MATCHED SUBSPACE DETECTORS FOR STOCHASTIC SIGNALS

Todd McWhorter*
Mission Research Corporation
3665 JFK Parkway
Building 1, Suite 206
Fort Collins, CO 80525
Phone: 970-282-4400 x25
Fax: 970-282-9444
Email: mcwhorter@aster.com

Louis L. Scharf†
Department of Electrical and Computer Engineering
Colorado State University
Ft. Collins, CO 80523
Phone: 970-491-2979
Fax: 970-491-2249
Email: scharf@engr.colostate.edu

ABSTRACT

Our aim in this paper is to extend the matched subspace detectors (MSDs) of [1–3] to the detection of stochastic signals. In [1–3] the signal to be detected was assumed to be placed deterministically at an unknown location in a known signal subspace. The basis for the subspace was irrelevant. In this paper the signal is assumed to be placed randomly at an unknown location in a known subspace. If nothing is known a priori about the second-order moments of the placement, then the generalized likelihood ratio test (GLRT) for a stochastic signal turns out to be identical to the GLRT for a deterministic signal. Consequently, the MSDs are more general than originally thought, applying to the detection of a signal whose mean value or covariance matrix is modulated by a subspace signal. Moreover, the invariance sets for stochastic MSDs are identical to those of the corresponding deterministic MSD. The results of this paper extend the theory of MSDs to radar and sonar problems where random target effects may be modeled, and to data communication problems where symbols are coded by subspaces, rather than coordinates of subspaces.

1. INTRODUCTION

Matched subspace detectors generalize matched filter detectors by allowing the signal to lie in a multi-dimensional subspace. Typically the subspace models the distribution of signal modes over a band of frequencies and/or wavenumbers, as in broadband array processing for the detection or decoding of spatially distributed sources. Prior work [1, 2, 4] on MSDs has proceeded under the assumption that the subspace signal is placed deterministically at an unknown location in a known subspace. That is, the subspace signal modulates the mean value of the noisy measurement. In this paper we ask what happens when the signal is placed randomly in a known subspace. That is, what happens to detector structures when the subspace signal modulates the covariance matrix of the noisy measurement? This is the kind of question that typically arises in radar and sonar when targets are randomly fluctuating. It can also arise in non-coherent data communication, when the symbol is coded with a subspace, rather than with a coordinate in a subspace [5].

Our most important finding is this: regardless of the distribution of the signal-coordinate vector, the optimal detector in the known noise case is a matched subspace detector minus an orientation-dependent term. In the case when the signal-coordinate vector is Gaussian, then the generalized likelihood ratio test (GLRT) for the detection of a random subspace signal is identical to the GLRT for the detec-
tion of a deterministic subspace signal. In this case, this conclusion holds whether or not the noise level in the measurement is known. That is, the MSD and the CFAR MSD of \([1, 2, 4]\), are GLRT for deterministic subspace signals \(\text{and} \) stochastic subspace signals. Moreover, the invariance sets for these detectors are the cylinders and cones of \([1, 2, 4]\), regardless of whether the underlying problem is deterministic or stochastic. These findings extend significantly the range of applications for which MSDs are optimum in the sense of GLRT.

2. MEASUREMENT MODEL

Consider the signal-plus-noise model

\[ y = x + n; \quad x = H\theta, \]

where \(H \in \mathbb{C}^{N \times p}\) is a basis for the \(p\)-dimensional subspace \((H)\) and \(\theta \in \mathbb{C}^p\) is, for now, a general complex-valued random vector with density function \(f_\theta(\theta)\). The signal-coordinate vector randomly determines the signal \(x\) in the subspace \((H)\). The noise \(n \in \mathbb{C}^N\) is distributed as \(n \sim CN_n[0, \sigma^2R_{nn}]\), with mean \(0\) and covariance \(\sigma^2R_{nn}\). It is understood throughout this paper that all complex Gaussian random vectors are proper, meaning \(E(nn^T) = 0\).

