Adaptive Filtering of Signals from Noise

Henry D. I. Abarbanel

January 1985
JSR-83-210
Revision I

Approved for public release, distribution unlimited.

JASON
The MITRE Corporation
1820 Dolley Madison Boulevard
McLean, Virginia 22102
## Adaptive Filtering of Signals from Noise

This is a primer on extracting signals from noise. Two situations are considered: (1) a reference signal is coming directly from the noise source itself (e.g., "own ship" noise generated by the machinery on a ship towing a sonar array); (2) a signal which has a correlation function in time, which falls off slowly in time compared to the correlation function of the noise.

A new idea in adaptive correlation, called instantaneous correlation, is introduced.
This is a primer on extracting signals from noise which has one potentially new idea and otherwise has as its goal a review — very brief review — of a subject with hundreds of papers to introduce it, explore it, develop it, and exploit it. This little primer is not a substitute for them. (Some are listed at the end of this note, i.e., the ones we found useful.)

The general problem is that a signal detected at times $kAt$, call it $D(k)$, is composed of the signal $s(k)$ one wishes to observe and unwanted additive noise, $n(k)$, one would like to remove. To accomplish this we want to feed into $D(k)$ a signal $y(k)$ with characteristics which isolate $s(k)$ from $n(k)$ according to some performance criterion. A typical performance criterion is that

$$P = \sum_{k=1}^{N} (D(k) - y(k))^2$$

be a minimum. $N$ is the total number of measurements of $D$.

Two cases arise naturally, one in which $y(k)$ is driven toward $n(k)$ as $P$ is minimized (this is called case R) and a second in which $y(k)$ is driven to $s(k)$ (case B):

(1) Case R. In case R we have a reference signal coming directly from the noise source itself. An important example of this
would be "own ship" noise generated by machinery on a surface ship or submarine which propagated from the turbines or whatever through elastic or waterborne paths to contaminate a signal, \( s(k) \), which an onboard or towed sonar array is trying to hear. In case R we choose for \( y(k) \) a linear combination of our knowledge of the reference noise \( r(k-1), r(k-2), \ldots, r(k-L) \)

\[
y(k) = \sum_{a=1}^{L} W(a) r(k-a)
\]

(2)

where \( r(k) \) is some linear combination of the noise samples themselves. In case R we want to choose the weights \( W(a) \) so \( P \) is minimum, i.e.

\[
\frac{\partial P}{\partial W(a)} = -2 \sum_{k=1}^{N} (D(k) - y(k)) r(k - a) = 0; a=1,\ldots,L
\]

(3)

If we succeed in choosing the \( W \)'s to minimize \( P \) in case R, then \( y \) has become our noise estimator chosen to minimize the correlation between \( D \) and \( y \). The part of \( D \) correlated with \( y \) is the noise (since we may assume the signal \( s(k) \) and the noise \( n(k) \) are uncorrelated), so \( D(k) - y(k) \) is our estimate of the signal alone.

---

(2) **Case B.** In this case we have a signal which has a correlation function in time
which falls off very slowly in time compared to the correlation function of the noise

\[ C_s(\tau) = \frac{1}{N} \sum_{k=1}^{N} s(k + \tau) s(k) \] (4)

\[ C_n(\tau) = \frac{1}{N} \sum_{k=1}^{N} n(k + \tau) n(k) . \] (5)

If the envelope of \( C_s \) or \( C_n \) behave as \( e^{-\ell/\Delta_s} \) or \( e^{-\ell/\Delta_n} \) respectively, case B is \( \Delta_s \gg \Delta_n \). Stated in other words, we have broadband (in frequency) noise.

In case B we want to choose as our reference signal, \( r \), the detected signal lagged by a time \( \Delta \) which satisfies \( \Delta_n \ll \Delta \ll \Delta_s \). Then \( r(k - \Delta) = s(k - \Delta) + n(k - \Delta) \) is correlated with \( s \) but uncorrelated with \( n \). If we choose the weights \( W(a) \), \( a=1, \ldots, M \) so

\[ y(k) = \sum_{a=1}^{M} W(a) r(k - a - \Delta) \] (6)

and minimize \( P \), we are in effect minimizing any correlation between \( D \) and \( y \). Since the part of \( D \) still correlated, after lag \( \Delta \), with \( y \) is the signal \( s \), it follows that \( y \) itself is our estimate of the signal.

-3-
Case R and B are indicated in Figure A.

