A Geometrical Interpretation of Signal Detection and Estimation

Bernard Picinbono

A Geometrical Interpretation of Signal Detection and Estimation

Bernard Picinbono, *Senior Member, IEEE*

The pdf copy of the final published text can be obtained from the author at the following address : bernard.picinbono@lss.supelec.fr

Reference of this text : IEEE Trans. Inf. Theory, Vol. 26, Num. 4 pp. 493 - 497, 1980.

Abstract

By introducing an appropriate representation of the observation, detection problems may be interpreted in terms of estimation. The case of the detection of a deterministic signal in Gaussian noise is associated with two orthogonal subspaces: the first is the signal subspace which is generally one dimensional and the second is called a reference noise alone (RNA) space because it contains only the noise component and no signal.The detection problem can then be solved in the signal subspace, while the use of the RNA space is reduced to the estimation of the noise in the signal subspace.This decomposition leads to a very simple interpretation of singular detection, even in the non-Gaussian case, in terms of perfect estimation. The method is also extended to multiple signal detection problems and to some special cases of detection of random signals.

Manuscript received

B. P. is with the Laboratoire des Signaux et Systèmes (L2S), a joint laboratory of the C.N.R.S. and the École Supérieure d'Électricité, Plateau de Moulon, 3 rue Joliot–Curie 91192, Gif sur Yvette, France. The L2S is associated with the University of Paris–Sud, France. E-mail: bernard.picinbono@lss.supelec.fr, .

I. Introduction

Detection and estimation are among the most important topics in the problems of signal processing. These problems are presented extensively in many textbooks [l], [2], and there are now a great number of efficient solutions, at least under the classical assumptions such as Gaussian noise. The aim of this correspondence is to give a new presentation of the classical detection problem by using an appropriate geometrical interpretation. The idea of a geometrical interpretation has already been used in detection problems, particularly the projection method in an appropriate reproducing kernel Hilbert space [3]-[5]. In the case of white noise this idea is the basis of the geometrical interpretation of many communications problems [6], [7]. The method presented here is very different from the previous ones and has many advantages, the most important being the following.

a) It permits a very simple physical interpretation of the relation between detection and estimation in the case of deterministic signals.

b) It makes systematic use of the notion of the reference noise alone (RNA) space which appears to be fundamental in the development of adaptive detection theory [8].

Moreover the problem of singular detection has a very simple interpretation, even in the non-Gaussian case. This correspondence is mostly concerned with the presentation of the general ideas and their physical interpretation. For this purpose we restrict the discussion to discrete time signals and a finite-dimensional observation space. The extension to continuous time signals changes only the mathematical aspects of the theory; the general results are the same.

II. Detection, Estimation, and Reference Noisie Alone

We first briefly recall the classical results concerning the detection of a deterministic signal in Gaussian noise. The observation is a vector $\mathbf{x}$ belonging to an observation space of finite dimension N. To any orthonormal bases $\mathbf{u}_i$ of this space we can associate the expansion

$$\mathbf{x} = \sum_{i=1}^{N} x_i \mathbf{u}_i, \tag{1}$$

which defines the components x_i of the vector $\mathbf{x}$. We suppose that the noise is a zero-mean Gaussian vector characterized by its covariance matrix

$$\mathbf{\Gamma} \triangleq E[\mathbf{x}\mathbf{x}^T], \tag{2}$$

which is assumed to be positive definite. The signal is a deterministic vector $\mathbf{s}$, and in this case the likelihood ratio is a monotonically varying function of the test function

$$T(\mathbf{x}) = \mathbf{s}^T \mathbf{\Gamma}^{-1} \mathbf{x}. \tag{3}$$

This function is linearly dependent on the observation and can be considered as the output of a matched filter [1, p. 122]. The test function can also be written as

$$T(\mathbf{x}) = \sigma^T \mathbf{x}, \tag{4}$$

when

$$\sigma = \mathbf{\Gamma}^{-1} \mathbf{s}. \tag{5}$$

In the case of a continuous time observation $x(t)$ and signal $s(t)$ on an interval ΔT the corresponding expressions are [1, p. 118], [2, p. 300-301]

$$T(x) = \int_{\Delta T} \sigma(\theta) x(\theta) d\theta, \tag{6}$$

where

$$\int_{\Delta T} \Gamma(t, \theta) \sigma(\theta) d\theta = s(t). \tag{7}$$

The mathematical problems concerning the validity of these expressions have already been extensively discussed [9].

