Channel equalization with rapid convergence based on $\varepsilon$–support vector machines

Li Chang
College of physics and electronic information engineering
WenZhou University
WenZhou, China
lichang@wzu.edu.cn

RUAN Xiu-Kai
College of Telecommunications and Information engineering
Nanjing University of Posts and Telecommunications
Nanjing, China

Abstract—In this paper, the problem of adaptive channel equalization is considered. We extend the previous work in a new direction with periodic training sequence and formulate the original channel equalization problem into the equivalent linear regression in which $\varepsilon$–support vector regression machines are proposed. Slack variables are introduced in our approach to solve the problem of quadratic programming derived from $\varepsilon$–support vector regression machines. Simulation results show that the bit error rate of the proposed equalizer based on $\varepsilon$–support vector regression machines comes near to that of the optimal equalizer and is better than that of wavelet neural network equalizer. Simulation results also support the rapid convergence of the proposed approach, which can meet the real-time requirement of time-varying channels effectively in modern digital communication systems.

Keywords: $\varepsilon$–support vector regression; channel equalization; quadratic programming; inter-symbol interference

I. INTRODUCTION

In modern wireless communication systems, symbols transmitted through channels are always corrupted severely by other symbols as long as their waveforms are aliasing in one symbol period, which is defined inter-symbol interference (ISI) in literatures [1]. In order to improve the efficiency and reliability of a communication system, distortion of symbols caused by inter-symbol interference must be suppressed. Adaptive channel equalization is well-known for the ability to reconstruct symbols transmitted based on periodical observations at the end of the channel. Recently, there has been an explosion in the research results considering the important and crucial role of equalization techniques in modern digital communication systems [2].

As a matter of fact, most of the existing standards on wireless communication implement channel equalization by training sequence, especially in the situation that channels are always varying rapidly. This traditional method gets the equalizer by inserting training sequences periodically. An excessively short training sequence will lead to inaccuracy of designed equalizer while an enough long training sequence will waste the precious bandwidth of the channel. How to improve the promptness and accuracy of adaptive channel equalization in wireless communication systems is still the focus of research work in recent years [3].

Traditionally, adaptive channel equalization is equivalent to the process of inverse filtering. The optimal solution known as the Bayesian method is based on maximum likelihood sequence estimation (MLSE), which has an unknown delay and computation complexity that grows exponentially with the dimension of the channel impulse response. Alternatively, technique of machine learning can be used to approximate MLSE decisions at a lower computational cost. Several nonlinear detection procedures have been proposed to address this problem: multi-layered perceptrons (MLPs), radial basis function networks (RBFNs) [4], recurrent RBFNs [5], self-organizing feature maps (SOFMs) [5], [6], wavelet neural networks [7]. However, these methods are still problematic because of slow convergence and computation complexity.

Support vector machines (SVMs) dose show good performance of rapid convergence and is widely used in various applications: image recognition, classification as well as control problems[8]. But there is much less work on channel equalization in which SVM-related algorithm is taken into account. In this article, $\varepsilon$–SVMs are adopted for adaptive channel equalization based on periodic training sequence, which formulates the problem of channel equalization into a quadratic programming (QP) problem. Equalizers could be obtained directly only within a very short training sequence. The organization of the paper is as follows. In section II, problem of regression equivalent to channel equalization based on periodic training sequence is formulated. Procedure of solving the regression based on $\varepsilon$–SVMs and corresponding algorithm are proposed in section III. Section IV shows the performance of the method by means of some experimental results. Finally, Section V includes some concluding remarks.

II. PROBLEM FORMULATION

A. Channel equalization based on periodic training sequence

Symbols are transmitted in frame composed of training sequence and payload in most of the wireless communication systems. Parameters of the fading channel are assumed to be invariable during one frame period.

Let $d(k)$, $s(k)$ and $h(k)$ denote the training sequence, data and channel pulse response function respectively. Periodic training sequence $d(k)$ means that $d(k) = d(k+P)$ with period P. Received training signal

978-1-4244-5668-0/09/$25.00 ©2009 IEEE
\( x(k) \) is the convolution of \( d(k) \) and \( h(k) \) while being added by AWGN. The goal of channel equalization is to obtain channel impulse response function or design a reasonable equalizer using received signal and training sequence directly.

