Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2017) 543-553

ECCOMAS
Proceedia

UNCECOMP 2017
2nd ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Rhodes Island, Greece, 15–17 June 2017

DIFFERENT VIEWS ON ADDITIONAL RANDOM PARAMETERS IN EXPERIMENT DESIGN FOR THERMOPHYSICAL PARAMETERS ESTIMATION (UNCECOMP 2017)

D. Jarušková and A. Kučerová

Czech Technical University in Prague, Faculty of Civil Engineering
Thakurova 7, CZ 166 29 Praha 6, Czech Republic
e-mail: daniela.jaruskova@cvut.cz

Keywords: nonlinear regression, common random factors, nuisance parameters, robust design of experiment

Abstract. *Our contribution deals with a design of experiment problem for estimating thermophysical parameters described by Ruffio et al. (2012). In particular, the partial differential model of the experiment is replaced here by nonlinear regression model where temperature is a dependent variable, time of measurement plays a role of independent variable and thermal conductivities along x and y axes are parameters to be estimated. The goal of statistical inference is to find optimal position(s) of sensor(s) to estimate the parameters of interest as accurately as possible.*

A statistical analysis becomes more complicated when some additional input parameters are random. In our contribution we discuss the situation, where sensor(s) are not placed exactly to designed position(s) due to an error modelled by random nuisance parameters or the so-called random factors. It appears that the very same problem can be viewed from various perspectives so that its solutions are completely different.

Peer-review under responsibility of the organizing committee of UNCECOMP 2017.
doi: 10.7712/120217.5390.16678

empty

1 INTRODUCTION

Recent developments in the field of uncertainty quantification open more possibilities to simulate the nonlinear systems with uncertain input parameters and moreover to design optimized and robust experiments for calibrating the models of such systems. Nevertheless, the difficulties connected to handling the uncertain parameters consist not only in computational requirements, but also in formulation of the problem with uncertain parameters itself, namely with respect to the considered source of uncertainty. The goal of this contribution is to demonstrate two different results obtained due to different perspectives on random parameters in experiment design problem. In particular, the difference originates from two different attitudes of a statistician analyzing the data from an experiment – whether he considers the randomness of parameters in his inverse analysis or not. Different perspectives of the statistician are reflected within the process of experiment design providing different optimal solutions. Particular scenarios are described for an illustrative example of non-stationary heat conduction in two-dimensional square domain inspired by Ruffio et al. [4], where the goal of experiment design problem is to find optimal position of a thermocouple measuring temperature in a set of time steps so as to reduce the uncertainties in estimated thermal conductivities in the two principal directions.

2 EXPERIMENT DESCRIPTION

Our paper was inspired by [4] that presents an analysis of a numerical experiment. In our simplified version the aim of the experiment is to identify two thermal parameters of an orthotropic homogeneous material - more precisely, a constant thermal conductivity λ_x along the $x-$axis and a constant thermal conductivity λ_y along the $y-$axis.

The experiment can be described as follows. A square sample is exposed to a constant and uniform heat flux φ on the left and bottom boundaries, while the right and top edges are insulated. The system obeys the following equation:

$$\begin{aligned} C\frac{\partial T}{\partial t} = \lambda_x \frac{\partial^2 T}{\partial x^2} + \lambda_y \frac{\partial^2 T}{\partial y^2}, \\ 0 \le x \le L_x,\, 0 \le y \le L_y, \\ 0 \le t \le \tau. \end{aligned}$$

The boundary and initial conditions are defined by:

$$\begin{aligned} -\lambda_x \frac{\partial T}{\partial x}(x=0) = \varphi, \quad -\lambda_y \frac{\partial T}{\partial y}(y=0) = \varphi, \\ -\lambda_x \frac{\partial T}{\partial x}(x=L_x) = 0, \quad -\lambda_y \frac{\partial T}{\partial y}(y=L_y) = 0, \\ T(x,y,0) = 0. \end{aligned}$$

One assumes that temperature is measured by one or more sensors at equidistant time points $t = 1, \ldots, 60$ (s) and $L_x = L_y = 0.05$ (m). The number of sensors is fixed. In Ruffio et al. [4] three sensors were considered. Jarušková and Kučerová [2] considered either one or three sensors. The aim of an analysis of the numerical model is to determine position(s) of sensor(s) to identify thermal conductivities λ_x and λ_y as accurately as possible while the thermal capacity

