A functional experimental design factor for optimum discrimination between multireponse models\textsuperscript{1}

D. Uciński and B. Bogacka\textsuperscript{2}

Abstract

We present the theoretical background for calculating optimum designs for non-linear model discrimination. The design support points consist of experimental conditions which vary continuously during the experimental run in addition to other quantitative factors. We implement the theory in a chemical kinetic model discrimination problem, for which we numerically find an optimum design.

1 Introduction

Optimum designs for discrimination between non-linear competing models of dynamic systems are usually considered in terms of quantitative design factors such as time. The optimum design then tells the experimenter at what levels of the factor to take the observations and at what proportions at each level. Various optimality criteria are applied, depending on the type of the models and on the information available prior to the experiment (for example, a distribution or a point prior for the unknown parameters), see [1], [8], [7], [4], [3], [5], [6]. In this work we use the criterion called T-optimality introduced by [1] in a linear model setup. The letter T is used for “testing” as the criterion maximizes the power of the test for the fit of a second model when the first one is assumed to be true.

In [11] we extended this criterion to the case of multireponse models with correlated errors and the variance depending on the unknown parameters. We calculated the optimum set of times and initial concentrations for discrimination between two models describing the kinetics of a complex chemical reaction. It may however be useful to determine other experimental conditions in an optimum way, particularly when observations are easily available at many time instants. We approached this problem in [10] and calculated an optimum temperature profile and the initial concentrations for the discrimination between these models. This can be further improved by finding a set of temperature profiles, together with the initial concentrations, and the proportions of the number of experiments for discrimination.

\textsuperscript{1}This work was supported by Polish-British Research Partnership Programme grant No. 225.

\textsuperscript{2}University of Zielona Góra, Department of Control and Computation Engineering, ul. Podgórna 50, 65-246 Zielona Góra, Poland. E-mail: D.Ucinski@issi.uz.zgora.pl and Department of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London E1 4NS, U.K. E-mail: B.Bogacka@qmul.ac.uk
each optimum set of the experimental conditions. We consider such designs in this paper.

In Section 2 we define the class of competing models and the structure of the experimental conditions yielding the design variables. We also present the theoretical background for calculating the optimum design for model discrimination. In Section 3 we calculate the optimum design for discriminating between two models of a chemical reaction. Finally, in Section 4 we give some general conclusions.

2 Model and Design Optimality Criterion

We assume that the observations can be replicated for different settings of experimental conditions \((T_i, \gamma_i) \in T \times \Gamma, \ i = 1, \ldots, n,\) where \(T\) is the set of all measurable functions satisfying \(T_{\min} \leq T(t) \leq T_{\max}\) for all \(t \in [0, t_f]\), \(T_{\min}\) and \(T_{\max}\) being some fixed constants, and \(\Gamma\) is a set of discrete experimental conditions. We consider such designs in this paper.

We assume that the observations can be replicated for different settings of experimental conditions \((T_i, \gamma_i) \in T \times \Gamma, \ i = 1, \ldots, n,\) where \(T\) is the set of all measurable functions satisfying \(T_{\min} \leq T(t) \leq T_{\max}\) for all \(t \in [0, t_f]\), \(T_{\min}\) and \(T_{\max}\) being some fixed constants, and \(\Gamma\) is a set of discrete experimental conditions. We consider such designs in this paper.

We assume that the observations can be replicated for different settings of experimental conditions \((T_i, \gamma_i) \in T \times \Gamma, \ i = 1, \ldots, n,\) where \(T\) is the set of all measurable functions satisfying \(T_{\min} \leq T(t) \leq T_{\max}\) for all \(t \in [0, t_f]\), \(T_{\min}\) and \(T_{\max}\) being some fixed constants, and \(\Gamma\) is a set of discrete experimental conditions. We consider such designs in this paper.

$$\gamma = \left(\gamma_1, \ldots, \gamma_s\right).$$

We consider the following statistical model of the sum of squares for the lack of fit of the second model defined as follows:

$$y_{kij} = \eta(t_k; T_i(\cdot), \gamma_i) + \varepsilon_{kij}, \quad k = 1, \ldots, K, \ i = 1, \ldots, n; \ j = 1, \ldots, r_i. \quad (1)$$

