AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Approximating Bayesian Posterior Means Using Multivariate Gaussian Quadrature

John A.L. Cranfield
Paul V. Preckel
Songquan Liu

Presented at Western Agricultural Economics Association 1997 Annual Meeting
July 13-16, 1997
Reno/Sparks, Nevada

July 1997

Approximating Bayesian Posterior Means using Multivariate Gaussian Quadrature

John A.L. Cranfield*

Paul V. Preckel

Songquan Liu

February 28, 1997

Abstract

Multiple integrals encountered when evaluating posterior densities in Bayesian estimation were approximated using Multivariate Gaussian Quadrature (MGQ). Experimental results for a linear regression model suggest MGQ provides better approximations to unknown parameters and error variance than simple Monte Carlo based approximations.

*The authors are Graduate Research Assistant, Professor and Graduate Research Assistant respectively, in the Department of Agricultural Economics, Purdue University, West Lafayette, Indiana, 47907-1145.

Approximating Bayesian Posterior Means using Multivariate Gaussian Quadrature

Abstract

Multiple integrals encountered when evaluating posterior densities in Bayesian estimation were approximated using Multivariate Gaussian Quadrature (MGQ). Experimental results for a linear regression model suggest MGQ provides better approximations to unknown parameters and error variance than simple Monte Carlo based approximations.

Approximating Bayesian Posterior Means

using Multivariate Gaussian Quadrature

Frequently, researchers wish to account for prior beliefs when estimating unknown parameters. For instance, prior beliefs regarding curvature restrictions may be incorporated during demand system estimation, or one may believe the space of an unknown parameter is bounded by a specific value. In such cases, Bayesian analysis may be used to incorporate prior beliefs.

In particular, consider the linear regression model denoted in matrix form as $\mathbf{Y}=\mathbf{X}\beta+\varepsilon$, and suppose that previous work has lead a researcher to form prior beliefs about the unknown parameter vector β and the error variance, σ^2. Now, additional data are available and new estimates of β and σ^2 are required. Moreover, the researcher wishes to explicitly account for their prior beliefs about β and σ^2. Consequently, a Bayesian approach to estimating the regression model may be used.

However, estimating parameters in a Bayesian framework, or using Bayesian analysis for inference and hypothesis testing, requires the use of multiple integration. Frequently, such integrals may be difficult, if not impossible, to evaluate analytically. In these circumstances, numerical integration techniques may be used to approximate the integrals. Such techniques include Monte Carlo integration with Importance Sampling (Kloek and van Dijk 1978, and Geweke 1989), Monte Carlo integration using antithetic acceleration (Geweke 1988) and univariate Gaussian Quadrature using a Cartesian product grid (Naylor and Smith 1988).

Recently, DeVuyst (1993) presented a new approach to numerical integration called Multivariate Gaussian Quadrature (MGQ). This approach has the advantage of being computationally inexpensive, and avoids independence problems encountered when using a Cartesian product grid to approximate multiple integrals.

The purpose of this paper is to demonstrate how MGQ can be used to approximate multiple integrals encountered in Bayesian analysis. The paper is organized as follows. First the notation used throughout is stated. Then, the Bayesian framework is illustrated in the context of estimating a linear regression model with unknown parameters and error variance. Following this, the data and experimental procedures are outlined. Results from the MGQ approximation of the regression coefficients and error variance are then presented, and juxtaposed to approximations computed using Monte Carlo integration, and also analytical solutions. Finally, results are summarized and discussed.

Notation

The following notation will be used in this study: $\mathbf{Y}=\mathbf{X}\beta+\varepsilon$ is a linear regression model, where $\mathbf{Y}$ is a tx1 vector of dependent variable observations, $\mathbf{X}=[\mathbf{x}_1\ \mathbf{x}_2\ ...\ \mathbf{x}_k]$ is a txk matrix of independent variable observations, $\mathbf{x}_i$ is a tx1 vector of observations of the ith explanatory variable (with every component of $\mathbf{x}_1$ equal to 1.0), $\beta=[\beta_1\ \beta_2\ ...\ \beta_k]$ is a k x1 vector of unknown parameters to be estimated ($\beta_k \in (-\infty,+\infty)$), ε is a tx1 vector of error terms (which is assumed to be an independently and identically distributed Normal with mean 0 and variance σ^2), σ^2 is the unknown error variance ($\sigma^2 \in (0,+\infty)$), t=1,2,...T represents the number of observations, i=1,2,...,k represents the number of explanatory variables, $\mathbf{b}=[b_1\ b_2\ ...\ b_k]$ is a kx1 vector of estimated values for β, s^2 is an estimate of the

