BLIND EQUALIZATION OF CONSTANT MODULUS SIGNALS VIA SUPPORT VECTOR REGRESSION

Ignacio Santamaría, Jesús Ibáñez, Luis Vielva, Carlos Pantaleón,

Department of Communications Engineering (DICOM), University of Cantabria
Avda. Los Castros s/n, Santander, 39005, SPAIN.
e-mail:nacho@gtas.dicom.unican.es

ABSTRACT

In this paper the problem of blind equalization of constant modulus (CM) signals is formulated within the support vector (SV) regression framework. The quadratic inequalities derived from the CM property are transformed into linear ones, thus yielding a quadratic programming (QP) problem. Then an iterative reweighted procedure is proposed to blindly restore the CM property. The technique can be generalized to nonlinear blind equalization using kernel functions. We present simulation examples showing that linear and nonlinear blind SV equalizers offer better performance than cumulant-based techniques, mainly in applications when only a small number of data samples is available, such as in packet-based transmission over fast fading channels.

1. INTRODUCTION

In many communication systems, digital signals are transmitted through an unknown bandlimited channel with severe intersymbol interference (ISI). When a training sequence is not available, blind equalization techniques must be used to recover the input signal. These techniques exploit the knowledge about the statistical properties of the input signal or the structure of the channel [1].

A number of blind algorithms are based on stochastic gradient descent (SGD) minimization (on-line techniques) of a specially designed non-MSE cost function (to this class belongs the widely used CMA [2]). Other algorithms collect a block of data (batch techniques) and iteratively maximize a cost function based on cumulants (for instance, the so-called "super-exponential" algorithm by Shalvi and Weinstein [3]).

In burst data transmission over fast fading channels, blind algorithms must be able to remove a sufficient level of ISI by using a short block of data. In this case stochastic gradient descent algorithms, which typically suffer from slow convergence, cannot be employed. Similarly, batch cumulant-based algorithms also offer poor performance in this situation.

In this paper we propose an alternative blind equalization technique for CM signals, which is expected to require fewer data samples than SGD and cumulant-based algorithms. Blind equalization is formulated as a support vector (SV) regression problem [4], and an iterative procedure, denoted as iterative reweighted quadratic programming (IRWQP), is proposed to find the optimal regressor. Support vector machines (SVM) have been successfully applied to linear and nonlinear supervised equalization problems [5, 6, 7]. In these works the equalization problem is viewed as a supervised classification problem (with a training sequence), and the corresponding SV classifier is derived. In this paper the problem is formulated as a nonsupervised regression problem: only the knowledge about the CM property of the input signal is exploited.

Recently, some techniques have been proposed that formulate the blind equalization problem either as a quadratic problem with binary constraints [8], or as a convex optimization method subject to some linear and semidefiniteness constraints [9]. In both cases, the problem is solved via semidefinite programming (SP) techniques. Similarly to these approaches, here we formulate a convex problem that has a global optimal solution, but, in addition to this, the proposed solution has several attractive features: it is derived from the powerful theory of SV machines; efficient quadratic programming (QP) implementations can be used [11]; and, finally, it can be readily extended to nonlinear blind equalizers. Some simulation examples show the advantages of this technique in comparison to cumulant-based algorithms.

2. PROBLEM FORMULATION

We consider a baud-rate sampled baseband representation of the digital communication system. A sequence of i.i.d. symbols belonging to a binary alphabet $\{s_k \in \pm 1\}$ is sent through a linear time-invariant channel with coefficients h_k (the extension to complex modulations is straightforward). The resulting channel output can be expressed as

$$x_k = \sum_n h_n s_{k-n} + e_k,$$

where e_k is a zero-mean white Gaussian noise.

The objective of a blind linear equalizer is to remove the ISI at its output without using any training sequence. Typically, the equalizer is designed as an FIR filter with M coefficients $\mathbf{w}$; then, its output is given by

$$y_k = \sum_{n=0}^{M-1} w_n x_{k-n} = \mathbf{w}^T \mathbf{x}_k.$$

The method proposed by Shalvi and Weinstein [3], which will be used for comparison purposes, maximizes $|K_y|$, subject to the constraint $E\left[|y_k|^2\right] = E\left[|s_k|^2\right]$, where K_y is the kurtosis of y_k defined as

$$K_y = E\left[|y_k|^4\right] - 2\left(E\left[|y_k|^2\right]\right)^2 - \left|E\left[y_k^2\right]\right|^2.$$

This work was supported in part by MCYT (Ministerio de Ciencia y Tecnología) under grant TIC2001-0751-C04-03.

