MIN - MAX DETECTION OF A WEAK SIGNAL IN STATIONARY MARKOV NOISE

George V. Kurkjian and John R. Thomas
Department of Electrical Engineering and Computer Science
Princeton University
Princeton, NJ 08544

ABSTRACT

The detection of a constant weak signal in stationary Markov noise is considered. The observation sequence is passed through a 2NL, and the sum of the outputs is compared to a fixed threshold. Using the efficiency of message, the optimally that leads to a minimum performance is obtained.

1. INTRODUCTION

Robert's ideas of robustness [1-2] are applied in [3,4, and 5] in order to obtain structures that are asymptotically robust signal detectors. It is assumed that the observation sequence is stationary and that the commonly known density belongs to an \( \mathcal{F} \) - contamination class. In [3,4] the common density is also symmetric, and the detection structure is a sum of a memory nonlinear transformations of the observation sequence, compared to a fixed threshold. Using the efficiency as the performance criterion, the robust structure turns out to be the non-linear nominal correlator detector when the signal is assumed small and the number of observations large. In [5] the assumption of symmetry is dropped instead symmetry within an interval around the origin is assumed. The same detection structure is used as in the other cases but here the threshold \( \eta \) must be fixed if desired level must be attained.

All of the above approaches assume that the observations are independent. In [5] the same detection problem is considered but the observation come from a mixing average type process and are weakly dependent. In our paper we are dealing with dependent observations too but we are using a different model of dependency. It is assumed that the observations form a stationary Markov sequence. As detection structure we use sums of memoryless nonlinear transformations and our goal is to optimize this structure. Obviously this structure is not the most optimal but it will give us an idea on how much the independence assumption structure changes under dependency and also if the performance changes drastically.

2. PRELIMINARIES

Let \( \{X_n\} \) be a strictly stationary Markov noise sequence. Since the statistics of such a sequence are well defined if we specify the generate distribution of two adjacent observations, we will assume that this favorite distribution belongs to a class \( \mathcal{F}_{\eta, \lambda} \) which is defined as follows

\[
 f(x, y) = f(x|y)(1 + \lambda|x, y|) \tag{15}
\]
\[ f(x) = \int f(x,y) \, dy = 0 \]

\[ f(x) = \left| \frac{1}{1-x} \right| |x| \, \text{sign}(x) \quad 0 < x < 1 \]

\[ f(x) \neq 0; \quad x \neq \eta < 1 \]

Condition (1) defines a common representation of a bivariate distribution. Condition (2) says that the function \( f(x) \) is the marginal density. With Condition (3) we define an \( \eta \)-contamination model for the marginal. The density \( g(x) \) is assumed to be known (symmetric, strongly smooth and not equal to zero). The density \( A(x) \) is unknown but symmetric and non-negative. Thus, \( f(x) \) has finite Fisher information and the set \( \{ f(x) \neq 0 \} \) has a \( \eta \)-measure zero. Condition (4) limits our dependency model: notice that whereas \( m=0 \) we are back to the i.i.d. case. The two constants \( x \) and \( m \) are assumed to be known. As we will see shortly, Condition (4) is not only important because it makes \( f(x) \) non-negative, but also because it makes the sequence \( \{ \eta_n \} \) a \( \eta \)-mixing sequence, a sufficient condition to guarantee asymptotic normality.

We now consider the detection of a constant signal. In particular, we would like to detect between the two hypotheses

\[ H_0: \quad \eta \leq \eta_0 \]

\[ H_1: \quad \eta > \eta_0 \]

where \( \{ \eta_n \} \) is the observation sequence and \( x \) a signal that tends to zero. Adopting the terminology from (5), our detection structure will use the nonlinear-quantizer (NQ) detector, which is of the form

\[ \tau_n(x) = \sum_{i=1}^n g_i(x) \]