unknown error variance, $\mathbf{\Phi}=[\beta, \sigma^2]'$ is a (k+1)x1 vector of true parameter values and error variance, $\phi=[\mathbf{b}, s^2]'$ is a (k+1)x1 vector of estimated parameter values and error variance, $\mathbf{\Omega}$ is the k+1 dimensional space of Φ, $L(\mathbf{\Phi},\phi)=[\mathbf{\Phi}-\phi]'\mathbf{I}[\Phi-\phi]$ is a quadratic loss function, where **I** is an identity matrix, $l(\mathbf{Y} \mid \beta, \sigma, \mathbf{X})$ is a likelihood function, $p(\beta, \sigma)$ is a prior density, $q(\beta, \sigma \mid \mathbf{Y}, \mathbf{X})$ is a posterior density, $U(x_L,x_U)$ is a uniform density over the interval (x_L,x_U), and $N(\theta,\tau^2)$ is a Normal density with mean θ and variance τ^2.

Bayesian Framework

The Bayesian approach to estimation assumes the data may be described via a likelihood function. Here, it is assumed that the likelihood function is a normal

distribution:

However, Bayesians account for prior beliefs concerning parameters of interest with a prior density. Frequently, however, prior information is limited, in which case a non-informative, or diffuse, prior is assumed. Following Zellner (1971), the joint prior density

$$p(\beta,\sigma) \propto \sigma^{-1}.$$

of β and σ^2 is assumed to be Jeffrey's prior (*i.e.*, diffuse), and is denoted as:

To form the posterior, multiply the likelihood function and prior together and integrate with respect to β and σ^2 to obtain a marginal density for **Y** and **X**. Now, apply Bayes' Theorem to obtain a density for β and σ^2, conditional on **Y** and **X**. This

conditional density is referred to as the posterior density in Bayesian analysis. The posterior reflects updated beliefs after combining prior beliefs with information contained in the sample (Berger 1985, p.126). For the assumed likelihood function and prior

density, the posterior is:

Judge *et al.* (1988 p.103) show $q(\beta,\sigma \mid \mathbf{Y},\mathbf{X})=q(\beta \mid \sigma,\mathbf{Y},\mathbf{X})\cdot q(\sigma \mid \mathbf{Y})$, where $q(\beta \mid \sigma,\mathbf{Y},\mathbf{X})$ is a multivariate normal density and $q(\sigma \mid \mathbf{Y})$ is an inverted gamma density. Consequently, the posterior can be shown to equal the product of two densities, each with an analytical solution to its mean. Thus, the exact form of $q(\beta,\sigma \mid \mathbf{Y},\mathbf{X})$ can be determined. Moreover, because the proportionality constant is contained in the numerator and denominator of the posterior and is not a variable of integration, it cancels out and the posterior can be stated as a strict equality.

In Bayesian analysis, the posterior is used in conjunction with a loss function to estimate the unknown parameters of interest. The framework is couched within an optimization problem where the objective is to minimize the expected loss of the estimate. Here, the loss function measures the difference between the true parameter and the estimated value, while the posterior serves as a weighting function in the expectation.

Thus, the Bayesian decision problem may be stated as:

Since a quadratic loss function is assumed, the Bayesian estimate of $\mathbf{\Phi}$, denoted as $\boldsymbol{\phi}$, is the mean of the marginal posterior for $\mathbf{\Phi}$. As well, individual estimates of Φ, ϕ_i, are

calculated from the marginal posterior for ϕ_i. So, for ϕ_i, the Bayesian estimate is: Thus, coefficients in $\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$, or σ^2, can be estimated by computing the expected value of the marginal posterior. Attention is now turned to the proposed MGQ framework.

Multivariate Gaussian Quadrature

As noted in the introduction, marginalization of the posterior requires multiple integration. Frequently, such integration is difficult, if not impossible to perform analytically. Consequently, numerical integration can be used to approximate these integrals. The approach proposed here is to approximate the marginal posteriors with the MGQ approach found in DeVuyst.

