An efficient numerical solution to the multinormal integral

S. Gollwitzer

Bogenstr.9, 8000 Munich 80, FRG

R. Rackwitz

Technical University of Munich, Arcisstr. 21, 8000 Munich 2, FRG

The multinormal integral which has important applications in structural reliability is evaluated by application of a recent, asymptotic formula for the probability content of general intersections in the standard normal space together with a recursive scheme proposed earlier by Hohenbichler. It is suitable for arbitrary correlations, computationally efficient even in higher dimensions and most accurate for small or large probabilities.

Key Words: multinormal probabilities, asymptotic analysis, structural reliability.

INTRODUCTION

The numerical computation of the multinormal integral is required in many problems of statistics and, recently, turned out to be a key problem in structural and operational system reliability[1]. If the problem is appropriately formulated the multinormal integral shows up whenever the probability of an intersection of events is to be determined. For example, an event tree type analysis of redundant systems requires the evaluation of the probability of the intersection of the dependent componental state change (failure) events along a given path to structural collapse[2]. Also, consideration of quality control or inspection results in structural reliability is intimately related to the evaluation of the probability of intersections of events[3,4].

Let $\mathbf{U}=(U_1,\ldots,U_n)^T$ be an independent standard normal vector and a domain D be given as $D=\bigcap_{j=1}^{m}\{g_j(\mathbf{U})\leqslant 0\}$ with $g_j(\mathbf{0})>0$ for at least one j and the functions g_j at least piecewise twice differentiable. If a random vector $\mathbf{X}$ with continuous joint distribution function is neither normal nor independent a probability preserving transformation $\mathbf{X}=\mathbf{T}(\mathbf{U})$ always exists such that the problem at hand can be cast into the form just described[5].

An asymptotic approximation for the probability content of D has been derived in Refs 6 and 7 which is given here for further reference:

$$P(D)\sim\Phi_k(\mathbf{c};\boldsymbol{R})(\det(\boldsymbol{I}-\boldsymbol{H}))^{-1/2}=\Phi_k(\mathbf{c};\boldsymbol{R})C \quad (\|u^*\|\to\infty) \tag{1}$$

Herein, Φ_k is the k-dimensional standard normal integral, $\mathbf{c}=\mathbf{A}^T\mathbf{u}^*$, $\mathbf{R}=\mathbf{A}^T\mathbf{A}$, $C=(\det(\mathbf{I}-\mathbf{H}))^{-1/2}$ a second-order correction term, and $\mathbf{u}^*$ a point defined by

$$\mathbf{u}^*=\min\{\|\mathbf{u}\|\} \qquad \text{for } \left\{\mathbf{u}:\bigcap_{j=1}^{m}g_j(\mathbf{u})\leqslant 0\right\} \tag{2}$$

$\boldsymbol{I}$ is the unit matrix and $\boldsymbol{H}$ a matrix to be defined below. $k(1\leqslant k\leqslant m)$ is the size of the index set J for 'active' constraints, i.e., for which $g_j(\mathbf{u}^*)=0$. $\boldsymbol{A}$ collects as columns the linearly independent, normalized gradients $(\|\mathbf{a}_j\|=\|\text{grad } g_j(\mathbf{u}^*)\|=1)$ of the active constraints. For simplicity of notation, an orthogonal transformation $\mathbf{w}=\boldsymbol{T}^T\mathbf{u}$ is performed such that $w_{k+1}^*=\ldots=w_n^*=0$ and $\mathbf{b}_i=\boldsymbol{T}^T\mathbf{a}_i$; $b_{ij}=0$ for $1\leqslant i\leqslant k$ and $k+1\leqslant j\leqslant m$. The first column of $\boldsymbol{T}$ is formed by any of the $\mathbf{a}_i, i\in J$. The other columns are found by a suitable orthogonalisation algorithm. Further, a negative vector $\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_k)$ to be determined from

$$\mathbf{w}^*=\boldsymbol{B}\boldsymbol{\gamma}; \qquad \boldsymbol{B}=\{\mathbf{b}_i\} \qquad (1\leqslant i\leqslant k) \tag{3}$$

exists. Then, the elements of the matrix $\boldsymbol{H}_{(m-k)\times(m-k)}$ in equation (1) are given by

$$h_{rs}=\sum_{t=1}^{k}\gamma_t\frac{\partial^2 g_t(\mathbf{w}^*)}{\partial w_r\,\partial w_s} \qquad k+1\leqslant r,s\leqslant m \tag{4}$$

