Communications for Statistical Applications and Methods
2020, Vol. 27, No. 3, 377–383

https://doi.org/10.29220/CSAM.2020.27.3.377
Print ISSN 2287-7843 / Online ISSN 2383-4757

Moments calculation for truncated multivariate normal in nonlinear generalized mixed models

Seung-Chun Lee[1,a]

[a]Department of Applied Statistics, Hanshin University, Korea

Abstract

The likelihood-based inference in a nonlinear generalized mixed model often requires computing moments of truncated multivariate normal random variables. Many methods have been proposed for the computation using a recurrence relation or the moment generating function; however, these methods rely on high dimensional numerical integrations. The numerical method is known to be inefficient for high dimensional integral in accuracy. Besides the accuracy, the methods demand too much computing time to use them in practical analyses. In this note, a moment calculation method is proposed under an assumption of a certain covariance structure that occurred mostly in generalized mixed models. The method needs only low dimensional numerical integrations.

Keywords: truncated multivariate normal, moment, nonlinear generalized mixed model

1. Introduction

Suppose $\mathbf{Y} = (Y_1, \ldots, Y_n)' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, a multivariate normal with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and we are interested in $\mathrm{E}(Y_1^{\kappa_1} \cdots Y_n^{\kappa_n} | \, \mathbf{a} < \mathbf{Y} < \mathbf{b})$ where κ_i's are nonnegative integers, $\mathbf{a} = (a_1, \ldots, a_n)'$ and $\mathbf{b} = (b_1, \ldots, b_n)'$. We wish to compute product moments in a truncated multivariate normal distribution with Y_i truncated at the lower limit a_i and the upper limit b_i. In this truncation, some or all of the a_i can be $-\infty$ and some or all of the b_i can be ∞. When all the b_i's are ∞, the random vector $\mathbf{Y}(> \mathbf{a})$ is called lower truncated, whereas $\mathbf{Y}(< \mathbf{b})$ is called upper truncated when all the a_i's are $-\infty$. Note that the computations of the moment for the upper and lower truncated multivariate normal distributions are the same. The other type is double truncation $\mathbf{a} < \mathbf{Y} < \mathbf{b}$, which can have both lower and upper truncation points.

There is extensive literature dealing with the moment under different conditions derived from the types of truncations (one-sided, double-sided) and the number of variables (univariate, bivariate, multivariate), for instance Rosenbaum (1961) provided a formula for the moments of upper truncated bivariate normal variables. Shah and Parikh (1964) extended this result. They gave recurrence relations between moments for computing doubly truncated bivariate normal variables. Begier and Hamdan (1971) gave an explicit formula for the moment of doubly truncated bivariate normal variables with the same lower limit points. A more general result for the truncated bivariate normal can be found in Muthén (1990).

The multivariate case, Tallis (1961) gave an explicit formula for the first two moments of an upper truncated standard multivariate normal distribution. His results was extended to a general covariance matrix case by Amemiya (1974) and Lee (1979). Using recurrence relations, they gave the first two

[1] Department of Applied Statistics, Hanshin University, 137 Hanshindae-gil, Osan, Gyeonggi-Do 18101, Korea.
E-mail: seung@hs.ac.kr

Published 31 May 2020 / journal homepage: http://csam.or.kr

moments of a one–sided truncated multivariate normal distribution. Manjunath and Wilhelm (2012) extended these results further to the doubly truncated case. Using the moment generating function, they provided an explicit expression for the mean and variance of a multivariate normal distribution with arbitrary double truncation. The results were implemented in an **R** package, **tmvtnorm** (Wilhelm and Manjunath, 2010).

