Statistical Analysis of Generalized Capon's Method

Jinho Choi[‡]**, Yong Up Lee**[*]**, and Iickho Song**[*]

[‡]R/D Center, Dacom Co.
34 Kajeong Dong, Yuseong Gu, Daejeon 305-350, Korea
Phone: +82-42-220-4225
[*]Department of Electrical Engineering
Korea Advanced Institute of Science and Technology (KAIST)
373-1 Guseong Dong, Yuseong Gu, Daejeon 305-701, Korea
Phone: +82-42-869-5445, Fax: +82-42-869-3410
e-mail: isong@Sejong.kaist.ac.kr

Abstract

We consider statistical properties of the generalized Capon's method. It is observed that the estimation error of the generalized Capon's method has almost the same variance as the MUSIC method, although the generalized Capon's method yields a slightly biased estimate.

1 Introduction

To estimate the direction of arrival (DOA) of signal sources, the multiple signal classification (MUSIC) method [1] has widely been used. The MUSIC method heavily depends on the subspaces decomposition which requires the estimate of the number of signal sources in the first place [2]. The estimation of the number of signal sources, however, is itself quite a difficult problem. On the other hand, there are some DOA estimation methods which do not require the number of signal sources. For example, the Capon's method [3] is a popular one. If the number of signal sources is known exactly, of course, the MUSIC method provides us with better performance than the Capon's method. In [4], a generalization of Capon's method is considered. In this paper we consider the statistical properties of the generalization of Capon's method.

Let us consider an L-element array of which output is $y(t) \in C^{L\times 1}$ with $C^{L\times 1}$ denoting the space of $L \times 1$ complex-valued vectors, and assume the standard model of observation:

$$y(t) = Ax(t) + n(t), \qquad t = 1, 2, \ldots, N. \tag{1}$$

In (1) it is assumed that the column vector $x(t)$ for M-signal sources is an $M \times 1$ zero mean complex normal random vector and the additive noise $n(t)$ is also a zero mean complex normal random vector with covariance matrix σI (the normal random vector assumption is not essential for the DOA estimation, but this assumption is used for the statistical analysis). The full-rank covariance matrix of $x(t)$ is $E[x(t)x^H(t)] = R_x$, where E denotes the statistical expectation and H denotes the Hermitian transpose. The matrix A is an $L \times M$ ($L > M$) complex matrix having the particular structure

$$A = [a(\theta_1), a(\theta_2), \ldots, a(\theta_M)],$$

where θ_i is the DOA of the i-th signal source. Here $a(\theta_i) \in C^{L\times 1}$ is called the steering vector.

If we denote the covariance matrix of $y(t)$ by R_y, it is easy to see that

$$R_y = AR_xA^H + \sigma I. \tag{2}$$

The eigenvalues and eigenvectors of R_y are denoted by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_L$ and $e_1, e_2, \ldots, e_L$, respectively. It is noteworthy that $\lambda_{M+1} = \lambda_{M+2} = \ldots = \lambda_L = \sigma$. The ranges of the matrices $S \triangleq [e_1, e_2, \ldots, e_M]$ and $G \triangleq [e_{M+1}, e_{M+2}, \ldots, e_L]$ are called the signal and noise subspaces, respectively.

2 Extension of the MUSIC Method

Let us first define two matrices

$$F \triangleq [\hat{S}\vdots\hat{G}] \tag{3}$$

and

$$W \triangleq \begin{pmatrix} 0_{M\times M} & 0_{M\times(L-M)} \\ 0_{(L-M)\times M} & I_{(L-M)\times(L-M)} \end{pmatrix}, \tag{4}$$

where $0_{N\times K}$ is the all-zero matrix of size $N \times K$. Then $f_{MU}(\theta) = a^H(\theta)FWF^Ha(\theta)$. Now let us consider a spectrum which is defined by

$$f_n(\theta) = a^H(\theta)F\Lambda^{-n}F^Ha(\theta), \qquad n \geq 0, \tag{5}$$

where $\Lambda = diag[(\frac{\hat{\lambda}_1}{\hat{\lambda}_L}), (\frac{\hat{\lambda}_2}{\hat{\lambda}_L}), \ldots, (\frac{\hat{\lambda}_L}{\hat{\lambda}_L})]$ with $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \ldots \geq \hat{\lambda}_L$ being the eigenvalues of $\hat{R}_y$. If $n \to \infty$ and $\hat{R}_y = R_y$, we have

$$\begin{aligned} f_\infty(\theta) &= a^H(\theta)GG^Ha(\theta) \\ &= f_{MU}(\theta) \\ &= a^H(\theta)GG^Ha(\theta), \end{aligned} \tag{6}$$

since $\lim_{n\to\infty} \Lambda^{-n} = W$ when $\hat{R}_y = R_y$.

