СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ AIC, MDL КРИТЕРИЕВ В ЗАДАЧЕ ОБНАРУЖЕНИЯ МНОГОМЕРНЫХ СИГНАЛОВ В СЛУЧАЕ КОРОТКОЙ ВЫБОРКИ

Ермолаев В.Т., Мальцев А.А., Родюшкин К.В.

Государственный университет Нижнего Новгорода,
603600, Н.Новгород, пр. Гагарина, 23, каф. бионики и статистической радиофизики.

В настоящее время существует множество методов оценивания числа источников сигналов, принимаемых антенной решеткой (АР). Наиболее известными из них являются методы, основанные на критериях AIC (Akaike Information Criterion) и MDL (Minimum Description Length). Исследованию статистических характеристик этих методов посвящено достаточно большое число работ [1,2]. Однако полученные в этих работах результаты справедливы для неизвестной мощности собственного шума и найденные теоретические оценки носят асимптотический характер. Анализ характеристик этих методов для случая коротких выборок, как правило, проводится путем численного моделирования. В данной работе находятся статистические характеристики AIC и MDL методов, для случая известного уровня собственного шума АР. Показано, что при такой постановке задачи характеристики обнаружения многомерных сигналов этими методами, могут быть рассчитаны на основе аналитических формул полученных авторами для распределения максимального собственного числа выборочной корреляционной матрицы гауссовских входных сигналов АР [3]. Получены аналитические выражения для расчета вероятности ложной тревоги обнаружения сигналов, справедливые для любой длины выборки.

Предположим, что на узкополосную N-элементную антенную решетку (АР), приходят простанственно-когерентные сигналы от K статистически независимых внешних источников. Тогда вектор сигналов $\vec{x}$ на выходах элементов АР можно представить в следующем виде:

$$\vec{x} = \sum_{i=1}^{K} a_i \vec{s}_i + \vec{n}\,, \tag{1}$$

где $\vec{x}$ [1x*N*] - вектор выходных сигналов, $\vec{s}_i$ - [1x*N*] вектор волнового фронта, а a_i - комплексная гауссовская амплитуда сигнала от i-ого источника, $\vec{n}$ - [1x*N*] вектор собственного шума элементов АР. Будем считать, что собственный шум в элементах АР не коррелирован и является комплексным гауссовским белым шумом с нулевым средним значением и единичной мощностью в каждом элементе. Задача оценки числа сигналов для модели (1) заключается в оценке числа K по заданной конечной выборке $\vec{x}_i$, $i = 1..L$ входного вектора $\vec{x}$ с помощью некоторого статистически обоснованного алгоритма или критерия.

В настоящее время наиболее известными критериями оценивания числа сигналов являются критерии AIC и MDL [1]. Согласно критериям AIC и MDL оценкой числа сигналов $\hat{K}$ является величина, которая минимизирует функцию $AIC(k)$ или соответственно $MDL(k)$.

$$AIC(k) = -2\max L_k + 2T(k)\,, \qquad MDL(k) = -\max L_k + 0.5\log(L)T(k)\,, \tag{2),(3}$$

где $\max L_k$ - максимум логарифма функции правдоподобия, при фиксированном числе сигналов k, $T(k)$ - число независимых параметров (например, неизвестные мощности источников, углы прихода сигналов и т.п.), определяющих модель (1).

Для рассматриваемой модели гауссовского входного сигнала АР (1) логарифм функции правдоподобия можно записать в следующем виде.

$$L_k = -LSp(M^{-1}\hat{M}) - L\log|M| - NL\log\pi \tag{4}$$

Здесь использовались следующие обозначения: $M = \langle \vec{x}\vec{x}^+ \rangle$ - корреляционная матрица вектора сигналов, $\hat{M} = \frac{1}{L}\sum_{i=1}^{L}\vec{x}_i\vec{x}_i^{\,+}$ - выборочная корреляционная матрица.

