
     6.4, -29.2, 14.8];
r = [1, -0.5704, 0.167;
     -0.5704, 1, -0.7486;
     0.167, -0.7486, 1];
[xbarOut, stdOut, sspOut, rOut, ifail] = g02cf(korder, xbar, std, ssp, r)
```

```
xbarOut =
    5.4000
    2.8000
    5.8000
stdOut =
    4.9800
    1.9240
    5.0695
sspOut =
   99.2000    6.4000  -57.6000
    6.4000   14.8000  -29.2000
  -57.6000  -29.2000  102.8000
rOut =
    1.0000    0.1670   -0.5704
    0.1670    1.0000   -0.7486
   -0.5704   -0.7486    1.0000
ifail =
           0
```