NAG Toolbox

nag_correg_linregm_estfunc (g02dn)

1 Purpose

nag_correg_linregm_estfunc (g02dn) gives the estimate of an estimable function along with its standard error.

2 Syntax

```
[est, stat, sestat, t, ifail] = nag_correg_linregm_estfunc(irank, b, cov, p, f, tol, 'ip', ip)
[est, stat, sestat, t, ifail] = g02dn(irank, b, cov, p, f, tol, 'ip', ip)
```

3 Description

nag_correg_linregm_estfunc (g02dn) computes the estimates of an estimable function for a general linear regression model which is not of full rank. It is intended for use after a call to nag_correg_linregm_fit (g02da) or nag_correg_linregm_update (g02dd). An estimable function is a linear combination of the arguments such that it has a unique estimate. For a full rank model all linear combinations of arguments are estimable.

In the case of a model not of full rank the functions use a singular value decomposition (SVD) to find the parameter estimates, $\hat{\beta}$, and their variance-covariance matrix. Given the upper triangular matrix R obtained from the QR decomposition of the independent variables the SVD gives

$$R = Q_* \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} P^{\mathrm{T}},$$

where D is a k by k diagonal matrix with nonzero diagonal elements, k being the rank of R, and Q_* and P are p by p orthogonal matrices. This gives the solution

$$\hat{\beta} = P_1 D^{-1} Q_{*_1}^{\mathrm{T}} c_1,$$

P_1 being the first k columns of P, i.e., $P = (P_1 P_0)$, Q_{*_1} being the first k columns of Q_*, and c_1 being the first p elements of c.

Details of the SVD are made available in the form of the matrix P^*:

$$P^* = \begin{pmatrix} D^{-1} P_1^{\mathrm{T}} \\ P_0^{\mathrm{T}} \end{pmatrix},$$

as given by nag_correg_linregm_fit (g02da) and nag_correg_linregm_update (g02dd).

A linear function of the arguments, $F = f^{\mathrm{T}}\beta$, can be tested to see if it is estimable by computing $\zeta = P_0^{\mathrm{T}} f$. If ζ is zero, then the function is estimable; if not, the function is not estimable. In practice $|\zeta|$ is tested against some small quantity η.

Given that F is estimable it can be estimated by $f^{\mathrm{T}}\hat{\beta}$ and its standard error calculated from the variance-covariance matrix of $\hat{\beta}$, C_β, as

$$\mathrm{se}(F) = \sqrt{f^{\mathrm{T}} C_\beta f}.$$

Also a t-statistic,

$$t = \frac{f^{\mathrm{T}}\hat{\beta}}{\mathrm{se}(F)},$$

can be computed. The t-statistic will have a Student's t-distribution with degrees of freedom as given by the degrees of freedom for the residual sum of squares for the model.

4 References

Golub G H and Van Loan C F (1996) *Matrix Computations* (3rd Edition) Johns Hopkins University Press, Baltimore

Hammarling S (1985) The singular value decomposition in multivariate statistics *SIGNUM Newsl.* **20(3)** 2–25

Searle S R (1971) *Linear Models* Wiley

5 Parameters

5.1 Compulsory Input Parameters

1: **irank** – INTEGER

k, the rank of the independent variables.

Constraint: $1 \leq$ **irank** $\leq$ **ip**.

2: **b**(**ip**) – REAL (KIND=nag_wp) array

The **ip** values of the estimates of the arguments of the model, $\hat{\beta}$.

3: **covar**(**ip** $\times$ (**ip** + **1**)/**2**) – REAL (KIND=nag_wp) array

The upper triangular part of the variance-covariance matrix of the **ip** parameter estimates given in **b**. They are stored packed by column, i.e., the covariance between the parameter estimate given in $\mathbf{b}(i)$ and the parameter estimate given in $\mathbf{b}(j)$, $j \geq i$, is stored in $\mathbf{cov}((j \times (j-1)/2 + i))$.

4: **p**(**ip** $\times$ **ip** + **2** $\times$ **ip**) – REAL (KIND=nag_wp) array

As returned by nag_correg_linregm_fit (g02da) and nag_correg_linregm_update (g02dd).

5: **f**(**ip**) – REAL (KIND=nag_wp) array

f, the linear function to be estimated.

6: **tol** – REAL (KIND=nag_wp)

η, the tolerance value used in the check for estimability.

tol ≤ 0.0
$\sqrt{\epsilon}$, where ϵ is the ***machine precision***, is used instead.

5.2 Optional Input Parameters

1: **ip** – INTEGER

Default: the dimension of the arrays **b**, **f**. (An error is raised if these dimensions are not equal.)

p, the number of terms in the linear model.

Constraint: **ip** ≥ 1.

5.3 Output Parameters

1: **est** – LOGICAL

Indicates if the function was estimable.

est $= true$
The function is estimable.

est $= false$
The function is not estimable and **stat**, **sestat** and **t** are not set.

