A Novel Differential Geometric Approach Toward Robust Signal Detection

M.W. Thompson and D.R. Halverson

1 November 1985

Reprint FROM 5/23/85 TO 5/22/86

Proc. of the 1986 Conf. on Information and System Sciences, March 19-21, 1986

Field Group Sub. Gr.

Differential geometric approach, robust detection.

We present a new approach toward robust signal detection which is based on techniques rooted in differential geometry. These methods, as opposed to the commonly employed classical saddlepoint criteria, readily admit the quantitative measure of the degree of robustness over very general classes of admissible noise distributions. Our approach thus is seen to make possible investigations of the quantitative tradeoff between optimal performance and robustness, and we illustrate the application of this differential geometric approach via various specific examples.
We present a new approach toward robust signal detection which is based on techniques rooted in differential geometry. These methods, as opposed to the commonly employed classical saddlepoint criteria, readily admit the quantitative measure of the degree of robustness over very general classes of admissible noise distributions. Our approach thus is seen to make possible investigations of the quantitative tradeoff between optimal performance and robustness, and we illustrate the application of this differential geometric approach via various specific examples.

I. Introduction

It is well known that there is increasing interest in the employment of robustness techniques for the discrete time detection of signals in imperfectly known noise. The traditional approach toward addressing questions within this rather broad area of research has been to rely heavily on the classical saddlepoint criterion of Huber (see, for example, [1]). A variety of work appearing in the engineering literature has verified that such an approach can lead to tractable results. However, it may be argued that the degree of robustness obtained owes much to the types of noise models admitted by the method. In reality it may not be easy to verify that the types of models appropriate to the saddlepoint robustness approach sufficiently represent the full extent of variation of the unknown perturbation of a distribution around the nominal. Moreover, although it is possible via the approach of [2] to obtain general representations of the noise model via Choquet capacities [3], it has yet to be seen if such elegant methods are capable of enhancing the denseness of the class of noise models beyond the relatively few standard models (see, for example, [1]). In addition, the saddlepoint criterion is inherently a nonquantitative approach toward imparting robustness. We intuitively might suspect that robustness is obtained by a judicious tradeoff with optimal performance, and we thus might desire a way to quantitatively measure the degree of robustness in order that a weighted combination of robustness and performance could be considered subject to some cost criterion. In this paper we present an entirely different approach which views the robustness question not from the saddlepoint perspective but from one which is rooted in differential geometry.

II. Development

Viewing the robustness problem from a slightly different perspective, let $D_n$ denote the class of $n$-dimensional distribution functions. From this point of view, the performance of the detector is thus expressed by considering the performance functional $P: D_n \rightarrow \mathbb{R}$; we then simply wish to choose the detector so that $P$ is reasonably high and doesn't vary much near the nominal element of $D_n$. Viewing $P$ as a height function over $D_n$, we could say that a robust detector would yield a "surface"
which above the nominal element is both relatively high and not strongly sloped.

Such a perspective thus would indicate that a geometric approach to robust detection might be appropriate. What would be needed would be to provide a differentiable structure to $\mathcal{D}$ so that the concept of slope would have the proper meaning. We would then be considering a height function over a differentiable manifold $\mathcal{M}$ which would result in a new manifold $\mathcal{M}'$, for which the Riemannian metric would yield a norm.

In this paper we present some specific applications of the above observations. Noting that a Neyman-Pearson approach involves comparing the sample vector to the appropriate $n$-dimensional Borel set $\mathcal{B}$ in $\mathbb{R}^n$, where $n$ is the number of samples, we then observe that in this robustness application we could in practice regard $\mathcal{B}$ as specified via the choice of nominal distribution under $H_0$; we then would be interested in analyzing the degree of variation in the false alarm probability $\alpha$ and/or detection probability $\beta$ as the underlying distribution varies about the nominal, thus fixing a choice of height function $h: \mathbb{R}^m \rightarrow \mathbb{R}$ for some natural number $m$, where $h(\cdot)$ corresponds to the value of $\alpha$ or $\beta$ for some fixed detector of interest.