Запишем корреляционную матрицу M и выборочную корреляционную матрицу $\hat{M}$ в виде разложений:

$$M = \sum_{i=1}^{N}\lambda_i\vec{\varphi}_i\vec{\varphi}_i^{\,+}, \quad \hat{M} = \sum_{i=1}^{N}\hat{\lambda}_i\hat{\vec{\varphi}}_i\hat{\vec{\varphi}}_i^{\,+}\,. \tag{5}$$

Здесь $\lambda_i, \vec{\varphi}_i$ и $\hat{\lambda}_i, \hat{\vec{\varphi}}_i$ - упорядоченные собственные числа и собственные векторы матриц M и $\hat{M}$ соответственно. С использованием (5) логарифм функции правдоподобия можно записать в следующем виде.

$$L_k = -L\sum_{i=1}^{N}\lambda_i^{-1}\vec{\varphi}_i^{\,+}\hat{M}\vec{\varphi}_i - L\sum_{i=1}^{N}\log\lambda_i - NL\log\pi\,. \tag{6}$$

Как видно, данное выражение состоит из трех слагаемых. Первое слагаемое зависит как от собственных чисел, так и от собственных векторов. Второе зависит только от собственных чисел, а третье слагаемое является константой, независящей от параметров модели. Для нахождения максимума всего выражения рассмотрим сначала, при каких собственных векторах $\vec{\varphi}_i$ достигается максимум первого слагаемого и соответственно всего выражения (6). Данное слагаемое представляет собой взвешенную сумму квадратичных форм. Очевидно, что такая сумма достигает своего минимума или максимума тогда, когда собственные вектора $\vec{\varphi}_i$ одним из N! способов совпадают с собственными векторами выборочной корреляционной матрицы $\hat{M}$, т.е. когда $\vec{\varphi}_i = \hat{\vec{\varphi}}_{s(i)}$, где $s(i)$ одна из N! перестановок. Таким образом, первое слагаемое в (6) может быть записано следующим образом:

$$-L\sum_{i=1}^{N}\lambda_i^{-1}\vec{\varphi}_i^{+}\hat{M}\vec{\varphi}_i = -L\sum_{i=1}^{N}\lambda_i^{-1}\hat{\lambda}_{s(i)} \tag{7}$$

Поскольку собственные числа упорядочены, то очевидно, что максимум этого выражения будет при $s(i) = i$. Следовательно, логарифм функции правдоподобия, максимизированный по собственным векторам $\vec{\varphi}_i$ корреляционной матрицы M, записывается в следующем виде.

$$L_k = -L\sum_{i=1}^{N}(\lambda_i^{-1}\hat{\lambda}_i + \log\lambda_i) - NL\log\pi \tag{8}$$

Данное выражение содержит сумму слагаемых, каждое из которых зависит от определенного собственного числа λ_i. Поэтому для нахождения максимума всего выражения следует найти максимум каждого слагаемого. При поиске максимума отдельного слагаемого необходимо учесть ограничения, налагаемые на собственные числа матрицы M. Для рассматриваемой модели входных сигналов (1), с учетом того, что мощность собственного шума каждого элемента АР считается априори заданной и равна единице, эти ограничения имеют вид.

$$\lambda_i > 1, i = 1..K \qquad \lambda_i = 1, i = K+1..N \tag{9}$$

Легко показать, что с учетом ограничений (9) максимум выражения (8) равен:

$$\max L_k = -L\sum_{i=1}^{\min(k,l)}(1 + \log\hat{\lambda}_i - \hat{\lambda}_i) - L\sum_{i=1}^{N}\hat{\lambda}_i - NL\log\pi\ , \tag{10}$$

где l - число собственных чисел выборочной корреляционной матрицы $\hat{M}$, превышающих единицу.

Число независимых параметров $T(k)$, входящих в выражения (2) и (3) в рассматриваемом случае равно

$$T(k) = k(2N-k)\ . \tag{11}$$

Таким образом, с учетом полученных выражений (10) и (11), функция (2) принимает следующий вид

$$AIC(k) = 2L\sum_{i=1}^{\min(k,l)}(1 + \log\hat{\lambda}_i - \hat{\lambda}_i) + 2k(2N-k) + 2L\sum_{i=1}^{N}\hat{\lambda}_i + 2NL\log\pi \tag{12}$$

Для нахождения вероятности переоценки числа сигналов методом AIC рассмотрим величину приращения $\Delta AIC(k) = AIC(k+1) - AIC(k)$ функции $AIC(k)$.

