

## 두 가지 판단 기준을 바탕으로 한 신호원 수 추정방법의 비교

正會員 이 주 식\*, 이 용 업\*\*, 송 익 호\*, 유 상 우, 김 홍 길\*\*

### Comparison of Estimation Methods for the Number of Signal Sources Based on Two Decision Criteria

Jooshik Lee\*, Yong Up Lee\*\*, Iickho Song\*, Sangwoo Ryu\*\*, and  
Hong Gil Kim\* *Regular Members*

#### 요 약

이 논문에서는 신호원 수를 추정하는 문제에 쓰는 아카이케 정보기준과 (아정기) 최소묘사길이 기준의 (최묘기) 성능 특징을 견주어 보았다. 일반적으로 신호원 수는 단순히 아정기나 최묘기로부터 구한 비용함수를 가장 작게 해서 얻지만, 늘 그런 것은 아니다. 모의실험을 통해 신호대잡음비와 표본크기에 관한 함수로 아정기와 최묘기의 성능을 견주어 보았다. 최묘기는 신호원 수를 더 정확하게 추정했지만, 아정기는 신호대잡음비가 높거나 표본크기가 클 때 신호원 수를 일정하지 않게 과추정했고, 신호대잡음비가 낮을 때 일정하지 않은 과소추정을 나타냈다.

#### ABSTRACT

In this paper, the performance characteristics of the Akaike information criterion (AIC) and minimum description length (MDL), which have been used in the problems of detecting the number of signal sources recently, are compared. Generally, the number of signal sources is obtained merely by minimizing the cost function obtained from the AIC or MDL criterion, but not always. Simulation results showing the performance of the AIC and MDL as a function of the signal to noise ratio (SNR) and sample size are presented. It is shown that the MDL method estimates the number of signal sources more reliably and exactly while the AIC overestimates the number of signal sources inconsistently at high SNR and sample size, and underestimates inconsistently at low SNR.

\*한국과학기술원 전기 및 전자공학과  
\*\*삼성전자 정보통신연구소 무선통신연구팀  
論文番號: 96030-0129  
接受日字: 1997年 1月 29日

## I. 머리말

신호처리분야에서 제시되는 많은 문제들 가운데에서 신호원 수를 추정하는 것은 매우 흥미있는 문제이다. 보기를 들어, 선형감지기 배열처리, 레이더 후방 분산으로부터 겹치는 반향들의 복귀, 자연응답으로부터의 시스템 극 (poles) 회복과 같은 문제들에서 신호원 수 추정은 중요한 문제이다. 일반적으로, 신호원 수 추정 방법은 신호원 공간에서 잡음이 더해진 한정된 신호원 관측벡터에 의해 형성된 공분산 행렬의 고유값으로부터 그 수가 추정될 수 있다는 것을 바탕으로 한다. 기존 방법의 한 보기인 가설검정법은 [1, 2] 각 가설에 대해 우도비 (likelihood ratio) 통계량이 계산되고 그 값이 문턱값보다 크면 가설을 받아들인다. 이 방법에서 문제는 문턱값 결정에 필요한 종속적 판단이다. 이와는 달리, 신호원 수를 추정하는 방법으로 널리 쓰이는 아정기와 [3, 4] 최요기에서는 [4, 5] 신호와 잡음의 통계적 특성을 완전히 알아야 쓸 수 있는데, 이들 방법에서는 결정처리에 어떠한 종속적 판단도 필요하지 않고, 단순히 신호원 수가 아정기나 최요기를 가장 작게 하는 값으로써 결정된다.

먼저, 이 논문에서는 아정기와 최요기의 수학적 모형화를 최대우도추정의 [4, 6, 7] 방법을 써서 나타냈다. 이것을 바탕으로 해서 아정기와 최요기에서 나타나는 문제점인 신호대잡음비에 따른 과추정과 (overestimate) [4] 과소추정 (underestimate) 현상을 모의실험을 통해서 알아보고, 표본크기에 따른 이들의 성능도 살펴보았다. 이렇게 해서 얻은 결과를 바탕으로 해서 신호원 수를 추정할 때 아정기와 최요기 방법 가운데에서 어느 것이 더 좋은지를 견주어 보았다.