The solution to the optimum filtering problem is that set of \( W_o(a), a = 1, \ldots, M \) satisfying (3):

\[
C_{Dr}(a) = \sum_{b=1}^{M} W_o(b) C_D(b,a),
\]

(7)

where

\[
C_{Dr}(a) = \frac{1}{N} \sum_{k=1}^{N} D(k) r(k-a),
\]

(8)

and

\[
C_r(b,a) = \frac{1}{N} \sum_{k=1}^{N} r(k-a) r(k-b)
\]

(9)

are the obvious correlation functions. If the processes involved are stationary in time, then \( C_D \) is a function only of \( |a-b| \).

Clearly at the optimum

\[
W_o(b) = \sum_{a=1}^{M} C_r^{-1}(b,a) C_{Dr}(a),
\]

(10)

which is the classical Wiener filtering solution. To use this solution one needs to accurately compute the correlation functions.
Signal Noise Contamination Measures

\[ r(k-l), r(k-2), \ldots \]

W Feed Back to Correct Weights

a) Reference Signal from Probe on Noise Source
b) Detected Signal D(k) Delayed by \( \Delta \)

R) Case R - Reference Signal is Filtered Noise.
B) Case B - Reference Signal is Time Delay of Detected Signals.

Figure A.
and \( C_r \) for as many lags as filter weights and then to do a \( M \times M \)
matrix inversion followed by a multiplication. All this is costly in computing time. Furthermore, if the nature of the signal or the noise changes during the observation time, the correlation functions remember too much about the past signal—so (10) is not likely to be very useful for time-varying situations; it doesn't learn enough.

The, by now almost traditional, approach is to correct the weights at each time step by adding an amount proportional to the instantaneous slope \( \frac{\partial P}{\partial W(a)} \) in weight space. If we give a time label to the weights so \( W_k(a) \) is the \( a \)th weight at time step \( k \), then the so-called LMS adaptive algorithm chooses \( W_{k+1}(a) \) as

\[
W_{k+1}(a) = W_k(a) + \mu(D(k) - y(k))r(k-a) .
\]  

(11)

The parameter \( \mu \) is up to the user. When the sequence of \( W_k(a) \) converges, one can expect it to find \( W_0(a) \) given by (10). Much less computation is involved.

We'll give an example of the use of (11) in a few paragraphs. Let us now comment that the use of the LMS prescription is numerically tricky. The time (in number of "cycles" of the basic signal) to learn a new signal and choose weights accordingly is proportional to \( \mu^{-1} \); this encourages the user to choose \( \mu \)
large. The convergence properties of the sequence, however, rapidly
go bad as \( \mu \) becomes too large. Unfortunately, this may happen at
very small \( \mu \). Although some of the references give some bounds on
\( \mu \) for convergence, in practice it is somewhat cut and try and
depends in a sensitive way on the scale of the noise. If the noise--
or \( r(k) \)--in (11) is large, \( \mu \) must be correspondingly decreased.

My colleague A. Peterson in JASON indicated to me that the
hardware versions of the LMS algorithm set \( \mu \) quite large to reduce
the learning time and then monitor the size of the \( W(a), a = 1 \)
to check for instabilities. If the \( W(a) \) get outside an acceptable
bound, they are reinitialized.

The second method of adaptive cancellation we suggest is the
one original piece of this note. The information in the correlation
functions \( C_r \) and \( C_{Dr} \) in (8) and (9) is contained in their
correlation lengths. This is the number of time steps in which
\( C_{Dr}(\xi) \), for example, falls by \( 1/e \) of its value at zero lag. So

\[
C_{Dr}(\xi) \sim e^{-\xi/L}
\]  

(12)

and \( L \) is the correlation length.
Our Instantaneous Correlation (IC) algorithm does the following: At time step $k$ compute

$$C_{Dr}^{(k)}(a) = \frac{1}{2L_1 + 1} \sum_{j = k - 2L_1}^{k} D(j) r(j - a), \quad (13)$$

and

$$C_{r}^{(k)}(b,a) = \frac{1}{2L_2 + 1} \sum_{j = k - 2L_2}^{k} r(j - a) r(j - b), \quad (14)$$

where $L_1$ and $L_2$ are correlation lengths to be adjusted.

Next form

$$w^{(k)}(a) = \sum_{b = 1}^{M} C_{r}^{(k)-1}(a,b) C_{Dr}^{(k)}(b), \quad (15)$$

which are the filter weights entering our determination of the signal.

One must learn by experimentation which $L_1$ and $L_2$ are appropriate. Since the correlation length of the noise is much less than that of the signal, we can expect $L_1 = L_2$.

