Rostock. Math. Kolloq. 30, 93 - 104 (1986)

62K05
62P10

Petra Duchrau, Kurt Frischmuth,
Dieter Raech, Erhard Schimke

Optimum Experimental Designs in Growth Curve Analysis

1. Introduction

In growth curve analysis often a growing trait is observed at time points t_i (i=1,...,n) (age of the growing organism) and modelled by a random variable $\underline{y}_i = \underline{y}(t_i)$ (random variables are underlined). We assume this situation in the present paper under the additional assumption that

$$\underline{y}_i = f(t_i, \vartheta) + \underline{e}_i, \quad i=1,\dots,n, \quad \vartheta \in \Omega, \quad \dim(\Omega) = p, \qquad (1)$$
$$t_i \in [x_l, x_u], \quad x_l < x_u, \quad n \geq p.$$

The function f in (1) is called the growth function, its graph a growth curve. The parameter vector $\vartheta = (\vartheta_1, \dots, \vartheta_p)$ is unknown and the $\underline{e}_i$ are i.i.d. error terms with expectation $E(\underline{e}_i) = 0$ and the unknown variance $V(\underline{e}_i) = \sigma^2$.

Furthermore the function f may be non-linear in at least one component of ϑ and twice differentiable with respect to ϑ. The aim of this paper is to define some criteria for the optimum choice of the t_i and to present methods for the construction of optimum designs as well as some optimum designs for special functions and parameter configurations.
The most important problems in growth curve analysis are

- the appropriate choice of a growth function f(model choice)
- the estimation of ϑ for a fixed function f from the observations $y(t_i)$
- the constructions of confidence intervals for the components ϑ_j of ϑ (j=1,...,p) and testing hypotheses on the ϑ_j.

We assume in this paper that the function f is given in advance and the n-tupel $(t_1,\dots,t_n) = T$ leads to a unique solution $\hat{\vartheta} = \hat{\vartheta}(y_1,\dots,y_n)$ of

$$\sum_{i=1}^{n}\left[y_i - f(t_i,\hat{\vartheta})\right]^2 = \inf_{\vartheta\in\Omega}\sum_{i=1}^{n}\left[y_i - f(t_i,\vartheta)\right]^2. \qquad (2)$$

The random variable $\hat{\underline{\vartheta}} = \vartheta(y_1,\ldots,y_n)$ defined by (2) is the least squares estimator of ϑ.

The problem of testing a hypothesis was discussed by Rasch and Schimke /10/, /11/.

Under mild regularity conditions it was shown by Jennrich /9/ that

$$\sqrt{n}(\hat{\underline{\vartheta}} - \vartheta) \qquad (3)$$

is asymptotically $N(0_p,\sum)$ distributed where $\sum$ is the limes (for $n\longrightarrow\infty$) of $n\cdot V_A(\hat{\vartheta}\mid\vartheta_0,T)$

with

$$V_A(\hat{\underline{\vartheta}}\mid\vartheta,T) := \sigma^2\left[F'(\vartheta,T)F(\vartheta,T)\right]^{-1}, \qquad (4)$$

$$F(\vartheta,T) := (f_{ij}(\vartheta,T)), \qquad (5)$$

and

$$f_{ij}(\vartheta,T) := \frac{\partial}{\partial\vartheta_j} f(t_i,\vartheta), \quad \begin{pmatrix} i=1,\ldots,n \\ j=1,\ldots,p \end{pmatrix}. \qquad (6)$$

We call $V_A(\hat{\underline{\vartheta}}\mid\vartheta,T)$ the asymptotic covariance matrix of the estimator $\hat{\underline{\vartheta}}$ and look for an n-tupel T which minimizes in some sense this matrix for given n.