## II. 신호원 수 추정방법

선형감지기 배열처리나 레이더 후방 분산으로부터 겹치는 반향의 회복과 같은 신호처리에서 크기  $p \times 1$  인 관측벡터,  $x(t)$ 는 다음과 같은 형태로 쓸 수 있다 [4, 8].

$$x(t) = \sum_{i=1}^q a(\Phi_i) s_i(t) + n(t) \quad (1)$$

여기서,  $q$ 는 ( $q < p$ ) 신호원 수,  $s_i(\cdot)$ 는  $i$ 제 신호의

복소파형이고,  $a(\Phi_i)$ 는  $i$ 제 신호와 관련된 매개벡터  $\Phi_i$ 에 의해 구성되는  $p \times 1$  복소벡터이며  $n(t)$ 는 부가적인 잡음의  $p \times 1$  복소벡터이다. 한편,  $q$  신호들,  $s_1(\cdot), s_2(\cdot), \dots, s_q(\cdot)$ 는 평균이 0이고 양의 한정된 공분산 행렬을 갖는 정규확률과정이라 하고, 잡음벡터  $n(\cdot)$ 은 신호들과 독립이고 평균이 0이며 공분산 행렬이  $\sigma^2 I$ 인 정규확률벡터라고 가정한다. 여기서  $\sigma^2$ 은 알려지지 않은 상수이고  $I$ 는 단위행렬이다.

이제, (1)로 나타낸 모형을 가지고 관측벡터들  $x(t_1), x(t_2), \dots, x(t_N)$ 로부터 신호원 수  $q$ 를 추정하는 문제를 생각할 수 있다.

(1)을 행렬의 꼴

$$x(t) = A s(t) + n(t) \quad (2)$$

로 쓸 수 있다. 여기서,  $p \times q$  행렬  $A = [a(\Phi_1), a(\Phi_2), \dots, a(\Phi_q)]$ 이고  $q \times 1$  벡터  $s^T(t) = [s_1(t) s_2(t) \dots s_q(t)]$ 이다.

따라서, 복소확률벡터  $x(\cdot)$ 의 공분산 행렬  $R_x = E\{(x - m_x)(x - m_x)^H\}$ 은 가정에 의해  $m_x = 0$ 이고 잡음은 신호들과 독립이므로

$$R_x = \Psi + \sigma^2 I \quad (3)$$

로 쓸 수 있다. 여기서,  $\Psi = A S A^H$ 이고 신호의 공분산 행렬  $S = E\{s(\cdot) s^H(\cdot)\}$ 이다.

신호 부공간에서 신호원들이 서로 독립이라면 행렬  $A$ 는 순위가 (rank)  $q$ 이고 신호원의 공분산 행렬  $S$ 는 정칙이다. 이 때, 행렬  $\Psi$ 의 고유값들 가운데에서  $p-q$ 개는 0이 된다. 곧,  $R_x$ 의 고유값들을  $\lambda_1, \lambda_2, \dots, \lambda_p$ 라고 했을 때 가장 작은  $p-q$ 개의 고유값들은 모두  $\sigma^2$ 이 된다 [9, 10].

따라서, 신호원 수는  $R_x$ 의 가장 작은 고유값들의 수로부터 추정될 수 있다. 그러나, 실제로  $R_x$ 는 알 수 없기 때문에 유한한 크기의 표본으로  $R_x$ 를 추정할 수밖에 없다. 이때, 모든 고유값들이 다르게 나오고 단순히 그 값들을 가지고 신호원 수를 추정하기란 어렵다. 바로 이 어려운 점을 풀기 위해서 제시된 방법들 가운데에서 널리 쓰이는 것이 아정기와 최요기 방법이다.

신호원 수를 추정하고자 할 때 정보이론기준에서 기본적으로 생각해야 하는 것은 자료로부터 얻어지

는 정보와 매개변수 추정에 따르는 불확실성 사이의 관계이다.

