Ambiguity Analysis in Source Localization with Unknown Signals

M. João D. Rendas
CAPS, Dep. Eng. Electr. e Comp.
Instituto Superior Técnico
1096 Lisboa Codex, Portugal

José M. F. Moura
LASIP, Depart. Elect. and Comp. Eng.
Carnegie Mellon University
Pittsburgh, PA, 15213

Abstract

In this paper we formulate a general definition of ambiguity function, based in the Kullback-Leibler directed divergence between probability densities. The ambiguity function here defined summarizes all the geometric aspects of the problem, measuring the difficulty in distinguishing between two different source locations. We show that by considering a particular model the classical RADAR ambiguity function is obtained. We illustrate the use of the new definition applying it to several problems, showing that it is a usefull tool for the analysis of passive location systems.

1. Introduction

There is an increasing interest in the study of the detection and location of targets in *inhomogeneous* mediums. In particular, in the context of SONAR applications, the ocean's complex *multipath* structure has been considered by many authors.

Besides the design of signal processing algorithms that can effectively deal with the peculiarities of multipath propagation, it is important to establish limits on the performance attainable in this situation. Two distinct kinds of performance analysis are usually considered: *local* and *global*. In [5], we did a local performance analysis, determining the Cramér-Rao Bound for the location of sources propagating over multipath channels. Here, we assess the problem of characterizing the expected global performance of localization mechanisms when there is lack of knowledge concerning the source signal. In the context of location problems, global performance analysis traditionally involves the computation of the *ambiguity function*. This function has been defined only for very simple models and is not adequate to the global analysis of source location in inhomogeneous mediums. Moreover, it assumes that the signal correlation function is known, which is not the case when studying passive systems.

We propose, in section 2, a general definition of ambiguity, based on the Kullback-Leibler directed divergence between probability densities [3]. This function summarizes, from the user's point of view, the geometry of the problem, measuring the difficulty in distinguishing any two points in the parameter space. Section 3 considers the particular case of normal densities. In section 4, we show how the classical definition is obtained from the generalized definition given here. Finally, section 5 considers the case of stationary observations, giving an expression in terms of the spectral densities of the involved processes, and applies it to the case of estimation of the location of a wideband stationary source propagating over a multipath channel, with observations in a single sensor.

2. Ambiguity Function

Consider a family $\mathcal{G}_\alpha$ of density functions, indexed by a parameter $\alpha \in A$:

$$\mathcal{G}_\alpha \triangleq \{p(x|\alpha), \quad \alpha \in A\}\,.$$

The Kullback-Leibler number (also called Kullback directed divergence or cross-entropy) between two members of $\mathcal{G}_\alpha$ is [3]:

$$I(\alpha_1,\alpha_2) \triangleq \mathrm{E}_{\alpha_1}\left\{\ln\frac{p(x|\alpha_1)}{p(x|\alpha_2)}\right\}.$$

This functional was introduced by Kullback [3] in the framework of information theory. Although it has some distance-like properties, it is not, in fact, a distance. As it can be easily seen, it is not symmetric and it does not satisfy, in general, the triangular inequality. However, $I(\alpha_1,\alpha_2) \geq 0$, with equality iff $\alpha_1 = \alpha_2$. Note that

$$I(\alpha_1,\alpha_2) = \mathrm{E}_{\alpha_1}\{\ln p(x|\alpha_1) - \ln p(x|\alpha_2)\}\,,$$

i.e., $I(\cdot,\cdot)$ is the mean value of the difference between the values of the log-likelihood function for two points in the parameter space, for observations x, conditioned on one

CH2977-7/91/0000-1285 $1.00 © 1991 IEEE

and following [6], with a discrete time representation, the observation vector is modelled as an N-dimensional complex Gaussian vector,

$$p(r|\theta) \sim \mathcal{N}(0, \gamma f(\theta) f(\theta)^H + \sigma^2 I) \tag{6}$$

where we assumed that the observation noise is white, and the power parameters σ^2 and γ are known and $f(\theta)$ is, for each θ a *known* N-dimensional vector. It is further assumed that $\|f(\theta)\|^2 = N$, i.e., the received signal energy is independent of θ.

