AN IMPROVED LS ALGORITHM FOR THE ESTIMATION OF AN IMPULSIVE NOISE CORRUPTED SIGNAL BY LINEAR PROGRAMMING TECHNIQUE

E.D. Di Claudio*, G. Orlandi**, F. Piazza, A. Uncini

Dept. of Electronics and Automatics, University of Ancona, Ancona, Italy
(*) ELASIS S.p.A., Chieti Scalo, Italy
(**) Dept. INFO-COM, University - "La Sapienza", Rome, Italy

ABSTRACT

Least Squares algorithms are often used in many spectrum estimation methods. However when the signals are contaminated by few strong noise spikes, the standard LS algorithm can easily lead to biased solutions characterized by a strongly reduced dynamic range of the estimated spectra. In order to face this problem the classical approach is to weight the prediction errors before applying the LS minimization algorithm.

In this paper a procedure for assigning optimal weights to the prediction equations is presented. The set of weights is computed, by linear programming techniques, in order to reduce the effects of strong impulsive noise. In order to demonstrate the capability of the proposed approach, the algorithm has been tested with signals corrupted by stationary white noise, impulsive additive spikes and a combination of both. The obtained results show a high degree of robustness that makes the method attractive for automatic analysis of real-world data.

1. Introduction

Least Squares (LS) estimation algorithms [1] are well known techniques for the computation of AR model parameters by linear prediction methods, for the adaptive filtering of signals, eventually multidimensional, and for the solution of other related problems (interpolation, deconvolution, etc.). However, as pointed out by Cadzow [2] and by Tufts and Kumaresan [3], the presence of noise can worsen the obtainable results. Noise increases the actual rank of the data matrix and perturbs the estimate of the signal subspace basis vectors.

In particular the presence of a small number of short-time strong pulses, can induce large projections of the predicted sample vector onto the noise subspace. This may increase the difficulty of LS parameter estimation by reducing the average difference between the signal and noise related eigenvalues of the involved data correlation matrix.

In this case, a LS estimate of the AR model parameters by linear prediction usually results in flattened spectra. In fact, since an AR model assumes the stationarity of the process during the observation time, noise pulses of very small duration can easily affect the results of the LS algorithm over the entire data segment. Another way to explain this behaviour is to observe that the norm of the prediction error vector is actually determined by the equations involving severely corrupted data samples.

In harmonic analysis, the traditional way to cope with the problem of the presence of noise, is that of setting an higher prediction order and decomposing data or correlation matrices by SVD or EVD techniques [2,3]. However, when the noise is impulsive, these methods cannot obtain satisfactory results since a large number of prediction equations is severely perturbed even by a single highly corrupted sample and therefore the whole correlation matrix results biased.

In order to face this problem, the corresponding equations can be weighted according to a suitable criterion of noise identification and suppression. One possible approach is first to apply the LS algorithm to the data and then compute a set of weights using suitable influence functions [4]. These weights are used in a further LS minimization stage and the whole procedure is repeated until the results stabilize. Another approach is to set the weights before applying the minimization procedure. These weights are chosen in order to reduce the influence of the equations corrupted by the noise.

Following the second approach, in this paper a procedure for assigning optimal weights to the prediction equations is presented. The set of weights is computed by linear programming techniques, reducing the effects of strong impulsive noise.

2. The Proposed Algorithm

For linear prediction problems, we can build the classical forward-backward equation matrix with N available signal samples \( \{x_i, \ i = 1, \ldots, N\} \) and \( p \) unknown regression parameters \( \{a_i, \ i = 1, \ldots, p\} \) in the following way:

\[
x_n = - \sum_{k=1}^{p} a_k x_{n-k} \quad n = (p+1),\ldots,N \quad \text{(forward)} \tag{1}
\]

\[
x^*_{n} = - \sum_{k=1}^{p} a_k x^*_{n+k} \quad n = 1,\ldots,(N-p) \quad \text{(backward)} \tag{2}
\]

or, in matrix notation:

\[
X^t \ a = - \ x^t \ \text{(matr)} \ (p \times 1) \quad \text{(mat1)} \tag{3}
\]
with obvious meaning of symbols. Then 2(N - p) = m > p
 equations are obtained, resulting in an overdetermined linear
 system in the unknown vector a.
 If a diagonal (mxm) weight matrix W is introduced, with
 nonnegative diagonal elements w_i (i=1,...,m), the
 overdetermined system of equations becomes:

\[(WX)a = - (Wx)\quad W = \text{diag}[w_i \geq 0, \quad i = 1,\ldots,m]\]  \hspace{1cm} (4)

