Conclusions: A new block iterative DFE has been presented. The new structure has a significantly lower complexity with respect to TD-DFE, while having a similar performance. Moreover, the new structures operate on blocks, thus allowing the use error correction on the feedback signal.

References

Blind identification of non-stationary MA systems

E. Alameda, D.P. Ruiz, D. Blanco and M.C. Carrion

A new adaptive algorithm for blind identification of time-varying MA channels is derived. This algorithm proposes the use of a novel system of equations derived by combining the third- and fourth-order statistics of the output signals of MA models. This overdetermined system of equations has the important property that it can be solved adaptively because of their symmetries via an over determined recursive instrumental variable-type algorithm. This algorithm shows good behaviour in arbitrary noisy environments and good performance in tracking time-varying systems.

Introduction: The use of adaptive algorithms is widely spread in several fields of science, from echo cancellation in communications to cardiology measures in medicine [1-3]. The least mean squares (LMS) algorithm could be considered the most common one; it is based on the minimisation of a quadratic cost function using the steepest descent method. Other algorithms, e.g. the recursive least squares (RLS) algorithm, impose a least squares condition. In both cases, the algorithm leads to a system of linear equations, where the coefficients bear the information of the statistics of the problem. The RLS algorithm only is able to solve systems of equations with square matrices involving second-order statistics. To include the properties and the information provided by higher-order statistics (HOS) e.g. Gaussian noise insensitivity and phase information, the recursive instrumental variable algorithm (RIV) was derived from RLS. However, these algorithms are not valid if overdetermined systems are to be solved. For this purpose the ORIV algorithm and its variants (as in [1]) has to be used. Overdetermined systems of equations based on HOS appear, for example, in the problem of blind identification of MA systems [2]. In this Letter, a new system of equations relating the coefficients of an MA model using the higher-order statistics of the output is derived and its resolution via ORIV is compared with another system of equations based on mixed statistics (second- and third-order). The most important property of the proposed adaptive algorithm is that it provides an estimate for every instant of time using all the information up to that time with better behaviour in terms of mean square deviation in noisy environments.

Problem and proposed method: The problem of identifying an MA channel only from output measurements, probably corrupted by additive Gaussian noise of unknown power spectrum, is considered. The data available is \( y(n) = x(n) + v(n) \), where \( v(n) \) is an additive Gaussian noise, independent of the input \( x(n) \), and \( x(n) \) is the output of an MA channel of order \( q \) defined as:

\[
x(n) = \sum_{i=0}^{q} b(i)w(n-i) \quad \text{with} \quad b_0 = 1
\]  

where \( b(i) \) is the model coefficients and \( w(n) \) is a fourth-order stationary and ergodic, non-Gaussian i.i.d process of zero mean and non-zero variance \( \sigma_w^2 \). In what follows, the mean value of the signal will be assumed to be zero.

Under the previous assumptions the Brillinger-Rosenblatt [1] equation holds

\[
c_{k}(t_1, t_2, \ldots, t_{k-1}) = \gamma_{k-1} \sum_{i=0}^{q} b(i)b(i+1+k)
\]  

for the special case where \( k = 4, t_2 = t_3 = 0 \) and \( t_1 \) is replaced by \( t_1 + t \), the form

\[
c_{4}(t_1 + t, 0, 0) = \gamma_{4} \sum_{j=0}^{q} b(j)b(j+1+t)
\]  

Multiplying both sides by \( b^2(j) \) and taking summation on \( j \) leads to

\[
\sum_{j=0}^{q} b(j) \sum_{i=0}^{q} b(i)b(j) = \gamma_{4} \sum_{j=0}^{q} b(j)\sum_{i=0}^{q} b(i)(j+1+t)
\]  

In (4) the summation over \( j \) in the left-hand side can be replaced by a third-order cumulant and using the symmetry properties of cumulants, we can get the desired expression

\[
\sum_{i=0}^{q} b(i)c_{3}(m-i) = \gamma_{4} \sum_{i=0}^{q} b(i)c_{3}(m-i)
\]  

with \(-q \leq m \leq 2q\) where for notational simplicity the following identities are assumed: \( c_{3}(m-i) = c_{3}(m-i) \), \( c_{4}(m-i) = c_{4}(m-i) \), \( e_{5}(m-i) = e_{5}(m-i) \), \( e_{6}(m-i) = e_{6}(m-i) \), which are the values of the third- and fourth-order cumulants along the main diagonal slice. In addition, \( t_1 \) is replaced by \( m \).

