Rostock. Math. Kolloq. 31, 95 - 101 (1987)

62K05
62P10

Kurt Frischmuth

An analytical result in experimental design

1. Introduction

In a previous paper /1/ Duchrau a.o. listed the commonly used criteria for optimal experimental designs and described a considerably fast algorithm for the calculation of those designs. A further acceleration of the search procedure may be expected, if the initial guess is made in a more sophisticated way. To this end an auxiliary restricted minimization problem is to be solved, and it turns out that this new optimization can be done very easily. Indeed, the number of unknowns in the auxiliary problem does not exceed five. However, in the present paper our aim is to derive an analytical solution for C_γ-optimality and a certain class of three-parameter growth functions. For this class containing the model of exponential regression as a most important case we dicussed the D-criterion in /2/. The C_γ-criterion seems to be much more complicated to handle with, since the quotient of some determinant and the D-criterion is to be minimized. However, from known numerical results /3/ we luckily guessed a feature of the solution, which enabled us to get the minimizer analytically. Afterwards we prove that our guess was right and, moreover, the only possible one.
The minimizer of the auxiliary problem, the so-called locally C_γ-optimal discrete experimental design, is even of some interest for itself. But our main interest is in finding exact optimal designs. The presented method fails to yield such designs, since differential calculus is used in an essential way. Nevertheless, a comparison of known exact optimal designs (cf. /3/) with the present results suggests that we have found a proper way for the choice of starting values.

2. The criterion

We start with the information matrix

$$M = \frac{\sigma^2}{n} F^T F$$

with $F = \begin{pmatrix} z(x_1) \\ \vdots \\ z(x_n) \end{pmatrix}$ and the rows defined by

$z(x) := \nabla_{\vartheta} g(x, \vartheta_0)$. We assume the given growth function

$$g : [x_1, x_u] \times R^3 \longrightarrow R$$

and ϑ_0 to fulfil i) $z_1(x) \equiv$ const.

ii) $z_2 : [x_1, x_u] \rightarrow [t_1, t_u]$ is a bijection.

iii) There exists a strictly convex smooth function $\varkappa : [t_1, t_u] \rightarrow R$ such that $z_3(x) = \varkappa(z_2(x))$,

where z_1, z_2, z_3 are the components of z.

For designs with a three-point support containing x_1 and x_u the following formula is true (cf. /2/)

$$|M| = n_1 \cdot n_m \cdot n_u \cdot \varphi(x_m). \tag{1}$$

Here x_m is the remaining point of the support and n_1, n_m and n_u are the corresponding frequencies. $\varphi : [x_1, x_u] \rightarrow R^+$ is strictly concave due to iii), further we have $\varphi(x_1) = \varphi(x_u) = 0$. Hence φ takes its unique maximum in (x_1, x_u). Now, C_γ-optimality consists in minimizing the third entry on the diagonal of M^{-1}. We obtain

$$C_\gamma \begin{pmatrix} x_1 & x_m & x_u \\ n_1 & n_m & n_u \end{pmatrix} = \frac{1}{|M|} |F_\gamma^T F_\gamma|, \tag{2}$$

where F_γ is obtained from F by rejecting the third column. Consequently, denoting $\varphi \cdot z_2 =: \psi$,

$$c_\gamma \begin{pmatrix} x_l & x_m & x_u \\ n_l & n_m & n_u \end{pmatrix} = \frac{n_l n_m (t_m - t_l)^2 + n_m n_u (t_m - t_u) + n_l n_u (t_l - t_u)^2}{n_l n_m n_u \psi(t_m)}. \quad (3)$$

As restrictions we have $x_m \in [x_l, x_u]$ and $n_l + n_m + n_u = n$, n_l, n_m, n_u being positive integers, $n \geq 3$ given.

Now, the objective function for the auxiliary problem is obtained by dropping the restriction that n_l, n_m and n_u be integers.

