AN ALTERNATIVE POINT OF VIEW ON LEPSKI'S METHOD

LUCIEN BIRGÉ

Université Paris VI

Lepski's method is a method for choosing a "best" estimator (in an appropriate sense) among a family of those, under suitable restrictions on this family. The subject of this paper is to give a nonasymptotic presentation of Lepski's method in the context of Gaussian regression models for a collection of projection estimators on some nested family of finite-dimensional linear subspaces. It is also shown that a suitable tuning of the method allows to asymptotically recover the best possible risk in the family.

AMS subject classifications: Primary: 62G07, Secondary 62G20.

Keywords and phrases: Adaptation, Lepski's method, Mallows' C_p, optimal selection of estimators.

1 Introduction

The aim of this paper is threefold. First we want to emphasize the importance of what is now called "Lepski's method", which appeared in a series of papers by Lepski (see Lepskii, 1990, 1991, 1992a and b). Then we shall present this method from an alternative point of view, different from the one initially developed by Lepski. Finally we shall introduce some generalization of the method and use it to prove some nice properties of it which, as far as we know, have not yet been considered, even by its initiator.

Let us first give a brief and simplified account of the classical method of Lepski. This method has been described in its general form and in great details in Lepskii (1991) and the interested reader should of course have a look at this milestone paper. Here we shall content ourselves to consider the problem within the very classical "Gaussian white noise model". According to Ibragimov and Has'minskii (1981, p.5), it has been initially introduced as a statistical model by Kotel'nikov (see Kotel'nikov, 1959). Since then, it has been extensively studied by many authors from the former Soviet Union (see for instance Ibragimov and Has'minskii, 1981, Pinsker, 1980, Efroimovich and Pinsker, 1984) and more recently by Donoho and Johnstone (1994a and b, 1995, 1996) and Birgé and Massart (1999), among many other references. Although not at all confined to this framework, the method has been often considered in the context of the Gaussian white noise model for the sake of simplicity. This model can be described by a stochastic differential equation of the form

$$dY_\varepsilon(t) = s(t)\,dt + \varepsilon\,dW(t), \qquad \varepsilon > 0, \quad 0 \le t \le 1, \tag{1.1}$$