Signal Processing 111 (2015) 50–60

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

The q-Least Mean Squares algorithm

Ubaid M. Al-Saggaf [a,b], Muhammad Moinuddin [a,b], Muhammad Arif [c], Azzedine Zerguine [d,*]

[a] Electrical and Computer Engineering Department, King Abdulaziz University, Saudi Arabia
[b] Center of Excellence in Intelligent Engineering Systems (CEIES), King Abdulaziz University, Saudi Arabia
[c] Electrical Engineering Department, PAF Karachi Institute of Economics and Technology University, Pakistan
[d] Electrical Engineering Department, King Fahd University of Petroleum & Minerals, Saudi Arabia

ARTICLE INFO

Article history:
Received 3 May 2014
Received in revised form
8 September 2014
Accepted 24 November 2014
Available online 3 December 2014

Keywords:
Adaptive filters
q-gradient
q-LMS algorithm
LMS algorithm
Transient analysis
Steady-state analysis

ABSTRACT

The Least Mean Square (LMS) algorithm inherits slow convergence due to its dependency on the eigenvalue spread of the input correlation matrix. In this work, we resolve this problem by developing a novel variant of the LMS algorithms based on the q-derivative concept. The q-gradient is an extension of the classical gradient vector based on the concept of Jackson's derivative. Here, we propose to minimize the LMS cost function by employing the concept of q-derivative instead of the convent ional derivative. Thanks to the fact that the q-derivative takes larger steps in the search direction as it evaluates the secant of the cost function rather than the tangent (as in the case of a conventional derivative), we show that the q-derivative gives faster convergence for $q>1$ when compared to the conventional derivative. Then, we present a thorough investigation of the convergence behavior of the proposed q-LMS algorithm and carry out different analyses to assess its performance. Consequently, new explicit closed-form expressions for the mean-square-error (MSE) behavior are derived. Simulation results are presented to corroborate our theoretical findings.

1. Introduction

The concept of adaptive filtering constitutes an important part in statistical signal processing. Whenever there is a requirement to process signals that result from unknown statistics of an environment, the use of an adaptive filter offers an attractive solution to the problem. Thus, adaptive filters are successfully applied in such diverse fields as equalization, noise cancelation, linear prediction, and in system identification [1,2]. The most widely used algorithm for adaptive filters is the Least Mean Squares (LMS) algorithm [3]. The conventional LMS algorithm is derived using the concept of the steepest descent approach with the aid of conventional gradient[1] whose weight update can be formulated as [1]

$$\mathbf{w}_{i+1} = \mathbf{w}_i - \frac{\mu}{2}\nabla_{\mathbf{w}} J(\mathbf{w}), \tag{1}$$

where $J(\mathbf{w}) = E[e_i^2]$ for the well known LMS algorithm [1,2] and e_i is the estimation error between the desired response, d_i, and its estimate, $\mathbf{u}_i^T\mathbf{w}_i$, produced by an adaptive filter for an input $\mathbf{u}_i$ at time instant i, that is,

$$e_i = d_i - \mathbf{u}_i^T\mathbf{w}_i. \tag{2}$$

Since the LMS algorithm belongs to the class of stochastic gradient type adaptive algorithms, it inherits their low

* Corresponding author.
E-mail addresses: usaggaf@kau.edu.sa (U.M. Al-Saggaf), mmsansari@kau.edu.sa (M. Moinuddin), mstel11@yahoo.com (M. Arif), azzedine@kfupm.edu.sa (A. Zerguine).

[1] For a function $f(\mathbf{x})$ of a real valued vector $\mathbf{x} = [x_1, \dots, x_M]^T$, the gradient is defined as $\nabla_{\mathbf{x}} f(\mathbf{x}) \triangleq [\partial f/\partial x_1, \dots, \partial f/\partial x_M]^T$.