We note that the function $T\longrightarrow V_A(\hat{\underline{\vartheta}}\mid\vartheta,T)$ is fully symmetric and therefore we restrict our considerations without loss of generality to the simplex

$$S = \left\{T\in[x_1,x_u]^n \;\middle|\; t_1\leq\ldots\leq t_n\right\} \text{ of ordered n-tuples T.}$$

As we will see later for optimum designs very often the number of different components t_i (i=1,...,n) does not exceed the number p of parameters. Therefore we represent the elements of S by tables D_n containing in the first row the ordered set $\{x_1,x_2,\ldots,x_m\}$ of the m different t_i-values and in the second row the numbers of their occurence. This leads to the following

Definition 1: A mapping D assigning a positive integer n_i to each time point x_i (i=1,...,m) from its domain Dom(D) $\subset [x_1, x_u]$ is called an exact experimental design. We call the domain Dom(D) the support and the number

$$n = \sum_{i=1}^{m} n_i$$

the size of the design D. If the cardinality of the support Dom (D) equals m then we will rate to D as a m-point design. The set of all designs of a given size n is denoted by D(n) and a subscript n indicates that a design D_n belongs to D(n).

It is a matter of course that the relation

$$D_n = \begin{pmatrix} x_1,\ldots,x_m \\ n_1,\ldots,n_m \end{pmatrix} = \mathbb{D}(t_1,\ldots,t_n) \text{ iff}$$

$$\left|\{i \mid t_i = x_1\}\right| = n_1, \quad 1 = 1,\ldots,m,$$

defines a one-to-one correspondence $\mathbb{D}$ between the simplex S and D(n). Frequently we will not distinguish between T and $\mathbb{D}(T)$.

2. Local optimality of exact experimental designs

In the case that f is linearly dependent on ϑ optimum designs are defined as designs minimizing some functional of the covariance matrix of the least squares estimator $\hat{\underline{\vartheta}}$. Those functionals and hence the optimum designs are independent of ϑ.
In the non-linear case we deal with in this paper the covariance matrix is unknown. We therefore base our optimality criteria on the asymptotic covariance matrix introduced in formula (4) using the same functionals as familiar in the linear case. This leads to the unsatisfactory situation that the functionals depend on the parameter ϑ to be estimated. Thus in practice we need some a priori information about ϑ for instance a tolerance interval where the optimum design is to be found for the least favourable case.

Definition 2: Let be given a mapping

$$Z : D(n) \to R^+$$

of the form

$$Z(D_n) := \psi(V_A(\hat{\underline{\vartheta}} \mid \vartheta_0, \mathfrak{M}^{-1}(D_n)) \tag{8}$$

with a certain $\psi: R^{p\times p} \to R^+$. Then D_n^* is called a locally Z-optimum exact experimental design of size n at $\vartheta = \vartheta_0 \in \Omega$ if

$$\inf_{D_n \in D(n)} Z(D_n) = Z(D_n^*). \tag{9}$$

In the following we deal with the functions ψ_r and Z_r, $r = 1,2,\ldots,p,p+1,p+2$, defined by (8) and

$$\psi_r(M) = \begin{cases} m_{rr} & \text{if } r\leq p, \\ |M| & \text{if } r=p+1, \\ tr(M) & \text{if } r=p+2, \end{cases}$$

with $M = (m_{ij})_{i,j=1,\ldots,p}$.

The corresponding designs D^* are called locally C_{ϑ_r}-optimum designs for $r\leq p$, locally D-optimum for $r=p+1$ and locally A-optimum designs for $r=p+2$.

Chernoff /2/ was the first who dealt with the problem of finding D-optimum exact designs in the case of a nonlinear function f and who mentioned the necessity of local optimality.