Поскольку AIC-критерий заключается в поиске минимума функции $AIC(k)$, то вероятность переоценки числа сигналов можно записать как сумму вероятностей следующих событий.

$$P\{\hat{K} > K\} = P\{\Delta AIC(K) < 0\} + P\{\Delta AIC(K+1) < -\Delta AIC(K) \mid \Delta AIC(K) \ge 0\} + ...$$

Можно показать, что величина слагаемых в этой сумме быстро убывает, поэтому будем считать, что

$$P\{\hat{K} > K\} = P\{\Delta AIC(K) < 0\}$$

Используя (12), выразим $\Delta AIC(K)$ через собственные числа выборочной корреляционной матрицы.

$$\Delta AIC(K) = -2(1 + \log\hat{\lambda}_{K+1} - \hat{\lambda}_{K+1}) + 2(2(N-K)-1)$$

Поскольку функция $\log\hat{\lambda}_{K+1} - \hat{\lambda}_{K+1}$ является монотонной, то можно утверждать, что вероятность переоценки числа сигналов для метода AIC равна вероятности превышения $K+1$ собственным числом выборочной корреляционной матрицы некоторого порога $h_{AIC}(K+1)$, значение которого вычисляется путем решения алгебраического уравнения.

$$\begin{gathered} P_{AIC}\{\hat{K} > K\} = P\{\hat{\lambda}_{K+1} > h_{AIC}(K+1)\} \\ h_{AIC}(K+1) - \log h_{AIC}(K+1) = 1 + \frac{2(N-K)-1}{L} \end{gathered} \tag{13}$$

Аналогичным образом можно получить выражение для вероятности переоценки числа сигналов и для

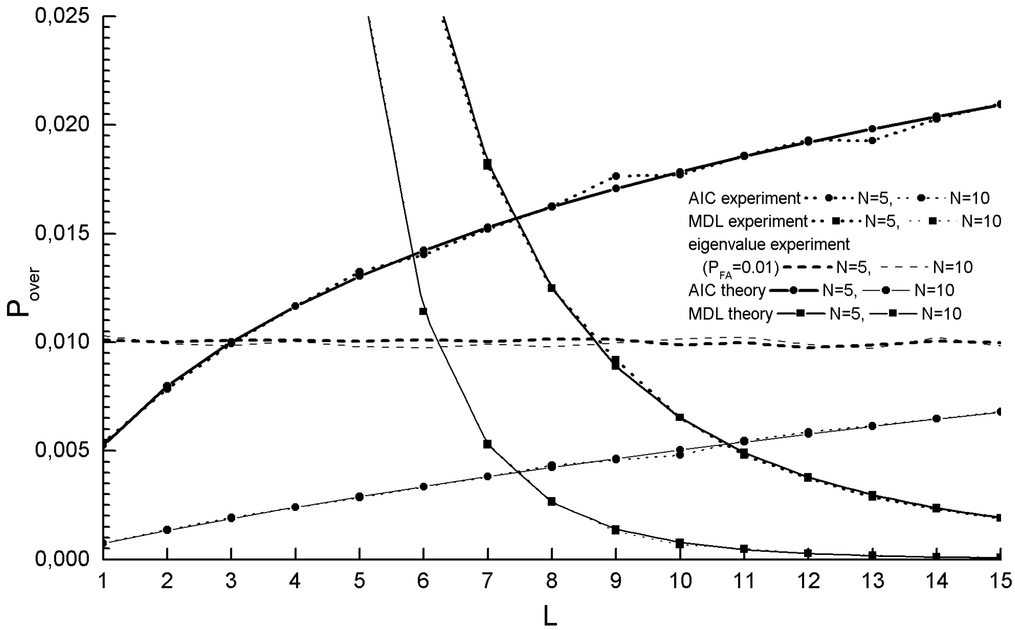

Рис.1 Эксперементальная и теоретическая вероятности ложной тревоги для методов AIC, MDL и метода основанного на анализе первого собственного числа, для N=5,10.