$C = 1700000\ \mathrm{Jm^{-3}K^{-1}}$ and the heat flux $\phi = 25000\ \mathrm{Wm^{-2}}$ are known. Ruffio et al. [4] claim that a solution of the model above may be found analytically in the form of a quickly converging series. In our paper, we assume that a final solution may be quite accurately approximated by the first term only, i.e.,

$$T(t;\lambda_x,\lambda_y;x,y) = \theta_x(t;\lambda_x;C;\varphi;x) + \theta_y(t;\lambda_y;C;\varphi;y),$$

$$\theta_x(t;\lambda_x;C;\varphi;x) = \frac{2\varphi}{\sqrt{C\lambda_x}}\sqrt{t}F\Big(\frac{\widetilde{x}}{\sqrt{t}}\Big), \quad \theta_y(t;\lambda_y;C;\varphi;y) = \frac{2\varphi}{\sqrt{C\lambda_y}}\sqrt{t}F\Big(\frac{\widetilde{y}}{\sqrt{t}}\Big),$$

$$F(z) = \frac{\exp(-z^2)}{\sqrt{\pi}} - z(1 - \frac{2}{\sqrt{\pi}}\int_0^z e^{-v^2}\,dv), \quad z \geq 0.$$

with $\widetilde{x} = (x/2)\sqrt{C/\lambda_x}$, $\widetilde{y} = (y/2)\sqrt{C/\lambda_y}$. Figure 1 presents description of our experiment in a graphical way.

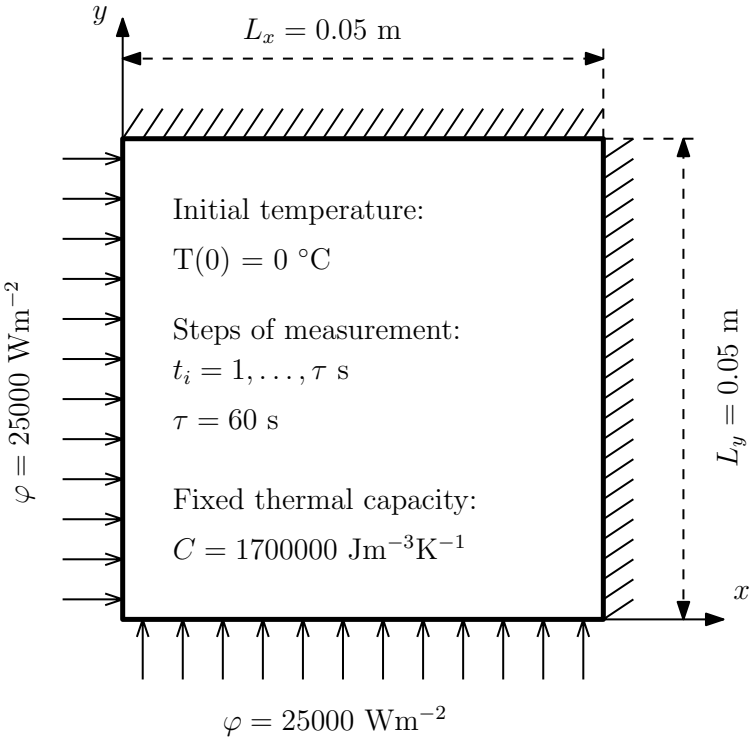

Figure 1: Description of the experiment.

Because a sensor situated at (x, y) measures temperatures $\{Y_i\}$ with measurement errors $\{e_i\}$, we may assume that the behavior of a measured temperature can be modeled with the help of a nonlinear regression:

$$Y_i = T(t_i;\lambda_x,\lambda_y;x,y) + e_i, \quad i = 1,\ldots,60,$$

where $\{e_i\}$ are i.i.d. with $E\,e_i = 0$ and $Var\,e_i = \sigma_m^2 = 0.1^2$. (In the case of three sensors we have 60 measurements in all three positions (x_1, y_1), (x_2, y_2), (x_3, y_3), i.e. 180 measurements in total.) The estimates $\widehat{\lambda}_x$ and $\widehat{\lambda}_y$ may be obtained using the least squares method.