The proposed algorithm allows to compute the matrix W in
order to reduce the influence of impulsive noise onto the
estimate of a. The algorithm is based on the following issues:
1) Better AR parameters are often obtained by weighting each
 equation proportional to the square root of the energy of the
 samples contained in the prediction window [5]. On the other
 hand, this energy can suddenly increase when a spike enters
 the window.
2) It is well known that the use of the L1 norm let better cope
 with the presence of spikes [6]. In fact, minimizing a cost
 function with this norm can easily evidentiate sharp
 discontinuities in it.
3) Predictable (correlated) signals show generally smaller
 envelope fluctuations than wide-band noise pulses.
 In order to find the optimal weights matrix W, the following
 quantities are computed:

\[E_i = \sqrt{\frac{1}{p} \sum_{j=1}^{p} |X_{ij}|^2} \quad P_i = |x_i| \]  \hspace{1cm} (5)

where i=1,...,m, X_{ij} is the (i-th,j-th) element of X and x_i is
the i-th element of X. These values represent respectively the
square root of the average signal energy in each prediction
window (E_i) and the target signal envelope (P_i). The
computation of E_i can be simplified using the symmetry
relationship reported in [5].
A strong difference between the E_i and P_i values indicates a
possible large prediction error due to a sudden change of the
signal statistics, for example the presence of a noise impulse.
Also a sudden non-stationarity of the signal itself can produce
such effect, in this case the AR estimation can be
impaired since the poles may move away from the unit circle.
The scope of the algorithm is to compute w_i values which
minimize the influence of the prediction equations that
 correspond to windows with small energy or/and with a
sudden change in the signal statistics. The problem can be
formalized as a linear programming one by choosing a cost
function as:

\[C = \sum_{i=1}^{m} w_i \cdot E_i \cdot P_i \]  \hspace{1cm} (6)

and set of suitable linear constraints:

\[w_i \leq E_i \quad i = 1,\ldots,m\]  \hspace{1cm} (7)

\[\sum_{i=1}^{m} w_i \geq \alpha \sum_{i=1}^{m} E_i \quad 0 < \alpha < 1\]  \hspace{1cm} (8)

The group of disequalities (7) imposes a set of constraints
 similar to that in [5]. It minimizes the influence of equations
 relating consecutive samples with a small average energy.
 It is expected that these signal segments have an higher
 stationary noise content. The last equation (8) simply avoids
the convergence to the trivial solution with all zero weights.
The parameter \(\alpha\) can be selected through the following

simple formula:

\[\alpha = \frac{m}{\sum_{i=1}^{m} E_i} \]  \hspace{1cm} (9)

where Int(.) indicates the floor operator and the E_i values
are ordered in a descending manner.
If the observed process is wide sense stationary, the highest
E_i values should indicate spike-contaminate prediction
windows, while the remaining ones should exhibit a small
variance with respect to the mean value of signal energy.
Therefore, a good estimate for the \(\alpha\) parameter should be a
number proportional to the ratio between the average of the
energy of uncontaminated windows and the total average
energy.
Expression (9) reports a simple estimate for the \(\alpha\) parameter,
obtained by dividing the prediction windows into two
different classes: the first comprising windows with high
probability to be spikes-contaminated (large values of E_i),
the second comprising windows with high probability to be
spike-free (small values of E_i). Typically the values for the K
parameter can be chosen around 0.5. In transient signal
analysis, a better estimate of \(\alpha\) can be obtained by clustering
the E_i's into more than two classes, for example by adding to
the previous two classes a third one comprising windows
where the signal content is very low.
The optimal weights \(w_i (i=1,\ldots,m)\) are then obtained by a
standard linear programming technique, for example with the
simplex method. The final weighted LS solution \(\hat{a}\) for
the prediction vector is computed through the pseudoinverse
\((WX)^{+}\) of the matrix \(WX\).