метода MDL.

$$P_{MDL}\{\hat{K} > K\} = P\{\hat{\lambda}_{K+1} > h_{MDL}(K+1)\}$$
$$h_{MDL}(K+1) - \log h_{MDL}(K+1) = 1 + \frac{2(N-K)-1}{L}\frac{\log L}{2} \qquad (14)$$

В работе [3] для случая когда число выборок L не меньше числа элементов решетки N получена аналитическая формула для интегральной функции распределения максимального собственного числа выборочной корреляционной матрицы в отсутствии внешних источников ($K=0$),

$$F(N,L,\lambda) = \det\left[\frac{\gamma(L-N+i+j-1, L\lambda)}{\Gamma(L-N+i)\Gamma(j)}\right], \quad i,j=1..N, \ L \ge N \qquad (15)$$

При $L < N$ можно показать, что интегральная функция распределения максимального собственного числа выборочной корреляционной матрицы примет следующий вид:

$$F(N,L,\lambda) = \det\left[\frac{\gamma(N-L+i+j-1, L\lambda)}{\Gamma(N-L+i)\Gamma(j)}\right], \quad i,j=1..L, \ L < N \qquad (16)$$

Таким образом, при решении простой задачи обнаружения внешнего сигнала ($K=0$) с помощью AIC и MDL критериев, вероятности переоценки числа сигналов (вероятности ложного обнаружения) легко выражаются через функции (15) и (16):

$$P_{AIC,MDL}\{\hat{K} > 0 \mid K=0\} = P\{\hat{\lambda}_1 > h_{AIC,MDL}(1) \mid K=0\} = 1 - F\left(N, L, h_{AIC,MDL}(1)\right), \qquad (17)$$

где пороги h_{AIC} и h_{MDL} для собственных чисел выборочной матрицы вычисляются в соответствии с уравнениями (13) и (14).

Нетрудно увидеть, что вероятность переоценки числа сигналов и для AIC и для MDL критериев не является постоянной, а зависит от значений параметров N и L. Если потребовать, чтобы система обнаружения сигнала имела постоянный уровень ложных тревог (рассчитывалась по критерию Неймана-Пирсона) необходимо определять порог h_{EIG} для максимального собственного числа в соответствии с уравнением

$$P_{FA} = 1 - F(N, L, h_{EIG}), \qquad (18)$$

исходя из заданной вероятности ложного обнаружения P_{FA}. В этом случае, вероятность ложной тревоги $P\{\hat{K} > 0 \mid K=0\}$ не будет зависеть ни от числа элементов решетки N, ни от числа выборок L.

Для проверки полученных результатов, было проведено численное моделирование методов обнаружения сигнала на основе критериев AIC и MDL для случая $K=0$. В результате моделирования получены экспериментальные зависимости вероятности ложного обнаружения сигнала от числа выборок L (см. рис. 1) для различного числа элементов АР (тонкие штриховые линии *N*=5, толстые штриховые линии *N*=10). На этом же рисунке приведены теоретические зависимости вероятности ложного обнаружения (показанные соответственно сплошными линиями), рассчитанные по формулам (13), (14), (15) и (16). Как видно из приведенных графиков, имеет место достаточно точное совпадение теоретических и экспериментальных результатов. На рисунке также представлена зависимость вероятности ложного обнаружения сигнала для метода, основанного на сравнении максимального собственного числа выборочной корреляционной матрицы с порогом, вычисленным с использованием формул (15), (16) и уравнения (18) для $P_{FA}=0.01$. Как видно, в отличие от методов AIC и MDL данный метод обеспечивает постоянный уровень ложных тревог.

Работа выполнена при поддержке грантов Российского фонда фундаментальных исследований № 00-02-17602, № 00-15-96620 и INTAS №96-2352.

Литература

1. M.Wax and T.Kailath, “Detection of signals by information theoretic criteria” //IEEE Trans. Acoust. Speech, Signal Processing, vol. ASSP-33, pp387-392, Apr. 1985.

2. W.Xu, M.Kaveh, “Analysis of the performance and the sensitivity of eigendecomposition-based detectors.” IEEE Transaction on signal processing vol. 43, No 6, June 1995.

3. В.Т.Ермолаев, К.В.Родюшкин "Функция распределения максимального собственного числа выборочной корреляционной матрицы собственного шума элементов антенной решетки" Известия вузов. Радиофизика. 1999г. №5 том 42, стр. 494.

