A COMBINED ALGORITHM FOR IDENTIFICATION AND APPROXIMATION*

DENG SHUHUI (邓述慧)
(Institute of Systems Science, Academia Sinica)

Abstract

The optimum point of an unknown function which is measured with noise and the unknown parameters in the measurement noise are estimated by a combined algorithm. The almost sure convergence and the convergence rate are established for both unconstrained and constrained problems.

§ 1. Introduction

Stochastic approximation method has been widely applied in control and optimization problems since its pioneer work appeared in 1951 [2] and the measurement noise considered there covers uncorrelated and some class of correlated noises [3, 4]. But to the author's knowledge the unknown parameters contained in the measurement noise are never estimated in literature except a recent work by Chen and Guo [5].

On the other hand, the convergence of parameter identification for dynamic systems is established only for exact ARMAX models (see, for example [6, 7]), while in stochastic approximation problems the unknown regression function, generally speaking, cannot be modelled as an ARMAX model.

This paper applies the extended least squares algorithm for estimating unknown parameters in noise and a stochastic approximation algorithm truncated at randomly varying bounds for approximating the unknown optimum point of a regression function. The rates of convergence of the estimates respectively for the optimum point and for the true parameters in the noise are established for both unconstrained and constrained problems.

§ 2. Statement of Problem and Algorithms

1) Unconstrained problem

Let \( h(\cdot) \) be an unknown continuous function defined and valued both in \( \mathbb{R}^d \) and let

\[
h(x^0) = 0, \quad x^0 \in \mathbb{R}^d.
\]

Assume that \( x_i \) is the \( i \)-th approximate to the unknown \( x^0 \) sought for, and that the observation at time \( t+1 \) is

\[
y_{t+1} = h(x_i) + \epsilon_{t+1} + \epsilon_{t+1}^0,
\]

Received June 26, 1986.
Project supported by the National Natural Science Foundation of China.
where \( e_{i+1} + e_{i+1}^* \) is the measurement error. Assume \( s_{i+1}^* \) is a moving average process

\[
e_{i+1} = w_{i+1} + c_1 w_i + \cdots + c_r w_{i-r+1}
\]

with an unknown matrix coefficient

\[
\theta^* = [c_0 \cdots c_r]
\]

where \( \{w_k, \mathcal{F}_k\} \) is an \( l \)-dimensional martingale difference sequence with respect to a nondecreasing family of \( \sigma \)-algebras such that

\[
E\{w_k/\mathcal{F}_{k-1}\} = 0, \quad E\{|w_k|^2/\mathcal{F}_{k-1}\} \leq \sigma^2, \quad \forall k,
\]

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^n w_k w_k^t = R > 0, \quad a.s.
\]

By \( s_{i+1}^* \), we mean the unmodelled part of the measurement error and \( s_{i+1}^* \) is assumed to be \( \mathcal{F}_r \)-measurable and bounded such that

\[
\|s_{i+1}^*\| \leq \frac{c}{\sqrt{t}}, \quad c > 0, \quad \forall i.
\]

Our purpose is to recursively estimate \( x^0 \) and \( \theta \), and to give the convergence rate.

We estimate \( \theta \) by the extended least squares algorithm which is defined by \([7]\)

\[
\theta_{k+1} = \theta_k + a_n P_n [y_{k+1}^* - \phi^* \theta_k],
\]

\[
P_{k+1} = P_k - a_n P_k \varphi_k \varphi_k^t P_k, \quad a_n = (1 + \varphi_k^t P_k \varphi_k)^{-1}, \quad P_0 = d I, \quad d = {\lambda},
\]

\[
\phi_k = [y_k^* - \phi_k \theta_k, \ldots, y_{k-r+1}^* - \phi_k \theta_{k-r+1}]
\]

with arbitrary \( \theta_0 \) and \( \phi_0 \).

The unknown zero \( x^0 \) is estimated by a stochastic approximation procedure truncated at randomly varying bounds which is defined as follows \([8-10]\). Take a sequence of real numbers \( M_k < M_{k+1} \). \( M_k \to \infty \) and let an \( l \)-dimensional vector \( x^* \) satisfy Condition a) described later on. Given any \( x_0, x_k \) is recursively defined by

\[
x_{k+1} = \left( x_k + \frac{1}{k+1} y_{k+1} \right) I \left[ \|x_k^* + \frac{1}{k+1} y_{k+1}\| < M_{x_0} \right]
\]

\[
+ x^* I \left[ \|x_k^* + \frac{1}{k+1} y_{k+1}\| > M_{x_0} \right],
\]

\[
\sigma(k) = \sigma(k-1) + I \left[ \|x_k + \frac{1}{k+1} y_{k+1}\| > M_{x_0} \right], \quad \sigma(0) = 0.
\]

2) Constrained problem

Let \( f(x) \) and \( g^*(x) = (g_1(x), \ldots, g_l(x)) \) be continuously differentiable \( R^l \to R \) and \( R^l \to R^m \) functions respectively with \( m < l \).

We want to solve the constrained optimization problem

\[
\begin{cases}
  f(x) = \text{min}, \\
  g(x) = 0,
\end{cases}
\]

but the measurement again is corrupted by noise

\[
y_{k+1} = -f_s(x_k) + e_{i+1} + e_{i+1}^*,
\]

where \( f_s(x) \) denotes the gradient of \( f(x) \), i.e.

\[
f_s(x) = \left[ \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_l} \right]^t.
\]