A decision where to place sensor(s) is an experiment design problem in the nonlinear regression, see [5]. We consider two criterions: the D criterion minimizes a determinant of the variance-covariance matrix of $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ and the F criterion minimizes $(Var\widehat{\lambda}_x/\lambda_x^2 + Var\widehat{\lambda}_y/\lambda_y^2)^{1/2}$. Due to independence of measurement errors $\{e_i\}$, the variance of the estimates $\widehat{\lambda}_x$ and $\widehat{\lambda}_y$ converges to zero with increasing number of measurements. As the number of observations is large (here 60 measurements), we may consider the variance of the estimates small enough to replace the nonlinear model by its linear approximation at their true values and the exact variance-covariance matrix by an asymptotically equivalent matrix:

$$\sigma_m^2 \left(F^{*T} F^* \right)^{-1}, \tag{1}$$

where the matrix F^* is a two-column matrix of partial derivatives of T with respect to λ_x and λ_y computed at their true values. It is important to notice that when one sensor is used, the matrix $F^{*T}F^*$ is singular when the sensor (x, y) is placed to the line $y = \sqrt{\lambda_y/\lambda_x}\, x$. Singularity is connected to the fact that we estimate two parameters by only one sensor. In case of three sensors the corresponding asymptotic variance-covariance matrix is singular when all sensors are situated at that line. Positions of sensor(s) yielding a singular approximate variance-covariance matrix used in the F or D criterion will be called inadmissible.

Ruffio et al. [4] as well as Jarušková and Kučerová [2] started an analysis supposing that $\lambda_x = 0.6$ and $\lambda_y = 4.7$. For one sensor there exist two optimal solutions with respect to the D criterion being $x_{op}^{(1)} = 0.0018$, $y_{op}^{(1)} = 0$ and $x_{op}^{(2)} = 0$, $y_{op}^{(2)} = 0.0050$, with the optimal value of the D criterion $3.1 \cdot 10^{-12}$, while there exists only one optimal position $x_{op} = 0.0024$, $y_{op} = 0$ with respect to the F criterion with the optimal value 0.0016, see Case 1 in [2]. Notice that the optimal positions are on the boundaries and relatively close to the point $(0, 0)$. It is clear that temperature is most sensitive to a change in thermal conductivities near the point $(0, 0)$, see Figure 2, but on the other hand a sensor must not be too close to the line $y = \sqrt{\lambda_y/\lambda_x}\, x$ to distinguish between the effect of λ_x and the effect of λ_y. For illustration Figure 2 shows dependence of the euclidean norm of the vector $((\partial T/\partial\lambda_x)(1; \lambda_x, \lambda_y; x, 0), \ldots, (\partial T/\partial\lambda_x)(60; \lambda_x, \lambda_y; x, 0))$ on the position of the x - coordinate of a sensor situated on the bottom boundary. Figure 3 shows the dependence of temperature on time for one sensor and different values of λ_x and λ_y.

The problem becomes more complicated when some input parameters of the model are random as described in the following section.

3 NONLINEAR REGRESSION WITH RANDOM PARAMETERS

In our contribution we assume that a sensor may not be placed exactly to a designed position. Then, instead of being measured at a point (x, y) temperature is measured at a point $(x+\Delta x, y+\Delta y)$, where Δx and Δy are uncertain. In what follows we describe the case of one sensor but the case where more sensors are used is analogue.

3.1 Nonlinear regression with common random factors

In the first scenario we assume that a statistician analyzing the data does not know about the position errors and estimates the parameters of interest λ_x and λ_y minimizing the least squares under the condition that $\Delta x = 0$ and $\Delta y = 0$, i.e.

$$(\widehat{\lambda}_x, \widehat{\lambda}_y) = \operatorname{argmin} \sum_{i=1}^{60} \left(Y_i - T(t_i; \lambda_x, \lambda_y; x, y) \right)^2. \tag{2}$$

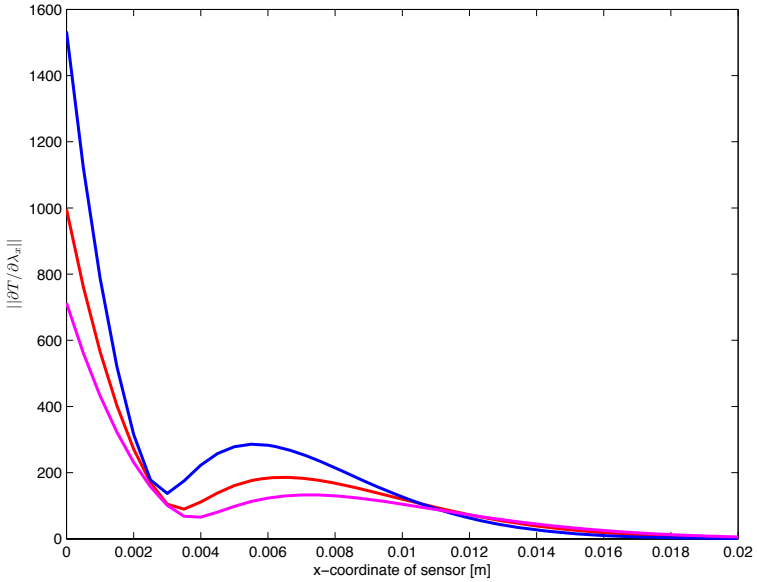

Figure 2: Dependence of $||\partial T/\partial \lambda_x||$ on the x - coordinate of a sensor situated at the bottom boundary for $\lambda_x = 0.6$ (red line), $\lambda_x = 0.45$ (blue line), $\lambda_x = 0.75$ (magenta line).