STATISTICAL CHARACTERISTICS OF AIC AND MDL CRITERIONS IN MULTIVARIATE SIGNALS DETECTION TASK IN CASE OF SHORT SAMPLE

Ermolaev V.T., Maltsev A.A., Rodyushkin K.V.

N.Novgorod state University,
603600, N.Novgorod, Gagarina 23, department of statistical radiophysics.

Now there are many methods of estimating number of sources of signals received by antenna array. Most known methods are those based on AIC (Akaike Information Criterion) and MDL (Minimum Description Length) criterions. There are many works [1], [2] related to investigation of the statistical characteristics of these methods. However, the results obtained in these works are suitable for unknown noise power and theoretical estimates found have asymptotic character. The analysis of these methods for case of a short sample is usually made by computer simulation. In this work, the AIC and MDL statistical characteristics are found in case of known array internal noise power. It is shown that for such task formulation, the characteristics of the multidimensional signal detection by these methods can be calculated analytically with help of results obtained by authors for maximal eigenvalue distribution of the sample correlation matrix [3]. In this paper, the analytical formulas for the false alarm probability are obtained for any size of the sample.

It is supposed that a space-coherent signals from the *K* statistical independent sources are received by the narrow-band *N* elements array. Then, the signal vector $\vec{x}$ of array elements can be written as:

$$\vec{x}=\sum_{i=1}^{K} a_i \vec{s}_i + \vec{n}\,, \tag{1}$$

where $\vec{x}$ [1x*N*] is output signals vector, $\vec{s}_i$ is [1x*N*] phase vector, and a_i is complex Gaussian amplitude of the signal from *i*-th-source, $\vec{n}$ is [1x*N*] array internal noise vector.

It is considered that array internal noise is uncorrelated, white, Gaussian noise with zero mean value and unity power in each element. The source number estimation task for model (1) is the number *K* estimation with finite sample $\vec{x}_i$, $i=1..L$ of signal vector $\vec{x}$ on the base of some statistical criterion or algorithm.

Now the most known criterions of source number estimation are AIC and MDL criterions [1]. By AIC and MDL criterions the source number estimation $\hat{K}$ is value which minimise the functions $AIC(k)$ or, correspondingly, $MDL(k)$.

$$AIC(k)=-2\max L_k + 2T(k)\,, \tag{2}$$

$$MDL(k)=-\max L_k + 0.5\log(L)T(k)\,, \tag{3}$$

where $\max L_k$ is maximum of a likelihood function logarithm, for fixed signal number k, $T(k)$ is the number of independent parameters to determinate model (1). The overestimation probability equals the probability of the function *AIC(k)* decreasing after point *k=K*

$$P\{\hat{K} > K\} = P\{\Delta AIC(K) < 0\}, \quad \Delta AIC(k) = AIC(k+1) - AIC(k) \tag{4}$$

For the considered model of the Gaussian input signal (1) with taking into account that array element noise power is equal to 1, the $\Delta AIC(k)$ is:

$$\Delta AIC(k) = -2(1 + \log \hat{\lambda}_{K+1} - \hat{\lambda}_{K+1}) + 2(2(N-K)-1) \tag{5}$$

where $\hat{\lambda}_i$ is the *i*-th eigenvalue of the sample correlation matrix $\hat{M} = \frac{1}{L}\sum_{i=1}^{L} \vec{x}_i \vec{x}_i^{\,+}$.

