CONSTRAINED OPTIMAL DESIGNS FOR POLYNOMIAL REGRESSION

Carl M.-S. Lee, Central Michigan University

1. INTRODUCTION

1.1 The design problem

The problem of taking account of several optimality criteria simultaneously has not been investigated systematically in the past decades. This article deals with the combinations of several criteria by determining the design that is optimal with respect to a particular criterion within a class of designs that achieve at least a minimal quality relative to other criteria. The major concern in this present paper is to introduce four useful constrained optimality criteria and demonstrate their application in the polynomial regression setting.

The univariate polynomial regression of degree m has the form:

$$y_i = \beta_0 + \beta_1 x_i + \ldots + \beta_m x_i^m + e_i, \quad i = 1, \ldots, N, \quad (1.1)$$

where $Y' = (y_1, \ldots, y_N)$ are the N observations of the response observed at points $\{x_1, \ldots, x_N\}$, $\beta'=(\beta_0,\ldots, \beta_m)$ is a vector of unknown parameters and $E' = (e_1, \ldots, e_N)$ is a vector of random errors with the assumption that $\{e_i\}$ are independent and have the mean 0 and common variance σ^2. This model is denoted by P_m. The matrix form of P_m is then $Y = X\beta + E$. A design for mth degree polynomial regression is a probability measure, denoted by ξ, defining on the Borel field generated by the open subsets of χ. The set of all possible design measures is denoted by Ξ. Given a design ξ in Ξ, the associated information matrix of design ξ is defined as $M_m(\xi) = [\mu_{i+j-2}]$, i, j = 1, ..., m+1, where $\mu_{i+j-2} = E(x^{i+j-2})$ with respect to the design ξ. The experimental region χ will be restricted to $\chi = [-1, 1]$ throughout the paper.

Section 2 defines four constrained D- and G- optimality criteria, namely, A-restricted and E-restricted D- and G-optimality criteria, and discuss their applications. In section 3, some general results based on convex analysis are given and a theorem that shows there exists a symmetric constrained D-optimal design with no more than (m+1) support points for any of these constrained D-optimal designs is proved. In section 4, some examples for quadratic polynomial regression are investigated.

1.2 Popular Optimality Criteria

Perhaps the most commonly used family of criteria is what Kiefer (1974) called Φ_{Ap}-optimality criteria, which is defined as the following:

$$\Phi_{Ap}(M_m(\xi)) = \{ 1/(s+1) \quad tr(AM_m^{-1}(\xi)A')^p \}^{1/p} \quad \text{for } 0 \le p \le \infty \qquad (1.2)$$

where A is an (s+1)×(m+1) matrix with rank s+1, $s \le m$, and $A\beta$ is estimable. It is seen that if $A = I_{m+1}$, the identity matrix of rank m+1, then $\Phi_p(M_m(\xi)) = \{1/(m+1) \quad tr(M_m^{-p}(\xi))\}^{1/p}$, which is the well known Φ_p-optimality criterion. The special cases Φ_0, Φ_1 and Φ_∞ are D-, A- and E-optimality criteria. The defects and advantages of these criteria can be found, e.g., in Kiefer (1959), Fedorov(1972), Silvey (1980).

The criteria introduced previously are applicable when the estimation of $A\beta$ is our main concern. There is another group of optimality criteria which provide designs minimizing some function of expected squared error of the fitted curve. The most popular one in this group is the G-optimality,

$$\max_{x\epsilon[-1,1]} \{E(f'(x)\hat{\beta}-f'(x)\beta)^2\},$$

where $f'(x) = (1, x, \ldots, x^m)$. If $\hat{\beta}$ is the least square estimator of β, then, a G-optimal design minimizes the maximum variance function, $var(f'(x)\hat{\beta}) = f'(x)M_m^{-1}(\xi)f(x)\sigma^2$. We shall write $d_m(x, \xi) = f'(x)M_m^{-1}(\xi)f(x)$ and $\overline{d_m}(\xi) = \max_{[-1,1]} d_m(x, \xi)$. The most important characteristics of D- and G-optimal designs are that both are invariant under linear transformation and equivalent to each other (Kiefer (1959, 1974)).

