Research Article

An Information Geometric Viewpoint on the Detection of Range Distributed Targets

Mengjiao Tang, Yao Rong, and Jie Zhou

College of Mathematics, Sichuan University, Chengdu, Sichuan 610064, China

Correspondence should be addressed to Jie Zhou; jzhou@scu.edu.cn

Received 1 December 2014; Accepted 16 April 2015

Academic Editor: Haranath Kar

Copyright © 2015 Mengjiao Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The paper adopts the information geometry, to put forward a new viewpoint on the detection of range distributed targets embedded in Gaussian noise with unknown covariance. The original hypothesis test problem is formulated as the discrimination between distributions of the measurements and the noise. The Siegel distance, which is exactly the well-known geodesic distance between images of the original distributions via embedding into a higher-dimensional manifold, is given as an intrinsic measure on the difference between multivariate normal distributions. Without the assumption of uncorrelated measurements, we propose a set of geometric distance detectors, which is designed based on the Siegel distance and different from the generalized likelihood ratio algorithm or other common criterions in statistics. As special cases, the classical optimal matched filter, Rao test, and Wald test, which have been proven to have the CFAR property, belong to the set. Moreover, it is also accessible to an intuitively geometric analysis about how strongly the data contradict the null hypothesis.

1. Introduction

Worthy of repetitive thought is the problem of detecting range distributed targets in the presence of Gaussian noise with unknown covariance. In the actual projects, measurements are collected via radar or sonar, which will be modeled as reflections from targets. With the collected data, a hypothesis test aims at distinguishing between target returns plus noise and noise only. In 1986, Kelly firstly proposed his famous generalized likelihood ratio test (GLRT) for detecting point targets in [1]. Later, a lot of advances on the detection of distributed targets are made based on the generalized likelihood ratio algorithm. There is no exception for the recent work that mainly focuses on the research of adaptive detection. In this field, one class is dealing with completely unknown signals [2, 3]. Specifically, paper [2] by Conte et al. provides a valuable statistical tool for ensuring the CFAR property. Others consider detecting a rank-one steering vector [4–6], or a spread-spectrum signal [7, 8], and so forth. All these detectors are reasonably testified to have perfect detection performances in corresponding situations.

However, a common problem is that the detectors designed by the generalized likelihood ratio algorithm are limited to the case of determinate signals, due to the given form of distribution under the alternative hypothesis. In this paper, the detection problem with completely unknown target returns is discussed. In addition, there is no assumption whether the measurements are correlated or not. The original hypothesis test problem is formulated as a measure on the difference between distributions of the measurements and the noise, which is intuitively described via the platform of information geometry.

The differential geometric approach was first introduced by Rao [9], in order to supply a way for the construction of distances between parametric density functions. Later related papers greatly improved Rao's concept. It is worth mentioning that Amari presented statistical models with a differential structure in [10, 11], which plays an important role in information geometry. The geodesic distance has been put forward as the shortest length between two distributions on the statistical manifold endowed with the Fisher metric and Levi-Civita connection. A very active field of

information geometry is about the manifold of multivariate normal distributions. On the multivariate normal manifold, the explicit expression of the geodesics has been derived by Eriksen [12] and later by Calvo and Oller [13], when the initial distribution and direction are given. However, the geodesic distance between multivariate normal distributions has not yet been obtained, except for on the submanifold of multivariate normal distributions with the same mean [14]. The lack of the explicit expression greatly limits the application of the geodesic distance. In 1990, Calvo and Oller proposed the Siegel distance on the multivariate normal manifold via embedding into the Siegel group [15]. The Siegel distance aims at calculating the geodesic distance between the images of the initial distributions and also provides a lower bound for the geodesic distance. Moreover, it is proven by Calvo et al. in [16] that the Siegel distance shows the similar behavior with the geodesic distance in special cases. This suggests that the Siegel distance is reasonable to be applied to measure the difference between the general multivariate normal distributions.

In this paper, the key point is how to calculate the Siegel distance between the distributions of the measurements and the noise. Since the target returns are completely unknown, the possible distributions of measurements form a submanifold of multivariate normal distributions. By calculating the Siegel distance $\mathcal{D}$ from some point on the multivariate normal manifold to the submanifold, a set of detectors are designed with the critical region

$$W_\epsilon = \{z_t \mid \mathcal{D} > \gamma_\epsilon\}, \quad (1)$$

where ϵ denotes the significance level and the constant γ_ϵ is chosen to satisfy

$$\text{Prob} \{z_t \in W_\epsilon \mid H_0\} \leq \epsilon. \quad (2)$$

From the viewpoint of information geometry, it also provides an intuitive analysis about how strongly the measurements contradict the null hypothesis. Moreover, the optimal matched filter, Rao test (or modified 2-step GLRT), and Wald test (or 2-step GLRT), which have been proven to have the CFAR property in [2], can also be derived from the set of detectors.

The outline of the paper is as follows: Section 2 shows the problem formulation. Section 3 reviews some important principles related to the geodesic distance between multivariate normal distributions. Section 4 is devoted to the Siegel distance and the design of geometric distance detectors and provides a discussion on the choice of the nominal noise covariance matrix in Section 5. Some intuitive interpretations are presented in Section 6. Section 7 gives a conclusion.

2. Problem Formulation

The problem under investigation is the detection of completely unknown target returns in the presence of Gaussian noise with unknown covariance. We assume the targets are spatially distributed across H range cells, from which the primary data $\{z_i\}$, $i = 1, \dots, H$, are collected. The primary data consist of the possible target returns plus Gaussian

noise. To acquire information about the noises, the secondary data coming from K cells arranged around the targets are supposed to be available. Or rather, the secondary data consist of a training set $\{z_i\}$, $i = H + 1, \dots, H + K$, of which each sample shares the same distribution as the Gaussian noise adherent to the primary data. Therefore, the detection problem is formulated as

$$\begin{aligned} H_0 : z_i &= \mathbf{n}_i, \quad i = 1, \dots, H + K, \\ H_1 : \begin{cases} z_i = \mathbf{p}_i + \mathbf{n}_i, & i = 1, \dots, H, \\ z_i = \mathbf{n}_i, & i = H + 1, \dots, H + K, \end{cases} \end{aligned} \quad (3)$$

where, $\mathbf{p}_i$, $i = 1, \dots, H$, denote the signals, and the noises $\mathbf{n}_i$, $i = 1, \dots, H + K$, are independent Gaussian vectors with mean zero and the same unknown covariance matrix $\mathbf{M}$. All $\mathbf{p}_i$ s, $\mathbf{n}_i$ s, and $\mathbf{z}_i$ s are p -dimensional vectors.