Consider first the case where the class of $n$-dimensional distribution functions is parameterized by $m$ parameters; this class can then be identified with a subset of $\mathbb{R}^m$, and the corresponding manifold $\mathcal{M}$ is a surface in $\mathbb{R}^{m+1}$. An appropriate metric tensor $g(\cdot, \cdot)$ is inherited from the standard inner product on $\mathbb{R}^{m+1}$ with the obvious choice of coordinate system $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ leading to the components of the metric tensor given by

$$
g_{ij} = g\left(\frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}, \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|}\right) = \begin{cases} \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{\|\mathbf{v}_i\|^2} & \text{if } i \neq j \\ 1 & \text{if } i = j \\ \end{cases}
$$

Associating the slope of the unit normal with the cosine of the angle of the unit normal to vertical (with $\mathbf{n}$ immersed in $\mathbb{R}^{m+1}$) it is then straightforward, although somewhat lengthy, to show that at the point corresponding to the nominal distribution this cosine is given by

$$
\cos \gamma = \left(1 + \sum_{i=1}^{m} \frac{\partial h}{\partial x_i} \right)^2 - 1/2
$$

Note that $\gamma$ provides a measure of local "first order" robustness; smaller values of $\gamma$ suggest less variation in $\alpha$ or $\beta$ near the nominal distribution.

Consider now the discrete time detection of a constant signal $s$ in additive i.i.d. Gaussian noise with mean $\mu$ and variance $\sigma^2$. We note that there may be some uncertainty in all of the values of $s, \mu$ and $\sigma^2$. Employing first the linear detector, we then choose $h(\cdot)$ to correspond to $\beta$ and then straightforwardly obtain (for $n$ samples)

$$
d \frac{\partial \alpha}{\partial s} \frac{\partial \beta}{\partial u} = \left(n/(2\pi\sigma^2)\right)^{\frac{1}{2}} \exp\left[-\left(\mathbf{u} + \mathbf{s} - \mathbf{T}\right)^2/(2n\sigma^2)\right]$$

where the threshold $T$ is specified for a given false alarm rate $\alpha$ by evaluating the detector at the nominal values of $s, \mu$ and $\sigma^2$. We next employ the robustified version of the linear detector, which replaces the identity function of the linear detector with the nonlinearity $g(\cdot)$ defined by

$$g(x) = \begin{cases} k_2 & \text{if } x > k_2 \\ x & \text{if } k_1 \leq x \leq k_2 \\ k_1 & \text{if } x < k_1 \end{cases}$$

It is well known that this detector resists the tractable development of closed form expressions for $\alpha$ or $\beta$ in the Gaussian case. In order to numerically compare the robustness of this "censored" version of the linear detector we therefore employ a large sample Gaussian approximation of the test statistic. The resultant lengthy analysis shows, for example, that with $n=50, \alpha=0.05$, $k_1=-0.4, k_2=0.6$ and nominally $\mu=0.1, \sigma=0.1$, $\gamma = -68.8^\circ$, which can be compared to $\gamma = 34.3^\circ$, which can be compared to $\gamma = 34.3^\circ$ for the linear detector.
The robustness of the detector which employs censoring is thus quantitatively demonstrated.

III. The General Case

It would also be important to consider more general classes of distributions; ideally we would wish to place virtually no constraints on the admissible distribution function. Such an approach is actually feasible in the i.i.d. case. Since \( \alpha \) and \( \beta \) are expressed via an integral over a Borel set \( B_n \) with respect to the appropriate \( n \)-dimensional distribution under \( H_0 \) and \( H_1 \) respectively, we can without loss of generality note the independence of the observations and investigate perturbations in \( \alpha \) and \( \beta \) by limiting consideration to the class of those univariate distributions given by step functions, i.e. those functions of form

\[
F(\cdot) = \sum_{i=0}^{m+1} a_i \mathbb{I}_{A_i}(\cdot),
\]

where the intervals \( A_i \) partition \( R \) and we take \( a_0 = 0 \) and \( a_{m+1} = 1 \). For a fixed finite partition \( P \) of \( R \), we note that the corresponding class of step functions can be viewed as parameterized by elements of \( R^m \). Letting \( F(\cdot) \) denote the nominal univariate distribution of the observations, we then can employ the aforementioned methods to obtain an expression for \( \cos y \), where for each partition a Stieltjes approximation to \( F(\cdot) \) is chosen and regarded as nominal for the parameterized case. We then define

