Equalisers for Digital Communications using Generalised Distance Measurement

C.P. Callender
Dept. of Electronic and Electrical Engineering
Loughborough University
Loughborough, Leics LE11 3TU, UK

C.F.N. Cowan
Dept. of Electronic and Electrical Engineering
Loughborough University
Loughborough, Leics LE11 3TU, UK

Abstract

An algorithm for non-linear equalisation involving the measurement of Mahalanobis distance has previously been published [1]. In this paper, we show that this algorithm makes the assumption that the clusters in the underlying observation space have a Gaussian distribution. If this assumption is violated, poor performance may be obtained. However, we show that the equaliser structure is capable of generating good approximations to the theoretical optimum decision boundary. It is the use of Mahalanobis distance which is inappropriate in the non-Gaussian case. By using a more general concept of distance, we demonstrate that it is possible to obtain significantly better results than those obtained using the Mahalanobis distance measure. The new method and previous algorithms are also extended to cover the case of a multi-level transmitted signals.

1 Non Linear Equalisation using Euclidean Distance

Figure 1 shows a model of a digital communications system. The transmitter generates a baseband signal which is chosen randomly from a set of \( L \) possible levels. The transmitted signal at time \( t \) will be denoted by \( a(t) \), so that \( a(t) \in \{ x_0, x_1, ..., x_{L-1} \} \). \( x_0 < x_1 < ... < x_{L-1} \) are the different levels in the transmitted code. The signal \( a(t) \) is passed through a linear communications channel which is modelled as a finite impulse response (FIR) filter. The output of the filter is given by the convolution

\[
\sum_{i=0}^{N_e-1} h_i a(t - i)
\]

(1)

\((h_0, h_1, ..., h_{N_e-1})\) is the sampled impulse response of the channel, which will be assumed to be time invariant. Noise in the communications system is modelled by adding a white Gaussian noise source to the output of this filter. Therefore the received signal is given by

\[
\Omega(t) = \sum_{i=0}^{N_e-1} h_i a(t - i) + \eta(t)
\]

(2)

where \( \eta(t) \) is the noise sample at time \( t \).

The function of an optimal equaliser is to provide an estimate of the transmitted symbol \( \hat{a}(t - \Delta) \) with the minimum probability that \( \hat{a}(t - \Delta) \neq a(t - \Delta) \). \( \Delta \) is a delay which needs to be introduced to simplify equalisation if the channel response has zeros which lie outside the unit circle (a non-minimum phase channel).

The information which the equaliser can use to perform this task is a finite number, \( N_e \), of observations of the received signal, \( \Omega(t) \). The equaliser therefore has an observation vector available to it given by:-

\[
X_t = \begin{bmatrix}
\sum_{i=0}^{N_e-1} h_i a(t - i) \\
\sum_{i=0}^{N_e-1} h_i a(t - i - 1) \\
\vdots \\
\sum_{i=0}^{N_e-1} h_i a(t - i - N_e + 1)
\end{bmatrix} + \begin{bmatrix}
\eta_t \\
\eta_{t-1} \\
\vdots \\
\eta_{t-N_e+1}
\end{bmatrix}
\]

(3)

Initially, the argument will be restricted to the noise free case, which is \( \eta(t) = \eta(t-1) = ... = 0 \). Then it is clear from (2) that the observation vector depends only on \( a(t), a(t-1), ... a(t-N_e-N_e+2) \). Since each of these symbols is randomly selected from a finite set of \( L \) levels, there are only \( L^{N_e+N_e-1} \) possible
observation points. This may be illustrated by an example. Consider a bipolar transmitted signal which is defined by \(a(t) \in \{-1,+1\}\). The example channel has \(N_c = 2\), \(h_0 = 1.0\) and \(h_1 = 0.5\) and an equaliser with \(N_c = 2\) will be used. In the noise free case, there are \(2^{2+2-1} = 8\) possible observations, corresponding to all possible values of \(a(t), a(t-1)\) and \(a(t-2)\). These points may be computed from (2) and are shown in figure 2.

When noise is added, such that \(\eta(t)\) has a normal distribution with zero mean and variance \(\sigma_n^2\), the effect is to replace each of the single observation points with a cloud of points. The centre of the cluster is at the same place as the original point in the noise free case, as a direct consequence of the noise having a mean of zero. The clusters for the example are illustrated in figure 3.

This graphical interpretation of the equalisation problem suggests the following equalisation strategy:

1. During training, the position of each of the \(L^{N_c+N_e-1}\) centres is identified, and labelled according to the transmitted sequence which generates it.
2. After training, when a new observation vector, \(X(t)\), is to be processed, the Euclidean distance from \(X(t)\) to each of the identified centres is computed. \(X(t)\) is then classified as belonging to the cluster to which it is closest. The label for this cluster may then be examined, yielding the transmitted symbol as required.