The likelihood-based inference in a nonlinear generalized mixed model requires computing up to the 4^{th} order moments of truncated multivariate normal random variables. That is, an algorithm meets our purpose only if it can compute $\mathrm{E}(Y_1^{\kappa_1} \cdots Y_n^{\kappa_n} |\ \mathbf{a} < \mathbf{Y} < \mathbf{b})$ for all the κ_i's satisfying $\sum_{i=1}^{n} \kappa_i \leq 4$. Arismendi (2013) derived a recursive formula for higher order moments in a one–sided truncated multivariate standard normal distribution. Kan and Robotti (2017) also developed excellent recurrence relations in a doubly truncated multivariate normal distribution, and claimed that their algorithm could compute higher order moments faster than Arismendi (2013). The result of Kan and Robotti (2017) can be applicable to the likelihood-based inference that has a capability to compute the necessary higher order moments. However, it seems that their algorithm is not fast enough for a large data analysis. For instance, they stated that the Matlab program bases on the algorithm computed all the product moments with $0 \leq \kappa_i \leq 4$ in less than 29 seconds on a PC when $n = 6$, but we have seen that the computing time is increasing rapidly as n increase. Actually, it requires computing 5^n product moments to get desired outcomes. Thus, even if n is moderately large, it is too time consuming. Note that the generalized mixed model (GMM) has some patterned covariance matrices. We may enjoy the patterned covariance structure to overcome this problem.

A general form of model equation for a mixed model is

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon} \tag{1.1}$$

with $\boldsymbol{\beta}$ and $\mathbf{u}$ representing fixed effects and random effects. $\mathbf{X}$ and $\mathbf{Z}$ are corresponding model matrices, and $\boldsymbol{\epsilon}$ is a vector of errors. The random part $\mathbf{Zu}$ could be partitioned as $\mathbf{Zu} = \sum_{i=1}^{r} \mathbf{Z}_i\mathbf{u}_i$. For notational convenience, define $\mathbf{u}_0 = \boldsymbol{\epsilon}$ and $\mathbf{Z}_0 = \mathbf{I}$. Customary, the random effects $\mathbf{u}_i$ have the properties $\mathrm{E}(\mathbf{u}_i) = \mathbf{0}$, $\mathrm{Var}(\mathbf{u}_i) = \sigma_i^2\mathbf{I}_{q_i}$ for $i = 0, \ldots, r$, and $\mathrm{Cov}(\mathbf{u}_i, \mathbf{u}_j') = \mathbf{0}$ for all $i \neq j$, where q_i denotes the dimension of $\mathbf{u}_i$. Then, under the normality assumption of random components, the distribution of $\mathbf{Y}$ is a multivariate normal $\mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = \sum_{i=0}^{r} \mathbf{Z}_i\mathbf{Z}_i'\sigma_i^2$.

A nonlinear generalized mixed model occurs when $\mathbf{Y}$ is an unobservable latent variable. For instance, we can only observe if the latent variable exceed a threshold of zero; that is, observed variable $\mathbf{W}$ is determined by $w_i = \mathbf{1}(y_i > 0)$, $i = 1, \ldots, n$, then $\mathbf{W}$ follows a mixed effects probit model. However, if we can observe $\mathbf{Y}$ only in limited range (ℓ, u), where ℓ and u are known lower and upper censored points, then the observed variable is defined by

$$w_i = \begin{cases} \ell, & \text{if } y_i \leq \ell, \\ y_i, & \text{if } \ell < y_i < u, \\ u, & \text{if } y_i \geq u. \end{cases}$$

This leads to a mixed effects censored regression model.

The most popular statistical method for the nonlinear GMM may be the maximum likelihood. Since the theory of maximum likelihood is well established, analyses based on the maximum likelihood can be done without theoretical problems, but there are certain computational issues inherent in the maximum likelihood method. Usually, the maximum likelihood estimate (MLE) is obtained by solving log-likelihood equations. Thus, to apply this approach, an analyzer must be able to calculate

log-likelihood equations, which requires computing the moments of a truncated multivariate normal distribution. See, for example, McCulloch (1994), and Lee and Lee (2019).