3. SV REGRESSION FOR BLIND EQUALIZATION

Suppose we are given a set of N observations at the channel output: $(\mathbf{x}_1, \cdots, \mathbf{x}_N)$, where $\mathbf{x}_i = (x_i, x_{i-1}, \cdots, x_{i-M+1})^T$. Then the goal of a linear blind equalizer is to restore at its output the CM property of the digital communications signal, i.e., $(\mathbf{w}^T\mathbf{x}_i)^2 = 1$, for $i = 1, \cdots, N$.

To apply the SVM concept, let us first introduce the so-called Vapnik's ϵ-insensitive loss function

$$|1 - (\mathbf{w}^T\mathbf{x})^2|_\epsilon = \max\{0, |1 - (\mathbf{w}^T\mathbf{x})^2| - \epsilon\}.$$

According to the Structural Risk Minimization principle [4], to estimate a linear equalizer with precision ϵ, one minimizes

$$J(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}|1 - (\mathbf{w}^T\mathbf{x}_i)^2|_\epsilon,$$

which can be rewritten as the following constrained optimization problem: for some penalty value $C > 0$, to minimize the objective function

$$L(\mathbf{w}, \xi, \xi^*) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*), \tag{1}$$

subject to

$$\begin{aligned} (\mathbf{w}^T\mathbf{x}_i)^2 - 1 &\leq \epsilon + \xi_i, && (2)\\ 1 - (\mathbf{w}^T\mathbf{x}_i)^2 &\leq \epsilon + \xi_i^*, && (3)\\ \xi_i, \xi_i^* &\geq 0, && (4) \end{aligned}$$

for all $i = 1, \cdots, N$.

In the conventional SVM approach for regression and function approximation, the inequality constraints are linear in the unknowns, $\mathbf{w}$, thus yielding a quadratic programming (QP) problem that can be efficiently solved [11]. The proposed constraints (2) and (3) for blind equalization of CM signals are, however, quadratic with respect to the coefficients of the equalizer. For this reason, a direct introduction of the constraints into the cost function, by means of Lagrange multipliers, does not render a QP problem.

To circumvent this drawback, we propose a procedure for finding the SV solution, which resembles the iterative reweighted least squares (IRWLS) technique used in some approximation and regression problems [10]. Let us first rewrite the squared modulus of the output of the equalizer as $(\mathbf{w}^T\mathbf{x}_i)^2 = y_i(\mathbf{w}^T\mathbf{x}_i)$. Now, considering y_i fixed, the inequalities (2) and (3) become linear in $\mathbf{w}$ and can be rewritten as

$$\begin{aligned} y_i(\mathbf{w}^T\mathbf{x}_i) - 1 &\leq \epsilon + \xi_i, && (5)\\ 1 - y_i(\mathbf{w}^T\mathbf{x}_i) &\leq \epsilon + \xi_i^*, && (6) \end{aligned}$$

in this way, the blind equalization problem can be formulated within the conventional support vector framework. In particular, the optimization problem reduces to the following: given C and ϵ, to find a saddle point of the quadratic problem

$$\begin{aligned} L(\mathbf{w}, \xi, \xi^*, \alpha, \alpha^*, \gamma, \gamma^*) &= \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*)\\ &- \sum_{i=1}^{N}(\gamma_i^*\xi_i^* + \gamma_i\xi_i) - \sum_{i=1}^{N}\alpha_i[1 - y_i(\mathbf{w}^T\mathbf{x}_i) + \epsilon + \xi_i]\\ &- \sum_{i=1}^{N}\alpha_i^*[y_i(\mathbf{w}^T\mathbf{x}_i) - 1 + \epsilon + \xi_i^*], \end{aligned} \tag{7}$$

minimum with respect to $\mathbf{w}$, ξ_i and ξ_i^*; and maximum with respect to Lagrange multipliers $\alpha_i \geq 0$, $\alpha_i^* \geq 0$, $\gamma_i \geq 0$ and $\gamma_i^* \geq 0$, for all $i = 1, \cdots, N$.

The solution for the linear equalizer can be expanded in terms of the outputs y_i, the input patterns $\mathbf{x}_i$, and the Lagrange multipliers α_i and α_i^* as

$$\mathbf{w} = \sum_{i=1}^{N}(\alpha_i^* - \alpha_i)\, y_i\mathbf{x}_i. \tag{8}$$

The Lagrange multipliers are obtained by maximizing the following quadratic form

$$\begin{aligned} W(\alpha, \alpha^*) &= -\epsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{N}(\alpha_i^* - \alpha_i)\\ &- \frac{1}{2}\sum_{i,j=1}^{N}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(y_iy_j)\langle\mathbf{x}_i, \mathbf{x}_j\rangle, \end{aligned} \tag{9}$$

subject to $0 \leq \alpha_i, \alpha_i^* \leq C$; and where $\langle\mathbf{x}_i, \mathbf{x}_j\rangle$ denotes the inner product between the inputs patterns.