\[
\cos y = \lim_{1 \to 0} \left( 1 + \sum_{j \in R} \frac{2^{-2} \beta_{ij}}{\beta_{ij}} \right) \cos y,
\]

where, as before, we limit consideration to the situation where the height function \( h(\cdot) \) is \( \mathbb{I}(\cdot) \) (with the corresponding univariate distribution of the observa-

tions \( F(\cdot) \)). Now let

\[
\delta^+_n = \{(x_1, x_2, \ldots, x_n) : \text{there exists } \epsilon > 0 \text{ such that }
\]

\[
(x_1, \ldots, x_i-1, y, x_i+1, \ldots, x_n) e B_n \quad \text{for } y \in (x_i-\epsilon, x_i) \text{ and }
\]

\[
(x_1, \ldots, x_i-1, z, x_i+1, \ldots, x_n) e B_n \quad \text{for } z \in (x_i, x_i+\epsilon)
\]

\[
\delta^-_n = \{(x_1, x_2, \ldots, x_n) : \text{there exists } \epsilon > 0 \text{ such that }
\]

\[
(x_1, \ldots, x_i-1, y, x_i+1, \ldots, x_n) e B_n \quad \text{for } y \in (x_i-\epsilon, x_i)
\]

and

\[
(x_1, \ldots, x_i-1, z, x_i+1, \ldots, x_n) e B_n \quad \text{for } z \in (x_i, x_i+\epsilon)
\]

\[
\delta^{++}_n = \{(x_1, \ldots, x_i-1, x_i+1, \ldots, x_n, y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_n) : \text{there exists } w \text{ such that }
\]

\[
(x_1, \ldots, x_i-1, w, x_i+1, \ldots, x_n) e B_n \quad \text{and }
\]

\[
(y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_n) e B_n.
\]

Similarly, we define \( \delta^{--}_n \) by replacing \( \delta^+_n \) in the \( \delta^{++}_n \) expression with \( \delta^-_n \). In an analogous manner, we also define \( \delta^{+-}_n \) and \( \delta^{-+}_n \). We then can establish the following result, which provides a closed form expression for \( \cos y \):

**Theorem:** Suppose that if

\[
x \in R^n, \quad \beta_{ij} = \begin{cases} 0 & \text{if } x_j = x_{j+1} \\ 1 & \text{otherwise} \end{cases}
\]

then \( x \in \text{int}(B_n) \cap \text{int}(B_n) \). We then have

\[
\cos y = \left(1 + \Delta^2 + \frac{1}{2}\right), \quad \text{where}
\]

\[
\Delta = \frac{1}{2} \left[ \sum_{j \neq i} \int_{B_i} dF(y_1) \ldots dF(y_{i-1}) dF(y_{i+1}) \ldots dF(y_n) - \int_{B_i} dF(y_1) \ldots dF(y_{i-1}) dF(y_{i+1}) \ldots dF(y_n) \right] \frac{1}{2}
\]

\[
\quad - \sum_{j \neq i} \left[ \int_{B_i} dF(x_1) \ldots dF(x_{i-1}) dF(x_{i+1}) \ldots dF(x_n) \right] \frac{1}{2}
\]

\[
\quad + \int_{B_i} dF(x_1) \ldots dF(x_{i-1}) dF(x_{i+1}) \ldots dF(x_n) \ldots dF(y_n) + \ldots dF(y_{j-1}) dF(y_{j+1}) \ldots dF(y_n). \]
of $B_n$ to its interior, whereas $\partial B_n$ is formed in an analogous manner with the rays moving in the negative direction from the exterior of $B_n$ to its interior. In addition, we note that the Theorem's hypothesis is simply a mild condition on the regularity of $B_n$ which is frequently easy to satisfy in this detection context (wherein $B_n$ arises by way of a threshold comparator). Moreover, the existence of the integrals in the Theorem is very often easy to verify since the boundary of $B_n$ is sufficiently well behaved in such cases.