Here the nonlinear \(m\)-dimensional function \(\eta\) constitutes the true model of the process, which is sampled at given time instants \(t_1, \ldots, t_K\). The terms \(\varepsilon_{kij} \in \mathbb{R}^m\) represent random errors of measurement such that \(E[\varepsilon_{kij}] = 0, E[\varepsilon_{kij}\varepsilon_{pqr}] = \sigma^2 I_m\) if \(k = p, \ i = q \) and \(j = r\) or \(E[\varepsilon_{kij}\varepsilon_{pqr}] = 0_m\) otherwise. Here \(I_m \in \mathbb{R}^{m \times m}\) is the identity matrix, \(0_m\) is the \((m \times m)\)-dimensional matrix of zeros, and \(\sigma^2\) is a positive constant. The additional index \(j\) is necessary when series of \(K\) observations in time are replicated \(r_i > 1\) times for each setting of experimental conditions \((T_i(\cdot), \gamma_i)\). Then the total number of experimental runs is \(\sum_{i=1}^n r_i = N\).

We denote the two competing models by \(\eta_1(t, \vartheta_1; T(\cdot), \gamma)\) and \(\eta_2(t, \vartheta_2; T(\cdot), \gamma)\), where \(\vartheta_1 \in \Theta_1 \subset \mathbb{R}^{p_1}\) and \(\vartheta_2 \in \Theta_2 \subset \mathbb{R}^{p_2}\) are vectors of unknown parameters and \(\Theta_1\) and \(\Theta_2\) are known compacts sets. Assuming that the true model (1) coincides with the response \(\eta_1\) for some known value of \(\theta_1^*\), i.e., \(\eta(\cdot; \cdot, \cdot) \equiv \eta_1(\cdot; \theta_1^*; \cdot, \cdot)\), we may formalize the problem of discriminating between \(\eta_1\) and \(\eta_2\) as maximization of the sum of squares for the lack of fit of the second model defined as follows:

$$T_{12}^0(\xi_N) = \min_{\vartheta \in \Theta_2} \sum_{i=1}^n w_i \sum_{k=1}^K \|\eta(t_k; T_i(\cdot), \gamma_i) - \eta_2(t_k, \vartheta_2; T_i(\cdot), \gamma_i)\|^2, \quad (2)$$

where \(\|\cdot\|\) stands for the Euclidean norm and where the collection of variables

$$\xi_N \overset{\text{def}}{=} \left\{ (T_1(\cdot), \gamma_1), \ldots, (T_n(\cdot), \gamma_n) \right\}$$

is called the normalized exact design of the experiment. The pairs \((T_i(\cdot), \gamma_i)\) are the design support points and the \(w_i\)'s are the weights representing the proportions of replications of each such design point, i.e., \(w_i = \frac{r_i}{N}, \quad \sum_{i=1}^n w_i = 1\).
For the purpose of numerical calculations we introduce a finite-dimensional parametrization of the function $T(\cdot)$:

$$T(t) = \sum_{i=1}^{M} \omega_i \phi_i(t),$$

(4)

where the functions $\phi_i(\cdot)$, $i = 1, \ldots, M$, form a suitable basis in $T$. The design variable $(T(\cdot), \gamma)$ can thus be replaced by $x \overset{\text{def}}{=} (\omega, \gamma) \in \Omega \times \Gamma \subset \mathbb{R}^{M+s}$, where the compact set $\Omega$ consists of all $\omega = (\omega_1, \ldots, \omega_M) \subset \mathbb{R}^M$ for which the bound constraint $T_{\text{min}} \leq T(t) \leq T_{\text{max}}$ holds on the whole interval $[0, t_f]$. Note that the resulting design space $X$ is finite-dimensional. The exact design takes now the standard form with $(M + s)$-dimensional support points $x$.