First analytical results for the criterion of local D-optimality were found by Box and Lucas /1/ for the special case $p = n$. In this case we have $|F'F| = |F|^2$. So the locally D-optimum design D_p^* is equivalently to (9) defined by

$$\sup_{D_p \in D(p)} |F(\vartheta_0, D_p)|^2 = |F(\vartheta_0, D_p^*)|^2. \tag{10}$$

Amongst other results Box and Lucas /1/ found for the support (x_1, x_2, x_3) of the local D-optimum design for the function

$$f(t,\vartheta) = \vartheta_1 + \vartheta_2 e^{\vartheta_3 t} \tag{11}$$

with $\vartheta = \vartheta_0$,

$$x_2 = -\frac{1}{\vartheta_{o_3}} + \frac{x_1 \exp(\vartheta_{o_3} x_1) - x_u \exp(\vartheta_{o_3} x_u)}{\exp(\vartheta_{o_3} x_1) - \exp(\vartheta_{o_3} x_u)},\ x_1 = x_l,\ x_3 = x_u \quad (12)$$

a solution depending only on ϑ_3. So D_3^* is given by

$$D_3^* = \begin{pmatrix} x_1 & x_2 & x_u \\ 1 & 1 & 1 \end{pmatrix} \quad (13)$$

with x_2 defined by (12).
A review of other results is given by Rasch a.o. /12/, most of them are special solutions for the criterion of local D-optimality and discussions of search algorithms.
In this paper we present some new results - analytical as well as numerical solution by search algorithms - which have been found by the authors. A more detailed version of the analytical results will be published in Frischmuth /7/, the search algorithm OPREG is described in Rasch a.o. /12/ and the search algorithm LIESA will be published by Duchrau and Frischmuth /8/.

3. Numerical search

3.1. The algorithm OPREG

Even for linear functions f exact D-optimum designs are constructed by search algorithms (see Cook and Nachtsheim /3/). By Schimke (see Rasch a.o. /12/) the following systematic search was programmed (OPREG). In the first step of this algorithm $l=p+6$ equidistant points were chosen in $[x_1, x_u]$ and a l-tupel

$(n_1^o, n_2^o, \ldots, n_l^o)$

with $\sum_{j=1}^{l} n_j^o = n$ ($n \geq p$ fixed), $0 \leq n_j^o < n$ which minimizes Z_r for this support is determined. This design is the starting point of an iteration. Let $(x_{l_1}, \ldots, x_{l_s})$ be the support of the s-th step ($l_o = l$) of the iteration. Then a new support is defined for the (s+1)-th step by all different v-values

$v_{q1} = x_q - 2\Delta x_s$,

$v_{q2} = x_q - \Delta x_s$,

$v_{q3} = x_q$,

$v_{q4} = x_q + \Delta x_s$,

$v_{q5} = x_q + 2\Delta x_s$

$(q=1,\ldots,l_s)$ with $x_1 \leq v_{qh} \leq x_q$. The ordered set of different v-values is the l_{s+1}-tupel $(x_1,\ldots,x_{l_{s+1}})$. For this l_{s+1}-tupel in the (s+1)-th step the l_{s+1}-tupel

$(n_1^{s+1},\ldots,n_{ls+1}^{s+1})$ is found which minimizes Z_r under the constraint $\sum_{u=1}^{ls+1} n_u = n$. For Δx_s we used

$$\Delta x_s = \frac{x_u - x_1}{2^{s+3}}.$$

If this algorithm leads to vanishing n_1^{s+1}-values the corresponding columns are dropped. OPREG is very succesful in finding the absolute minimum of Z_r in $[x_1,x_o]$ at $\vartheta = \vartheta_o$ but it costs an enormous computing time for $n > 20$. Therefore Duchrau /4/, /5/ developed the algorithm LIESA described in the following.

3.2. The algorithm LIESA

In order to speed up the search procedure we are going now to introduce the assumption that the functional $\tilde{Z}_r := Z_r' \circ ID$ is smooth on S. Note that all of the above mentioned functionals Z_r, (r=1,....,p+2) yield smooth functions on S. Now, our basic algorithm consists of the following steps

- choice of an initial point T^o,
- choice of a suitable direction Δ of descent for Z_r (if possible),
- choice of a stepsize λ.