Let $\mathbf{Z}_p^T = [\mathbf{z}_1^T, \dots, \mathbf{z}_H^T]$ and $\mathbf{Z}_s^T = [\mathbf{z}_{H+1}^T, \dots, \mathbf{z}_{H+K}^T]$. In fact, the hypotheses (3) carry out a comparison between the distributions of $\mathbf{z}_i$ and the noise. As usual, considering the distribution of the whole measurement set $\mathbf{Z}_p$ is verified to be more effective [4].

Suppose the joint probability distribution of the measurements $\mathbf{Z}_p$ is Gaussian with an unknown covariance matrix Σ . Let $\mathbf{Z}_p^0$ be the realization of random vector $\mathbf{Z}_p$. In the real-time processing such as radar target detection, owing to the unknown mean of $\mathbf{Z}_p$, the multivariate normal sampling model for $\mathbf{Z}_p$ is regarded as

$$\mathcal{N}(\mathbf{Z}_p^0, \Sigma), \quad \Sigma \in \mathbb{S}_+^{Hp}, \quad (4)$$

where $\mathbb{S}_+^n$ denotes the set of $n \times n$ symmetric positive-definite matrices. As a comparison, an array of noise with the same size is considered. Let $\mathbf{N}_p^T = [\mathbf{n}_1^T, \dots, \mathbf{n}_H^T]$; then we can specify the sampling distribution of $\mathbf{N}_p$ as

$$\mathcal{N}(\mathbf{0}_{Hp}, \Lambda), \quad (5)$$

where $\Lambda \in \mathbb{S}_+^{Hp}$ denotes the block diagonal matrix having the same block $\mathbf{S} \in \mathbb{S}_+^p$. The matrix $\mathbf{S}$, which is obtained from the collected data, contains information about the unknown noise covariance. We will have a discussion in Section 5.

The remaining problem is how to compare the two distributions above. It is essentially a property of distance and can be intuitively described via the platform of information geometry. In this paper, our major concern is to measure the difference between distributions (4) and (5) with methods of information geometry and also to give out a geometric criterion of how strongly the data contradict the hypothesis H_0 .

Remark 1. In [2, 4, 5], since the distribution of $\mathbf{Z}_p$ under H_1 is specified as

$$\mathcal{N}\left(\begin{pmatrix} \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_H \end{pmatrix}, \begin{pmatrix} \mathbf{M} & & \\ & \ddots & \\ & & \mathbf{M} \end{pmatrix}\right), \quad (6)$$

it is limited to discuss the case of determinate signals. Thus, many other cases cannot be described by the process of generalized likelihood ratio algorithm; however, the derived detectors still have perfect performances in the simulations of [2]. For our development, the methods of information geometry enlarge the scope of our discussion, since the covariance Σ in (4) has no constraints on the matrix structure.

Remark 2. It is common in literature works to use distance as a test statistic. Please refer to [14, 15]. In fact, the Mahalanobis distance can also be used to derive a perfect test, but it is not an intrinsic distance in geometry.

3. The Riemannian Geometry of the Multivariate Normal Model

We hereafter review some important notions in information geometry, in order to induce a measure of distance on the multivariate normal model.

3.1. Statistical Manifold. The statistical model describes a family of parameterized probability distributions:

$$\mathcal{G} = \{p(\mathbf{x}; \xi) \mid \xi = [\xi^1, \dots, \xi^n] \in \Xi\}, \quad (7)$$

where Ξ is a subset of $\mathbb{R}^n$. Under the condition of C^∞ diffeomorphism, the family $\mathcal{G}$ is considered as an n -dimensional statistical manifold [11]. Each probability distribution $p(\mathbf{x}; \xi)$ is represented by a point on the manifold $\mathcal{G}$ with the corresponding coordinate ξ .

3.2. The Fisher Metric and the Geodesic Distance. For a statistical manifold $\mathcal{G}$, the Fisher metric aims at defining an inner product on the tangent space $T_\xi(\mathcal{G})$. It is given by the Fisher information matrix $G(\xi)$ widely used in information science as

$$g_{ij}(\xi) = \mathbb{E}_\xi \left[\frac{\partial}{\partial \xi^i} \log p(\mathbf{x}; \xi) \frac{\partial}{\partial \xi^j} \log p(\mathbf{x}; \xi) \right], \quad (8)$$

where $\mathbb{E}_\xi$ denotes the expectation with respect to the distribution $p(\mathbf{x}; \xi)$.

With the Fisher metric, the length of a curve $\xi(t)$, $t \in [a, b]$, on the manifold is defined in [11] as

$$\|\xi(t)\| = \int_a^b \sqrt{\sum_{i,j} g_{ij} \dot{\xi}^i \dot{\xi}^j} dt, \quad (9)$$

where the dot denotes the differentiation with respect to the variable t and the indices indicate the corresponding components in the given coordinate parameterisations. Suppose a curve joining $p(\mathbf{x}; \xi_1)$ and $p(\mathbf{x}; \xi_2)$; the length between the endpoints depends on the choice of the curve. The geodesic distance between $p(\mathbf{x}; \xi_1)$ and $p(\mathbf{x}; \xi_2)$ is defined as the minimum length over all possible curves connecting the two points on the manifold $\mathcal{G}$ [17]; that is,

$$\begin{aligned} \mathcal{D}(p(\mathbf{x}; \xi_1), p(\mathbf{x}; \xi_2)) \\ \triangleq \min \{\|\xi(t)\| \mid \xi(a) = \xi_1, \xi(b) = \xi_2\}. \end{aligned} \quad (10)$$

We must emphasize that the geodesic distance is not only the intrinsic metric on the manifold of parametric probability distributions but also a genuine metric in the sense that it satisfies the nonnegative, symmetry properties and the triangle inequality. As a measure between probability distributions, the Kullback-Leibler divergence is also popular in information science, which is well-known as the relative entropy [18]:

$$\begin{aligned} \mathcal{D}_{kl}(p(\mathbf{x}; \xi_1), p(\mathbf{x}; \xi_2)) \\ = \mathbb{E}_{\xi_1} [\log p(\mathbf{x}; \xi_1) - \log p(\mathbf{x}; \xi_2)]. \end{aligned} \quad (11)$$

However, it is not a genuine metric, since it fails to be symmetric and satisfies the triangle inequality [19].