For the second- and third-order case, an analogous expression can be obtained:

\[
\sum_{i=0}^{q} b(i)c_{3}(m-i) = \sum_{i=0}^{q} b(i)c_{3}(m-i)
\]  

with \(-q \leq m \leq 2q\). This equation was previously obtained using a different framework in [3] and is known as the Giannakis-Mendel equation. Both (5) and (6) share the same property: they can span as an overdetermined linear system of equations, the matrices of which have
certain symmetries, mainly because they are a cross-correlation between the input vector and a defined instrumental variable vector $d(t)$

$$E[x(t), y(t)] = E[x(t)]$$  (7)

that are exploited by ORIV [1] and references therein) to compute a solution in an efficient way for every instant of time. Depending on the system of equations to be solved, the instrumental variable vector, the input vector and the desired response that appear in ORIV have to be defined. For the proposed equation the following definitions hold: $x(t) = [x_1(t), \ldots, x_{n-1}(t), x_n(t)]^T$, $y(t) = [y_1(t), \ldots, y_{n-1}(t), y_n(t)]^T$ and $d(t) = [d_1(t), \ldots, d_{n-1}(t), d_n(t)]^T$ for system (5) and $y(t) = [y_n(t)]^T$ for system (6) where $d(t)$ is an estimator of the variance of the output $y(t)$ at time $t$. In this way ORIV gives a vector of estimates which are function of the MA coefficients: $\theta_0 = [\theta_1, \ldots, \theta_{n-2}, \theta_{n-1}, \theta_n]$ with $\theta_{n+1} = \theta_n/\sigma^2$ for system (5) and $\theta_0 = [\theta_1, \ldots, \theta_{n-2}, \theta_{n-1}, \theta_n]$ with $\theta_{n+1}^2 = \sigma^2/\gamma_1$ for system (6). From these, estimates of the coefficients can be easily obtained.

Simulations results: Two set of simulations have been carried out: the first studies the ability of the method for tracking time-varying systems; the second studies the influence of additive Gaussian noise on the estimates. In either case, the coefficients of the MA model are obtained using the proposed system of equation (5) combining third- and fourth-order statistics and solved via the ORIV algorithm; this method is named ORIV-C3C4. To compare the goodness of the behaviour of the proposed method, the system of equation (6) is also solved by the same algorithm (ORIV) and this method is called ORIV-C2C3.

The first set of simulations show the possibility of tracking time-varying systems. Since the proposed algorithm can provide an estimate for every instant of time using all the information up to that time, it is possible to plot how the algorithm learns with time, i.e. how the algorithm reduces the mean square deviation $\text{MSD}(\theta) = E[\theta(t) - \theta_0]^2$ of the estimators. A non-stationary MA model is now considered and a forgetting factor $\lambda = 0.9985$ used. The MA system is supposed to change linearly with time according to $b(n) = b_0 + \rho n$ where $\rho$ is the vector of the slope of the linear change. Fig. 1 shows the convergence in the mean of parameter $b_2$ for a model with $b_0 = [1 - 0.75]$ and $\rho = 0 - 1/16000 \{1/16000\}$. This figure shows the results for both methods ORIV-C3C4 and ORIV-C2C3. As can be seen from the figure, both methods suitably track the parameter evolution of the time-varying system. The method ORIV-C2C3 is faster to track the true evolution of the parameter, i.e. it is faster to evolve linearly and therefore it has a smaller convergence time. This is because this is that it requires less data to estimate the second-order statistics than higher-order statistics and therefore ORIV-C2C3 convergence is quicker.

Fig. 1 Evolution of $b_2$, true and estimated via ORIV-C2C3 and ORIV-C3C4.

Fig. 2 Evolution of MSD for several SNR: ORIV-C2C3 and ORIV-C3C4.

Conclusion: The symmetries derived from the proposed equation allows use of iterative algorithms like ORIV to estimate system parameters. For $\text{SNR} \leq 10 \text{dB}$, the proposed method ORIV-C3C4 is preferred compared to methods using second-order statistics because it gives better estimates of the model parameters.

Acknowledgement: This work was partially supported by the 'Ministerio de Ciencia y Tecnología' of Spain under Project TIC 2001-2902.