The acquisition and tracking of the positions of the centres of the clusters is achieved using averaging techniques.

### 2 Mahalanobis Distance Equalisation

The Euclidean distance algorithm gives near to the optimum solution to the equalisation problem, but the computational cost is high. It depends on the computation of \(L^{N_c+N_e-1}\) distances. If \(L, N_c\), and \(N_e\) are moderately large, this will represent a great computational burden. Moreover, \(N_c\) is a-priori unknown. A solution to this problem is to overestimate \(N_c\), which will result in some of the centres identified by the algorithm being identical. This will give good performance, but is an additional source of computation.

To solve these problems, we seek to group together a number of adjacent clusters in the observation space into a single cluster. However, the distribution of the larger clusters are no longer spherical and so Euclidean distance is no longer appropriate. An assumption will be made that the new clusters are elliptical. Under this approximation, the Mahalanobis distance, defined by

\[d_j^2 = (X(t) - c_j)^T \Sigma^{-1}(t)(X(t) - c_j)\]  

is the appropriate measure, where \(\Sigma(t)\) is the covariance matrix of dimension \(N_c \times N_e\). It is well known from the RLS adaptive filter algorithm [3] that \(\Sigma^{-1}(t)\) may be computed recursively without the need for explicit matrix inversion by

\[y_j(t) = (X(t) - c_j)^T\]

\[\Sigma^{-1}(t) = \frac{1}{\lambda} \Sigma^{-1}(t-1) - \Sigma^{-1}(t-1)y_j(t)\]

\[\cdot \left[1 + y_j^T \Sigma^{-1}(t-1)y_j\right]^{-1} y_j^T \Sigma^{-1}(t-1)\]

### 3 Failure of the Mahalanobis Distance Equaliser under certain circumstances

To illustrate a problem which can arise when using the Mahalanobis distance equaliser, consider again the example \(a(t) \in \{-1,+1\}\), \(N_c = 2\), \(h_0 = 1.0\), \(h_1 = 0.5\) and \(N_e = 2\). For the equaliser we will choose \(C=4\), so that each cluster contains two subclusters. It is then possible to calculate the centres, which are

\[c_0 = \begin{bmatrix} -1.5 \\ -1.0 \end{bmatrix}, c_1 = \begin{bmatrix} -0.5 \\ 1.0 \end{bmatrix}\]

\[c_2 = \begin{bmatrix} 0.5 \\ -1.0 \end{bmatrix}, c_3 = \begin{bmatrix} 1.5 \\ 1.0 \end{bmatrix}\]

and the covariance matrices

\[\Sigma_0^{-1} = \Sigma_1^{-1} = \Sigma_2^{-1} = \Sigma_3^{-1} = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{\sigma^2+\delta^2} \end{bmatrix}\]

These values may be substituted into (3) to give equations for the Mahalanobis distance metrics \(d_0^2, d_1^2, d_2^2\) and \(d_3^2\). The decision boundary, which separates regions of the observation space where a -1 decision is made from regions where a +1 decision is made may then be calculated by solving the equations \(d_0^2 = d_1^2(A), d_1^2 = d_2^2(B)\) and \(d_2^2 = d_3^2(C)\). (A) has the solution \(\Omega_t = -0.5\), (B) has the solution \(\Omega_t = \frac{0.25+\delta^2}{\delta^2}\), and (C) gives \(\Omega_t = 0.5\). These three line segments are shown in figure 4 for \(\sigma^2 = 0.025\). The Bayes optimum solution to the problem is shown in figure 3. It is clear that, for this example, the solution obtained using the Mahalanobis distance equaliser is not a good...
approximation to the optimum solution. As \( \sigma_i^2 \), the additive noise power becomes small, the gradient of line segment (B) becomes much steeper than the correct solution.

To examine more deeply the reasons for the failure, it is necessary to consider the justification for the use of Mahalanobis distance. In using a distance measurement to perform classification, it is assumed that the distance represents in some way the probability of observing an event. Observations which are a small distance from a centre are deemed to have a high probability of belonging to the cluster around that centre. Conversely, observations which are a large distance from a centre should have a low probability of belonging to that centre. The necessary condition is that distance should be a monotonically decreasing function of probability. If the assumption is made that the clusters have a multivariate normal distribution, then the probability of a point \( X(t) \) belonging to the cluster around \( c_j \) is

\[
p(X(t) \text{ belongs to cluster around } c_j) = e^{-(X(t) - c_j)^T \Sigma^{-1} (X(t) - c_j)}
\]

Taking the logarithm of both sides,

\[
(X(t) - c_j)^T \Sigma^{-1} (X(t) - c_j) = -ln(p(X(t) \in c_j))
\]

So, under the multivariate normal assumption, the Mahalanobis distance measurement represents the probability of an observation in the required way. This assumption is valid when the additive noise is high, but under low noise conditions, when each cluster is composed of a number of non-cohesive sub-clusters, it is invalid. This is the reason for the incorrect decision surface obtained in the previous example.