For the computation of the multimormal integral in equation (1) in higher dimensions and with arbitrary arguments $\mathbf{c}$ and $\boldsymbol{R}$ no analytical solution exists and only a few practical computation schemes are available. Numerical integration as proposed in Refs 8 or 9 soon becomes computationally inefficient as the numerical effort grows exponentially with the dimension. The primary purpose of approximations like equation (1) also is just to avoid numerical integration. A bounding scheme[10] combining a result for special correlation matrices due to Dunnet/Sobel[11] and an inequality due to Sidak[12] requires relatively little numerical work but is too inaccurate for many reliability applications. In Ref. 1 and further in Ref. 13 computationally more efficient approximations are proposed and shown to be fairly accurate. Finally, Ruben derived an asymptotic

formula[14]

$$\Phi_k(\mathbf{c};\boldsymbol{R}) \sim (\det(\boldsymbol{R}))^{-1/2} \prod_{i=1}^{k} \frac{\varphi(u_i^*)}{(\gamma_i)} \quad (\|u^*\| \to \infty) \tag{5}$$

where the $\mathbf{c}$ must be negative and the γ_i, obtained from $\gamma = \boldsymbol{R}^{-1}\mathbf{c}$ strictly positive. In Ref. 7 equation (5) is shown to correspond to a linear approximation of the functions $g_i(\mathbf{u})=0$ $(1 \leqslant i \leqslant k)$ in $\mathbf{u}^*$.

DERIVATION OF MAIN RESULT

In this paper the multinormal integral is approximated by applying the foregoing general result (1) for the probability content of arbitrary intersections to the multinormal integral in equation (1) on the basis of a recursive scheme suggested in Refs 1 and 13. In other words, equation (1) is applied to itself since any intersection of half-domains in correlated standard normal variables can be represented as an equivalent intersection of half-domains in uncorrelated standard normal variables. This is especially suitable because many operations turn out to be analytic in this case. Let Z be a standard normal vector with correlation matrix $\boldsymbol{R}=(\rho_{ij})$. It can be represented in terms of an independent standard normal vector $\mathbf{U}$ by

$$Z_i = \sum_{j=1}^{i} \alpha_{ij} U_j \qquad (1 \leqslant i \leqslant n) \tag{6}$$

with $\alpha_{11}=1$ and the other coefficients determined such that the left-hand and the right-hand side have the same correlation matrix, e.g., by Cholesky's triangularization procedure for positive definite matrices. The multinormal integral can be rewritten as

$$\begin{aligned} P(\mathbf{Z} \leqslant \mathbf{c}) &= P(Z_1 \leqslant c_1) P\left(\bigcap_{i=2}^{n} (Z_i \leqslant c_i \,|\, Z_1 \leqslant c_1) \right) \\ &= \Phi(c_1) P\left(\bigcap_{i=2}^{n} \{\alpha_{i1}\bar{Z}_1 + \sum_{j=2}^{i} \alpha_{ij} U_j - c_i \leqslant 0\} \right) \\ &= \Phi(c_1) P\left(\bigcap_{i=2}^{n} \{g_i(\mathbf{U}) \leqslant 0\} \right) \end{aligned} \tag{7}$$

where the variable $\bar{Z}_1$ is again standard normal but truncated at c_1 and independent of the other U's. $\bar{Z}_1$ can be represented by its Rosenblatt-transformation[5] as follows:

$$\bar{Z}_1 = \Phi^{-1}[\Phi(c_1)\Phi(U_1)] \tag{8}$$

Hence, the g-functions in equation (7) are:

$$\begin{aligned} g_i(\mathbf{U}) &= \alpha_{i1}\Phi^{-1}[\Phi(c_1)\Phi(U_1)] + \sum_{j=2}^{i} \alpha_{ij} U_j - c_i \\ &= \alpha_{i_1}\Phi^{-1}[\Phi(c_1)\Phi(U_1)] + (1-\alpha_{i1}^2)^{1/2} V - c_i \end{aligned} \tag{9}$$

The second term on the right-hand side of equation (9) is replaced by an equivalent term in one standard normal variable V.