Table 1: Estimates of MA parameters after 5000 iterations for ORIV-C3C4

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>$b_1$</th>
<th>Bias ($b_1$)</th>
<th>Var($b_1$)</th>
<th>$b_2$</th>
<th>Bias ($b_2$)</th>
<th>Var($b_2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>-1.1348</td>
<td>0.0034</td>
<td>0.5626</td>
<td>0.0774</td>
<td>0.2609</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-1.1361</td>
<td>0.0047</td>
<td>0.7091</td>
<td>0.1301</td>
<td>0.4180</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-1.1159</td>
<td>0.0155</td>
<td>0.1487</td>
<td>0.1209</td>
<td>0.3586</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>-0.8125</td>
<td>0.3398</td>
<td>0.4942</td>
<td>0.7699</td>
<td>0.1299</td>
<td>0.5588</td>
</tr>
</tbody>
</table>

Table 2: Estimates of MA parameters after 5000 iterations for ORIV-C2C3

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>$b_1$</th>
<th>Bias ($b_1$)</th>
<th>Var($b_1$)</th>
<th>$b_2$</th>
<th>Bias ($b_2$)</th>
<th>Var($b_2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>-1.1256</td>
<td>0.0054</td>
<td>0.6259</td>
<td>0.0141</td>
<td>0.0370</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-0.0210</td>
<td>0.1104</td>
<td>0.3810</td>
<td>0.2590</td>
<td>0.1820</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.7010</td>
<td>0.4304</td>
<td>0.0890</td>
<td>0.0120</td>
<td>0.2240</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>-0.0889</td>
<td>1.0425</td>
<td>0.3820</td>
<td>0.0079</td>
<td>0.6321</td>
<td>0.7730</td>
</tr>
</tbody>
</table>
Modifier formula on mean square convergence of LMS algorithm

Yuantao Gu, Kun Tang, Huijuan Cui and Wen Du

In describing the mean square convergence of the LMS algorithm, the update formula based on independence assumption will bring explicit errors, especially when step-size is large. A modifier formula that describes the convergence well is proposed. Simulations support the proposed formula in different conditions.

Introduction: The LMS algorithm has widespread use in many applications [1]. Convergence analysis on LMS has been researched in detail for many years. Independence theory, which may be traced back to Widrow and Mazo in the 1970’s, is a powerful tool to perform the analysis, for the deduction can be simplified greatly by assuming no independence existing between the filter coefficient error and the input vector. The disadvantage of independence assumption is obvious. Though it figures out the steady-state mean square error (MSE) exactly, it loses accuracy while describing the convergence process. Especially when a larger step-size is used to accelerate the convergence, the difference between theory and experiment becomes more obvious. In this Letter, we propose a modifier formula, which describes the MSE convergence well.

Modifier formula: Based on LMS criterion, the update of filter coefficient vector \( \mathbf{w}(n) \) is

\[
\mathbf{w}(n+1) = \mathbf{w}(n) + \mu(n)\mathbf{e}(n)
\]

where \( \mathbf{e}(n) \) is the input vector, \( \mu \) is the step-size, and \( \mathbf{e}(n) \) is the estimation error:

\[
\mathbf{e}(n) = \mathbf{x}^T(n) \mathbf{c} - \mathbf{x}^T(n) \mathbf{w}(n) + \mathbf{v}(n)
\]

where \( \mathbf{c} \) is the unknown system coefficient (without specification, the vectors in this Letter are all \( N \) length), \( \mathbf{v}(n) \) is the additive noise in observation. We define the filter coefficient error as \( \mathbf{g}(n) \triangleq \mathbf{w}(n) - \mathbf{c} \). Using (1) and (2), we get the update equation of filter coefficient vector

\[
\mathbf{g}(n+1) = A(n)\mathbf{g}(n) + b(n)
\]

where \( A(n) \triangleq 1 - \mu(n)\mathbf{e}(n) \mathbf{e}(n)^T \), \( b(n) \triangleq \mu(n)\mathbf{e}(n) \mathbf{c} \).