3.3. Riemannian Connection and Geodesics. Besides the Fisher metric, another fundamental tensor in Riemannian geometry is the Riemannian connection. The Riemannian connection is a symmetric affine connection that defines a linear one-to-one mapping between tangent spaces. For the Fisher metric, the Riemannian connection exists uniquely. The Riemannian connection coefficients are given in the form of the Christoffel symbols Γ_{ij}^k by the Fisher metric [11] as

$$\Gamma_{ij}^k = \sum_l \frac{1}{2} g^{kl} \left(\frac{\partial g_{il}}{\partial \xi^j} + \frac{\partial g_{jl}}{\partial \xi^i} - \frac{\partial g_{ij}}{\partial \xi^k} \right). \quad (12)$$

Geodesics are defined as the autoparallel curves on the manifold. Usually with the Christoffel symbols, a geodesic $\xi(t)$ is calculated by the equations [11]:

$$\ddot{\xi}^k(t) + \sum_{i,j} \dot{\xi}^i(t) \dot{\xi}^j(t) \Gamma_{ij}^k(\xi(t)) = 0, \quad k = 1, \dots, n. \quad (13)$$

It must be pointed out that, on Riemannian manifold, the geodesic distance between two points equals the length of the shortest geodesic segment between them [14].

3.4. Riemannian Geometry of the Multivariate Normal Manifold. The multivariate normal manifold consists of n -dimensional normal distributions parameterized by the mean and covariance matrix and is denoted by

$$\mathcal{G}^n = \{\mathcal{N}(\mu, \Sigma) \mid \mu \in \mathbb{R}^n, \Sigma \in \mathbb{S}_+^n\}. \quad (14)$$

As detailed in [14], the coordinate system of $\mathcal{G}^n$ is given by $\{\mu_i, i = 1, \dots, n; \sigma_{ij}, i \leq j\}$. Moreover, the tangent space $T_\xi(\mathcal{G}^n)$ coincides with the product space of n -dimensional vectors and $n \times n$ symmetric matrices, which is denoted by $\mathbb{R}^n \times \mathbb{S}^n$. According to (8), the Fisher metric is given in the form of inner product:

$$\begin{aligned} \langle \eta_1, \eta_2 \rangle_\xi &= \eta_1^T \Sigma^{-1} \eta_2, \\ \langle \eta_1, \mathbf{A} \rangle_\xi &= 0, \\ \langle \mathbf{A}, \mathbf{B} \rangle_\xi &= \frac{1}{2} \text{tr}(\Sigma^{-1} \mathbf{A} \Sigma^{-1} \mathbf{B}), \end{aligned} \quad (15)$$

where ξ denotes the point (μ, Σ) and $\eta_1, \eta_2 \in \mathbb{R}^n$, $\mathbf{A}, \mathbf{B} \in \mathbb{S}^n$. According to (12) and (13), the geodesics on the multivariate normal manifold $\mathcal{G}^n$ are calculated by solving the following equations:

$$\begin{aligned} \frac{d^2 \mu}{dt^2} - \left(\frac{d\Sigma}{dt} \Sigma^{-1} \frac{d\mu}{dt} \right) &= \mathbf{0}, \\ \frac{d^2 \Sigma}{dt^2} + \left(\frac{d\mu}{dt} \right) \left(\frac{d\mu}{dt} \right)^T - \left(\frac{d\Sigma}{dt} \right) \Sigma^{-1} \left(\frac{d\Sigma}{dt} \right) &= \mathbf{0}. \end{aligned} \quad (16)$$

The geodesic distance is calculated by integral (9) along the shortest geodesic. Specifically, when $n = 1$, the geodesic is explicitly given in [14, 20, 21]. The geodesic distance on the univariate normal manifold is given by

$$\begin{aligned} \mathcal{D}(\mathcal{N}(\mu_1, \sigma_1^2), \mathcal{N}(\mu_2, \sigma_2^2)) \\ = \sqrt{2} \cosh^{-1} \left(\frac{(\mu_1 - \mu_2)^2 + 2(\sigma_1^2 + \sigma_2^2)}{4\sigma_1\sigma_2} \right). \end{aligned} \quad (17)$$

Besides the aforementioned basic properties, the geodesic distance defined on the multivariate normal manifold owns the invariance property under the group transformation $\mathbb{R}^n \times \text{GL}_n(\mathbb{R})$ which acts as

$$\begin{aligned} g_{\mathbf{P}, \boldsymbol{\eta}}(\mu, \Sigma) &= (\mathbf{P}\mu + \boldsymbol{\eta}, \mathbf{P}\Sigma\mathbf{P}^T), \\ \forall (\mathbf{P}, \boldsymbol{\eta}) &\in \text{GL}_n(\mathbb{R}) \times \mathbb{R}^n, \end{aligned} \quad (18)$$

where $\text{GL}_n(\mathbb{R})$ denotes the group of n -dimensional regular matrices.

Unfortunately, in the general case $n > 1$, the analytical solutions to the geodesic equations (16) are complicated. The recent development is the achievement of deriving the explicit formula of the geodesic curves in [12, 13, 22], when the initial point and initial direction of the geodesic are given. Nevertheless, it remains unsolved how to obtain the geodesic curves given the endpoints and further to calculate the geodesic distance between them. However, efforts are made on the set of multivariate normal distributions with constant mean vectors, and another well-known geometry is defined.