The computation of truncated moments is tedious. To avoid this difficulty, most statistical packages dealing with a nonlinear GMM maximize numerically approximated likelihood rather than exact likelihood by some numerical methods. For example, a SAS procedure **NLMIXED** (SAS, 2015) approximates the likelihood by integrating over the random components using the adaptive Gaussian quadrature, and then maximize it by a dual quasi-Newton method. **pglm** (Croissant, 2017), an R package approximate the likelihood function by Gauss-Hermite quadratures, while **Rchoice** (Sarrias, 2016), another R package use a Monte Carlo integration method for the approximation. As a result, it can be observed that different packages give different estimates and standard errors.

2. The result

In what follows, we will use shorthand notations

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{x}^{\boldsymbol{\kappa}} f(\mathbf{x})\mathrm{d}\mathbf{x} = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} x_1^{\kappa_1} \cdots x_n^{\kappa_n} f(x_1, \ldots, x_n)\mathrm{d}x_n \cdots \mathrm{d}x_1$$

and $\mathrm{E}(\mathbf{Y}^{\boldsymbol{\kappa}} |\, \mathbf{a} < \mathbf{Y} < \mathbf{b}) = \mathrm{E}(Y_1^{\kappa_1} \cdots Y_n^{\kappa_n} |\, \mathbf{a} < \mathbf{Y} < \mathbf{b})$. Let $\phi_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\Phi_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denote joint density and distribution functions of a n-variate normal with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Define $m^{\kappa}_{(a,b)}(\mu, \sigma^2) = \mathrm{E}(Y^{\kappa} |\, a < Y < b)$ and $\Phi(b, a; \mu, \sigma^2) = \Pr[a < Y < b]$ where $Y \sim N(\mu, \sigma^2)$, the κ-th truncated moment of a univariate normal distribution with mean μ, variance σ^2 and truncation from below at a and above at b. Also, we rewrite (1.1) as

$$\mathbf{Y} = \boldsymbol{\mu} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\mathbf{u} \sim \mathcal{N}_q(\mathbf{0}, \mathbf{D})$.

Theorem 1. *Under (2.1), the product moment* $\mathrm{E}(\mathbf{Y}^{\boldsymbol{\kappa}} |\, \mathbf{a} < \mathbf{Y} < \mathbf{b})$ *is equal to*

$$\frac{1}{\Pr[\mathbf{a} < \mathbf{Y} < \mathbf{b}]} \int_{-\infty}^{\infty} \prod_{i=1}^{n} m^{\kappa_i}_{(a_i,b_i)}\left(\tilde{\mu}_i, \sigma_\epsilon^2\right) \Phi\left(b_i, a_i; \tilde{\mu}_i, \sigma_\epsilon^2\right) \phi_q\left(\mathbf{u}; \mathbf{0}, \mathbf{D}\right) \mathrm{d}\mathbf{u}, \tag{2.2}$$

where $\tilde{\mu}_i = \mu_i + \mathbf{z}_i'\mathbf{u}$*, and* $\mathbf{z}_i'$ *is the* i^{th} *row of* $\mathbf{Z}$*.*

Proof: The marginal distribution of $\mathbf{Y}$ in (2.1) is $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and the conditional distribution of $\mathbf{Y}$, given $\mathbf{u}$ is $\mathcal{N}(\boldsymbol{\mu} + \mathbf{Z}\mathbf{u}, \sigma_\epsilon^2\mathbf{I})$ where $\boldsymbol{\Sigma} = \sigma_\epsilon^2\mathbf{I}_n + \mathbf{Z}'\mathbf{D}\mathbf{Z}$. Therefore,

$$\phi_n(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \int_{-\infty}^{\infty} \phi_n\left(\mathbf{y}; \boldsymbol{\mu} + \mathbf{Z}\mathbf{u}, \sigma_\epsilon^2\mathbf{I}_n\right) \phi_q\left(\mathbf{u}; \mathbf{0}, \mathbf{D}\right) \mathrm{d}\mathbf{u}$$