먼저, 정보이론기준의 대수우도함수는 어느 정도의 정보를 알고 있는지의 척도로 생각할 수 있다. 신호원 수의 추정과 관련이 있는 행렬  $\Psi$ 의 순위가  $k$ 라고 할 때, (3)에서 주어진 공분산 행렬은 Hermitian 행렬이므로 대각행렬화 하면,

$$\begin{aligned} R_x &= PDP^H \\ &= [P_S \ P_N] D [P_S \ P_N]^H \\ &= [P_S \ P_N] \begin{bmatrix} D_S & 0 \\ 0 & D_N \end{bmatrix} \begin{bmatrix} P_S^H \\ P_N^H \end{bmatrix} \\ &= P_S D_S P_S^H + P_N D_N P_N^H \end{aligned} \quad (4)$$

이다. 여기서,  $P_S$ 는 신호 부공간을 정의하는 고유벡터들  $e_1, e_2, \dots, e_k$ 로 이루어진  $p \times k$  행렬이고  $P_N$ 는 잡음 부공간을 [11, 12] 정의하는 고유벡터들  $e_{k+1}, e_{k+2}, \dots, e_p$ 로 이루어진  $p \times (p-k)$  행렬이다. 한편,  $D_S$ 는 신호 부공간에 대응하는 고유값들  $\lambda_1, \lambda_2, \dots, \lambda_k$ 을 원소로 하는  $k \times k$  대각행렬,  $D_N$ 은 잡음 부공간에 대응하는 고유값들  $\lambda_{k+1}, \lambda_{k+2}, \dots, \lambda_p$ 을 원소로 하는  $(p-k) \times (p-k)$  대각행렬이다. 그리고  $k(0 \leq k \leq p-1)$ 는 가능한 신호원 수를 나타낸다. 한편, 위 모형의 매개변수 집합은  $\theta = (\lambda_1, \lambda_2, \dots, \lambda_k, \sigma^2, P_S)$ 로 쓸 수 있다.

이제, 정보이론기준의 대수우도함수를 이끌어 보자.

관측벡터들  $x(t_1), \dots, x(t_N)$ 이 통계학적으로 평균 0을 가지는 독립 복소정상확률 벡터라고 할 때, 그것들의 결합 확률밀도함수는

$$f(x_1, x_2, \dots, x_N | \theta) = \prod_{i=0}^N \frac{1}{\pi^p \det R_x} \exp(-x_i^H R_x^{-1} x_i) \quad (5)$$

이다. (5)의 양변에 대수를 취하고 상수항을 무시하면 대수우도함수  $L(\theta)$ 는

$$L(\theta) = N \ln(\det R_x) - N \operatorname{tr}(R_x^{-1} \hat{R}) \quad (6)$$

로 쓸 수 있다. 여기서,  $\hat{R} = \frac{1}{N} \sum_{i=1}^N x_i x_i^H$ 는 표본공분산 행렬이다.

한편,  $R_x^{-1} = P_S D_S^{-1} P_S^H + P_N D_N^{-1} P_N^H$ 를 (6)의

둘째 항에 대입하면,

$$\begin{aligned} \operatorname{tr}(R_x^{-1} \hat{R}) &= \operatorname{tr}((P_S D_S^{-1} P_S^H + P_N D_N^{-1} P_N^H) \hat{R}) \\ &= \operatorname{tr}(P_S^H \hat{R} P_S D_S^{-1}) + \operatorname{tr}(P_N^H \hat{R} P_N D_N^{-1}) \quad (7) \\ \therefore \operatorname{tr}(AB) &= \operatorname{tr}(BA), \operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B) \end{aligned}$$

이다. 여기서,  $P_S^H \hat{R} P_S$ 와  $P_N^H \hat{R} P_N$ 을 각각  $\tilde{\Sigma}$ 와  $\tilde{N}$ 라 하면, (4)와 (7)에서

$$L(\theta) = -N \ln \det D_S - N \ln \det D_N - N \operatorname{tr}(\tilde{\Sigma} D_S^{-1}) - N \operatorname{tr}(\tilde{N} D_N^{-1}) \quad (8)$$

