Dynamics of Continuous, Discrete and Impulsive Systems
Series **B:** *Applications & Algorithms* **10 (2003)** 195-206

STOCHASTIC APPROXIMATION OF A SIMPLE NEURAL NETWORK-TYPE LEARNING ALGORITHM VIA COMPUTER SIMULATION

Israel Ncube[1], S. A. Campbell[2] and E. R. Vrscay[3]

[1]Department of Mathematics and Statistics
York University, Toronto, Ontario M3J 1P3

[2,3]Department of Applied Mathematics
University of Waterloo, Ontario N2L 3G1

Abstract. We present a computer algorithm for estimating the probability of convergence of a random neural network-type learning algorithm to a particular fixed point. Our algorithm is particularly useful for predicting the asymptotic behaviour of random algorithms which can converge to more than one fixed point. Our results suggest that the estimated probabilities are independent of the initialisation of the random learning algorithm.
Key words. neural networks, learning, dynamical systems, stochastic approximation.
AMS subject classification. 92B20, 34K20, 34K15.

1 Introduction

Consider a recursive random algorithm of the form

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \gamma_n h[\mathbf{x}_n, \varphi_n] \,, \tag{1}$$

where the $\mathbf{x}(n)$'s and $\varphi(n)$'s are in $\mathbb{R}^m$, $\{\gamma_n\}$ is a sequence of positive decreasing-to-zero real numbers such that $\sum_n \gamma_n = \infty$, $h : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ is a continuous function, and $\varphi(n)$ is a sequence of random variables that are distributed according to some given law. The subject of stochastic approximation is concerned with characterisation of the long term behaviour of recursive random algorithms. Different approaches to this problem have been proposed [1-10].

A particular issue of interest may be posed by the following question: *Does the algorithm converge to a unique fixed point, independent of initialisation?* Ljung [1] and Kushner *et al.* [2] addressed this question by considering the ordinary differential equation (ODE)

$$\frac{d\mathbf{z}}{dt} = \overline{h}[\mathbf{z}(t)] \,, \tag{2}$$

where

$$\overline{h}(\mathbf{z}) = E[h(\mathbf{z},\ \varphi_n) \mid \varphi_0, \cdots, \varphi_{n-1}] \,,$$

[1]To whom correspondence should be addressed. Email: israel_n@mathstat.yorku.ca