GTLS의 ARMA시스템식별에의 적용 및 적응 GTLS 알고리듬에 관한 연구

ARMA System Identification Using GTLS method and Recursive GTLS Algorithm

김 재 인*, 김 진 영**, 이 태 원***
(Jae In Kim*, Jin Young Kim**, Tae Won Rhee***)

요 약

일반화된 완전최소자승법(generalized total least squares method, GTLS)의 ARMA 시스템 식별에의 적용과 GTLS의 적응알고리듬에 대하여 논한다. 일반화된 완전최소자승법은 입력과 출력을 알고 있는 시스템식별(system identification)문제에서, 출력이 잡음에 의하여 오염된 경우, 편이되지 않은 해를 구하기 위하여 사용되는 방법이다. 본 논문에서는 먼저 GTLS를 ARMA 시스템 식별에 적용하기 위한 formulation을 하고, 일반화된 완전최소자승법의 일반 해의 성질과 역행렬 정리(matrix inverse lemma)를 이용하여 적응 GTLS 방법을 제안한다. 다음 제안된 방법을 통하여 시스템식별에 적용하여 그 성능을 평가한다. 또한 GTLS 알고리듬과 제안한 적응 GTLS 알고리듬의 성능을 수학적으로 해석하고 컴퓨터 시뮬레이션을 통하여 이를 검증한다.

ABSTRACT

This paper presents an estimation of ARMA coefficients of noisy ARMA system using generalized total least square(GTLS) method. GTLS problem for ARMA system is defined as minimizing the errors between the noisy output vectors and estimated noisy-free output.

The GTLS problem is solved in closed form by eigen-problem and the perturbation analysis of GTLS is presented. Also its recursive solution(recursive GTLS) is proposed using the power method and the covariance formula of the projected output error vector into the input vector space. The simulation results show that GTLS ARMA coefficients estimator is an unbiased estimator and that recursive GTLS achieves fast convergence.

I. 서 론

ARMA(auto regressive moving average) 시스템의 식별이나 역모델링 문제에 있어서 파라미터 추정방법으로서 대표적으로 사용되고 있는 최소자승법(least squares, LS) 방법은 시스템의 출력에 부가적인 잡음(additive noise)이 섞이게 될 때 편이된 추정기(biased estimator)가 되므로 S/N비가 어느 한계 이하로 저하될 때에는 시스템 식별에 사용되어 질 수

* 한국통신 S/W연구소 음성언어연구팀
** 전남대학교 공과대학 전자공학과
*** 고려대학교 공과대학 전자공학과
접수일자 : 1994년 12월 27일

없다. 이러한 문제점을 극복하기 위하여 일반화된 최소자승법(generalized LS)의 방법이 적용되고 있으나[2] 계산량이 많은 것이 단점이다. 한편, 최근에 완전최소자승법(total least squares, TLS) 방법이 LS의 문제점을 극복하기 위하여 시스템 식별에 사용되었으나, 출력에만 잡음이 섞인 경우에는 TLS도 편이된 추정기가 된다[3,4]. 왜냐하면, 완전최소자승법은 입력과 출력 모두 잡음에 의하여 오염되었다고 가정하고 있기 때문이다.

본 논문에서는 편이 되지 않은 시스템 파라미터 추정기(unbiased estimator)로서 일반화된 완전최소자승법(generalized TLS, GTLS) 방법을 도입하여 ARMA시스템 식별에 적용하고 GTLS 방법의 타당성을 검증한다. GTLS 방법은 $y = A\theta$의 시스템식(system equation)의 식을 푸는데 있어서 LS 방법과는 달리 행렬 A의 일부 행벡터에도 잡음이 섞여 있을 수 있다는 전제조건하에서 해를 구하는 방법으로서, 편이가 되어 있지 않은 시스템 파라미터를 풀어내게 된다[5].

한편, 시스템 식별을 반복적으로 하기 위하여 여러 적응 알고리듬들이 제안되어 왔다. 그림 1은 본 논문에서 다루고자 하는 시스템 식별문제를 적응알고리듬의 관점에서 나타낸 것이다. 적응알고리듬의 대표적인 것으로는 LMS(least mean squares) 알고리듬과 RLS(recursive least squares)이 있다[6]. 물론 그 외에도 많은 적응 알고리듬이 있으나 대부분이 RLS 방법들 변형이라고 할 수 있다. 그런데, 그림에 보인 바와 같이 시스템의 출력 z가 잡음에 의하여 오염되어 있다고 한다면, LMS나 RLS 방법은 참값으로부터 편이된 시스템 파라미터를 구하게 된다. 왜냐하면, 위에서 간략히 살펴보았듯이 LS방법은 출력이 오염된 경우 편이된 해를 구하기 때문이다. 따라서 편이되지 않은 파라미터를 구하는 적응알고리듬이 요구되는데, 본 논문에서는 한 방법으로서 적응 GTLS 알고리듬에 대하여 논한다.

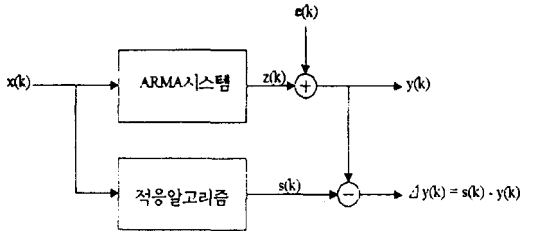


그림 1. 적응 시스템식별의 구성도
Figure 1. Block diagram of adaptive system identification.

