Blind Equalization

Introduction

In the case of a highly nonstationary communications environment (i.e. wireless/mobile communications), it is impractical to consider the use of a training sequence since the channel may change before the training period is over and/or may be too costly in terms of data throughput.

In such a situation, the adaptive filter has to equalize the communications channel in a self-organized (unsupervised) manner. Thus "blind" adaptive equalizers do not rely on training sequences and hence can begin self-adaptation without transmitter assistance. This ability of blind startup also enables a blind equalizer to self-recover from system breakdowns. This self-recovery ability is critical in broadcast and multicast systems where channel variation often occurs.

Figure: Blind equalizer in wireless communication application

We have already considered the *property restoration approach* for constant envelope (modulus) communications signals. We can base our adaptive adjustment algorithm on restoring the constant modulus property to the input signal. This of course easily leads to the successful Constant Modulus Algorithm (CMA). We consider generalized approaches to blind equalization.

Bussgang Algorithm for Blind Equalization of Real Baseband Channels

We begin by considering the baseband model of a digital communications system. The model consists of the cascade connection of a linear channel (transmit filter, medium, receive filter), $\mathbf{h}$ and a blind equalizer, $\hat{\mathbf{w}}$.

Figure 18.2

Problem Statement: Given the received signal $u(n)$, reconstruct the original data sequence $x(n)$. Alternatively, design a blind equalizer, $\hat{\mathbf{w}}$ that is the inverse (or approximate inverse), $\mathbf{w}$ of the unknown channel, $\mathbf{h}$ i.e.

$$\sum_{\kappa} w_k h_{n-\kappa} = \delta_{\nu}$$

$$\sum_{\kappa} \hat{\omega}_{\kappa} \eta_{\nu-\kappa} \approx \delta_{\nu}$$

with the channel input being unobservable *and* with no desired response available.

Assumptions on the Communication System

1) We assume a real-valued impulse response (*M*-ary pulse amplitude modulation). The channel output is then

$$u(n) = \sum_{\kappa=-\infty}^{\infty} \eta_{\kappa} \xi(\nu-\kappa), \quad \nu = 0, \pm 1, \pm 2, \ldots$$

where $x(n)$ is the input data sequence and $u(n)$ is the resulting channel output.

2) The impulse response, h_k is unknown and may be (slowly) time varying, i.e. $h_k(n)$.

3) We assume the noise gain of the filter is unity, i.e.

$$\sum_k h_k^2 = 1 \, .$$

4) The channel is noncausal

$$h_n \neq 0, \; \nu < 0$$

5) We ignore receiver noise in order to maintain simplicity and since performance is dominated by ISI.

Assumptions on the Data Sequence

1) The data sequence $x(n)$ is zero mean and uncorrelated

$$E[x(n)] = 0$$

$$E[\xi(\nu)\xi(\kappa)] = \begin{cases} 1, & \kappa = \nu \\ 0, & o.\omega. \end{cases}$$

2) The probability density function (pdf) of the data sequence $x(n)$ is given by

$$f_X(x) = \begin{cases} \dfrac{1}{2\sqrt{3}}, & -\sqrt{3} \le \xi < \sqrt{3} \\ 0, & o.\omega. \end{cases}$$

This distribution has the merit of being independent of the number M of amplitude levels employed in the modulation process.

Assumption on the Equalizer, $\hat{\mathbf{w}}$

We assume the equalizer, $\hat{w}_k$ is FIR of length $2L + 1$ and symmetric about the midpoint $k = 0$.

The output of the equalizer is given by

$$\begin{aligned} y(n) &= \sum_k \hat{w}(k)u(n-k) \\ &= \sum_k w(k)u(n-i) + \sum_k [\hat{w}(k) - w(k)]u(n-k) \\ &= x(n) + v(n) \end{aligned}$$

where $v(n)$ is called the convolutional noise representing the residual ISI that results from using an approximation, $\hat{\mathbf{w}}$ to the channel inverse, $\mathbf{w}$.

In the digital communication system, the equalizer output is applied to a zero-memory nonlinear estimator (decision device), $g(\bullet)$ which produces the estimate, $\hat{x}(n)$ of the data sequence $x(n)$

$$\hat{x}(n) = g(y(n)) \, .$$

Example: In decision-directed mode $g(\bullet) = \text{sgn}(\bullet)$ for binary systems (+/– 1), i.e. make decisions based on a positive or negative match (output from the matched filters). For *M*-ary systems we use a *slicer*.

We view the estimate, $\hat{x}(n)$ as a "desired response" (decision-directed) and employ an LMS update (Bussgang Algorithm)

$$\begin{aligned} \hat{x}(n) &= g(y(n)) \\ e(n) &= \hat{x}(n) - y(n) \\ \hat{\mathbf{w}}(n+1) &= \hat{\mathbf{w}}(n) + \mu \mathbf{u}(n) e(n) \end{aligned}.$$

Figure 18.5

Nonconvexity of the Cost Function

The cost function for the blind equalizer is defined by

$$\begin{aligned} J(n) &= E[e^2(n)] \\ &= E[(\hat{x}(n) - y(n))^2] \\ &= E[(g(y(n)) - y(n))^2] \end{aligned}.$$

In our study of adaptive filters, the MSE cost function is a quadratic (convex) function of the adaptive filter coefficients and therefore has a single, global minimum. However, the cost function for the blind equalizer is a nonconvex function and therefore may have many local minima in addition to global minima. The nonconvex cost function may result in ill-convergence of the equalizer.

Convergence of Bussgang Algorithm

The condition for convergence in the mean of the proposed blind equalizer is

$$E[u(n-i)\,y(n)] = E[u(n-i)\,g(y(n))], \text{ large } L; \quad i = 0, \pm 1, \ldots, \pm L.$$

It can be shown that this condition is equivalent to

$$E[y(n)\,y(n-k)] \approx E[g(y(n))\,y(n-k)], \text{ large } n \text{ and large } L.$$

Definition: A stochastic process $y(n)$ is said to be a Bussgang process if it satisfies the condition

$$E[y(n)\,y(n-k)] = E[g(y(n))\,y(n-k)]$$

where $g(\bullet)$ is a zero-memory nonlinearity. In other words, a Bussgang process has the property that its autocorrelation function is equal to the cross-correlation between that process and the output of a zero-memory nonlinearity produced by that process, with both correlations being measured for the same lag.

For the equalizer we have proposed, the process $y(n)$ is approximately a Bussgang process, provided L is large; the approximation becomes better as L is made bigger.

Convergence of the Bussgang algorithm is not guaranteed. In fact, the cost function of the Bussgang algorithm operating with a finite L is nonconvex and may therefore have false minima.

Benveniste-Goursat-Ruget (BGR) theorem

For the ideal case of a doubly infinite equalizer, convergence of the Bussgang algorithm is guaranteed provided the following hold.

1) The probability distribution of the data sequence $x(n)$ is sub-Gaussian, i.e. the kurtosis is less than 3

$$\kappa_x \equiv \frac{E[x^4]}{\{E[x^2]\}^2} < 3$$

2) The second derivative of

$$\Psi(y) = g(y) - y$$

is negative on the interval $[0, \infty)$.

Notes:

1) No zero-memory nonlinear function $g(\bullet)$ is known which would result in global convergence for the finite length equalizer.

2) There is a conjecture (Li and Ding) that the Bussgang algorithm will converge to a desired global minimum if the equalizer is long enough and initialized with a non-zero center tap, $\hat{w}_0(0) = 1$.