According to the Karush-Kuhn-Tucker (KKT) condition, the input patterns that appear in the expansion (8) are points where exactly one of the Lagrange multipliers is greater than zero: these input patters are called support vectors.

The difference with the conventional formulation for SV regression problems is that the linear kernel $\langle\mathbf{x}_i, \mathbf{x}_j\rangle$ in the quadratic form (9) now is weighted by the factor y_iy_j, and that the solution is expanded in terms of a set of weighted Lagrange multipliers, which we define as $\beta_i = (\alpha_i^* - \alpha_i)\, y_i$

$$\mathbf{w} = \sum_{i=1}^{N}\beta_i\mathbf{x}_i. \tag{10}$$

Obviously, the procedure can be readily extended to nonlinear regression just by replacing in (9) the linear dot product by a nonlinear kernel. For example, the polynomial kernel $K(\mathbf{x}_i, \mathbf{x}_j) = (\langle\mathbf{x}_i, \mathbf{x}_j\rangle + 1)^p$, or the radial basis function kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\|\mathbf{x}_i - \mathbf{x}_j\|^2/(2\sigma^2)\right)$ could be used. The output of the nonlinear equalizer in this case is given by

$$y_k = \sum_{i=1}^{N}\beta_iK(\mathbf{x}_i, \mathbf{x}_k), \tag{11}$$

where β_i are again the weighted Lagrange multipliers.

4. ITERATIVE REWEIGHTED QP

The optimal blind regressor cannot be found in a single step because the weighted Lagrange multipliers depend on the solution $\beta_i = \beta_i(\mathbf{w})$. Therefore, we need to apply an iterative procedure, which due to its similarity with the IRWLS technique is called iterative reweighted quadratic programming (IRWQP). The procedure is as follows:

1. Solve the QP problem (9) considering y_i fixed.
2. Obtain the new equalizer as (10) and compute the corresponding new output y_i.
3. Repeat until convergence.

In order to complete the algorithm it is necessary to smooth somehow the equalizer coefficients from iteration to iteration. The reason for this smoothing is the following: suppose that the initial output of the equalizer is y_i, then, in the first step we perform a linear (or nonlinear) regression trying to fit as desired output $1/y_i$, forcing in this way a constant modulus signal. If we apply a new iteration of the IRWQP procedure, the new weights applied to the Lagrange multipliers will be close to $1/y_i$, whereas the new desired output will be again close to y_i. Then, to avoid a limit cycle oscillation between these two outputs, some type of smoothing must be introduced. In particular, we propose to smooth the weighted Lagrange multipliers according to

$$\beta_k = \lambda \beta_{k-1} + (1-\lambda)\,\beta^*, \tag{12}$$

where β^* are the weighted Lagrange multipliers obtained by solving (9) at the k-th iteration, and λ is a constant parameter close to one. This smoothing procedure can be applied to linear or nonlinear blind SVM equalizers.

The initial β can be obtained by solving the following linear problem

$$\mathbf{D}\beta_0 = \mathbf{y}_d, \tag{13}$$

where $\mathbf{D}$ is an $N \times N$ kernel matrix with elements given by $D_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$, and $\mathbf{y}_d = (x_{1-d}, \cdots, x_{N-d})^T$. In this way the initial equalizer delays the input d samples. This delay is typically chosen at the center of the equalizer coefficient vector.

Finally, the proposed algorithm can be summarized in the following steps

1. Choose C, ϵ and λ.
2. Initialize β_0 according to (13).
3. For $k = 0, 1, \cdots$, maxiter
   3.1. Obtain the output of the equalizer, y_i, using β_k.
   3.2. Solve the QP problem (9) and obtain β^*.
   3.3. $\beta_{k+1} = \lambda \beta_k + (1-\lambda)\,\beta^*$.
   End.

5. SIMULATION RESULTS

In this section we compare the performance of the proposed blind (linear or nonlinear) SVM and the batch super-exponential algorithm proposed by Shalvi and Weinstein (denoted as SW) [3], which is based on fourth-order cumulants. The QP problem at each step of the IRWQP procedure has been solved using the Matlab SVM toolbox available at [11].