The IC method may require a lot of computing, but it clearly learns very well. As the signal or the noise changes, so will the
Power Spectrum of
Noisy Signal

Gain 3.0
2^6 Points FT
2^6 Points Shown

Figure 10
Power Spectrum of $a_i(t)$

Gain  1.0
$\mu$  0.01
Lag   3
16 Weights

Figure 9
Correlated Time Series
2' Points
16 Weights
\( \mu = 0.01 \)
Lag  5
Gain  1.0

Figure 8
Corrected Time Series
2° Corrected Points
16 Weights
\( \mu = 0.01 \)
Lag \( = 3 \)
Noise Gain \( = 1.0 \)

Figure 7
Corrected Time Series
2' : 2500 - 4nw Points on Orbit
2' Corrected Points
16 Weights
μ  0.01
lag  3
Noise Gain  1.0

Figure 6
$a(t)$

![Graph](image)

Figure 5

g = 3.041
2 Points
Gain = 1.0
Noise Correlation Function

$2^{13}$ Points on Orbit
$2^2$ Points of $C_n(t)$

Figure 4
$g : 3.041$

$2^{12}$ Points on Orbit
$2^{12}$ Points of $C_d(t)$ Shown
Correlation Function

Figure 3
\( g = 3.041 \)

Power Spectrum
Gain = 0.0
2^2 Points FT
2^2 Points Displayed

Figure 2
Figure 1

\[ g = 3.041 \]

\[ 2^\circ \text{ Points} \]

\[ \text{Gain} = 0.0 \]
the adaptive filter through more cycles of the signal to let the filter learn better. We had to choose $\mu$ small and therefore adaptation times long to insure stability of the algorithm. This is the fundamental snag of the LMS technique. One should be aware of it, not detered by it.

(2) We could have taken a number of samples of noisy signals, filtered them, and then averaged the power spectra. This will quite likely improve matters, and is very straightforward we simply didn't have time.

Furthermore and finally, time prevented us from exploring the IC algorithm in any detail. It should be easy to do — we did it running in a few hours, but hadn't enough experience by this writing to "learn" from it.
Next we corrupted the signal with gaussian random noise with zero mean and variance unity. The noisy signal is shown in Figure 5. Using the LMS algorithm we then corrected the noisy signal using 16 weights, an LMS parameter \( \mu = 0.01 \) (chosen by trial and error to be small enough to be stable), and a lag \( \Delta = 3 \). The result of the calculation shown in Figure 6 shows the initial learning and then the corrected signal. In Figure 7 we suppressed the learning period and exhibit only the corrected signal. This is to be directly compared to the noisy signal in Figure 5. Our final time series display, Figure 8, is identical to Figure 7 but with the lag \( \Delta = 5 \) which does not produce a visible improvement.

Next we show three power spectra. Figure 9 is the power spectrum of Figure 7. It should be compared with the spectrum in Figure 2 of the clean signal. A few lines are still visible. In the last two figures (10 and 11) we show the power spectra of the even more noisy signal (signal variance = 3 times larger) first without correction, then LMS corrected with \( \mu = 0.008 \), 16 weights, and a lag of 25.

What conclusions shall we draw from this. Clearly, to the eye, the LMS algorithm has cleaned up the time series by reducing the variance of the noise. Nevertheless, the visible improvement in the power spectra is not much. Several things should be tried: (1) run
might not. Using one of the filtering algorithms (LMS or IC or whatever) guarantees that the appropriate correlation is driven as small as possible, thus providing one with an objective goal which will serve even when box averaging fails. Perhaps the way to look at the weighted average is that it knows a little bit more about the noise than its average. It also knows that it must kill its correlations with the signal and that it must reduce the variance due to the noise which creeps, unwanted, into the estimated signal.

A LITTLE EXPERIMENT

To examine the ease with which one can use these ideas we created a signal from the solution of a set of nonlinear differential equations and then, by hand, added Gaussian random noise to it. Using the LMS algorithm we then tried to extract the signal back out of noise.

The signal is shown in Figure 1. It is periodic but not sinusoidal. Its power spectrum is given in Figure 2. Clearly it has many harmonics. To see that we can apply Case B to this signal we evaluated its autocorrelation function (unnormalized) and show it in Figure 3. On the same horizontal scale we show our noise autocorrelation function in Figure 4. An expanded view of Figure 4 indicates that $\Delta_n = 2$ while from Figure 3 we see $\Delta_s = 10^3$. Clearly Case B applies.
$C^{(k)}$ and the $W^{(k)}$, and within a correlation time of the changed signal. Further the method is intermediate between the LMS algorithm and the full Wiener $W_0$ in amount of computation required. It is stable, while the LMS approach may not be.