For the performance criteria we will use the efficiency, in order for the efficiency to exist and to be a well criterion we have to impose restrictions on the non-centrality \( \eta(x) \), which will determine the allowable class \( \psi \) of noncentralities \( \eta(x) \). We assume that \( \eta(x) \) is an odd symmetric, zero mean, and second order function such that

\[ \eta(x) = E_\psi \eta(x), \quad \eta(x) = 0 \quad \sum_{i=1}^n g_i(x) \eta(x) \eta_i(x) > 0 \]

We also assume that \( \eta(x) \) satisfies conditions that are sufficient for the validity of the Pitman - Norther theorem. Such conditions are given in [7]. The assumption that \( \eta(x) \) is odd symmetric is reasonable since the marginal densities are assumed even symmetric functions.

Proposition 1. Let \( \eta(x) \) be a strictly stationary Markov sequence with biivariate density \( f(x,y) \) and \( g(y) \), also let \( \eta(x) \) \( \in \mathbb{R}^n \) and \( \eta(0) = 0 \). Then \( \eta(x) \) is a \( \eta \)-mixing sequence and \( g(y) \) defined in (5) is absolutely summable.

Proof. Because the sequence is stationary Markov, the biivariate density between \( y \) and \( N_{0} \) is given recursively by

774
\[ f_1(x, y) = f(x) f(y) \] where \( f(x) \) and \( f(y) \) are the marginal densities. Using the definition of the product measure and the equation for joint density, we have

\[ f_1(x, y) = \int_{-\infty}^{\infty} f_2(x, y) \, dy \]

Now, we can use the joint density function and the equation for marginal densities to show that the density defined by \( f_1 \) has the following form:

\[ f_1(x) = f(x) \int_{-\infty}^{\infty} f(y) \, dy \]

Now, let \( \lambda_2(x) = f(x) \) and \( \lambda_1(y) = \int_{-\infty}^{\infty} f(y) \, dy \).

We also have that \( \lambda_3(x, y) = \lambda_1(x) \lambda_2(y) \).

Using \( A, B \) and noting that the first and the last terms in (1) are the marginal densities of \( (X_1, X_2, X_3) \) and \( (X_1, X_2, X_3) \), respectively, we have

\[ P(A, B, C) = \int_{\Omega} f(x, y) \, dx \, dy \]

Now, let \( \lambda_2(x, y) = f(x, y) \) and \( \lambda_1(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \).

Using (6.10) and noting that the first and the last terms in (1) are the marginal densities of \( (X_1, X_2, X_3) \) and \( (X_1, X_2, X_3) \), respectively, we have

\[ P(A, B, C) = \int_{\Omega} f(x, y) \, dx \, dy \]

So the sequence is symmetrically \( x \)-mixing with \( \alpha = \mu \). Clearly since \( m < 1 \), we have that \( \alpha = 0 \) for \( x = \infty \) and also that \( \sum_{m}^{\infty} \alpha_i = \infty \). And this takes care of the first part of Proposition 1.

To prove now that \( \psi(|V|) \) is absolutely summable, we use Equations (6.10) and we have

\[ \psi(|V|) = \frac{1}{2} \left[ \int_{-\infty}^{\infty} \psi(x) \, dx + \int_{-\infty}^{\infty} \psi(x) \, dx \right] \]

\[ \psi(|V|) = \frac{1}{2} \left[ \int_{-\infty}^{\infty} \psi(x) \, dx + \int_{-\infty}^{\infty} \psi(x) \, dx \right] \]

715
and finally, using Schurts inequality and that \( \psi(x) \) is second order we have
\[
\sigma_0(\psi) = \min_{\lambda \geq 0} \int \psi(x) f(x) dx 
\]

and this proves the second part of Proposition 1.