For the following discussion we decompose the observation space into two orthogonal subspaces. The first, H_s, is spanned by a unit vector $\mathbf{u}_1$, in the signal direction:

$$\mathbf{u}_1 = s^{-1} \mathbf{s} \tag{8}$$

where s is the energy of the signal, or

$$s^2 = \mathbf{s}^T \mathbf{s} = \sum_1^N s_i^2. \tag{9}$$

This space is called the *signal subspace*. The second, $H_\perp$, is orthogonal to H_s, and is called the RNA space. There is no signal component in this space, and thus the projection of the observation vector $\mathbf{x}$ in this space has exactly the same value under the hypotheses H_0 (noise alone) and H_1 (signal plus noise). This space is assumed to have a basis consisting of the orthonormal vectors $\mathbf{u}_2$, $\mathbf{u}_2$, ..., $\mathbf{u}_N$.

If the noise $\mathbf{n}$ is white the projections $\mathbf{n}_1$ and $\mathbf{n}_2$ of $\mathbf{n}$ on H_s and $H_\perp$ are independent, and thus the space $H_\perp$ is irrelevant for the detection of $\mathbf{s}$. That is the basic idea used in the geometric representation of communication problems [6], [7]. But if $\mathbf{n}_2$ is correlated with $\mathbf{n}_1$, this correlation must be used in the detection process, as we now discuss.

The observation is always given by (2.1), but the component x_1 has now a particular meaning. Indeed

$$x_1 = \mathbf{u}_1^T \mathbf{x} = (1/s) \mathbf{s}^T \mathbf{x}. \tag{10}$$

and $\mathbf{s}^T\mathbf{x}$ is the output of the matched filter in the case of white noise (WNMF). Clearly we have the same results in the case of continuous time signals where the vectors $\mathbf{u}_i$ are time functions $u_i(t)$ and N is in general infinite.

Now let us write all the vectors in terms of the particular bases for H_s and $H_\perp$. The observation vector x is

$$\mathbf{x}^T = [x_1 \vdots x_2, \ldots, x_N] = [x_1 \vdots \mathbf{x}_2^T], \tag{11}$$

where x_1 is the first component of $\mathbf{x}$ and $\mathbf{x}_2$ is the vector of $H_\perp$, with the components $x_2, x_3, \ldots, x_N$. The signal vector is evidently

$$\mathbf{s}^T = [s \vdots \mathbf{0}^T], \tag{12}$$

and the noise vector

$$\mathbf{n}^T = [n_1 \vdots \mathbf{n}_2^T]. \tag{13}$$

(Since the signal subspace is one dimensional there is no difference between the vectors $\mathbf{x_1}$, $\mathbf{s}_1$, $\mathbf{n}_1$ and the components x_1, s_1, n_1.)

The covariance matrix given by (2) can be partitioned in the form

$$\mathbf{\Gamma} = \begin{pmatrix} \mathbf{\Gamma}_1 & \vdots & \mathbf{\Gamma}_{12} \\ \ldots & \ldots & \ldots \\ \mathbf{\Gamma}_{21} & \vdots & \mathbf{\Gamma}_2 \end{pmatrix}, \tag{14}$$