As an example of an application of the Theorem, consider again the linear detector. For $n=2$ it then follows that (where the detector threshold is $T$)

\begin{align*}
\mathcal{S}_2 &= \{ (x,y) : y = T-k \} \\
\mathcal{S}_2^\perp &= \mathcal{S}_2
\end{align*}

We therfore have $\cos\gamma = 1/(1+1+1+0)^{1/3} = 1/3$, i.e. $\gamma = 54.7^\circ$, regardless of the nominal distribution. This may be generalized to show that for $n$ samples $\cos\gamma = 1/(1+1+1+\cdots)^{1/(n+1)}$, regardless of the nominal distribution. Note that $\lim_{n \to \infty} \cos\gamma = 0$, i.e. the linear detector becomes completely unrobust (as measured by $\gamma$) as the number of samples approaches infinity.

On the other hand, consider the classical robustified version of the linear detector, wherein a "censored" detector non-linearity $g(\cdot)$ of form

\[ g(x) = \begin{cases} 
  x & \text{if } |x| \leq k \\
  -k & \text{if } x < -k \\
  k & \text{if } x > k 
\end{cases} \]

It then can be shown that

\begin{align*}
\mathcal{S}_2^\perp &= \{ (x,y) : y = T-k \} & \text{or } y = T-x, \\
\mathcal{S}_2 &= \{ (x,y) : y = T-k, x > k \} & \text{or } y = T-x, \\
\mathcal{S}_2^\perp &= \mathcal{S}_2
\end{align*}

For this case the exact value of $\cos\gamma$
depends on the values of $k$, $T$, and the choice of nominal distribution. However, we can make some general conclusions when the amount of censoring approaches maximal ($k=0$). In this case we have, when $n=2$,
\[
\lim_{k \to 0} \cos \gamma \approx (1+2(1-F(0))^2+2(1-F(0))^2)^{-\frac{1}{2}}.
\]

For the common case where $F(0) \leq \frac{1}{2}$, we then obtain
\[
\lim_{k \to 0} \cos \gamma \approx 1/2, \text{ i.e. } \lim_{k \to 0} \gamma \geq 45^\circ,
\]
which may be compared to the linear detector's $\gamma = 54.7^\circ$ for $n=2$. This may be generalized through a lengthy analysis to conclude that for $n$ samples,
\[
\lim_{k \to 0} \cos \gamma = (1+n(1-F(0))^2)^{n(n-1)+n(n-1)}.
\]

Note that for $F(0)>0$ we have $\lim \lim \cos \gamma = 1$, that is, the detector approaches possessing complete robustness (as measured by $\gamma$) as the number of samples tends to infinity. For $F(0) \leq \frac{1}{2}$ we have
\[
\lim_{k \to 0} \cos \gamma \approx (1+n \cdot 2(n-1)+n(n-1))^{-\frac{1}{2}}.
\]

Note that the upper bound can be approached arbitrarily closely for $F(0)$ near $\frac{1}{2}$. For $n=3$ this upper bound becomes 0.8 (corresponding to $\gamma = 36.1^\circ$), whereas for $n=10$ it becomes 0.96 (corresponding to $\gamma = 16.3^\circ$). This can be compared to the case of the linear detector, where for $n=3$ we have $\gamma = 60^\circ$ and for $n=10$ we have $\gamma = 72.5^\circ$. For the larger values of $n$ the robustness advantages of the classical robustified linear detector are thus quantitatively demonstrated.

**III. Conclusion**

We have presented a new approach toward robust signal detection which is based on differential geometric methods as opposed to classical saddlepoint criteria. These techniques are seen to admit a quantitative measure of robustness through the geometric concepts of unit normal slope and scalar curvature, thus allowing the consideration of a weighted combination of performance, first order robustness (via unit normal slope), and second order robustness (via scalar curvature) subject to some cost criterion of interest. Our techniques are additionally illustrated in the paper through various specific examples.

**ACKNOWLEDGEMENT**

This research was supported by the Air Force Office of Scientific Research under Grant AFOSR-82-0033.

**References**


END

DTIC

8-86