NAG Toolbox for MATLAB

g01hb

1 Purpose

g01hb returns the upper tail, lower tail or central probability associated with a multivariate Normal distribution of up to ten dimensions.

2 Syntax

```
[result, ifail] = g01hb(tail, a, b, xmu, sig, 'n', n, 'tol', tol)
```

3 Description

Let the vector random variable $X = (X_1, X_2, \ldots, X_n)^{\mathrm{T}}$ follow an n-dimensional multivariate Normal distribution with mean vector μ and n by n variance-covariance matrix Σ, then the probability density function, $f(X : \mu, \Sigma)$, is given by

$$f(X : \mu, \Sigma) = (2\pi)^{-(1/2)n} |\Sigma|^{-1/2} \exp\left(-\tfrac{1}{2}(X - \mu)^{\mathrm{T}} \Sigma^{-1} (X - \mu)\right).$$

The lower tail probability is defined by:

$$P(X_1 \leq b_1, \ldots, X_n \leq b_n : \mu, \Sigma) = \int_{-\infty}^{b_1} \cdots \int_{-\infty}^{b_n} f(X : \mu, \Sigma)\, dX_n \cdots dX_1.$$

The upper tail probability is defined by:

$$P(X_1 \geq a_1, \ldots, X_n \geq a_n : \mu, \Sigma) = \int_{a_1}^{\infty} \cdots \int_{a_n}^{\infty} f(X : \mu, \Sigma)\, dX_n \cdots dX_1.$$

The central probability is defined by:

$$P(a_1 \leq X_1 \leq b_1, \ldots, a_n \leq X_n \leq b_n : \mu, \Sigma) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(X : \mu, \Sigma)\, dX_n \cdots dX_1.$$

To evaluate the probability for $n \geq 3$, the probability density function of $X_1, X_2, \ldots, X_n$ is considered as the product of the conditional probability of $X_1, X_2, \ldots, X_{n-2}$ given X_{n-1} and X_n and the marginal bivariate Normal distribution of X_{n-1} and X_n. The bivariate Normal probability can be evaluated as described in g01ha and numerical integration is then used over the remaining $n - 2$ dimensions. In the case of $n = 3$, d01aj is used and for $n > 3$ d01fc is used.

To evaluate the probability for $n = 1$ a direct call to g01ea is made and for $n = 2$ calls to g01ha are made.

4 References

Kendall M G and Stuart A 1969 *The Advanced Theory of Statistics (Volume 1)* (3rd Edition) Griffin

5 Parameters

5.1 Compulsory Input Parameters

1: **tail – string**

Indicates which probability is to be returned.

tail = 'L'

The lower tail probability is returned.

tail = 'U'

The upper tail probability is returned.

tail = 'C'

The central probability is returned.

Constraint: **tail** = 'L', 'U' or 'C'.

2: **a(n) – double array**

If **tail** = 'C' or 'U', the lower bounds, a_i, for $i = 1, 2, \ldots, n$.

If **tail** = 'L', **a** is not referenced.

3: **b(n) – double array**

If **tail** = 'C' or 'L', the upper bounds, b_i, for $i = 1, 2, \ldots, n$.

If **tail** = 'U' **b** is not referenced.

Constraint: $\mathbf{a}(i) < \mathbf{b}(i)$ if **tail** = 'C', for $i = 1, 2, \ldots, n$.

4: **xmu(n) – double array**

μ, the mean vector of the multivariate Normal distribution.

5: **sig(ldsig,n) – double array**

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

Σ, the variance-covariance matrix of the multivariate Normal distribution. Only the lower triangle is referenced.

Constraint: Σ must be positive-definite.

5.2 Optional Input Parameters

1: **n – int32 scalar**

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

n, the number of dimensions.

Constraint: $1 \leq \mathbf{n} \leq 10$.

2: **tol – double scalar**

If $n > 2$ the relative accuracy required for the probability, and if the upper or the lower tail probability is requested then **tol** is also used to determine the cut-off points, see Section 7.

If $n = 1$, **tol** is not referenced.

Suggested value: **tol** = 0.0001.

Default: 0.0001

Constraint: if **tol** > 0.0, **n** > 1.

5.3 Input Parameters Omitted from the MATLAB Interface

ldsig, wk, lwk

5.4 Output Parameters

1: **result – double scalar**

The result of the function.

2: **ifail – int32 scalar**

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

6 Error Indicators and Warnings

Note: g01hb may return useful information for one or more of the following detected errors or warnings.

If on exit **ifail** = 1, 2 or 3, then g01hb returns zero.

ifail = 1

On entry, **n** < 1,
or **n** > 10,
or **ldsig** < **n**,
or **tail** ≠ 'L', 'U' or 'C',
or **n** > 1 and **tol** ≤ 0.0,
or **lwk** is too small.

ifail = 2

On entry, **tail** = 'C' and **a**$(i) \geq$ **b**(i), for some $i = 1, 2, \ldots, n$.

ifail = 3

On entry, Σ is not positive-definite, i.e., is not a correct variance-covariance matrix.

ifail = 4

The requested accuracy has not been achieved, a larger value of **tol** should be tried or the length of the workspace should be increased. The returned value will be an approximation to the required result.

ifail = 5

Round-off error prevents the requested accuracy from being achieved; a larger value of **tol** should be tried. The returned value will be an approximation to the required result. This error will only occur if $n = 3$.

7 Accuracy

The accuracy should be as specified by **tol**. When on exit **ifail** = 4 the approximate accuracy achieved is given in the error message. For the upper and lower tail probabilities the infinite limits are approximated by cut-off points for the $n - 2$ dimensions over which the numerical integration takes place; these cut-off points are given by $\Phi^{-1}(\textbf{tol}/(10 \times n))$, where Φ^{-1} is the inverse univariate Normal distribution function.

8 Further Comments

The time taken is related to the number of dimensions, the range over which the integration takes place ($b_i - a_i$, for $i = 1, 2, \ldots, n$) and the value of Σ as well as the accuracy required. As the numerical integration does not take place over the last two dimensions speed may be improved by arranging X so that the largest ranges of integration are for X_{n-1} and X_n.

9 Example

```
tail = 'c';
a = [-2;
     -2;
     -2;
     -2];
b = [2;
     2;
     2;
     2];
xmu = [0;
     0;
     0;
     0];
sig = [1, 0.9, 0.9, 0.9;
     0.9, 1, 0.9, 0.9;
     0.9, 0.9, 1, 0.9;
     0.9, 0.9, 0.9, 1];
[result, ifail] = g01hb(tail, a, b, xmu, sig)
```

```
result =
    0.9142
ifail =
           0
```