2. CONSTRAINED D- AND G-OPTIMALITY

2.1 A-restricted D- and G-optimality

If we assume normality of errors in the model P_m, then the volume of the confidence ellipsoid of β, which has the form $\{\beta \mid (\beta-\hat{\beta})' M_m(\xi)(\beta-\hat{\beta}) \le w\}$, is minimized by the usual D-optimal design. The constant w depends on the given confidence coefficient and residual sum of squares, and $\hat{\beta}$ is the least square estimate of β. The A-optimal design minimizes the sum of the squared principal axes. To counterpoise these two criteria, we introduce A-restricted D- and G-optimality criteria.

Definition 2.1 A design ξ_{AD} is called an A-restricted D-optimal design if it maximizes $|M_m(\xi)|$ among all designs in the set S_A,

$$S_A = \{\xi \epsilon \Xi \mid tr(M^{-1}(\xi)) \le c\} \qquad (2.1)$$

A design ξ_{AG} is called an A-restricted G-optimal design if it minimizes $\bar{d}_m(\xi)$ among all designs in S_A.

We note that an A-restricted D-optimal design minimizes the volume of the confidence ellipsoid of β among all designs for which the sum of squared principal axes is no larger than a given constant. The constant c must be in the interval $[tr(M_m^{-1}(\xi_A)), tr(M_m^{-1}(\xi_D))]$, where ξ_A is the usual A-optimal design for P_m. If $c < tr(M_m^{-1}(\xi_A))$, then the set S_A is null. On the other hand, if $c > tr(M_m^{-1}(\xi_D))$ then the D-optimal design ξ_D is feasible and thus it is optimal for the constrained problem. Defining the D- and A-efficiencies as follows, respectively :

$$e_m^D(\xi) = \left[\frac{|M_m(\xi)|}{|M_m(\xi_D)|} \right]^{1/(m+1)}, \qquad (2.2)$$

$$e_m^A(\xi) = \frac{tr(M_m^{-1}(\xi_A))}{tr(M_m^{-1}(\xi))}, \qquad (2.3)$$

we see that the design ξ_{AD} maximizes $e_m^D(\xi)$ among all designs for which $e_m^A(\xi)$ is at least ρ, where $\rho = tr(M_m^{-1}(\xi_A))/c$.

2.2 E-restricted D- and G-optimality

It is well known that an E-optimal design minimizes the maximum principal axis of the confidence ellipsoid of β. Geometrically, it makes the shape of the ellipsoid as spherical as possible, and thus, the variances of $\{\hat{\beta}_i\}$ are as close as possible. However, the E-optimality criterion is not differentiable, nor invariant in linear transformation. On the other hand, the design that minimizes the volume of the ellipsoid may have very large variations among the variances of $\{\hat{\beta}_i\}$. Therefore, to counterpoise these two criteria, we introduce E-restricted D- and G-optimality.

Definition 2.2 An E-restricted D-optimal design, ξ_{ED}, is a design that maximizes $|M_m(\xi)|$ among all designs in the set S_E, where

$$S_E = \{\xi \epsilon \Xi \mid \max\{\lambda_i(\xi)\}/\min\{\lambda_i(\xi)\} \le c\}, \qquad (2.4)$$

where $\{\lambda_i(\xi)\}$ are the eigenvalues of $M_m(\xi)$.

An E-restricted G-optimal design minimizes $\bar{d}_m(\xi)$ among all designs in S_E.

