DETECTION OF THE NUMBER OF SIGNALS USING PREDICTIVE STOCHASTIC COMPLEXITY

Shahrokh Valaee [1] and Peter Kabal [1,2]

[1] Electrical Engineering
McGill University
Montreal, Quebec, H3A 2A7

[2] INRS-Télécommunications
Université du Québec
Verdun, Quebec, H3E 1H6

Abstract

In this paper, we propose a new algorithm for the processing of signals by an array of sensors. The objective is to find the number and the Directions Of Arrival (DOA) of signals impinging on a linear array. The Predictive Stochastic Complexity (PSC) criterion of Rissanen is used to select the best model order. To reduce the computational load, the algorithm operates with a sub-optimal estimator while maintaining the consistency of the estimator. The proposed method is on-line and can be utilized in time-varying systems for target tracking. The method can be used for both correlated and uncorrelated signals.

1. Introduction

In this paper we address the problem of finding the number and Directions Of Arrival (DOA) of signals impinging on a linear array. Suppose we have a linear omnidirectional array of p sensors which are exposed to $q < p$ far-field sources. The signals from the sources can be partially or fully correlated. The fully correlated case, also called the *coherent* case, stems from the multipath propagation or smart jamming and is of practical importance in signal processing. We assume narrowband signals with known center frequency. Signals arrive at the array from distinct directions $\theta_1, \ldots, \theta_q$. The objective is to find the number of sources, q, and their directions of arrival, $\theta^q = [\theta_1, \ldots, \theta_q]$.

The approaches taken so far are based on Akaike's Information Criterion (AIC) and Rissanen's Minimum Description Length (MDL) [1,2,3]. These two criteria are described by

$$\text{AIC}(N) = -\log f(\mathbf{x}^N, \hat{\theta}_N) + k, \qquad (1)$$

$$\text{MDL}(N) = -\log f(\mathbf{x}^N, \hat{\theta}_N) + \frac{k}{2}\log N, \qquad (2)$$

where $\mathbf{x}^N$ is the observation up to time N, $\hat{\theta}_N$ is the Maximum Likelihood (ML) estimate of the parameter vector and k is the number of free elements of the parameter vector. In these formulae, the first term is the log likelihood function of the observation. The generating model class is shown by f.

Asymptotic study of the AIC estimator indicates that it is not consistent. It tends to over-estimate the true order of the system as the number of observation increases. The advantage of Rissanen's MDL is its consistency. In both methods, buffering of data is essential. Thus, neither of them can be used on-line. In this paper, we develop an algorithm based on Rissanen's Predictive Stochastic Complexity (PSC), [4,5]. This algorithm is consistent and its structure makes it suitable for on-line use. The PSC is based on the concept of predictive coding. Rissanen proves in [5] that this criterion achieves the shortest code length of the data relative to the generating model class. The concept of the codelength minimization stems from the fact that the best model which fits to the data is the one that gives the most information about it. The Minimum Description Length (MDL) is also based on minimization of the codelength with a special method of coding. For the MDL, the restriction of the coding method increases the codelength. This increase in the codelength makes us skeptical about its optimality.

The predictive stochastic complexity of the observation vector $\mathbf{x}_i$, $i = 1, \ldots, N$, is defined by

$$\text{PSC}(N) = -\sum_{i=1}^{N} \log f(\mathbf{x}_i, \hat{\theta}_{i-1}), \qquad (3)$$

where $\hat{\theta}_{i-1}$ is the ML estimate of the parameter vector with respect to the observations up to time $(i-1)$. The shortcoming of the PSC is its high computational complexity. In this paper we modify the criterion by using a sub-optimal recursive estimator. We will discuss the strength of the estimator. We have also proved that the estimator is consistent, but this issue will not be discussed in the present paper. By using this estimator, the multivariate nonlinear optimization is decomposed into several one-variable nonlinear problems. The simulation results show that with the new estimator, PSC performs better than MDL.

The paper is organized as follows. The problem is formulated in the next section. The concept of subspace decomposition which is widely used in array processing is discussed in Section 3. Detection and estimation parts of the problem are the issues which are described in Sections 4 and 5. A recursive estimator that can be utilized to reduce the computational complexity is also introduced in Section 5. In Section 6, we present the simulation results and compare two methods, MDL and PSC.