먼저 GTLS 방법을 ARMA 시스템에 적용하기 위한 formulation 및 그 해법에 대하여 살펴본 후에 GTLS 방법의 일반 해의 성질에 대하여 논한다. 다음 GTLS 해의 성질과 행렬역변환정리를 이용하여 GTLS의 적응알고리듬을 제안한다. 제안된 적응 GTLS 알고리듬의 성능을 시뮬레이션을 통하여 검증한다. 그리고 수학적 해석에 의하여 적응 GTLS 알고리듬의 수렴성에 대하여 논하고, 수학적 결과와 시뮬레이션의 결과와의 비교를 통하여 수학적 해석의 타당함을 증명한다.

Ⅱ. 일반화된 최소자승법과 일반해

백색잡음에 의하여 오염된 ARMA 모델에 의하여 모델링된 잡음시스템은, 그림 1을 참조하여 기술하면, 다음과 같이 주어진다.

$$\sum_{i=0}^{p} a_i z_{k-i} + \sum_{j=0}^{q} b_j x_{k-j} = 0 \tag{1.1}$$

$$y_k = z_k + e_k, \tag{1.2}$$

단, x_k는 입력, z_k는 시스템의 출력,
y_k는 잡음이 섞인 관측출력, e_k는 백색잡음,
a_k는 AR 계수, b_k는 MA 계수, $a_0 = 1$,
p는 AR 차수 그리고 q는 MA 차수.

만약 N개의 관측값들을 갖고 있다고 한다면, 즉, 시간열 k가 1에서 N까지 변한다고 한다면 위의 식을 벡터 관계식으로 다음과 같이 나타낼 수 있다.

$$a^T z_k + b^T x_k = 0. \tag{2.1}$$

$$y_k = z_k + e_k, \tag{2.2}$$

단, $z_k = \begin{bmatrix} z_k \\ z_{k-1} \\ \vdots \\ z_{k-p} \end{bmatrix}$, $x_k = \begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-q} \end{bmatrix}$, $y_k = \begin{bmatrix} y_k \\ y_{k-1} \\ \vdots \\ y_{k-p} \end{bmatrix}$,

$z_k = \begin{bmatrix} z_k \\ z_{k-1} \\ \vdots \\ z_{k-p} \end{bmatrix}$, $e_k = \begin{bmatrix} e_k \\ e_{k-1} \\ \vdots \\ e_{k-p} \end{bmatrix}$, $a = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{bmatrix}$

그리고 $\boldsymbol{b}=\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_q \end{bmatrix}$ 이다.

이 때 입력과 잡음이 섞인 출력 그리고 시스템의 차수 p와 q를 알고 있다면, GTLS 문제를 ARMA 시스템 식별에 적용하였을 때, GTLS ARMA 식별문제는 다음과 같은 식으로 기술될 수 있다.

$$\min_{\Delta Y,a,b} \|\Delta Y\|_F \text{ subject to } [Y+\Delta Y \mid X]\begin{bmatrix} a \\ b \end{bmatrix}=0,$$

단, $Y=\begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}$ 이고 $X=\begin{bmatrix} x_1^T \\ \vdots \\ x_N^T \end{bmatrix}$ 이다.

위와 같이 정의된 GTLS 문제의 해는 Largrange multiplier(LM)를 사용하여 쉽게 풀어낼 수 있는데, LM을 이용한 간단한 수학적 계산에 의하여 위의 GTLS 문제는 다음의 문제와 같음을 보일 수 있다.

$$Minimize_{a,b} \frac{[a^T\ b^T]\begin{bmatrix} R_{yy} & R_{yx} \\ R_{xy} & R_{xx} \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}}{[a^T\ b^T]\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}},$$

단, $R_{yy}=\sum_k y_k y_k^T$, $R_{xx}=\sum_k x_k x_k^T$,

$R_{yx}=R_{xy}^T=\sum_k y_k x_k^T$

위의 문제는 다음의 등식을 만족하는 최소의 λ와 λ에 해당하는 $\boldsymbol{a}, \boldsymbol{b}$를 구하는 문제와 동일하다[8].

$$\left(\begin{bmatrix} R_{yy} & R_{yx} \\ R_{xy} & R_{xx} \end{bmatrix} -\lambda\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}\right)\begin{bmatrix} a \\ b \end{bmatrix}=0$$

위 등식을 풀면 다음과 같은 GTLS ARMA 모델링의 해를 구할 수 있는데, 참고문헌 [8]에 따른다.

(GTLS ARMA 모델링의 해)

만약 입력, 출력 그리고 입출력간의 공분산행렬(covariance matrix)을 알고 있다면, 시스템의 AR 파라미터는 행렬 $R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}$의 고유벡터중 최소의 고유치에 해당하는 고유벡터로 나타난다. 즉, $R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}$의 고유벡터를 $v_0, v_1, \cdots, v_p$라하고 고유치를 $\lambda_0, \lambda_1, \cdots, \lambda_p$이라고 하자. 그리고 $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_p$이라고 한다면, λ_p는 잡음의 파우어가 되며 AR파라미터는 다음과 같다.

$$a=\frac{1}{v_{p0}} V_p, \tag{3.1}$$

단, v_{p0}는 v_p의 첫번째 원소.