Independence theory assumes the input vector \( \mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(n) \) are statistically independent and draws a conclusion that independence also exists between \( \mathbf{g}(n) \) and \( \mathbf{x}(n) \). This assumption decreases the complexity in analysing (3) considerably. If the input signal \( \mathbf{x}(n) \) is i.i.d Gaussian signal, we can draw the MSE update equation easily [2]

\[
E[\|\mathbf{g}(n+1)\|^2] = \beta E[\|\mathbf{g}(n)\|^2] + \gamma
\]

where

\[
\beta = 1 - 2\mu^2 + (N + 2)\mu^2 \sigma_v^2, \quad \gamma = N\mu^2\sigma_v^2\sigma_c^2
\]

Notice that \( E[A(n)\mathbf{g}(n)] = \beta \mathbf{I} \) while in most applications, \( \mathbf{x}(n) \) is a successive vector, which means that the oldest \( N-1 \) samples of \( \mathbf{x}(n) \) are identical with the newest \( N-1 \) samples of \( \mathbf{x}(n-1) \). Therefore the independence assumption is far-fetched, which lead to a large error existing in predicting the convergence using (4). We propose a modifier formula to make the prediction more accurate:

\[
E[\|\mathbf{g}(n+1)\|^2] = f\beta E[\|\mathbf{g}(n)\|^2] + \gamma + k
\]

where

\[
f = 1 - \frac{1}{N} (a_0 + a_1\mu + a_2\mu^2 + a_3\mu^3)
\]

Notice that \( a_0, a_1, a_2, a_3 \) is the parameter vector, which may be traced back to Widrow and Mazo in the 1970’s, is a powerful tool to perform the analysis, for the deduction can be simplified greatly by assuming no independence existing between the filter coefficient error and the input vector. The disadvantage of independence assumption is obvious. Though it figures out the steady-state mean square error (MSE) exactly, it loses accuracy while describing the convergence process. Especially when a larger step-size is used to accelerate the convergence, the difference between theory and experiment becomes more obvious. In this Letter, we propose a modifier formula, which describes the MSE convergence well.

Explanation: We can rewrite MSE without recursion

\[
E[\|\mathbf{g}(n+1)\|^2] = E[\|\mathbf{g}(n)\|^2] + \gamma
\]

where

\[
s(n) = E[\|\mathbf{g}(n)\|^2]
\]

\[
l(n, j) = E[\mathbf{b}^T(n)\mathbf{A}^T(j + 1)\mathbf{A}(n)\mathbf{A}(n)\mathbf{A}(j + 1)\mathbf{b}(j)]
\]

With independence assumption, the following relation is deduced:

\[
s(n) = \beta s(n-1) + l(n, j)
\]

However, with its definition, we know \( A(n) \) and \( \mathbf{A}(n) \) are correlated while \( |n - n'| < N \). Therefore (13) will produce a large error because \( A(n)\mathbf{A}(n) \) inside the expectation of the matrix multiplication series (both in (11) and (12)) is extracted. The expectation is obtained first, which is why (4) loses accuracy. However with this hypothesis, we can still presume that the relation between \( s(n) \) and \( s(n-1) \) is closely connected with \( \beta \). In addition, we consider that the relation between \( s(n, j) \) and \( s(n-1, j) \) in very similar to that between \( s(n) \) and \( s(n-1) \).

We then speculate the following assumption:

\[
s(n) = f s(n-1) + k(n, j)
\]

where \( f \) is a modifier constant, and \( k(n, j) \) and \( k(n, j) \) are modifier variables.

Though no analytical explanation can be provided, we still take it for granted that the error produced by (14) is smaller than that by (13). Using (14) in (10) we get

\[
E[\|\mathbf{g}(n+1)\|^2] = f\beta E[\|\mathbf{g}(n)\|^2] + \gamma + k
\]

where

\[
k(n, j) = \frac{\sum_{m=0}^\infty (1 - \mu^2)^m s(m, j)}{(1 - \mu^2)^N - 1}
\]

Since the original formula (4) predicts the steady state MSE well, we force (15) to have the same steady-state MSE as (4). We then get \( k = (1 - f\beta)(1 - 1/N) \). Where \( k \) is the immediate outcome, the specific structure of (7) and the parameters of (9) are got experimentally. What is remarkable is that \( \mu^2 \) in (8), culled relative step-size, is defined to remove the influence of \( N \) and \( \sigma_v^2 \), and it will be used instead of \( \mu \) in the following Section.

Simulation: In each special experiment, 10 unknown system responses are modelled by 50 input series separately. Each system response is an \( N \) points random vector, and \( |\mathbf{e}|^2 = 1 \), while the input series is i.i.d Gaussian random signal, and \( \sigma_v^2 = 1 \). We average all of these 500 study curves as the experimental curve, which is denoted by \( E(n) \). Conversely, the theoretical curve updated with (4) is denoted by \( \tilde{E}(n) \) and that updated with (6) by \( \tilde{E}_2(n) \). The mean error and maximum error between the experiment and the theory are defined

References