Suppose that the set of received training data samples is \( T = \{(x_1,d_1),(x_2,d_2), \cdots , (x_l,d_l)\} \) with different elements consist of one column vector \( x_i = (x_{i-1},x_{i-1}, \cdots ,x_{i-L+i})^T \) and one training symbol \( d_i \) known to the receiver, where \( L \) is length of the designed equalizer and \( L \geq P \). The goal of channel equalization is to design reasonable coefficients of the equalizer which will make the output \( y_i \) equal to the expected symbol \( d_i \) shown as following:

\[
\mathbf{w}^T \mathbf{x}_i = y_i = d_i, \quad i = 1 \cdots l
\]

(1)

B. Problem of regression equivalent to channel equalization

When errors in practical process of equalization are considered, signal estimation error \( e_i \) is introduced into formula (1). This leads to that the issue of equalization denoted as formula (1) can be converted to a problem of linear regression as follows:

\[
y_i = f(x_i) = \mathbf{w}^T \cdot \mathbf{x}_i + e_i
\]

(2)

As we know that \( x_i \in \mathbb{R}^n \) and \( y \in \mathbb{R} \). A general expression of linear regression based on formula (2) is deduced as follows:

\[
y = f(x) = \mathbf{w}^T \cdot \mathbf{x} + b
\]

(3)

From the view of geometry, we can figure out the way to solve a problem of channel equalization based on training sequence by the following equivalent linear regression method:

Since a training set \( T = \{(x_i,y_i)\}_{i=1}^l \in (X \times Y)^l \) is given, the purpose for linear regression is to find the classifier \( f(x) = \mathbf{w}^T \cdot \mathbf{x} + b \) so that we can separate the symbols received.

III. PROCEDURE OF SOLVING THE REGRESSION BASED ON \( \varepsilon \)-SVMs

A. Introduction of \( \varepsilon \)-support vector regression machines

Support vector machines have been successfully applied in classification and function estimation problems[9]. The standard SVM is constructed by Vapnik to separate training data into two classes with goal to find the hyper-plane that maximizes the minimum distance between any data point. Following the basic theory of SVM, we can solve the regression problem in formula (3) using SVM[9]:

\[
\begin{align*}
\mathbf{x} \in \mathbb{R}^n, y_i \in Y = \{\pm 1\}, \\
\min_{w,b} & \quad \frac{1}{2} ||w||^2, \\
\text{s.t.} & \quad y_i \cdot (w^T \mathbf{x}_i + b) \geq 1, \quad i = 1, \cdots , l
\end{align*}
\]

(4)

If the optimal solution is founded, the output of the equalizer might be given by the following decision function:

\[
f(x) = \text{sgn} \left( (w^* \cdot x + b^*) \right)
\]

(5)

When possible errors are considered during classification of the linear regression, positive point and negative point are introduced into the linear regression above-mentioned in formula (4). These two classes of points restrict the estimation error within boundary \( \pm \varepsilon \) so we call this method is \( \varepsilon \)-SVMs regression:

\[
D^+ = \left\{ (x_i^+, y_i^+ + \varepsilon) \right\}, \quad i = 1, \cdots , l
\]

(6)

\[
D^- = \left\{ (x_i^-, y_i^- - \varepsilon) \right\}, \quad i = 1, \cdots , l
\]

(7)

\[-\varepsilon \leq y_i - (w^T \cdot x_i) + b \leq \varepsilon, \quad i = 1, 2, \cdots , l
\]

(8)

As the training sequence is supposed to be periodic, we can draw the following conclusion:

\[
\left( x_{i+l}, y_{i+l} - \varepsilon \right) = \left( x_i, y_i - \varepsilon \right), \quad i = 1, \cdots , l.
\]

Then, the original training set can be adapted as follows:

\[
Q = \left\{ (x_i, y_i + \varepsilon), (x_i, y_i - \varepsilon) \right\}, \quad i = 1, \cdots , l
\]