Basically, the constant c can be chosen from the interval $[1, \infty)$. However, if $c > \max\{\lambda_i(\xi_D)\}/\min\{\lambda_i(\xi_D)\}$, the design ξ_D is feasible, and hence, it is optimal to the constrained problem. On the other hand, the situation that $c = 1$ is usually unattainable, since otherwise, all $\lambda_i(\xi)$'s are equal and therefore the design must be D-optimal as well as E-optimal. In fact, the lower boundary of c is the minimum of the ratio, $\max\{\lambda_i(\xi)\}/\min\{\lambda_i(\xi)\}$. Thus the values of c are restricted to the interval,

$$[\min_{\xi}\{\max \lambda_i(\xi) / \min \lambda_i(\xi)\}, \; \max\{\lambda_i(\xi_D)/\min \lambda_i(\xi_D)\}].$$

Geometrically an E-restricted D-optimal design minimizes the volume of the confidence ellipsoid of β among all designs of which the ratio of the maximal principal axis to the minimal principal axis is no larger than a given constant. Defining the E-efficiency as the following :

$$e_m^E(\xi) = \frac{\min\{\lambda_i(\xi)\}}{\min\{\lambda_i(\xi_E)\}} \qquad (2.5)$$

where ξ_E is an E-optimal design for the model P_m, we see that an E-restricted D-optimal design maximizes $e_m^D(\xi)$ among all designs with $e_m^E(\xi) \geq \rho$, where $\rho = \max\{\lambda_i(\xi)\}/\, c \min\{\lambda_i(\xi_E)\}$.

3. SOME GENERAL RESULTS

The usual D-optimal design for the model P_m is a design which puts equal mass on $\{\pm 1, x_i, i = 1, \ldots, m-1, x_i \in (-1, 1)\}$, where x_i' s are the zeros of the first derivative of the Legendre polynomial of degree m (Hoel, 1958). This is no longer true for constrained cases. However, the properties of symmetry and finite support still hold for constrained D- and G-optimal designs.

The constrained D- and G-optimal design problems are generalized as the following :

(D) Maximize $|M_m(\xi)|$ subject to $\xi \in S$,

(G) Minimize $\bar{d}_m(\xi)$ among all $\xi \in S$,

where S is a given convex subset of Ξ.

Lemma 3.1 Let $S^{(D)}$ ($S^{(G)}$) be the set of all constrained D- (G-) optimal designs for problem (D) ((G)), then $S^{(D)}$ ($S^{(G)}$) is a convex set.

Proof: Using concavity of $|M(\xi)|$ and convexity of max(.), it is trivial to show $S^{(D)}$ and $S^{(G)}$ are both convex. □

Lemma 3.2
Among all $\xi \in S^{(D)}$ ($S^{(G)}$), there exists a symmetric design ξ_0 in $S^{(D)}$ ($S^{(G)}$).

Proof. See Stigler (1971) □

Lemma 3.3 Given any design for P_m, there exists a design ξ such that both designs share the same moments of order 1, ..., 2m and ξ has support on at most m+1 distinct points.

Proof. See Escobar and Cornette (1983). □

Combining these three Lemmas, we obtain the following Theorem.

Theorem 3.1 If the problem (D) ((G)) has solutions, then, there exists a symmetric and finite supported optimal design for the problem (D) ((G)). The support of this design is on at most m+1 distinct points.

Proof: The existence of a symmetric optimal design follows from Lemma 3.1 and 3.2. Now, if this symmetric design has more than m+1 distinct support points, then, by Lemma 3.3, there exists a design ξ_0 with the same first 2m moments which has exactly m+1 support points. Furthermore, the design ξ_0 can be chosen to be symmetric , since the corresponding odd moments of ξ_0 are all zero. Thus this theorem is proved. □

We note that the sets S_A and S_E are both convex sets. Theorem 3.1 is thus applicable for our constrained problems.