로 쓸 수 있다. 최대우도추정을 이용해서  $L(\theta)$ 를 가장 크게 하는  $\theta$ 의 값을 찾기 위해 (8)의 양변을 매개변수 벡터  $\theta$ 의 원소인 신호 부공간의 고유값  $\lambda_i$ 와  $\sigma^2$ 에 대해서 미분을 한다. 이 때, 얻어진 추정값들은

$$\frac{\partial L(\theta)}{\partial \lambda_i} = -N \frac{1}{\lambda_i} + N \frac{l_i}{\lambda_i^2} = 0 \quad \therefore \hat{\lambda}_i = l_i, i=1, 2, \dots, k \quad (9a)$$

$$\begin{aligned} \frac{\partial L(\theta)}{\partial \sigma^2} &= -N(p-k) \frac{1}{\sigma^2} + N(l_{k+1} + l_{k+2} + \dots + l_p) \frac{1}{(\sigma^2)^2} = 0 \\ \therefore \hat{\sigma}^2 &= \frac{1}{p-k} \sum_{i=k+1}^p l_i \end{aligned} \quad (9b)$$

이다.

여기서,  $l_1 > l_2 > \dots > l_p$ 는 표본공분산 행렬  $\hat{R}$ 의 고유값들이다. (9a)와 (9b)를 (8)에 대입하여 얻은 식에서 상수항을 무시하면,

$$\begin{aligned} L(\hat{\theta}) &= -N \ln(l_1 l_2 \dots l_k) - N \ln \left( \frac{1}{p-k} \sum_{i=k+1}^p l_i \right)^{p-k} \\ &= -N \ln \left( \frac{\prod_{i=1}^k l_i}{\left( \frac{1}{p-k} \sum_{i=k+1}^p l_i \right)^{(p-k)}} \right) \\ &= -N \ln \left( \frac{\prod_{i=1}^k l_i / \prod_{i=k+1}^p l_i}{\left( \frac{1}{p-k} \sum_{i=k+1}^p l_i \right)^{(p-k)}} \right) \end{aligned} \quad (10)$$

$$\approx \ln \left( \frac{\prod_{i=k+1}^p l_i \frac{1}{p-k}}{\frac{1}{p-k} \sum_{i=k+1}^p l_i} \right)^{N(p-k)}$$

이다. (10)은 표본공분산 행렬의 가장 작은  $p-k$  고유값들의 기하평균과 산술평균의 비임을 알 수 있다.

매개변수 추정의 불확실성은  $\theta$ 의 매개변수 가운데에서 독립적인 매개변수의 수를 뜻한다. 따라서, 이것은  $\theta$ 로 나타나는 공간의 자유도의 수로부터 얻을 수 있다. 복소공분산 행렬의 고유값들은 실수이고, 고유벡터들은 원소로 복소수를 가지므로  $\theta$ 가 가지는 매개변수 수는  $k+1+2pk$ 가 된다. 그러나 복소고유벡터들은 각각 크기가 1이어야 하고 서로 수직이어야 한다. 이런 제한은 각각의 복소고유벡터들에 적용되는 전자의 경우,  $2k$ 개만큼 자유도 수의 감소를 뜻하며 복소고유벡터들 가운데에서 서로 다른 두 개의 벡터들에 적용되는 후자의 경우,  $2_k C_2$ 개만큼의 감소를 뜻한다. 따라서,  $\theta$ 에서 독립적인 매개변수의 수는

$$\mu = (k+1+2pk) - 2k - k(k-1) = k(2p-k) + 1 \approx k(2p-k) \quad (11)$$

이다.

### 2.1. 아정기 방법을 쓴 신호원 수 추정

일반적으로 정보이론기준은  $N$  관측벡터들  $x_1, x_2, \dots, x_N$ 과 이들에 대한 매개변수에 의존하는 확률밀도 함수들의 집합이 주어질 때, 자료에 가장 잘 접근한 모형을 고르는데 쓰인다. 아정기는 Akaike에 의해 제안된 정보이론기준인데 관측벡터들에 대한 우도함수를 써서

$$AIC = -2 \ln f(X|\hat{\theta}) + 2k \quad (12)$$

로 정의한다. 여기서, 첫째 항은 앞에서 세운 모형의 매개변수 벡터 원소들의 최대우도추정을 쓴 대수우도함수를 나타내며, 둘째 항은 모형화된 확률밀도 함수  $f(X|\theta)$ 와 추정된 확률밀도 함수  $f(X|\hat{\theta})$  사이의 차이를 보상해 주는 항을 나타낸다. 그리고,  $\hat{\theta}$ 는 매개변수 벡터  $\theta$ 의 최대우도추정값을 나타내며,  $k$ 는 매개변수 벡터  $\theta$ 의 공간에서 서로 독립적으로 자유롭게 생

각할 수 있는 매개변수의 수를 나타낸다.