The spectrum $f_n(\theta)$ does not require the number of signal sources in estimating DOA and as we can see from (5) it is expected to have almost the same performance as the MUSIC null-spectrum when n goes to infinity if $\hat{R}_y \simeq R_y$. In addition if $n = 1$, $f_n(\theta) = f_1(\theta)$ is Capon's spectrum scaled by $\hat{\lambda}_L$, which can be easily shown by noting $\hat{\lambda}_L\hat{R}_y^{-1} = F\Lambda^{-1}F^H$.

3 Statistical Properties

The estimation errors $(\hat{\theta}_i - \theta_i), i = 1, 2, \ldots, M$, when we use $f_{MU}(\theta)$, are asymptotically jointly normally distributed with zero means and covariances given by [5]

$$\begin{aligned} \sigma^2_{MU}(i,j) &= E[(\hat{\theta}_i - \theta_i)(\hat{\theta}_j - \theta_j)] \\ &= \frac{2}{N}\frac{Re\{d^H(\theta_j)GG^H d(\theta_i)a^H(\theta_i)U a(\theta_j)\}}{\bar{f}''_{MU}(\theta_i)\bar{f}''_{MU}(\theta_j)}, \end{aligned} \tag{7}$$

where

$$U = \sum_{k=1}^{M} \frac{\sigma\lambda_k}{(\sigma - \lambda_k)^2} e_k e_k^H, \tag{8}$$

$$\bar{f}''_{MU}(\theta_i) = 2d^H(\theta_i)GG^H d(\theta_i), \tag{9}$$

and

$$d(\theta) = \frac{d}{d\theta}a(\theta). \tag{10}$$

First let us consider the asymptotic (for a large value of N) statistical properties of the estimates of DOA's obtained from $f_n(\theta)$. Consider the spectrum $f_n(\theta)$ of (5), which may be expressed as

$$f_n(\theta) = a^H(\theta)\hat{G}\hat{\Lambda}_G^{-n}\hat{G}^H a(\theta) + a^H(\theta)\hat{S}\hat{\Lambda}_S^{-n}\hat{S}^H a(\theta) \tag{11}$$

with $\hat{\Lambda}_G = diag[\frac{\hat{\lambda}_{M+1}}{\hat{\lambda}_L}, \frac{\hat{\lambda}_{M+2}}{\hat{\lambda}_L}, \ldots, \frac{\hat{\lambda}_L}{\hat{\lambda}_L}]$ and $\hat{\Lambda}_S = diag[\frac{\hat{\lambda}_1}{\hat{\lambda}_L}, \frac{\hat{\lambda}_2}{\hat{\lambda}_L}, \ldots, \frac{\hat{\lambda}_M}{\hat{\lambda}_L}]$. Asymptotically (i.e., for a large value of N) we can replace $\hat{\Lambda}_G$ and $\hat{\Lambda}_S$ by Λ_G and Λ_S, respectively [5], without affecting the asymptotic properties. We thus have asymptotically

$$f_n(\theta) = a^H(\theta)\hat{G}\hat{G}^H a(\theta) + a^H(\theta)\hat{S}\Lambda_S^{-n}\hat{S}^H a(\theta), \tag{12}$$

because $\Lambda_G = I$. It is noteworthy that this replacements are also valid when n is small, if $\hat{\Lambda}_G \simeq \Lambda_G$ and $\hat{\Lambda}_S \simeq \Lambda_S$. Noting that the first M eigenvalues depend on the signal power, it is easy to see that the spectrum $f_n(\theta)$ is asymptotically almost identical to the MUSIC null-spectrum $f_{MU}(\theta)$ when SNR is high or when the value of n is large. This is because of the fact that the matrix Λ_S^{-n} can be approximated by the $M \times M$ zero matrix when the value of n is large (in which case $(\frac{\lambda_L}{\lambda_i})^n, i = 1, 2, \ldots, M$, approach 0) or when the SNR is high (in which case $\frac{\lambda_L}{\lambda_i}, i = 1, 2, \ldots, M$, are quite close to 0). Thus the asymptotic statistical properties of the estimates of DOA's obtained from $f_n(\theta)$ should be almost identical to those obtained from $f_{MU}(\theta)$ when the SNR is high or when the value of n is large.