4. CONSTRAINED OPTIMAL DESIGNS FOR P_2

We now consider the special case, P_2. The constrained optimal designs ξ_{AD} ξ_{AG} ξ_{ED}, ξ_{EG} and their performances are investigated. The designs that are interesting to us can be restricted to the set $\{\xi = (d_1, d_2, d_3) \mid d_1 = d_3\}$ with the corresponding supports { -1, 0, 1 }. The information matrix for P_2 is

$$M_2(\xi) = \begin{bmatrix} 1 & 0 & 1-d_2 \\ 0 & 1-d_2 & 0 \\ 1-d_2 & 0 & 1-d_2 \end{bmatrix}.$$

We then obtain $|M_2(\xi)| = (1-d_2)^2 d_2$, $tr(M_2^{-1}(\xi)) = 2/((1-d_2)d_2)$, the eigenvalues of $M_2(\xi)$ are $\{\lambda_1, \lambda_2, \lambda_3\} = [\{(2-d_2) - (5d_2^2 - 8d_2+4)^{1/2}\}/2,\ 1-d_2,\ \{(2-d_2) + (5d_2^2 - 8d_2+4)^{1/2}\}/2]$ with $\lambda_1 \leq \lambda_2 \leq \lambda_3$, and $d_2(x, \xi) = [x^4 + (3d_2-2)x^2 + (1-d_2)]/((1-d_2)d_2)$.

4.1 A-restricted D- and G-optimal designs

The A-restricted D-optimal design problem is to maximize $(1-d_2)^2 d_2$ subject to $2/((1-d_2)d_2) \leq c$ and $d_2 \in [0, 1]$. For c>9, we see that $d_1 = d_2 = d_3 = 1/3$ is the optimal design. For $c \in [8, 9]$, the solution is given at $d_2 = 1/2 - (1/4 - 2/c)^{1/2}$.

The A-restricted G-optimal design problem is to minimize $\max\{1/d_2,\ 2/(1-d_2)\}$ subject to $2/((1-d_2)d_2) \leq c$ and $d_2 \in [0, 1]$. The solution is seen to be the same as the A-restricted D-optimal design. Thus we obtain the following Lemma:

Lemma 4.1 The A-restricted D- and G-optimal design for P_2 is given by

$$d_1 = d_2 = d_3 = 1/3 \qquad \text{for } c > 9$$

$$\begin{cases} d_1 = d_3 = 1/4 + \{(1/4 - 2/c)^{1/2}\}/2 \\ \qquad\qquad\qquad \text{for } c \in [8, 9] \\ d_2 = 1/2 - (1/4 - 2/c)^{1/2} \end{cases} \qquad \square$$

The D-, G- and A-efficiencies of this design are

$$e_2^D = 3\,[\,(2/c)\{1/2 + (1/4 - 2/c)^{1/2}\}/4\,]^{1/3}\ .$$

$$e_2^G = 3 \min [\,1/2 - (1/4 - 2/c)^{1/2},\ 1/4 + \{(1/4 - 2/c)^{1/2}\}/2\,],$$

$$e_2^A = 8/c\ .$$

To compare the usual D-, G- and A-optimal designs with the A-restricted D-optimal design, let us consider the problem with equality constraint, $2/((1-d_2)d_2) = c$, and denote the optimal design as $\xi_0^{(1)}$. Then, the design $\xi_0^{(1)}$ is given by $d_2 = 1/2 - (1/4 - 2/c)^{1/2}$, $d_1 = d_3 = (1-d_2)/2$.

Table 4.1 shows the characteristics of $\xi_0^{(1)}$ and **Figure 4.1** plots the efficiencies of $\xi_0^{(1)}$ versus the constant c.

The design having equal D- and A-efficiency $e_2^D = e_2^A = 0.9805$, denoted as ξ_{AD}^*, is attained at c = 8.1587 with $d_1 = d_3 = 0.285$ and $d_2 = 0.43$. The design having equal G- and A-efficiency $e_2^G = e_2^A = 0.9375$, denoted as ξ_{AG}^*, is attained at c = 8.533 with $d_1 = d_3 = 0.3125$ and $d_2 = 0.375$.

