MIG Median Detectors with Manifold Filter

Xiaoqiang Hua\textsuperscript{1} and Linyu Peng\textsuperscript{2}\textsuperscript{*}

1. College of Meteorology and Oceanography, National University of Defense Technology, Changsha 410073, China
2. Department of Mechanical Engineering, Keio University, Hiyoshi 3-14-1, Yokohama 223-8522, Japan

Abstract

In this paper, we propose a class of median-based matrix information geometry (MIG) detectors with a manifold filter and apply them to signal detection in nonhomogeneous environments. As customary, the sample data is assumed to be modeled as Hermitian positive-definite (HPD) matrices, and the geometric median of a set of HPD matrices is interpreted as an estimate of the clutter covariance matrix (CCM). Then, the problem of signal detection can be reformulated as discriminating two points on the manifold of HPD matrices, one of which is the HPD matrix in the cell under test while the other represents the CCM. By manifold filter, we map a set of HPD matrices to another set of HPD matrices by weighting them, that consequently improves the discriminative power by reducing the intra-class distances while increasing the inter-class distances. Three MIG median detectors are designed by resorting to three geometric measures on the matrix manifold, and the corresponding geometric medians are shown to be robust to outliers. Numerical simulations show the advantage of the proposed MIG median detectors in comparison with their state-of-the-art counterparts as well as the conventional detectors in nonhomogeneous environments.

Keywords: Matrix information geometry (MIG) detector, geometric median, manifold filter, signal detection, clutter covariance matrix

1 Introduction

Signal detection in nonhomogeneous environments is an important and fundamental problem in the studies of radar, sonar and communications. Typically, to achieve satisfactory detection performance, the unknown clutter covariance matrix (CCM) should be estimated accurately. The problem of CCM estimation has been extensively studied during the latest decades \cite{1,3,4,25,27,39,46}. One major limitation of the classical CCM estimators is that the accuracy depends heavily on the number of homogeneous secondary data, which is independent and identically distributed and is collected from the range cells adjacent to the range cell under test. In practical applications, the number of homogeneous secondary data is often very limited in nonhomogeneous environments. In addition, the secondary data is often contaminated by outliers due to the environmental nonhomogeneity caused by the interfere signal or the variation of clutter power. Both factors make the classical signal detection techniques, such as Kelly’s detector \cite{36}, adaptive matched filter (AMF) \cite{21,47} and adaptive coherence estimator (ACE) \cite{37} derived by exploiting the CCM estimator suffer from a severe performance degradation.

To address these drawbacks, lots of attempts are made to improve the estimation accuracy of CCM. For instance, De Maio et al. \cite{23} exploited a priori information on the covariance structure to design an adaptive detection of point-like target in interference environments. Specifically, a symmetrically structured power spectral density is assumed to agree with the clutter properties. Based on this assumption, three adaptive decision schemes leveraging on the symmetric power spectral density structure for the interference are proposed, whose performance advantage was confirmed by simulation.

\textsuperscript{*}Corresponding author. E-mail: l.peng@mech.keio.ac.jp.
as well as experimental analysis using real radar data. In [4], a special structure that the covariance matrix is described as the sum of an unknown positive semi-definite matrix and a matrix proportional to the identity one, and a condition number upper-bound constraint are enforced on the estimated matrix. The constrained structure both accounting for the clutter and the white interference can achieve signal-to-interference-plus-noise ratio (SINR) improvements over their conventional counterparts, especially in the case of limited data. Similar studies concerning the priori information on the structure of the estimated covariance matrix can also be found in [5,6,20,35]. Other possible strategies to improve the performance of CCM estimation are mainly focused on resorting to the Bayesian covariance matrix estimation method [12,13], using knowledge-based covariance models [7,22,42,43,49], or considering shrinkage estimation methods [17,18]. A fatal drawback on these methods is that the performance gain relies on the known or assumed statistical characteristics of clutter environments, which is often difficult to be acquired in advance in practical applications.

Recently, a new signal detection scheme designed in the framework of matrix information geometry (MIG) was proposed by Lapuyade-Lahorgue and Barbaresco [38] and developed by Hua et al. [31], etc. This technique is often referred to as matrix constant false alarm rate detector or MIG detector. MIG studies intrinsic properties of matrix models, and many information processing problems can be transformed into discriminational or optimization problems on matrix spaces equipped with a metric. Based on the MIG theory, an MIG detector designed by using the Riemannian metric or affine invariant Riemannian metric (AIRM) has been successful applied to target detection in coastal X-band and HF surface wave radars [2,8,9], burg estimation of radar scatter matrix [24], the analysis of statistical characterization [10] and the monitoring of wake vortex turbulences [11,40]. The detection framework has been extended by replacing the AIRM distance with other different geometric measures. In [32], the authors developed the MIG detector based on the symmetrized Kullback–Leibler divergence (SKLD) and total Kullback–Leibler divergence. In [30,33], the total Bregman and total Jensen–Bregman divergences were used to design effective MIG detectors. It is noticed that MIG detectors designed by using different measures have different detection performances. As MIG detectors do not rely on the statistical characteristics of clutter environment and take into account the underlying geometry of matrix manifolds, they have the potential to achieve a performance improvement over the conventional detection techniques in nonhomogeneous environments.

It is well known that a filter can often remove unwanted information from a signal, which can be the noise or the clutter in signal detection. As a consequence, discrimination of the signal and the clutter/noise will be enhanced and hence the signal can be detected more easily. To improve the discriminative power on matrix manifolds, we propose a class of MIG median detectors with a manifold filter, that will improve detection performances by reducing the intra-class distances while increasing the inter-class distances. The main contributions are described as follows:

1. We define a map (see Eq. (4.5)) on the manifold of Hermitian positive-definite (HPD) matrices by a manifold filter. Consequently, each HPD matrix is mapped to a weighted average of several surrounding HPD matrices, which is again Hermitian positive-definite.

2. We reformulate the problem of signal detection into discriminating two points on the HPD matrix manifold, and then design three MIG median detectors by resorting to the Log-Euclidean metric (LEM), SKLD and the Jensen–Bregman LogDet (JBLD) divergence, respectively. In particular, the median matrices related to the three metrics for a set of HPD matrices are computed by using the fixed-point algorithm and are used as the CCM estimate. Moreover, the anisotropy indices associated with the three metrics are defined to analyze the difference in discriminative power on the HPD manifold and the other manifold.

3. Simulation results verify that the geometric median matrix is robust and reliable, and the proposed MIG median detectors with manifold filter can achieve performance improvements over state-of-the-art counterparts as well as the AMF in nonhomogeneous environments.