If the measurement \(y\) includes subspace interference of the form \(SS\phi\), then the GLRT methodology may be applied by making various assumptions about what is known in the covariance model \(SR_{\phi\phi}S\). This approach complicates the problem considerably. As an alternative, we may force any detector we derive to be invariant to the statistics of \(\phi\). That is we require the detector to be a function of the pre-processed data \((I - P_S)y\). In this case, all of the results in the sequel go through, with \(y\) replaced by \((I - P_S)y\) and \(H\) by \((I - P_S)H\).

We assume throughout that the covariance matrix \(R_{nn}\) is known and full rank, meaning it may be factored as \(F^HR_{nn}F = I\). The vector \(z = F^Hy\) can be used in the detection system assuming that the additive noise is white. This being the case, we shall proceed as if \(y\) were conditioned distributed as \(y \sim CN_y[H\theta, \sigma^2I]\), with \(H\) being a whitened version of the original subspace.

The problem is to design a detector to distinguish between the two hypotheses \(H_1: \theta \neq 0\) and \(H_0: \theta = 0\). That is we would like a statistic, that when compared to a threshold, allows us to decide between the case when “signal” is present and when it is not.

3. GLRT FOR GENERAL SIGNAL AMPLITUDES

In this section we initially do not assume a particular statistical description for the coordinate vector \(\theta\). The likelihood function for the measured data when the signal is present is denoted by \(l_1(y; \sigma^2, \beta)\). The parameters \(\sigma^2\) and \(\beta\) represent the parameters that characterize the various density functions. These parameters will be clarified shortly. Our program will be to find the maximum likelihood estimates of the unknown parameters, and then insert these estimates into the likelihoods to obtain compressed likelihood functions under each hypothesis. The ratio of these likelihoods is then used as a GLRT for testing which model is more likely.

We begin by partitioning the measured data into two statistically independent vectors. Define

\[ z = (H^HH)^{-1/2}H^Hy \]

and

\[ w = A^Hy; \quad A \in \mathbb{C}^{N \times (N-p)} \quad A^H[H A] = [0 I]. \]

Now \(w\) has density function \(w \sim CN(0, \sigma^2I)\) regardless of which hypothesis is in effect. Consequently, if \(\sigma^2\) is known then \(w\) is of no use in the detection problem and we only use \(z\). For now assume that we know \(\sigma^2\). Under hypothesis \(H_1\) the conditional density of \(z\) is

\[ f_1(z | \theta; \sigma^2) = \frac{1}{(\pi\sigma^2)^p} \exp\left(-\frac{1}{\sigma^2}\|z - (H^HH)^{1/2}\theta\|^2\right). \]

Define a “colored” amplitude vector

\[ \phi = (H^HH)^{1/2}\theta. \]

The conditional likelihood for \(z\) under \(H_1\) is then

\[ l_1(z | \phi; \sigma^2) = \frac{1}{(\pi\sigma^2)^p} \exp\left(-\frac{1}{\sigma^2}\|z - \phi\|^2\right). \]

The conditional likelihood under \(H_0\) is

\[ l_0(z | \theta; \sigma^2) = \frac{1}{(\pi\sigma^2)^p} \exp\left(-\frac{1}{\sigma^2}\|z\|^2\right). \]
Since $l_0$ does not depend on $\phi$ we can form a conditional likelihood ratio before solving for any unknown parameters $\beta$. The resulting conditional GLRT statistic is

$$v(z \mid \phi; \sigma^2) = \frac{l_1(z \mid \phi; \sigma^2)}{l_0(z \mid \phi = 0; \sigma^2)}$$

$$= \exp\left(\frac{z^H z}{\sigma^2}\right) \exp\left(\frac{1}{2\sigma^2} \|z - \phi\|^2\right).$$

A likelihood-ratios test (LRT) statistic can be obtained by integrating over the distribution of $\phi$.

$$\tilde{v}(z, \sigma^2, \beta) = \int v(z \mid \phi; \sigma^2) f_\phi(\phi; \beta) d\phi$$

$$= \exp\left(\frac{z^H z}{\sigma^2}\right) \times \int \exp\left(-\frac{1}{\sigma^2} \|z - \phi\|^2\right) f_\phi(\phi; \beta) d\phi.$$