One experiment consists of 60 temperature measurements and one assumes that we perform many experiments. A position of sensor(s) varies from one experiment to the other according to some prescribed distribution, but in one experiment it does not change and thus it plays a role of the so-called common factor. The least squares estimates $\widehat{\lambda}_x$ and $\widehat{\lambda}_y$ are some functions of $\Delta x, \Delta y, e_1, \ldots, e_{60}$, which will be different in repeated experiments and thus the estimates $\widehat{\lambda}_x, \widehat{\lambda}_y$ will attain different values as well. We are indeed interested in a distribution of $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ induced by the given distribution of $\Delta x, \Delta y, e_1, \ldots, e_{60}$.

We assume that position errors of $\{\Delta x\}$ and $\{\Delta y\}$ in different experiments are (in agreement with [4]) i.i.d. variables distributed according to a zero mean normal distribution with a variance $\sigma^2 = 0.0005^2$. Jarušková and Kučerová [2] pointed out that due to a large number of measurements the effect of measurement errors $\{e_i\}$ on the distribution of $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ is practically negligible in comparison with the effect of random displacements Δx and Δy that remain constant in one experiment.

When designing an optimal experiment, we assume that the distribution of measurement errors as well as distribution of position errors is known to the designer who derives the optimality criterion based on the variance-covariance matrix of $(\widehat{\lambda}_x, \widehat{\lambda}_y)$. If the regression function T is an approximately linear function of λ_x, λ_y, Δx and Δy, then $\widehat{\lambda}_x$ and $\widehat{\lambda}_y$ are approximately linear function of Δx and Δy and $E\,(\widehat{\lambda}_x - \lambda_x) \approx 0$ and $E\,(\widehat{\lambda}_y - \lambda_y) \approx 0$ and a variance-covariance matrix of $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ is approximately equal to:

$$\sigma^2 \left(F^{*T} F^*\right)^{-1} F^{*T} Z^* Z^{*T} F^* \left(F^{*T} F^*\right)^{-1} + \sigma_m^2 (F^{*T} F^*)^{-1}, \tag{3}$$

where F^* is a matrix of partial derivatives of T with respect to components of λ_x and λ_y and Z^* is a matrix of partial derivatives of T with respect to Δx and Δy, both computed at true values of λ_x and λ_y and $\Delta x = 0$ and $\Delta y = 0$.

Ruffio et al. [4] suggest looking for an optimal position by applying either the D criterion or the F criterion to the approximate variance-covariance matrix (3). Then, an optimal position of one sensor with the respect to the F criterion is $x_{op} = 0.0165$, $y_{op} = 0.0158$ with the optimal