The remainder of this paper is organized as follows. Relevant fundamental theories of MIG are introduced in Section 2. Section 3 is devoted to formulation of detection problems on matrix manifolds. Derivation of the detection architecture is presented in Section 4 with numerical simulations given in Section 5. Finally, we conclude in Section 6.
Notations: In the sequel, scalars, vectors and matrices are denoted by lowercase, boldface lowercase and boldface uppercase letters, respectively. The symbols \((\cdot)^*, (\cdot)^T\) and \((\cdot)^H\) stand for the conjugate, the transpose, and the conjugate transpose, respectively. The operators \(\det(\cdot)\) and \(\text{tr}(\cdot)\) denote the determinate and trace of a matrix, respectively. The identity \(n \times n\) matrix is denoted by \(I_n\) or \(I\) if no confusion would be caused. We use \(\mathbb{C}^n\) and \(\mathbb{C}^{n \times n}\) to denote the set of \(n\)-dimensional complex vectors and \(n \times n\) complex matrices, respectively.

2 Preliminary of matrix information geometry

MIG studies intrinsic properties of matrix manifolds, whose elements, e.g., matrix-valued data, can be discriminated by a geometric measure. Many problems in information science related to matrix-valued data can be equivalently formulated as measuring the difference of two points in a matrix manifold. In this section, we briefly introduce the theory of HPD manifolds and define several geometric measures as the discrimination functions using the Frobenius metric.

2.1 HPD matrix manifolds

Let \(\mathcal{H}(n, \mathbb{C})\) be the set of \(n \times n\) Hermitian matrices, given by
\[
\mathcal{H}(n, \mathbb{C}) = \{ A \mid A = A^H, A \in \mathbb{C}^{n \times n} \}. \tag{2.1}
\]
A Hermitian matrix \(A \in \mathcal{H}(n, \mathbb{C})\) is positive-definite, expressed as \(A \succ 0\), if the quadratic expression \(x^H Ax > 0\) holds for all \(x \in \mathbb{C}^n, x \neq 0\). The set of \(n \times n\) HPD matrices is
\[
\mathcal{P}(n, \mathbb{C}) = \{ A \mid A \succ 0, A \in \mathcal{H}(n, \mathbb{C}) \}. \tag{2.2}
\]
The subspace \(\mathcal{P}(n, \mathbb{C})\) becomes a Riemannian manifold of non-positive curvature equipped with an AIRM \[15\], which has been extensively studied during the recent few decades \[41, 50, 51\]. The HPD matrix manifold \(\mathcal{P}(n, \mathbb{C})\) has many interesting properties \[48\]. For instance,

1. Its closure forms a closed, nonpolyhedral, self-dual convex cone.
2. It embodies a canonical higher-rank symmetric space.

Each element of the matrix manifold \(\mathcal{P}(n, \mathbb{C})\) denotes an HPD matrix. Given any two HPD matrices, their difference on the HPD matrix manifold \(\mathcal{P}(n, \mathbb{C})\) can be measured by a metric, a divergence or other measures. Different measures amount to different geometric structures. Thus, it is important to choose an appropriate measure function for a specific application. In the following, we introduce several geometric measures on the HPD matrix manifold.

2.2 Geometric measures on HPD matrix manifolds

We first introduce the AIRM and AIRM distance on the HPD matrix manifold \(\mathcal{P}(n, \mathbb{C})\). The tangent space at a point \(P \in \mathcal{P}(n, \mathbb{C})\) is \(T_P \mathcal{P}(n, \mathbb{C}) = P \times \mathcal{H}(n, \mathbb{C})\). A well-known AIRM defined in \(T_P \mathcal{P}(n, \mathbb{C})\) is
\[
\langle A, B \rangle_P := \text{tr} \left( P^{-1} A P^{-1} B \right), \quad A, B \in T_P \mathcal{P}(n, \mathbb{C}). \tag{2.3}
\]

There are natural exponential map and logarithm map on the tangent bundle \(T \mathcal{P}(n, \mathbb{C}) = \bigcup_P T_P \mathcal{P}(n, \mathbb{C})\) using the geodesics. They are related to matrix exponentials and matrix logarithms. Matrix exponential for a general matrix \(X\) is given by
\[
\exp(X) = \sum_{i=0}^{\infty} \frac{X^i}{i!}. \tag{2.4}
\]
The following proposition defines the principle logarithm of an invertible matrix.
Lemma 2.1 ([28,44]). Suppose \( \mathbf{X} \) is an invertible matrix that does not have eigenvalues in the closed negative real line. Then, there exists a unique logarithm with eigenvalues lying in the strip \( \{ z \in \mathbb{C} \mid -\pi < \text{Im}(z) < \pi \} \). This logarithm is called the principle logarithm and denoted by \( \text{Log} \mathbf{X} \).

Moreover, the matrix \( \mathbf{X} \) satisfies the following properties.

(i) Both \( \mathbf{X} \) and \( \text{Log} \mathbf{X} \) commute with \( ([\mathbf{X} - \mathbf{I}]s + \mathbf{I})^{-1} \) for any real number \( s \).

(ii) The following identity holds that

\[
\int_0^1 [(\mathbf{X} - \mathbf{I})s + \mathbf{I}]^{-2} ds = (\mathbf{I} - \mathbf{X})^{-1}[(\mathbf{X} - \mathbf{I})s + \mathbf{I}]^{-1}
= \mathbf{X}^{-1}.
\]

(iii) Suppose \( \mathbf{A}(\varepsilon) \) is an invertible matrix which does not have eigenvalues lying in the closed real line. Then, we have

\[
\frac{d}{d\varepsilon} \text{Log} \mathbf{A}(\varepsilon) = \int_0^1 [(\mathbf{A}(\varepsilon) - \mathbf{I})s + \mathbf{I}]^{-1} \frac{d}{d\varepsilon} \mathbf{A}(\varepsilon) [(\mathbf{A}(\varepsilon) - \mathbf{I})s + \mathbf{I}]^{-1} ds.
\]

The AIRM distance of two points \( \mathbf{X}, \mathbf{Y} \in \mathcal{P}(n, \mathbb{C}) \) is defined as length of the local geodesic connecting them, which is given by

\[
d_R^2(\mathbf{X}, \mathbf{Y}) = \left\| \text{Log} \left( \mathbf{X}^{-1/2} \mathbf{Y} \mathbf{X}^{-1/2} \right) \right\|^2
= \sum_{i=1}^n \ln^2 \lambda_i,
\]

where \( \lambda_i \) is the \( i \)-th eigenvalue of \( \mathbf{X}^{-1/2} \mathbf{Y} \mathbf{X}^{-1/2} \). Here \( \| \cdot \| \) is the Frobenius norm \( \| \mathbf{A} \|^2 = \text{tr}(\mathbf{A}^\mathsf{H} \mathbf{A}) \) of a matrix \( \mathbf{A} \), induced from the Frobenius metric.