위 식에서 v_{p0}로 나누는 이유는 $a_0=1$의 조건을 만족시키기 위해서이다. 그러면, MA 파라미터 $\boldsymbol{b}$는 다음의 식으로 주어진다.

$$\boldsymbol{b}=-R_{xx}^{-1}R_{xy}\boldsymbol{a} \tag{3.2}$$

Ⅲ. 적응 GTLS 알고리듬

위의 GTLS 문제 및 해를 고찰하여 볼 때, GTLS 문제의 해를 반복적으로(recursive) 구하는 적응 GTLS 방법의 구현은 행렬 $R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}$의 고유벡터를 updating하는 문제로 간주할 수 있다. 특히, 최소의 고유값에 해당하는 고유벡터를 updating하는 문제이다. 임의 행렬이 주어졌을 때 고유벡터를 updating하는 문제는 간단한 문제가 아니다. 지금까지 알려진 고유벡터를 구하는 방법 중에 반복적으로 구하는데 가장 적절한 방법은 power method(PM)이다[9]. PM은 그 구현이 매우 간단한데 이 방법을 이용하면, 임의의 행렬에 대하여 최대의 고유치에 해당하는 고유벡터를 쉽게 구할 수 있으며, 그 외의 고유벡터도, 조금 많은 계산량이 필요하지만, 구할 수 있다. 특히 원점-이동 PM(origin-shift PM) 방법[9]이 적응 GTLS 문제에 적절히 응용될 수 있다. 그러나, 적응 GTLS 문제가 주어진 행렬 $R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}$에 대하여 최소의 고유치에 해당하는 고유벡터를 구하는 문제이므로 이러한 방법을 곧바로 적용했을 때 원점-이동 PM 방법을 최소의 고유치에 해당하는 고유벡터를 구하고자 할 때 수렴속도가 매우 느리기 때문에 적응 GTLS 알고리듬은 수렴속도상 사용하기가 힘들어진다. 그런데 PM이 최대의 고유치에 해당하는 고유벡터를 구할 때는 수렴속도가 빨라질 수 있는데, 이는 최대의 고유치가 다른 고유치에 비하여 상대적으로 큰 값을 가질 때이다. 그런데, 시스템 식별문제에서 S/N비가 크다면 행렬 $R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}$의 역행렬에 해당하는 행렬의 최대고유치는 부가잡음의 역(in-

verse)값이되어 다른 고유치에 비하여 항상 크게 된다. 따라서 행렬 $R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}$의 역행렬을 구할 수 있다면, 적응 GTLS 문제는 쉽게 풀릴 수 있을 것이다. 그런데, 다음의 정리는 $R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}$의 역행렬을 쉽게 구할 수 있도록 하며, 이 행렬의 updating도 매우 간단하게 구현할 수 있도록 하여 준다.

정리 : 행렬 $R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}$은 다음의 벡터 시퀀스의 공분산행렬이다. 즉, 다음의 벡터

$$\tilde{y}_k = y_k - R_{yx}R_{xx}^{-1}x_k, \tag{4}$$

를 정의하면, 분산행렬은 다음과 같다.

$$cov(\tilde{y}_k, \tilde{y}_k) = R_{yy}-R_{yx}R_{xx}^{-1}R_{xy} \tag{5}$$

증명 :

N개의 신호가 주어진 다음의 입출력 행렬을 정의하자

$$X=[x_1\ x_2\ \dots\ x_N]$$

$$Y=[y_1\ y_2\ \dots\ y_N]$$

그리고 식(4)에 의해 정의된 벡터 시퀀스의 행렬을 생각하여보자.

$$\tilde{Y}=[\tilde{y}_1\ \tilde{y}_2\ \dots\ \tilde{y}_N]$$

그러면 다음과 같이 행렬 $R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}$를 표현하여 위의 정리를 증명할 수 있다.

$$\begin{aligned}
&R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}\\
&\quad = YY^T - YX^T(XX^T)^{-1}XY^T\\
&\quad = Y(I-X^T(XX^T)^{-1}X)Y^T\\
&\quad = Y(I-X^T(XX^T)^{-1}X)(I-X^T(XX^T)^{-1}X)Y^T\\
&\qquad (\because I-X^T(XX^T)^{-1}X \text{ is idempotent})\\
&\quad = (Y-YX^T(XX^T)^{-1}X)(Y^T-X^T(XX^T)^{-1}XY^T)\\
&\quad = (Y-R_{yx}R_{xx}^{-1}X)(Y-R_{yx}R_{xx}^{-1}X)^T\\
&\quad = \tilde{Y}\tilde{Y}^T\\
&\quad = R_{\tilde{y}\tilde{y}}
\end{aligned}$$

따라서 $R_{yy}-R_{yx}R_{xx}^{-1}R_{xy}$는 벡터 시퀀스 $\tilde{y}_k=y_k-R_{yx}R_{xx}^{-1}x_k$의 공분산행렬이다.

위의 증명에서 $(I-X^T(XX^T)^{-1}X)$는 입력벡터 공간으로의 투영행렬(project matrix)이 된다. 따라서 벡터 $\tilde{y}_k$는 출력 벡터가 입력벡터 공간에 투영되었을 때의 에러 벡터에 해당한다.

그러면 벡터 $\tilde{y}_k$의 공분산행렬의 역행렬은 역행렬 레마(matrix inversion lemma)를 사용하여 쉽게 updating할 수 있다[6]. 이러한 방법은 RLS 알고리듬에서의 방법과 같은 방법이다. 이제 PM과 공분산행렬 등식 (5)를 이용하여 다음과 같은 적응 GTLS (recursive GTLS, RGTLS) 알고리듬을 다음과 같이 제안한다.

적응 GTLS 알고리듬

1. w, I, λ, $R_{yx}(k)$, $R_{xx}^{-1}(k)$ and $R_{\tilde{y}\tilde{y}}^{-1}(k)$를 k=0일 때 초기화한다. 여기서 벡터 w는 PM방법을 적용하기 위한 초기벡터이며, λ는 forgotten factor이고 I는 PM의 반복수(iteration number)이다.