Second, let us consider the case when the value of N is not (very) large and $n \to \infty$. In this case the replacements of $\hat{\Lambda}_G$ and $\hat{\Lambda}_S$ by Λ_G and Λ_S, respectively, can not be justified, since $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_L$ for finite N, and we have

$$\lim_{n\to\infty} f_n(\theta) = a^H(\theta)\hat{e}_L\hat{e}_L^H a(\theta)$$

$$\triangleq \quad h(\theta), \tag{13}$$

which is totally different from $f_{MU}(\theta)$. In fact, we may use this spectrum $h(\theta)$ to obtain the estimates of DOA's. In this case, however, since only the eigenvector $\hat{e}_L$ is used in the estimation, the variance of the estimation error will be greater than that obtained from $f_{MU}(\theta)$ which uses the whole noise subspace, $Range(\hat{G})$. In summary the statistical properties of the estimates of DOA's obtained from $f_n(\theta)$ can be analyzed by using the approximation (12) when the value of N is large or when the value of n is finitely large, and by using the approximation (13) when the value of N is not large and $n \to \infty$ (the value of n is very large).

In order to obtain the estimation error of DOA's when we use $f_n(\theta)$, we use a Taylor expansion,

$$\begin{aligned} f_n'(\hat{\theta}_i) &= 0 \\ &\simeq f_n'(\theta_i) + f_n''(\theta_i)(\hat{\theta}_i - \theta_i), \end{aligned} \tag{14}$$

where $f_n'(\theta) = \frac{d}{d\theta} f_n(\theta)$ and $f_n''(\theta) = \frac{d^2}{d\theta^2} f_n(\theta)$. Thus in order to obtain the statistical properties of $(\hat{\theta}_i - \theta_i)$, those of $f_n'(\theta_i)$ and $f_n''(\theta_i)$ are necessary.

It can be shown that the estimation errors are jointly normal random variables with means

$$\mu_n(i) \triangleq E[\hat{\theta}_i - \theta_i] = -\frac{m(\theta_i)}{f_{MU}''(\theta_i) + \bar{\epsilon}_n''(\theta_i)} \tag{15}$$

and covariances

$$\begin{aligned} \sigma_n^2(i,j) &\triangleq E[(\hat{\theta}_i - \theta_i)(\hat{\theta}_j - \theta_j)] - \mu_n(i)\mu_n(j) \\ &\simeq \frac{r(i,j) + c(i,j) + f_{MU}''(\theta_i) f_{MU}''(\theta_j) \sigma_{MU}^2(i,j)}{(f_{MU}''(\theta_i) + \bar{\epsilon}_n''(\theta_i))(f_{MU}''(\theta_j) + \epsilon_n''(\theta_j))}, \end{aligned} \tag{16}$$

where

$$\bar{\epsilon}_n''(\theta_i) = 2d^H(\theta_i) S \Lambda_S^{-n} S^H d(\theta_i) + 2Re\{d^H(\theta_i) S \Lambda_S^{-n} S^H a(\theta_i)\}, \tag{17}$$

and $m(\theta) = E[\epsilon_n'(\theta)]$, and $r(i,j) \triangleq E[\epsilon_n'(\theta_i)\epsilon_n'(\theta_j)] - m_\epsilon(\theta) m_\epsilon(\phi)$, and $c(i,j) \triangleq E[f_{MU}'(\theta_i)\epsilon_n'(\theta_j)] + E[f_{MU}'(\theta_j)\epsilon_n'(\theta_i)]$.

Now let us consider the variance of the ith estimation error

$$\sigma_n^2(i,i) = \frac{r(i,i) + c(i,i) + (f_{MU}''(\theta_i))^2 \sigma_{MU}^2(i,i)}{(f_{MU}''(\theta_i) + \epsilon_n''(\theta_i))^2}. \tag{18}$$

When $n \to \infty$ or when the SNR is high, $\bar{\epsilon}_n''(\theta_i)$, $c(i,i)$, and the covariance $r(i,i)$ tend to vanish, because $r(i,i) \propto (\frac{\sigma}{\lambda_M})^{2n}$, $c(i,i) \propto (\frac{\sigma}{\lambda_M})^n$, and $\bar{\epsilon}_n''(\theta_i) \propto (\frac{\sigma}{\lambda_M})^n$, from (17), respectively. Thus we have

$$\sigma_n^2(i,i) \simeq \sigma_{MU}^2(i,i). \tag{19}$$

In addition we have $\mu_n(i) \simeq 0$, because $m(\theta_i) \propto (\frac{\sigma}{\lambda_M})^n$, from (15). Thus, when the value of N is large, the statistical properties of the estimates obtained from $f_{MU}(\theta)$ and $f_n(\theta)$ are expected to be almost the same if the SNR is high or if the value of n is large.