2. Problem Formulation

Let us define the observation vector ($p \times 1$) by $\mathbf{x}(t)$, the signal vector ($q \times 1$) by $\mathbf{s}(t)$ and the noise vector ($p \times 1$) by $\mathbf{n}(t)$. For the linear multisensor problem the following equation holds,

$$\mathbf{x}(t) = \mathbf{A}(\theta)\mathbf{s}(t) + \mathbf{n}(t), \qquad (4)$$

where $\mathbf{A}(\theta) = [\mathbf{a}(\theta_1) \ldots \mathbf{a}(\theta_q)]$ is the $p \times q$ matrix of steering vectors, and $\mathbf{a}(\theta_i) = [1 \ \mu_i \ \mu_i^2 \ \ldots \ \mu_i^{p-1}]^T$, where $\mu_i = \exp(j\omega_0 d \sin(\theta_i)/c)$. In this formula, the distance between two consecutive sensors is represented by d, the wave speed is shown by c and the center frequency of the source signal is given by ω_0. It is further assumed that every p collection of the steering vectors $\mathbf{a}(\theta_i)$, $i = 1, \ldots, p$ is independent, for all possible θ_i's.

It is also assumed that the signal samples are i.i.d. sequence of complex Gaussian random vectors with an unknown covariance matrix $\mathbf{S}$. The noise samples are i.i.d. sequence of Gaussian random vectors with unknown covariance matrix $\sigma^2\mathbf{I}$ and are

V-345

independent of the signal samples. With these assumptions, the observations will be the samples of a complex Gaussian process with zero mean. The probability density function of this process is given by

$$f(\mathbf{x}|\theta^q) = \frac{1}{\pi^p \, \det \mathbf{R}^q} \exp\{-\mathbf{x}^H[\mathbf{R}^q]^{-1}\mathbf{x}\}, \qquad (5)$$

where the correlation matrix of observation vector, is defined by,

$$\mathbf{R}^q = E[\mathbf{x}\mathbf{x}^H|\theta^q]. \qquad (6)$$

It is straightforward to show that,

$$\begin{aligned} \mathbf{R}^q &= \mathbf{A}(\theta^q)\mathbf{S}\mathbf{A}^H(\theta^q) + \sigma^2\mathbf{I}, \\ &= \mathbf{Q} + \sigma^2\mathbf{I}. \end{aligned} \qquad (7)$$

The correlation matrix $\mathbf{R}^q$ is a function of q, θ^q, σ^2 and $\mathbf{S}$. It might be thought that the unknown parameters of $\mathbf{R}^q$ can be found by jointly maximizing the likelihood function. However, the maximum value of the likelihood function is an increasing function of the order of the system. In other words, direct ML estimation always gives the maximum value for the number of sources, q. For this reason, in the AIC and MDL a second term is added to the likelihood function to penalize the over-estimation.

It is assumed that $q < p$, which yields $q \in \mathcal{P} = \{0, 1, \ldots, p-1\}$. For any $k \in \mathcal{P}$, based on an observed sequence, an appropriate model of order k can be constructed. In order selection methods, p models run simultaneously. The corresponding model for a given $k \in \mathcal{P}$ is described by

$$\mathbf{x}(t) = \mathbf{A}(\theta^k)\mathbf{s}^k(t) + \mathbf{n}(t), \qquad (8)$$

where $\mathbf{A}(\theta^k)$ is the parameterized steering matrix with respect to the parameter vector θ^k, and $\mathbf{s}^k(t)$ represents the signal vector with dimension k. The probability density function of the observation vector for model k is given by

$$f(\mathbf{x}|\theta^k) = \frac{1}{\pi^p \, \det \mathbf{R}^k} \exp\{-\mathbf{x}^H[\mathbf{R}^k]^{-1}\mathbf{x}\}, \qquad (9)$$

where $\mathbf{R}^k$ is the correlation matrix of the observation vector and

$$\mathbf{R}^k = E[\mathbf{x}\mathbf{x}^H|\theta^k]. \qquad (10)$$

For each model, a quantity which can be AIC or MDL is computed. The best value of q is obtained by minimizing this quantity. In this work we apply the PSC criterion.