Widening the class of admissible designs to all probability measures $\xi$ over $X$ which are absolutely continuous with respect to the Lebesgue measure, and denoting the set of all such measures by $\Xi(X)$, we can introduce the following continuous generalization of the T-optimality criterion:

$$T_{12}(\xi) = \min_{\vartheta_2 \in \Theta_2} \Delta_{12}(\xi),$$

(5)

where

$$\Delta_{12}(\xi) = \int_X \sum_{k=1}^{K} ||\eta(t_k; x) - \eta_2(t_k, \theta_2; x)||^2 \xi(dx).$$

Then the following generalized version of the equivalence theorem holds.

**Theorem 2.1** Assume that the minimization problem defined in (5) possesses a unique solution $\vartheta_2^* \in \Theta_2$ for a design measure $\xi^* \in \Xi(X)$. Then a necessary and sufficient condition for $\xi^*$ to be $T_{12}$-optimal is that for each $x \in X$

$$\sum_{k=1}^{K} ||\eta(t_k; x) - \eta_2(t_k, \theta_2^*; x)||^2 \leq T_{12}(\xi^*).$$

(6)

The equality in (6) is attained at all support points of $\xi^*$. Furthermore, the set of all optimal measures $\xi^*$ is convex.

This is a useful tool in checking that a (usually numerically) calculated design is indeed optimum.

### 3 Optimal Temperature Profiles

Consider two chemical reactions: a reaction where substance $A$ changes into substance $B$, which in turn changes into substance $C$, but the first part of the reaction may be reversible, that is $A \rightleftharpoons B \rightarrow C$, with rate constants $k_1^{(1)}$, $k_2^{(1)}$. 


and \( k^{(1)}_3 \) (for the rate of the reverse reaction) and a very similar, but irreversible process \( A \rightarrow B \rightarrow C \), with rates \( k^{(2)}_1 \) and \( k^{(2)}_2 \). The rate constants depend on the process temperature \( T \) in accordance with the Arrhenius relation (c.f. [2])

\[
k_i(T) = \alpha_i e^{-E_i/RT},
\]

where \( \alpha_i \) is the so-called pre-exponential factor, \( E_i \) is the activation energy and \( R \) is the universal gas constant of 1.987 cal/g-mole °K. The parameters \( \alpha_i \) and \( E_i \) are unknown and they are estimated from the experimental data. Their values will be different for different models. We are also interested in estimating the unknown orders of the reaction denoted by \( \lambda_i \).

The concentrations of any or of all of the reactants, \( A, B \) or \( C \), can be measured. The changes in concentrations are governed by ordinary differential equations (for their exact forms see [10]) with the initial conditions for both models given by \( \gamma = (a_0, b_0, c_0)' \), where \( a_0, b_0 \) and \( c_0 \) denote the initial reactant concentrations. We assume that \( \gamma = (a_0, b_0, c_0)' \) is selected to lie in the set \( \Gamma \) being the intersection of \( \Gamma_0 = [a_{0\text{min}}, a_{0\text{max}}] \times [b_{0\text{min}}, b_{0\text{max}}] \times [c_{0\text{min}}, c_{0\text{max}}] \) and the plane defined by the constraint \( a_0 + b_0 + c_0 = 1 \), where \( a_{0\text{min}}, b_{0\text{min}}, c_{0\text{min}}, a_{0\text{max}}, b_{0\text{max}} \) and \( c_{0\text{max}} \) are given non-negative constants.

The solutions of the systems of the differential equations, subject to the initial condition, yield expected responses of the form

\[
\eta_\ell(t, \vartheta; T(\cdot), \gamma) = ([A](t), [B](t), [C](t))', \quad \ell = 1, 2,
\]

where

\[
\vartheta_1 = (\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_3^{(1)}, E_1^{(1)}, E_2^{(1)}, E_3^{(1)}, \lambda_1^{(1)}, \lambda_2^{(1)}, \lambda_3^{(1)}),
\]

\[
\vartheta_2 = (\alpha_1^{(2)}, \alpha_2^{(2)}, E_1^{(2)}, E_2^{(2)}, E_3^{(2)}, \lambda_1^{(2)}, \lambda_2^{(2)}).
\]

represent the unknown parameters of the reversible and irreversible processes, respectively.