and

$$\begin{aligned}
\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{y}^{\boldsymbol{\kappa}} \phi_n\left(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \mathrm{d}\mathbf{y} &= \int_{\mathbf{a}}^{\mathbf{b}} \int_{-\infty}^{\infty} \mathbf{y}^{\boldsymbol{\kappa}} \phi_n\left(\mathbf{y}; \boldsymbol{\mu} + \mathbf{Z}\mathbf{u}, \sigma_\epsilon^2\mathbf{I}_n\right) \phi_q(\mathbf{u}; \mathbf{0}, \mathbf{D})\mathrm{d}\mathbf{u}\mathrm{d}\mathbf{y} \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{n} \int_{a_i}^{b_i} y_i^{\kappa_i} \phi_1\left(y_i; \tilde{\mu}_i, \sigma_\epsilon^2\right) \phi_q(\mathbf{u}; \mathbf{0}, \mathbf{D})\mathrm{d}y_i\mathrm{d}\mathbf{u} \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{n} m^{\kappa_i}_{(a_i,b_i)}\left(\tilde{\mu}_i, \sigma^2\right) \Phi_1\left(b_i, a_i; \tilde{\mu}_i, \sigma_\epsilon^2\right) \phi_q(\mathbf{u}; \mathbf{0}, \mathbf{D})\mathrm{d}\mathbf{u}
\end{aligned} \tag{2.3}$$

Thus, dividing (2.3) by $\Pr[\mathbf{a} < \mathbf{Y} < \mathbf{b}]$, we can get (2.2). □

The truncated moment of a univariate normal distribution can be calculated by an explicit formula, see for example Burkardt (2014). Thus, these simple observations can transition high dimensional integration problem into a relatively simple low dimensional integration problem. Theorem 1 is useful for computing $\Pr[\mathbf{a} < \mathbf{Y} < \mathbf{b}]$ as well. Let $\boldsymbol{\kappa} = \mathbf{0}$, then (2.3) is directly applicable to calculate the probability. This calculation could be simpler than the randomized Quasi Monte Carlo procedure by Genz (1992, 1993). See also the computation of $L_n(\mathbf{a}, \mathbf{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ in Kan and Robotti (2017).

Example 1. Consider the one-way classification model:

$$Y_{ij} = \mu_{ij} + \alpha_i + \epsilon_{ij}, \quad i = 1, \ldots, q;\ j = 1, \ldots, n_i$$

with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_q)' \sim \mathcal{N}(\mathbf{0}, \sigma_\alpha^2 \mathbf{I}_q)$. Then, the integral part of (2.2) becomes

$$\begin{aligned}
&\int_{-\infty}^{\infty} \prod_{i=1}^{q} \prod_{j=1}^{n_i} m_{(a_{ij},b_{ij})}^{\kappa_{ij}} \left(\mu_{ij} + \alpha_i, \sigma_\epsilon^2\right) \Phi_1 \left(b_{ij}, a_{ij}; \mu_{ij} + \alpha_i, \sigma_\epsilon^2\right) \phi_q \left(\boldsymbol{\alpha}; \mathbf{0}, \sigma_\alpha^2 \mathbf{I}_q\right) \mathrm{d}\boldsymbol{\alpha} \\
&= \prod_{i=1}^{q} \left\{ \int_{-\infty}^{\infty} \prod_{j=1}^{n_i} m_{(a_{ij},b_{ij})}^{\kappa_{ij}} \left(\mu_{ij} + \nu, \sigma_\epsilon^2\right) \Phi_1 \left(b_{ij}, a_{ij}; \mu_{ij} + \nu, \sigma_\epsilon^2\right) \phi_1 \left(\nu; 0, \sigma_u^2\right) \mathrm{d}\nu \right\}. \qquad (2.4)
\end{aligned}$$