Let’s take take the logarithm of both sides to obtain an equivalent detector

$$s(z, \sigma^2, \beta) = \ln \tilde{v}$$

$$= \frac{z^H z}{\sigma^2} + \ln \left[ \int \exp\left(-\frac{1}{\sigma^2} \|z - \phi\|^2\right) f_\phi(\phi; \beta) d\phi \right]$$

$$= \frac{y^H P_H y}{\sigma^2} - p_r(z, \sigma^2, \beta)$$

(1)

where $P_H = H (H^H H)^{-1} H^H$ and

$$p_r(z, \sigma^2, \beta) = \ln \left[ \int \exp\left(-\frac{1}{\sigma^2} \|z - \phi\|^2\right) f_\phi(\phi; \beta) d\phi \right].$$

(2)

Equations (1) and (2) hold regardless of the distribution of $\phi = (H^H H)^{1/2} \theta$ and constitute our most general result when the noise power is known. This result says that the appropriate detector in this case is a matched subspace detector reduced by a penalty term. We call this penalty term a resolving penalty because if $y$, when projected into the space spanned by $(H)$, lies far from the “favored” orientations dictated by $f_\phi(\theta)$, then the penalty is high. Succinctly, both the subspace power and orientation matter in the stochastic case.

If $\beta$ is known we can evaluate $p_r$ at the given parameters and $s$ in (1) serves at the log-likelihood ratio detector. In the case where $\beta$ is unknown, the GLRT principle can be applied and $p_r$ is minimized over the unknown parameters and then compressed with the corresponding ML estimates.

4. DETECTOR FOR GAUSSIAN SIGNAL AMPLITUDES AND KNOWN NOISE POWER

In this section we assume that the random vector $\theta$ is distributed as $\theta \sim CN_p[0, R_{\theta \theta}]$, with mean $\theta$ and covariance $R_{\theta \theta}$. Recall that the random vector $\phi$ was defined as $\phi = (H^H H)^{1/2} \theta$. Then $\phi \sim CN_p[0, R_{\phi \phi}]$ where

$$R_{\phi \phi} = (H^H H)^{1/2} R_{\theta \theta} (H^H H)^{1/2}.$$

The parameters that describe the covariance matrix $R_{\phi \phi}$ are summarized by the parameter vector $\beta$ in the likelihood functions described above. In this instance the penalty term becomes

$$p_r = -\ln E_\phi \left[ \exp\left(-\frac{1}{\sigma^2} \|\phi - z\|^2\right) \right]$$

$$= -\ln \left[ \int \exp\left(-\frac{1}{\sigma^2} \|\phi - z\|^2\right) \times \frac{1}{\pi^p \det(R_{\phi \phi})} \exp(-\phi^H R_{\phi \phi}^{-1} \phi) d\phi \right]$$

$$= \ln \left[ \det R_{\phi \phi} \det\left(\frac{I}{\sigma^2} + R_{\phi \phi}^{-1}\right) + z^H (\sigma^2 I + R_{\phi \phi})^{-1} z \right]$$

$$= \ln \left[ \det(I + R_{\phi \phi}/\sigma^2) \right] + z^H (\sigma^2 I + R_{\phi \phi})^{-1} z.$$

In deriving the detector, we re-parameterize the covariance matrix $R_{\phi \phi}$ as follows:

$$R_{\phi \phi} = (H^H H)^{1/2} R_{\theta \theta} (H^H H)^{1/2}$$

$$= GD^2 G^H; \quad G \text{ unitary}; \quad D^2 = \text{diag}[\beta_1^2, ..., \beta_p^2].$$

The penalty term can now be written as

$$p_r = \sum_{i=1}^p \ln(1 + \beta_i^2/\sigma^2) + \sum_{i=1}^p \frac{z_i^H P_{\phi i} z_i}{\sigma^2 + \beta_i^2}.$$  

(3)

Here $g_i$ are the orthonormal columns of $G$ and $P_{\phi i}$ is the corresponding orthogonal projection matrix. Define the “resolved signal plus noise to noise ratios”

$$r_i = 1 + \beta_i^2/\sigma^2$$
and write the penalty term as
\[ p_r = \sum_{i=1}^{p} \ln(r_i) + \sum_{i=1}^{p} \frac{(\tilde{z}^H P_{g_i} z / \sigma^2)}{r_i}. \] (4)

This is a general result for a Gaussian coordinate-vector with known \( R_{\phi \phi} \) and known additive noise power \( \sigma^2 \).