The main problem is that $\tilde{Z}_r$ takes its minimum on the boundary ∂S. The direction Δ is calculated from grad $\tilde{Z}_r$ using a quasi-Newton-formula (cf. Schwetlick, /13/) and projections on the tangent space to the actual part of the boundary. The stepsize $\lambda \in R^+$ is chosen such that

$$\tilde{Z}_r(T^k + \lambda\Delta) = \min\left\{\tilde{Z}_r(T^k + \alpha\lambda\Delta), \tilde{Z}_r(T^k + \lambda\Delta), \tilde{Z}_r(T^k + \frac{\lambda}{\alpha^*}\Delta)\right\} \qquad (14)$$

where $\alpha \in (0,1)$ and $\alpha^* = \inf\left\{a > \alpha \mid T^k + \frac{\lambda}{a}\Delta \in [x_1, x_u]^n\right\}$.

We start with $\tilde{\lambda} = 1$ and multiply or divide by α or α^*, respectively, until (14) becomes true. If a number λ satisfying (14) is found, then T^{k+1} is defined by

$$T^{k+1} = T^k + \lambda\Delta. \qquad (15)$$

Of course, T^{k+1} may not belong to S, but it always belongs to $[x_1, x_u]^n$. Thus by reordering one gets an element of S with the same values of $\tilde{Z}_r$. In order to avoid redundant calculations and to save storage it is useful to store $\mathbb{D}(T)$ instead of T and to represent Δ analogously. Hence there is no problem with reordering. In a first version of our program (Duchrau /4/) the initial guess was left to the user or made randomly. Unfortunately, the results indicated a strong dependence of the calculated optimum design on the initial guess. Even repeated calculations with random starting points showed a high error rate, taking the results of Rasch a.o. /12/ found by OPREG for $n = 3,\ldots,13$ as a criterion. But the type of errors was very suggestive: writing the results as tables the first row was almost correct in all considered cases. Hence it is quite natural to try some modifications of the second row having the descent finished. To this end the following procedure was introduced (Duchrau /5/).

Let $D_n = \begin{pmatrix} x_1 \ldots x_m \\ n_1 \ldots n_m \end{pmatrix}$, $\sum n_1 = n$. We define a set $U_e(D_n)$

by

$$U_e(D_n) = \left\{ \begin{pmatrix} x_1 \dots x_m \\ \hat{n}_1 \dots \hat{n}_m \end{pmatrix} \;\middle|\; \sum \hat{n}_1 = n,\ \hat{n}_1 \geq 0,\ \|n - \hat{n}\| \leq e \right\}.$$

Here again columns with a zero in the second row are dropped.

We considered the norms $\|n - \hat{n}\| = \max_{0<1\leq m} |n_1 - \bar{n}_1|$ with $e = 1$ and

$\|n-\hat{n}\| = \sum_{I=1}^{m} |n_1 - \hat{n}_1|$ with $e = 2$. In both cases the cardinality of U_e is small, so that the problem

$$\tilde{Z}(D_n^*) = \min_{U_e(D_n)} \tilde{Z}(\tilde{D}_n) \qquad (16)$$

is easy to solve. If $D_n^* = D_n$ is the unique solution of (16) then we stop, otherwise we restart the descent algorithm with D_n^* as a new initial guess.

Of course, there is no guaranty, that this algorithm finds the absolute minimum of Z_r on D(n). On the other hand, tests with the Z_1-criterion for simple exponential regression in all known cases gave the correct results - independently of the choice of T^o. Even for larger sizes n we obtained results in considerably small times.

n	n_1	n_2	n_3	x_1	x_2	x_3	time	m_o
3	1	1	1	0	12,35	65	20"	3
10	1	2	7	0	16,80	65	1'	10
64	5	15	44	0	17,38	65	2'	10
100	7	24	69	0	18,42	65	3'	10

Table 1 Results for locally Z_1 optimum designs
(Exponential regression, $\langle x_1, x_u \rangle = \langle 0,65 \rangle$, $\vartheta^*_{o_3} = -0,05$)

The computation time per step is increasing proportionally to $m^2 n$. For the optimal design the number m usually equals the number of parameters. If no exception of this experience is expected, a starting point with a low number of different components is to be preferred.

