Journal of Statistical Planning and Inference 186 (2017) 28–41

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

K-optimal designs for parameters of shifted Ornstein–Uhlenbeck processes and sheets

Sándor Baran *

Faculty of Informatics, University of Debrecen, Kassai street 26, H-4028 Debrecen, Hungary

ARTICLE INFO

Article history:
Received 19 April 2016
Received in revised form 31 January 2017
Accepted 9 February 2017
Available online 21 February 2017

Keywords:
D-optimality
K-optimality
Optimal design
Ornstein–Uhlenbeck process
Ornstein–Uhlenbeck sheet

ABSTRACT

Continuous random processes and fields are regularly applied to model temporal or spatial phenomena in many different fields of science, and model fitting is usually done with the help of data obtained by observing the given process at various time points or spatial locations. In these practical applications sampling designs which are optimal in some sense are of great importance. We investigate the properties of the recently introduced K-optimal design for temporal and spatial linear regression models driven by Ornstein–Uhlenbeck processes and sheets, respectively, and highlight the differences compared with the classical D-optimal sampling. A simulation study displays the superiority of the K-optimal design for large parameter values of the driving random process.

1. Introduction

Continuous random processes and fields are regularly applied to model temporal or spatial phenomena in many different fields of science such as agriculture, chemistry, econometrics, finance, geology or physics. Model fitting is usually done with the help of data obtained by observing the given process at various time points or spatial locations. These observations are either used for parameter estimation or for prediction. However, the results highly depend on the choice of the data collection points. Starting with the fundamental works of Hoel (1958) and Kiefer (1959), a lot of work has been done in the field of optimal design. Here by a design we mean a set $\boldsymbol{\xi} = \{x_1, x_2, \ldots, x_n\}$ of distinct time points or locations where the investigated process is observed, whereas optimality refers to some prespecified criterion (Müller, 2007). In case of prediction, one can use, e.g., the Integrated Mean Square Prediction Error criterion, which minimizes a functional of the error of the kriging predictor (Baldi Antognini and Zagoraiou, 2010; Baran et al., 2013) or maximize the entropy of observations (Shewry and Wynn, 1987). In parameter estimation problems, a popular approach is to consider information based criteria. An A-optimal design minimizes the trace of the inverse of the Fisher information matrix (FIM) on the unknown parameters, whereas E-, T- and D-optimal designs maximize the smallest eigenvalue, the trace and the determinant of the FIM, respectively (see, e.g., Pukelsheim, 1993; Abt and Welch, 1998; Pázman, 2007). The latter design criterion for regression experiments has been studied by several authors both in uncorrelated (see, e.g., Silvey, 1980) and in correlated setups (Müller and Stehlík, 2004; Kiseľák and Stehlík, 2008; Zagoraiou and Baldi Antognini, 2009; Dette et al., 2015). However, there are several situations when D-optimal designs do not exist, for instance, if one has to estimate the covariance parameter(s) of an Ornstein–Uhlenbeck (OU) process (Zagoraiou and Baldi Antognini, 2009) or sheet (Baran et al., 2015). This deficiency can obviously be corrected by choosing a more appropriate design criterion. In case of regression models a recently introduced approach, which optimizes the condition number of the FIM, called K-optimal design (Ye and Zhou, 2013), might be a

* Fax: +36 52 512996.
E-mail address: baran.sandor@inf.unideb.hu.

http://dx.doi.org/10.1016/j.jspi.2017.02.003
0378-3758/

reasonable choice. K-optimal designs try to minimize the error sensitivity of experimental measurements (Maréchal et al., 2015) resulting in more reliable least squares estimates of the parameters. However, one can also consider the condition number of the FIM as a measure of collinearity (Rempel and Zhou, 2014), thus minimizing the condition number avoids multicollinearity.

In contrast to the standard information based design criteria, the condition number (and the corresponding optimization problem) is not convex, only quasiconvexity holds (Maréchal et al., 2015). Hence, finding a K-optimal design usually requires non-smooth algorithms. Ye and Zhou (2013) consider polynomial regression models and solve the K-optimal design problem with nonlinear programming, whereas in Rempel and Zhou (2014) simulated annealing is applied. In this class of models K-optimal designs are quite similar to their A-optimal counterparts. Further, Maréchal et al. (2015) investigate Chebyshev polynomial models and suggest a two-step approach to find a probability distribution approximating the K-optimal design.