3. Experimental Results
In order to demonstrate the capability of the proposed
method, several experiments have been performed, whose
results are briefly reported here.
At first, a simple AR parameter estimation experiment is
presented. The signal is composed by 40 samples of two
complex exponentials with no white additive noise:

\[x_n = e^{j(2\pi f_1 n)} + e^{j(2\pi f_2 n + \pi/4)}\]  \hspace{1cm} (10)

with \(n=1,\ldots,40, f_1=0.15\) and \(f_2=0.2\); corrupted by adding 3
spikes \(S_k, k=1,2,3\):

\[S_1 = 10 + j \quad \text{for } n = 5\]  \hspace{1cm} (11)

\[S_2 = -10 + j7 \quad \text{for } n = 9\]

\[S_3 = 10 - j5 \quad \text{for } n = 22\]

Two linear prediction vectors, respectively of order \(p=2\) and
\(p=4\), have been computed by means of either the standard
Forward Backward Least Squares Algorithm (FBLS) [7], or
by the proposed weighted LS approach. In both cases, the
pseudoinverse of the data matrix was computed by the SVD
algorithm reported in [8]. For uncorrupted data the rank of
X remains fixed at two for any greater prediction order
[3]. In our case, however, the effective rank of the matrix
WX is determined by the spread between the largest singular
values (belonging to the signal subspace) and the smallest
ones (introduced by the noise).
After adding the three spikes, the FBLS algorithm is unable to detect and resolve the two complex exponentials for a selected order ps. The singular values (Table 1) show no clear distinction between signal and noise subspaces. The resulting AR spectrum is therefore severely perturbed (see Fig. 1). Several other trials with higher prediction orders do not exhibit any improvement in detectability of signals, because several spurious peaks appear in the spectra which mask those generated by the noiseless signal.
Using the proposed method, the eigenvector W has been computed after selecting a value of K=0.5 (this choice is shown to be not critical) and prediction orders p=4 and p=2.
As shown in Figure 2, in both cases the algorithm assigns very small or zero weights to the equations contaminated by the impulsive noise. The resulting AR spectra are almost ideal, as shown in Figure 3 for order 4. The corresponding singular values are reported in Table 1.
In the case of p=4, two singular values of the matrix WX are very small, indicating that the impulse noise has been perfectly suppressed and the ideal rank of the noiseless matrix has been restored. The AR spectrum was computed after suppressing these small singular values and projecting the vector x onto the estimated signal subspace, spanned by the two dominant singular vectors [3].
In the second experiment, the robustness of the method with respect to additive white Gaussian noise is tested. When such noise is present, the standard FBLS covariance estimate should give the best results in a maximum likelihood sense [4]. Due to the presence of this wideband stationary noise, the prediction order has been increased to 10 to achieve the desired frequency resolution. In all trials, two clearly dominant singular values have been obtained. Then, the rank of each pseudo-inverse has been forced to a value of 2. Table 2 summarizes the results for the signal described by Eq. (10) (without any spike) in terms of sample mean and variance of the estimated AR coefficients on 50 trials with a SNWR of 10 dB.
As last experiment, the proposed method has been applied to the test signal (10) corrupted by the spikes (11) and a single realization of an additive white noise with a SNWR of 10 dB. Figure 5 shows the spectrum estimated by a 10th order AR model (k=0.7) and a reduced rank of 2.

4. Limits on applicability

The proposed approach is more effective when the gaussian white noise component is smaller than both spikes and useful signal. If this hypothesis is satisfied, the algorithm will work properly, since in most cases the number of severely underweighted equation is small and it performs actually as the basic Covariance Least Squares algorithm of [5].

With the suppression of noise pulses and localized nonstationarities, the effective rank of the matrix WX is often reduced. This fact allows to reduce also the order of the AR model, thus increasing the number of "useful" equations since a lower number of prediction windows are now contaminated by spikes. However, although an optimized simplex algorithm exploiting the quasi-banded nature of the involved matrix is used, the computational cost remains several times higher than that required by the classical LS minimization.

Experimental results indicate that the LS matrix (WX)

* (WX) can be nearly singular for the selected order p, requiring the use of truncated pseudo-inverses through SVD or EVD decompositions.
More advantages are expected by applying the proposed weighting approach to data matrices used in multichannel filtering or in direction finding problems.

REFERENCES


Table 1. AR prediction coefficients (first block of rows) and singular values (second block of rows), computed setting a prediction order of 2 (first 3 columns) and of 4 (last 3 columns) by using respectively the noiseless signal, the signal plus impulsive noise with the standard FBLS algorithm and the signal plus impulsive noise with the proposed algorithm.