Using elementary facts from linear algebra,

$$R(\theta)^{-1} = \frac{1}{\sigma^2}\left(I - \frac{\gamma}{\sigma^2 + N\gamma} f(\theta) f(\theta)^H\right),$$

where I denotes the identity matrix of order N, and $|R(\theta)| = \sigma^{2K}\left(1 + \frac{N\gamma}{\sigma^2}\right)$. Using eq. (5),

$$I(\theta_1, \theta_2) = \frac{1}{2}\frac{\gamma^2}{\sigma^2(\sigma^2 + N\gamma)}\left[N^2 - |f(\theta_2)^H f(\theta_1)|^2\right]. \tag{7}$$

Since (Schwartz's inequality) $0 \leq |f(\theta_2)^H f(\theta_1)|^2 \leq N^2$, we can determine $I_{MAX} = N^2\gamma^2/2\sigma^2(\sigma^2 + N\gamma)$, from which and (1) we get

$$\mathcal{A}(\theta_1, \theta_2) = \frac{1}{N^2}|f(\theta_2)^H f(\theta_1)|^2, \tag{8}$$

which is the classical definition of ambiguity.

Note that the model that leads to the classical definition of ambiguity belongs to the first class of estimation problems mentioned in section 2, i.e., there is no uncertainty regarding the signal parameters, in this case, the Rayleight coefficient γ, and, for each θ, the vector $f(\theta)$.

Consider now that γ *is not known*. We denote, from now on, the covariance matrix of the observed field by $R(\theta, \gamma)$. Simple calculations lead to:

$$\tilde{\gamma}(\theta_2) = \gamma\frac{|f(\theta_2)^H f(\theta_1)|^2}{N^2},$$

and,

$$\begin{aligned}\tilde{I}_\gamma(\theta_1, \theta_2) &= \frac{1}{2}\left[\frac{\gamma}{N\sigma^2}\left(N^2 - |f(\theta_2)^H f(\theta_1)|^2\right)\right. \\ &+ \left.\ln\frac{\sigma^2 + \frac{\gamma}{N}|f(\theta_2)^H f(\theta_1)|^2}{\sigma^2 + N\gamma}\right].\end{aligned}$$

Note that this expression is always smaller than the Kullback–Leibler number for known γ, given by eq. (7), i.e., the ambiguity is always larger in this case.

For each value of γ, the following bound on $\tilde{I}(\cdot,\cdot)$can be found:

$$I_{MAX}(\gamma) = \frac{1}{2}\frac{N\gamma}{\sigma^2},$$

and we get the following *conditional* measure of ambiguity

$$\begin{aligned}\mathcal{A}(\theta_1, \theta_2)_\gamma = \frac{1}{I_{MAX}(\gamma)}&\left[\frac{\gamma}{N\sigma^2}|f(\theta_2)^H f(\theta_1)|^2\right. \\ &- \left.\ln\frac{N\sigma^2 + \gamma|f(\theta_2)^H f(\theta_1)|^2}{N\sigma^2 + N^2\gamma}\right].\end{aligned}$$

A study of this function reveals that its maxima in the parameter space occur in the same points as the maxima of the classical ambiguity function. However, the relative heights of the maxima are not the same, as it is obvious from the previous equation. Note also that while in the previous case the ambiguity was independent of the signal-to-noise ratio, this is not true in this more general situation. In fact, for very high signal-to-noise ratios, the influence of the second term is negligible, and the two expressions predict the same behaviour. In the low signal-to-noise ratio limit, the second term dominates over the first.

A more radical change is obtained when, instead of known complex envelop $f(\theta)$ we know only that $f(\theta) \in \mathcal{M}_\theta$, where $\mathcal{M}_\theta$ is a known proper $P(\theta)$ dimensional subspace of $\mathbb{C}^N$. This is the relevant model for the case of location of multiple perfectly correlated narrowband sources.

Assume that $\mathcal{M}_\theta = \text{Sp}\{a_i(\theta)\}_{i=1}^{P_\theta}$, i.e.,

$$R(\theta, s) = \sigma^2 I + A(\theta) s s^H A(\theta)^H$$

where $s \in \mathbb{C}^{P_\theta}$, is an unknown vector, and $A(\theta)$ is the matrix that gathers the basis vectors $a_i(\theta)$. In this case,

$$\begin{aligned}I(\theta_1, s_1; \theta_2, s_2) = \frac{1}{2}&\left[\ln\frac{\sigma^2 + s_2^H A^H(\theta_2)A(\theta_2)s_2}{\sigma^2 + s_1^H A^H(\theta_1)A(\theta_1)s_1}\right. \\ +\frac{s_1^H A^H(\theta_1)A(\theta_1)s_1}{\sigma^2} &- \left.\frac{s_2^H A^H(\theta_2)R(\theta_1, s_1)A(\theta_2)s_2}{\sigma^2(\sigma^2 + s_2^H A^H(\theta_2)A(\theta_2)s_2)}\right].\end{aligned}$$