2. $R_{yx}(k)$, $R_{xx}^{-1}(k)$를 update 한다. 다음의 두 식을 사용한다.

$$R_{yx}(k)=\lambda R_{yx}(k-1)+y_k x_k^T \tag{6}$$

$$R_{xx}^{-1}(k)=\lambda^{-1}R_{xx}(k-1) - \frac{\lambda^{-2}R_{xx}^{-1}(k-1)x_k x_k^T R_{xx}^{-1}(k-1)}{1+\lambda^{-1}x_k^T R_{xx}^{-1}(k-1)x_k} \tag{7}$$

3. 다음의 식을 이용하여 벡터 $\tilde{y}_k$를 계산한다.

$$\tilde{y}_k = y_{k-1} - R_{yx}(k)R_{xx}^{-1}(k)x_k \tag{8}$$

4. 행렬 $R_{\tilde{y}\tilde{y}}^{-1}(k)$를 다음과 같이 update 한다.

$$R_{\tilde{y}\tilde{y}}^{-1}(k)=\lambda^{-1}R_{\tilde{y}\tilde{y}}^{-1}(k-1) - \frac{\lambda^{-2}R_{\tilde{y}\tilde{y}}^{-1}(k-1)\tilde{y}_k\tilde{y}_k^T R_{\tilde{y}\tilde{y}}^{-1}(k-1)}{1+\lambda^{-1}\tilde{y}_k^T R_{\tilde{y}\tilde{y}}^{-1}(k-1)\tilde{y}_k} \tag{9}$$

5. 다음과 같이 PM을 이용하여 고유벡터를 계산한다.

do i = 1, l

$$\boldsymbol{v} = \boldsymbol{R}_{\dot{y}\dot{y}}^{-1}(k)\,\boldsymbol{w} \tag{10.1}$$

$$\boldsymbol{w} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \tag{10.2}$$

continue

6. ARMA 계수를 계산한다.

$$\boldsymbol{a}_k = \frac{1}{w(0)}\ \boldsymbol{w}. \tag{11.1}$$

$$\boldsymbol{b}_k = -\boldsymbol{R}_{xx}^{-1}(k)\,\boldsymbol{R}_{xy}(k)\,\boldsymbol{a}_k. \tag{11.2}$$

7. k를 증가하고 (k = k + 1) 단계 2로 간다.

한편, 지금까지의 적응 IIR 필터 문제를 살펴보면 equation error와 output error formulation의 두 가지가 있다[8]. 위의 두 가지 formulation에 기반한 방법은 추정된 계수가 편이 된다는 점과 국부적 최소값(local minimum)에 수렴한다는 두 가지의 단점을 가지고 있다. 즉, 시스템의 출력에 잡음이 더하여진 실제의 시스템을 고려하면, 위의 두 formulation에 기반한 적응 알고리듬은 참 계수 값으로부터 편이된 값에 수렴한다. 또한 output error formulation에 기반한 적응알고리듬의 에러표면(error surface)은 여러 개의 국부적 최소값을 가지고 있어서 국부적 최소값에 수렴할 가능성이 있는 것이다. 이들 지금까지의 적응 IIR 필터 알고리듬과 본 논문에서 제안된 적응 알고리듬 RGTLS방법을 비교하여 보면, 제안한 방법인 RGTLS알고리듬은 항상 PM방법의 특성상 최적의 해에 수렴하며[6], 입출력 관찰 sequence가 무한대로 갈 때 추정된 계수의 값은 편이되지 않는다. 즉, 편이되지 않은 추정기(unbiased estimator)이다.

Ⅳ. RGTLS 알고리듬의 섭동(perturbation) 분석

본 장에서는 적응 GTLS 알고리듬의 근사적인 통계적 특성에 대해서 기술한다. 추정된 ARMA 계수의 통계적인 편이와 분산이 어떻게 되는지 분석한다. 다음의 분석은 참고문헌 [9]에서 유도된 가우시안 프로세스의 샘플행렬의 고유벡터를 위한 고유벡터의 근사적 통계와 부록 A와 B에서 유도된 결과를 이용한다.

Ⅱ 절에서 AR 계수는 다음과 같이 주어졌다.

$$\boldsymbol{a} = \frac{1}{\boldsymbol{v}_{p0}}\ \boldsymbol{v}_p. \tag{12}$$

이때 $\hat{\boldsymbol{v}}_p$이 샘플행렬 $\hat{\boldsymbol{R}}_{yy} - \hat{\boldsymbol{R}}_{yx}\,\hat{\boldsymbol{R}}_{xx}^{-1}\,\hat{\boldsymbol{R}}_{xy}$의 최소 고유치에 해당하는 고유벡터라고 하면, 추정된 AR계수는 다음과 같다.

$$\hat{\boldsymbol{a}}_p = \frac{1}{\hat{\boldsymbol{v}}_{p0}}\ \hat{\boldsymbol{v}}_p. \tag{13}$$

그러면, N번의 추정결과의 각 AR계수에 대한 평균과 분산의 평균은 다음의 두 식에 의하여 주어진다. 이 두 식의 유도과정은 부록 A에 있다.

$$E[\hat{a}_l] \cong \frac{v_{pl}}{v_{p0}} + \sum_{j=0,\, j\neq p}^{n-1} \frac{\lambda_p \lambda_j}{(\lambda_p - \lambda_j)^2 N} \left(\frac{v_{pl} v_{j0}^2}{v_{p0}^3} - \frac{v_{jl} v_{j0}}{v_{p0}^2}\right) \tag{14}$$