In this example the time horizon \( t_f = 20 \) was assumed, and the admissible temperatures were restricted to satisfy the following simple box constraint \( 300 \leq T(t) \leq 550, \quad t \in [0, t_f] \). For the temperature parametrization (4) we apply B-cubic spline basis functions \( \phi_i \) and the coefficients \( \omega_i, \quad i = 1, \ldots, M \), \( M = 8 \), are to be determined as components of the design \( \xi \). Furthermore, for the set of admissible initial concentrations, \( \Gamma_0 = [0.6, 1.0] \times [0.0, 0.2] \times [0.0, 0.2] \) was adopted. Consequently, the support points are 11-dimensional vectors. For the true model, the parameters (7) were assumed to be \( \vartheta_1 = (0.7, 0.2, 0.1, 2000, 2000, 2000, 2000, 2000, 2000, 1.0, 1.0, 1.0) \). In the alternative model, the vector of unknown parameters (8) was assumed to belong to the set

\[
\Theta_2 = \{ \vartheta_2 : \vartheta_{2\text{min}} \leq \vartheta_2 \leq \vartheta_{2\text{max}} \},
\]

where

\[
\vartheta_{2\text{min}} = (0.45, 0.10, 1800, 1800, 1.5, 0.8), \quad \vartheta_{2\text{max}} = (0.95, 0.20, 2200, 2200, 2.5, 1.2).
\]

All the responses were sampled at time moments \( t_k = k, \quad k = 1, \ldots, 20 \). A maximum of two support points were assumed in the sought optimal design.

Finding a T-optimum design reduces to solving a maximin problem, in which we wish to determine the value of the design variable \( \xi \) which maximizes \( T_{12}(\xi) \).
Figure 1: T-optimal temperature profiles (a), (c), and the corresponding responses (continuous and dashed lines denote the responses of the true and alternative models, respectively) (b), (d).
Each evaluation of $T_{12}(\xi)$ involves minimizing $\Delta_{12}(\xi, \vartheta_2)$ with respect to $\vartheta_2 \in \Theta_2$ and so the problem can be cast as the following semi-infinite programming one, c.f. [9]:

$$\begin{align*}
\text{Maximize} & \quad \varsigma \\
\text{subject to} & \quad \Delta_{12}(\xi, \vartheta_2) \geq \varsigma, \quad \vartheta_2 \in \Theta_2.
\end{align*}$$

(10)

The maximization above is to be performed with respect to $\xi \in \Xi(X)$ and the auxiliary variable $\varsigma \in \mathbb{R}$.

In our implementation, a Fortran 95 program based on an exchange method was written and run using the Lahey/Fujitsu compiler v5.7 and the IMSL Fortran 90 MP Library 4.0. The program is available on request from the first author. The results obtained after some forty minutes of calculations on a PC equipped with Pentium 4 CPU 2.40 GHz running Windows 2000 (11 iterations of the relaxation scheme) are presented in Fig. 1. The corresponding optimal weights are $w_1 = 0.215907$ and $w_2 = 0.784093$. The value of $\vartheta_2$ giving the most competing (closest) models is $\vartheta_2^* = (0.7119, 0.1816, 2055, 2128, 2.500, 0.891)$.

4 Conclusions

In this paper we have shown how nonstandard support points consisting of a continuous bounded function and a quantitative design factor on several levels can be introduced in an optimum design. In some experiments, it may be more practical to control a continually changing quantitative factor (such as temperature of a chemical process) and take observations at fixed levels of other design factors (such as time) rather than to observe at some optimally chosen time instants while keeping all other experimental conditions fixed. It would be interesting to widen applicability of the optimum design theory for non-linear models further by including qualitative factors into the design support points as well. Here we considered an optimum optimum design for model discrimination, however, it would be also useful to calculate such designs for parameter estimation.

References


