Chung Yuan Journal
Vol. 27. No. 1, 1999. pp. 61-67

Bayesian Estimation for the Optimum in Single Factor Quadratic Regression

T.H. FAN

Graduate Institute of Statistics
National Central University
Chung Li, Taiwan, R.O.C.

M.J. KARSON

Decision Sciences Department
University of New Hampshire
Durham, New Hampshire, U.S.A.

H.S. WANG AND C.C. LEE

Graduate Institute of Industrial Engineering
Chung Yuan Christian University
Chung Li, 32023, Taiwan, R.O.C.

(Received: June 25, 1998; Accepted: February 10, 1999)

ABSTRACT

This article deals with the problem of estimating the optimal location and magnitude of a single factor quadratic response function. A Bayesian approach is employed so that the information accumulated by a subject matter specialist can be fully utilized. Two families of prior distributions are considered. One is informative, and the other is noninformative. For a given set of data, the form of the marginals, and the posterior mean, mode, variance and 95% highest posterior density intervals are reported. A robustness analysis is undertaken. Comparisons are also made between two noninformative priors.

Key words: *Quadratic Regression, Response Surface, Bayes, Estimation, Optimization.*

I. INTRODUCTION

Dealing with a guadratic response model, people are interested in inference, usually estimation, for the optimum response, maximum or minimum, and the point in the factor space where the optimum occurs. Consider the single regressor quadratic regression model

$$y = \delta_1 + \delta_2 x + \delta_3 x^2 + \varepsilon \tag{1}$$

where δ_1, δ_2, δ_3 are unknown parameters, and the random error $\varepsilon \sim N(0,\sigma^2)$. In the linear model (1), interest is generally in inference on the unknown regression parameters δ_1, δ_2, δ_3 and the error variance σ^2. However, for our purposes the parameters of interest are α and γ, where γ is the coordinate of the carrier x at which the optimum expected value of the response variable, maximum or minimum, takes on its ordinate α. In the context of control setting, γ is the planning parameter. Therefore, as in Mandel (1978) the linear model (1) is transformed to the nonlinear model

$$y = \alpha + \beta(x-\gamma)^2 + \varepsilon \tag{2}$$

where

$$\alpha = \delta_1 - \frac{\delta_2^2}{4\delta_3},$$
$$\beta = \delta_3, \tag{3}$$
$$\gamma = -\frac{\delta_2}{2\delta_3}.$$

The parameters α, β, γ are referred to as characteristic parameters, whereas δ_1, δ_2, δ_3 are referred to as linear model parameters. Note that model (2) is reparameterized in terms of parameters that have physical meaning. Since we use the Bayesian paradigm for inference, any non-Bayesian complications that may arise from a nonlinear regression model that emphasizes explicitly the parameters of interest and thereby leads naturally to Bayesian estimation. Classical inference for α and γ is based on least squares estimation of the linear model parameters, and therefore small sample properties of these inferences are not generally best. The classical estimator of the optimum point may be traced to Hotelling (1941), and perhaps earlier, where the ordinate and the coordinate of the optimum are estimated by applying the necessary and sufficient conditions to the estimated response function obtained through the ordinary least squares methodology. The usual procedure, as in for example Khuri and Cornell (1987), for finding coordinates of stationary points is to calculate the derivative of the estimating response surface equation and to solve equations obtained by equating derivatives to zero. Thus, if $\hat{\delta}_1$, $\hat{\delta}_2$ and $\hat{\delta}_3$ are the ordinary least squares estimators of δ_1, δ_2 and δ_3, respectively, then the classical estimators of α and γ are $\hat{\alpha} = \hat{\delta}_1 - \hat{\delta}_2^2/(4\hat{\delta}_3)$ and $\hat{\gamma} = -\hat{\delta}_2/(2\hat{\delta}_3)$, respectively.

In this paper we are interested in Bayesian estimation of α and γ in the model (2). Bayesian estimation is not uncommon in nonlinear regression, as discuessed in Seber and Wild (1989). However, there has been little work on Bayesian estimation of α and γ, besides Buonaccorsi and Gatsonis (1988), who used a prior distribution of Hoadley (1970) for estimating γ in the context of Bayesian inference for ratios of linear model coefficients, and Fan, Karson and Wang (1996), who estimated α, β and γ by transforming a trivariate-t posterior distribution of $\hat{\delta}_1$, $\hat{\delta}_2$ and $\hat{\delta}_3$.