Let \( \eta \) denote a new component added to weight vector \( \mathbf{w} \), the new weight vector \( (w^T, \eta)^T \) is introduced in formula (4), the optimal problem based on \( \varepsilon \)-SVMs regression can be described as follows:

\[
\begin{align*}
\min_{w,\eta} & \quad \frac{1}{2} ||w||^2 + \frac{1}{2} ||\eta||^2, \\
\text{s.t.} & \quad w^T \mathbf{x}_i + \eta (y_i + \varepsilon) + b \geq 1, \quad i = 1, \cdots , l, \\
& \quad w^T \mathbf{x}_j + \eta (y_j - \varepsilon) - b \leq -1, \quad j = l+1, \cdots , 2l.
\end{align*}
\]

(9)

Let \( z_i = \left\{ 1, \quad i = 1, \cdots , l, \\
-1, \quad i = l+1, \cdots , 2l \right\} \) be substituted into formula (9)
\[
\min_{w,\eta} \frac{1}{2} \|w\|_2^2 + \frac{1}{2} \eta^2 \tag{10}
\]
\[
\text{s.t. } z_i(w^T x_i + \eta(y_i + e) + b) \geq 1, i = 1, \ldots, 2l
\]

If the optimal values denoted as \(w^*, \eta^*\) and \(b^*\) are founded, a decision surface can be constructed as follows:
\[
\left(w^*\right)^T x + \eta^* y + b^* = 0 \tag{11}
\]

B. Solution to the problem based on \(\varepsilon\)-support vector regression machines

Because the linear decision surface described in formula (11) may be hard to construct or non-exist, slack variable \(\xi_i \geq 0\) should be imported to loose the constraints:
\[
z_i(w^T x_i + \eta(y_i + e) + b) + \xi_i \geq 1
\]

Then the quadratic programming problem is updated as follows:
\[
\min_{w,\eta,b} \frac{1}{2} \|w\|_2^2 + \frac{1}{2} \eta^2 + C \sum_{i=1}^{2l} \xi_i
\]
\[
\text{s.t. } z_i(w^T x_i + \eta(y_i + z_i e) + b) + \xi_i \geq 1, i = 1, \ldots, 2l
\]
\[
\xi_i \geq 0, i = 1, \ldots, 2l
\]

Where \(C\) is the trade-off parameter between the error and margin, let’s suppose that \(w^*, \eta^*, b^*, \xi^*\) denote the optimal values of formula (13) and substitute \(e^* = e - \frac{1}{\eta}\) in formula (13), the problem is updated as follows:
\[
\min_{w,\eta,b} \frac{1}{2} \|w\|_2^2 + \frac{1}{2} \eta^2 \tag{14}
\]
\[
\text{s.t. } w^T x_i + b - y_i \leq e^* + \xi_i
\]

We can see that the solution satisfies:
\[
w = \frac{w^*}{\eta^*}, \quad b = -\frac{b^*}{\eta^*}
\]

The quadratic programming problem is solved by considering the dual problem:
\[
\min_{\alpha, \xi} \frac{1}{2} \sum_{i=1}^{2l} (\tilde{\alpha}_i - \alpha_i)(\tilde{\alpha}_j - \alpha_j)K(x_i, x_j) + e \sum_{i=1}^{2l} (\tilde{\alpha}_i + \alpha_i) - \sum_{i=1}^{2l} \xi_i (\tilde{\alpha}_i - \alpha_i) \tag{15}
\]
\[
\text{s.t. } \sum_{i=1}^{2l} (\tilde{\alpha}_i - \alpha_i) = 0
\]
\[
0 \leq \alpha_i, \quad \tilde{\alpha}_i \leq C, i = 1, \ldots, l
\]