A final remark: in case B the signal estimate is

$$y(k) = \sum_{a=1}^{M} W(a) \left[ s(k - \Delta - a) + n(k - \Delta - a) \right]. \quad (16)$$

One might have thought to "smooth" the detected signal by extracting the broadband noise through an averaging procedure over the "recent" past. Namely, one might take as the the estimate for the signal

$$s_e(k) = \frac{1}{K} \sum_{j=1}^{K} \left[ s(k - j) + n(k - j) \right]. \quad (17)$$

This kind of "box" averaging will eliminate high frequency noise and is really a choice of uniform weights $W(a) = 1/K$ in our optimization scheme. The procedure outlined chooses $W(a)$ according to a prescription which filters out noise at different rates at different frequencies to smooth the detected signal in an optimum way over the whole spectrum without a priori bias as to the relative magnitude of the $W(a)$. As given, the prescription (17) might well work - and it

-9-
Power Spectrum of Corrected Signal

$g = 3.041$

2\textsuperscript{nd} Points FT

2\textsuperscript{nd} Points Shown

$\mu = 0.008$

16 Weights

Lag = 25\hspace{1em} Gain = 3.0

Figure 11
DISTRIBUTION LIST

Dr. Mary Atkins
Deputy Director, Science & Tech.
Defense Nuclear Agency
Washington, D.C. 20305

National Security Agency
Attn RS: Dr. N. Addison Ball
Pt. George G. Meade, MD 20755

Professor Peter M. Banks
928 Casa Nueva Place
Stanford, CA 94305

Dr. Charles Buffalano
DARPA
Deputy Director for Research
1400 Wilson Boulevard
Arlington, VA 22209

Dr. Robert Cooper [2]
Director, DARPA
1400 Wilson Boulevard
Arlington, VA 22209

Cameron Station
Alexandria, VA 22314

The Honorable Richard DeLauer
Under Secretary of Defense (R&E)
Office of the Secretary of Defense
The Pentagon, Room 3E1006
Washington, D.C. 20301

Dr. Alvin M. Despain
1192 Grizzly Peak Boulevard
Berkeley, CA 94708

CAPT Craig E. Dorman
Department of the Navy, OP-095T
The Pentagon, Room 5D576
Washington, D.C. 20350

CDR Timothy Dugan [2]
NFOIO Detachment, Suitland
4301 Suitland Road
Washington, D.C. 20390

Dr. Craig Fields
DARPA/DSO
1400 Wilson Boulevard
Arlington, VA 22209

Dr. Stanley M. Flatté
1126 Muirlands Vista Way
La Jolla, CA 92037

Director [2]
National Security Agency
Fort Meade, MD 20755
ATTN: Mr. Richard Foss, A05

Dr. Larry Gershwin
NIO for Strategic Programs
P.O. Box 1925
Washington, D.C. 20505

Dr. S. William Gouse, W300
Vice President and General Manager
The MITRE Corporation
1820 Dolley Madison Blvd.
McLean, VA 22102

Dr. Edward Harper
SSBN, Security Director
OP-021T
The Pentagon, Room 4D534
Washington, D.C. 20350

Mr. R. Evan Hineman
Deputy Director for Science & Technology
P.O. Box 1925
Washington, D.C. 20505
DISTRIBUTION LIST (Cont'd.)

Mr. Ben Hunter [2]
1917 Westmoreland Street
McLean, VA 22101

The MITRE Corporation [25]
1820 Dolley Madison Blvd.
McLean, VA 22102
ATTN: JASON Library, W002

Mr. V. Larry Lynn
Deputy Director, DARPA
1400 Wilson Boulevard
Arlington, VA 22209

Mr. John McMahon
Dep. Dir. Cen. Intelligence
P.O. Box 1925
Washington, D.C. 20505

Dr. Oscar P. Manley
Eng., Math, & Geosciences Division
Office of Basic Energy Research
Department of Energy
Mail Stop G-226
Washington, DC 20545

Dr. Claire E. Max
617 Grizzly Peak Boulevard
Berkeley, CA 94708

Dr. George A. Keyworth
Director
Office of Science & Tech. Policy
Old Executive Office Building
17th & Pennsylvania, N.W.
Washington, D.C. 20500

MAJ GEN Donald L. Lamberson
Assistant Deputy Chief of Staff (RD&A) HQ USAF/RD
Washington, D.C. 20330

Dr. Donald M. LeVine, W385 [3]
The MITRE Corporation
1820 Dolley Madison Blvd.
McLean, VA 22102

Mr. John F. Kaufmann
Dep. Dir. for Program Analysis
Office of Energy Research, ER-31
Room F326
U.S. Department of Energy
Washington, D.C. 20545

Dr. Marvin Moss
Technical Director
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Dr. Julian Nall [2]
P.O. Box 1925
Washington, D.C. 20505