Assume that data, (x_i, y_i); $i = 1,2,\ldots,n$ are available in the form of a random sample. Let ,

$$y_i = \alpha + \beta(x_i - \gamma)^2 + \varepsilon_i, \tag{4}$$

where the ε_i are i.i.d. $N(0,\sigma^2)$ random variables. We shall denote the y_i's by $\mathbf{y}$ and we shall not distinguish notationally between realizations and random variables. We shall write the likelihood as

$$L(\alpha,\beta,\gamma,\sigma^2;\mathbf{y}) = \frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[y_i-\alpha-\beta(x_i-\gamma)^2]^2\right\} \tag{5}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left\{-\frac{1}{2\sigma^2}[\beta^2 a_1(\gamma) + a_3(\alpha) - 2\beta a_2(\alpha,\gamma)]\right\} \tag{6}$$

where

$$a_1(\gamma) = \sum_{i=1}^{n}(x_i-\gamma)^4,\ a_2(\alpha,\gamma) = \sum_{i=1}^{n}(y_i-\alpha)(x_i-\gamma)^2, a_3(\alpha) = \sum_{i=1}^{n}(y_i-\alpha)^2$$

We shall also use $\pi(\bullet)$ to denote a prior distribution, $m(\bullet)$ a marginal distribution, and $p(\bullet|\bullet,\mathbf{y}) = P(\bullet)$a posterior, where the posterior distribution is conditional on $\mathbf{y}$, and perhaps other random variables. In addition we shall sometimes write $\pi(\bullet)$ = as a symbolic designation of a density.

The balance of this paper is organized into three sections. In Section 2 we discuss our prior distributions and in Section 3 the Bayesian analysis. In Section 4 we illustrate the analysis.

II. PRIOR DISTRIBUTIONS

Two families of prior distributions are considered, one is informative, and the other noninformative. In each case the four-variate prior distribution for $\alpha,\beta,\gamma,\sigma^2$ is denoted $\pi(\alpha,\beta,\gamma,\sigma^2)$. The informative prior is a normal inverse gamma (NIG), and the noninformative prior is the Jeffreys prior. It is also understood that the supports of the priors are $-\infty<\alpha<\infty$, $-\infty<\beta<\infty$, $-\infty<\gamma<\infty$, $\sigma^2>0$.

A. The Informative Prior

Assume that α and γ are independent of each other, and of σ^2, and are independent of β for a given σ^2. Therefore, the four-variate prior distribution of α, β, γ, and σ^2 is $\pi(\alpha,\beta,\gamma,\sigma^2) = \pi(\alpha)\pi(\gamma)\pi(\beta|\sigma^2)\pi(\sigma^2)$. The NIG prior assumes $\pi(\alpha) = N(\mu_\alpha,\sigma_\alpha^2)$, $\pi(\beta|\sigma^2) = N(\mu_\beta,\sigma^2\sigma_\beta^2)$, $\pi(\gamma) = N(\mu_\gamma,\sigma_\gamma^2)$ and $\pi(\sigma^2) = \mathrm{IG}(a,b)$, where N denotes a normal density and IG an inverse gamma. As is in usual practice, it is assumed that subjective information is available on the eight hyperparameters. As $\sigma_\beta^2 \to \infty$,and/or a and $b \to 0$, the priors on β and σ^2 become, in fact, vague priors. We shall write the NIG prior as

$$\begin{aligned}\pi(\alpha,\beta,\gamma,\sigma^2) = {} & \frac{1}{(2\pi\sigma_\alpha^2)^{1/2}}\exp\left\{-\frac{1}{2\sigma_\alpha^2}(\alpha-\mu_\alpha)^2\right\} \\ & \times \frac{1}{(2\pi\sigma^2\sigma_\beta^2)^{1/2}}\exp\left\{-\frac{1}{2\sigma^2\sigma_\beta^2}(\beta-\mu_\beta)^2\right\} \\ & \times \frac{1}{(2\pi\sigma_\gamma^2)^{1/2}}\exp\left\{-\frac{1}{2\sigma_\gamma^2}(\gamma-\mu_\gamma)^2\right\} \\ & \times \frac{b^a}{\Gamma(a)}\frac{1}{(\sigma^2)^{a+1}}\exp\left\{-\frac{b}{\sigma^2}\right\}I_{(0,\infty)}\end{aligned} \tag{7}$$

$$\propto \frac{1}{(\sigma^2)^{(a+3/2)}}\exp\left\{-\frac{1}{2\sigma^2}\left[2b + \frac{1}{\sigma_\beta^2}(\beta-\mu_\beta)^2 + \frac{\sigma^2}{\sigma_\alpha^2}(\alpha-\mu_\alpha)^2 + \frac{\sigma^2}{\sigma_\gamma^2}(\gamma-\mu_\gamma)^2\right]\right\} \tag{8}$$