4.2 E-restricted D- and G- optimal designs

The associated E-restricted D-optimal design problem for P_2 is to maximize $(1-d_2)^2 d_2$ subject to $\lambda_3 / \lambda_1 \le c$ and $d_2 \in [0, 1]$, where λ_1 and λ_3 are the smallest and the largest eigenvalues of $M_2(\xi)$, respectively. The ratio of λ_3 / λ_1 for the usual D- and E-optimal designs are $(21 + 5\sqrt{17})/4 \approx 10.4$ and 6, respectively. The smallest ratio among all designs is found to be $3+\sqrt{8} \approx 5.828$. Thus the range of c is $[3 + \sqrt{8},\ \infty]$. The usual D-optimal design is also the E-restricted D-optimal deign when the value $c > (21 + 5\sqrt{17})/4$. The corresponding range of d_2 for $c \in [5.828, 10.4]$ is [1/3, 2/3]. Using the concavity of $(1-d_2)^2 d_2$ for $d_2 \le 2/3$ and the convexity of the set S_E, d_2 has the form given in Lemma 4.2.

The associated E-restricted G-optimal design problem is to minimize max $\{1/d_2,\ 2/(1-d_2)\}$ subject to $\lambda_3 / \lambda_1 \le c$ and $d_2 \in [0, 1]$. By the same argument as the E-restricted D-optimal design problem, the range of c must be limited to $[3 + \sqrt{8},\ \infty]$, and $d_1 = d_2 = d_3 = 1/3$ is the optimal design for $c > (21 + 5\sqrt{17})/4$. For $c \in [3+\sqrt{8},\ (21+5\sqrt{17})/4]$, max $\{1/d_2, 2/(1-d_2)\}$ $= 2/(1-d_2)$. Hence, the solution is the same as the E-restricted D-optimal design. We then obtain the following Lemma:

Lemma 4.2 The E-restricted D- and G- optimal design for P_2 is given by

$$d_1 = d_3 = \frac{(c^2+1) + \{(c+1)^2(c^2-6c+1)\}^{1/2}}{4(c^2+3c+1)}$$

$$d_2 = \frac{(c^2+6c+1) - \{(c+1)^2(c^2-6c+1)\}^{1/2}}{2(c^2+3c+1)}$$

$$\text{for } c \in [3+\sqrt{8},\ (21+5\sqrt{17})/4] \qquad (4.1)$$

$$d_1 = d_2 = d_3 = 1/3 \qquad \text{for } c > (21 + 5\sqrt{17})/4 \qquad \square$$

We note that if we consider the equality constraint only, the E-restricted optimal design for P_2, denoted as $\xi_0^{(2)}$, is given by equation (4.1) for all $c \in [5+\sqrt{8}, \infty]$. The characteristics of $\xi_0^{(2)}$ is given in **Table 4.2** and the graph of efficiencies versus constant c is given in **Figure 4.2**. The D-, G- and E-efficiencies are:

$$e_2^D = (\,27(1-d_2)^2 d_2/4\,)^{1/3}\ ,$$

$$e_2^G = 3 \min \{\,d_2,\ (1-d_2)/2\,\}\ ,$$

$$e_2^E = 2.5\,[\,(2-d_2) - (5d_2^2 - 8d_2 + 4)^{1/2}\,]\ ,$$

where the value of d_2 is given in equation (4.1).

The design having equal D- and E-efficiency $e_2^D = e_2^E = 0.9491$, denoted as ξ_{ED}^*, is given at

$c = 6.93755$ with $d_1 = d_3 = 0.25333$ and $d_2 = 0.49334$. The design having equal G- and E-efficiency $e_2^G = e_2^E = 0.86782$, denoted as ξ_{EG}^*, is given at $c = 8.0945$ with $d_1 = d_3 = 0.28924$ and $d_2 = 0.42153$.

5. DISCUSSION

There are many considerations in choosing a design for a particular experimental situation. Box and Draper (1975) proposed fourteen properties for a good design. A single optimality criterion usually results in optimal designs that satisfy one or very few properties. Thus, it is necessary to take into account of several criteria. It may not be possible to optimize several criteria simultaneously. By optimizing a criterion among the designs that achieve a minimal quality with respect to other criteria, we are able to obtain optimal designs that satisfy more properties as well as reduce the sensitivity of the defects of each criterion. Thus, a constrained optimal design is robust in the sense that it makes good compromises with respect to all involved criteria.