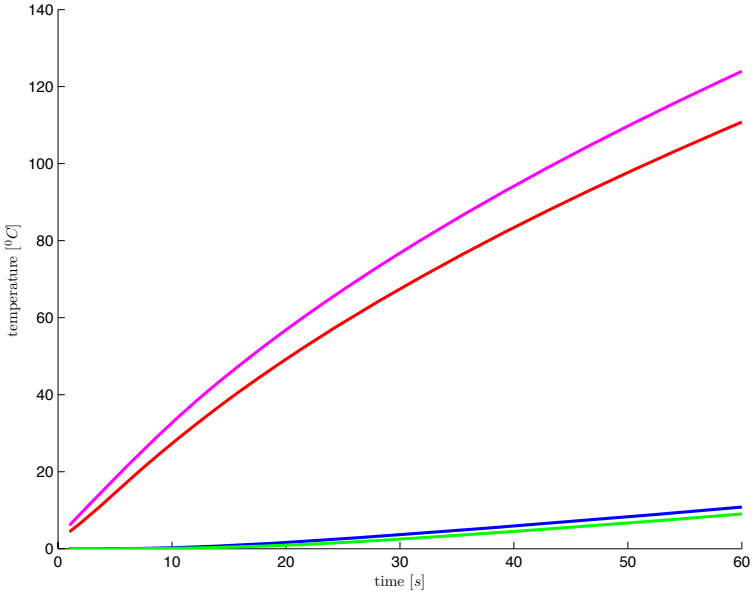

Figure 3: Dependence of temperature on time. Red line corresponds to the position $(x = 0.0018, y = 0)$ and $\lambda_x = 0.6, \lambda_y = 4.7$, magenta line corresponds to the position $(x = 0.0018, y = 0)$ and $\lambda_x = 0.45, \lambda_y = 3.2$. Blue line line corresponds to the position $(x = 0.0165, y = 0.0158)$ and $\lambda_x = 0.6, \lambda_y = 4.7$, green line corresponds to the position $(x = 0.0165, y = 0.0158)$ and $\lambda_x = 0.45, \lambda_y = 3.2$.

value $F = 0.04$ and an optimal position with the respect to the D criterion is $x_{op} = 0.0167$ and $y_{op} = 0.0310$ with a value $D = 8.4 \cdot 10^{-5}$. The optimal position(s) of sensor(s) were found numerically on a grid $[0 : 0.0001 : 0.05] \times [0 : 0.0001 : 0.05]$. When positions of three sensors are to be found we have to use a faster method that provides us with suboptimal solutions. In our contribution we applied an evolutionary GRADE algorithm extended by niching strategy CERAF, see [1].

However, it may happen that for some (x, y) the regression function is a strongly nonlinear function and then the only way how to get an approximate distribution of $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ is by Monte Carlo simulations, see [2].

3.1.1 Note on the specific distribution of random parameters

Ruffio et al. [4] assume that random displacements Δx and Δy are independent, distributed according to a normal distribution $N(0, 0.0005^2)$. Clearly, if the distribution of Δx and Δy is normal, then for all possible designed positions (x, y) there exist displacements Δx and Δy such that the sensor is shifted to the line of inadmissible positions. This is not a serious problem because we may, for instance, replace the two-dimensional normal distribution by its trimmed version being zero outside a circle $(\Delta x^2 + \Delta y^2)^{1/2} \leq 3 \cdot 0.0005$. This changes the original distribution very slightly. When temperature is measured by one sensor and we calculate values of the criteria using the variance-covariance matrices obtained by Monte Carlo simulations and by numerical minimization of the least squares, we consider a narrow band around the line to be a set of inadmissible solutions and we look for an optimal solution outside this band.

A more serious problem might occur when a designed position of a sensor is at one of the boundaries or close to it. This often happens when three or more sensors are used, see [4] or [2]. We suggested a way how to replace a normal distribution by a distribution that shifts sensors to an interior of the heated square sample, see [2], but we do not know how much such

a replacement changes an optimality of the solutions.

3.2 Nonlinear regression with nuisance parameters

In the second scenario, it is supposed that the statistician analyzing the data from one experiment knows that real positions of sensors might be different from the designed positions due to random displacements Δx and Δy. He considers them being nuisance parameters and estimates them together with the parameters of interest (λ_x, λ_y).

The most simple way for estimating all unknown parameters is again the least squares method:

$$(\widehat{\lambda}_x, \widehat{\lambda}_y, \widehat{\Delta x}, \widehat{\Delta y}) = \operatorname{argmin} \sum_{i=1}^{60} \Big(Y_i - T(t_i; \lambda_x, \lambda_y; x + \Delta x, y + \Delta y)\Big)^2. \tag{4}$$

The estimates are indeed maximum likelihood estimates because measurement errors are normal.

We introduce the four column matrix G^* of the partial derivatives of the function T with respect to λ_x, λ_y, Δx and Δy. The matrix

$$\frac{1}{\sigma_m^2} G^{*T} G^* \tag{5}$$

is an expected Fisher information matrix under the assumption that λ_x, λ_y, Δx and Δ_y are true values of the parameters. Its inverse $V_{(x,y)}$ is asymptotically equivalent with the variance-covariance matrix of the least squares estimates $(\widehat{\lambda}_x, \widehat{\lambda}_y, \widehat{\Delta x}, \widehat{\Delta y})$ supposing that the position (x, y) was chosen. Denote $V^{\lambda}_{(x,y)} = V_{(x,y)}(1:2, 1:2)$ an approximate variance-covariance matrix of $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ (for n large). The matrix $V^{\lambda}_{(x,y)} = V^{\lambda}_{(x,y)}(\lambda_x, \lambda_y, \Delta x, \Delta y)$ depends on true values of parameters λ_x and λ_y as well as on true values of displacements Δx and Δy.