3.5. The Submanifold with Fixed Mean Vector. Consider the submanifold $\mathcal{G}_{\mu_0}^n$ in $\mathcal{G}^n$ defined by

$$\mathcal{G}_{\mu_0}^n = \{ \mathcal{N}(\mu_0, \Sigma) \mid \Sigma \in \mathbb{S}_+^n \}, \quad (19)$$

where $\mu_0 \in \mathbb{R}^n$ is a constant vector. The structure of $\mathcal{G}_{\mu_0}^n$ has been studied in [14, 15, 23]. The geodesics have been explicitly given and the geodesic distance has been computed. In particular, it is shown in [14] that $\mathcal{G}_{\mu_0}^n$ is a totally geodesic submanifold of $\mathcal{G}^n$. That means, on the manifold $\mathcal{G}^n$, the geodesic joining two points of $\mathcal{G}_{\mu_0}^n$ will line entirely in $\mathcal{G}_{\mu_0}^n$.

Thus, if given $\Sigma_1, \Sigma_2 \in \mathbb{S}_+^n$, the geodesic from $\mathcal{N}(\mu_0, \Sigma_1)$ to $\mathcal{N}(\mu_0, \Sigma_2)$ is given in [17], by

$$\begin{aligned} \Sigma(t) &= \Sigma_1^{1/2} \exp \{ t \log(\Sigma_1^{-1/2} \Sigma_2 \Sigma_1^{-1/2}) \} \Sigma_1^{1/2}, \\ t &\in [0, 1], \end{aligned} \quad (20)$$

and the initial vector is in the direction of

$$\dot{\Sigma}(0) = \Sigma_1^{1/2} \log(\Sigma_1^{-1/2} \Sigma_2 \Sigma_1^{-1/2}) \Sigma_1^{1/2}. \quad (21)$$

For any $\mathbf{A} \in \mathbb{S}^n$, define $\|\mathbf{A}\|_{\Sigma}^2 = \langle \mathbf{A}, \mathbf{A} \rangle_{\Sigma} = \text{tr}[(\Sigma^{-1} \mathbf{A})^2]$. Then the geodesic distance is

$$\begin{aligned} \mathcal{D} &= \|\dot{\Sigma}(0)\|_{\Sigma_1} = \sqrt{\text{tr}(\log^2(\Sigma_1^{-1/2} \Sigma_2 \Sigma_1^{-1/2}))} \\ &= \sqrt{\frac{1}{2} \sum_{i=1}^n \ln^2 \lambda_i}, \end{aligned} \quad (22)$$

where $\lambda_i, i = 1, \dots, n$, are the eigenvalues of $\Sigma_1^{-1} \Sigma_2$.

4. The Design of Geometric Distance Detector

As described in Section 2, it is possible to apply the results of geodesic distance to compare the distributions of $\mathbf{Z}_P$ and $\mathbf{N}_P$, further to construct test statistics for problem (3).

Let $n = H \times p$. Note $\Sigma \in \mathbb{S}_+^n$ in (4); thus the possible distributions of $\mathbf{Z}_P$ constitute a submanifold $\mathcal{G}_{\mathbf{Z}_P}^n$ in $\mathcal{G}^n$. We denote the distribution in (5) as the point ξ_1 in the manifold $\mathcal{G}^n$. Since in general case $n > 1$, it has been mentioned before that the direct acquisition of the geodesic distance between multivariate normal distributions with different mean vectors is hard. Here we introduce the Siegel distance given by Calvo and Oller in [15], which attempts to calculate the geodesic distance between the images of the original distributions based on embedding into $\mathcal{G}_0^{n+1}$:

$$\begin{aligned} f: \mathcal{G}^n &\longrightarrow \mathcal{G}_0^{n+1} \\ (\mu, \Sigma) &\longmapsto \begin{pmatrix} \Sigma + \mu\mu^T & \mu \\ \mu^T & 1 \end{pmatrix}. \end{aligned} \quad (23)$$

Then the Siegel distance is defined in [15] as

$$\mathcal{D}_s(\xi_1, \xi_2) \stackrel{\Delta}{=} \mathcal{D}(f(\xi_1), f(\xi_2)). \quad (24)$$

Similarly, the Siegel distance between ξ_1 and the submanifold $\mathcal{G}_{\mathbf{Z}_P}^n$ is defined as

$$\begin{aligned} \mathcal{D}_s(\xi_1, \mathcal{G}_{\mathbf{Z}_P}^n) \\ &\stackrel{\Delta}{=} \min \{ \mathcal{D}_s(\xi_1, \xi_2) \mid \xi_2 \in \mathcal{G}_{\mathbf{Z}_P}^n \} \\ &= \min \{ \mathcal{D}_s(\xi'_1, \xi'_2) \mid \xi'_2 \in \mathcal{G}_0^n \} \\ &= \min \{ \mathcal{D}(f(\xi'_1), f(\xi'_2)) \mid \xi'_2 \in \mathcal{G}_0^n \}, \end{aligned} \quad (25)$$

where ξ'_1 denotes the normal distribution $\mathcal{N}(-\mathbf{Z}_P^0, \mathbf{A})$ and $\mathcal{G}_0^n = \{ \mathcal{N}(\mathbf{0}, \Sigma) \mid \Sigma \in \mathbb{S}_+^n \}$. The second equality is due to the invariance properties of the Siegel distance, which is the same as the geodesic distance. The last equality follows from the definition of the Siegel distance in (24).

It has been proven that $\mathcal{G}^n$ is isometric to $f(\mathcal{G}^n)$. However, $f(\mathcal{G}^n)$ is not a geodesic submanifold in $\mathcal{G}_0^{n+1}$. That is to say, the geodesic curve joining the images $f(\xi_1)$ and $f(\xi_2)$ in $\mathcal{G}_0^{n+1}$ may contain points outside $f(\mathcal{G}^n)$. Thus, the Siegel distance provides a lower bound for the geodesic distance between multivariate normal distributions, which has also been verified as a distance measure on the multivariate normal manifold.