<table>
<thead>
<tr>
<th>AR coeff.</th>
<th>noless signal system order = 2</th>
<th>standard FBLS system order = 2</th>
<th>proposed method system order = 2</th>
<th>noless signal system order = 4</th>
<th>standard FBLS system order = 4</th>
<th>proposed method system order = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.868+1.7601</td>
<td>0.0584 -0.00915</td>
<td>-0.868+1.7601</td>
<td>-0.4395 -0.8625</td>
<td>-0.02493 -0.0841</td>
<td>-0.4395 -0.8625</td>
<td></td>
</tr>
<tr>
<td>-0.5878+1.08090</td>
<td>0.0438 -0.00959</td>
<td>-0.5878+1.08090</td>
<td>0.2851 -0.3923</td>
<td>0.0561 -0.0629</td>
<td>0.2850 -0.3923</td>
<td></td>
</tr>
<tr>
<td>-0.1500-2.4801</td>
<td>0.2782 -0.1453</td>
<td>-0.1500-2.4801</td>
<td>0.2782 -0.1453</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.1175</td>
<td>30.9573</td>
<td>23.0328</td>
<td>22.9061</td>
<td>33.3782</td>
<td>33.5513</td>
<td></td>
</tr>
<tr>
<td>1.3885</td>
<td>27.4027</td>
<td>0.3921</td>
<td>4.3040</td>
<td>27.5275</td>
<td>2.0704</td>
<td></td>
</tr>
<tr>
<td>2.85e-6</td>
<td>27.3861</td>
<td>3.51e-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.54e-6</td>
<td>27.3861</td>
<td>3.51e-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Mean and variance, on 50 trials, of the 10th order AR prediction coefficients, computed with the proposed algorithm by using the test signal corrupted by additive white noise with SNWR of 10 dB.

| coef.  | noiseless values | variance |  | noiseless values | variance |  |  |  |  |  |  |
|--------|------------------|----------|  | mean |  |  | mean |  |  |  |  |  |  |  |  |
|        | real part | imaginary part |  | real part |  |  | imaginary part |  |  |  |  |  |  |  |  |
| coef. 1 | -1.516494E-01 | -2.976287E-01 |  | -1.474619E-01 |  |  | -2.838264E-01 |  |  |  |  |  |  |  |  |
| coef. 2 | 1.691808E-01 | -2.328574E-01 |  | 1.610047E-01 |  |  | -2.238031E-01 |  |  |  |  |  |  |  |  |
| coef. 3 | 2.316440E-01 | 3.668880E-02 |  | 2.373917E-01 |  |  | 3.405505E-02 |  |  |  |  |  |  |  |  |
| coef. 4 | 5.422020E-02 | 1.668726E-01 |  | 5.495876E-02 |  |  | 1.619391E-01 |  |  |  |  |  |  |  |  |
| coef. 5 | -7.924453E-02 | 7.924452E-02 |  | -7.704347E-02 |  |  | 7.783294E-02 |  |  |  |  |  |  |  |  |
| coef. 6 | -4.367026E-02 | -1.418933E-02 |  | -4.93688E-02 |  |  | -1.596901E-02 |  |  |  |  |  |  |  |  |
| coef. 7 | 3.342070E-03 | 2.110099E-02 |  | 5.773560E-03 |  |  | 1.322820E-02 |  |  |  |  |  |  |  |  |
| coef. 8 | -7.129037E-02 | 5.179549E-02 |  | -6.120918E-02 |  |  | 4.672480E-02 |  |  |  |  |  |  |  |  |
| coef. 9 | -1.360617E-01 | -6.932688E-02 |  | -1.227249E-01 |  |  | -6.082497E-02 |  |  |  |  |  |  |  |  |
| coef. 10 | 1.0E-08 | -2.135312E-01 |  | -2.751472E-03 |  |  | -1.951660E-01 |  |  |  |  |  |  |  |  |

Fig.1. Spectrum of a 4th order AR model, estimated by the standard FBLS algorithm applied to the test signal.

Fig.2. Weight vector computed by the proposed algorithm applied to the test signal, in the case of order equal to 4 and K=0.5.

Fig.3. Spectrum of a 4th order AR model, estimated by the proposed algorithm applied to the test signal.

Fig.4. Spectrum of a 10th order AR model, estimated by the proposed algorithm applied to the test signal (10), corrupted by the spikes (11) and by an additive white noise with SNWR of 10 dB.