(10)과 (11)로 (12)를 나타내면,

$$AIC(k) = -2 \ln \left( \frac{\prod_{i=k+1}^p l_i \frac{1}{p-k}}{\frac{1}{p-k} \sum_{i=k+1}^p l_i} \right)^{N(p-k)} + 2k(2p-k) \quad (13)$$

이다.

아정기 방법은 아정기를 가장 작게 하는  $k$ 값을 신호원 수라고 추정한다. 그러나, 이 방법은 신호대잡음비나 표본크기에 관계없이 신호원 수를 과추정한다.

### 2.2. 최요기 방법을 쓴 신호원 수 추정

최요기는 Schwartz와 Rissanen에 의해 소개된 정보이론기준이다. Schwartz는 베이즈 추정 (Bayes estimation) 바탕으로 해서 각 모형에 사전 확률을 주고 사후 확률을 가장 크게 하는 모형을 고르는 방법을 제시했다. Rissanen은 정보 이론에 근거해서 관측된 자료를 암호화하는데 각 모형을 쓴 뒤 최소 암호길이를 나타내는 모형을 고르는 방법을 제시했다 [4]. 표본크기가 클 때 Schwartz와 Rissanen의 방법들은 다음과 같은 기준을 나타낸다.

$$MDL = -\ln f(X|\hat{\theta}) + \frac{1}{2} k \ln N \quad (14)$$

위 식에서 첫째 항은 인수 2를 제외하고는 아정기의 첫째 항과 같지만 둘째 항은 곱해지는 인수가 다르다.

(10)과 (11)로 (14)를 나타내면,

$$MDL(k) = -\ln \left( \frac{\prod_{i=k+1}^p l_i \frac{1}{p-k}}{\frac{1}{p-k} \sum_{i=k+1}^p l_i} \right)^{N(p-k)} + \frac{1}{2} k(2p-k) \ln N \quad (15)$$

이다.

최요기 방법도 아정기와 마찬가지로 최요기를 가장 작게 하는  $k$ 값을 신호원 수라고 추정한다. 신호대잡음비가 낮을 때 최요기는 과소추정을 나타내지만, 표본크기가 작을 때는 과추정을 나타낸다.

### III. 모의 실험

앞에서 제시한 방법들을 선형감지기 배열처리에 적용했을 때, 그 성능을 모의실험을 통해 알아본다.