4 Generalised Distance Equaliser

Interestingly, the Mahalanobis distance equaliser is capable of forming good decision boundaries in high noise conditions. This demonstrates that the basic "structure" of the equaliser can give good results, but that the use of the inverse of the covariance matrix in the distance computation equation (4) is inappropriate, due to the close links between the covariance matrix and the normal distribution (7). In this section, a more general measure of distance is developed which is more appropriate for the low noise environment.

The development begins by noting the similarity in form of a general Volterra series to the Mahalanobis distance of equation (4). A quadratic Volterra filter acting on an input vector \( u_j \) to produce an output \( y_{volts} \) may be defined by

\[
y_{volts} = b_j + u_j^T l_j + u_j^T Q_j u_j
\]

where \( b_j \) is a bias coefficient, \( l_j \) is a vector of linear filter coefficients and \( Q_j \) is a (diagonally symmetric) matrix of quadratic filter coefficients. It is clear that this is equivalent to Mahalanobis distance measurement if the substitutions \( b_j = 0, l_j = 0, Q_j = \Sigma^{-1} \) and \( v_j = (X(t) - c_j) \) are made. However, it has already been demonstrated that these substitutions do not always lead to good equaliser performance. Instead the more general form of a Volterra series (equation 6) will now be used. An adaptive scheme will be developed to find the unknown parameters \( b_j, l_j \) and \( Q_j \).

To derive the algorithm, it is necessary to consider the purpose for which the distance measurement is to be used. Its purpose is to classify observations as to whether they belong to a cluster around a particular centre \( c_j \) or not. The use of the Volterra series as a classifier has previously been examined for a different problem in equalisation [2]. In this work, it was found to be necessary to use a sigmoidal non-linearity at the output of the Volterra series to enhance the classification performance. Our results have confirmed this finding. The generalised distance measure is therefore defined by

\[
d_j^2 = f(y_{volts}) = f(b_j + u_j^T l_j + u_j^T Q_j u_j)
\]

where

\[
f(\theta) = \frac{1}{1 + e^{-\theta}}
\]

The parameters \( b_j, l_j \) and \( Q_j \) are updated using the LMS adaptive algorithm. Since the Volterra series is operating as a classifier and the sigmoid non-linearity gives an output of between 0.0 and 1.0, it is reasonable to define a desired response signal during training by:

\[
\zeta = 0.0 \text{ if the current observation } X(t) \text{ belongs to cluster } j.
\]

\[
\zeta = 1.0 \text{ otherwise}
\]

The LMS update algorithm may then be expressed as:

\[
\text{Compute } d_j^2 \text{ using (9)}
\]

\[
e_j = d_j^2(1 + d_j^2) (c_j(t) - d_j^2) \\
b_j = b_j + 2 \mu e_j \\
l_j = l_j + 2 \mu e_j (X(t) - c_j) \\
Q_j = Q_j + 2 \mu e_j (X(t) - c_j)^T (X(t) - c_j)
\]
5 Comparison of the Algorithms

Figure 5 compares the error rates obtained using the three non-linear algorithms as well as that of a linear equaliser. The transmitted signal, channel and equaliser are the same as for the previous examples in this paper. The Euclidean distance equaliser gives the best performance. The underlying assumption which is made in using this algorithm is that the clusters are spherical, which is valid as the additive noise is white. However, this algorithm may be very computationally intensive over more realistic channels which would yield a large number of centres. The Mahalanobis distance equaliser gives poor performance at high signal to noise ratios, as would be expected from the previous discussion on the decision boundary. Indeed, at high signal to noise ratios, it performs less well than the linear equaliser. This particular channel example does tend to exaggerate the problem, and this result does not follow for all channels. Nevertheless, it is indicative of the problems which may arise when using Mahalanobis distance classification in equalisation. The generalised distance algorithm does not suffer from these problems, since no assumptions are made about the distribution of the underlying clusters. It gives virtually the same performance as the Euclidean distance equaliser but allows grouping of adjacent clusters to form larger clusters, which is necessary to obtain a reasonable computational complexity when the channel has a long impulse response.

6 Conclusions

The generalised distance equaliser developed in this paper has been shown to give superior performance compared to the Mahalanobis distance equaliser for a specific channel example. Its computational complexity and methodology are similar to the Mahalanobis equaliser, but it does not rely on the probability distributions of the clusters, and so its use may be preferable.

References


Figure 5: Comparison of the error rates obtained using the three non-linear algorithms and a linear equaliser.