그리고

$$\begin{aligned} COV[\hat{\boldsymbol{a}}, \hat{\boldsymbol{a}}]_{lm} &= E[(\hat{a}_l - E[\hat{a}_l])(\hat{a}_m - E[\hat{a}_m])] \\ &\cong \sum_{j=0,\, j\neq p}^{n-1} \frac{\lambda_p \lambda_j}{(\lambda_p - \lambda_j)^2 N} \left(\frac{v_{pl} v_{pm} v_{j0}^2}{v_{p0}^4} + \frac{v_{jl} v_{jm}}{v_{p0}^2} \right. \\ &\quad \left. - \frac{v_{j0}(v_{pl} v_{jm} + v_{pm} v_{jl})}{v_{p0}^3}\right). \end{aligned} \tag{15}$$

여기서 v_i는 행렬 $\boldsymbol{R}_{yy} - \boldsymbol{R}_{yx}\,\boldsymbol{R}_{xx}^{-1}\,\boldsymbol{R}_{xy}$의 i 번째 고유벡터의 j 번째 고유벡터이다. 위 식에서 평균간 분산이 $O(N^{-1})$임을 알 수 있다. 그런데

$$\boldsymbol{R}_{yy} - \boldsymbol{R}_{yx}\,\boldsymbol{R}_{xx}^{-1}\,\boldsymbol{R}_{xy} = \boldsymbol{R}_{zz} - \boldsymbol{R}_{yx}\,\boldsymbol{R}_{xx}^{-1}\,\boldsymbol{R}_{xy} + \sigma_e^2 \boldsymbol{I} \tag{16}$$

이므로, 만약 잡음이 섞이지 않은 경우의 고유치를 λ_j^o라 하면, 고유치의 성질에 의하여

$$\lambda_j = \lambda_j^o + \sigma_e^2 \tag{17}$$

가 되어 다음의 등식을 얻을 수 있다[8, 9].

$$E[\hat{a}_l] \simeq \frac{v_{pl}}{v_{p0}} + \sum_{j=0}^{p-1} \frac{\sigma_e^2(\lambda_j^o + \sigma_e^2)}{\lambda_j^{o2} N} \left(\frac{v_{pl} v_{j0}^2}{v_{p0}^3} - \frac{v_{jl} v_{j0}}{v_{p0}^2} \right) \tag{18}$$

그리고

$$COV[\hat{\boldsymbol{a}}, \hat{\boldsymbol{a}}]_{lm}$$

$$= E[(\hat{a}_l - E[\hat{a}_l])(\hat{a}_m - E[\hat{a}_m])$$

$$\simeq \sum_{j=0}^{p-1} \frac{\sigma_e^2(\lambda_j^o + \sigma_e^2)}{\lambda_j^{o2} N} \left(\frac{v_{pl} v_{pm} v_{j0}^2}{v_{p0}^4} + \frac{v_{jl} v_{jm}}{v_{p0}^2} \right.$$

$$\left. - 2 \frac{v_{j0}(v_{pl} v_{jm} + v_{pm} v_{jl})}{v_{p0}^3} \right) \tag{19}$$

한편 추정된 MA 계수들의 평균과 분산은 다음의 두식으로 주어지는데, 부록 B에 유도과정이 있다.

$$E[\hat{\boldsymbol{b}}] = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{U}\Delta\boldsymbol{a} \tag{20}$$

$$COV[\hat{\boldsymbol{b}}, \hat{\boldsymbol{b}}] = E[(\hat{\boldsymbol{b}} - E[\hat{\boldsymbol{b}}])(\hat{\boldsymbol{b}} - E[\hat{\boldsymbol{b}}])]$$

$$= \boldsymbol{U} COV[\hat{\boldsymbol{a}}, \hat{\boldsymbol{a}}]\ \boldsymbol{U}^T + E[\delta\ \boldsymbol{U}\boldsymbol{a}\boldsymbol{a}^T\ \delta\ \boldsymbol{U}^T] \tag{21}$$

여기서 $\Delta\boldsymbol{a}$는 AR 계수의 편이이며 즉 $\Delta\boldsymbol{a} = \boldsymbol{a} - E[\hat{a}]$ 이고, $U = R_{xx}^{-1} R_{xy}$. 식 (21)에서 $E[\delta\ \boldsymbol{U}\boldsymbol{a}\boldsymbol{a}^T\ \delta\ \boldsymbol{U}^T]$는 간단히 계산되지 않으며 부록 A에 구체적인 식이 주어져 있다. 식 (18), (19), (20) 그리고 (21)에서 다음과 같은 사실을 알 수 있다. 적응 GTLS알고리듬에 의하여 추정된 ARMA 계수의 편이와 분산은 S/N비 즉 잡음 파우어가 증가할수록 커진다. 또한 N이 증가할 수록 편이와 분산은 감소하는데, 편이와 분산값은 1/N에 비례한다. N이 무한대 일 때 편이는 0이므로 GTLS가 출력이 잡음에 오염된 ARMA의 파라미터 출력기로서 편이되지 않은 추정기임을 알 수 있다.

V. 컴퓨터 모의실험 결과

본 절에서는 컴퓨터 모의실험을 통해서 GTLS 알고리듬을 ARMA 시스템 식별에 적용했을 때의 유용성과 제안한 적응 RGTLS알고리듬의 유용성에 대하여 검증할 것이다.