A fundamental problem is how the minimal quality should be determined. This is apparently quite subjective. The goal is to make the efficiencies as high as possible. One way to obtain the best design is to determine a set of constrained optimal designs by giving different minimal qualities, then choose the one that is best for the particular experimental situation. The difficulty of this method is the computational complexity. However, with the help of prior information, if possible, and a highly developed computer system, this should not be a serious problem. If there is no prior information available, then, the one having equal efficiencies for all the involved criteria is a generally good design.

The examples discussed in this article are based on P_2. For higher order polynomial regressions, the closed-form D-restricted D-optimal designs can be obtained by using canonical moment (Lee, 1984) . In fact, Studden (1982) has obtained closed-form solutions for C-restricted D-optimal designs (Stigler,1971). However, computational methods are necessary in general. Bohning (1981) proposed an approach based on the penalty method. Futher investigation on numerical algorithms is needed.

It is interesting that constrained D-optimal designs are equivalent to constrained G-optimal designs for both A-restricted and E-restricted cases for P_2. Unfortunately, this is not generally true (Stigler,1971). However, Kiefer's equivalence theorem (1974) can be extended to constrained optimal design problems by using the Lagrangian theory (Lee 1984).

Table 4.1 A-restricted D- and G-optimal designs, $\xi_0^{(1)}$

c	d_2	$d_1=d_3$	e_2^D	e_2^G	e_2^A
8.0	0.50 (ξ_A)	0.250	0.945	0.750	1.0000
8.159	0.43 (ξ_{AD}^*)	0.285	0.9805	0.855	0.9805
8.4	0.391	0.3045	0.981	0.9135	0.9524
8.533	0.375 (ξ_{AG}^*)	0.3125	0.996	0.9375	0.9375
8.8	0.349	0.3255	0.999	0.976	0.9090
9.0	0.333 (ξ_D)	0.3333	1.000	1.000	0.8890
10.0	0.276	0.362	0.992	0.829	0.8000
13.0	0.158	0.421	0.912	0.475	0.5330
20.0	0.113	0.4435	0.842	0.338	0.4000
30.0	0.072	0.464	0.748	0.218	0.2670
100.0	0.020	0.490	0.509	0.061	0.0800

Table 4.2 E-restricted D- and G-optimal designs, $\xi_0^{(2)}$

c	d_2	$d_1=d_3$	e_2^D	e_2^G	e_2^E
5.83	0.667	0.167	0.794	0.500	0.976
6.00	0.600 (ξ_E)	0.200	0.865	0.600	1.000
6.938	0.4933 (ξ_{ED}^*)	0.2533	0.949	0.760	0.949
7.50	0.4545	0.2727	0.970	0.818	0.909
8.095	0.4215 (ξ_{EG}^*)	0.2892	0.984	0.868	0.868
9.00	0.381	0.3095	0.995	0.928	0.809
10.40	0.333 (ξ_D)	0.333	1.000	1.000	0.731
11.00	0.317	0.3415	0.999	0.950	0.701
20.00	0.183	0.4085	0.938	0.550	0.433
50.00	0.077	0.4615	0.761	0.231	0.189
100.00	0.039	0.4805	0.625	0.118	0.097