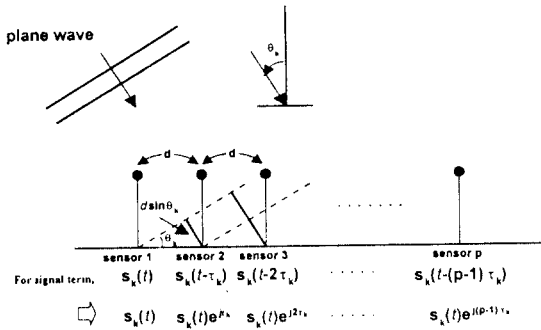


그림 1. 모의실험을 위한 센서배열처리 모형

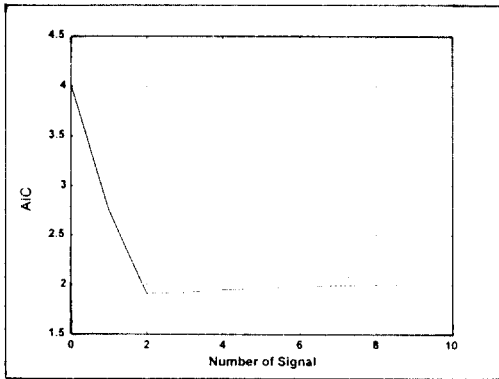


그림 2. 아정기 방법을 써서 신호원 수를 결정한 경우

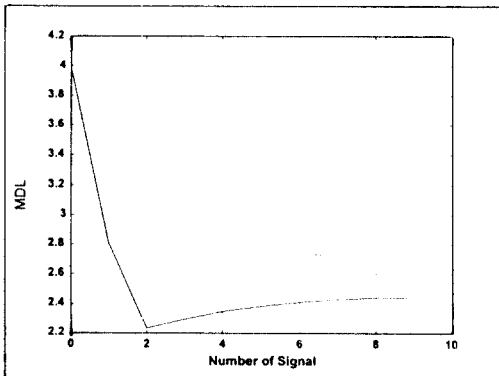


그림 3. 최요기 방법을 써서 신호원 수를 결정한 경우

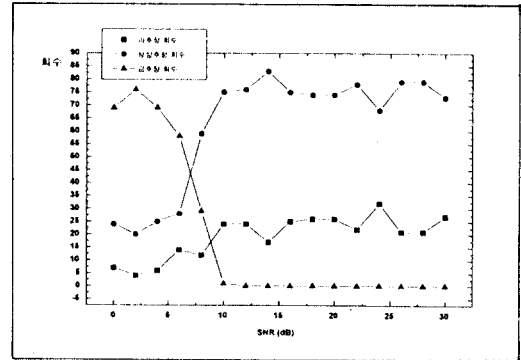


그림 4. 신호대잡음비에 따라 아정기 방법이 신호원 수를 과추정, 정상추정, 과소추정하는 회수

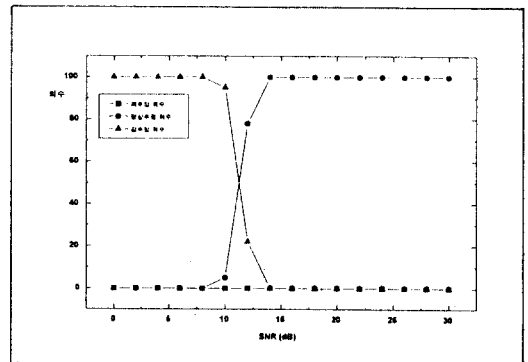


그림 5. 신호대잡음비에 따라 최요기 방법이 신호원 수를 과추정, 정상추정, 과소추정하는 회수

균일한 간격으로 놓인 선형감지기  $p$ 로 이루어진 안테나에 평면파  $q$ 가 서로 다른 방향에서 입사한다고 할 때,  $\theta_k$ 의 방향에서 입사하는 신호에 대해 시각  $t$ 에 감지기  $p$ 를 거쳐 얻은 값을 그림 1에서 보여준다. 여기서,  $d$ 는 감지기들 사이의 간격이고  $\tau_k = \frac{d \sin \theta_k}{c}$ 는 평면파가 감지기에 도달할 때, 이웃한 감지기 사이에서 생기는 지연이다.  $c$ 는 빛의 속도이고, 공간에서 전파 신호는 안테나와 신호원 사이가 거리  $r > \frac{2l^2}{\lambda}$  이면 그 신호가 중심주파수  $f_0$ 를 갖고 협대역 평면파의 형태로 안테나에 도달한다고 볼 수 있다. 여기서,  $l$ 은 안테나의 길이이고  $\lambda$ 는 입사 전자기파의 파장이다.

따라서, (1)의  $p \times 1$  행렬  $A(\tau_k)$ 는  $A(\tau_k) = [1 e^{j\tau_k} e^{j2\tau_k} \dots e^{j(p-1)\tau_k}]^T$  이 된다.

한편, 입사하는 평면파에 대해서  $s(t) = Ae^{-j(2\pi f_c t + \gamma)}$  라고 가정했다. 여기서, 위상  $\gamma$ 는  $[0, 2\pi]$  사이에 균일하게 분포된 확률변수이고,  $A$ 는 진폭이다.

이제, 10개의 감지기를 가진 안테나에 ( $p=10$ ) 2개의 신호원이 ( $q=2$ ) 각각  $5^\circ$ 와  $7^\circ$ 의 방향에서 도달한다고 할 때, 아정기와 최요기 방법을 써서 신호원 수를 추정한 모의 실험의 결과를 살펴보자.

그림 2와 그림 3은 표본크기가 256이고 신호대잡음비가 20dB일 때, 50회를 반복한 뒤 평균해서 각각 아정기와 최요기 방법이 추정하는 신호원 수를 보여준다. 여기서, 아정기와 최요기 방법은 정확하게 신호원 수를 추정했다.