본 컴퓨터 모의실험에서는 다음의 ARMA(2,1) 시스템[7] 을 대상으로 하였다.

$$z_k = 1.1214 z_{k-1} - 0.25 z_{k-2} + 0.05 x_k - 0.4 x_{k-1} \tag{22.1}$$

$$y_k = z_k + e_k \tag{22.2}$$

모의실험에서 입력으로는 백색 가우시안 잡음이 사용되었으며, 부가잡음(additive noise)으로도 가운시안 확률분포함수를 갖는 잡음을 사용하였다

A. LS 그리고 TLS와 GTLS의 비교

LS, TLS 그리고 GTLS의 세가지 추정방법이 위에서 언급한 ARMA 시스템의 계수를 추정하는데 사용되었다. 100 번의 Monte Carlo 시험이 시도되었으며 각 500 개의 샘플을 사용하여 시험하였다. 표 1은 추정된 ARMA 계수들의 평균과 분산을 보여주고 있다. 표에서 괄호안에 들어 있는 값이 표준편차를 나타낸다. 표준편차는 다음의 식을 사용하여 계산되었다.

$$var = \sqrt{E[\hat{c} - E[\hat{c}]]^2}$$

여기서 $\hat{c}$는 ARMA 계수중의 임의의 하나를 가리킨다. 표 1로부터 GTLS에 의해 추정한 계수들의 편이 값의 차수가 10^{-2}임을 알 수 있는데, 이는 추정한 시도값이 500개의 샘플을 사용했을때와 일치한다. 표에 의하면 GTLS의 편이값은 매우 작아서 무시할만함을 알 수 있다. 그러나 LS나 TLS의 경우에 있어서는 추정된 계수들의 편이값이 S/N 비가 떨어지면서 급격히 증가하는 것을 알 수 있다. 특히 LS 방법에 있어서는 추정된 AR 계수들의 값의 편이가 너무 커서 LS 방법이 잡음이 섞인 시스템의 추정에 사용될 수 없음을 알 수 있다.

한편, GTLS방법에 의하여 추정된 계수들의 분산을 생각하여 보자. 표준편차는 S/N 비가 작아지면서 증가하고 있다. 물론 이 표준편차는 식(19)와 식(21)를 사용하여 추측될 수 있다. 입력 파우어가 1 이고 잡음 파우어가 0 일 때, 투영된 출력 에러 벡터들의 공분산 행렬인 $R_{yy} - R_{yx} R_{xx}^{-1} R_{xy}$의 고유벡터들과 고유치들은 간단히 계산되는데 다음과 같다.

$V_0^T = [\ -0.5440,\ -0.6085,\ -0.5777\ ]$,
$V_1^T = [\ 0.5264,\ 0.2886,\ -0.7997\ ]$,
$V_2^T = [\ -0.6534,\ -0.7392,\ -0.1633\ ]$,
$\lambda_0 = 2.1090$, $\lambda_1 = 0.1406$ and $\lambda_3 = 0.0$

표 1. 추정된 ARMA 계수들의 평균과 분산
Table 1. Means and variances of estimated ARMA coefficients.

S/N		GTLS	LS	TLS
30dB	a1	-1.13143(0.0041)	-1.12499(0.0040)	-1.12714(0.0099)
	a2	0.24997(0.0042)	0.24402(0.0041)	0.24540(0.0103)
	b0	-0.04988(0.0019)	-0.04984(0.0019)	-0.04984(0.0025)
	b1	0.39979(0.0013)	0.04037(0.0013)	0.39977(0.0016)
24dB	a1	-1.13142(0.0080)	-1.10637(0.0076)	-1.12624(0.0116)
	a2	0.24989(0.0083)	0.22680(0.0079)	0.24438(0.0122)
	b0	-0.04976(0.0034)	-0.04962(0.0038)	-0.04971(0.0040)
	b1	0.39959(0.0026)	0.39831(0.0026)	0.40045(0.0028)
18dB	a1	-1.13125(0.0156)	-1.04049(0.0145)	-1.12219(0.0174)
	a2	0.24954(0.0160)	0.16653(0.0148)	0.24040(0.0179)
	b0	-0.04951(0.0077)	-0.04898(0.0075)	-0.04987(0.0077)
	b1	0.39918(0.0052)	0.39457(0.0053)	0.40338(0.0055)
12dB	a1	-1.13143(0.0292)	-0.86380(0.0259)	-1.0867(0.0336)
	a2	0.24930(0.0299)	0.01153(0.0255)	0.22766(0.0343)
	b0	-0.04500(0.0154)	-0.04731(0.0142)	-0.05111(0.0160)
	b1	0.39811(0.0105)	0.38495(0.0105)	0.41614(0.0109)
6dB	a1	-1.13136(0.0680)	-0.59510(0.0344)	-1.04300(0.0727)
	a2	0.24810(0.0722)	-0.18095(0.0322)	0.16708(0.0738)
	b0	-0.04788(0.0308)	-0.03416(0.0257)	-0.05703(0.0368)
	b1	0.39690(0.0210)	0.37061(0.0264)	0.47831(0.0250)

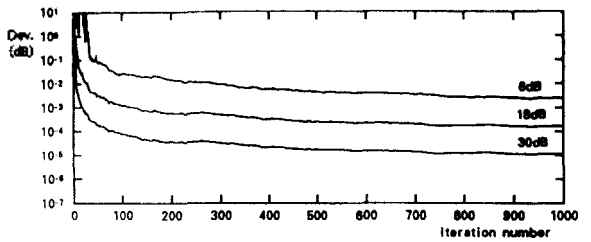
식(19)와 위의 고유벡터 및 고유치들을 사용하여 AR 계수들의 이론적 표준편차를 계산하였다. 이론적 표준편차는 표2에 나타내 보였다. 컴퓨터 모의실험결과와 표2의 이론적 표준편차와 비교하여 보면, S/N 비가 높을 때는 이론치가 실험치와 일치하나 S/N가 저하되면서 이론치와 실험치사이의 차이가 커짐을 알 수 있다. 표준편차식 (19)는 타당하다고 볼 수 있다. MA 계수들에 대한 이론치에 대한 계산은 하지 않았는데, 이를 위해서는 다중적분의 복잡한 계산을 필요로 한다. 하여간 실험대상으로 사용한 시스템에 대해서는 MA 계수의 분산이 AR 계수의 것에 비하여 작음을 알 수 있었다.