The MGQ framework allows one to reduce the number of points needed to numerically approximate a multiple integral. The approximation works by selecting points from the support of the density being approximated. The integrand is then evaluated at each point and multiplied by a corresponding probability weight. The probability weight is determined from a discrete distribution used to approximate the underlying density. The sum of the weighted evaluations is then used as the approximation to the integral. Details of which points to choose, and what weighting values to use are contained in Appendix 1.

In general, MGQ provides better approximations than Monte Carlo integration.

Moreover, if the variables of integration are correlated (or not independent), then MGQ is preferred to a Cartesian product grid of univariate GQ approximation for each variable, since the latter tend to place probability mass in negligible regions of the approximating distribution. In such instances, poor approximations may result. Next, the data and experimental procedures are outline.

Data and Procedure

To evaluate this framework, 10 replications were completed, each consisting of 25 observations of pseudo-random data for $\mathbf{Y}$, $\mathbf{X}$ and $\boldsymbol{\varepsilon}$, and a vector of coefficients, β. It is assumed that the regression model consists of an intercept and two explanatory variables. In all 10 replications, the same set of randomly drawn explanatory variables was used. These data were generated as follow: $\mathbf{x}_1$ was a vector of ones, $\mathbf{x}_2$ was drawn from $U(5,15)$ and $\mathbf{x}_3$ from $U(1,10)$. However, in each replication, the error term, $\boldsymbol{\varepsilon}$, was drawn from $N(0,1)$, β_1 from $U(2.0,5.0)$, β_2 from $U(3.5,6.5)$, and β_3 from $U(-5.0,-1.0)$. The data, parameters and error term were then combined to compute $\mathbf{Y}$. Thus, each replication consists of 25 observations of same values for $\mathbf{X}$, but different values for $\mathbf{Y}$.

The approximation begins by assuming the approximating distribution is centered on the OLS estimates.[1] However, since $\boldsymbol{\Omega}$ is the product of three infinite parameter spaces and one parameter space bounded below by zero, one must reduce the area of integration in order to make the approximation feasible using a uniform density function (Arndt 1996). This is done by assuming that most of the mass in the approximating distribution is

[1] This centering is arbitrary. In fact, one could use other values, such as Maximum Likelihood estimates.

contained within a fixed number of standard deviations of the OLS estimates.

To simplify the computations, the area of integration is broken into smaller blocks. Each block is determined by forming a grid over the area of integration, where each block initially has a dimension of one standard deviation by one standard deviation (this approximation is referred to as MGQ-1). In each block, 59 points were used to evaluate the integrand in (1). These points correspond to an order 5 MGQ for 4 variables using a uniform weight function.

The sensitivity of MGQ approximation to the limits of integration was investigated by changing the size of the grid considered. Specifically, the grid was increased to ± two standard deviations of the OLS estimates (MGQ-2), and then ± three standard deviations (MGQ-3). This increases the total number points used in the approximation because the size of each block was maintained, while the number of blocks increased.

To show the efficacy of the MGQ framework versus a competing approach, simple Monte Carlo integration was used to compute the mean of the marginal posteriors in each replication. The Monte Carlo approximations were based on 10,000 draws. Since the variables of integration are regression coefficients and the error variance, each draw had an element corresponding to a regression coefficient and the error variance. The intercept term, β_1, was drawn from $U(-20,20)$, β_2 from $U(0.05,20)$, β_3 from $U(-20,-0.05)$, and σ^2 from $U(0.05,1)$. The corresponding 4-tuple was used in conjunction with $\mathbf{Y}$ and $\mathbf{X}$ to evaluate the integrand in (1), with equal weights assigned to each evaluation.

To compare the MGQ and Monte Carlo approximations, the analytical solution to the Bayesian decision problem is needed. For the assumed likelihood function and prior

density, the analytical solution to (1) is well known. In particular, the marginal posterior for β is a multivariate Student t density, while the marginal posterior for σ^2 is an inverted gamma density (Judge *et al*. 1988, pp.219-220).

The analytical solutions were computed in each replication, and used to judge the MGQ and Monte Carlo approximations. This comparison is done in three ways: 1) compare the mean of the approximations to the mean of the analytical solutions, 2) compare the standard deviation of the approximations to the standard deviation of the analytical solutions, 3) compare the MGQ and Monte Carlo approximations on the basis of mean absolute percent error (MAPE). The latter criteria provides a percent measure of the absolute deviation between the approximation and the analytical solution.