Suppose that \( G \) and \( \{\beta_i^2\}_{i=1}^p \) are unknown (recall we have assumed to this point that \( \sigma^2 \) is known). The GLRT detector can be obtained by finding the estimates of \( G \) and \( \{\beta_i^2\}_{i=1}^p \) and compressing the statistic in (1). In (1), the unknown parameter vector \( \beta \) is comprised of the parameters that define \( G \) and \( \{\beta_i^2\}_{i=1}^p \). In this case, we can equivalently find a set of normal equations by differentiating \( p_r \) in (4) with respect to the \( r_i \). Solving these normal equations yields the estimates
\[ \hat{r}_i = \max \{1, z^H P_{g_i} z / \sigma^2\}. \]

We will assume that “resolved signal plus noise to noise ratio”
\[ \hat{r}_i = \frac{y^H P_{g_i} y}{\sigma^2} \] (5)
is greater than 1 in the following derivations. That is, we assume there exists a unitary \( G \) such that this constraint can be met. Compressing the penalty term with these estimates yields the partially compressed statistic
\[ s(y; \sigma^2, \{\beta_i^2\}, G) = \frac{y^H P_{H} y}{\sigma^2} - p_r(y; \sigma^2, \{\beta_i^2\}, G) \]
\[ = \frac{y^H P_{H} y}{\sigma^2} - \left[ p + \ln \left( \prod_{i=1}^{p} \hat{r}_i \right) \right]. \]

This statistic may be compressed one more time with respect to the basis \( G \). However, as a preliminary to this we first minimize the product \( \prod \hat{r}_i \) subject to the constraints that \( \hat{r}_i \geq 1 \) and the constraint that the \( \hat{r}_i \) sum to \( r \) where \( r \) is the unresolved signal plus noise to noise ratio:
\[ r = \sum_{i=1}^{p} \hat{r}_i = \frac{y^H P_{H} y}{\sigma^2} \geq p. \]

The sum constraint simply says that the energy in the individual subspaces must account for all of the energy in the overall subspace. The solution is
\[ \hat{r}_i = 1, i = 1, \ldots, p-1 \quad \& \quad \hat{r}_p = r - (p-1). \]

Compressing the statistic with these estimates gives
\[ s = r - [p + \ln(r - (p - 1))]. \]

The term \( r - (p - 1) \) in the penalty term of this formula for compressed likelihood is guaranteed to be non-negative, by virtue of the constraint that all of the \( \hat{r}_i \geq 1 \). Consequently \( s(r) \) is monotonic in \( r \) for \( r > p \) and we can just use \( r \) as the detection statistic. That is, the GLRT statistic, when the coordinate-vector is Gaussian with unknown covariance, is simply the matched subspace detector. The only question that remains in this compression of likelihood is whether there exists a corresponding choice of coordinate system \( \hat{G} \) that achieves these maximizing values for the \( \hat{r}_i \). The following paragraph clarifies this point.

The construction of \( \hat{G} \) may be described as follows: the measurement \( y \) is resolved onto the subspace \( \langle H \rangle \) where its energy \( y^H P_{H} y \) is computed. Then the subspace \( \langle G \rangle \) is rotated so that the energy resolved onto the one-dimensional subspace \( \langle g_{i_p} \rangle \) equals \( y^H P_{H} y - (p - 1)\sigma^2 \). The rest of the subspace is then rotated around \( \langle g_{i_p} \rangle \) until equal energy of \( \sigma^2 \) is resolved onto each of the \( p - 1 \) dimensions. Why should the maximum likelihood estimator of the subspace \( G \) work like this? The answer is that ML treats the total energy \( y^H P_{H} y \) as if it were signal energy plus \( p \) units of noise variance: \( y^H P_{H} y = E_H + p\sigma^2 \), which it resolves by placing all of the signal energy, plus one unit of noise power, in one coordinate \( \langle g_{i_p} \rangle \) and the remaining \( p - 1 \) units of noise power in the remaining \( p - 1 \) coordinates, equally. That is, it rotates the coordinate system \( \langle G \rangle \) so that the total energy resolved in the subspace \( \langle G \rangle \), or equivalently \( \langle H \rangle \), is decomposed as follows:
\[ y^H P_{H} y = E_H + p\sigma^2 = [E_H + \sigma^2] + [(p - 1)\sigma^2] \]
\[ = [y^H P_{g_{i_p}} y] + \sum_{i=1}^{p-1} [\sum_{i=1}^{p} y^H P_{g_{i_p}} y]. \]