The Predictive Stochastic Complexity for model k, at time instant N, is computed as

$$\mathrm{PSC}_k(N) = -\sum_{i=1}^{N} \log f(\mathbf{x}_i|\hat{\theta}^k_{i-1}), \qquad (11)$$

$$= \sum_{i=1}^{N} (\log \det \hat{\mathbf{R}}^k_{i-1} + \mathbf{x}_i^H[\hat{\mathbf{R}}^k_{i-1}]^{-1}\mathbf{x}_i), \qquad (12)$$

where $\hat{\mathbf{R}}^k_{i-1}$ is the ML estimate of the correlation matrix based on the observations up to time $(i-1)$ and with respect to the system of order k. The correlation matrix is estimated from the sample correlation matrix

$$\bar{\mathbf{R}}_{i-1} = \frac{1}{i-1}\sum_{l=1}^{i-1} \mathbf{x}_l\mathbf{x}_l^H. \qquad (13)$$

Most of the high resolution methods for the estimation of the angles of arrival use the subspace decomposition method. This issue is of great importance in array processing. In the following section we apply this concept to our problem.

3. Subspace Decomposition

Let us assume that the p-dimensional complex vector space is represented by $\mathcal{C}^p$. In the subspace decomposition approach, this vector space is split into two orthogonal subspaces. For a given model of order k, the columns of the matrix $\mathbf{A}(\theta^k)$ span a sub-space of $\mathcal{C}^p$ called the *signal subspace*. An estimate of the signal subspace can be obtained through estimation of the angles of arrival, $\hat{\theta}^k$. The signal subspace for the model k at time instant i is shown by $\mathcal{C}^k_{s_i}$. This is the column span of the matrix $\mathbf{A}(\hat{\theta}^k_i)$. The dimension of this subspace is determined by the number of sources. For a model of order k dimension of the signal subspace is equal to k. The *noise subspace*, $\mathcal{C}^k_{n_i}$, is the orthogonal space to the signal subspace. The dimensionality of the noise subspace for a model of k signals is $(p-k)$. The true signal and noise subspaces are represented by $\mathcal{C}^q_s$ and $\mathcal{C}^q_n$, respectively. It should be noted that once the signal subspace is given, the corresponding DOA can be determined uniquely.

Decomposition of the observation vector, $\mathbf{x}$, into its components in the two subspaces yields,

$$\begin{aligned} \mathbf{x} &= \mathbf{G}\begin{bmatrix} \mathbf{x}_{ss} \\ \mathbf{x}_{nn} \end{bmatrix}, \\ &= \mathbf{x}_s + \mathbf{x}_n, \end{aligned} \qquad (14)$$

where $\mathbf{G}$ is a unitary coordinate transformation and

$$\mathbf{x}_s = \mathbf{G}\begin{bmatrix} \mathbf{x}_{ss} \\ 0 \end{bmatrix}, \qquad \mathbf{x}_n = \mathbf{G}\begin{bmatrix} 0 \\ \mathbf{x}_{ss} \end{bmatrix}. \qquad (15)$$

We present a lemma which proves the uniqueness of the decomposition.

Lemma 1. *For noncoherent sources, a fixed θ^q and independent steering vectors, $\mathbf{a}(\theta_i)$, $i = 1, \ldots, q$, the subspace decomposition is unique.*

Proof: It is possible to show that for every space, $\mathcal{C}^q = \mathrm{span}\{\mathbf{a}(\theta_i); \ i = 1, \ldots, q\}$, there is a unique set, $\theta = \{\theta_i; \ i = 1, \ldots, q\}$, that generates this space and vice versa. Due to the uniformity of the array and the far–field sources the basis is Vandermonde. Hence the steering vectors $\mathbf{a}(\theta_i)$ are independent. We claim that this basis is unique. If it is not the case, there will be another set, $\mathcal{A}(\psi) = \{\mathbf{a}(\psi_i); i = 1, \ldots, q\}$, with Vandermonde structure, which is also a basis for $\mathcal{C}^q$. Then every vector in $\mathcal{C}^q$ could be represented with respect to this basis. Take for instance $\mathbf{a}(\theta_1)$. The independence of the steering vectors suggests that θ_1 has to be in the set $\psi = \{\psi_i; \ i = 1, \ldots, q\}$. The continuation of this argument implies $\psi = \theta$. ∎

It is also worth noting that almost surely for $i \geq p$ the sample correlation matrix, $\bar{\mathbf{R}}_i$, is of full rank. Thus, the columns of $\bar{\mathbf{R}}_{i-1}$, $i > p$, span $\mathcal{C}^p$.