Dr. Oscar P. Manley
Eng., Math, & Geosciences Division
Office of Basic Energy Research
Department of Energy
Mail Stop G-226
Washington, DC 20545

Dr. Claire E. Max
617 Grizzly Peak Boulevard
Berkeley, CA 94708

Dr. George A. Keyworth
Director
Office of Science & Tech. Policy
Old Executive Office Building
17th & Pennsylvania, N.W.
Washington, D.C. 20500

MAJ GEN Donald L. Lamberson
Assistant Deputy Chief of Staff (RD&A) HQ USAF/RD
Washington, D.C. 20330

Dr. Donald M. LeVine, W385 [3]
The MITRE Corporation
1820 Dolley Madison Blvd.
McLean, VA 22102
DISTRIBUTION LIST (Cont'd.)

Prof. William A. Nierenberg  
Scripps Institution of Oceanography  
University of California, S.D.  
La Jolla, CA 92093

Dr. Fred Saalfeld  
Director of Research Programs  
Office of Naval Research  
800 N. Quincy  
Arlington, VA 22217

Mr. C. Wayne Peale  
Office of Research and Development  
P.O. Box 1925  
Washington, DC 20505

Dr. Phil Selwyn  
Technical Director  
Office of Naval Technology  
800 N. Quincy Street  
Arlington, VA 22217

Dr. Allen M. Peterson  
SRI International, (EL-154)  
333 Ravenswood Avenue  
Menlo Park, CA 94025

Dr. Eugene Sevin [2]  
Defense Nuclear Agency  
6801 Telegraph Road  
Room 244  
Alexandria, VA 22310

The MITRE Corporation  
Records Resources  
Mail Stop W971  
McLean, VA 22102

Dr. Joel A. Snow [2]  
Senior Technical Advisor  
Office of Energy Research  
U.S. DOE, M.S. E084  
Washington, D.C. 20585

Mr. Alan J. Roberts  
Vice President & General Manager  
Washington C&I Operations  
The MITRE Corporation  
1820 Dolley Madison Boulevard  
McLean, VA 22102

COMD William O. Studeman  
Director, Long Range Planning Group  
Office of C&O (OP-09X)  
Navy Department  
Washington, DC 20350

Mr. Alan J. Roberts  
Vice President & General Manager  
Washington C&I Operations  
The MITRE Corporation  
1820 Dolley Madison Boulevard  
McLean, VA 22102

COMD William O. Studeman  
Director, Long Range Planning Group  
Office of C&O (OP-09X)  
Navy Department  
Washington, DC 20350

Los Alamos Scientific Laboratory  
ATTN: C. Paul Robinson  
P.O. Box 1000  
Los Alamos, NM 87545

Mr. Richard Ross [2]  
P.O. Box 1925  
Washington, D.C. 20505

Mr. Alexander J. Tachmindji  
Senior Vice President & General Manager  
The MITRE Corporation  
P.O. Box 208  
Bedford, MA 01730

Los Alamos Scientific Laboratory  
ATTN: C. Paul Robinson  
P.O. Box 1000  
Los Alamos, NM 87545

Mr. Richard Ross [2]  
P.O. Box 1925  
Washington, D.C. 20505

Mr. Alexander J. Tachmindji  
Senior Vice President & General Manager  
The MITRE Corporation  
P.O. Box 208  
Bedford, MA 01730

Mr. Richard Ross [2]  
P.O. Box 1925  
Washington, D.C. 20505

Mr. Alexander J. Tachmindji  
Senior Vice President & General Manager  
The MITRE Corporation  
P.O. Box 208  
Bedford, MA 01730

Dr. Oscar S. Rothaus  
106 Devon Road  
Ithaca, NY 14850

Dr. Vigdor Teplitz  
ACDA  
320 21st Street, N.W.  
Room 4484  
Washington, D.C. 20451
DISTRIBUTION LIST (Concl'd.)

Dr. Al Trivelpiece
Director, Office of Energy
Research, U.S. DOE
M.S. 6E084
Washington, D.C. 20585

Dr. John F. Vesecky
Center for Radar Astronomy
233 Durand Building
Stanford University
Stanford, CA 94305

Mr. James P. Wade, Jr.
Prin. Dep. Under Secretary of
Defense for R&E
The Pentagon, Room 3E1014
Washington, D.C. 20301

Dr. Kenneth M. Watson
2191 Caminito Circulo Norte
La Jolla, CA 92037

Mr. Leo Young
OUSDRE (R&AT)
The Pentagon, Room 3D1067
Washington, D.C. 20301
END

FILMED

4–85

DTIC