B. The Noninformative Prior

Following the proposal by Eaves(1983), who used a Jeffreys noninformative prior for a three parameter non-linear regression model, we shall use the Jeffreys prior as our noninformative prior. It is, of course, important to note various concerns about Jeffreys' priors in multiparameter problems that were raised by Box and Tiao (1973), Berger (1985), and Clarke and Wasserman (1993). We shall, wherever appropriate, compare results based on the Jeffreys prior with those based on the vague noninformative prior, $\pi(\alpha,\beta,\gamma,\sigma^2) = \frac{1}{\sigma^2}$.

It is straightforward to derive the Jeffreys prior, since the log likelihood is $-\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}[y_i - \alpha - \beta(x_i - \gamma)^2]^2$, and the Fisher information matrix is, with $w_i = x_i - \gamma$,

$$I(\alpha,\beta,\gamma,\sigma^2) = -\begin{bmatrix} -\frac{n}{\sigma^2} & -\frac{\sum_{i=1}^{n} w_i^2}{\sigma^2} & \frac{2\beta\sum_{i=1}^{n} w_i}{\sigma^2} & 0 \\ -\frac{\sum_{i=1}^{n} w_i^2}{\sigma^2} & -\frac{\sum_{i=1}^{n} w_i^4}{\sigma^2} & \frac{2\beta\sum_{i=1}^{n} w_i^3}{\sigma^2} & 0 \\ \frac{2\beta\sum_{i=1}^{n} w_i}{\sigma^2} & \frac{2\beta\sum_{i=1}^{n} w_i^3}{\sigma^2} & \frac{4\beta\sum_{i=1}^{n} w_i^2}{\sigma^2} & 0 \\ 0 & 0 & 0 & -\frac{n}{2(\sigma^2)^2} \end{bmatrix}$$

The determinant is

$$\frac{2n\beta^2}{\sigma^{10}}\left[n\left(\sum_{i=1}^{n} w_i^4 \sum_{i=1}^{n} w_i^2\right) + 2\left(\sum_{i=1}^{n} w_i^3 \sum_{i=1}^{n} w_i^2 \sum_{i=1}^{n} w_i^1\right) - \left(\sum_{i=1}^{n} w_i^4\right)\left(\sum_{i=1}^{n} w_i\right)^2 - n\left(\left(\sum_{i=1}^{n} w_i^3\right)^2 - \left(\sum_{i=1}^{n} w_i^2\right)^3\right)\right]$$

$$= \frac{2n\beta^2}{\sigma^{10}} a(\gamma) say,$$

which is greater than zero. The Jeffreys noninformative prior is therefore,

$$\pi(\alpha,\beta,\gamma,\sigma^2) = \sqrt{2n}\left[\frac{\beta a(\gamma)}{(\sigma^2)^5}\right]^{1/2}. \tag{9}$$

We will rewrite (9) as

$$\pi(\alpha,\beta,\gamma,\sigma^2) = \frac{\sqrt{2n}(a(\gamma))^{1/2}|\beta|}{\sigma^5}. \tag{10}$$

Recall that the support of $\pi(\alpha,\beta,\gamma,\sigma^2)$ is $-\infty<\alpha<\infty$, $-\infty<\beta<\infty$, $-\infty<\gamma<\infty$, $\sigma^2>0$. Note also that the Jeffreys prior depends on the design points, x_i.

III. BAYESIAN ANALYSIS

For a given prior distribution we shall denote the joint posterior distribution as $P(\alpha,\beta,\gamma,\sigma^2)\propto L(\alpha,\beta,\gamma,\sigma^2;\mathbf{y})\times\pi(\alpha,\beta,\gamma,\sigma^2)$. Treating β and α as nuisance parameters and integrating them out of the joint posterior leads to the following kernels, $K(\alpha,\gamma)$ of the bivariate posterior distribution of α and γ, $P(\alpha,\gamma)\propto K(\alpha,\gamma)$. In these kernels $a_1(\gamma)$, $a_1(\alpha,\gamma)$, and $a_3(\alpha)$ are given in (6).