Unfortunately, the computational cost of the AIRM distance is expensive in practical applications. An alternative selection, the Log-Euclidean metric (LEM), is derived by defining the metric in the tangent space \( T_\mathbf{P} \mathcal{P}(n, \mathbb{C}) \ni \mathbf{A}, \mathbf{B} \) using the Frobenius metric as follows:

\[
\langle \mathbf{A}, \mathbf{B} \rangle_L^\mathcal{E} := \langle D_A \text{Log} \mathbf{P}, D_B \text{Log} \mathbf{P} \rangle,
\]

where \( D_A \text{Log} \mathbf{P} \) denotes the directional derivative of the logarithm function along direction \( \mathbf{A} \) at a point \( \mathbf{P} \). The LEM metric is more efficient in the sense of computational cost, compared with the AIRM metric. In addition to them, many geometric measures that possess good properties have also been introduced on the manifold \( \mathcal{P}(n, \mathbb{C}) \), for instance, the JBLD divergence [19], the KLD [14], and the SKLD [32]. Among these geometric measures, only the AIRM and LEM metrics lead to a true geodesic distance on the manifold \( \mathcal{P}(n, \mathbb{C}) \). It should also be noted that only the AIRM, the JBLD divergence and the SKLD are invariant w.r.t. affine transformations. In this paper, we will mainly be focused on the LEM, the JBLD divergence and the SKLD, defined in the manifold \( \mathcal{P}(n, \mathbb{C}) \).

Definition 2.2. The LEM distance between two matrices \( \mathbf{X}, \mathbf{Y} \in \mathcal{P}(n, \mathbb{C}) \) is defined as

\[
d_L^2(\mathbf{X}, \mathbf{Y}) = \left\| \text{Log} \mathbf{X} - \text{Log} \mathbf{Y} \right\|^2
\]

Definition 2.3. The JBLD divergence of two matrices \( \mathbf{X}, \mathbf{Y} \in \mathcal{P}(n, \mathbb{C}) \) is defined as

\[
d_J^2(\mathbf{X}, \mathbf{Y}) = \ln \text{det} \left( \frac{\mathbf{X} + \mathbf{Y}}{2} \right) - \frac{1}{2} \ln \text{det}(\mathbf{X} \mathbf{Y}).
\]

Definition 2.4. The SKLD divergence of two matrices \( \mathbf{X}, \mathbf{Y} \in \mathcal{P}(n, \mathbb{C}) \) is defined as

\[
d_S^2(\mathbf{X}, \mathbf{Y}) = \text{tr} \left( \mathbf{Y}^{-1} \mathbf{X} + \mathbf{X}^{-1} \mathbf{Y} - 2 \mathbf{I} \right).
\]
3 Problem formulation

Consider the detection of a target signal embedding into a clutter environment. We assume that a set of $K$ secondary data, free of signal components, is available. As customary, the detection problem can be formulated as the following binary hypotheses test,

$$
\begin{align*}
\mathcal{H}_0 : & \quad x = c \\
& \quad x_k = c_k, k = 1, 2, \ldots, K \\
\mathcal{H}_1 : & \quad x = c \\
& \quad x_k = \alpha p + c_k, k = 1, 2, \ldots, K
\end{align*}
$$

(3.1)

where $x \in \mathbb{C}^n$ and $x_k \in \mathbb{C}^n$, $k = 1, 2, \ldots, K$ denote the primary data and secondary data, respectively, $c \in \mathbb{C}^n$ and $c_k \in \mathbb{C}^n$, $k = 1, 2, \ldots, K$ are the clutter data, $\alpha \in \mathbb{R}$ is a deterministic but unknown real number, which stands for the target property and channel propagation effects, and $p \in \mathbb{C}^n$ is a known signal steering vector. The purpose of signal detection is to discriminate the target signal from the clutter data. Generally, this problem is considered as the estimation of the logarithm likelihood function of the statistical models of $\mathcal{H}_0$ and $\mathcal{H}_1$, and can be solved by resorting to the generalized likelihood ratio test criteria. One crucial drawback of this method is that it treats the statistical models independently regardless of their intrinsic connections due to the manifold structure. At the moment, we assume that the sample data $x$ is modeled as an HPD matrix and then reformulate the problem of signal detection as discriminating the target signal and clutter on the HPD matrix manifold $\mathcal{P}(n, \mathbb{C})$. Specifically, the decision on the presence or absence of a target is made by comparing the geometric distance (or difference) between the matrix in the range cell under test (CUT) $R_{\text{CUT}}$ and the estimated CCM $\hat{R}$ with a given threshold $\gamma$. Consequently, the rule of signal detection can be formulated as follows:

$$
d(R_{\text{CUT}}, \hat{R}) \overset{\mathcal{H}_1}{\geq} \gamma,
$$

(3.2)

where $d(\cdot, \cdot)$ denotes a geometric measure, and $\gamma$ is the detection threshold that is set to maintain a constant false alarm rate. It can be noted from (3.2) that the detection performance greatly depends on the measure function used to discriminate two points on the HPD matrix manifold. Moreover, robustness of the estimated CCM to outliers also has an effect on the detection performance. The commonly used CCM estimator is the sample covariance matrix (SCM). For a set of secondary data $\{x_1, x_2, \ldots, x_K\}$, the SCM is the maximum likelihood estimation (MLE) of secondary data given by

$$
\hat{R}_{\text{SCM}} = \frac{1}{K} \sum_{k=1}^{K} x_k x_k^H.
$$

(3.3)

Note that the SCM estimator can be viewed as the arithmetic mean of $K$ autocorrelation matrices $\{x_k x_k^H\}_{k=1}^{K}$ with rank one. The matrix $x_k x_k^H$ is the MLE of the secondary data $x_k$. To analyze the matrix data on the HPD matrix manifold, for any a sample data $x = (x_0, x_1, \ldots, x_{n-1})^T$, a suboptimal estimation that is positive-definite can be derived as

$$
\hat{R} = \begin{bmatrix}
  r_0 & r_1^* & \cdots & r_{n-1}^* \\
  r_1 & r_0 & \cdots & r_{n-2}^* \\
  \vdots & \ddots & \ddots & \vdots \\
  r_{n-1}^* & \cdots & r_1 & r_0
\end{bmatrix}
$$

(3.4)

with entries

$$
\begin{align*}
r_l &= \mathbb{E}[x_l x_{l+i}^*], \quad 0 \leq l \leq n - 1, 0 \leq i \leq n - l - 1,
\end{align*}

(3.5)

where $x_l$ is the $l$-th correlation coefficient of the data $x$ and $\mathbb{E}[\cdot]$ denotes the statistical expectation. According to the ergodicity of stationary Gaussian process, the correlation coefficient of the data $x$ can be calculated by averaging over time instead of its statistical expectation, as follows:

$$
r_l = \frac{1}{n} \sum_{i=0}^{n-l-1} x_l x_{l+i}^*, \quad 0 \leq l \leq n - 1.
$$

(3.6)
The estimation in (3.4) allow us to model each sample data as an HPD matrix. Then, the CCM can be estimated according to a set of HPD matrices \{\hat{R}_1, \hat{R}_2, \ldots, \hat{R}_K\}. It is known that the SCM is sensitive to outliers in nonhomogeneous environments. Take the nonlinear geometric structure of HPD matrix manifolds into consideration, a robust CCM estimator can be chosen as the geometric mean or median. The geometric mean estimator has been discussed in details [29]. In this paper, we devote to studying geometric median estimators. The geometric median is defined as follows.