4 Simulation Results

Now let us consider the variance of the estimation error. When $L = 10$ and $M = 2$, the variances are given in Tables 1 and 2, when the four spectra, $f_{MU}(\theta)$, $f_2(\theta)$, $f_4(\theta)$, and $h(\theta)$, are used in the DOA estimation. In Table 1, the variances of estimation errors using $f_{MU}(\theta)$, $f_2(\theta)$, $f_4(\theta)$, and $h(\theta)$ are shown, which are computed from the results of Section 3. The value of the variance decreases as the SNR or the number of snap shots N increases, but it does not depend on a specific null-spectrum used for DOA estimation except $h(\theta)$. Although we did not include in Table 1, the mean of the estimation errors of $f_n(\theta)$ turned out to be almost zero.

In Table 2 the variances of the estimation errors obtained from simulation results are given. If we compare Tables 1 and 2 we may conclude that the value of variance tends to increase in practice as n increases; the effect of n becomes less as N increases. When $N = 100$, the difference between the variance obtained from simulation and that computed by (16) is larger than that when $N = 500$. In addition, we can see that the values of variance obtained from simulation and (16) are close to each other when N is large. As n increases, the variances obtained from simulation become more close to each other, which implies that when the parameter n is large, the statistical properties of estimation error obtained from $f_n(\theta)$ would follow those obtained from $h(\theta)$ as we discussed in Section 3.

SNR	$N = 100$				$N = 500$			
(dB)	Var. (f_{MU})	Var. (f_2)	Var. (f_4)	Var. (h)	Var. (f_{MU})	Var. (f_2)	Var. (f_4)	Var. (h)
15	4.31E-5	4.31E-5	4.31E-5	1.50E-4	8.63E-6	8.63E-6	8.63E-6	3.01E-5
20	9.79E-6	9.79E-6	9.79E-6	1.01E-4	1.95E-6	1.95E-6	1.95E-6	2.02E-5
25	3.08E-6	3.08E-6	3.08E-6	3.23E-5	6.16E-7	6.16E-7	6.16E-7	6.46E-6
30	1.20E-6	1.20E-6	1.20E-6	1.25E-5	2.40E-7	2.40E-7	2.40E-7	2.51E-6
35	3.62E-7	3.62E-7	3.62E-7	7.57E-6	7.24E-8	7.24E-8	7.24E-8	1.51E-6

Table 1. Variances of estimation errors using $f_{MU}(\theta)$, and $f_2(\theta)$, and $f_4(\theta)$, and $h(\theta)$ (theoretical results).

SNR	$N = 100$			$N = 500$		
(dB)	Var. (f_2)	Var. (f_4)	Var. (f_{50})	Var. (f_2)	Var. (f_4)	Var. (f_{50})
15	5.19E-5	9.09E-5	6.05E-3	9.49E-6	1.07E-5	1.07E-4
20	1.63E-5	2.62E-5	3.56E-4	2.26E-6	2.57E-6	2.55E-5
25	5.76E-6	9.33E-6	1.12E-4	9.05E-7	1.02E-6	1.25E-5
30	1.41E-6	2.38E-6	3.67E-5	2.47E-7	2.93E-7	3.08E-6
35	5.02E-7	7.92E-7	1.88E-5	7.77E-8	8.79E-8	1.49E-6

Table 2. Variances of estimation errors using $f_2(\theta)$, and $f_4(\theta)$, and $f_{50}(\theta)$ (simulation results).

5 Concluding Remarks

In this paper we consider the statistical properties of the generalized Capon's method to estimate direction of arrivals without a priori knowledge of the number of sources. A statistical analysis showed that the estimation error of the proposed method has almost the same variance as that of the MUSIC method.

Acknowledgements

This research was supported by Korea Advanced Institute of Science and Technology (KAIST) under an Internal Grant and Korea Science and Engineering Foundation (KOSEF) under Grant 941-0900-040-2, for which the authors would like to express their thanks.

References

[1] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Ant., Prop.*, vol. AP-34, pp. 276-280, March 1986.

[2] M. Wax and T. Kailath, "Detection of signal sources by information theoretic criteria," *IEEE Trans. Acous., Speech, Signal Proc.*, vol. ASSP-33, pp. 387-392, April 1985.

[3] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408-1418, August 1969.

[4] J. Choi, Y. U. Lee, J. S. Yun, and I. Song, "A method of estimation of direction of arrival with number of signal sources unspecified," *Jour. Acoustic Soc. Korea*, vol. 12, pp. 56-61, June 1993.

[5] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acous., Speech, Signal Proc.*, vol. ASSP-37, pp. 720-740, May 1989.