Given $\xi'_2 \in \mathcal{G}_0^n$, it must be mentioned that $f(\xi'_1)$ and $f(\xi'_2)$ denote two points

$$\begin{aligned} f(\xi'_1) &: \mathcal{N}\left(\mathbf{0}_{n+1}, \begin{pmatrix} \Lambda + \mathbf{Z}_P^0 (\mathbf{Z}_P^0)^T & -\mathbf{Z}_P^0 \\ -(\mathbf{Z}_P^0)^T & 1 \end{pmatrix}\right), \\ f(\xi'_2) &: \mathcal{N}\left(\mathbf{0}_{n+1}, \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}\right) \end{aligned} \quad (26)$$

in $\mathcal{G}_0^{n+1}$. By (20), the geodesic from $f(\xi'_2)$ to $f(\xi'_1)$ is given as

$$\Sigma(t) = \begin{pmatrix} \Sigma^{1/2} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix} \exp\{t\mathbf{Q}\} \begin{pmatrix} \Sigma^{1/2} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}, \quad t \in [0, 1], \quad (27)$$

where

$$\mathbf{Q} = \log \begin{pmatrix} \Sigma^{-1/2} (\Lambda + \mathbf{Z}_P^0 (\mathbf{Z}_P^0)^T) \Sigma^{-1/2} & -\Sigma^{-1/2} \mathbf{Z}_P^0 \\ -(\mathbf{Z}_P^0)^T \Sigma^{-1/2} & 1 \end{pmatrix}. \quad (28)$$

In expression (25), the initial direction at $f(\xi'_2)$ is

$$\dot{\Sigma}(0) = \begin{pmatrix} \Sigma^{1/2} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix} \mathbf{Q} \begin{pmatrix} \Sigma^{1/2} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}. \quad (29)$$

The solution to (25) is easy to be found according to the following lemma.

Lemma 3 (Pythagorean theorem). *Let ξ be a point in $\mathcal{G}$ and let $\mathcal{G}_0$ be a submanifold of $\mathcal{G}$. A necessary and sufficient condition for a point $\xi' \in \mathcal{G}_0$ to be a stationary point of the function $\mathcal{D}(\xi, \cdot) : \mathcal{G}_0 \rightarrow \mathbb{R}$ is for the geodesic connecting ξ and ξ' to be orthogonal to $\mathcal{G}_0$ at ξ' .*

The proof of the Pythagorean theorem is common in differential geometry. Then the lemma follows.

Theorem 4. *The point ξ'_2 in $\mathcal{G}_0^n$ that achieves the minimum of (25) must satisfy*

$$\mathbf{Q} = \begin{pmatrix} \mathbf{0} & \mathbf{b} \\ \mathbf{b}^T & a \end{pmatrix}, \quad (30)$$

for some $a \in \mathbb{R}$ and $\mathbf{b} \in \mathbb{R}^n$.

Proof. By the Pythagorean theorem, we have

$$\begin{aligned} 0 &= \langle \dot{\Sigma}(0), df(\xi'_2) \rangle_{f(\xi'_2)} \\ &= \text{tr} \left(\begin{pmatrix} \Sigma^{-1} & \\ & 1 \end{pmatrix} \dot{\Sigma}(0) \begin{pmatrix} \Sigma^{-1} & \\ & 1 \end{pmatrix} \begin{pmatrix} d\Sigma \\ 0 \end{pmatrix} \right) \\ &= \text{tr} \left(\begin{pmatrix} \Sigma^{-1} d\Sigma \\ 0 \end{pmatrix} \mathbf{Q} \right). \end{aligned} \quad (31)$$

This theorem follows due to the arbitrariness of $d\Sigma$. $\square$

The matrix $\mathbf{Q}$ given in (30) has at most two nonzero eigenvalues λ_1, λ_2 , which must satisfy

$$\begin{aligned} \lambda_1 + \lambda_2 &= \text{tr}(\mathbf{Q}) = a, \\ \lambda_1^2 + \lambda_2^2 &= \text{tr}(\mathbf{Q}^2) = 2\mathbf{b}^T \mathbf{b} + a^2. \end{aligned} \quad (32)$$

Note that the eigenvalues λ_1 and λ_2 are opposite sign while $\mathbf{b} \neq \mathbf{0}$.

Theorem 5. *If λ_1 and λ_2 are a solution of the following equations,*

$$\begin{aligned} \frac{\lambda_1 - \lambda_2}{\lambda_1 \lambda_2} \frac{\lambda_2 e^{\lambda_1} - \lambda_1 e^{\lambda_2}}{(e^{\lambda_1} - e^{\lambda_2})^2} &= \frac{1}{\mathbf{Z}_P^0 \Lambda^{-1} \mathbf{Z}_P^0} + 1, \\ \frac{\lambda_1 e^{\lambda_1} - \lambda_2 e^{\lambda_2}}{\lambda_1 - \lambda_2} &= 1, \end{aligned} \quad (33)$$

then the minimum in (25) is achieved by

$$\begin{aligned} \Sigma &= \Lambda + \left(1 - \frac{\lambda_1 - \lambda_2}{\lambda_1 \lambda_2} \frac{\lambda_2 e^{\lambda_1} - \lambda_1 e^{\lambda_2} + (\lambda_1 - \lambda_2)}{(e^{\lambda_1} - e^{\lambda_2})^2} \right) \\ &\quad \cdot \mathbf{Z}_P^0 (\mathbf{Z}_P^0)^T, \end{aligned} \quad (34)$$

and the Siegel distance between ξ_1 and $\mathcal{G}_0^n$ is

$$\mathcal{D}_s(\xi_1, \mathcal{G}_0^n) = \sqrt{\lambda_1^2 + \lambda_2^2}. \quad (35)$$

Proof. For the eigenvalues λ_1, λ_2 , the corresponding eigenvectors of $\mathbf{Q}$ in (30) can be given as $\mathbf{u}_i = [\mathbf{b}^T, \lambda_i]^T$, $i = 1, 2$. Taking $(n+1) \times (n+1)$ orthogonal matrix

$$\mathbf{U} = \left[\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \mathbf{U}_0 \right], \quad (36)$$

we have

$$\mathbf{U}_0 \mathbf{U}_0^T = \mathbf{I}_{n+1} - \sum_{i=1}^2 \frac{1}{\|\mathbf{u}_i\|^2} \mathbf{u}_i \mathbf{u}_i^T. \quad (37)$$