For the informative prior, we have

$$K(\alpha,\gamma) = \frac{\left(a_1(\gamma) + \frac{1}{\sigma_\beta^2}\right)^{(a+\frac{n-1}{2})} \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma_\alpha^2}(\alpha-\mu_\alpha)^2 + \frac{1}{\sigma_\gamma^2}(\gamma-\mu_\gamma)^2\right]\right\}}{\left[\left(a_3(\alpha) + 2b + \frac{\mu_\beta^2}{\sigma_\beta^2}\right)\left(a_1(\alpha) + \frac{1}{\sigma_\beta^2}\right) - \left(a_2(\alpha,\gamma) + \frac{\mu_\beta}{\sigma_\beta^2}\right)^2\right]^{(a+\frac{n}{2})}}. \tag{11}$$

For the Jeffreys noninformative prior, we obtain

$$K(\alpha,\gamma) = a(\gamma)^{1/2}\int_{-\infty}^{\infty} \frac{|\beta|}{(a_1(\gamma)\beta^2 - 2a_2(\alpha,\gamma)\beta + a_3(\alpha))^{\frac{n+3}{2}}} d\beta \tag{12}$$

In addition, the kernel for the vague noninformative prior, $\pi(\alpha,\beta,\gamma,\sigma^2)=1/\sigma^2$, takes the following form,

$$K(\alpha,\gamma) = \frac{a_1(\gamma)^{\frac{n-2}{2}}}{[a_3(\alpha)a_1(\gamma) - a_2(\alpha,\gamma)^2]^{\frac{n-1}{2}}}. \tag{13}$$

It is of interest to compare the form of the last kernel to the kernel obtained by Buonaccorsi and Gatsonis (1988). In our linear model context and notation, if $\rho=\delta_2/\delta_1$ and if $h(\rho)$is a proper prior for ρ, and treating $\delta_1,\delta_2,\sigma^2$ as nuisance parameters, their posterior kernel for ρ is

$$K(\rho) = h(\rho)\frac{[c_1(\rho)]^{\frac{n-3}{2}}}{[c_1(\rho)c_3 - c_2^2(\rho)]^{\frac{n-2}{2}}}, \tag{14}$$

where c_1,c_2,c_3 are second-, first-, and zero-order polynomials in ρ.

Our Bayesian analysis, for given data and a given prior, consists of 1) plotting the joint posterior kernel of α and γ, 2) obtaining the marginal posteriors of α and γ, either by numerical integration or by Markov Chain Monte Carlo methods. We have used both techniques. In this paper we report results based on implementing the Markov Chain Monte Carlo method by the Metropolis algorithm as in Lee (1992), and 3) reporting from the marginal posteriors, the posterior mean, mode and variance, and 95% highest posterior density (HPD) intervals. Certain kinds of robustness analyses are also recommended. Comparisons are also made between the two noninformative priors. The next section illustrates the Bayesian analysis.

IV. EXAMPLE

We illustrate different aspects of the Bayesian analysis with an example. Consider the experimental data reported in Box and Draper (1987), where the response is the percentage yield of a certain dyestuff and the regressor is the reaction temperature. These data were also studied by Fan *et. al.* (1993). They are given in Table1, and the scatter plot is presented in Figure 1.

Figure 2 shows plots of the bivariate posterior kernels of α and γ for three priors: the noninformative vague and Jeffreys priors, and the NIG prior with hyperparameters $\mu_\alpha = 80$, $\sigma^2_\alpha = 8$, $\mu_\beta = -1.2$, $\sigma^2_\beta = 1.2$, $\mu_\gamma = 60$, $\sigma^2_\gamma = 6$, $a = 2$, and $b = 1$. The parameters are assumed arbitrarily. Note the shape of the kernels is of relatively bivariate normal

The Gibbs sampler was used to obtain the marginal posterior distributions of α and γ. By using the Metropolis algorithm, each marginal distribution is empirically approximated by 30 iterations repeated 1000 times. Table 2 gives posterior means, modes, variances and 95% HPD intervals for the three priors, for α and γ. Note the classical estimates of α and γ are 80.53 and 61.3, respectively.