**Definition 3.1.** For HPD matrices \{\hat{R}_1, \hat{R}_2, \ldots, \hat{R}_K\}, the geometric median estimator associated with a geometric measure is the unique solution of the minimizer of the sum of the geometric measure

\[ \hat{R}_g = \arg\min_{R \in \mathcal{P}(n, C)} \sum_{k=1}^{K} d(R, \hat{R}_k). \] (3.7)

Based on the geometric median estimator \( \hat{R}_g \), we can design the MIG detector by replacing the CCM with \( \hat{R}_g \). Different geometric measures possess different discriminative power. Besides, the robustness of different geometric estimators to outliers are also different. These will be clarified later.

4 The proposed detectors

According to the principle of signal detection mentioned above, we present the block-scheme of the proposed detector in Fig. 1. We first model each sample data as an HPD matrix, and then map each HPD matrix to another HPD matrix. The detection statistic is computed by the geometric distance between the HPD matrix in the CUT and the CCM estimated by the geometric median related to a measure. Finally, the decision on the existence or absence of a target is made by comparing the detection statistic with a given threshold. Here, the manifold-to-manifold map and the geometric median are the two essential tools for designing the detector. In the following, we introduce a manifold-to-manifold map by a manifold filter and deduce geometric medians for a set of HPD matrices. We also define the anisotropy index to analyze the discriminative power of a geometric measure.

**4.1 Manifold filter: Weighting a set of HPD matrices**

In this subsection, a weighting map in the HPD manifold is defined to derive a more discriminative measure of a given set of HPD matrices.

Suppose that the HPD matrix modeled using the sample data contains some redundant information, which would often limit the discriminative power of the intra-class and the inter-class. As illustrated by Fig. 2 given the data collected from two classes on an HPD manifold, marked as the blue circle and the red square, we propose a manifold filter that cuts down the redundant information contained in the sample data to improve the discriminative power. The improvement is achieved by reducing the intra-class distance while increasing the inter-class distance.

Figure 1: Block-scheme of the proposed detector.

Figure 2: Manifold filter.
The manifold filter is designed by mapping an HPD matrix to the weighted average of its surrounding HPD matrices. Suppose that \( R \) is an HPD matrix in one cell, and its surrounding \( m \) HPD matrices is denoted as \( \{ R_1, R_2, \ldots, R_m \} \). The weighted average matrix \( \tilde{R} \) is given by

\[
\tilde{R} := g \{ w_1 R_1, w_2 R_2, \ldots, w_m R_m \},
\]

(4.1)

where \( w_i \) is the \( i \)-th weight. The weighted average matrix \( \tilde{R} \) can be chosen as the arithmetic mean, geometric mean or geometric median. In each case, the matrix \( \tilde{R} \) is still HPD and this defines a map from a HPD matrix to another HPD matrix: \( R \mapsto \tilde{R} \). Analogous to the filter in image denoising, the arithmetic mean can be viewed as a mean filter \([26]\) and the geometric mean or median can be viewed as nonlocal mean filters \([45]\). In this paper, we choose the arithmetic mean as the weighted average matrix. Other means or medians will be reported separately. Thus, the weighted average matrix reads

\[
\tilde{R} = \sum_{i=1}^{m} w_i R_i,
\]

(4.2)

where a further constraint of the weights are imposed, namely

\[
\sum_{i=1}^{m} w_i = 1.
\]

(4.3)

We define each weight \( w_i \) to represent the similarity between the matrix \( R_i \) and the matrix \( R \) such that the smaller their difference, the greater the weight. For simplicity, we use an exponential function to define the weights as follows

\[
w_i = \frac{\exp\left(-\frac{d^2(R_i, R)}{h^2}\right)}{\sum_{j=1}^{m} \exp\left(-\frac{d^2(R_j, R)}{h^2}\right)},
\]

(4.4)

where \( h \) is a constant parameter. Substituting (4.4) into (4.2), we obtain

\[
\tilde{R} = \sum_{i=1}^{m} \frac{\exp\left(-\frac{d^2(R_i, R)}{h^2}\right)}{\sum_{j=1}^{m} \exp\left(-\frac{d^2(R_j, R)}{h^2}\right)} R_i.
\]

(4.5)

In the manifold filter, \( m \) and \( h \) are free parameters, which can be adjusted properly in practical applications to monitor the filter. In general, bigger \( m \) leads to better filter results as more information is included. The parameter \( h \) is related to the influence of data on the filter weight: a bigger \( h \) usually means a smaller variation of the weights.

### 4.2 Geometric medians

Geometric median of the weighted HPD matrices is considered as the estimate of CCM. According to (3.7), we can obtain the geometric median related to a measure for a set of HPD matrices. The fixed-point iteration algorithms for computing the LEM, JBLD divergence and SKLD medians are given below. Be noted that fixed-point iteration for a given fixed-point problem is not unique.

To find the median minimizing the function

\[
F(R) = \sum_{k=1}^{K} d(R, \tilde{R}_k), \quad R \in \mathcal{P}(n, C),
\]

(4.6)

it suffices to solve the vanishment of its gradient, i.e., \( \nabla F(R) = 0 \), which is defined using the Frobenius metric as (see, e.g., \([34]\))

\[
\langle \nabla F(R), X \rangle := \frac{d}{d\varepsilon} \bigg|_{\varepsilon=0} F(R + \varepsilon X), \quad \forall X \in \mathcal{P}(n, C).
\]

(4.7)

In this paper, we show fixed-point iteration algorithms for solving the algebraic equation \( \nabla F(R) = 0 \). Note that the algorithms may not be unique.
**Proposition 4.1.** The LEM median of a set of HPD matrices \( \{ \tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_K \} \) can be computed via the following fixed-point iteration
\[
R_{t+1} = \exp \left\{ \sum_{k=1}^{K} \frac{\log \tilde{R}_k}{\left\| \log R_t - \log \tilde{R}_k \right\|} \right\},
\]
where \( t \) is a nonnegative integer. Uniqueness of the median is proved in [16].

**Proof.** It was shown in [16] that median related to the function
\[
F(R) = \sum_{k=1}^{K} \left\| \log \tilde{R}_k - \log R \right\|
\]
is characterized by
\[
\sum_{k=1}^{K} \frac{\log \tilde{R}_k - \log R}{\left\| \log \tilde{R}_k - \log R \right\|} = 0.
\]
The fixed-point iteration follows immediately.

**Proposition 4.2.** The JBLD median of a set of HPD matrices \( \{ \tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_K \} \) can be computed by using the fixed-point iteration
\[
R_{t+1} = \frac{1}{2} \left( \sum_{k=1}^{K} \frac{1}{\sqrt{\ln \det \left( \frac{R_t + \tilde{R}_k}{2} \right) - \frac{1}{2} \ln \det \left( R_t \tilde{R}_k \right)}} \right)^{-1} \left( \sum_{k=1}^{K} \frac{\left( \frac{R_t + \tilde{R}_k}{2} \right)^{-1} - R^{-1}}{\sqrt{\ln \det \left( \frac{R_t + \tilde{R}_k}{2} \right) - \frac{1}{2} \ln \det \left( R_t \tilde{R}_k \right)}} \right). \]

**Proof.** Denote \( F(R) \) as the function to be minimized, namely
\[
F(R) = \sum_{k=1}^{K} \left[ \ln \det \left( \frac{R + \tilde{R}_k}{2} \right) - \frac{1}{2} \ln \det \left( R\tilde{R}_k \right) \right].
\]

Its gradient can be computed via the definition (4.7) and we obtain
\[
\nabla F(R) = \frac{1}{4} \sum_{k=1}^{K} \frac{\left( \frac{R + \tilde{R}_k}{2} \right)^{-1} - R^{-1}}{\sqrt{\ln \det \left( \frac{R + \tilde{R}_k}{2} \right) - \frac{1}{2} \ln \det \left( R\tilde{R}_k \right)}}.
\]

Moving the \( R^{-1} \) terms of the equation \( \nabla F(R) = 0 \) to one side, we obtain the fixed-point iteration.