It is seen that the dimension of the intersection in equation (1) is reduced by one in equation (7). Let the point $\mathbf{u}^*$ defined as in equation (2) be found by some suitable search algorithm. Note that due to the reduction of dimension in equation (9) the search for $\mathbf{u}^*$ is only a search in two variables. An approximation for the second factor in equation (7) then is by using equation (1)

$$P\left(\bigcap_{i=2}^{n} \{g_i(\mathbf{U}) \leqslant 0\} \right) \sim P\left(\bigcap_{i \in J_2} \{\mathbf{a}_i^T(\mathbf{u}-\mathbf{u}^*) \leqslant 0\} \right) C_2 \tag{10}$$

where, as before, $J_2 \subseteq \{2,\ldots,n\}$ is the subset of $k \leqslant n-1$ active constraints at $\mathbf{u}^*$ and C_2 the correction factor as in equation (1). Therefore, equation (7) can be rewritten as

$$P(\mathbf{Z} \leqslant \mathbf{c}) = \Phi(c_1) C_2 P\left(\bigcap_{i \in J_2} \{\mathbf{a}_i^T(\mathbf{u}-\mathbf{u}^*) \leqslant 0\} \right) \tag{11}$$

which is the main improvement as compared to Ref. 13.

For the last factor in equation (11) one, in turn, proceeds as described in equation (7) with new Z_i in equation (6) given by $Z_i = \mathbf{a}_i^T\mathbf{U}$ being standard normal with correlations $\rho_{ij} = \mathbf{a}_i^T\mathbf{a}_j$ and the new c_i defined by $c_i = \mathbf{a}_i^T\mathbf{u}^*$. Repeated at most $(n-1)$-fold application of this scheme finally yields the result. Note that the points $\mathbf{u}^*$ and the gradients $\mathbf{a}_i$ are different in each step.

TECHNICALITIES

The computation of the correction factors $C_t (t \geqslant 2)$ is facilitated by the following observations. Equation (9) shows that the g-functions are nonlinear only in the first variable but linear in the second variable. With $\bar{Z}_1$ as in equation (8) their gradients have the elements:

$$\frac{\partial g_i(\mathbf{u})}{\partial u_1} = \frac{\alpha_{i1}}{\varphi(\bar{z}_1)} \Phi(c_1)(\varphi(u_1) \tag{12a}$$

and

$$\frac{\partial g_i(\mathbf{u})}{\partial u_j} = \alpha_{ij} \qquad (2 \leqslant j \leqslant i) \tag{12b}$$

The Hessian matrix of second derivatives has only one nonzero entry, i.e.,

$$\frac{\partial^2 g_i(\mathbf{u})}{\partial u_1^2} = \alpha_{i1}\Phi(c_1)\frac{\varphi(u_1)}{\varphi^2(\bar{z}_1)}[\bar{z}_1\Phi(c_1)\varphi(u_1) - u_1\varphi(\bar{z}_1)] \tag{13}$$

In particular, the (directional) second-order derivatives at $\mathbf{w}^* = \boldsymbol{T}^T\mathbf{u}^*$ in the directions orthogonal to the active gradients to be used in equation (4) can be given analytically as:

$$\frac{\partial^2 g_i(\mathbf{w}^*)}{\partial w_r \partial w_s} = b_{1r} b_{1s} \frac{\partial^2 g_i(\mathbf{u}^*)}{\partial u_1^2} \tag{14}$$