Denoting $\frac{n_l}{n} = \lambda_l$, $\frac{n_u}{n} = \lambda_u$, $t := \frac{t_m - t_l}{t_u - t_l}$, and $h(t) = \psi(t_m)$ we get

$$\frac{n}{(t_u - t_l)^2} \tilde{C}_\gamma = f(t, \lambda_l, \lambda_u) := \frac{\lambda_l (1 - \lambda_l - \lambda_u) t^2 + \lambda_u (1 - \lambda_l - \lambda_u)(1-t)^2 + \lambda_l \lambda_u}{\lambda_l \lambda_u (1 - \lambda_l - \lambda_u) h(t)}. \quad (4)$$

Hence the following optimization problem is stated

$$f(t, \lambda_l, \lambda_u) \overset{!}{=} \underset{P}{\mathrm{Min}}, \quad (5)$$

$$P = \left\{ (t, \lambda_l, \lambda_u) \mid t \in [0,1],\ \lambda_l \in (0,1),\ \lambda_l + \lambda_u \in (\lambda_l, 1) \right\}. \quad (6)$$

3. Conditions for a minimum

The necessary conditions for a local extremum of f are rather cumbersome, and at the first glance there is little hope to solve them analytically. However, numerical experience suggests that the solution makes the denominator of f a maximum with respect to t and the numerator a minimum, simultaneously. Let us assume this for a while as a working hypothesis.

The maximum of the denominator with respect to t is uniquely determined from $h'(t) = 0$. Denoting the solution by t_0, the minimum condition for the numerator reads

$$\lambda_l t_o = \lambda_u (1 - t_o) \quad (7)$$

or equivalently

$$t_o = \frac{\lambda_u}{\lambda_l + \lambda_u}, \quad 1 - t_o = \frac{\lambda_l}{\lambda_l + \lambda_u}. \quad (8)$$

So, if there is a solution satisfying our above working hypothesis, then it is contained in the common curve of the surfaces

$$t = t_o \text{ and } \frac{\lambda_1}{\lambda_u} = \frac{1 - t_o}{t_o} .$$

We parametrize this curve by $\lambda := \lambda_1 + \lambda_u$ and substitute (8) into f.
We obtain along the curve

$$f(t,\lambda_1,\lambda_u) = \frac{\lambda_1\lambda_u}{(\lambda_1+\lambda_u)^2} \frac{\lambda^2+(1-\lambda_1-\lambda_u)(\lambda_1+\lambda_u)}{\lambda_1\lambda_u(1-\lambda_1-\lambda_u)h(t)} = \frac{1}{h(t)} \frac{1}{\lambda(1-\lambda)} . \quad (9)$$

Consequently, f takes its minimum along the curve for $\lambda = \frac{1}{2}$.
Hence the solution is given by

$$t = t_o = h'^{-1}(0), \quad \lambda_1 = \frac{1 - t}{2}, \quad \lambda_u = \frac{t}{2}, \quad (10)$$

provided our hypothesis was true.

4. Verification and uniqueness

We know that the derivatives of f in the directions (1,0,0) and (0,1,1) vanish. Thus it remains to show that the derivative in an arbitrary independent direction vanishes, too. We choose the direction (0,1,0) of λ_1, keeping t and λ fixed. We obtain after some transformations

$$h(t)\frac{d}{d\lambda_1} f\Big|_{\substack{t=const\\ \lambda=const}} = \frac{t^2}{(\lambda-\lambda_1)^2} - \frac{(1-t)^2}{\lambda_1^2} . \quad (11)$$

Substituting (10) into this relation we infer that indeed

$\frac{d}{d\lambda_1} f = 0$. Furthermore, the last equality (11) implies

$$\frac{df}{d\lambda_1}\Big|_{\substack{t=const\\ \lambda=const}} = 0 \quad \text{iff} \quad \frac{t}{\lambda_u} = \frac{1 - t}{\lambda_1} , \quad (12)$$

but this in turn is equivalent to our working hypothesis. Consequently, the relations (10) are necessary and sufficient for ∇f to vanish.

Now a discussion of ∂P is in order. The domain P is bounded by the surfaces $t = 0$, $t = 1$, $\lambda_1 = 0$, $\lambda_u = 0$, $\lambda_1 + \lambda_u = 1$. Regarding the original form of f, we see that the denominator vanishes, while the numerator is nonzero, if exactly one of the above equations holds. We omit the details of the proof that f tends to infinity for arbitrary convergence of the argument to ∂P.
Now we can state the main result as an obvious consequence of the above considerations.