4. Some analytical considerations

The main objective of this section is to illustrate the difficulties of the numerical search mentioned above by the example of D-optimality. Further we point out a less expensive way to calculate optimum designs for this criterion and some functions f. To this end we generalize the result (11), (12), (13) of Box and Lucas to the case $n \geq p = m$. Then we have for the determinant of the inverse of $V_A(\hat{\vartheta}^* | \vartheta_o^*, D_n)$ cf. Frischmuth /7/

$$|V_A^{-1}(\hat{\vartheta}^*/\vartheta_o^*, D_n)| = |F'F| = \prod_{l=1}^{m} n_l |M|^2$$

with M being a regular p x p matrix obtained from F by dropping repeated rows. Consequently the problem splits into two independent ones, the maximization of the product $\prod_{l=1}^{m} n_l$ under the constraint $\sum_{l=1}^{m} n_l = n$ and the maximization of $|M|^2$. The first part is quite easy - the product takes its maximum iff $\max_{l,k} |n_l - n_k| \leq 1$. The maximization of $|M|^2$ was managed by Box and Lucas /1/ since in the case $n = p$ the matrix M equals F. So we can formulate

Theorem 1: All locally Z_4 optimum designs with n obeservations at $m = 3$ different time points from $I = \langle t_1, t_u \rangle$ for the function

$$f(t, \vartheta^*) = \vartheta_1^* + \vartheta_2^* e^{\vartheta_3^* t}$$

have the form

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ \left[\frac{n}{3}\right] + \delta_1 & \left[\frac{n}{3}\right] + \delta_2 & \left[\frac{n}{3}\right] + \delta_3 \end{pmatrix} \qquad (17)$$

with x_1, x_2, x_3 from (12) and $n - 3\left[\frac{n}{3}\right]$ of the δ_1 being equal to 1, the remainder vanishing.
Consequently, for $n = 3k$ the solution is unique, otherwise there are 3 designs with the same value of the criterion.

In Frischmuth /7/ it was shown, that all points from S related to designs of the form

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ n_1 & n_2 & n_3 \end{pmatrix} \quad \text{with} \quad x_1, x_2, x_3 \quad \text{from (12)} \quad \text{and} \quad n_1 + n_2 + n_3 = n$$

are critical points for $\tilde{Z}_4$. The eigenvalues of the corresponding matrix of second order derivatives of $\tilde{Z}_4$ were calculated and their signs discussed.
From those results we infer for the exponential growth function $f(t,\vartheta) = \vartheta_1 + \vartheta_2 e^{\vartheta_3 t}$ and the criterion Z_4 the following

Theorem 2: Starting from an arbitrary initial design D_{n_o} with

$$D_{n_o} = \begin{pmatrix} x_1^o & x_2^o & x_3^o \\ n_1^o & n_2^o & n_3^o \end{pmatrix}$$

the algorithm from the previous section reaches a design of the form (17) in not more than $\max\left\{1,\ n_1^o - n_k^o \mid 1,\ k=1,2,3\right\}$ steps.
The time points take their terminal values already in the first step.

Now the question arises, what generalizations of the above theorems are possible. In Frischmuth /7/ the class of growth functions of the form

$$f(t,\vartheta) = \vartheta_1 + \phi(t,\vartheta_2,\vartheta_3) \tag{18}$$

was discussed. Under some regularity conditions for ϕ it can be proven that Theorem 1 holds for f from this class with the only modification that x_2 has to be calculated from a single equation $\mathcal{C}(x_2) = 0$. In general there is no explicit solution of this equation, but $\mathcal{C}$ turns out to be monotonous, hence numerically there is no problem.
Theorem 2 can be generalized as well. Of course, for those situations in which we are able to prove a theorem of this type, we will solve the equation $\mathcal{C}(x_2) = 0$ and use Theorem 1 rather than the search algorithm. Generalizations to other classes of growth functions are more cumbersome. In order to use a theorem analogous to Theorem 1 the solution of a p-dimensional optimi-

zation problem is required. To this end one should put n = m = p and use the previous search algorithm.