Then from $\|\mathbf{u}_i\|^2 = \lambda_i^2 + \mathbf{b}^T \mathbf{b} = \lambda_i^2 - \lambda_1 \lambda_2$ and

$$\mathbf{Q} = \begin{pmatrix} \mathbf{0} & \mathbf{b} \\ \mathbf{b}^T & a \end{pmatrix} = \mathbf{U} \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \mathbf{0}_{n-1} \end{pmatrix} \mathbf{U}^T, \quad (38)$$

we have

$$\begin{aligned} \exp(\mathbf{Q}) &= \mathbf{U} \begin{pmatrix} e^{\lambda_1} & & \\ & e^{\lambda_2} & \\ & & \mathbf{I}_{n-1} \end{pmatrix} \mathbf{U}^T \\ &= \sum_{i=1}^2 \frac{e^{\lambda_i}}{\|\mathbf{u}_i\|^2} \mathbf{u}_i \mathbf{u}_i^T + \mathbf{U}_0 \mathbf{U}_0^T \\ &= \mathbf{I}_{n+1} + \sum_{i=1}^2 \frac{e^{\lambda_i} - 1}{\|\mathbf{u}_i\|^2} \mathbf{u}_i \mathbf{u}_i^T \\ &= \mathbf{I}_{n+1} + \begin{pmatrix} g_1(\lambda_1, \lambda_2) \mathbf{b} \mathbf{b}^T & g_2(\lambda_1, \lambda_2) \mathbf{b} \\ g_2(\lambda_1, \lambda_2) \mathbf{b}^T & g_3(\lambda_1, \lambda_2) \end{pmatrix}, \end{aligned} \quad (39)$$

where

$$\begin{aligned} g_1(\lambda_1, \lambda_2) &= \sum_{i=1}^2 \frac{e^{\lambda_i} - 1}{\|\mathbf{u}_i\|^2} = \frac{\lambda_2 e^{\lambda_1} - \lambda_1 e^{\lambda_2}}{\lambda_1 \lambda_2 (\lambda_1 - \lambda_2)} + \frac{1}{\lambda_1 \lambda_2}, \\ g_2(\lambda_1, \lambda_2) &= \sum_{i=1}^2 \frac{e^{\lambda_i} - 1}{\|\mathbf{u}_i\|^2} \lambda_i = \frac{e^{\lambda_1} - e^{\lambda_2}}{\lambda_1 - \lambda_2}, \\ g_3(\lambda_1, \lambda_2) &= \sum_{i=1}^2 \frac{e^{\lambda_i} - 1}{\|\mathbf{u}_i\|^2} \lambda_i^2 = \frac{\lambda_1 e^{\lambda_1} - \lambda_2 e^{\lambda_2}}{\lambda_1 - \lambda_2} - 1. \end{aligned} \quad (40)$$

From (28) and (39), we can get

$$\Sigma^{-1/2} (\Lambda + \mathbf{Z}_p^0 (\mathbf{Z}_p^0)^T) \Sigma^{-1/2} = \mathbf{I}_n + g_1(\lambda_1, \lambda_2) \mathbf{b} \mathbf{b}^T, \quad (41)$$

$$-\Sigma^{-1/2} \mathbf{Z}_p^0 = g_2(\lambda_1, \lambda_2) \mathbf{b}, \quad (42)$$

$$0 = g_3(\lambda_1, \lambda_2). \quad (43)$$

Therefore,

$$\Sigma = \Lambda + \left(1 - \frac{g_1(\lambda_1, \lambda_2)}{g_2^2(\lambda_1, \lambda_2)}\right) \mathbf{Z}_p^0 (\mathbf{Z}_p^0)^T, \quad (44)$$

$$\mathbf{b} = -\frac{1}{g_2(\lambda_1, \lambda_2)} \Sigma^{-1/2} \mathbf{Z}_p^0. \quad (45)$$

From (41)-(42), we have

$$\frac{1}{(\mathbf{Z}_p^0)^T \Lambda^{-1} \mathbf{Z}_p^0} = -1 + \frac{-(1/\lambda_1 \lambda_2) + g_1(\lambda_1, \lambda_2)}{g_2^2(\lambda_1, \lambda_2)}. \quad (46)$$

Equations (43) and (46) imply (33). Since $\mathcal{G}_0^{n+1}$ is a complete manifold, there exists a solution to (33) of the two variables λ_1 and λ_2 .

Equation (34) follows from (44). By (22), the Siegel distance between ξ_1 and $\mathcal{G}_{\mathbf{Z}_p^0}^n$ becomes

$$\begin{aligned} &\text{Tr} \left(\log^2 \left(\begin{pmatrix} \Sigma^{-1/2} & \\ & 1 \end{pmatrix} \begin{pmatrix} \Lambda + \mathbf{Z}_p \mathbf{Z}_p^T & -\mathbf{Z}_p \\ -\mathbf{Z}_p^T & 1 \end{pmatrix} \begin{pmatrix} \Sigma^{-1/2} & \\ & 1 \end{pmatrix} \right) \right) \\ &= \text{tr} \left(\log^2 \begin{pmatrix} \Sigma^{-1/2} (\Lambda + \mathbf{Z}_p^0 (\mathbf{Z}_p^0)^T) \Sigma^{-1/2} & -\Sigma^{-1/2} \mathbf{Z}_p^0 \\ -(\mathbf{Z}_p^0)^T \Sigma^{-1/2} & 1 \end{pmatrix} \right) \\ &= \text{tr}(\mathbf{Q}^2) = \lambda_1^2 + \lambda_2^2. \end{aligned} \quad (47)$$

That is (35). This theorem thus holds. $\square$

The equations in (33) are nonlinear which can be solved numerically, and then we can calculate the Siegel distance $\mathcal{D}_s$ by (35). Let $T^2 = (\mathbf{Z}_p^0)^T \Lambda^{-1} \mathbf{Z}_p^0$. From Theorem 5, it is easy to find the Siegel distance $\mathcal{D}_s$ closely related to T^2 . As stated in Section 1, with the Siegel distance $\mathcal{D}_s$, the detection problem (3) can be carried out with the critical region defined as (1). Note that the matrix $\mathbf{S}$ in Λ is undeterminate, which does not affect the derivation of $\mathcal{D}_s$. Therefore, a set of geometric distance detectors are given via different choices of the matrix $\mathbf{S}$.