Results

Table 1 shows summary statistics (mean, standard deviation, maximum and minimum) for the analytical solutions, and the MGQ and Monte Carlo approximations, and MAPE of the MGQ and Monte Carlo approximations.[2] Compared to the analytical solutions, the MGQ framework seems to provide better approximations to (1) than Monte Carlo approximations.

For MGQ-1, a total of 944 points were used in the approximation. Note that the mean approximation of b_1 and s^2 were below those of the analytical solution, while the mean approximation for b_2 and b_3 were slightly above the mean of the analytical solution. In addition, there was little difference between the standard deviations of the MGQ-1

[2] Maximum and minimum values are shown to further indicate the spread of the analytical solution and approximations, but are not be discussed.

approximations and the standard deviations of the analytical solutions. As well, the approximation of s^2 had the largest MAPE, followed by b_1, then b_3 and finally b_2.

When the area of integration was increased to ± two standard deviations from the OLS estimates (*i.e.*, MGQ-2), the total number of points increased to 15,104. In this case, the mean of the MGQ-2 approximation of the regression coefficients were below the mean of the analytical solutions. However, when compared to MGQ-1, the mean of the MGQ-2 approximations of b_1 and b_2 increased, while those for b_3 and s^2 decreased slightly. The standard deviation of the MGQ-2 approximations for b_2, b_3 and s^2 were below those corresponding to the MGQ-1 approximations. Compared to the MGQ-1 results, the MGQ-2 approximation of b_1 and s^2 had higher MAPE values, but lower MAPE values for b_2 and b_3.

Increasing the area of integration to ± three standard deviations from the OLS estimates resulted in approximations based on a total of 79,464 points. Except for s^2, the mean of MGQ-3 approximations were very precise compared to the mean of the analytical solutions. In particular, any differences in the means were in the second or third decimal place. As well, the standard deviations associated with the MGQ-3 approximations of b_2, b_3 and s^2 were not appreciably different from MGQ-1 and MGQ-2. As well, MGQ-3 approximations of b_2, b_3 and s^2 lower MAPE values than the other MGQ approximations.

Results from the Monte Carlo integration were not as good. The mean approximation of b_1 and s^2 differed substantially from the mean of the analytical solution and the MGQ values. Although, b_2 and b_3 appear to have been relatively close to the analytical solutions. In general, the standard deviation of these approximations suggests

that Monte Carlo integration results in greater variability. Realize, of course, that such a result is conditional on the size of the interval from which samples are being drawn. Finally, notice that the MAPE values for the Monte Carlo approximations were higher than those associated with the MGQ approximations.

Summary and Conclusions

Results from this study indicate that a Multivariate Gaussian Quadrature approach to evaluating multiple integrals provides an alternative to Monte Carlo methods for Bayesian analysis. In particular, MGQ produced better estimates of linear regression coefficients and error variance than Monte Carlo integration, but using only about 10% of the integrand evaluations.

Moreover, as the number of points used in the MGQ approximation was increased, the precision of the slope coefficients and error variance, relative to the analytical solution, did not change appreciably. At the same time, however, the approximated intercept terms appear to diverge from the analytical solution. This latter result is troubling, and is an area of further investigation. In addition, it would be illuminating to compare the MGQ framework to other techniques when there are many variables of integration.

	Analytical Solution	MGQ-1[a]	MGQ-2[b]	MGQ-3[c]	Monte Carlo[d]
No. of points		944	15,104	76,464	10,000
Mean[e]					
b_1	3.5107	3.3215	3.3989	3.5285	2.4319
b_2	4.6217	4.6293	4.6187	4.6111	4.6604
b_3	-2.8906	-2.8994	-2.8872	-2.8995	-2.9132
s^2	1.1939	1.1375	1.1309	1.1258	1.6119
Standard Deviation[f]					
b_1	0.7891	0.8417	0.8708	1.1392	2.0832
b_2	0.6984	0.6882	0.6857	0.6947	0.7609
b_3	0.9948	0.9873	0.9788	0.9848	0.7355
s^2	0.0961	0.1053	0.1046	0.1064	0.2498
Maximum					
b_1	4.4708	4.5975	4.9611	5.2832	4.8839
b_1	5.8222	5.8322	5.8390	5.8388	5.7251
b_1	-1.3472	-1.2832	-1.2535	-1.3004	-1.7149
s^2	1.3565	1.2664	1.2631	1.2655	1.8869
Minimum					
b_1	1.9734	1.7437	1.9529	1.8221	-0.4185
b_1	3.8169	3.7859	3.8687	3.8222	3.5155
b_1	-4.5785	-4.5468	-4.5317	-4.5656	4.2156
s^2	1.0476	0.8776	0.8869	0.8936	1.1738
Mean Absolute Percent Error (MAPE)[e]					
b_1		6.0174	9.6004	21.8903	55.7686
b_2		1.7376	0.7653	0.3883	5.9220
b_3		1.9541	1.3863	1.0087	11.7076
s^2		8.4430	8.6293	8.4012	35.7187