where $\mathbf{\Gamma}_{ij} = E[\mathbf{n}_i\mathbf{n}_j^T]$, with $i,\ j = 1$ or 2. Evidently $\mathbf{\Gamma}_{12} = \mathbf{\Gamma}_{21}^T$. In order to calculate the test function (3) we must decompose the matrix $\mathbf{\Gamma}_{-1}$ with the same partitioning. After elementary calculations we obtain

$$\mathbf{\Gamma} = \begin{pmatrix} \mathbf{A} & \vdots & \mathbf{C} \\ \ldots & \ldots & \ldots \\ \mathbf{C}^T & \vdots & \mathbf{B} \end{pmatrix}, \tag{15}$$

with

$$\mathbf{A} = [\mathbf{\Gamma}_1 - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_2^{-1}\mathbf{\Gamma}_{21}]^{-1}, \tag{16}$$

$$\mathbf{B} = [\mathbf{\Gamma}_2 - \mathbf{\Gamma}_{21}\mathbf{\Gamma}_1^{-1}\mathbf{\Gamma}_{12}]^{-1}, \tag{17}$$

$$\mathbf{C} = -\mathbf{A}\mathbf{\Gamma}_{12}\mathbf{\Gamma}_2^{-1}. \tag{18}$$

These relations are valid for any symmetric positive matrix $\mathbf{\Gamma}$ (even when $\mathbf{\Gamma}_1$ is not a scalar).

The matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are directly connected with linear mean-square (LMS) estimation, and this point is fundamental for the following. Consider for example the LMS estimation of the vector $\mathbf{n}_1$ in terms of $\mathbf{n}_2$ By application of the projection theory it is well-known [10] that the LMS estimate is

$$\hat{\mathbf{n}}_1 = \mathbf{\Gamma}_{12}\mathbf{\Gamma}_2^{-1}\mathbf{n}_2. \tag{19}$$

The variance matrix is the correlation matrix of the error

$$\tilde{\mathbf{n}}_1 = \mathbf{n}_1 - \hat{\mathbf{n}}_1, \tag{20}$$

and can be written

$$\epsilon^2 \stackrel{\triangle}{=} E[\tilde{\mathbf{n}}_1\tilde{\mathbf{n}}_1^T] = \mathbf{\Gamma}_{11} - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_2^{-1}\mathbf{\Gamma}_{21}. \tag{21}$$

By comparing with the previous equations we see that

$$\hat{\mathbf{n}}_1 = -\mathbf{A}^{-1}\mathbf{C}\mathbf{n}_2 \tag{22}$$

and

$$\epsilon^2 = \mathbf{A}^{-1}. \tag{23}$$

Let us now apply these general results to our particular case where $\mathbf{n}$ is given by (13), which implies that n_1, and A are scalars. In this case ϵ^2 is the minimum mean-square error in the estimation of the noise in the signal subspace H_s in terms of the noise in the RNA subspace $H_\perp$. Moreover if we introduce the quantity

$$d^2 = \mathbf{s}^T\mathbf{\Gamma}^{-1}\mathbf{s} \tag{24}$$

we obtain

$$A = \frac{1}{\epsilon^2} = \frac{d^2}{s^2}\ . \tag{25}$$

At this point we recall that the problem of regular or singular detection is completely specified by d^2 [1, p. 121], [2, p. 99]. Thus this problem is also one of regular or singular estimation. More precisely singular detection occurs when $\epsilon^2 = 0$, which means physically that the noise in the signal subspace can be estimated without error, and thus suppressed, only by consideration of the noise in the RNA space. This point will be discussed more precisely in the following.