Lemma 2. *For noncoherent sources the column span of $\mathbf{A}(\theta^q)$ coincides with the span of the eigenvectors of $\mathbf{R}^q$ which correspond to the q largest eigenvalues.*

This is a standard result (see [6].)

The projection matrices onto the two subspaces are given by,

$$\mathbf{P}_s(\theta^k) = \mathbf{A}(\theta^k)\left(\mathbf{A}^H(\theta^k)\mathbf{A}(\theta^k)\right)^{-1}\mathbf{A}^H(\theta^k), \qquad (16)$$

$$\mathbf{P}_n(\theta^k) = \mathbf{I} - \mathbf{P}_s(\theta^k). \qquad (17)$$

The projection of the sample correlation matrix onto the signal and noise subspaces yields,

$$\bar{\mathbf{R}}^k_{s_i} = \mathbf{P}_s(\theta^k)\bar{\mathbf{R}}_i\mathbf{P}_s(\theta^k), \qquad (18)$$

$$\bar{\mathbf{R}}^k_{n_i} = \mathbf{P}_n(\theta^k)\bar{\mathbf{R}}_i\mathbf{P}_n(\theta^k). \qquad (19)$$

The assumptions for the noise vector makes this representation possible. The noise samples are generated by a white process.

Thus, the components of the noise in the noise subspace are independent of the components of the noise and the signal in the signal subspace. It is then possible to find a unitary transformation $\mathbf{G}$, which decomposes the correlation matrix into two independent parts in the orthogonal subspaces. Thus the following structure can be constructed for $\mathbf{R}^q$,

$$\mathbf{R}^q = \mathbf{G}\begin{bmatrix} \mathbf{R}^q_{ss} & 0 \\ 0 & \mathbf{R}^q_{nn} \end{bmatrix}\mathbf{G}^H, \qquad (20)$$

where square matrix $\mathbf{R}^q_{ss}$ has q eigenvalues presented by $\lambda_1, \ldots, \lambda_q$ and $\mathbf{R}_{nn}$ has an eigenvalue, σ^2, with multiplicity $(p - q)$.

An alternative formulation for the correlation matrix is possible based on the subspace decomposition. Specifically, it is possible to split $\mathbf{R}^q$ into two parts in the signal and noise subspaces,

$$\mathbf{R}^q = \mathbf{R}^q_s + \mathbf{R}^q_n, \qquad (21)$$

where

$$\mathbf{R}^q_s = \mathbf{P}_s(\theta^q)\mathbf{R}^q\mathbf{P}_s(\theta^q), \qquad (22)$$

$$\mathbf{R}^q_n = \mathbf{P}_n(\theta^q)\mathbf{R}^q\mathbf{P}_n(\theta^q). \qquad (23)$$

Uniqueness of the decomposition (21) implies that

$$\mathbf{R}^q_s = \mathbf{G}\begin{bmatrix} \mathbf{R}^q_{ss} & 0 \\ 0 & 0 \end{bmatrix}\mathbf{G}^H, \qquad (24)$$

$$\mathbf{R}^q_n = \mathbf{G}\begin{bmatrix} 0 & 0 \\ 0 & \mathbf{R}^q_{nn} \end{bmatrix}\mathbf{G}^H, \qquad (25)$$

Thus, maximum likelihood estimation of the correlation matrix can be based on the estimation of two correlation matrices in the signal and noise subspaces. We shall discuss this issue in the next section.

4. Detection

The problem which was described earlier can be split into *detection* and *estimation* parts. The detection problem concerns finding the true number of sources, q. The determination of the DOA of the signals is the goal of the estimator. In the proposed algorithm, on arrival of each sample, detection and estimation are done consecutively. If the number of sources is fixed, the problem collapses to an estimation procedure. The algorithm includes p models running in parallel. The PSC criterion for all p cases is computed. At each time instant, the task is to choose the best model through minimization of the PSC.