5. Final comments and possible directions

The analytical results for D-optimality allow to omit numerical search in a high dimensional space and suggest a good behaviour of our algorithm. On the other hand it should be mentioned that in general the assumption m = p does not lead to the absolute minimum of Z_{p+1}. However, the absolute minimum cannot be "much better" than the relative minimum (17), the ratio being bounded by 1:4.5 for p = 3 (cf. Frischmuth /7/). For other criteria an analytical discussion seems to be much more complicated. For discrete designs, exponential growth and local Z_3 optimality Bock obtained formulae for relative minima, Frischmuth /6/ proved the uniqueness of those minima and gave a generalization to functions of the form (18). Such results are valuable to the choice of starting points for numerical search, especially for large n they help to keep the computation time short.
Our present interest is focussed on the application of LIESA to other nonlinear growth functions like

$$f(t,\vartheta) = \vartheta_1 + \vartheta_2 \arctan \vartheta_3[t+\vartheta_4],$$

$$f(t,\vartheta) = \left[\vartheta_1 + \vartheta_2 e^{\vartheta_3 t}\right]^{\vartheta_4},$$

$$f(t,\vartheta) = \frac{\vartheta_1}{1+\vartheta_2 e^{\vartheta_3 t}},$$

$$f(t,\vartheta) = (\vartheta_1 + \vartheta_2 e^{\vartheta_3 t})^3,$$

and all criteria mentioned above. We further look for more analytical results.

References

/1/ Box, G. E. P., and Lucas, H. L.: Design of experiments in non-Linear estimations. Biometrika 49, 77 - 90 (1959)

/2/ Chernoff, H.: Locally optimal designs for estimating parameters. Ann. Math. Statist. 24, 586 - 602 (1953)

/3/ Cook, R. D., and Nachtsheim, Ch. J.: A comparison of algorithms for constructing exact D-optimal designs. Technometrics 22, 315 - 324 (1980)

/4/ Duchrau, P.: Ein Verfahren zur Suche nach lokal-optimalen Versuchsplänen in der nichtlinearen Regression. Praktikumsarbeit FZT Dummerstorf-Rostock, 1985

/5/ Duchrau, P.: Ein Suchalgorithmus für die Versucheplanung in der nichtlinearen Regression. Jahresarbeit, Sektion Mathematik, Wilhelm-Pieck-Universität, Rostock, 1986

/6/ Frischmuth, K.: An analytical result in experimental design. (to appear)

/7/ Frischmuth, K.: On locally D-optimal experimental designs. (to appear)

/8/ Duchrau, P., and Frischmuth, K.: A numerical search algorithm for experimental designs. (to appear)

/9/ Jennrich, R.: Asymptotic properties of non-linear least squares estimators. Ann. Math. Statist. 40, 633 - 643 (1969)

/10/ Rasch, D., and Schimke, E.: Distribution of estimators in exponential regression - a simulation study. Scand. J. Statist. 10, 293 - 300 (1983)

/11/ Rasch, D., and Schimke, E.: A test for exponential regression and its robustness. In Rasch, D., and Tiku, M.L. (eds.): Robustness of Statistical Methods and Nonparametric Statistics, pp. 104 - 112. Berlin 1984 and Dortrecht 1985

/12/ Rasch, D., Rudolph, P. R., and Schimke, E.: Optimale Versuchspläne in der nichtlinearen Regression. In Rasch, D. (ed.) Robustheit IV. Probl. der angew. Statist. 13, 93 - 121 (1985)

/13/ Schwetlick, H.: Numerische Lösung nichtlinearer Gleichungen. Berlin 1979

received: August 1, 1986

Authors addresses:

stud. math. P. Duchrau;
Dr. K. Frischmuth
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Prof. Dr. D. Rasch;
Dipl.-Math. E. Schimke
Akademie der Landwirtscheftswissenschaften der DDR
Forschungszentrum für Tierproduktion Dummerstorf-Rostock
DDR-2551 Dummerstorf