표 2. GTLS 방법에 의해 추정된 AR 계수의 이론적 분산
Table 2. Theoretical variances of AR coefficients estimated by GTLS method.

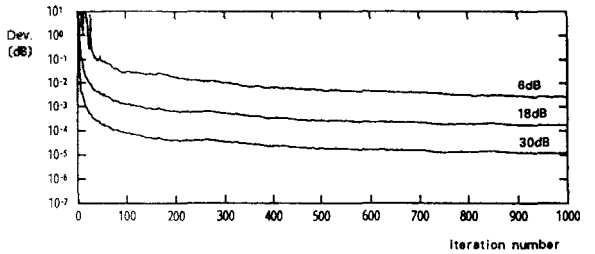
S/N	a1	a2
30dB	0.0041	0.0035
24dB	0.0084	0.0072
18dB	0.0176	0.0154
12dB	0.0416	0.0382
6dB	0.1206	0.1182

B. 적응 GTLS 알고리듬의 모의실험결과

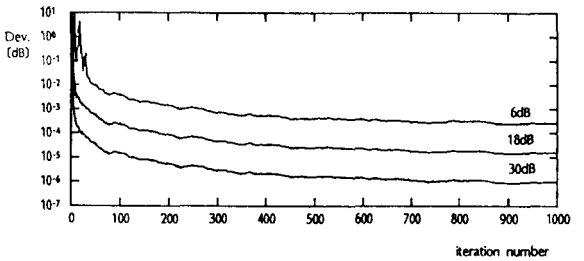
식(11)로 주어진 시스템에 대하여 적응알고리듬에 의한 식별실험을 수행하였다. 모의실험에서 PM을 위한 반복횟수는 2로서 작은 값을 사용하였다. 그림 2는 S/N 비가 6, 18 그리고 30일 때의 ARMA 계수들의 편차값을 보여주고 있다. 그림3은 추정된 계수들의 참값으로부터의 편이를 보여주고 있다. 그림들에서 y축 스케일은 일반지수함수 스케일이다. 표준편차와 편이는 50번의 시도로부터 구해진 것이다. 그림들은 RGTLS 알고리듬이 매우 빠른 수렴성을 가지고 있음과 S/N비가 커짐에 따라 편이가 매우 작아지고 표준편차가 줄어듬을 알 수 있다. 또한 표준편차는 시도 횟수 N에 비례함을 알 수 있다. 컴퓨터 모의실험 A와 B의 결과를 비교하여 볼 때 우리는 N이 500일 때의 표준편차와 공분산이 거의 같음을 관찰할 수 있다. 한가지 적응 GTLS알고리듬이 초기 적응과정에서 추정계수들이 불안해짐을 알 수 있었다. 특히 S/N비가 낮을 때에 이러한 현상이 생겼다. 이러한 현상에 대한 연구가 앞으로 필요하겠고 이러한 경우에 대처할 수 있는 견인한 방법이 요구된다.



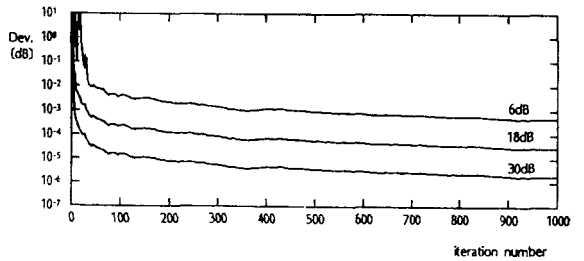
(a)



(b)



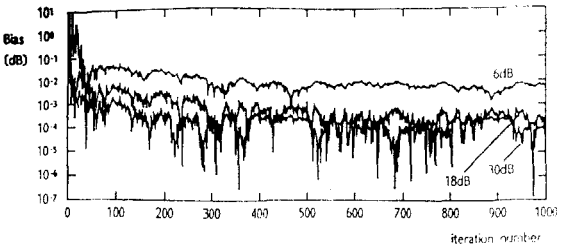
(c)



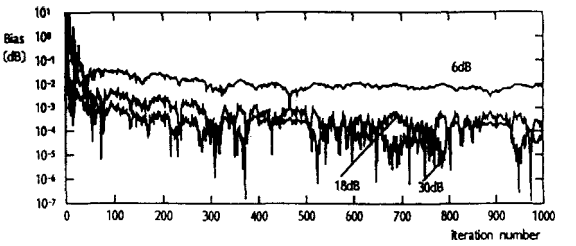
(d)

그림 2. 추정된 ARMA 계수의 편차
(a) a_1의 편차 (b) a_2의 편차
(c) b_0의 편차 (d) b_1의 편차

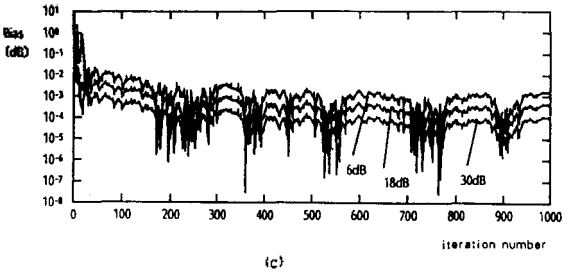
Figure 2. Deviations of estimated ARMA coefficients
(a) deviation of a_1 (b) deviation of a_2
(c) deviation of b_0 (d) deviation of b_1