In the first example we consider a linear blind SVM. A binary signal is sent through the channel $H_1(z) = (0.4 + z^{-1} - 0.7z^{-2} + 0.6z^{-3} + 0.3z^{-4} - 0.4z^{-5} + 0.1z^{-6})$ (used in [3]) and, at the channel output, white Gaussian noise is added. We have used an equalizer of length $M = 17$, which was initialized as $\mathbf{w} = \delta[n-8]$. As a measure of equalization performance we use the ISI defined as

$$\text{ISI} = 10\log_{10} \frac{\sum_n |\theta_n|^2 - \max_n |\theta_n|^2}{\max_n |\theta_n|^2},$$

where $\theta = \mathbf{h} * \mathbf{w}$ is the combined channel-equalizer impulse response, which is a delta function for a zero-forcing equalizer. The initial ISI for the selected channel and for an equalizer initialized with a centered spike is 1.03 dB.

Similarly to other cumulant-based algorithms, the SW algorithm provides poor results with very short blocks of data [3]. This could happen, for instance, in burst TDMA transmissions (without any training sequence) over fast fading channels. It is in this situation when the proposed blind SVM technique is expected to offer some advantages. To corroborate this point, we have tested the SW and blind SVM algorithms for blocks of $N = 50$, 100 and 200 samples; and different noise levels. For the blind SVM we used a penalty factor of $C = 10$, a precision term of $\epsilon = 0.1$, a smoothing term of $\lambda = 0.9$, and the Vapnik's ϵ-insensitive loss function (the results were similar for the quadratic loss function). A maximum of 30 iterations of the IRWQP procedure were carried out and, if the final ISI was below -5 dB, we considered that the channel was successfully equalized, since with this level of ISI it is already possible to switch to a decision-directed mode. This convergence criterion was also applied for the SW algorithm. For each data block size and noise level, both algorithms were tested in 200 Monte-Carlo trials. The results are summarized in Table 1, that shows the percentage of trials in which each algorithm converged, and, for the successful trials, the mean ISI level after convergence.

		Shalvi-Weinstein		blind SV	
SNR	N	% converg.	ISI	% converg.	ISI
30 dB	50	6%	-9.4	66.5%	-14.1
	100	67%	-12.2	96.5%	-15.2
	200	98%	-14.5	100%	-15.7
10 dB	50	0%	–	14.0%	-7.2
	100	9.8%	-7.1	50.5%	-8.6
	200	69.6%	-10.2	86.5 %	-10.6

Table 1. Performance of the SW and blind SVM algorithms. For each method, the first column shows the percentage of convergence and the second column the mean final ISI.

The blind SVM obtains better results than the SW method, mainly for very short data registers $N = 50, 100$. The price to be paid is obviously an increase in the required computational burden. The average ISI versus the IRWPQP iterations is plotted in Fig. 1.

In our second example we test the nonlinear blind SVM. We consider a nonlinear channel composed of a linear channel followed by a memoryless nonlinearity. Such a nonlinear channel can be encountered in digital satellite communications and in digital magnetic recording. The linear channel considered is $H(z) = 1 - 0.5z^{-1}$, and the nonlinear function applied is $z = x + 0.2x^2 - 0.9x^3$, where x is the linear channel output. Finally, white Gaussian noise for a SNR=10 dB was added. A short sequence of N=100 samples was considered. For the nonlinear blind SVM we use a polynomial kernel of degree $p = 3$ and the dimension

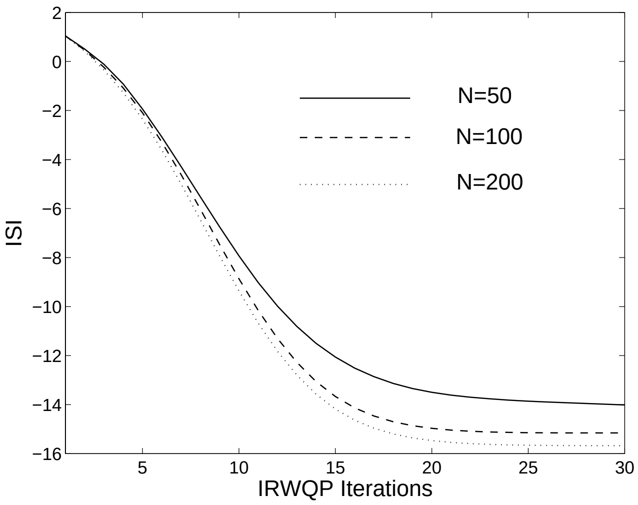

Fig. 1. Average convergence curves for the IRWQP procedure ($N = 50, 100$, and 200; SNR=30 dB).

of the input patterns is $M = 5$. We also selected a penalty factor value of $C = 10$, a precision term of $\epsilon = 0.01$, a smoothing factor of $\lambda = 0.9$, and the Vapnik's ϵ-insensitive loss function. Figure 2 shows the solution obtained by the nonlinear SVM (for which 89 patterns became support vectors) with solid line, and the true bits depicted with asterisks. We have also included for comparison the solution provided by the SW algorithm, although obviously a linear FIR filter cannot remove the distortion introduced by a nonlinear channel. Nevertheless, we can see that the blind SVM is able to restore the CM property of the input signal.