The direction Hessian matrix in the directions $(\mathbf{b}_{k+1},\ldots,\mathbf{b}_n)$ is

$$\boldsymbol{H}_i = \frac{\partial^2 g_i(\mathbf{u}^*)}{\partial u_1^2} \hat{\boldsymbol{H}} \tag{15}$$

with

$$\hat{\boldsymbol{H}} = \mathbf{b}_1\mathbf{b}_1^T \tag{16}$$

It is easy to show that H_i has the $n-k$ eigenvalues:

$$\lambda_{i,k+1}=\frac{\partial^2 g_i(\mathbf{u}^*)}{\partial u_1^2}\sum_{j=k+1}^{n} b_{1j}^2 \tag{17a}$$

$$\lambda_{i,k+2}=\ldots\lambda_{i,n}=0 \tag{17b}$$

Further, if there are several active constraints $g_i(\mathbf{u})=0$ $(i=2,\ldots,k)$ in $\mathbf{u}^*$, it is according to equation (4) with the γ_i's determined from equation (3):

$$\lambda_{k+1}=\sum_{i=1}^{k}\gamma_i\left[\frac{\partial^2 g_i(\mathbf{u}^*)}{\partial u_1^2}\sum_{j=k+1}^{n} b_{1j}^2\right] \tag{18a}$$

$$\lambda_{k+2}=\ldots=\lambda_n=0 \tag{18b}$$

Therefore, the asymptotic correction factor in equation (10) or (11) becomes:

$$C_2=(1-\lambda_{k+1})^{-1/2} \tag{19}$$

It can be shown that for smaller $\|\mathbf{u}^*\|$ the correction factor can be improved especially for not very small probabilities by[15]:

$$C_2=\left(1-\frac{\lambda_{k+1}\varphi(\|\mathbf{u}^*\|)}{\|\mathbf{u}^*\|\Phi(-\|\mathbf{u}^*\|)}\right)^{-1/2} \tag{20}$$

REMARKS

In equation (10) the inactive constraints are omitted and similarly in all consecutive steps. It is possible to retain those constraints if one can justify reasonable expansion points. The simplest choice is to use $\mathbf{u}^*$ but better results are obtained by determining points defined by

$$\|\mathbf{u}^i\|=\min\{\|\mathbf{u}\|\}\quad\text{for}\left\{\mathbf{u}:\left(\bigcap_{\substack{j=2\\ j\neq i}}^{n} g_j(\mathbf{u})\leqslant 0\right)\cap(g_i(\mathbf{u})=0)\right\}$$

where i runs over all indices not contained in J_2. In these points a linear expansion of the corresponding surface may be used.

A crucial assumption for the foregoing results to hold is that the coordinate origin is not contained in D. If, however, $\mathbf{0}\in D$ or $g_i(\mathbf{0})<0$ for all i (which implies that D is a 'large' probability domain), it is proposed to use the so-called equivalent plane concept outlined in Refs 13 and 16. A simple, but accurate procedure is as follows. Referring to the last two lines in equation (7), all points $\mathbf{u}_i^*(1\leqslant i\leqslant m)$ for which

$$\|\mathbf{u}_i^*\|=\min\{\|\mathbf{u}\|\}\qquad\text{for }\{\mathbf{u}:\mathbf{g}_i(\mathbf{u})\leqslant 0\}$$

are determined. The individual events in equation (7) have probabilities involving the binormal distribution function. In particular, it is

$$P(X_i\leqslant c_i\,|\,X_1\leqslant c_1)=\frac{P((X_i\leqslant c_i)\cap(X_1\leqslant c_1))}{P(X_1\leqslant c_1)}$$

$$=\frac{P\left(\left(\sum_{j=1}^{i}\alpha_{ij}U_j\leqslant c_1\right)\cap(U_1\leqslant c_1)\right)}{P(U_1\leqslant c_1)}$$

$$=\frac{\Phi_2(c_i,c_1;\alpha_{i1})}{\Phi(c_1)}$$

where Φ_2 is best evaluated by numerical integration from:

$$\Phi_2(c_i,c_1;\alpha_{i1})=\Phi(c_i)\Phi(c_1)+\int_0^{\alpha_{i1}}\varphi(c_i,c_1;\rho)\,\mathrm{d}\rho$$

with $\varphi(c_i,c_1;\rho)$ the two-dimensional standard normal density. This suggests to use improved c_i's in equation (7) given by

$$c_i'=\Phi^{-1}\left[\frac{\Phi_2(c_i,c_1;\alpha_{i1})}{\Phi(c_1)}\right]$$

and to be interpreted as the distances to the origin of equivalent hyperplanes with the same gradients in $\mathbf{u}_i^*$ as the surfaces $g_i(\mathbf{u})$ and the same (individual) probability contents, respectively. It is easy to show that this procedure is asymptotically exact, i.e., for all c_i's approaching infinity. The proposal in Ref. 13 to determine equivalent gradients, too, by matching the derivatives of the probability of the given and the equivalent domain with respect to the coordinates can be shown to produce only slight if any improvement in the nonasymptotic case but requires substantial more numerical effort. If $\bar{\Phi}(\mathbf{c};\boldsymbol{R})=P(\bigcup_{i=1}^{m}\{X_i>c_i\})$ needs to be computed one simply passes over to the complement, i.e., $\bar{\Phi}(\mathbf{c};\mathbf{R})=1-\Phi(-\mathbf{c};\mathbf{R})$.

Another assumption in the approach is the linear independence of the active gradients $\mathbf{a}_i(i\in J)$ at $\mathbf{u}^*$. In the case of the multinormal integral this can be achieved by checking in each step of the procedure whether there are 'components' in the intersection (see equation (11)) for which $|\rho_{ij}|=1$. For $\rho_{ij}=+1$ the component with $\min\{c_i,c_j\}$ is omitted; if $\rho_{ij}=-1$, there is $\Phi_k(\ldots)=0$ for $c_i\leqslant -c_j$. In each step there are at most $p-1$ active

Table 1. Comparisons in terms of $\beta=-\Phi^{-1}(\mathrm{P})$

		$c=-4;\ n=10$		
ρ	Eq. (5)	Ref. 13	Eq. (11)	Exact[11]
0.0	14.11	14.15	14.15	14.15
0.2	8.69	8.95	8.94	8.93
0.4	6.46	7.10	7.08	7.05
0.6	4.75	5.97	5.97	5.92
0.8	2.67	5.11	5.10	5.06
0.9	−0.26	4.70	4.69	4.65
0.99	–	4.18	4.17	4.17
		$c=-4;\ \rho=0.6$		
n	Eq. (5)	Ref. 13	Eq. (11)	Exact[11]
2	4.68	4.73	4.73	4.73
5	5.17	5.49	4.49	5.48
10	4.75	5.97	5.97	5.92
20	0.53	6.42	6.39	6.28
50	–	7.00	6.92	6.67
100	–	7.25	7.20	6.92
		$\rho=0.6;\ n=10$		
c	Eq. (5)	Ref. 13	Eq. (11)	Exact[11]
0	–	1.16	1.14	1.13
−2	–	3.56	3.56	3.51
−4	4.75	5.97	5.97	5.92
−6	7.81	8.39	8.38	8.34
−8	10.48	10.81	10.80	10.79

components in a p-dimensional U- space which will have linear independent gradients provided that this checking is done.

The accuracy of the method described before has been tested at the examples given in Ref. 13. The quality of the numerical results is at least as good as the best approximation in Ref. 13. Theoretically, the results are asymptotically exact, i.e., for all values $\|\mathbf{u}^*\|$ being large during reduction of equation (11). Practically, the whole range of probabilities is calculated with sufficient accuracy as demonstrated in Table 1 for the equicorrelated case which can be checked by Dunnet/Sobel's[11] integration formula in one dimension. It is seen that Ruben's formula is inaccurate except for relatively small ρ's and large negative c's. Equation (11) turns out to be only slightly more accurate than the method in Ref. 13 despite the somewhat dubious theoretical basis of the latter. The error in both methods increases moderately with the problem dimension and is worst for medium correlations. Further numerical comparisons are given in Ref. 17.