This shows $\mathrm{E}(\mathbf{Y}^{\boldsymbol{\kappa}} |\ \mathbf{a} < \mathbf{Y} < \mathbf{b}) = \prod_{i=1}^{q} \mathrm{E}(\mathbf{y}_i^{\boldsymbol{\kappa}_i} |\ \mathbf{a_i} < \mathbf{y}_i < \mathbf{b}_i)$, where $\mathbf{y}_i, \boldsymbol{\kappa}_i, \mathbf{a}_i$ and $\mathbf{b}_i$ are appropriate partitions of corresponding vectors. This is somewhat obvious, because if $i \neq i'$, $\mathbf{y}_i$ and $\mathbf{y}_{i'}$ are independent. Once we note $\mathrm{Var}(\mathbf{y}_i) = \sigma_\epsilon^2 \mathbf{I}_{n_i} + \sigma_\alpha^2 \mathbf{J}_{n_i}$ where $\mathbf{J}_n$ is a $n \times n$ square matrix of 1's, then Corollary 1 is the consequence of (2.4).

Corollary 1. *Suppose* $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \sigma_\epsilon^2 \mathbf{I}_n + \sigma_\alpha^2 \mathbf{J}_n)$*, then*

$$\begin{aligned}
&\mathrm{E}(\mathbf{Y}^{\boldsymbol{\kappa}} |\ \mathbf{a} < \mathbf{Y} < \mathbf{b}] \qquad (2.5) \\
&= \frac{1}{\Pr[\mathbf{a} < \mathbf{Y} < \mathbf{b}]} \int_{-\infty}^{\infty} \prod_{i=1}^{n} m_{(a_i,b_i)}^{\kappa_i} \left(\mu_i + \nu, \sigma_\epsilon^2\right) \left[\Phi_1 \left(b_i; \mu_i + \nu, \sigma_\epsilon^2\right) - \Phi_1 \left(a_i; \mu_i + \nu, \sigma_\epsilon^2\right)\right] \phi_1 \left(\nu; 0, \sigma_u^2\right) \mathrm{d}\nu.
\end{aligned}$$

The covariance matrix considered in Corollary 1 can be observed in various statistical models such as the random effects panel regression model. A numerical method, such as the Gauss-Hermit quadrature might work well since (2.5) requires only a one-dimensional integration. Therefore, it seems that Corollary 1 is useful for limited dependent models in the random effects panel regression model.

Example 2. Suppose now a random nested model:

$$Y_{ijk} = \mu_{ijk} + \alpha_i + \gamma_{ij} + \epsilon_{ijk}, \quad i = 1, \ldots, q;\ j = 1, \ldots, r;\ k = 1, \ldots, m \qquad (2.6)$$

with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_q)' \sim \mathcal{N}(\mathbf{0}, \sigma_\alpha^2 \mathbf{I}_q)$ and $\boldsymbol{\gamma} = (\gamma_{11}, \ldots, \gamma_{qr})' \sim \mathcal{N}(\mathbf{0}, \sigma_\gamma^2 \mathbf{I}_{qr})$. Let $\mathbf{y}_i = (\mathbf{y}_{i1}, \ldots, \mathbf{y}_{ir})'$ and $\mathbf{y}_{ij} = (Y_{ij1}, \ldots, Y_{ijm})'$. As before, $\mathbf{y}_i$'s are independent, it suffice to calculate the truncated moment of $\mathbf{y}_i$, which have a covariance matrix of the form $\sigma_\epsilon^2 \mathbf{I}_{qr} + \sigma_\alpha^2 \mathbf{J}_{qr} + \sigma_\gamma^2 (\bigoplus_{i=1}^{q} \mathbf{J}_r)$, where $\mathbf{A} \bigoplus \mathbf{B}$ denotes

the direct sum of two matrices $\mathbf{A}$ and $\mathbf{B}$, for $\mathrm{E}(\mathbf{Y}^{\boldsymbol{\kappa}}|\ \mathbf{a} < \mathbf{Y} < \mathbf{b}]$. In view of Theorem 1, it is easy to check that