The vector $\tilde{s}_2(\theta_2)$ that minimizes this expression satisfies

$$A(\theta_2)\tilde{s}_2(\theta_2) = \Pi_{\mathcal{M}_{\theta_2}}\left[A(\theta_1)s_1\right],$$

where $\Pi_{\mathcal{M}_{\theta_2}}$ denotes the orthogonal projection operator into $\mathcal{M}_{\theta_2}$. Define

$$\gamma_s^2(\theta_1, \theta_2) \triangleq \|\Pi_{\mathcal{M}_{\theta_2}}\left[A(\theta_1)s\right]\|^2.$$

Then,

$$\begin{aligned}\tilde{I}_s(\theta_1, \theta_2) = \frac{1}{2}&\left[\ln\frac{\sigma^2 + \gamma_s^2(\theta_1, \theta_2)}{\sigma^2 + \|A(\theta_1)s\|^2} + \frac{\|A(\theta_1)s\|^2}{\sigma^2}\right. \\ &\left. - \frac{\gamma_s^2(\theta_1, \theta_2)}{\sigma^2}\right].\end{aligned}$$

In this case, since $0 \leq \gamma_s^2(\theta_1, \theta_2) \leq \|A(\theta_1)s\|^2$, we can find the bound

$$I_{MAX}(\theta_1)_s = \frac{1}{2}\frac{\|A(\theta_1)s\|^2}{\sigma^2}.$$

of those points. The value of $I(\cdot,\cdot)$ depends, naturally, on the size of the observation interval. Here, we consider only the assymptotic case of very long observation interval.

Heuristically, $I(\alpha_1, \alpha_2)$ is a measure of the resemblance, or closseness, of the two models described by $p(x|\alpha_1)$ and $p(x|\alpha_2)$. The values of α_2 that yield small values of $I(\alpha_1, \alpha_2)$ indicate possible erroneous estimates of α when the true value of the parameter is α_1. It can be shown, [3], that the Kullback-Leibler number is the exponent of the minimum error probability of the binary decision of α_2 against α_1, when α_1 is the true value of the parameter and the error probability under α_2 is held constant.

Based on these arguments, we define ambiguity between two points (α_1, α_2) in the parameter space as

$$\mathcal{A}(\alpha_1, \alpha_2) \triangleq \frac{I_{MAX}(\alpha_1) - I(\alpha_1, \alpha_2)}{I_{MAX}(\alpha_1)} \tag{1}$$

where $I_{MAX}(\alpha_1)$ denotes an upper bound on the value of $I(\alpha_1, \alpha_2)$ over $\alpha_2 \in A$. Since $I(\cdot,\cdot)$ is not symmetric, $\mathcal{A}(\alpha_1, \alpha_2)$ willl not be, in general, a symmetric function of its two arguments.

In the context of source location with unknown source signals, the density of the observations is parametrized by two distinct sets of parameters: those that describe the source location ($\theta \in \Theta$) and those describing the source signal, or its statistics ($\gamma \in \Gamma$). Let α denote the complete set of parameters:

$$\alpha = [\theta, \gamma], \qquad \theta \in \Theta, \gamma \in \Gamma.$$

Define the following sub-families of $\mathcal{G}_\alpha$:

$$\mathcal{G}_\theta^\gamma \triangleq \{p(x|\theta,\gamma), \theta \in \Theta\}, \quad \mathcal{G}_\gamma^\theta \triangleq \{p(x|\theta,\gamma), \gamma \in \Gamma\}.$$

Note that in the case of *known source signal*, the data is modeled by $\mathcal{G}_\theta^\gamma$, where γ is the actual value of the source parameters.