**Proposition 2.** Assume that \( \{\psi_n\} \) is a noise sequence defined as above and that \( \int \psi(x) N(x) dx = 0 \) and \( \int \psi(x) (x) dx = \psi(x) \), where \( \psi(x) \) is a measurable function. Then
\[
\sum_{n=1}^{\infty} \psi_n(x) \rightarrow (\psi(x)) \quad (6)
\]

**Proof:** It is a direct consequence of (6, Theorem 211).

### III. MIN-NOSE DETECTION

As our performance criterion we will be using the efficiency. Under the preceding assumptions it takes the following form

\[
eff(\psi(x) f(x, y)) = \frac{\int f(x, y) f(x, y) dx}{\int f(x, y) dx} \quad (7)
\]

The problem we want to solve is the following. We would like to find a nonlinear \( \psi(x) \) and a density \( f(x, y) \) on \( X_n \) such that

\[
\max_{\psi(x)} \min_{f(x, y)} \int f(x, y) f(x, y) dx = \int f(x, y) f(x, y) dx \quad (7)
\]

The first step is to try for a given \( \psi(x) \) to minimize (7) over the density \( f(x, y) \). Notice that the density depends on two functions, the marginal \( f(x) \) and the function \( f(x, y) \). Since it makes no difference, we minimize over \( f(x, y) \) fixed. From (3) we have that

\[
\sigma_0(\psi) = \int f(x, y) f(x, y) dx + \min_{\lambda \geq 0} \int f(x, y) f(x, y) dx 
\]

Equality is achieved when \( f(x, y) \) is given by

\[
\lambda(x, y) = m_0 \lambda(x) m_0(y) \quad (9)
\]

where \( m_0(x) \) is an odd symmetric function for \( x > 0 \) is given by

\[
\lambda(x) = \begin{cases} 
-1 & \text{if } \psi(x) < 0 \\
0 & \text{if } \psi(x) = 0
\end{cases}
\]

(30)
Notice that the definition leads to a legitimate bivariate density since all conditions \( (\gamma \geq x) \), \( \delta \) are satisfied. Also, applying (9), we have
\[
\psi(x, y) = \int \psi(x, y) \, dx = m \eta_3(\psi(x, y)) \quad \gamma \geq x
\]
Thus if \( \alpha(x, y) \) is given from (19), the efficacy becomes only a function of \( \psi(x, y) \) and \( f(x) \) and takes the form
\[
\text{eff}^* \left( \psi(x), f(x) \right) = \frac{\int \psi(x) f'(x) \, dx}{\int \psi(x) f(x) \, dx} + \frac{\int \psi(x) f(x) \, dx}{\int \psi(x) f'(x) \, dx}
\]
In order to continue we have to maximize (20) over \( f(x) \) and then maximize the result over \( \psi(x) \). This is a min-max problem in itself with a criterion function the \( \text{eff}^* \) given by (20). It turns out that this new problem has a saddle point, in other words, there exist a pair \( \psi(x) \) and \( f(x) \) that satisfies the following saddle point relation
\[
\text{eff}^\prime \left( \psi(x), f(x) \right) = \text{eff}^\prime \left( \psi(x), f(x) \right) = \text{eff}^\prime \left( \psi(x), f(x) \right)
\]
for any allowable \( \psi(x) \) and \( f(x) \). As we know, any pair that satisfies (20) also satisfies the min-max relation is (17), so the pair \( \psi(x), f(x) \) is the one we are seeking. The left side inequality in (20) says that \( \psi(x) \) is the optimum linearity for \( f(x) \) and is the one that will maximize the \( \text{eff}^* \) over the class. The form of the optimum linearity is given by the following theorem.