For a chosen position of a sensor (x, y) the matrix $V^{\lambda}_{(x,y)}(0.6, 4.7, \Delta x, \Delta y)$ expresses variability of the estimates $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ when a position of a sensor remain in all experiments the same $(x + \Delta x, y + \Delta y)$ (supposing the true values $\lambda_x = 0.6$ and $\lambda_y = 4.7$.) The elements of the matrix $V^{\lambda}_{(x,y)}(0.6, 4.7, \Delta x, \Delta y)$ are functions of $(\Delta x, \Delta y)$ so that if position errors differ the approximate variance-covariance matrix may be also different. If the position errors $(\Delta x, \Delta y)$ are random variables so are the elements of the matrix $V^{\lambda}_{(x,y)}(0.6, 4.7, \Delta x, \Delta y)$. It corresponds to the situation that many experiments are performed and the position errors vary randomly from one experiment to the other. It is not a big surprise that under an assumption that $\lambda_x = 0.6$, $\lambda_y = 4.7$ and $\Delta x = 0$, $\Delta y = 0$, the optimal positions can be found on the boundaries. The optimal positions with respect to the D criterion are $x^{(1)}_{op} = 0.0044$, $y^{(1)}_{op} = 0$ and $x^{(2)}_{op} = 0$, $y^{(2)}_{op} = 0.0123$ with the D value $2.3 \cdot 10^{-8}$, while there is only one optimal position with a respect to the F criterion $x_{op} = 0.057$, $y_{op} = 0$ with the F value 0.012.

Table 1 summarizes values of D and F criteria for different situations.

4 Optimality criteria using prior knowledge on parameters

In nonlinear regression an exact as well as an approximate variance-covariance matrices of estimates of parameters of interest depend generally on the very same parameters that are to be estimated. Therefore, we always need some prior information, which commonly takes a form of some expert guess about the estimated parameter values, which can be then used for evaluation of approximate variance-covariance matrices.

	D-optimality	F-optimality
no position errors	$3.1 \cdot 10^{-12}$	0.0016
position errors in scenario 1	$8.4 \cdot 10^{-5}$	0.04
position errors in scenario 2 when $\Delta x = 0, \Delta y = 0$	$2.3 \cdot 10^{-8}$	0.012

Table 1: Values of optimality criteria achieved when nuisance parameters are present and/or considered or not present at all.

4.1 Prior information in terms of feasible intervals - worst case approach

Sometimes, we have some prior information in the form of feasible bounded intervals for estimated parameters. For instance, we may know that $0.45 \leq \lambda_x \leq 0.75$ and $3.2 \leq \lambda_y \leq 6.2$. Our desire is to find a position that is the best for the "worst" values of λ_x and λ_y in a sense that for any other λ_x and λ_y from the corresponding intervals we get smaller values of the considered criteria. For any admissible position (x, y) we calculate the values of the F or D criterion over a dense grid in $(\lambda_x, \lambda_y) \in [0,45; 0.75] \times [3.2; 6.2]$. The value of the F_R criterion is a maximum of the corresponding F values and similarly the value of the D_R criterion is a maximum of the corresponding D values. The position (x, y) with the smallest F_R value, respectively the smallest D_R value, is optimal. When a position of one sensor is to be designed and the values of the F and D criterion are calculated using the approximate variance-covariance matrix (3), we consider a set $\{[x, y] \in [0; 0.05] \times [0; 0.05]; \sqrt{3.2/0.75}\, x \leq y \leq \sqrt{6.2/0.45}\, x\}$ to be inadmissible and we are looking for an optimal position outside this set. Following the first scenario and using (3) the optimal position with the respect of the F_R criterion is $x_{op} = 0.0155$ and $y_{op} = 0.0260$. The optimal value $F_R = 0.102$ is attained for $\lambda_x = 0.45$ and $\lambda_y = 3.2$. The optimal position with the respect of the D_R criterion is $x_{op} = 0.0174$ and $y_{op} = 0.0307$. The optimal value $D_R = 1.94 \cdot 10^{-4}$ is attained for $\lambda_x = 0.45$ and $\lambda_y = 6.2$. If the variance-covariance matrices are obtained by Monte Carlo simulations and numerical minimization of the least squares, we look for an optimal solution outside the set $\{[x, y] \in [0; 0.05] \times [0; 0.05]; \sqrt{3.2/0.75}(x - 0.0016) \leq y \leq \sqrt{6.2/0.45}(x + 0.0016)\}$, see Figure 4.