5. The Choice of Matrix $\mathbf{S}$

Let $\mathbf{z}_i^0$ be the realization of random vector $\mathbf{z}_i$ for each $i = 1, \dots, H + K$. Note that

$$\begin{aligned} T^2 &= (\mathbf{Z}_p^0)^T \Lambda^{-1} \mathbf{Z}_p^0 \\ &= \begin{pmatrix} \mathbf{z}_1^0 \\ \vdots \\ \mathbf{z}_H^0 \end{pmatrix}^T \begin{pmatrix} \mathbf{S}^{-1} & & \\ & \ddots & \\ & & \mathbf{S}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{z}_1^0 \\ \vdots \\ \mathbf{z}_H^0 \end{pmatrix} \\ &= \sum_{i=1}^H (\mathbf{z}_i^0)^T \mathbf{S}^{-1} \mathbf{z}_i^0. \end{aligned} \quad (48)$$

Figure 1 illustrates how the Siegel distance varies with the logarithm of T^2 increasing. The curve is obvious to demonstrate a monotonic increase. Thus the critical region in (1) is rewritten as

$$W_\epsilon = \{z_t \mid \mathcal{D} = h(T^2) > \gamma_\epsilon\} = \{z_t \mid T^2 > h^{-1}(\gamma_\epsilon)\}, \quad (49)$$

where h is an increasing function of T^2 .

If the underlying covariance $\mathbf{M}$ is known, it is natural to take $\mathbf{S}$ as $\mathbf{M}$ directly. In this case, the optimal matched filter is derived. Otherwise, efficient information about $\mathbf{M}$ is given by the maximum likelihood estimate based on the received data. When only with the secondary data, $\mathbf{S}$ is replaced by

$$\mathbf{S}_1 = \frac{1}{K} \sum_{i=H+1}^{H+K} \mathbf{z}_i^0 (\mathbf{z}_i^0)^T, \quad (50)$$

FIGURE 1: The Siegel distance $\mathcal{D}$ versus the logarithm of T^2 .

thus the Wald test belongs to the set of geometric distance detectors. In addition, substituting the maximum likelihood estimate based on the whole set of data

$$\mathbf{S}_2 = \frac{1}{H+K} \sum_{i=1}^{H+K} \mathbf{z}_i^0 (\mathbf{z}_i^0)^T \quad (51)$$

in place of $\mathbf{S}$, we can find it is the same as the Rao test.

In summary, it is verified that the referred three classical tests are members of the set of geometric distance detectors designed based on the Siegel distance.

6. Interpretations

In this section, some figures are presented, in order to give intuitive interpretations about how the geometric distance detectors work. As special cases, the optimal matched filter, the Rao test, and Wald test are described in the form of geometric distance detectors.

Figure 2 illustrates the definition of the geodesic distance from some point on the univariate normal manifold to the submanifold with the same mean. A group of geodesics from $\mathcal{N}(0, 1)$ are shown. Gradient colors along the geodesics represent different geodesic distances given by (17) from $\mathcal{N}(0, 1)$ to the current point. The geodesic distance from $\mathcal{N}(0, 1)$ to $\mathcal{S}_{0.55}^1$ is given by the integral along the geodesic marked red. As a geometric distance detector, a threshold is required for the hypothesis test problem. H_0 is not rejected if the Siegel distance between the distribution of noise and the submanifold determined by the measurements is less than the threshold. The critical region in (1) is obtained by a projection of submanifolds that are outside of the threshold on the observation space. The same principle applies to the Siegel distance on the multivariate normal manifold. As without visualization features, the multidimensional cases are not presented here.

Figure 3 shows a simulation with $H = 1$, $p = 2$, since the examples with a higher dimension have similar

FIGURE 2: Illustration of the geodesic distance from $\mathcal{N}(0, 1)$ on the univariate normal manifold to the submanifold $\mathcal{S}_{0.55}^1$. The blue straight line shows the submanifold with the fixed mean 0.55. The geodesic curves from the initial point $\mathcal{N}(0, 1)$ with different initial directions are displayed with gradient colored curves, and the one orthogonal to the submanifold $\mathcal{S}_{0.55}^1$ is marked red.

behaviors but no visualization features. The distributions of noise are displayed at the origin, with the corresponding noise covariance matrix: the underlying covariance $\mathbf{M}$ and the sample covariance $\mathbf{S}_1$ or $\mathbf{S}_2$. All curves were drawn as the projections on the measurement space. It is obvious that the part outside a contour curve represents the critical region of the geometry detector with the corresponding alarm probability listed in the legend. As the intrinsic merit of geometric measure, the magnitude of the Siegel distance reflects how strongly the collected data contradict the null hypothesis H_0 . In Figure 3(d), contour curves with the same alarm probability $\alpha = 10^{-4}$ are presented. We can observe that the contour curves of the Siegel distance based on $\mathbf{S}_1$ and $\mathbf{S}_2$ appear to overlap. This is due to the equivalency of the Rao test and Wald test when $H = 1$ [24]. In fact, if we denote the eigenvalues of the $p \times p$ matrix $\mathbf{\Omega} = \sum_{i=1}^H \mathbf{z}_i \mathbf{S}_1^{-1} \mathbf{z}_i^T$ as $\omega_1, \dots, \omega_p$, then, according to [2, 3], the Rao test and Wald test are, respectively, rewritten as

$$T_{\text{Rao}}^2 = \sum_{i=1}^p \frac{(H+K) \omega_i}{K + \omega_i}, \quad (52)$$

$$T_{\text{Wald}}^2 = \sum_{i=1}^p \omega_i.$$

When $H = 1$, $\mathbf{\Omega} = \mathbf{z}_1 \mathbf{S}_1^{-1} \mathbf{z}_1^T$ has only one nonzero eigenvalue. Thus, both the Rao test and the Wald test coincide to test for the nonzero eigenvalue. However, they yield different Siegel distances because of the distinct covariance matrices which are neglected in projection mapping. The acceptance regions of both the Rao test and the Wald test can be improved by increasing the number of secondary data, so as to approach that of the geometric distance detector based on the underlying noise covariance.