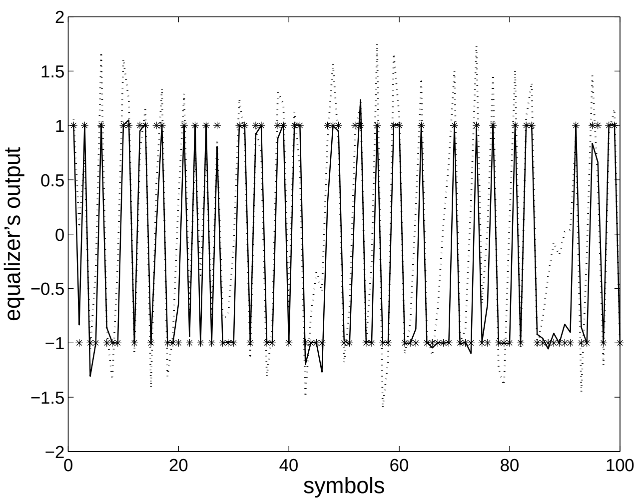

Fig. 2. Output of a nonlinear blind SVM equalizer (solid line), a linear equalizer trained with the SW algorithm (dotted line) and the true bits ('*').

6. CONCLUSIONS

In this paper, blind equalization of CM signals has been formulated as a regression problem and the powerful SVM technique has been applied to solve it. An iterative reweighted quadratic programming (IRWQP) procedure has been proposed to find the optimal regressor. It is shown in the paper that blind SVM equalization has several attractive features: the quadratic programming problem solved at each iteration is convex and has a globally optimal solution; it simultaneously exploits all the information in the given block of data, thus requiring fewer data samples than other standard blind algorithms; and, finally, linear and nonlinear equalizers can be treated in an unified manner and within a powerful machine-learning framework.

7. REFERENCES

[1] Z. Ding, Y. Li, *Blind Equalization and Identification*, New York: Marcel Dekker, 2001.

[2] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, R. A. Casas, "Blind equalization using the constant modulus criterion: A review", *Proc. IEEE*, vol. 86, no. 10, pp. 1927-1950, Oct. 1998.

[3] O. Shalvi, E. Weinstein, "Universal methods for blind deconvolution," in *Blind Deconvolution*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall,1994.

[4] V. Vapnik, *The Nature of Statistical Learning Theory*, New York: Springer Verlag, 1995.

[5] S. Chen, S. Gunn, C. J. Harris, "Decision feedback equaliser design using Support Vector Machines", *IEE Proc. Vis. Image Signal Process.*, vol. 147, no.3, pp. 213-219, June 2000.

[6] D. J. Sebald, J. A. Bucklew, "Support Vector Machine techniques for nonlinear equalization", *IEEE Trans. on Signal Processing*, vol. 48, no. 11, pp. 3217-3226, Nov. 2000.

[7] F. Pérez-Cruz, A. Navia-Vázquez, P.L. Alarcón-Diana, A. Artés-Rodríguez, "SVC-based equalizer for burst TDMA transmissions", *Signal Processing*, vol. 81, pp. 1681-1693, 2001.

[8] Q. Li, E.-W Bai, Y. Ye, "Blind channel equalization and ϵ-approximation algorithms, *IEEE Trans. Signal Processing*, vol. 49, no. 11, pp. 2823-2831, Nov. 2001.

[9] B. Maricic, Z.-Q Luo, T. N. Davidson, "Blind equalization of constant modulus signals via restricted convex optimization", *Proc. of the 2001 International Conference on Acoustics, Speech, and Signal Processing*, Salt Lake City, UT, USA, May 2001.

[10] P.W. Holland, R. E. Welch, "Robust regression using iterative reweighted least squares," *Commun. Statist. Theory Methods*, vol. A 6, no. 9, pp. 813-827, 1997.

[11] S. R. Gunn, *MATLAB Support Vector Machine Toolbox*, available at `http://www.isis.ecs.soton.ac.uk/isystems/kernel/`, University of Southampton, Image Speech and Intelligent Systems Research Group, UK, 1998.