**Theorem 1.** Let \( f(x) \) be a symmetric density function with finite Fisher's information and such that the set \( \{ \psi(x) = \int f(x) \, dx \} \) has \( f \)-measure zero. Then the nonlinear \( \psi(x) \) that maximizes (21) is given by
\[
\psi(x) = \psi(x) - m \psi_0(x)
\]
where \( \psi(x) \) is defined as
\[
\psi(x) = \begin{cases} \frac{1}{\mu} \psi_0(x) & \text{when} \quad \mu \psi_0(x) < 0 \\ 0 & \text{when} \quad \mu \psi_0(x) = 0 \\ -1 & \text{when} \quad \mu \psi_0(x) > 0 \end{cases}
\]
and \( \mu \) is a constant that satisfies
\[
N(\mu) = \mu + 2m [ \int \psi(x) \psi_0(x) \, dx ] + 0
\]
Theorem 2. The density \( f_s(x) \) that gives the solution to the saddle point problem is the following:

\[
f_s(x) = \begin{cases} 
1 - \varphi(x), & \text{for } x < 1, \\
\varphi(x), & \text{for } x > 1 
\end{cases}
\]

where \( \varphi(x) \geq 0 \) and such that \( f_s(x) \) has total mass equal to unity.

Proof. This density is nothing else than the one defined by Huber in [1, 2]. It belongs to the \( \varphi \) contamination family with a legitimate density \( \lambda(x) \) (see [1]). To find \( \psi_s(x) \), we apply Theorem 1 and if we denote by \( \psi_s(x) \) the locally optimum nonlinearity of \( f_s(x) \), we get:

\[
\psi_s(x) = \begin{cases} 
0, & \text{for } x = 0, \\
\psi_s(x) - \psi_s(x), & \text{for } 0 < x < 1, \\
\psi_s(x) - \psi_s(x), & \text{for } x > x
\end{cases}
\]

For \( x = 0 \) we recall that \( \psi_s(x) \) is an odd function. The constant \( x_0 \) is defined as

\[
\psi_s(x_0) = \mu
\]

and in order for (28) to be valid it has to satisfy \( 0 < x_0 < x \). In the Appendix we show that such an \( x_0 \) always exists.

Let us now for convenience define:

\[
y = \psi_s(x) - \psi_s(x)
\]

Since \( \psi_s(x) \) is non-decreasing we will have that \( \psi_s(x) < M \). Up to now we have shown that \( \psi_s(x) \) and \( f_s(x) \) satisfy the left inequality in (25). To prove the right one is straightforward. If we define as \( m(f) \) and \( d(f) \) the numerator and denominator of the \( \text{off} \) then:

\[
m(f) = \int_{-\infty}^{x_0} f(x) \varphi(x) dx = \int_{-\infty}^{x_0} f(x) \phi(x) dx + \int_{x_0}^{\infty} f(x) \phi(x) dx
\]

and

\[
d(f) = \int_{-\infty}^{x_0} f(x) \varphi(x) dx + \int_{x_0}^{\infty} f(x) \varphi(x) dx
\]

\[
\Xi_1
\]
\[
\begin{align*}
&= \frac{-1}{1 - |x^-|^2} \int_{|x|^2} \psi(x) \|\psi(x)\| \, dx^2 + \frac{1}{1 - |x^-|^2} \int_{|x|^2} \left( \frac{1}{2} |x|^2 - \frac{|x|^4}{1 - |x|^2} \right) \psi(x) \, dx
\end{align*}
\]

From (31) and (32) we see that \( f_\epsilon(x) \) simultaneously maximizes the numerator and minimizes the denominator of \( \epsilon(f') \), so that the pair \((\psi_\epsilon(x), f_\epsilon(x))\) satisfies also the right inequality in (23). The complete case proof

Summarizing the results, the solution to our min-max problem is the following; the density \( f_\epsilon(x) \) is given by

\[
f_\epsilon(x) = f_\epsilon(x|f_\epsilon(x)) \quad \text{if} \quad x, m_\epsilon(x), m_\epsilon(y) \quad \in \quad (35)
\]

where \( f_\epsilon(x) \) is defined in (27). The nonlinearity \( \psi_\epsilon(x) \) is defined in (26) and these are all of the things we were looking for.