Taking determinant of both sides of (20), recall that $\mathbf{G}$ is unitary, we get

$$\det(\mathbf{R}^q) = \det(\mathbf{R}^q_{ss})\det(\mathbf{R}^q_{nn}). \qquad (26)$$

Correspondingly, the sample correlation matrix is written as

$$\bar{\mathbf{R}}_{i-1} = \mathbf{G}\begin{bmatrix} \bar{\mathbf{R}}_{ss_{i-1}} & \bar{\mathbf{R}}_{sn_{i-1}} \\ \bar{\mathbf{R}}_{ns_{i-1}} & \bar{\mathbf{R}}_{nn_{i-1}} \end{bmatrix}\mathbf{G}^H. \qquad (27)$$

As i increases, $\bar{\mathbf{R}}_{i-1}$ tends to $\mathbf{R}^q$ and the matrices $\bar{\mathbf{R}}_{ss_{i-1}}$ and $\bar{\mathbf{R}}_{nn_{i-1}}$ approach $\mathbf{R}_{ss}$ and $\mathbf{R}_{nn}$, respectively. In such a case, $\det\hat{\mathbf{R}}^k_{i-1}$ can be decomposed into two parts in the signal and noise subspaces, i.e.

$$\det(\hat{\mathbf{R}}^k_{i-1}) = |\hat{\mathbf{R}}^k_{s_{i-1}}| \; |\hat{\mathbf{R}}^k_{n_{i-1}}|, \qquad (28)$$

where $|\mathbf{B}|$ is the multiplication of the nonzero eigenvalues of the matrix $\mathbf{B}$. Since $\mathbf{x}_s$ and $\mathbf{x}_n$ are independent,(see (21)),

$$\hat{\mathbf{R}}^k_{i-1} = \hat{\mathbf{R}}^k_{s_{i-1}} + \hat{\mathbf{R}}^k_{n_{i-1}}. \qquad (29)$$

The PSC is a function of the parameter vector $\hat{\theta}^k_{i-1}$. The PSC criterion uses the ML estimate. If $\hat{\theta}^k_{i-1}$ is the ML estimate of the true DOA, it can be shown [7] that,

$$\hat{\mathbf{R}}^k_{s_{i-1}} = \bar{\mathbf{R}}^k_{s_{i-1}}, \qquad (30)$$

$$|\hat{\mathbf{R}}^k_{n_{i-1}}| = (\hat{\sigma}^2(\hat{\theta}^k_{i-1}))^{p-k}, \qquad (31)$$

where,

$$\hat{\sigma}^2(\hat{\theta}^k_{i-1}) = \frac{1}{p-k}\mathrm{tr}\bar{\mathbf{R}}^k_{n_{i-1}}. \qquad (32)$$

The projected sample correlation matrices $\bar{\mathbf{R}}^k_{s_{i-1}}$ and $\bar{\mathbf{R}}^k_{n_{i-1}}$ are given by (18) and (19), respectively. With ML estimation, it is also possible to show that $\hat{\mathbf{R}}^k_{n_{i-1}}$ has a single eigenvalue of order $(p-k)$ which can be calculated from (32). The eigenvectors of $\hat{\mathbf{R}}^k_{n_{i-1}}$ and $\mathbf{R}^k_{n_{i-1}}$ are identical. Thus $\hat{\mathbf{R}}^k_{n_{i-1}}$ can be obtained by a linear transformation $\mathbf{T}^k_i$, on the matrix $\bar{\mathbf{R}}^k_{n_{i-1}}$. Being more specific, we propose the following estimator,

$$\hat{\mathbf{R}}^k_{n_{i-1}} = \mathbf{T}^k_{i-1}\bar{\mathbf{R}}^k_{n_{i-1}}, \qquad (33)$$

$$\mathbf{T}^k_{i-1} = \mathbf{V}_{p-k}\mathrm{diag}[\frac{\hat{\sigma}^2(\hat{\theta}^k_{i-1})}{\lambda_l}]\mathbf{V}^H_{p-k}, \qquad (34)$$

where λ_l, $l = 1, \ldots, p-k$, are non-zero eigenvalues of $\bar{\mathbf{R}}_{n_{i-1}}$ and $\mathbf{V}_{p-k}$ is the $p \times (p-k)$ matrix of the corresponding eigenvectors. The ML estimate of the noise power $\hat{\sigma}^2(\hat{\theta}^k_{i-1})$ is given by (32).