Further, one should also mention that K-optimal design is invariant to the multiplication of the FIM by a scalar, so it does not measure the amount of information on the unknown parameters. Besides this, K-optimality obviously does not have meaning for one-parameter models, but in this case multicollinearity does not appear either.

All regression models where K-optimality has been investigated so far consider uncorrelated errors, but there are no results for correlated processes. In the present paper we derive K-optimal designs for estimating the regression parameters of simple temporal and spatial linear models driven by OU processes and sheets, respectively, and compare the obtained sampling schemes with the corresponding D-optimal designs. Both increasing domain and infill equidistant designs are investigated and the key differences between the two approaches are highlighted. Our aim is to give a first insight into the behavior of K-optimal designs in a correlated setup, but many results presented here can be generalized to models with different base functions and/or correlation structures (see, e.g., Näther, 1985; Dette et al., 2016). This is a natural direction for further research.

2. Ornstein–Uhlenbeck processes with linear trend

Consider the stochastic process

$$Y(s) = \alpha_0 + \alpha_1 s + U(s) \tag{2.1}$$

with design points taken from a compact interval $[a, b] \subset \mathbb{R}$, where $U(s),\ s \in \mathbb{R}$, is a stationary OU process, that is a zero mean Gaussian process with covariance structure

$$\mathsf{E}\, U(s)U(t) = \frac{\sigma^2}{2\beta}\exp\big(-\beta|s-t|\big), \tag{2.2}$$

with $\beta > 0,\ \sigma > 0$. We remark that $U(s)$ can also be represented as

$$U(s) = \frac{\sigma}{\sqrt{2\beta}}\mathrm{e}^{-\beta s}\mathcal{W}\big(\mathrm{e}^{2\beta s}\big), \tag{2.3}$$

where $\mathcal{W}(s),\ s \in \mathbb{R}$, is a standard Brownian motion (see, e.g., Shorack and Wellner, 1986; Baran et al., 2003). In the present study the parameters β and σ of the driving OU process U are assumed to be known. However, a valuable direction for future research will be the investigation of models where these parameters should also be estimated. We remark that the same type of regression model appears in Müller and Stehlík (2004), where the properties of D-optimal design under a different driving process are investigated.

For model (2.1), the FIM $\mathcal{I}_{\alpha_0,\alpha_1}(n)$ on the unknown parameters α_0 and α_1 based on observations $\big\{Y(s_i),\ i = 1, 2, \ldots, n\big\},\ n \geq 2$, equals

$$\mathcal{I}_{\alpha_0,\alpha_1}(n) = H(n)C(n)^{-1}H(n)^\top, \quad \text{where } H(n) := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ s_1 & s_2 & \cdots & s_n \end{bmatrix},$$

and $C(n)$ is the covariance matrix of the observations (see, e.g., Xia et al., 2006; Pázman, 2007). Without loss of generality, one can set the variance of U to be equal to one, which reduces $C(n)$ to a correlation matrix. Due to the particular structure of $C(n)$ resulting in a special form of its inverse (see Appendix A.1 or Kiseľák and Stehlík, 2008), a short calculation shows that

$$\mathcal{I}_{\alpha_0,\alpha_1}(n) = \begin{bmatrix} L_1(n) & L_2(n) \\ L_2(n) & L_3(n) \end{bmatrix},$$

with

$$\begin{aligned} L_1(n) &:= 1 + \sum_{i=1}^{n-1}\frac{1-p_i}{1+p_i}, \qquad L_2(n) := s_1 + \sum_{i=1}^{n-1}\frac{s_{i+1}-s_ip_i}{1+p_i}, \\ L_3(n) &:= s_1^2 + \sum_{i=1}^{n-1}\frac{(s_{i+1}-s_ip_i)^2}{1-p_i^2}, \end{aligned} \tag{2.4}$$