2: **stat** – REAL (KIND=nag_wp)

If **est** $= true$, **stat** contains the estimate of the function, $f^{\mathrm{T}}\hat{\beta}$.

3: **sestat** – REAL (KIND=nag_wp)

If **est** $= true$, **sestat** contains the standard error of the estimate of the function, $\mathrm{se}(F)$.

4: **t** – REAL (KIND=nag_wp)

If **est** $= true$, **t** contains the t-statistic for the test of the function being equal to zero.

5: **ifail** – INTEGER

ifail $= 0$ unless the function detects an error (see Section 5).

6 Error Indicators and Warnings

Note: nag_correg_linregm_estfunc (g02dn) may return useful information for one or more of the following detected errors or warnings.

Errors or warnings detected by the function:

ifail $= 1$

On entry, **ip** < 1,
or **irank** < 1,
or **irank** $>$ **ip**.

ifail $= 2$ (*warning*)

On entry, **irank** $=$ **ip**. In this case **est** is returned as true and all statistics are calculated.

ifail $= 3$

Standard error of statistic $= 0.0$; this may be due to rounding errors if the standard error is very small or due to mis-specified inputs **cov** and **f**.

ifail $= -99$

An unexpected error has been triggered by this routine. Please contact NAG.

ifail $= -399$

Your licence key may have expired or may not have been installed correctly.

ifail $= -999$

Dynamic memory allocation failed.

7 Accuracy

The computations are believed to be stable.

8 Further Comments

The value of estimable functions is independent of the solution chosen from the many possible solutions. While nag_correg_linregm_estfunc (g02dn) may be used to estimate functions of the arguments of the model as computed by nag_correg_linregm_constrain (g02dk), β_c, these must be expressed in terms of the original arguments, β. The relation between the two sets of arguments may not be straightforward.

9 Example

Data from an experiment with four treatments and three observations per treatment are read in. A model, with a mean term, is fitted by nag_correg_linregm_fit (g02da). The number of functions to be tested is read in, then the linear functions themselves are read in and tested with nag_correg_linregm_estfunc (g02dn). The results of nag_correg_linregm_estfunc (g02dn) are printed.

9.1 Program Text

```
     function g02dn_example

fprintf('g02dn example results\n\n');

x = [1, 0, 0, 0;
     0, 0, 0, 1;
     0, 1, 0, 0;
     0, 0, 1, 0;
     0, 0, 0, 1;
     0, 1, 0, 0;
     0, 0, 0, 1;
     1, 0, 0, 0;
     0, 0, 1, 0;
     1, 0, 0, 0;
     0, 0, 1, 0;
     0, 1, 0, 0];
y = [33.63;     39.62;     38.18;     41.46;     38.02;     35.83;
     35.99;     36.58;     42.92;     37.80;     40.43;     37.89];

[n,m]  = size(x);
isx    = ones(m,1,nag_int_name);
mean_p = 'M';
ip     = nag_int(m+1);

% Fit general linear regression model
[rss, idf, b, se, covar, res, h, q, svd, irank, p, wk, ifail] = ...
  g02da(mean_p, x, isx, ip, y);

% Display initial results
fprintf('Estimates from g02da\n\n');
fprintf('Residual sum of squares = %12.4e\n', rss);
fprintf('Degrees of freedom      = %4d\n', idf);
fprintf('\nVariable   Parameter estimate   Standard error\n\n');
ivar = double([1:ip]');
fprintf('%6d%20.4e%20.4e\n',[ivar b se]');

% Estimable functions
f = [1 1  0 0 0;
     0 1 -1 0 0;
     0 1  0 0 0];
nf = size(f,1);
tol = 1e-05;

% Loop over estimable functions
for j = 1:nf
  [est, stat, sestat, t, ifail] = ...
  g02dn(irank, b, covar, p, f(j,:), tol);

  % Display results
   fprintf('\nFunction %2d\n\n', j);
```

```
   fprintf('%8.2f', f(j,:));
   if est
     fprintf('\n\nstat = %10.4f, se = %10.4f, t = %10.4f\n', stat, sestat, t);
   else
     fprintf('\n\nFunction not estimable\n');
   end
end
```

9.2 Program Results

```
     g02dn example results

Estimates from g02da

Residual sum of squares =   2.2227e+01
Degrees of freedom      =    8

Variable   Parameter estimate   Standard error

     1          3.0557e+01          3.8494e-01
     2          5.4467e+00          8.3896e-01
     3          6.7433e+00          8.3896e-01
     4          1.1047e+01          8.3896e-01
     5          7.3200e+00          8.3896e-01

Function  1

    1.00    1.00    0.00    0.00    0.00

stat =    36.0033, se =     0.9623, t =    37.4119

Function  2

    0.00    1.00   -1.00    0.00    0.00

stat =    -1.2967, se =     1.3610, t =    -0.9528

Function  3

    0.00    1.00    0.00    0.00    0.00

Function not estimable
```