Therefore the PSC criterion will be equal to

$$\mathrm{PSC}_k(N) = \sum_{i=1}^{N}[\log|\bar{\mathbf{R}}^k_{s_{i-1}}| + (p-k)\log(\hat{\sigma}^2(\hat{\theta}^k_{i-1}))+$$
$$\mathbf{x}^H_i(\bar{\mathbf{R}}^k_{s_{i-1}} + \mathbf{T}^k_{i-1}\bar{\mathbf{R}}^k_{n_{i-1}})^{-1}\mathbf{x}_i]. \qquad (35)$$

It is seen that the computation of PSC relies on the estimation of the angles of arrival, $\hat{\theta}^k_{i-1}$. In the following section we shall propose an algorithm to estimate $\hat{\theta}^k_{i-1}$.

5. Estimation

To obtain $\hat{\theta}^k_{i-1}$ an estimation problem has to be solved. A suitable choice for the cost function is given by

$$\hat{\theta}^k = \arg\min_{\theta}\{\log(\hat{\sigma}^2(\hat{\theta}))\}, \qquad (36)$$

$$= \arg\min_{\theta} \mathrm{tr}[\mathbf{P}_n(\hat{\theta})\bar{\mathbf{R}}_i]. \qquad (37)$$

If the dimensionality of the signal subspace is fixed, the trace of the projected matrix is minimum if and only if the projection is done onto the space spanned by those eigenvectors of $\mathbf{R}_i$ which correspond to the smallest eigenvalues. In other words, minimization of the criterion (36) for a model of order k is the same as choosing the eigenvectors corresponding to the largest k eigenvalues of $\bar{\mathbf{R}}_i$ as the basis for the projection operator, $\mathbf{P}_s(\hat{\theta}^k)$. From the observation vectors, we can compute the sample correlation matrix, $\bar{\mathbf{R}}_i$. Therefore according to Lemma 1 and Lemma 2, minimization of the criterion (36) gives a unique vector of angles of arrival and a subspace which is an estimate of the actual signal subspace. Because of the nature of (36) we call it the *noise estimator*.

It should be noted that for stochastic signals, the ML estimator is given by

$$\hat{\theta}^k = \arg\min_{\theta}\{\log[|\bar{\mathbf{R}}^k_s| \; (\hat{\sigma}^2(\hat{\theta}))^{p-k}]\}. \qquad (38)$$

As a comparison of the two cost functions, (36) and (38), we considered two sources and computed the costs for them. The results are shown in Fig. 1 and Fig. 2. Both criteria have global optimum points on the true values of the angles of arrival. The ML algorithm is played by prominent local minima which can cause problems with hill climbing optimization procedures.

Estimation of the angles of arrival is a multivariate nonlinear optimization problem. To reduce the computational complexity we introduce a recursive algorithm which is very similar to the Alternating Maximization method of Ziskind and Wax [8].
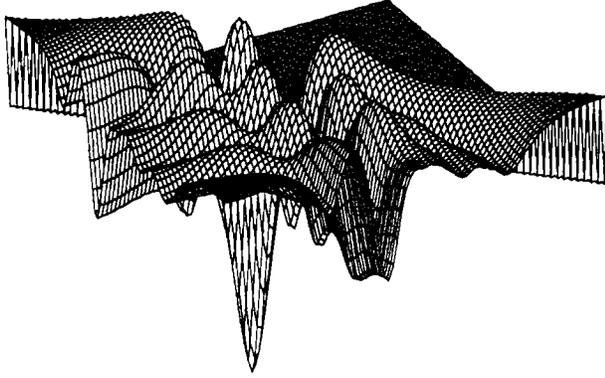
Fig. 1 The cost for the ML estimator as a function of the angles θ_1 and θ_2.
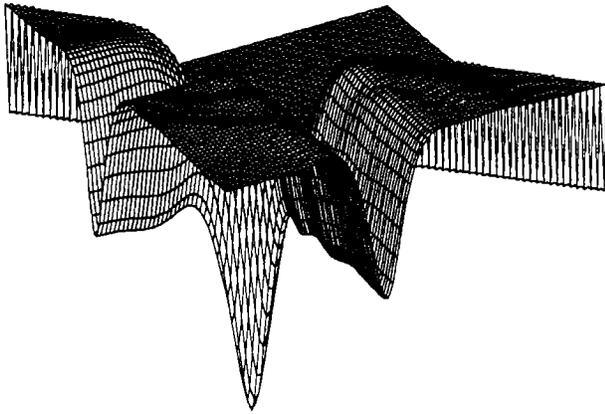

Fig. 2 The cost for the noise estimator as a function of the angles θ_1 and θ_2.