This equation may be divided through by \( \sigma^2 \) to produce the corresponding formula
\[ r = \hat{r}_p + \sum_{i=1}^{p-1} \hat{r}_i, \]
which is the required constraint. This solution allows the detector to account for \( p \) units of noise.
variance, which is known apriori to be uniformly distributed in $C^N$, and to resolve all of the remaining signal energy, which it takes to be the measured energy plus the apriori noise power, in a one-dimensional subspace.

5. UNKNOWN NOISE POWER

In this section we assume that the noise power $\sigma^2$ is unknown. Consequently, the random vector $w$ will be used, along with $z$ to estimate the unknown parameters. We implicitly use many of the results from the previous section in deriving the detector in this case.

Under hypothesis $H_0$ the noise power estimate is $\hat{\sigma}^2 = y_H^*y/N$. We can use this to obtain the partially compressed, conditional GLRT statistic

$$v = e^{N\left(\frac{y_H^*y}{N}\right)} \left[\frac{1}{(\sigma^2)^{N-p}} \exp(-\frac{1}{\sigma^2}w^Hw)\right] \times \left[\frac{1}{(\sigma^2)^p} \exp(-\frac{1}{\sigma^2}\|z - \phi\|^2)\right].$$

Using the same methods as above, we obtain the unconditional partially compressed GLRT statistic

$$\tilde{v} = e^{N\left(\frac{y_H^*y}{N}\right)} \left[\frac{1}{(\sigma^2)^{N-p}} \exp(-\frac{1}{\sigma^2}w^Hw)\right] \times \frac{1}{(\sigma^2)^p} \left(\prod_{i=1}^{p} \frac{1}{r_i}\right) \exp(-\sum_{i=1}^{p} \frac{z_i^H P_{\ell_i} z / \sigma^2}{r_i}).$$

where $r_i = 1 + \beta_i^2/\sigma^2$. The ML normal equations are

$$-\frac{1}{r_i} + \frac{(z_i^H P_{\ell_i} z / \sigma^2)}{r_i} = 0,$$

$$-\frac{(N-p)}{\sigma^2} + \frac{w_i^H w}{(\sigma^2)^2} - \frac{1}{\sigma^2} \sum_{i=1}^{p} \frac{z_i^H P_{\ell_i} z / (\sigma^2)^2}{r_i} = 0.$$

Again, if we ignore the case where the data are such that

$$r_i = \frac{z_i^H P_{\ell_i} z}{w_i^H w} = \frac{z_i^H P_{\ell_i} z}{y_H^*(I - P_H)y},$$

$$= \frac{z_i^H P_{\ell_i} z}{\sigma^2},$$

is less than 1, then we can compress the GLRT with the estimates implicitly defined in (6) to obtain a detector

$$\tilde{s} = \left(1 + \frac{y_H^* P_H y}{y_H^*(I - P_H)y}\right)^N.$$

In the above we have removed known constants. Since $\tilde{s}$ is a monotonic function of

$$s = \frac{y_H^* P_H y}{y_H^*(I - P_H)y}$$

we can use $s$ in (7) as the detector. This detector is identical to the CFAR matched subspace detector described for first-order models in [2].

