Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 126 (2016) 3952–3967

stochastic processes and their applications

www.elsevier.com/locate/spa

Estimation of low-rank covariance function

V. Koltchinskii[a], K. Lounici[a], A.B. Tsybakov[b,*]

[a] Georgia Institute of Technology, 686 Cherry St, Atlanta GA 30332, USA
[b] Laboratoire de Statistique, CREST-ENSAE, 3, av. P.Larousse, 92240 Malakoff, France

Received 31 March 2015; accepted 9 July 2015
Available online 29 April 2016

Abstract

We consider the problem of estimating a low rank covariance function $K(t, u)$ of a Gaussian process $S(t), t \in [0, 1]$ based on n i.i.d. copies of S observed in a white noise. We suggest a new estimation procedure adapting simultaneously to the low rank structure and the smoothness of the covariance function. The new procedure is based on nuclear norm penalization and exhibits superior performances as compared to the sample covariance function by a polynomial factor in the sample size n. Other results include a minimax lower bound for estimation of low-rank covariance functions showing that our procedure is optimal as well as a scheme to estimate the unknown noise variance of the Gaussian process.

Keywords: Gaussian process; Low rank covariance function; Nuclear norm; Empirical risk minimization; Minimax lower bounds; Adaptation

1. Introduction

Let $X(t), t \in [0, 1]$ be a Gaussian process satisfying the following stochastic differential equation:

$$dX(t) = S(t)dt + \sigma dW(t), \quad t \in [0, 1], \tag{1}$$

* Corresponding author.
E-mail addresses: vlad@math.gatech.edu (V. Koltchinskii), klounici@math.gatech.edu (K. Lounici), alexandre.tsybakov@ensae.fr (A.B. Tsybakov).

http://dx.doi.org/10.1016/j.spa.2016.04.006
0304-4149/

where W is the standard Brownian motion, $\sigma > 0$ is the noise level, and

$$S(t) = \sum_{k=1}^{r} \sqrt{\lambda_k}\xi_k\varphi_k(t), \quad t \in [0, 1].$$

Here ξ_k are i.i.d. standard Gaussian random variables independent of the Brownian motion W, $\{\varphi_k\}_{k=1}^r$ are unknown orthonormal functions in $L_2[0, 1]$, possibly, with $r = \infty$, and the coefficients $\lambda_k > 0$ are unknown and such that $\sum_{k=1}^r \lambda_k < \infty$. The value of r is also unknown.

Assume that we observe n i.i.d. copies $X_1(t), \ldots, X_n(t)$ of the process $X(t)$. In this paper, we study the problem of estimation of the covariance function of the stochastic process $S(\cdot)$,

$$K(t, u) = \mathbb{E}(S(t)S(u)) = \sum_{k=1}^{r} \lambda_k\varphi_k(t)\varphi_k(u), \quad t, u \in [0, 1], \tag{2}$$

based on the observations $\{X_1(t), \ldots, X_n(t), t \in [0, 1]\}$. If $r = \infty$, the sum in (2) is understood in the sense of $L_2([0, 1] \times [0, 1])$-convergence. In short, (1) is a model of a "signal" (Gaussian stochastic process S) observed in a Gaussian white noise and the goal is to estimate the covariance of the signal based on a sample of such observations.

Statistical estimation of covariance functions has already received some attention in the literature. However, somewhat different setting was considered where the trajectories $X_i(\cdot)$ are observed at discrete time locations:

$$Y_{i,j} = S_i(T_{i,j}) + \sigma\eta_{i,j}, \quad 1 \le i \le n,\ 1 \le j \le m,$$

where S_i are i.i.d. copies of S, $\eta_{i,j}$ are i.i.d. $\mathcal{N}(0, 1)$ and, for each i, the points $T_{i,j}$, $1 \le j \le m$, are equispaced in the interval $[0, 1]$ or independent random variables with uniform distribution on $[0, 1]$. In this setting, Yao et al. [14] proposed a local smoothing estimation procedure assuming that the trajectories $X_i(\cdot)$ are well approximated by the projection on the linear span of functions $\varphi_1, \ldots, \varphi_k$ for some known fixed k chosen by cross-validation. This procedure is computationally intensive as it requires to compute the eigenvalues and the inverse for n distinct $m \times m$ empirical covariance matrices of the trajectories X_i, $1 \le i \le n$, at each of the cross-validation steps. The results in [14] provide theoretical guarantees for estimation of the covariance function and its eigenfunctions under the condition that the previous approximation is sufficiently precise. Hall et al. [3] consider the same methodology and study the effect of the sampling rate on the estimation rate of the eigenfunctions. In a similar framework, Bunea and Xiao [2] propose a simpler procedure to estimate the eigenfunctions and obtain theoretical guarantees on the estimation error. Their approach involves a dimension reduction step where the selection of the relevant eigenfunctions is performed by thresholding the eigenvalues of a correctly constructed empirical covariance matrix. In a similar setting, Bigot et al. [1] consider the estimation of the covariance matrix of the process S at sample points rather than that of the covariance function. This problem can be reduced to multivariate regression and Bigot et al. [1] develop a model selection approach to it resulting in some oracle inequalities.

Noteworthy, strong regularity conditions are usually imposed on the eigenfunctions φ_k in the existing literature. In [3] the eigenfunctions are assumed to admit bounded derivatives of order at least two. In addition, the optimal bandwidth choice in the local smoothing approach used in [3,14] requires the knowledge of smoothness degree of the eigenfunctions. In [2], the eigenfunctions are assumed to be continuously differentiable with bounded derivatives, the