Table 1. Data on percentage dyestuff yield, y and reaction temperature x.
(From Box and Draper (1987), p.46)

y	x
45.0	56
79.8	60
78.9	61
77.1	63
62.5	65

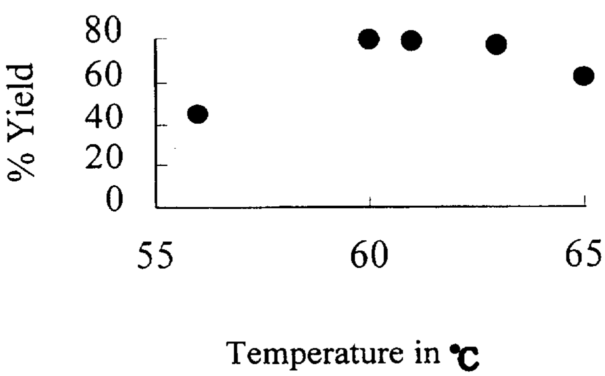
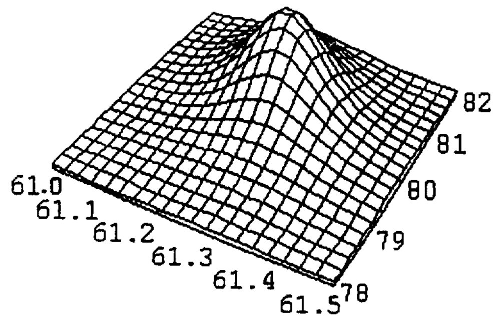
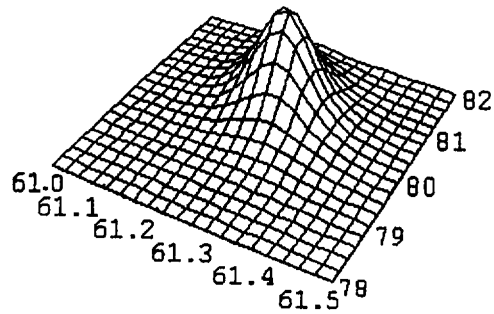


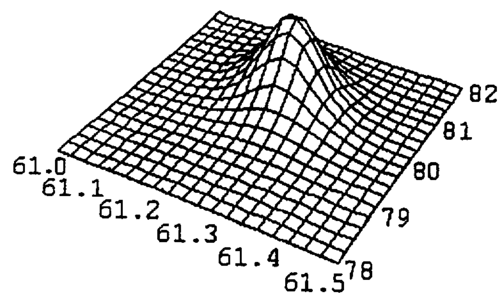
Fig. 1. Scatter plot of y vs. x (Box and Draper data).



(a)



(b)



(c)

Fig. 2. Plots of bivariate posterior kernels of α and γ, Priors: a) Vague, b) Jeffreys, c) NIF (Box and Draper Data)

Robustness in a Bayesian analysis generally deals with sensitivity of the posterior analysis to possible misspecification of the prior distribution, either through misspecification of hyperparameters within a family of priors or through, in some sense, misspecification of the prior family itself. We shall give some results for both

Table 2. Summaries of the Marginal Posteriors of α and γ.

	Prior	Mode	Mean	Variance	95% HPD Interval
	Vague	80.75	80.30	43.79	[75.10, 83.57]
α	Jeffreys	80.50	80.53	0.54	[80.49, 80.58]
	NIG	80.46	80.45	0.44	[80.41, 80.49]
	Vague	61.15	61.40	15.28	[60.68, 62.07]
γ	Jeffreys	61.25	61.23	0.0059	[61.23, 61.24]
	NIG	61.23	61.23	0.0054	[61.23, 61.24]

kinds of robustness analyses.