**Proposition 4.3.** The SKLD median of a set of HPD matrices \( \{ \tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_K \} \) can be computed by the fixed-point iteration
\[
R_{t+1} = P_t^{-1/2} \left( P_t^{1/2} Q_t P_t^{1/2} \right)^{1/2} P_t^{-1/2},
\]
where
\[
P_t = \sum_{k=1}^{K} \frac{\tilde{R}_k^{-1}}{\sqrt{\text{tr} \left( \tilde{R}_k^{-1} R_t + R_t^{-1} \tilde{R}_k \right) - 2n}}
\]
and
\[
Q_t = \sum_{k=1}^{K} \frac{\tilde{R}_k}{\sqrt{\text{tr} \left( \tilde{R}_k^{-1} R_t + R_t^{-1} \tilde{R}_k \right) - 2n}}.
\]
Proof. For the function

\[ F(R) = \sum_{k=1}^{K} \sqrt{\text{tr} \left( \tilde{R}_k^{-1} R + R^{-1} \tilde{R}_k \right) - 2n} \]  

(4.17)

direct computation using definition (4.7) gives

\[ \nabla F(R) = \frac{1}{2} \sum_{k=1}^{K} \frac{\tilde{R}_k^{-1} - R^{-1} \tilde{R}_k R^{-1}}{\sqrt{\text{tr} \left( \tilde{R}_k^{-1} R + R^{-1} \tilde{R}_k \right) - 2n}}. \]  

(4.18)

The equation \( \nabla F(R) = 0 \) can be rewritten as follows

\[ RPR = Q, \]  

(4.19)

which looks like a special algebraic Riccati equation but note that the matrices \( P \) and \( Q \) also depend on \( R \), given by

\[ P = \sum_{k=1}^{K} \frac{\tilde{R}_k^{-1}}{\sqrt{\text{tr} \left( \tilde{R}_k^{-1} R + R^{-1} \tilde{R}_k \right) - 2n}} \]  

(4.20)

and

\[ Q = \sum_{k=1}^{K} \frac{\tilde{R}_k}{\sqrt{\text{tr} \left( \tilde{R}_k^{-1} R + R^{-1} \tilde{R}_k \right) - 2n}}. \]  

(4.21)

Since both \( P \) and \( Q \) are HPD matrices, we multiply on both sides of (4.19) by \( P^{1/2} \) from both the left and the right, amounting to

\[ P^{1/2}RPRP^{1/2} = P^{1/2}QP^{1/2}. \]  

(4.22)

The matrix \( R \) can hence be solved in terms of \( P \) and \( Q \) as follows

\[ R = P^{-1/2} \left( P^{1/2}QP^{1/2} \right)^{1/2} P^{-1/2}. \]  

(4.23)

The fixed-point iteration is proposed according to the final equation.

\[ \square \]

4.3 Anisotropy index

The discriminative power of weighted HPD matrices can be shown through the corresponding anisotropy indices. The anisotropy index associated with a measure, e.g., a metric or a divergence, at any point of a matrix manifold reflects the local geometric structure around this point.

Definition 4.4. The anisotropy index related to a geometric measure at a point \( R \in \mathcal{P}(n, \mathbb{C}) \) is defined by

\[ a(R) := \min_{\varepsilon > 0} d^2(R, \varepsilon I). \]  

(4.24)

Namely, the anisotropy index \( a \) is the minimum square of a geometric distance between \( R \) and \( \varepsilon I \) (\( \varepsilon > 0 \)). This distance represents the projection distance from matrix \( R \) to the subspace \( \{ \varepsilon I \mid \varepsilon \in \mathbb{R}^+ \} \). Bigger the distance, lager the anisotropy index. Next, we will show the anisotropy indices for the LEM, JBLD divergence and SKLD, respectively. They can be simply calculated by considering minimum of the function

\[ f(\varepsilon) = d^2(R, \varepsilon I), \quad \forall R \in \mathcal{P}(n, \mathbb{C}). \]  

(4.25)

Proposition 4.5. The anisotropy index related to the LEM metric at \( R \in \mathcal{P}(n, \mathbb{C}) \) is given by

\[ a_L(R) = \| \log R - \log(\varepsilon^* I) \|^2, \]  

(4.26)

where \( \varepsilon^* = \sqrt{\det(\tilde{R})} \).
Proof. Now the function $f(\varepsilon)$ reads
\[ f(\varepsilon) = \| \log R - \log(\varepsilon I) \|^2, \]  
whose derivative can immediately be computed:
\[ f'(\varepsilon) = - \frac{2}{\varepsilon} \text{tr} (\log R - \log(\varepsilon I)) \]
\[ = - \frac{2}{\varepsilon} (\ln \det(R) - \ln \varepsilon^n). \]  
Note that Lemma 2.1 is used here. Stationary condition that $f'(\varepsilon) = 0$ gives the unique positive solution, denoted by $\varepsilon^*$:
\[ \varepsilon^* = \sqrt{\det(R)}. \]  

Proposition 4.6. The anisotropy index associated with the JBLD divergence at a point $R \in \mathcal{P}(n, \mathbb{C})$ is given by
\[ a_J(R) = \ln \det \left( \frac{R + \varepsilon^* I}{2} \right) - \frac{1}{2} \ln \det(R) \det(\varepsilon^* I), \]  
where $\varepsilon^*$ is the unique positive solution of the equation
\[ 2 \sum_{i=1}^{n} \frac{1}{\lambda_i + \varepsilon} = \frac{n}{\varepsilon} \]  
for all $n \geq 1$, where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of $R$.

Proof. See Appendix A.

Proposition 4.7. The anisotropy index associated with the SKLD divergence at a point $R \in \mathcal{P}(n, \mathbb{C})$ is
\[ a_S(R) = \text{tr} \left( \varepsilon^* R^{-1} + \frac{1}{\varepsilon^*} R - 2I \right), \]  
where
\[ \varepsilon^* = \sqrt{\frac{\text{tr}(R)}{\text{tr}(R^{-1})}}. \]  

Proof. The proof is similar by considering stationary condition of the function
\[ f(\varepsilon) = \text{tr} \left( \varepsilon R^{-1} + \frac{1}{\varepsilon} R - 2I \right), \]  
whose derivative is
\[ f'(\varepsilon) = \text{tr} \left( R^{-1} - \frac{1}{\varepsilon^2} R \right). \]  
By solving $f'(\varepsilon) = 0$, the proof is complete.