As a general rule, the Rao test and Wald test have different acceptance regions. Figure 4 displays another simulation with

FIGURE 3: Contour curves of the Siegel distance based on (a) the underlying covariance matrix $\mathbf{M}$; (b) the sample covariance matrix $\mathbf{S}_1$ or (c) $\mathbf{S}_2$. The values marked on the contour curves denote different Siegel distances from the origin to the submanifold specified by the current point. α_i s listed in the legends denote the corresponding alarm probabilities of the geometric distance detectors; (d) contour curves of the Siegel distances based on different choice of the noise covariance matrix, with the same alarm probability $\alpha = 10^{-4}$. All curves were drawn as the projections on the measurement space.

$H = 2$, $p = 2$. For flat visualizations, the contour curves of the Siegel distance are projected on the space of Ω .

7. Conclusion

In this paper, an information geometric viewpoint on how to deal with the detection problem of range distributed targets embedded in Gaussian noise with unknown covariance is put forward. More precisely, we have derived a set of geometric distance detectors, of which the optimal matched filter, the

Rao test, and Wald test are members. This establishes a link between the information geometry and the hypothesis testing. As a future research, other choices of $\mathbf{S}$ can be tested, and it might also be of interest to find one with a better performance among the set of geometric distance detectors.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

FIGURE 4: The contour curves of the Siegel distance based on the sample covariance matrices S_1 and S_2 , with the same probability $\alpha = 10^{-4}$. All curves are drawn as the projections on the eigenvalues space of Ω .

Acknowledgments

This work was supported in part by the National Key Basic Research Program of China under Grant 2013CB329405, the National Natural Science Foundation of China under Grant 61374027, and the Specialized Research Fund for the Doctoral Program of Higher Education under Grant 20130181110042.

References

- [1] E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, no. 2, pp. 115–127, 1986.
- [2] E. Conte, A. De Maio, and C. Galdi, "CFAR detection of multi-dimensional signals: an invariant approach," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 142–151, 2003.
- [3] W. Liu, W. Xie, and Y. Wang, "Rao and wald tests for distributed targets detection with unknown signal steering," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1086–1089, 2013.
- [4] K. Gerlach and M. J. Steiner, "Adaptive detection of range distributed targets," *IEEE Transactions on Signal Processing*, vol. 47, no. 7, pp. 1844–1851, 1999.
- [5] E. Conte, A. De Maio, and G. Ricci, "GLRT-based adaptive detection algorithms for range-spread targets," *IEEE Transactions on Signal Processing*, vol. 49, no. 7, pp. 1336–1348, 2001.
- [6] O. Besson, S. Kraut, and L. L. Scharf, "Detection of an unknown rank-one component in white noise," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2835–2839, 2006.
- [7] H. Wang and L. Cai, "On adaptive multiband signal detection with GLR algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 27, no. 2, pp. 225–233, 1991.
- [8] H.-R. Park, J. Li, and H. Wang, "Polarization-space-time domain generalized likelihood ratio detection of radar targets," *Signal Processing*, vol. 41, no. 2, pp. 153–164, 1995.
- [9] C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters," *Bulletin of the Calcutta Mathematical Society*, vol. 37, no. 3, pp. 81–91, 1945.
- [10] S.-I. Amari, "Geometrical theory of asymptotic ancillarity and conditional inference," *Biometrika*, vol. 69, no. 1, pp. 1–17, 1982.
- [11] S.-I. Amari, *Differential Geometrical Methods in Statistics*, Springer, Berlin, Germany, 1985.
- [12] P. S. Eriksen, "Geodesics connected with the Fisher metric on the multivariate normal manifold," in *Proceedings of the GST Workshop: Geometrization of Statistical Theory*, pp. 225–229, University of Lancaster, Lancaster, UK, October 1987.
- [13] M. Calvo and J. M. Oller, "An explicit solution of information geodesic equations for the multivariate normal model," *Statistics & Decisions*, vol. 9, no. 1-2, pp. 119–138, 1991.
- [14] L. T. Skovgaard, "A Riemannian geometry of the multivariate normal model," *Scandinavian Journal of Statistics. Theory and Applications*, vol. 11, no. 4, pp. 211–223, 1984.
- [15] M. Calvo and J. M. Oller, "A distance between multivariate normal distributions based in an embedding into the Siegel group," *Journal of Multivariate Analysis*, vol. 35, no. 2, pp. 223–242, 1990.
- [16] M. Calvo, A. Villarroja, and J. M. Oller, "A biplot method for multivariate normal populations with unequal covariance matrices," *Test*, vol. 11, no. 1, pp. 143–165, 2002.
- [17] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras, "Statistics on the manifold of multivariate normal distributions: theory and application to diffusion tensor MRI processing," *Journal of Mathematical Imaging and Vision*, vol. 25, no. 3, pp. 423–444, 2006.
- [18] K. M. Carter, R. Raich, W. G. Finn, and A. O. Hero III, "FINE: fisher information nonparametric embedding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 2093–2098, 2009.
- [19] Y. Cheng, X. Wang, M. Morelande, and B. Moran, "Information geometry of target tracking sensor networks," *Information Fusion*, vol. 14, no. 3, pp. 311–326, 2013.
- [20] L. Madsen, *The geometry of statistical models [Ph.D. thesis]*, University of Copenhagen, Copenhagen, Denmark, 1978.
- [21] S.-I. Amari, "Theory of information space: a differential-geometrical foundation of statistics," Post RAAG Reports 106, 1980.
- [22] T. Imai, A. Takaesu, and M. Wakayama, "Remarks on geodesics for multivariate normal models," *Journal of Math-for-Industry*, vol. 3, no. 6, pp. 125–130, 2011.
- [23] J. Burbea, "Informative geometry of probability spaces," *Expositiones Mathematicae*, vol. 4, no. 4, pp. 347–378, 1986.
- [24] W. Liu, W. Xie, J. Liu, and Y. Wang, "Adaptive double subspace signal detection in Gaussian background—part I: homogeneous environments," *IEEE Transactions on Signal Processing*, vol. 62, no. 9, pp. 2345–2357, 2014.