The computation times of the new method presented herein are significantly smaller than for the method in Ref. 13 (roughly by a factor of 10) but still proportional to n^2 and depend to a certain degree on the minimization routine used for searching the points(s) $\mathbf{u}^*$ in equation (2) resp. equation (10). The analytical form of first- and second-order derivatives of both the objective function and of the constraints in equation (2) suggest to use a search algorithm for $\mathbf{u}^*$ which directly uses this information.

ACKNOWLEDGEMENTS

The authors highly appreciate discussions with Dr. M. Hohenbichler in which he contributed essential ideas.

REFERENCES

1 Hohenbichler, M. and Rackwitz, R. First-Order Concepts in System Reliability, *Structural Safety*, 1983, **1**, 117–188

2 Gollwitzer, S. and Rackwitz, R. First-Order Reliability of Structural Systems in Proc. ICOSSAR'85, 4th Conf. on Struc. Saf. and Reliab., Kobe, Vol. I, 171–180

3 Rackwitz, R. and Schrupp, K. Quality Control, Proof Testing and Structural Reliability, *Structural Safety*, 1985, **2**, 239–244

4 Madsen, H. O. Model Updating in Reliability Theory, Proc. ICASP-5, Vol. 1, Vancouver, 1987, 564–577

5 Hohenbichler, M. and Rackwitz, R. Non-normal Dependent Vectors in Structural Safety, *Journal of the Eng. Mech. Div., ASCE*, 1981, **107**(6), 1227–1249

6 Hohenbichler, M. An Asymptotic Formula for the Probability of Intersections in Berichte zur Zuverlaessigkeitstheorie der Bauwerke, SFB 96, 69, Technical University of Munich, 1984, 21–48

7 Hohenbichler, M., Gollwitzer, S., Kruse, W. and Rackwitz, R. New Light on First- and Second-Order Reliability Methods, *Structural Safety*, 1987, **4**, 267–284

8 Milton, D. C. Computer Evaluation of the Multivariate Normal Integral, *Technometrics*, 1972, **14**(4), 881–889

9 Bohrer, R. E. and Schervish, M. J. An Error-bounded Algorithm for Normal Probabilities of Rectangular Regions, *Technometrics*, 1981, **23**, 297–300

10 Rackwitz, R. Close Bounds for the Reliability of Structural Systems, Berichte zur Zuverlaessigkeitstheorie der Bauwerke, SFB 96, 29, Technical University of Munich, 1978, 68–75

11 Dunnet, C. W. and Sobel, M. Approximation to the Probability Integral and Certain Percentage Points of Multivariate Analogue of Student's Distribution, *Biometrika*, 1955, **42**, 258–260

12 Sidak, Z On Multivariate Normal Probabilities of Rectangles, their Dependence and Correlations, *Ann. Math. Statist.*, 1968, **39**, 1425–1434

13 Hohenbichler, M. and Rackwitz, R. A Bound and an Approximation to the Multivariate Normal Distribution Function, *Math. Japonica*, 1985, **30**(5), 821–828

14 Ruben, H. An Asymptotic Expansion for the Multivariate Distribution and Mill's Ratio, *Journal of Research*, National Bureau of Standards, 1964, **68b**(1), 1–11

15 Hohenbichler, M. and Rackwitz, R. Improvement of Second-Order Reliability Estimates by Importance Sampling, to be published in *Journal of Eng. Mech., ASCE*

16 Gollwitzer, S. and Rackwitz, R. Equivalent Components in First-Order System Reliability, *Reliability Engineering*, 1983, **5**, 99–115

17 Gollwitzer, S. and Rackwitz, R. Comparison of Numerical Schemes for the Multinormal Integral, IFIP-TC7 Conf., Aalborg, May, 1987 in Lecture Notes in Engineering, (Ed. P. Christensen), Springer-Verlag, Berlin, 1987, 157–174