5 Numerical results

The class of MIG detectors we propose in this paper is based on medians of a set of HPD matrices with a manifold filter, corresponding to various geometric measures. To illustrate the effectiveness of those MIG median detectors, in this section, numerical simulations are provided to analyze them from three aspects: robustness to outliers; the discriminative power on the matrix manifold before and after taking the manifold filter, e.g., matrix weighting, into consideration (see Section 4.1); and detection performances.
5.1 Robustness of geometric medians

In this subsection, we investigate the robustness of geometric medians in terms of the sensitivity to the number of interferences as well as to the number of samples, respectively. The sample dataset is generated according to a multivariate complex Gaussian distribution with zero mean and the following covariance matrix

\[ \Sigma = \sigma_n^2 I + \sigma_c^2 \Sigma_c, \]  

where \( \sigma_n^2 I \) and \( \sigma_c^2 \Sigma_c \) denote the noise component with power \( \sigma_n^2 \) and the clutter component with power \( \sigma_c^2 \), respectively. The clutter-to-noise ratio (CNR) is defined by \( \text{CNR} = \sigma_c^2 / \sigma_n^2 \). Entries of \( \Sigma_c \) are given by

\[ \Sigma_c(j, k) = \rho^{|j-k|} \exp(i 2\pi f_c(j-k)), \quad j, k = 1, 2, \ldots, n, \]  

where \( \rho \) is the one-lag correlation coefficient, and \( f_c \) denotes the clutter normalized Doppler frequency. In the simulations, the parameter are fixed with \( \rho = 0.95, \sigma_n^2 = 1, \text{CNR} = 20 \text{ dB}, f_c = 0.1 \). A set of 40 sample data is generated without interferences, and then, a number of 1, 2, \ldots, 15 interferences are respectively injected into the sample data. The normalized Doppler frequency of all interferences is set to be 0.2. The SCNR is defined as

\[ \text{SCNR} = |\alpha|^2 p \Sigma^{-1} p, \]  

where \( \alpha \) denotes the amplitude coefficient and \( p \) is the steering vector. Here, SCNR is set to be 15 dB. An offset error \( L_{\text{error}} \) that denotes the difference between the median \( R_0 \) of the sample data without interferences and the median \( R_{\text{interf}} \) of the sample data with interferences is defined by

\[ L_{\text{error}} = \left\| R_0 - R_{\text{interf}} \right\| / \| R_0 \|. \]  

Fig. 3 shows offset errors of the SCM and geometric medians under different number of interferences. It is clear that geometric medians are more robust than the SMC, and among them, the JBLD median is the most robust against interferences.

![Figure 3: Offset errors under different number of interferences.](image)

We further analyze their robustness through energy distribution analysis based on either of the following criteria: 1) a method is more robust if smaller change about its energy distribution occurs w.r.t. the same number of interferences; 2) a method is more robust if its energy distribution is more compact, namely with smaller covariance. Fig. 4 and Fig. 5 show the energy distributions without and with 10 interferences, respectively. From criterion 1), the JBLD median and the LEM median are the most robust, while the SKL median is less robust but more robust compared with the SCM. From criterion 2), it is obvious that the geometric medians are more compact and hence more robust than the SCM. In particular, the JBLD median and the LEM median are again the most robust. In
summary, energy distribution analysis shows that the geometric medians are more robust than the SCM and the JBLD and LEM medians are the most robust.

Next, we are going to show their robustness against the number of sample data. The sample dataset is generated according to a multivariate complex Gaussian distribution with zero mean and the covariance matrix. Denote the covariance matrix of those samples by $\hat{R}$, and define the error against the number of sample data by

$$T_{\text{error}} = \frac{\| \hat{R} - \Sigma \|}{\| \Sigma \|}.$$  \hspace{1cm} (5.5)

It is shown in Fig. 6 that the SKL median is the most robust and geometric medians are much
Figure 6: Offset errors of the SCM and geometric medians under different number of sample data.

Figure 7: Energy distributions with 8 sample data.

more robust than the SCM.

In Fig. 7 we show the energy distributions with 8 sample data. It is clear that the geometric medians are more compact and hence more robust compared with the SCM.

5.2 The discriminative power

In this subsection, we analyze the discriminative power by comparing the statistics with and without manifold filter. First of all, we compare the normalized detection statistics before and after weighting the matrices; see Eq. 4.5 in Section 4.1. A number of 40 sample data is simulated and an interference is injected into the 20-th sample with SCNR = 15 dB and $f_d = 0.2$. Parameters $m$ and $h$ are chosen to be $m = 11, h = 1.5$, $m = 11, h = 2$ and $m = 13, h = 1.5$, respectively. The detection statistic, namely the distance between the CMM and the HPD matrix in the CUT, is computed with and without manifold filter. The normalized detection statistics are plotted in Fig. 8. Obviously, the detection statistic after manifold filter is smaller than that before filter in the cell without target signal. In other words, after manifold filter, the distance between clutter and clutter becomes smaller while the distance between target and clutter becomes larger.
Anisotropy analysis is also used to illustrate the discriminative power. To analyze the difference between the target signal and the CCM in terms of the anisotropy on the matrix manifold, we define an anisotropy index of discrimination (AD) to measure the difference on the matrix manifold as

\[ AD = \frac{AI_{\text{signal}}}{AI_{\text{CCM}}}. \]  

(5.6)

Here, \( AI_{\text{signal}} \) and \( AI_{\text{CCM}} \) denote the target and clutter anisotropies, respectively reflecting the local geometric structures at locations of the target and the CCM on the matrix manifold. AD represents the difference in the anisotropy for the target and the CCM. The larger the value of AD, the greater the difference between the two points. Fig. 9 shows the AD with 100 Monte Carlo trials. It can be seen that after manifold filter, the values of AD become larger, implying that the difference in local geometric structures between target signal and clutter becomes larger after manifold filter.

5.3 Detection performances

In this subsection, numerical simulations are conducted to validate performances of the proposed detectors in comparison with their counterparts and the AMF. To ease from huge computational load, the probability of false alarm \( P_{FA} \) is chosen to be \( 10^{-3} \). For the Monte Carlo simulations, the detection threshold and the detection probability \( P_d \) are obtained from \( 100/P_{FA} \) and 2000 independent trials, respectively. In the following simulations, the dimension of matrices is \( n = 8 \); two interferences are randomly added to the supplementary data with SCNR = 10 dB and the normalized frequency \( f_d = 0.2 \). Figs. 10, 11 and 12 show the detection performances of the proposed MIG median detectors in Gaussian and non-Gaussian clutter with two interferences. The number of sample data is \( K = 8 \) or 12. When \( K = 8 \), the AMF is much worse compared with the MIG median detectors with or without manifold filter. Moreover, the MIG median detectors with manifold filter have better performance than the MIG median detectors without manifold filter. When \( K = 1.5n \), the AMF and the geometric median detectors behave very similarly while geometric median detectors with manifold filter are slightly better than the AMF.