Figures 3 and 4 contain graphs that enable a robustness analysis of the posterior means, denoted by $\bar{\alpha}$ and $\bar{\gamma}$, i.e., Bayes estimates of α and γ, of the NIG prior. They reveal that in this prior case, the posterior mean of α is robust with respect to μ_γ and σ_γ^2 and that the posterior mean of γ is robust with respect to μ_α and σ_α^2. The graphs show how the posterior means vary with changes in the hyperparameters, μ_α, μ_γ, σ_α^2 and σ_γ^2. The hyperparameters μ_β, σ_β^2, a and b are kept fixed. While considering Figure 3 and 4, recall the original values of the hyperparameters, μ_α and μ_γ are 80 and 60, respectively and the flatter the graph the more robust the Bayes estimate to the corresponding hyperparameter. Figure 3 studies robustness of $\bar{\alpha}$ and $\bar{\gamma}$ to μ_α and σ_α^2, and Figure 4 studies robustness of $\bar{\alpha}$ and $\bar{\gamma}$ to μ_γ and σ_γ^2. For example, in Figure 3a, μ_β= -1.2, σ_β^2 = 0.12, a = 2, b = 1, and μ_γ= 60, σ_γ^2 = 6,, and μ_γ= 30 in Figure 3b. The upper plots show robustness of $\bar{\alpha}$ and the lower plots, robustness of $\bar{\gamma}$. It is interesting to observe from the upper plot of Figure 3a that for μ_γ chosen near the classical estimate of γ, $\bar{\alpha}$ is robust to changes in μ_α when it is far away from the classical estimates of α and σ_α^2 is small. In this case the prior and sample information are consistent. As soon as changes in μ_α shift away from 80 and exceed about 10%, robustness decreases relatively rapidly, and it is especially true for a smaller σ_α^2, where the prior belief is stronger. On the other hand, from the upper plot of Figure 3b where μ_γ = 30 is much smaller than the classical estimate of γ, $\bar{\alpha}$ is not robust to drift in μ_α from its prior value near the classical estimate of α. Similar arguments apply for robustness of $\bar{\gamma}$ from the lower plots of Figure 4, where μ_α= 80 near the classical estimate α in Figure 4a, and μ_γ= 40, in Figure 4b.

We conclude that for the NIG prior, as long as the prior information is close to the sample information, in terms of the prior means of α and γ, the Bayes estimates of and γ are robust provided the respective prior variance is small. The more the prior information conflicts with the sample information, the more the influence of the prior, with the prior virtually dominating for values of μ_α and μ_β extreme to the classical estimates.