IV. NUMERICAL EXAMPLE.

Let the density \( p(x) \) be the \( N(0,1) \) normal. Then the \( \psi_\epsilon(x) \) is

\[
\psi_\epsilon(x) = \begin{cases} 
0 & \text{for } |x| \leq \epsilon \\
\frac{x - x_0}{\epsilon m_\epsilon(x)} & \text{for } x_0 < x < x_1 \\
\frac{x - x_1}{\epsilon m_\epsilon(x)} & \text{for } x > x_1
\end{cases}
\]

(36)

For the density \( f_\epsilon(x) \), we have the usual definition, only here the function \( m_\epsilon(x) \) can be equal to \( \omega_\epsilon(x) \). In the following we give the table. Table I contains the values of \( \psi_\epsilon(x) \) for different values of \( \epsilon \) and \( x \). Table II is the ASS of \( \psi_\epsilon(x) \) and \( f_\epsilon(x) \) when under the density \( f_\epsilon(x) \) values for \( x \) are not given. These \( \psi_\epsilon(x) \) depends only on \( x \); these values are the same with those given in (36) under the name \( \psi_\epsilon(x) \).

<table>
<thead>
<tr>
<th>Table I</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x )</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>0.00</td>
</tr>
<tr>
<td>0.10</td>
</tr>
<tr>
<td>0.20</td>
</tr>
<tr>
<td>0.30</td>
</tr>
<tr>
<td>0.40</td>
</tr>
<tr>
<td>0.50</td>
</tr>
<tr>
<td>0.60</td>
</tr>
<tr>
<td>0.70</td>
</tr>
<tr>
<td>0.80</td>
</tr>
<tr>
<td>0.90</td>
</tr>
<tr>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table II</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x )</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>0.00</td>
</tr>
<tr>
<td>0.10</td>
</tr>
<tr>
<td>0.20</td>
</tr>
<tr>
<td>0.30</td>
</tr>
<tr>
<td>0.40</td>
</tr>
<tr>
<td>0.50</td>
</tr>
<tr>
<td>0.60</td>
</tr>
<tr>
<td>0.70</td>
</tr>
<tr>
<td>0.80</td>
</tr>
<tr>
<td>0.90</td>
</tr>
<tr>
<td>1.00</td>
</tr>
</tbody>
</table>
APPENDIX

Proof of Theorem 1. Notice that the value of the $\mathcal{S}_{\text{aff}}$ does not change if we multiply $\psi(x)$ by a constant. From one of the conditions for the validity of the Poincaré-Bendixson theorem (Condition a) in [17] we conclude that

$$\int \psi(x)f(x)dx < 0.$$ 

Thus maximizing Expression (20) is equivalent to maximizing

$$H(\psi) = \psi_0 f(x)dx - \rho \left\{ \int \psi_0^2 f(x)dx + \frac{\mathcal{S}_m}{\mathcal{S}_n} \int \psi_0 f(x)dx \right\}$$

where $\rho$ is a Lagrange multiplier. We will show that (30) is maximized by

$$\psi(x) = \frac{1}{\mathcal{S}_m} \left[ f(x) - \mu_0 x \right]$$

(37)

where $\mu_0$ and $\mu_0 x$ are defined in Equations (25, 26). Let now $\psi(x)$ be any other nonlinearity from the class $\psi$. We define the following variation.

$$J(\gamma) = \int \left[ f(x) - \mu_0 x \right] \psi(x)dx - \rho \left\{ \int \left[ f(x) - \mu_0 x \right] \psi_0 f(x)dx \right\} + \frac{\mathcal{S}_m}{\mathcal{S}_n} \left\{ \int \left[ f(x) - \mu_0 x \right] \psi_0 f(x)dx \right\}$$

(38)

where $\gamma(0, 1]$. Notice that $J(0) = H(\psi_0(x))$ and $J(1) = H(\psi_0(x))$. By manipulating (38) we can rewrite it as

$$J(\gamma) = \frac{1}{\mathcal{S}_m} \int f(x)\left[ \psi_0 f(x)dx - \mu_0 \psi(x)dx \right] + \frac{\mathcal{S}_m}{\mathcal{S}_n} \int \psi_0 f(x)dx \psi_0 f(x)dx \psi_0 (x)dx$$