6 Conclusions

In this paper, we designed a class of MIG median detectors with manifold filter to detect signals embedded in nonhomogeneous environments. The sample data was modeled as an HPD matrix with the Toeplitz structure. Particularly, the clutter covariance matrix was estimated as a geometric median for a set of HPD matrices w.r.t. various geometric measures. Then, the problem of signal detection was formulated into the discrimination of two points on the HPD matrix manifold. Furthermore, a manifold-to-manifold map was defined by using the manifold filter to alter the manifold structure, that consequently improved the discrimination power on the matrix manifold. Numerical simulations showed the performance advantages of those MIG median detectors compared with the AMF and their state-of-the-art counterparts. Similar to the AMF, MIG detectors can also be widely applied in practical problems, for instance, target detection in sea or ground clutter, sonar signal detection, and signal detection in communication system. Potential future research will concern the subspace signal detection and their applications for range distributed target detection and synthetic aperture radar image target detection.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China under Grant No. 61901479 and JST-CREST. L. Peng is an adjunct faculty member of Waseda Institute for Advanced Study, Waseda University, Japan, and School of Mathematics and Statistics, Beijing Institute of Technology, China.
Figure 8: Normalized statistics with and without manifold filter

References

[1] A. Ali, N. González-Prelcic, and R. W. Heath. Spatial covariance estimation for millimeter wave hybrid systems using out-of-band information. *IEEE Transactions on Wireless Communications*, 18(12):5471–5485, 2019.

[2] M. Arnaudon, F. Barbaresco, and L. Yang. Riemannian medians and means with applications to radar signal processing. *IEEE Journal of Selected Topics in Signal Processing*, 7(4):595–604, 2013.
Figure 9: Discriminative power.

[3] A. Aubry, A. De Maio, and L. Pallotta. A geometric approach to covariance matrix estimation and its applications to radar problems. *IEEE Transactions on Signal Processing*, 66(4):907–922, 2018.

[4] A. Aubry, A. De Maio, L. Pallotta, and A. Farina. Maximum likelihood estimation of a structured covariance matrix with a condition number constraint. *IEEE Transactions on Signal Processing*, 60(6):3004–3021, 2012.

[5] A. Aubry, A. D. Maio, L. Pallotta, and A. Farina. Covariance matrix estimation via geometric barycenters and its application to radar training data selection. *IET Radar, Sonar & Navigation*, 7(6):600–614, July 2013.

[6] A. Aubry, A. D. Maio, L. Pallotta, and A. Farina. Median matrices and their application to radar training data selection. *IET Radar, Sonar & Navigation*, 8(4):265–274, 2014.

[7] F. Bandiera, O. Besson, and G. Ricci. Knowledge-aided covariance matrix estimation and adaptive detection in compound-Gaussian noise. *IEEE Transactions on Signal Processing*, 58(10):5391–5396, 2010.

[8] F. Barbaresco. Innovative tools for radar signal processing Based on Cartan’s geometry of SPD matrices & Information Geometry. In *2008 IEEE Radar Conference*, pages 1–6, 2008.

[9] F. Barbaresco. RRP 3.0: 3rd generation Robust Radar Processing based on Matrix Information Geometry (MIG). In *2012 9th European Radar Conference*, pages 42–45, 2012.

[10] F. Barbaresco. Coding statistical characterization of radar signal fluctuation for Lie group machine learning. In *2019 International Radar Conference (RADAR)*, pages 1–6, 2019.

[11] F. Barbaresco and Uwe Meier. Radar monitoring of a wake vortex: Electromagnetic reflection of wake turbulence in clear air. *Comptes Rendus Physique*, 11(1):54–67, 2010.

[12] O. Besson, J. Tourneret, and S. Bidon. Knowledge-aided Bayesian detection in heterogeneous environments. *IEEE Signal Processing Letters*, 14(5):355–358, 2007.

[13] S. Bidon, O. Besson, and J. Tourneret. A Bayesian approach to adaptive detection in nonhomogeneous environments. *IEEE Transactions on Signal Processing*, 56(1):205–217, 2008.

[14] N. Bouhlel and A. Dziri. Kullback–Leibler divergence between multivariate generalized Gaussian distributions. *IEEE Signal Processing Letters*, 26(7):1021–1025, 2019.
Figure 10: $P_d$ vs SCNR with JBLD divergence

[15] M. R. Bridson and A. Häfliger. *Metric Spaces of Non-Positive Curvature*, volume 319. Springer Science & Business Media, 2013.

[16] M. Charfi, Z. Chebbi, M. Moakher, and B. C. Vemuri. Bhattacharyya median of symmetric positive-definite matrices and application to the denoising of diffusion-tensor fields. In *Proc IEEE Int Symp Biomed Imaging*, pages 1227–1230, 2013.

[17] X. Chen, Z. J. Wang, and M. J. McKeown. Shrinkage-to-tapering estimation of large covariance matrices. *IEEE Transactions on Signal Processing*, 60(11):5640–5656, 2012.

[18] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero. Shrinkage algorithms for MMSE covariance estimation. *IEEE Transactions on Signal Processing*, 58(10):5016–5029, 2010.

[19] A. Cherian, S. Sra, A. Banerjee, and N. Papanikolopoulos. Jensen–Bregman LogDet divergence with application to efficient similarity search for covariance matrices. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(9):2161–2174, 2013.

[20] D. Ciouonzo, A. De Maio, and D. Orlando. On the statistical invariance for adaptive radar detection in partially homogeneous disturbance plus structured interference. *IEEE Transactions on Signal Processing*, 65(5):1222–1234, 2017.

[21] E. Conte, M. Lops, and G. Ricci. Adaptive matched filter detection in spherically invariant noise. *IEEE Signal Processing Letters*, 3(8):248–250, 1996.
(a) Gaussian, $K = 8, n = 8$

(b) Gaussian, $K = 12, n = 8$

(c) Non-Gaussian, $K = 8, n = 8$

(d) Non-Gaussian, $K = 12, n = 8$

Figure 11: $P_d$ vs SCNR with LEM

[22] A. De Maio, A. Farina, and G. Foglia. Knowledge-aided Bayesian radar detectors & their application to live data. *IEEE Transactions on Aerospace and Electronic Systems*, 46(1):170–183, Jan 2010.

[23] A. De Maio, D. Orlando, C. Hao, and G. Foglia. Adaptive detection of point-like targets in spectrally symmetric interference. *IEEE Transactions on Signal Processing*, 64(12):3207–3220, 2016.

[24] A. Decurninge and F. Barbaresco. Robust Burg estimation of radar scatter matrix for autoregressive structured SIRV based on Fréchet medians. *IET Radar, Sonar & Navigation*, 11(1):78–89, 2017.

[25] J. Duník, O. Straka, and M. Šimandl. On autocovariance least-squares method for noise covariance matrices estimation. *IEEE Transactions on Automatic Control*, 62(2):967–972, 2017.

[26] R. G. Gavaskar and K. N. Chaudhury. Fast adaptive bilateral filtering. *IEEE Transactions on Image Processing*, 28(2):779–790, 2019.