http://dx.doi.org/10.1016/j.sigpro.2014.11.016
0165-1684/

computational complexity and their slow convergence, especially when operating on highly correlated signals like speech. One approach to overcome the slow convergence problem of the LMS algorithm is by employing a time varying step size in the standard LMS algorithm [4–9]. This is based on using a large step size when the algorithm is far from the optimal solution, thus speeding up the convergence rate, and when the algorithm is near the optimum, a small step size is used to achieve a low level of misadjustment, thus achieving a better overall performance. This can be obtained by adjusting the step size in accordance to some criterion. Several criteria have been used, such as squared instantaneous error [4], sign changes of successive samples of the gradient [5], cross correlation of input and error [6], gradient of squared error cost function [7], and square of the time averaged estimate of the correlation of the error [8], just to name a few. The second approach to improve the convergence speed is to use a normalization in the weight update of the LMS or the Least Mean Fourth (LMF) algorithms, such as used in the normalized LMS (NLMS) algorithm [10] and in the variable XE-NLMF algorithm [11]. Unlike the previous two approaches, a third approach relies on adding a proper constraint to the cost function of the LMS or LMF algorithms [12–15]. Or, more recently, the kernel-based non-linear kernel LMS variants such as the Kernel LMS algorithm for real-valued input [16], the Complex Kernel LMS (CKLMS) algorithm [17] and a modified CKLMS based on modified Wirtinger's Calculus [18] have also been investigated. All these variants of the LMS algorithm improve convergence speed and/or reduce the mean-square-error at the expense of an increase in the computational complexity. In order to improve more the convergence performance of the conventional LMS algorithm while retaining its simplicity, here we propose to utilize a novel concept based on the q-calculus which is introduced in the ensuing section, and eventually yield the q-LMS algorithm.

1.1. Overview of the q-calculus and the q-gradient

In the last few decades, the q-calculus has gained a lot of interest in various fields of science, mathematics, physics, quantum theory, statistical mechanics, and signal processing [19]. Jackson introduced the concepts of the q-derivative [20] (well known as Jackson's derivative) and the q-integral [21]. The q-derivative of a function $f(x)$ with respect to variable x, denoted by $D_{q,x}f(x)$, is defined as [22]

$$D_{q,x}f(x) \triangleq \begin{cases} \frac{f(qx)-f(x)}{qx-x} & \text{if } x \neq 0, \\ \frac{df(0)}{dx}, & x=0, \end{cases} \tag{3}$$

where q is a real positive number different from 1. In the limiting case of $q \to 1$, the q-derivative reduces to the classical derivative. Thus, as an example, the q-derivative of a function of the form x^n is

$$D_{q,x}x^n = \begin{cases} \frac{q^n-1}{q-1}x^{n-1} & \text{if } q \neq 1, \\ nx^{n-1} & \text{if } q=1. \end{cases} \tag{4}$$

Extending this idea to the q-gradient of a function $f(\mathbf{x})$ of n variables, where $\mathbf{x}=[x_1,x_2,\ldots,x_n]^T$, the q-gradient in this case is defined as

$$\nabla_{\mathbf{q},\mathbf{x}}f(\mathbf{x}) \triangleq [D_{q_1,x_1}f(\mathbf{x}), D_{q_2,x_2}f(\mathbf{x}), \ldots, D_{q_n,x_n}f(\mathbf{x})]^T, \quad \text{for } q \neq 1, \tag{5}$$

where $\mathbf{q}=[q_1,q_2,\ldots,q_n]^T$.

Using the concept of q-gradient, it is shown in [23] that the use of the negative of the q-gradient of the objective function as the search direction for unconstrained global optimization gives better results than the one obtained by the conventional gradient. This motivates us to investigate the q-gradient-based adaptive algorithms.

1.2. Paper contributions and organization

The main contributions of the paper are as follows:

(1) In this work, we introduce a new class of adaptive filtering based on q-calculus. More specifically, we derive a novel variant of the LMS algorithm by replacing the conventional gradient in (1) by the q-gradient which we named as q-LMS algorithm.
(2) We provide a geometrical interpretation of the q-gradient to justify the proposed design. This also offers us a better understanding that how the q-gradient can improve the convergence speed of an adaptive filter.
(3) We show an interesting attribute of the q-gradient based LMS algorithm that it can whiten the colored input of the adaptive filter by employing proper selection of its q-parameters. Consequently, it improves the convergence speed of the algorithm.
(4) We carry out a thorough analytical investigation of the proposed algorithm by studying both its transient and steady-state convergence behaviors. Consequently, both the MSE and MSD learning curves are evaluated and expressions for the steady-state EMSE and the MSD are derived.
(5) We also develop an efficient mechanism to make the q parameter time varying such that variable q-LMS algorithm should give a faster convergence while attaining a lower steady-state EMSE.
(6) We perform extensive simulations to show the superiority of the q-LMS algorithms over the conventional LMS and the NLMS algorithms and to validate the analytical results.

The paper is organized as follows. Following this introduction, the q-steepest descent algorithm is developed in Section 2. A geometrical interpretation of the q-gradient is presented in Section 3. Section 4 introduces the proposed q-LMS algorithm. In Section 5, whitening property of the q-LMS algorithm is investigated. A thorough performance analysis is carried out for the developed q-LMS algorithm in Section 6. In Section 7, an efficient time varying q-LMS algorithm is designed. While the simulation results are presented in Section 8, Section 9 summarizes this work.