Let $\tilde{\gamma}(\theta_2)$ determine the member of $\mathcal{G}_\gamma^{\theta_2}$ closest to $p(x|\theta_1, \gamma)$:

$$I((\theta_1,\gamma),(\theta_2,\gamma_2)) \geq I((\theta_1,\gamma),(\theta_2,\tilde{\gamma}(\theta_2))), \qquad \gamma_2 \in \Gamma$$

and define

$$\tilde{I}_\gamma(\theta_1,\theta_2) \triangleq I((\theta_1,\gamma_1),(\theta_2,\tilde{\gamma}(\theta_2))).$$

We define ambiguity between two points in the space of the parameter of interest, Θ, conditioned on the value of the unwanted parameter γ, in the following way:

$$\mathcal{A}(\theta_1,\theta_2)_\gamma \triangleq \frac{I_{MAX}(\theta_1)_\gamma - \tilde{I}_\gamma(\theta_1,\theta_2)}{I_{MAX}(\theta_1)_\gamma}, \tag{2}$$

where $I_{MAX}(\theta)_\gamma$ is a bound on the value of $\tilde{I}$:

$$I_{MAX}(\theta)_\gamma \geq \tilde{I}_\gamma(\theta,\theta_2).$$

This definition reflects the central issue that distinguishes the situation of known and unknown signal, namely, the necessity of estimating the signal parameters in order to estimate the source location.

Since $\tilde{I}_\gamma(\theta_1,\theta_2) \leq I((\theta_1,\gamma),(\theta_2,\gamma))$, we conclude that the presence of unwanted parameters can only increase the ambiguity, as it should be expected.

From this conditional definition of ambiguity, we can derive global measures of ambiguity, independent of the particular point γ in the space of unwanted parameters:

$$\mathcal{A}(\theta_1,\theta_2) = A_\gamma\left\{\mathcal{A}(\theta_1,\theta_2)_\gamma\right\}. \tag{3}$$

where $A_\gamma\{\cdot\}$ is an operator on Γ. The definition of the operator $A_\gamma\{\cdot\}$ can be done in different ways, leading to different global characterizations of ambiguity. For instance, we can define it to be a mean value operator, or, alternatively, search for the pairs ($\theta_1 \neq \theta_2$) where it takes the maximum value (worst case analysis).

3. Normal densities

In this section, we give the expression of the Kullback-Leibler number for normal densities. We consider the problems of parametrized mean and covariance.

In the first one, denote by μ_θ the mean value, and by R the *fixed* covariance matrix. It is easily seen that

$$I(\theta_1,\theta_2) = \frac{1}{2}(\mu_{\theta_1} - \mu_{\theta_2})^T R^{-1}(\mu_{\theta_1} - \mu_{\theta_2}). \tag{4}$$

The Kullback-Leibler distance is in fact a distance and we can identify the model manifold with the N-dimensional Hilbert space with metric defined by the covariance matrix R.

In the case of normal observations with information on the covariance, $R = R_\theta$, a few lines of algebra show that the Kullback-Leibler number is :

$$I(\theta_1,\theta_2) = \frac{1}{2}\left[\mathrm{tr}(R_{\theta_2}^{-1}R_{\theta_1}) - N - \ln\left|R_{\theta_2}^{-1}R_{\theta_1}\right|\right]. \tag{5}$$

In this case, $I(\cdot,\cdot)$ does not have the properties of a distance. In fact, as it is easily seen, it is not even symmetric.

4. The Classical Definition

The Woodward ambiguity function was introduced in the context of active RADAR systems, for the problem of simultaneous estimation of delay and Doppler shift in a narrowband signal of ***known*** complex envelope, transmitted through a Rayleigh channel. Let θ denote the vector that gathers the wanted parameters, i.e., the delay τ and the Doppler shift, ω. In complex notation, From the expressions given above, we conclude that the role of the classical ambiguity function is here played by the projection of $\mathcal{A}(\theta_1)s$ into $\mathcal{M}_{\theta_2}$.

Note that while in the previous case the location of the peaks of the ambiguity function was independent of the unwanted parameter γ, in this case this is not so. Different directions of the complex vector s will result in different shapes of the conditional ambiguity function. In an effort to globally characterize the problem, we can define a *mean* ambiguity function, dependent only on the location parameters:

$$\overline{\mathcal{A}(\theta_1,\theta_2)} = \frac{1}{S_{S_{P(\theta_1)}}} \int_{S_{P(\theta_1)}} \mathcal{A}(\theta_1,\theta_2)_s ds,$$

where $S_{P(\theta_1)}$ denotes the unit spherical surface in $C^{P(\theta_1)}$, and $S_{S_{P(\theta_1)}}$ its area. This approach leads to the use of the principal angles between the subspaces $\mathcal{M}_{\theta_1}$ and $\mathcal{M}_{\theta_2}$, as defined in [2]. We do not pursue it here due to lack of space, referring the interested reader to [4].