Finally, we make a brief comparison of the

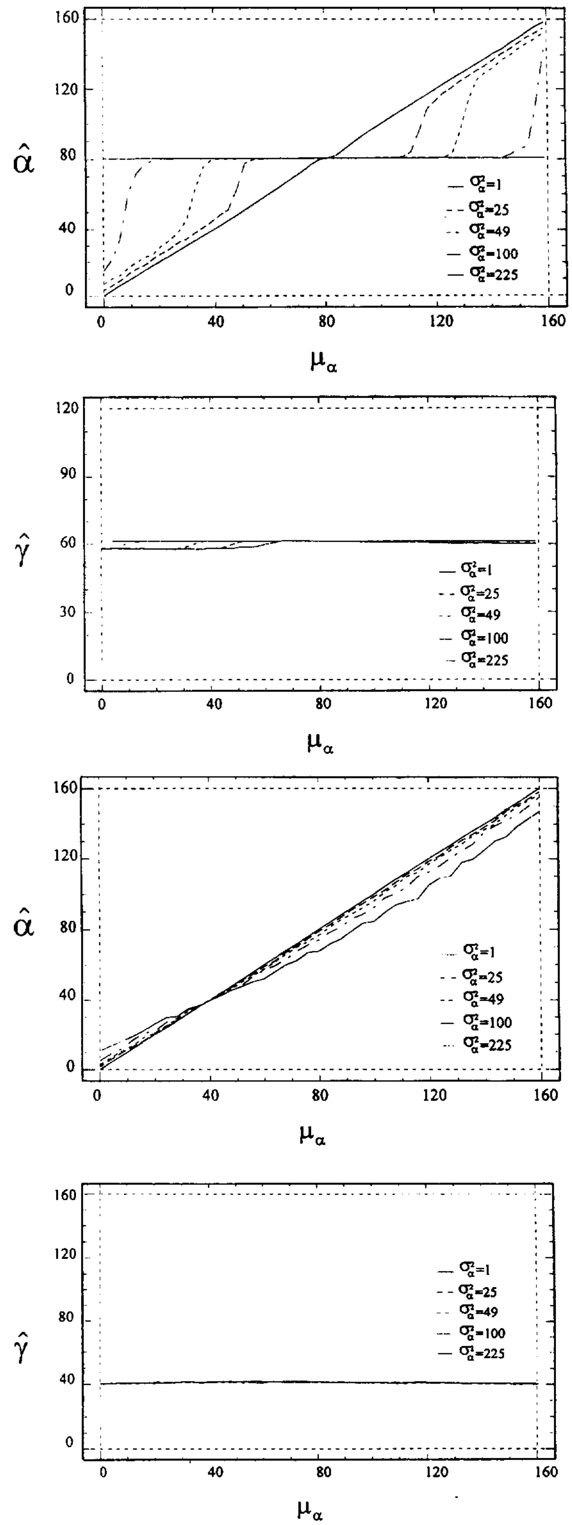


Fig. 3 (a)Robustness of the Posterior Means of α and γ(NIG prior: μ_γ=60, $\sigma_\gamma^2 = 6$, μ_β=-1.2, σ_β^2 = -1.2,, a=2, b=1), (b) Robustness of the Posterior Means of α and γ(NIG prior: μ_γ=30, $\sigma_\gamma^2 = 6$, μ_β=-1.2, σ_β^2 = -1.2, a=2, b=1)

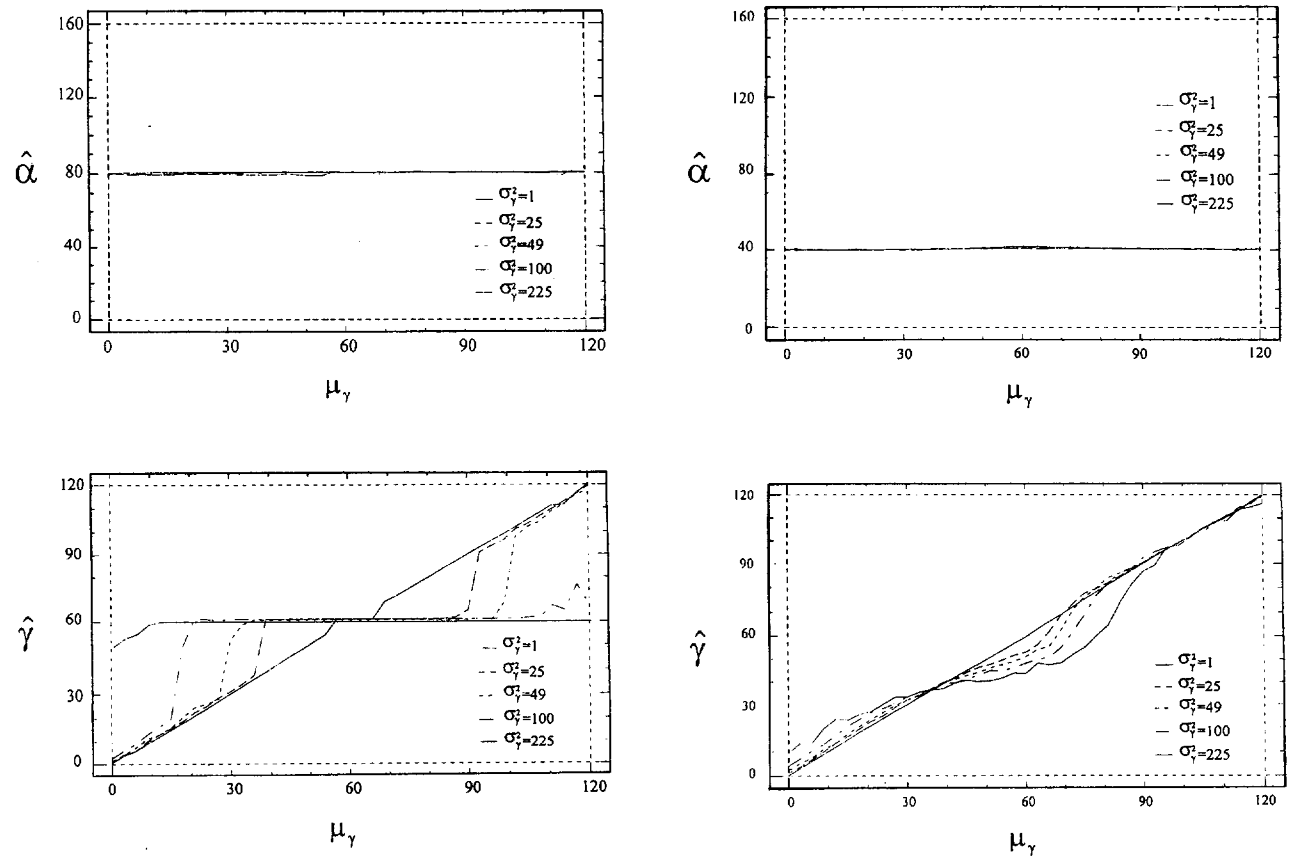


Fig. 4 (a)Robustness of the Posterior Means of α and γ(NIG prior: μ_α=80, $\sigma^2_\alpha = 8$, μ_β=-1.2, $\sigma^2_\beta = 0.12$, a=2, b=1), (b)Robustness of the Posterior Means of α and γ(NIG prior: μ_α=40, $\sigma^2_\alpha = 8$, μ_β=-1.2, $\sigma^2_\beta = 0.12$, a=2, b=1)