Notice, that the "worst" case occurs for values of λ_x and λ_y in vertices of the set $[0,45; 0.75] \times [3.2; 6.2]$. We call the criteria based on the worst case approach the robust criteria. Ruffio et al. [4] suggest to calculate the maximum over the edge centers according to the so-called star design.

Similarly, if in the second scenario we knew that position errors take values from some bounded intervals we might take the worst case approach not only for the parameters of interest but also for the nuisance parameters, i.e. to minimize either the D or F criterion based on $V^{\lambda}_{(x,y)}(\lambda_x, \lambda_y, \Delta x, \Delta y)$ with respect to (x, y) for the maximum value with respect to $(\lambda_x, \lambda_y, \Delta x, \Delta y)$ over the corresponding hypercube. We did not take this approach as Ruffio et al. [4] assume that the intervals for feasible values of Δx and Δy are theoretically unbounded.

4.2 Prior information in terms of probability densities

Very often we may formulate our prior information in terms of some given probability densities. For instance, we may suppose that λ_x, λ_y, $(\Delta x, \Delta y)$ are independent with respective densities $p_1(\lambda_x), p_2(\lambda_y), f(\Delta x, \Delta y)$.

We may employ this knowledge by considering the mean value of the D or F criterion based

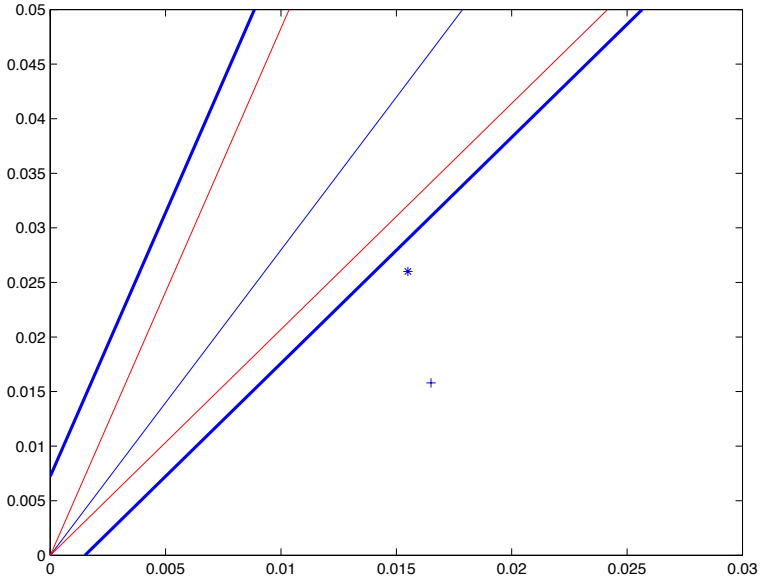

Figure 4: Set of inadmissible solutions and the optimal solutions using the F and F_R criteria.

on $V^{\lambda}_{(x,y)}(\lambda_x, \lambda_y, \Delta x, \Delta y)$ with respect to prior distribution of λ_x, λ_y, $(\Delta x, \Delta y)$. In the other words we weight the values of the criterion over the prior distribution:

$$\int \Big(D(\widehat{\lambda}_x, \widehat{\lambda}_y)\Big) p_1(\lambda_x) p_2(\lambda_y) f(\Delta x, \Delta y) d\,(\Delta x)\, d\,(\Delta y)\, d\,\lambda_x\, d\,\lambda_y, \tag{6}$$

see (23) in [3], or

$$\int \Big(F(\widehat{\lambda}_x, \widehat{\lambda}_y)\Big)\, p_1(\lambda_x) p_2(\lambda_y) f(\Delta x, \Delta y) d\,(\Delta x)\, d\,(\Delta y)\, d\,\lambda_x\, d\,\lambda_y, \tag{7}$$

see (20) in [3]. The functions $p_1(\lambda_x)$ and $p_2(\lambda_y)$ might be for instance the densities of uniform distributions on the corresponding intervals. The first criterion is thus a weighted average of determinants of an approximate variance-covariance matrix of the estimates $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ where the "weights" correspond to the prior distribution of $\lambda_x, \lambda_y, \Delta x, \Delta y$. Similarly, the second criterion is a weighted average of the sum of standardized approximate variances of $\widehat{\lambda}_x$ and $\widehat{\lambda}_y$ with the same weights.