그림 4와 5는 표본크기가 256일 때, 100번 시행을 통해 신호대잡음비에 따라 아정기와 최요기 방법이 신호원 수를 과추정, 정상추정, 과소추정하는 회수를 나타낸다. 한편, 그림 6은 표본크기가 256일 때, 1000번 시행을 통해 신호대잡음비에 따라 아정기와 최요기 방법이 신호원 수를 정확하게 추정하는 확률을 보여준다. 아정기는 신호대잡음비가 낮을 때 (0-8 dB범위) 일정하지 않은 과소추정을 하고, 신호대잡음비가 높을 때 (10dB이상) 일정하지 않은 과추정을 하기 때문에 일관성 있게 신호원 수를 추정하지 못한다는 것을 알 수 있다. 최요기는 신호대잡음비가 낮을 때에는 (0-10 dB범위) 일정한 과소추정을 하고 신호대잡음비가 문턱값을 넘어서면 일정하게 정확한 신호원 수를 추정한다는 것을 알 수 있다.

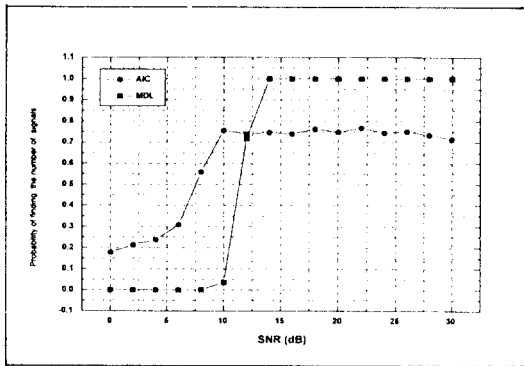


그림 6. 신호대잡음비에 따라 아정기와 최요기 방법이 신호원 수를 정확하게 추정할 확률

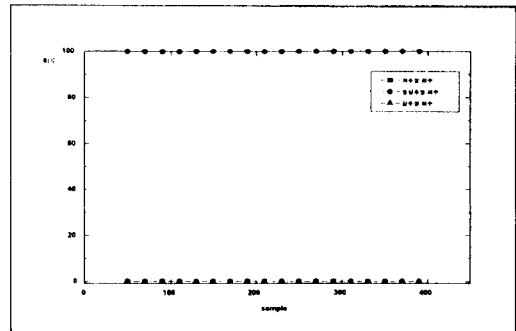


그림 8. 표본크기에 따라 최요기 방법이 신호원 수를 과추정, 정상추정, 과소추정하는 회수

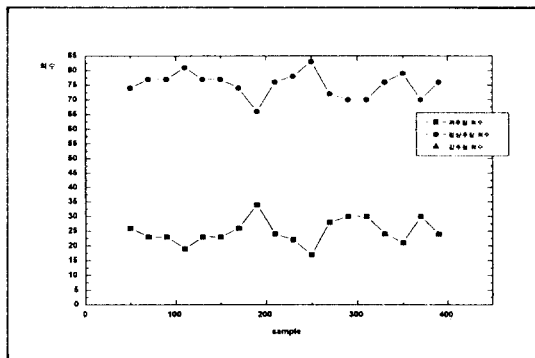


그림 7. 표본크기에 따라 아정기 방법이 신호원 수를 과추정, 정상추정, 과소추정 하는 회수

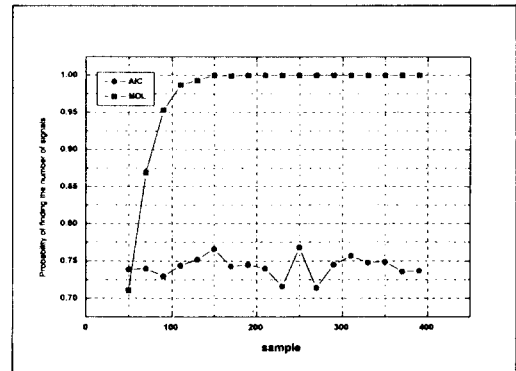


그림 9. 표본크기에 따라 아정기와 최요기 방법이 신호원 수를 정확하게 추정할 확률