$$\mathrm{E}\left(\mathbf{y}_i^{\boldsymbol{\kappa}_i}|\ \mathbf{a}_i < \mathbf{y}_i < \mathbf{b}_i\right] = \frac{1}{\Pr\left[\mathbf{a}_i < \mathbf{y}_i < \mathbf{b}_i\right]} \times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{\prod_{j=1}^{q}\prod_{k=1}^{m} m_{\left(a_{ijk},b_{ijk}\right)}^{\kappa_{ijk}}\left(\tilde{\mu}_{ijk},\sigma_\epsilon^2\right)\Phi_1\left(b_{ijk},a_{ijk};\tilde{\mu}_{ijk},\sigma_\epsilon^2\right)\right\}$$
$$\phi_1\left(\nu;0,\sigma_\alpha^2\right)\mathrm{d}\nu\,\boldsymbol{\phi}_r\left(\boldsymbol{\gamma}_i;\mathbf{0},\sigma_\gamma^2\mathbf{I}_r\right)\mathrm{d}\boldsymbol{\gamma}_i,$$

where $\tilde{\mu}_{ijk} = \mu_{ijk} + \nu + \gamma_{ij}$.

This integration requires $(r+1)$-dimensional integration, which may be easier than rm-dimensional integration.

3. Numerical examples

To evaluate the performance of Corollary 1, we built an R program using Rcpp (Eddelbuettel *et al.*, 2018), and calculate the truncated mean and covariance matrix of $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{I}_n + \mathbf{J}_n)$ with truncation from above at the same point 1. We takes equal spaced n values including end points between -1 and 1 for $\boldsymbol{\mu}$. When $n = 5$, the calculated values are shown below.

```
$tmean
[1] -1.8767852 -1.4108813 -0.9786409 -0.5947036 -0.2695688

$tvar
          [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 1.4373311 0.4591293 0.4217783 0.3702806 0.3098274
[2,] 0.4591293 1.3412275 0.4028950 0.3542322 0.2968225
[3,] 0.4217783 0.4028950 1.1919713 0.3271588 0.2747230
[4,] 0.3702806 0.3542322 0.3271588 1.0004813 0.2433471
[5,] 0.3098274 0.2968225 0.2747230 0.2433471 0.7935028
```

These values coincide with those of **tmvtnorm** (Wilhelm and Manjunath, 2010) up to the 4^{th} digit after the decimal point, but Corollary 1 provides the result much faster than **tmvtnorm**. The mean running time of our program was 5.148414 milliseconds on our PC, which is about 15 times faster than **tmvtnorm** in 100 iterations. The advantage stands out when n is large. For example, when $n = 10$, Corollary 1 is about 200 times faster than **tmvtnorm**. These computations are done under **R** environment (R Core Team, 2019).

The computation of distribution function of a multivariate normal distribution may be another issue in computational statistics. For instance, most computational methods for the truncated moment are based on the distribution function. As the truncated moment, the multiple integral problem is involved in this computation. Genz (1992, 1993) gave a randomized Quasi Monte Carlo procedure for computing the distribution function of a multivariate normal distribution. Genz's method seems adequate for the computation since a Quasi Monte Carlo dominates quadrature methods for higher dimensional problems; therefore, we can compare our computation with those since his method was implemented in an **R** package called **mvtnorm** (Genz *et al.*, 2020). At the same configuration used before, we have observed that two methods give essentially the same values, matched up to the 4^{th} digit after the decimal point, for $n = 5, 10, 20, 100$; however, Corollary 1 is faster than **mvtnorm** about 100 times when $n = 10$, but 600 times when $n = 100$.

Kan and Robotti (2017) provided a Matlab package, **ftnorm** for computing the product moments of a truncated multivariate normal distribution. We could obtain a very similar mean and covariance matrix from **ftnorm**; however, it is much slower than **tmvtnorm**. It is hard to use **ftnorm** in real data analysis despite the capacity to compute the higher order moments because it is unendurably slow when n is large, say $n \geq 10$.