728
\[\begin{align*}
\gamma \sum_{i=0}^{m} \left( \int_{\chi_i(z_i)} f(x) \, dx \right) & \left( \int_{\chi_i(z_i)} - \phi(z_i) \left[ f(x) \right] \, dx \right) \\
- \int_{\chi_i(z_i)} \phi(z_i) \left[ f(x) \right] \, dx & \left( \int_{\chi_i(z_i)} - \phi(z_i) \left[ f(x) \right] \, dx \right)
\end{align*}\]

(31)

It is trivial to show that \(J(0) = 0\).

From the way that \(\phi(x)\) is defined in (20) we can see that

\[\phi(z) = \phi(z_0) \quad \text{for all } z \in \chi_i(z_i) \quad \text{(40)}\]

\[\left| \phi(z) \right| < 1 \quad \text{(41)}\]

By multiplying (39) by \(\phi(z_i) f(x)\) and integrating we can show, using (38) and (40) that

\[\int_{\chi_i(z_i)} \phi(z_i) f(x) \, dx = \frac{\mu_i}{\gamma_i} \quad \text{(42)}\]

If we substitute (42) in the first term of the difference \(J(y) - J(0)\) and also use (37), we get the first term of the second term using (40) becomes

\[\left( \int_{\chi_i(z_i)} f(x) \, dx \right) \left( \int_{\chi_i(z_i)} - \phi(z_i) f(x) \, dx \right) \quad \text{(13)}\]

and because of (13), the quantity in the bracket is non-negative. Thus for \(\mu > 0\), the above expression becomes non-negative. The second term for \(\mu > 0\) is clearly non-negative too. And we have that the difference \(J(y) - J(0)\) is non-negative and in particular \(J(y) \leq J(0)\). If we also define \(\mu \leq \xi\), Equation (17) becomes the same as (36).

For the existence of \(\mu \) that satisfies (38) notice the following: since by assumption \(\phi_0(x) = 0\), except on sets of \(f \)-measure zero we have that as \(\mu \to 0\) then \(\mu \phi_0(x) \to 0\) and thus \(\phi(x) = \text{sgn}(\phi_0(x)) \to \text{sgn}(f(x))\). Thus from (28) we get

\[S(0) = - \sum_{i=0}^{m} \int_{\chi_i(z_i)} f(x) \, dx = 0 \quad \text{(43)}\]

Using Schwartz's inequality it is easy to show that the integral in (44) is bounded by Grosky's expansion. As \(\mu \to 0\), the second term in \(S(0)\) remains bounded so that \(J(0) \to 0\). By continuity, there exist a \(\mu \) that satisfies

\[S(0) = 0 \quad \text{(44)}\]
$S(\mu) = 0$. Also (by the proof of Theorem 1).

Suppose $n = \infty$. To prove that there exists an $\alpha$ that satisfies $S(\mu) = 0$, for the density defined in (29) it is enough to show that $F(Z(\mu)) = 0$. Notice that when $\mu = \mu_0, \ldots, \mu_n$ then

$$S(\mu_0, \ldots, \mu_n) = \frac{\hat{\nu} \cdot \nu}{(1.5\hat{\nu} + 2\nu)^2} = 0$$

On applying this to the expression for $S(x)$ and manipulating we get

$$S(\mu, \ldots, \mu_n) = \frac{\nu_1}{(1.5\nu_1 + 2\nu)^2} = 0$$

whereby $F$ we denote Fisher's information. Thus the solution to $S(\mu) = 0$ must be seen that $\nu_1 \equiv \nu$ and since $\nu_1(x)$ is a nondecreasing function we get $\alpha = \nu$. And hence complete the proof.

ACKNOWLEDGMENT

This research was supported by the U.S. Army under Grant DAAA 29-05-K-0020, and by the National Science Foundation Grant from EECS 78-09495.

SUGGESTIONS