noninformative priors to each other based on Table 2. Without entering a debate on a Jeffreys vs. a vague prior, we recommend for this problem the Jeffreys noninformative prior. Note that the posterior variances of α and γ and corresponding HPD intervals for the vague prior are large in comparison to the variance and HPD intervals for the Jeffreys prior. The large variances in the vague prior case may be attributed to the large second moments of the posterior marginals of α and γ. In fact, we have observed heavy tailed marginal for both α and γ. Furthermore, the kernel of the vague prior, in (13), is such that the posteriors may have no finite moments. However no convergence check has been undertaken on this aspect.

V. ACKNOWLEDGMENT

The authors are indebted to the National Science Council ROC for the grant supporting the research undertaken. Sincere gratitude is also expressed to Mr. Hammond K.C. Huang for his tireless assistance in computational support and processing of the manuscript.

REFERENCES

1. Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis, Second Edition*, Springer-Verlag, New York.
2. Box, G.E.P. and Draper, N.R. (1987). *Empirical Model Building and Response Surfaces*, Wiley, New York.
3. Box, G.E.P. and Tiao, G.C. (1973). *Bayesian Inference in Statistical Analysis*, Wiley, New York.
4. Buonaccorsi, J.P. and Gatsonis, C.A. (1988). Bayesian Inference for Ratios of Coefficients in a Linear Model, *Biometrics*, 44, 87-101.
5. Clarke, B. and Wasserman, L. (1993). Noninformative Priors and Nuisance Parameters, *J. Amer. Statist. Assoc.*, 88, 1427-1432.
6. Eaves, D.M. (1983). On Bayesian Nonlinear Regression with an Enzyme Example, *Biometrika*, 70, 373-379.
7. Fan, T.H., Karson M.J. and Wang, H.S., (1996) A Bayesian Estimator of the Optimum for a Single Factor Quadratic Response Model, *Biometrical Journal*, 38, 163-172.
8. Hoadley, B. (1970). A Baysian Look at Inverse Linear Regression, *J. Amer. Statist. Assoc.*, 65, 357-369.
9. Hotelling, H. (1941). Experimental Determination of the Maximum of a Function, *Ann. Math. Statist.*, 12, 20-45.

10. Khuri, A.I. and Cornell, J.A. (1987). Response Surfaces. New York: Marcel Dekker, Inc.
11. Lee, T.M. (1992). *Comparison of Sampling Based Approaches for Bayesian Computations.* unpublished Ph.D. dissertation, The University of Connecticut.
12. Mandel, N.K. (1978). On Estimation of the Maximal Point of s Single Factor Function, *Calcut. Stat. Assoc.* Bull., 27, 119-125.
13. Seber, G.A.F. and Wild, C.J. (1989). *Nonlinear Regression*, Wiley, New York.

單因子二次迴歸駐點之貝氏估計

樊采虹
中央大學統計學系
台灣省中壢市五權里

M.J. Karson
Decision Sciences Department
University of New Hampshire
Durham, New Hampshire, U.S.A.

王晃三　李長洽
中原大學工業工程學系
台灣省中壢市普忠里普仁22號

摘　要

本文處理單因子二次反應函數最佳點之位置與幅度的估計問題。爲求充份運用專家經驗累積之資訊，本論文使用貝氏估計方法，爲此，二種先驗分佈在文中列爲研究對象。其一爲帶訊(informative)先驗分佈，另一則爲無訊(non-informative)先驗分佈。針對一組已知數據，文中提出邊際分佈之形態、後設平均值、衆數、變異數及95%最高後設密度區間並提出韌度分析。此外，二種無訊先驗分佈亦經加以比較。

關鍵詞：二次迴歸，反應曲面，貝氏，估計，最佳化。