4. Conclusion

In this paper, an algorithm is provided to compute the higher order moment of a truncated multivariate normal distribution in a nonlinear generalized mixed model. The algorithm requires a relatively lower dimensional integration. For instance, it demands only one-dimension numerical integral in the one-way layout model. In this situation, we have shown numerically that the algorithm can compute the product moment in accuracy using the advantage of computing time. The algorithm would dominate other existing methods to calculate the higher order product moments of a truncated multivariate normal distribution with some covariance matrices. The algorithm also has an ability to compute the distribution function of a multivariate normal distribution with an apparent advantage.

The method may be useful in many areas. For example, we can use it to derive the exact maximum likelihood estimate in the random effects panel probit or censored data models.

References

Amemiya T (1974). Multivariate regression and simultaneous equation models when the dependent variables are truncated normal, *Econometrica*, **42**, 999–1012.

Arismendi JC (2013). Multivariate truncated moments, *Journal of Multivariate Analysis*, **117**, 41–75.

Begier MH and Hamdan MA (1971). Correlation in a bivariate normal distribution with truncation in both variables, *Australian Journal of Statistics*, **13**, 77–82.

Burkardt J (2014). The truncated normal distribution, Online Document. Available from: http://people.sc.fsu.edu/jburkardt/presentations/truncated normal.pdf

Croissant Y (2017). Package pglm: Panel generalized linear models, R package version 0.2-1. Available from: https://cran.r-project.org/web/packages/pglm/pglm.pdf

Eddelbuettel D, Francois R, Allaire J, Ushey K, Kou Q, Russell N, Bates D, and Chambers J (2018). *Seamless R and C++ Integration*. Available from: http://www.rcpp.org, http://dirk.eddelbuettel.com/code/rcpp.html

Genz A (1992). Numerical computation of multivariate normal probabilities, *Journal of Computational and Graphical Statistics*, **1**, 141–150.

Genz A (1993). Comparison of methods for the computation of multivariate normal probabilities, *Computing Science and Statistics*, **25**, 400–405.

Genz A, Bretz F, Miwa T, Mi X, Leisch F, Scheipl F, Bornkamp B, Maechler M, and Hothorn T (2020). mvtvnorm: Multivariate Normal and t Distributions, R package version 1.0-12. Available from: https://cran.r-project.org/web/packages/mvtnorm/mvtnorm.pdf

Kan R and Robotti C (2017). On moments of folded and truncated multivariate normal distributions, *Journal of Computational and Graphical Statistics*, **26**, 930–934.

Manjunath BG and Wilhelm S (2012). Moments calculation for the doubly truncated multivariate normal density, arXiv preprint arXiv:1206.5387.

Lee LF (1979). On the first and second moments of the truncated multi-normal distribution and a simple estimator, *Economics Letters*, **3**, 165–169.

Lee M and Lee SC (2019). Restricted maximum likelihood estimation of a censored random effects

panel regression model, *Communications for Statistical Applications and Methods*, **26**, 371–383.
McCulloch CE (1994). Maximum likelihood variance components estimation for binary data, *Journal of the American Statistical Association,* **89**, 330–335.
Muthén B (1990). Moments of the censored and truncated bivariate normal distribution, *British Journal of Mathematical and Statistical Psychology*, **43**, 131–143.
R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
Rosenbaum S (1961). Moments of a truncated bivariate normal distribution, *Journal of the Royal Statistical Society: Series B (Methodological)*, **23**, 405–408.
Sarrias M (2016). Discrete choice models with random parameters in R: The Rchoice Package, *Journal of Statistical Software*, **74**, 1–31.
Shah SM and Parikh NT (1964). Moments of singly and doubly truncated standard bivariate normal distribution, *Vidya*, **7**, 81–91.
SAS (2015). *SAS/STAT 14.1 Users Guide: The NLMIXED Procedure*.
Tallis GM (1961). The moment generating function of the truncated multi-normal distribution, *Journal of the Royal Statistical Society: Series B (Methodological)*, **23**, 223–229.
Wilhelm S and Manjunath BG (2010). tmvtnorm: Truncated multivariate normal and Student t distribution, *The R Journal,* **2**.

Received February 10, 2020; Revised March 7, 2020; Accepted March 7, 2020