In the new algorithm, optimization is done with respect to one element of the parameter vector, keeping the rest of the elements constant. Therefore, ML estimation is decomposed into several one-variable nonlinear optimization problems.

To give more insight into the algorithm, we present a simple example. Suppose we are dealing with a model of order 3. We use the value of $\hat{\theta}_{i-1}^{(3)}$ to determine $\hat{\theta}_i^{(3)}$.

Let us assume $\hat{\theta}_{i-1}^{(3)} = [a_{i-1}\ b_{i-1}\ c_{i-1}]^T$. To obtain $\hat{\theta}_i^{(3)}$, a minimization problem is solved 3 times to get a_i, b_i and c_i. At each step we use the latest computed values of the elements of the parameter vector. In other words, optimization is accomplished in the following manner,

$$a_i = \arg\min_{a \in D_a} C(a, b_{i-1}, c_{i-1})$$

$$b_i = \arg\min_{b \in D_b} C(a_i, b, c_{i-1})$$

$$c_i = \arg\min_{c \in D_c} C(a_i, b_i, c)$$

where C is the cost function defined earlier. $D_a = [-\frac{\pi}{2}, b_{i-1}]$, $D_b = [a_i, c_{i-1}]$ and $D_c = [b_i, \frac{\pi}{2}]$ are the intervals of search. It is seen that for a model of order k, the cost function is optimized k times to get $\hat{\theta}^k$.

Once the optimum value of $\hat{\theta}_{i-1}^k$ is determined, the PSC can be computed. The computer simulations, which are given in the following section, show that the method works in both correlated and uncorrelated cases.

SNR	uncorrelated sources		correlated sources	
(dB)	MDL	PSC	MDL	PSC
1	80	68	11	7
2	70	40	7	4
3	57	39	6	3
4	44	30	9	5
6	23	15	9	7
8	20	0	9	1
10	2	0	6	3

Table 1 Percent failure to detect for MDL and PSC

6. Simulation Results

A configuration with 6 omnidirectional sensors is investigated. The spacing between two consecutive sensors is taken to be equal to half the wave length. In the first experiment there are three uncorrelated sources. These sources are in the far field and at the angles 0, 10 and 20 degrees. The signals have unit power and the SNR is defined as $10\log(\frac{1}{2\sigma^2})$. We perform 100 Monte-Carlo runs and count the number of times that MDL and PSC, fail to detect the true order of the system. The data window is 10 samples. We have chosen a small width for the data window to emphasize the differences between the two methods. The performances of both methods for large data windows are almost equal. The results for uncorrelated and correlated sources are given in Table 1. For correlated case, a scenario of two sources in the angles 10 and 20 degrees is considered. The phase difference between the signals at the reference point of the array is equal to -30 degrees. The superiority of the PSC is clear from the data for the uncorrelated case. However, in correlated situation the differences are less dramatic.

7. Summary

In this paper we introduced a new method for the detection of the number of signals, using a linear array. Our approach is based on Rissanen's Predictive Stochastic Complexity. To apply PSC for array processing, we utilized the concept of subspace decomposition. We also used a hill climbing method for the estimation of the direction of arrivals. This method reduces the computational complexity involved in the use of PSC. Simulation results show that the PSC performs better than the MDL in both correlated and uncorrelated cases.

References

1. M. Wax, "Detection of coherent and noncoherent signals via the stochastic signal model," *ICASSP 91*, pp. 3541–3544.
2. M. Wax and I. Ziskind, "Detection of the number of coherent signals by the MDL principle," *IEEE Trans. on ASSP*, vol. 37, pp. 1190–1196, August 1989.
3. M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. on ASSP*, vol. 33, pp. 387–392, April 1985.
4. J. Rissanen, "Stochastic complexity and modeling," *Annals of statistics*, vol. 14, pp. 1080–1100, 1986.
5. J. Rissanen, "*Stochastic complexity in statistical inquiry*," World Scientific Publisher, 1989.
6. R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. on Antennas and Prop.* vol. AP–34, pp. 276–280, March 1986.
7. T. W. Anderson, "Asymptotic theory for principal component analysis," *Ann. J. Math. Stat.*, vol. 34, pp. 122–148, 1963.
8. I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. on ASSP*, vol. 36, pp. 1553–1560, Oct. 1988.