<u>Theorem:</u> If the conditions i), ii) and iii) hold, then the relations (10) describe the absolute minimum of the auxiliary problem (5), (6). The solution is unique.

<u>Remark 1:</u> Analogous considerations yield locally C_α- and C_β-optimal discrete experimental designs provided appropriate modifications of i), ii) and iii) are true.

<u>Remark 2:</u> The assumptions ii) and iii) may be weakened. It suffices to assume i) and

ii') φ takes a unique maximum at $x_m \in (x_1, x_u)$,

iii') $z_2(x_1) < z_2(x_m) < z_2(x_u)$ or

$z_2(x_1) > z_2(x_m) > z_2(x_u)$.

<u>5. Example</u>

For the growth function $g(x,\vartheta) = \alpha + \beta \exp(\gamma x)$, $\vartheta = (\alpha,\beta,\gamma)$, $\alpha > 0$, $\beta < 0$, $\gamma < 0$ we have

$$z(x) = (1, \beta\cdot\exp(\gamma x),\ \beta\cdot x\cdot\exp(\gamma x)). \tag{13}$$

Hence $z_3(x) = x\cdot z_2(x) = \frac{1}{\gamma}\cdot\ln\frac{z_2}{\beta}\cdot z_2(x)$. The assumptions i) and ii) are fulfilled, iii) holds with $\varkappa$ defined by

$$\varkappa(\xi) = \xi\cdot\gamma^{-1}(\ln\xi - \ln\beta).$$

The derivative $\varkappa'(\xi) = \gamma^{-1}(\ln\xi - \ln\beta + 1)$ is obviously increasing. After some less interesting calculations the minimum of the corresponding f is found at

$$t = \frac{\exp\left(\gamma \frac{x_u \exp(\gamma x_u) - x_1 \exp(\gamma x_1)}{\exp(\gamma x_u) - \exp(\gamma x_1)} - 1\right) - \exp(\gamma x_1)}{\exp(\gamma x_u) - \exp(\gamma x_1)} \qquad (14)$$

and λ_1, λ_u from (10).

In /3/ the interval $[x_1, x_u] = [0, 65]$ and $\gamma_o = -0.03$ $(-0.05, -0.07, -0.09)$ were chosen. For these parameters we obtain

t	0.573179	(0.604008, 0.620358, 0.62766)
λ_1	0.213411	(0.197996, 0.189821, 0.18617)
λ_m	0.5	(0.5, 0.5, 0.5)
λ_u	0.286589	(0.302004, 0.310179, 0.31383).

In table 1 we compare the values $\lambda_u n$, $\lambda_1 n$ with the optimum designs for several choices of n (cf. /3/). In the last column the optimum x_m-values are given. From our formulae we obtain $x_m = 22.5516$, $(\gamma_o = -.03)$.

n	$\lambda_1 n$	n_1	$\lambda_u n$	n_u	x_m
3	.64	1	.86	1	21.96
4	.85	1	1.15	1	21.72
5	1.07	1	1.43	1	21.59
6	1.28	1	1.71	2	23.72
7	1.49	2	2.01	2	21.82
8	1.70	2	2.29	2	21.72
9	1.92	2	2.58	3	22.84
10	2.13	2	2.87	3	22.87
11	2.34	2	3.15	3	22.90
12	2.56	3	3.44	3	21.72
13	2.77	3	3.73	4	22.53.

Table 1

In all cases but one rounding gives already the optimum frequencies. Further it is worthwhile to remark that with increasing n the approximation of x_m does not become worser. Hence the overall cost of the algorithm from /1/ with the actual initial guess may be expected to behave in a desirable manner.

Refrences

/1/ Duchrau, P., Frischmuth, K., Rasch, D., and Schimke, E.: Optimum experimental designs in growth curve analysis. Rostock. Math. Kolloq. 30, 93 - 104 (1986)

/2/ Frischmuth, K.: On locally D-optimal experimental designs. (to appear)

/3/ Rasch, D., Rudolph, E., and Schimke, E.: Optimale Versuchspläne in der nichtlinearen Regression. In: Rasch, D. (ed.): Robustheit IV. Probleme der angewandten Statistik 13, 93 - 121 (1985)

received: 05. 01. 1987

Author:

Dr. K. Frischmuth
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
Rostock
DDR-2500