As n_1 is a scalar we can write (19) as

$$\hat{n}_1 = \mathbf{h}^T\mathbf{n}_2 \tag{26}$$

and from (22) we have

$$\mathbf{h}^T = -A^{-1}\mathbf{C} = -(s^2/d^2)\mathbf{C}. \tag{27}$$

Let us consider the test function given by (4) and (5). The vector σ^T which is equal to $\mathbf{s}^T\mathbf{\Gamma}^{-1}$ can be calculated from (12) and (15), and we obtain

$$\sigma^T = [As\vdots\mathbf{C}s] \tag{28}$$

where As is a scalar. Thus it can also be written as

$$\sigma^T = As[1\vdots A^{-1}\mathbf{C}] = (d^2/s)[1\vdots -\mathbf{h}^T]. \tag{29}$$

We deduce from (4) that the test function is

$$T(\mathbf{x}) = (d^2/s)(x_1 - \mathbf{h}^T\mathbf{x}_2). \tag{30}$$

But $H_\perp$ is an RNA space, since it is orthogonal to the signal. Thus the component $\mathbf{x}_2$ of the observation is only a noise component $\mathbf{n}_2$ and by using (26) we deduce the final and simpler expression for the test function

$$T(\mathbf{x}) = (d^2/s)(x_1 - \hat{n}_1). \tag{31}$$

This structure means that the optimal detection of a deterministic signal in Gaussian noise can be decomposed into two different operations:

1) projection of the observation on the signal subspace: this can be carried out by the WNMF,

2) estimation of the noise component in this signal space in terms from the noise in the RNA space: this can be done in many ways well-known in the engineering literature.

Thus we see the strong connection between detection and estimation problems, a point which has been already widely discussed in the case of random signals, but in a completely different way.

III. Structure and Characteristic Properties of the Optimal Receiver

In this section we will discuss more carefully the consequences of the structure of the test function given by (31). First we notice that the factor d^2/s is irrelevant and can also be associated with the value of the threshold t with which $T(\mathbf{x})$ is compared in order to make the optimal decision. Thus this decision can also be made by comparing the test function

$$T'(\mathbf{x}) = x_1 - \hat{n}_1 \tag{32}$$

with a fixed threshold t'. This value clearly depends on the allowable false alarm probability.

Let us now indicate how this test function can be deduced from the observation $\mathbf{x}$. In particular we clearly see the separation of the matched filtering for white noise and the estimation of the RNA signal

$\mathbf{x}_2$. The advantage of this structure is that in adaptive detection problems, it is possible to adapt the estimator even in the presence of a signal, because the signal never appears in the RNA channel.

Let us now consider the case where the input noise is white. In this case it is clear than n_1 and $\mathbf{n}_2$ are uncorrelated, and we have $\hat{n}_1 = 0$. That means that the estimation procedure is useless, and therefore the receiver is reduced to the WNMF. Thus the estimator channel is only of interest in the case where the input noise is colored. In order to specify the receiver which computes the test function $T'(\mathbf{x})$, it is interesting to calculate the statistical properties of this function. As the noise is Gaussian and the receiver linear, $T'(\mathbf{x})$ is a Gaussian random variable under the two hypotheses H_0 and H_1. In the absence of a signal obviously we have $E[T'|H_0] = 0$. Moreover in this case

$$T'(\mathbf{x}) = n_1 - \hat{n}_1 = \tilde{n}_1, \tag{33}$$

and the variance of $T'(\mathbf{x})$ is the mean-square error of the estimate of n_1. By comparing with (23) and (25), we can write

$$\text{Var}\ T'(\mathbf{x}) = \epsilon^2 = A^{-1} = s^2/d^2. \tag{34}$$

In the presence of a signal the variance is the same, and the mean value is

$$E[T'(\mathbf{x})|H_1] = [x_1|H_1] = s. \tag{35}$$

Thus the probability densities of $T'(\mathbf{x})$ under H_0 and H_1 are two Gaussian curves of the same variance ϵ^2 and separated by a distance s. So the separation between the curves is due only to the signal while the estimation modifies their width.

In this representation we see very clearly why the performance of the optimal receiver is improved by reducing the estimation error for the same signal.

As noticed before, the detection becomes singular if $\epsilon^2 = 0$, i.e., if it possible to estimate the noise in the signal subspace from the observation without error.