[27] L. I. Gliga, H. Chafouk, D. Popescu, and C. Lupu. A method to estimate the process noise covariance for a certain class of nonlinear systems. *Mechanical Systems and Signal Processing*, 131:381–393, 2019.

[28] N. J. Higham. *Functions of Matrices: Theory and Computation*. SIAM, Philadelphia, 2008.
Figure 12: $P_d$ vs SCNR with SKLD

[29] X. Hua, Y. Cheng, H. Wang, and Y. Qin. Robust covariance estimators based on information divergences and Riemannian manifold. *Entropy*, 20(4), 2018.

[30] X. Hua, Y. Cheng, H. Wang, Y. Qin, and D. Chen. Geometric target detection based on total Bregman divergence. *Digital Signal Processing*, 75:232–241, 2018.

[31] X. Hua, Y. Cheng, H. Wang, Y. Qin, and Y. Li. Geometric means and medians with applications to target detection. *IET Signal Processing*, 11(6):711–720, 2017.

[32] X. Hua, Y. Cheng, H. Wang, Y. Qin, Y. Li, and W. Zhang. Matrix CFAR detectors based on symmetrized Kullback–Leibler and total Kullback–Leibler divergences. *Digital Signal Processing*, 69:106–116, 2017.

[33] X. Hua, H. Fan, Y. Cheng, H. Wang, and Y. Qin. Information geometry for radar target detection with total Jensen–Bregman divergence. *Entropy*, 20(4), 2018.

[34] X. Hua, Y. Ono, L. Peng, Y. Cheng, and H. Wang. Target detection within nonhomogeneous clutter via total Bregman divergence-based matrix information geometry detectors. 2020. arXiv:2012.13861.

[35] B. Kang, V. Monga, and M. Rangaswamy. Computationally efficient Toeplitz approximation of structured covariance under a rank constraint. *IEEE Transactions on Aerospace and Electronic Systems*, 51(1):775–785, 2015.
A Anisotropy index of the JBLD divergence: Proof of Proposition 4.6

Denote $f(\varepsilon)$ as the function to be minimized, namely

$$f(\varepsilon) = \ln \det \left( \frac{R + \varepsilon I}{2} \right) - \frac{1}{2} \ln \det(\varepsilon R),$$
whose derivative can be computed directly and reads

\[ f'(\varepsilon) = \text{tr} \left( (R + \varepsilon I)^{-1} - \frac{1}{2\varepsilon} I \right). \]

Let \( \{\lambda_1, \lambda_2, \ldots, \lambda_n\} \) be the eigenvalues of the matrix \( R \), and we have

\[ \text{tr} \left( (R + \varepsilon I)^{-1} \right) = \sum_{i=1}^{n} \frac{1}{\lambda_i + \varepsilon}. \]

Then, the equation \( f'(\varepsilon) = 0 \) becomes

\[ 2 \sum_{i=1}^{n} \frac{1}{\lambda_i + \varepsilon} = \frac{n}{\varepsilon}. \]

It can be rearranged as an \( n \)-th order polynomial for the unknown variable \( \varepsilon \), which has one unique positive solution for all \( n \geq 1 \), denoted by \( \varepsilon^* \). The unique existence is verified in the next Lemma A.1.

**Lemma A.1.** The rational equation

\[ 2 \sum_{i=1}^{n} \frac{1}{\lambda_i + \varepsilon} = \frac{n}{\varepsilon} \]

has only one positive solution \( \varepsilon^* \) for all \( n \geq 1 \).

**Proof.** First of all, we rearrange the equation into the vanishment of a polynomial \( g_n(\varepsilon) = 0 \) (\( n \geq 1 \)) where

\[ g_n(\varepsilon) = 2\varepsilon \left( \sum_{i=1}^{n} (\lambda_1 + \varepsilon)(\lambda_2 + \varepsilon) \cdots (\lambda_i + \varepsilon) \cdots (\lambda_n + \varepsilon) \right) - n(\lambda_1 + \varepsilon)(\lambda_2 + \varepsilon) \cdots (\lambda_n + \varepsilon), \]

satisfying

\[ 2 \sum_{i=1}^{n} \frac{1}{\lambda_i + \varepsilon} - \frac{n}{\varepsilon} = \frac{g_n(\varepsilon)}{\varepsilon \prod_{i=1}^{n}(\lambda_i + \varepsilon)}. \]

Here \( \lambda_i + \varepsilon \) means this term does not appear in the calculation. As all eigenvalues are positive, searching for positive solutions of the rational equation is equivalent to finding positive solutions of \( g_n(\varepsilon) = 0 \).

By mathematical induction, we have

\[ g_n(\varepsilon) = n\varepsilon^n + (n-2)\varepsilon^{n-1} \sum_{i=1}^{n} \lambda_i + (n-4)\varepsilon^{n-2} \sum_{i_1 < i_2} \lambda_{i_1}\lambda_{i_2} \]

\[ + \cdots + (n-2k)\varepsilon^{n-k} \sum_{i_1 < i_2 < \cdots < i_k} \lambda_{i_1}\lambda_{i_2} \cdots \lambda_{i_k} + \cdots + (-n) \prod_{i=1}^{n} \lambda_1\lambda_2 \cdots \lambda_n \]

\[ = \sum_{k=0}^{n} (n-2k)\varepsilon^{n-k} \sum_{i_1 < i_2 < \cdots < i_k} \lambda_{i_1}\lambda_{i_2} \cdots \lambda_{i_k}. \]

We will consider when \( n \) is even only; the proof for an odd \( n \) is analogous. Note that there is no \( \varepsilon^{n/2} \) term in \( g_n(\varepsilon) \), and all coefficients for the terms \( \varepsilon^{n-k} \), \( k = 0, 1, 2, \ldots, n/2 - 1 \), are positive while all coefficients for the terms \( \varepsilon^{n-k} \), \( k = n/2 + 1, n/2 + 2, \ldots, n \), are negative. Starting from \((n/2)\)-th order derivative of \( g_n(\varepsilon) \), we have

\[ g_n^{(n/2)}(\varepsilon) > 0, \text{ for all } \varepsilon > 0. \]

Furthermore, since

\[ g_n^{(n/2-1)}(0) < 0 \text{ and } \lim_{\varepsilon \to +\infty} g_n^{(n/2-1)}(\varepsilon) = +\infty, \]

21
we know $g_{n/2-1}(\varepsilon)$ passes the $\varepsilon$-axis once and only once for $\varepsilon > 0$. Again, since

$$g_{n/2-2}(0) < 0 \text{ and } \lim_{\varepsilon \to +\infty} g_{n/2-2}(\varepsilon) = +\infty,$$

the function $g_{n/2-2}(\varepsilon)$ decreases from a negative value and then increases to positive infinity. Consequently, $g_{n/2-2}(\varepsilon)$ also passes the $\varepsilon$-axis once and only once for $\varepsilon > 0$.

Continuing this procedure of analysis, we conclude that the function $g_n(\varepsilon)$ passes the $\varepsilon$-axis only once for $\varepsilon > 0$. Namely, $g_n(\varepsilon) = 0$ has only one positive solution.