

Robust Identification of AR Models Based on Empirical Risk Minimization

Vojislav Filipović

Abstract: In this paper the problem of parameters estimation of AR process in the presence of noise uncertainty is considered. Off-line estimation method is based on empirical risk minimization (method important in theory of learning and generalization). It is supposed that properties of stochastic process is not known exactly. More realistic assumptions is that we have apriori information about the class of distributions to which stochastic process belongs. In such situation philosophy of robust statistics is used. For estimates consistency and asymptotic normality is proved.

Keywords: Empirical risk, robust statistics, consistency, asymptotic normality

I. INTRODUCTION

The representation of a stochastic process by model dates back to an idea that was originated by Yule [1]. Next step was the observation that stationary stochastic process is decomposed into the sum of a deterministic and purely random component [2]. The deterministic component is perfectly predictable from the infinite past and has the form of finite combination of sinusoids. The purely random component can be represented as the output of linear system driven by white noise. Very large class of a stochastic processes can be represented by autoregressive models (AR). That kind of models is related to fundamental theorem in the decomposition of time series [2] and parameters of model can be computed by solving Yule-Walker equations (system of linear equations) [3]. The AR models have applications in: geophysics, speech processing, radars, weather prediction and in many others areas [4], [5]. It is extremely important prediction that theory of stochastic processes and recursive procedure for estimation will play in close future very important role in the development of quantum computers [6].

In this paper we will consider autoregressive parameter estimation under the different assumptions in comparison with the standard approach. Namely, the most commonly used assumption is that the stochastic disturbance has Gaussian model. Beliefs in existence a kind of continuity principle according to which the results of inference would change only a small amount if the actual model deviated only a small amount from the assumed model are unjustified [7]. Estimation algorithms, based on the Gaussian model, have been found to be especially inefficient when the real distribution belongs to the heavy tailed variety, giving rise to the occasionally very large outliers [8]. According to outliers type which can occur in

practice in this paper we will consider disturbance uncertainty in the form of innovation outliers [9].

In this paper we will consider identification of AR models when stochastic process is non-Gaussian. We use the iterative procedure (off-line identification). The minimized functional is nonlinear and includes a priori information about the probability distribution of observations. Namely, using game theory it is possible to find least favourable pdf within a prespecified pdf class \mathcal{P} to which the real noise pdf belongs [7]. For such pdf asymptotic estimation error covariance matrix has a saddle point. In theoretical investigation of iterative algorithm we extend the results for scalar parameters [10], [11] on the case of AR models. It is proved consistency and asymptotic normality for estimated parameters.

II. PROBLEM FORMULATION

Let the system under consideration be described by a linear single input-single output AR model [12]

$$A(q^{-1})y(i) = e(i) \quad (1)$$

where

$$A(q^{-1}) = 1 + a_1 q^{-1} + \dots + a_n q^{-n}, \quad n \geq 0 \quad (2)$$

characteristic polynomial in backwards shift operator q^{-1} with unknown coefficients

$$\theta^T = [a_1 \dots a_n]$$

but with known orders n . Here $y(i) \in R^1$ and $e(i) \in R^1$ and $e(i)$ is stochastic process. The stochastic process $\{e(i)\}$ has the properties

$$E\{e(i)\} = 0, \quad E\{e^2(i)\} < \infty$$

where $E(\cdot)$ is mathematical expectation operator.

Autoregressive model (1) can be rewritten in the next form

$$y(i) = Z^T(i)\theta + e(i) \quad (4)$$

where $Z^T(i) = [-y(i-1) \dots -y(i-n)]$ is vector of measurement. The main goal is parameters estimation based on minimization of next functional

$$J(\theta) = E\{H(y(i) - Z^T(i)\theta)\} \quad (5)$$

where $H: R^1 \rightarrow R^1$ is a function which depends from distribution of stochastic process $e(i)$. When exact