At this point in the discussion it is interesting to investigate the possible extension to the non-Gaussian case. We decompose the observation vector by the same procedure as in Section II, and the likelihood ratio can also be written as

$$L(\mathbf{x}) = \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} = \frac{p(x_1 - s, \mathbf{x}_2)}{p(x_1, \mathbf{x}_2)} \tag{36}$$

because there is no signal component outside the signal subspace. By introducing the a posteriori probability density $p(x_1|\mathbf{x}_2)$, we obtain

$$L(\mathbf{x}) = \frac{p(x_1 - s|\mathbf{x}_2)}{p(x_1|\mathbf{x}_2)}\ . \tag{37}$$

In particular we see from this expression that the a priori distribution of the observation in the RNA space, $\mathbf{x}_2$, does not play any role in the detection problem.

Let us now explore more precisely the problem of singular detection in the-non-Gaussian context. First we suppose that the linear estimate of n_1 in terms of an RNA observation $\mathbf{x}_2$ is singular. This means that

$$\epsilon^2 = E[\tilde{n}_1^2] = E[(n_1 - \hat{n}_1)^2] = 0, \tag{38}$$

and so $\tilde{n}_1$ is nonrandom. As the test function $T'(\mathbf{x})$ can be written as

$$T'(\mathbf{x}) = s\delta_{0i} + \tilde{n}_1 \tag{39}$$

under the hypothesis H_i, $i = 0, 1$, we deduce directly that the detection is also singular because the two hypotheses can be separated from the observation without error.

Conversely suppose that the detection problem is singular for every value of the amplitude s of the signal. That means that for every s and $\mathbf{x}_2$ the probability densities $p(x_1 - x|\mathbf{x}_2) = p(x_1|\mathbf{x}_2)$ as functions of x_1 are nonoverlapping. Thus we have

$$p(x_1|\mathbf{x}_2) = \delta[x_1 - f(\mathbf{x}_2)] \tag{40}$$

where

$$f(\mathbf{x}_2) = E[x_1|\mathbf{x}_2] = \hat{x}_1 \tag{41}$$

which is the best mean-square estimate of x_1 in terms of $\mathbf{x}_2$, in general nonlinear. But (40) shows that the variance of $\hat{x}_1$, is zero, which means that the estimation problem is singular in the sense that there is zero mean-square error. This shows directly the connection between singular detection and singular estimation.

All these results can be easily extended to the case where the signal belongs to a subspace of dimension higher that one (multiple signals detection) and even in some cases of random signals. These extensions clearly show the strong connection between detection ad estimation problems in the sense analyzed in this paper.

References

[1] C. W. Helstrom, *Statistical Theory of SignaI Detection*, Oxford: Pergamon, 1968.

[2] H. L. Van Trees, *Estimation and Modulation Theory, Part I*, New York: Wiley, 1968.

[3] E. Parzen, *Times Series Analysis Papers*, San Francisco: Holden Day, 1967, p. 477.

[4] T. Kailath, "A projection method for signal detection in colored Gaussian noise," *IEEE Trans. Information Theory*, IT-13, pp. 441-447, July 1967.

[5] T. Kailath, “RKHS approach to detection and estimation problems - Part I: Deterministic signals in Gaussian noise,’ *IEEE Trans. Information Theory*, vol. IT-17, pp. 530-549, Sept. 1971.

[6] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering.*, New York: Wiley, 1967.

[7] D. Sakrison, *Communication Theory: Transmission of Waweforms and Digital Information*, New York: Wiley, 1968.

[8] B. Picinbono, “Adaptive signal processing for detection and communication in communication systems and random process theory,’ in *Nato Advanced Study Institute*, J. K. Skwirzynski, Ed., Sijthoff and Noordhoff 1978. pp. 639-660.

[9] T. T. Kadota, “Differentiation of Karhunen-Loève expansion and application to optimum reception of sure signals in noise, *IEEE Trans. Information Theory*, vol. IT-13, pp. 255-260, Apr. 1967.