6. RANK-ONE SIGNALS, GENERALIZED RAYLEIGH AMPLITUDES

In the rank-one case, the log-likelihood ratio test becomes

$$s = \frac{y_H^* P_H y}{\sigma^2} - p_r(y; \sigma^2; \beta)$$

where

$$p_r = -\ln \int \exp(-\frac{\|z - \phi\|^2}{\sigma^2}) f_\phi(\phi, \beta).$$

We assume that the signal subspace vector $h$ is unit-norm so $\phi = (h^H h)^{1/2} \theta = \theta$. Write $\theta$ as $\theta = Me^{j\psi}$ where we assume that the magnitude $M$ and the phase $\psi$ are independent. Moreover, let’s assume that the phase is uniformly distributed $U[0, 2\pi]$. Define $z = m_x e^{j\alpha}$. The penalty term can be written as

$$p_r = -\ln \left[\int \exp(-\frac{M^2 + m_x^2}{\sigma^2} \times \int \exp(\frac{2M m_x \cos(\psi - \alpha)}{\sigma^2}) f_M(M) f_\psi(\psi) dM d\psi\right]$$

$$= -\ln \left[\int \exp(-\frac{2m_x^2}{\sigma^2}) \int \exp(-\frac{M^2}{\sigma^2}) \times I_0(2M m_x / \sigma^2) f_M(M) dM\right]$$

where $I_0(.)$ represents a modified Bessel function of order 0. The series representation of this function
can be used to write

\[ p_r = -\ln \left[ \exp(-\frac{m_z^2}{\sigma^2}) \sum_{n=0}^{\infty} \int_0^{\infty} \frac{(Mm_z/\sigma^2)^{2n}}{n!} \right] \exp(-\frac{M^2}{\sigma^2}) f_M(M) dM \]

Now suppose that the signal magnitude had a generalized Rayleigh distribution (with L a non-negative integer)

\[ f_M(M) = \frac{2M}{\beta^2} \left( \frac{M^2}{\beta^2} \right)^L e^{-M^2/\beta^2} \frac{1}{L!} \]

We can write the penalty term as

\[ p_r = -\ln \left[ \exp(-\frac{m_z^2}{\sigma^2}) \left( \frac{\sigma^2}{\sigma^2 + \beta^2} \right)^{L+1} \sum_{n=0}^{\infty} \left( \frac{m_z^2 \beta^2}{\sigma^2(\sigma^2 + \beta^2)} \right)^n \frac{(n + L)}{n! L!} \right] \]

This expression can be reduced to

\[ p_r = -\ln \left[ \exp(-\frac{m_z^2}{\sigma^2}) \left( \frac{\beta^2}{\sigma^2 + \beta^2} \right)^{L+1} \left[ \frac{d^L}{dx^L} \left( x^L \sum_{n=0}^{\infty} \frac{x^n}{n! L!} \right) \bigg|_{x=(m_z^2 \beta^2)/(\sigma^2(\sigma^2 + \beta^2))} \right] \right] \]

\[ = -\ln \left[ \exp(-\frac{m_z^2}{\sigma^2}) \left( \frac{\beta^2}{\sigma^2 + \beta^2} \right)^{L+1} \frac{1}{L!} \left[ \frac{d^L}{dx^L} (x^L e^x) \bigg|_{x=(m_z^2 \beta^2)/(\sigma^2(\sigma^2 + \beta^2))} \right] \right] \]

This reduces to (with \( r = 1 + \beta^2/\sigma^2 \))

\[ p_r = (L + 1) \ln(r) + \frac{(m_z^2/\sigma^2)}{r} \ln \sum_{k=0}^{L} \left( \frac{L}{k} \right) \frac{1}{k!} \left( \frac{(m_z/\sigma^2)(r - 1)}{r} \right)^k \]

This equation can be minimized with respect to \( r \) to obtain the ML penalty term for the given \( L \). Note we solved the \( L = 0 \) case in Section 3.

7. CONCLUSIONS

In this paper we have derived generalized likelihood ratios (GLRTs) for detecting stochastic subspace signals. For an important class of such signals the GLRT is, in fact, a matched subspace detector (MSD) of the type previously reported in [2]. This means MSDs are more generally applicable than originally thought. We also have shown that in the known noise case the GLRT detector is composed of a matched subspace detector reduced by a penalty term that depends on the distribution of the coordinate-vector. This penalty term arises because one presumes to know something, in a statistical sense, about the distribution of signal power and coordinates. If the estimated signal-coordinate vector is unlikely in this model, then the penalty term is large.

8. REFERENCES