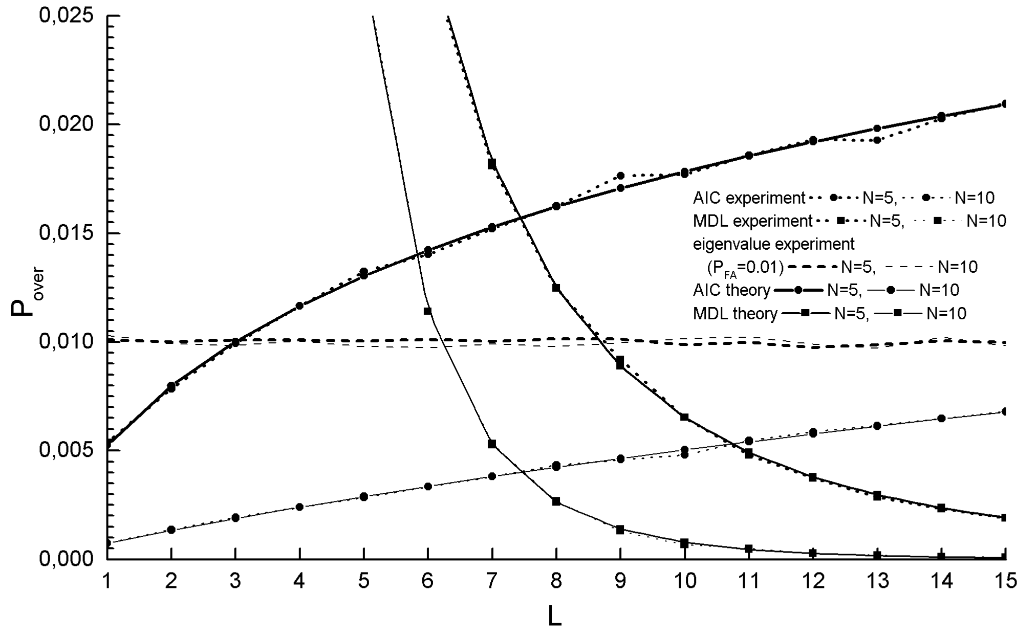

Fig.1

The function $\log \hat{\lambda}_{K+1} - \hat{\lambda}_{K+1}$ is monotone, hence the source number overestimation probability for AIC and MDL methods equals the probability of the threshold $h_{AIC}(K+1)$ exceeding by $\hat{\lambda}_{K+1}$ eigenvalue. The threshold value is calculate by the equation solution.

$$\begin{gathered} P_{AIC}\{\hat{K} > K\} = P\{\hat{\lambda}_1 > h_{AIC}(K+1)\} \\ h_{AIC}(K+1) - \log h_{AIC}(K+1) = 1 + \frac{2(N-K)-1}{L} \end{gathered} \tag{6}$$

In the work [3], the cumulative distribution function $F(N, L, \lambda)$ of the maximal eigenvalue of sample correlation matrix was obtained for case of source absence ($K = 0$). Using this result and equations (6), the false alarm probability of signal detection by AIC criterion can be written:

$$P_{AIC}\{\hat{K} > 0 \mid K = 0\} = P\{\lambda_1 > h_{AIC}(1) \mid K = 0\} = 1 - F\big(N, L, h_{AIC}(1)\big) \tag{7}$$

This function is not constant and depends on parameters N и L. If the constant false alarm probability is needed for a detection system (Neyman-Pearson criterion), the threshold h_{EIG} must be calculated by equation

$$P_{FA} = 1 - F(N, L, h_{EIG}) \tag{8}$$

using defined false alarm probability P_{FA}. In this case, the false alarm probability $P\{\hat{K} > 0 \mid K = 0\}$ will not be dependent on *N* and *L*. By analogy the MDL criterion can be analysed.

For checking theoretical results, the computer simulation of the AIC and MDL methods was made for case $K = 0$. The experimental dependence of the false alarm probability from snapshots number *L* was obtained for *N*=5 (thin lines) and *N*=10 (thick lines) (see fig. 1.) On the same figure are shown the theoretical false alarm probability for AIC and MDL calculated by (6) and (7). It is seen that the theoretical and experimental curves coincident. Also the dependence of the false alarm probability for the method based on the maximal eigenvalue comparison with the threshold calculated by (8) for $P_{FA} = 0.01$ is depicted. It is seen that this method gives constant level of false alarm probability in contrast to the AIC and MDL methods.

This work was supported by grants of RFFI № 00-02-17602, №00-15-96620, and INTAS №96-2352

References

1. M.Wax and T.Kailath, “Detection of signals by information theoretic criteria” //IEEE Trans. Acoust. Speech, Signal Processing, vol. ASSP-33, pp387-392, Apr. 1985.

2. W.Xu, M.Kaveh, “Analysis of the performance and the sensitivity of eigendecomposition-based detectors.” IEEE Transaction on signal processing vol. 43, No 6, June 1995.

3. V.T.Ermolaev, K.V.Rodyushkin "The distribution functions of the maximum eigenvalue of a sample correlation matrix of internal noice of antenna array elements "Radiophysiscs and Quantum Electronics 1999г. №5 vol 42, p. 439.