aMGQ-1 is Multivariate Gaussian Quadrature based on 10 replications, and evaluating (1) within one standa
he OLS estimate.

bMGQ-2 is Multivariate Gaussian Quadrature based on 10 replications, and evaluating (1) within two stan
the OLS estimate.

cMGQ-3 is Multivariate Gaussian Quadrature based on 10 replications, and evaluating (1) within three
ations of the OLS estimate.

dReported values are the mean and standard deviation of the approximation (or exact solution) over all

eMAPE is mean absolute percent error, and is calculated as:

$$MPAE = \frac{1}{R}\sum_{r=1}^{R} |Actual - Appr.| over Actual$$

where *Actual* is the analytical solution, *Appr.* is the numerical approximation, and R is the number of replications.

Appendix 1: Choice of sample points and corresponding probability weights.

The choice of points and weights used to evaluate the integrand is not arbitrary. The approach followed here is to form a uniform GQ of order *d* for each variable of integration. Then, create a Cartesian product GQ of order *d* using the uniform GQ density. The number of points required to approximate the underlying density in each uniform GQ is determined from $2n-1=d$, where *n* is the number of points. So, the

$$-\frac{d+1}{2}-^{k+1}$$

Cartesian product has at most *m* points, where *m* is:

Now, use the (*k*+1)-tuples determined above to compute the corresponding weights at each point. This is done by solving the following system of linear simultaneous

$$\sum_{i=1}^{m} p_i \prod_{j=1}^{k+1} \beta_{ij}^{\gamma_j} = \int_{\Omega_1}\int_{\Omega_2}\ldots\int_{\Omega_{k+1}} w(\beta)\prod_{j=1}^{k+1}\beta_j^{\gamma_j} d\,\beta_{k+1}\ldots d\,\beta_2 d\,\beta_1 \quad \forall|\xi|\le d$$

equations:

where $p_i \in [0,1]$ is the probability mass used to weight the integrand evaluated at the *i*th point, β_{ij} is the *j*th element of the (*k*+1)-tuple at the *i*th point, γ_j is an interger value, $\xi=\Sigma_j\gamma_j$, $w(\beta)$ is a weighting function, and the right hand side are moments (including cross-moments) of the polynomial function in the left hand side. All points satisfying the above system, and have non-negative probabilities are used in evaluating the multiple integral in (1).

Bibliography

Arndt, Channing. *Three Essays in the Efficient Treatment of Randomness.* Unpublished Ph.D. Dissertation, Department of Agricultural Economics, Purdue University, West Lafayette, Indiana, 1996.

Berger, James O. *Statistical Decision Theory and Bayesian Analysis* (Second Edition). New York: Springer-Verlag, 1985.

DeVuyst, Eric A. *Moment Preserving Approximations of Multivariate Probability Distributions for Decision and Policy Analysis: Theory and Application.* Unpublished Ph.D. Dissertation, Department of Agricultural Economics, Purdue University, West Lafayette, Indiana, 1993.

Geweke, John. Antithetic Acceleration of Monte Carlo Integration in Bayesian Analysis. *Econometrica* 56(1988): 73-89.

Geweke, John. Bayesian Inference in Econometric Models Using Monte Carlo Integration. *Econometrica* 57(1989): 1317-1339.

Kloek, T. and H.K. van Dijk. Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo. *Econometrica* 46(1978): 1-19.

Judge, George G., W.E. Griffiths, R. Carter Hill, Helmut Lütkepohl, and Tsoung-Chao Lee. *The Theory and Practice of Econometrics* (Second Edition). New York: John Wiley and Sons, 1988.

Naylor, J.C. and A.F.M. Smith. Econometric Illustrations of Novel Numerical Integration Strategies for Bayesian Inference. *J. of Econometrics* 38(1988):103-125.

Zellner, Arnold. *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley and Sons, 1971.