5. Stationary scalar observations

We treat here the case of normal stationary zero-mean scalar observations, with spectral density function $S_\theta(\omega), \theta \in \Theta$. Our results are based on the work [1]. Under the assumption that the observed series has a strictly positive spectral density $S_\theta(\omega)$, the Kullback - Leibler number can be written as

$$I(\theta_1,\theta_2) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left[\ln\frac{S_{\theta_2}(\omega)}{S_{\theta_1}(\omega)} - 1 + \frac{S_{\theta_2}(\omega)}{S_{\theta_1}(\omega)}\right]d\omega.$$

For the specific case of a stationary wideband source at location θ, with spectral density $S(\omega)$ propagating to a single sensor over a channel with transfer function $H_\theta(\omega)$, the received signal spectral density is, for white observation noise,

$$S_\theta(\omega) = \sigma^2 + S(\omega)|H_\theta(\omega)|^2.$$

For *known source signal spectral density*, the ambiguity in the estimation of the location parameter θ involves the following Kullback-Leibler number:

$$I_{S(\omega)}(\theta_1,\theta_2) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left[\ln\frac{\sigma^2+S(\omega)|H_{\theta_2}(\omega)|^2}{\sigma^2+S(\omega)|H_{\theta_1}(\omega)|^2} - 1 + \frac{\sigma^2+S(\omega)|H_{\theta_1}(\omega)|^2}{\sigma^2+S(\omega)|H_{\theta_2}(\omega)|^2}\right]d\omega$$

which depends on $S(\omega)$. Using a variational approach, we determine the spetral density function that minimizes $I_{S(\omega)}(\cdot,\cdot)$, subject to the restriction of fixed source power, obtaining

$$S_{MIN}^{\theta_1,\theta_2}(\omega) = \frac{C}{\left(|H_{\theta_2}(\omega)|^2 - |H_{\theta_1}(\omega)|^2\right)^2} \quad (9)$$

for all the points ω where $|H_{\theta_2}(\omega)|^2 \neq |H_{\theta_1}(\omega)|^2$, and where C is a constant. Note that eq. (9) corresponds to having maximum power in the frequencies where the values of the medium's power transfer function are more close to each other, which is an intuitively pleasing result.

For the case of *unknown source spectrum*, the relevant distance is

$$\mathrm{I}(\theta_1,\mathrm{S}_1,\theta_2,\mathrm{S}_2) = \frac{1}{4\pi}\int_{\pi}^{\pi}\left[\ln\frac{\sigma^2+\mathrm{S}_2(\omega)|H_{\theta_2}(\omega)|^2}{\sigma^2+\mathrm{S}_1(\omega)|H_{\theta_1}(\omega)|^2} - 1 + \frac{\sigma^2+\mathrm{S}_1(\omega)|H_{\theta_1}(\omega)|^2}{\sigma^2+\mathrm{S}_2(\omega)|H_{\theta_2}(\omega)|^2}\right]d\omega.$$

Imposing no constraints on $S(\omega)$ other than it being a non-negative function leads to the following:

$$\tilde{S}_2(\omega:\theta_2)|H_{\theta_2}(\omega)|^2 = S_1(\omega)|H_{\theta_1}(\omega)|^2,$$

which, in turn, implies $\tilde{I}_{S_1}(\theta_1,\theta_2) \equiv 0$. We conclude thus, that with a single sensor, we must be able to impose conditions on the shape of the spectral density of the source to be able to estimate the source location.The local version of this result had already been derived using the Cramer-Rao lower bound expressions, in [5]. Without no prior knowledge about the source spectrum, the model is not informative.

References

[1] K. Dzhaparidze. *Parameter Estimation and Hypotheses Testing in Spectral Analysis of Stationary Time Series.* Springer Series in Statistics, Springer–Verlag, 1986.

[2] Gene H. Golub and Charles F. Van Loan. *Matrix Computations.* John Hopkins, 1989.

[3] Solomon Kullback. *Information Theory and Statistics.* Peter Smith, 1978.

[4] M. João D. Rendas and José M. F. Moura. Ambiguity analysis of source location systems with nuisance parameters. *To be submitted*, 1991.

[5] M. João D. Rendas and José M. F. Moura. Cramer-Rao bounds for source location systems in multipath environments. *To appear in the IEEE Trans. on Acoustic, Speech and Signal Processing*, Dec. 1991.

[6] H.L. Van Trees. *Detection, Estimation, and Modulation Theory, Part III.* John Willey & Sons, Inc., 1971.