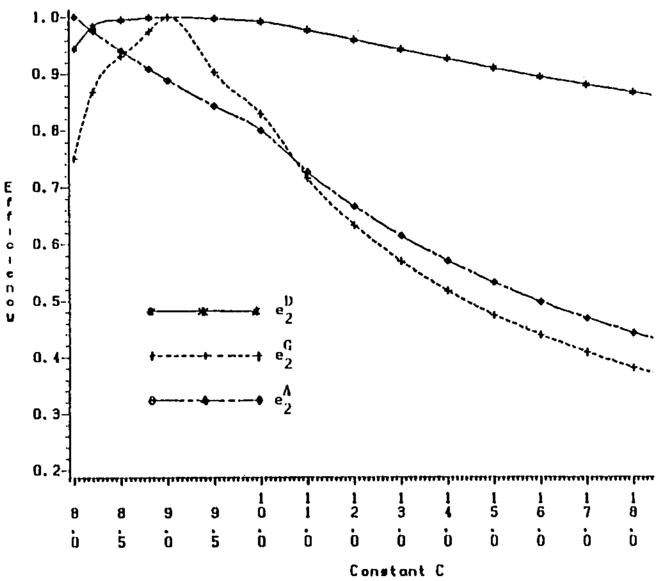

Figure 4.1 Efficiencies of the A-restricted optimal design, $\xi_0^{(1)}$

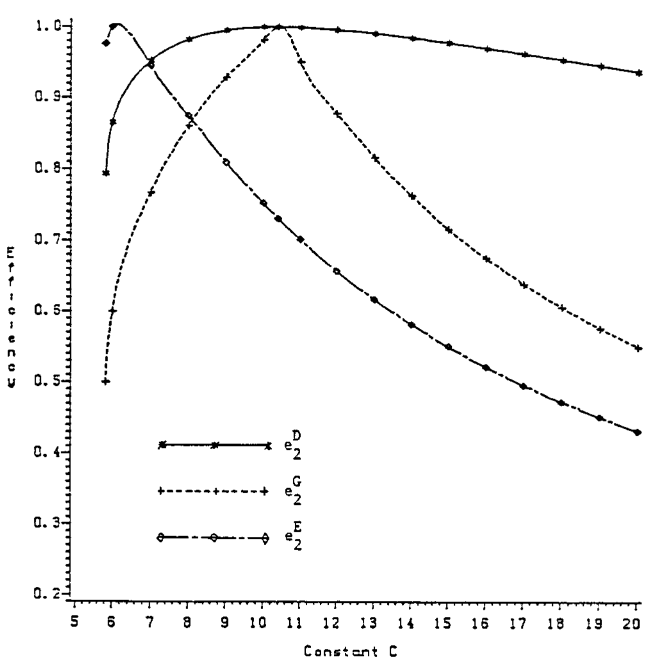

Figure 4.2 Efficiencies of the E-restricted optimal design, $\xi_0^{(2)}$

Acknowledgements :

The author wishes to thank two reviewers, K. Constantine and T. S. Lau, for their time and helpful comments.

REFERENCES

Bohning, D. (1981) On the construction of optimal experimental design : a penalty approach. Math. Oper. Stat. , Series Stat.,12:487-495.

Box, G. E. P. and Draper, N. R.(1975) Robust designs. Biometrika, 62:347-352.

Escobar, L. A. and Cornette, J. (1983) Spacing of information in polynomial regression: a simple solution. Comm. Stat. - T. and M., 12(17):2035-2042.

Fedorov, V. V. (1972) Theory of Optimal Experiments. Academic Press, N.Y.

Hoel, P. G. (1958) Efficiency problems in polynomial estimation. Ann. of Math. Stat. 29 : 1134-1145.

Kiefer, J. (1959) Optimal experimental designs. JRSS , B21:273-319.

Kiefer, J. (1974) General equivalence theory for optimal designs (approximate theory). Annals of Statistics, 2:849-879.

Lau, T. S. (1983) Theory of canonical moments and its applications in poly- nomial regression. Technical Report #83-23. Dept. of Statistics, Purdue University, West Lafayette, Indiana.

Lee, C. (1984) Constrained optimal designs. Ph.D. Thesis, I.S.U. , Ames, Iowa.

Silvey, S. D. (1980) Optimal Design. Chapman and Hall, London.

Stigler, S. (1971) Optimal experimental designs for polynomial regression. JASA, 66:311-318.

Studden, W. J. (1982) Some robust-type D-optimal designs in polynomial regression. JASA, 77:916-921.