For any $\lambda_x \in [0.45, 0.75]$ and any $\lambda_y \in [3.2, 6.2]$ the optimal positions with respect to the D and F criteria are either on the left or bottom boundary. Therefore, it seems reasonable to look for optimal positions with respect to (6) and (7) on these boundaries. However, it is again not clear how to define the distribution of Δx and Δy on the boundaries. After a discussion we decided to consider a trimmed normal distribution on the interval $[-c\,\sigma; c\,\sigma]$ for one coordinate and a trimmed half normal distribution on the interval $[0, c\,\sigma]$ for the second one. Such a prior distribution ensures that the "incorrectly" placed sensor will be always inside the heated sample. The numerical optimization shows that the new optimal positions are shifted from the points obtained by least square methods presented in the subsection 3.2 to points that are slightly more distant from the origin $(0, 0)$. (A size of shift depends on c and a numerical approximation of integrals in (6) and (7).)

5 CONCLUSION

In our paper we deal with a nonlinear regression describing dependence of temperature measured by one or more sensors on time. Additionally to the parameters of interest, here the thermal conductivities, the regression function contains some other input parameters, e.g. displacements of sensors from their design positions $\{(\Delta x(j), \Delta y(j))\}$. The distribution of these input parameters is assumed to be known. The aim of statistical inference is to design optimal position(s) of sensor(s).

It is shown that in the scope of mathematical statistics the problem may be viewed in two different ways at least. In the first model the thermal conductivities are estimated under an assumption that the positions of the sensors are exact. In reality this assumption is true only in average but not for all performed experiments. On the contrary the positions vary from one experiment to the other according to an assumed distribution. In the second model the statistician knows that the sensors positions may be not exact and estimates their x and y coordinates together with the thermal conductivities.

In our first model it is assumed that one experiment is performed many times and the considered criteria express a variability of the estimates $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ in many repeated experiments. A crucial point in our example is that the number of measurements is large. Therefore, in the first model variability of thermal conductivities estimates is to a large extent determined by a variability of common random factors, i.e. random displacements of sensors. To minimize the estimates variability, the optimal positions of sensors are chosen in a way that their random displacements affect the estimates of the parameters of interest (thermal conductivities) as little as possible but they also have to be sensitive to the estimated parameters.

In the second model we are interested in a variability of the estimates $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ in one experiment where the parameters $\{(\Delta x(j), \Delta y(j))\}$ are fixed but unknown and they are estimated together with the parameters of interest $(\widehat{\lambda}_x, \widehat{\lambda}_y)$. The variability of $(\widehat{\lambda}_x, \widehat{\lambda}_y)$ is caused by measurement errors. How large the variability is, depends not only on the true values of (λ_x, λ_y), but also on the true values of $\{(\Delta x_j, \Delta y_j)\}$. In case the true values of the input parameters $\{(\Delta x_j, \Delta y_j)\}$ vary from one experiment to the other experiment according to a known distribution, the value of any criterion based on the asymptotic variance-covariance matrix is a random variable. We may choose a characteristic of its distribution to be a design criterion. In the subsection 4.2 we have chosen a mean value.

In practice the true values of the parameters of interest are also unknown but we have some prior knowledge about their values. Therefore, we may consider a criterion being the mean with respect to the distribution of nuisance parameters as well as to the parameters of interest. In our example with one sensor the variance-covariance matrix depends on the true parameters values relatively slightly and the optimal design is not too far from the optimal design for a model with the parameters fixed at their mean values. Indeed, we could also consider a robust version of the criteria with the respect to (λ_x, λ_y), as it is described in the subsection 4.1.

As in two considered models a different type of "errors" in estimating thermal conductivities are considered, it is not surprising that the optimal designs are completely different.

This work was supported by the Czech Science Foundation under grants 15-09663S and 15-07299S.

REFERENCES

[1] A. Ibrahimbegovic, C. Knopf-Lenoir, A Kučerová, P. Villon, Optimal design and optimal control of structures undergoing finite rotations and elastic deformations. *Int. J. Numer. Meth. Engng*, **61**, 2428 – 2460, 2004.

[2] D. Jarušková and A. Kučerová, Estimation of thermophysical parameters revisited from the point of view of nonlinear regression with random parameters, *Int. Heat Mass Transfer*, 2016. http://dx.doi.org/10.1016/j.ijheatmasstransfer.2016.10.044

[3] K. Chaloner and I. Verdinelli, Bayesian experimental design: A review. *Statistical Science*, **10**, 273 – 304, 1995.

[4] E. Ruffio, D. Saury, D. Petit, Robust experiment design for the estimation of thermophysical parameters using stochastic algorithms. *Int. Heat Mass Transfer*, **55**, 2901 – 2915, 2012.

[5] G. A. F. Seber and C. J. Wild, *Nonlinear regression*. John Wiley ans Sons, 1989.