As channel equalization based on training sequence is considered, \(y_i = d_i\) is substituted in formula (15):
\[
\min_{a,d,\alpha} \frac{1}{2} \sum_{i=|a|j=1}^{l} (\tilde{\alpha}_i - \alpha_i)(\tilde{\alpha}_j - \alpha_j)K(x_i, x_j) + e \sum_{i=1}^{2l} (\tilde{\alpha}_i + \alpha_i) - \sum_{i=1}^{2l} d_i (\tilde{\alpha}_i - \alpha_i) \tag{16}
\]
\[
\text{s.t. } \alpha \geq 0, \tilde{\alpha}_i \geq 0, \quad i = 1, \ldots, l
\]

It can be simplified into a quadratic form as follows:
\[
\min_{a,d,\alpha} -\frac{1}{2} A^T QA - A^T E \tag{17}
\]
\[
\text{s.t. } \alpha \geq 0, \tilde{\alpha}_i \geq 0, \quad i = 1, \ldots, l
\]

Where \(A = \left(\begin{array}{c} a \\ \tilde{a} \end{array}\right)\) is a 2l-dimensional column vector with elements \(\alpha_i\) and \(\tilde{\alpha}_i\), \(Q = \left(\begin{array}{cc} -K & K \\ K & -K \end{array}\right)\) is a block matrix composed by a Gram matrix \(K\) with element \(K(x_i, x_j)\), \(E = \left(\begin{array}{cc} d - \varepsilon \cdot 1 \\ -d - \varepsilon \cdot 1 \end{array}\right)\) is a 2l-dimensional column vector.

According to the Karush-Kuhn-Tucker(KKT) conditions, if optimal solution is obtained as follows[9]:
\[
a^* = \left(\begin{array}{c} \alpha^*_1, \tilde{\alpha}^*_1, \ldots, \alpha^*_l, \tilde{\alpha}^*_l \end{array}\right)^T
\]

The expected equalizer would be
\[
w_{QP} = \sum_{i=1}^{l} (\tilde{\alpha}_i - \alpha_i) x_i \tag{18}
\]

IV. SIMULATION RESULTS

In this section, we will demonstrate the validity of the proposed method in a communication system. Parameters are supposed that \(C = 20\), \(\varepsilon = 0.05\), number of Monte Carlo simulation is 200. Channel adopted in the simulation is worse than the one used in literatures [3][6];
\[
h = 0.4z^{-1} - 0.7z^{-2} + 0.9z^{-3} + 0.3z^{-4} - 0.4z^{-5} + 0.1z^{-6}
\]

If the discrete-time model of this channel is depicted, the ISI caused by the channel is obvious.

A. Experiment 1: Average ISI in our proposed method vs. data block size in the situation of different SNR

Let the signal-to-noise ratio be 15dB, 20dB, 25dB and 30dB respectively and training data block size grow from 20 to 200 with interval 20. Performance of channel equalization measured with average ISI vs. data block size in the situation of different SNR is shown in figure 1. According to this figure, It is obvious that the proposed method can converge rapidly as long as a short training data block is given.
B. Experiment 2: Performance comparison of SVMs-based equalizers VS other methods.

In this experiment, SNR is set to be 20dB and the length of training data grows from 20 to 2000 with non-uniform interval. We test the performances of Bayesian equalizer[3], Wavelet neural network(WNN)-based on equalizer[6] and the proposed SVMs-based equalizer. BER and convergence are compared in the experiment with results shown in the following figures. We can see from figure 2 that Bayesian equalizer is the most optimal method since convergence is out of consideration. Figure 2 also shows that BER of the proposed SVMs-based method is close to Bayesian equalizer and better than that of WNN equalizer. As convergence performances are shown in figure 3, we can draw the conclusion that our method can converge rapidly and gain reasonable effects of equalization if only length of training data is bigger than 50. Results of figure 3 also show that the convergence performance of the proposed method is insensitive to the growth of the length of the training data as long as it is long enough.

V. CONCLUSION

In this paper, mathematical model of channel equalization is formulated into a quadratic programming problem of linear regression. Solution by means of ε-SVMs is deduced and equalizer with rapid convergence is obtained. Simulation results show that the bit error rate of the proposed method is close to the optimal equalizer and better than WNN equalizers. Rapid convergence of this novel method is verified through the experiment. Weather it is suitable for non-linear channel still needs to be further studied.

REFERENCES