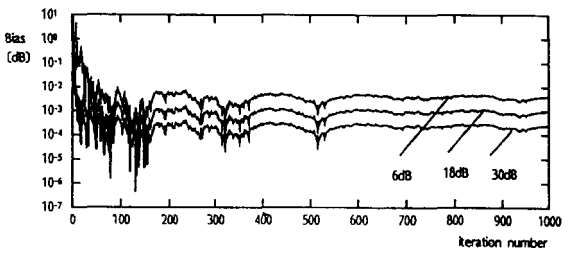
(a)



(b)



(c)



(d)

그림 3. 추정된 ARMA 계수의 편이
(a) a_1의 편이 (b) a_2의 편이
(c) b_0의 편이 (d) b_1의 편이

Figure 3. Biases of estimated ARMA coefficients
(a) bias of a_1 (b) bias of a_2
(c) bias of b_0 (d) bias of b_1

C. 적응 GTLS 알고리듬과 RLS(recursive least squares) 알고리듬과의 비교

RLS 알고리듬은 파라미터의 적응추정방법으로 지금까지 가장 널리 사용되고 있는 방법이다. 따라서 본 절에서는 제안한 RGTLS 방법을 RLS 방법과 간단한 비교를 통하여 제안한 방법의 우수함을 보이고자 한다. 먼저 우리는 제 Ⅳ장에서 RGTLS 알고리듬의 섭동분석을 통하여 RGTLS가 편이되지 않은 추정기임을 보였으며 (식19)와 (식20)으로 추정된 파라미터의 분산을 구한 바 있다. 한편, RLS 추정방법의 bias는 만약 지수가중인자(exponential weighting factor) λ가 1 이라면[6, p383], LS 방법과 동일하게 되어 LS 방법의 bias와 동일하다. 따라서 추정반복회수가 500 인 경우의 bias는 LS의 결과인 표 1의 추정결과와 동일하여, RLS의 경우에는 추정된 파라미터의 bias가 S/N비가 증가함에 따라서 상당히 큼을 알 수 있다. 한편 RLS 방법에서 추정된 파라미터의 상관행렬은 다음과 같은 근사식으로 주어진다.

$$E[(\hat{\theta}_l-\theta_l)^2] \simeq \frac{\sigma^2}{n} R^{-1}{}_{LS}(l, l) \tag{23}$$

여기서 $\theta=[a_1\ a_2 \cdots a_p\ b_0\ b_1 \cdots b_q]^T$ R_{LS}는 LS방법을 풀기 위해 사용되는 $(p+q+1)\times(p+q+1)$의 상관행렬인데 Ⅱ장에서 정의한 행렬 $R=\begin{bmatrix} R_{yy} & R_{yx} \\ R_{xy} & R_{xx} \end{bmatrix}$에서 첫번째 행과 열을 제외시킨 행렬이고, 즉,

$$R_{LS}=\begin{bmatrix} R_{2,2} & R_{2,3} & \cdots & R_{2,p+q+2} \\ R_{3,2} & R_{3,3} & \cdots & R_{3,p+q+2} \\ \vdots & \vdots & \cdots & \vdots \\ R_{p+q+2,2} & R_{p+q+2+3,3} & \cdots & R_{p+q+2,p+q+2} \end{bmatrix}$$

이고 σ^2은 추정 오차의 전력(power)인데 부가잡음의 전력이 아니라 LS방법을 사용하여 파라미터를 추정했을 때의 오차의 전력이며, $R^{-1}{}_{LS}(l, l)$은 $R^{-1}{}_{LS}$의 (l, l) 원소이다[6]. 이때, LS 추정오차 σ^2은 다음의 식으로 구하여 진다.

$$\sigma^2=[\sigma_x^2+\sigma_e^2]-\boldsymbol{r}^T \boldsymbol{R}_{LS}^{-1}\boldsymbol{r} \tag{24}$$

여기서 $\boldsymbol{r}$은 행렬 $\boldsymbol{R}$의 첫번째 행벡터에서 첫번째 원소를 제외시킨 벡터이고, σ_e^2는 부가잡음의 파우어 그리고 σ_x^2는 잡음이 섞이지 않은 출력의 파우어이다.

(식23)으로 주어진 RLS 방법의 수렴특성을 S/N 비가 6 dB이고 n이 500인 경우에 계산하여 보면 표 3과 같다. 표 3을 표 1의 결과와 비교하여 보면 n이 500일 때 (식22)로 주어진 RLS의 수렴특성의 결과가 n이 500일 때의 LS의 시뮬레이션의 결과와 같음을 알 수 있다. 즉 RLS와 RGTLS의 수렴특성은 표 1의 결과를 사용하여 유추할 수 있다. 표 1의 결과를 분석하여 보면 다음과 같이 정리할 수 있다.

표 3. RLS 방법에 있어서 추정파라미터의 표준편차
Table 3. Standard deviation of estimated coefficients by RLS method
(S/N Ratio = 6dB, 반복횟수 = 500)

파라미터	이론적 표준편차
a0	0.0342
a1	0.0342
b0	0.0239
b1	0.0240

S/N 비가 높을 때에는 RLS 방법과 RGTLS 방법의 수렴특성이 유사하다. 그러나 S/N 비가 낮을 때에는 수렴특성의 측면에서 고찰할 때 RLS의 성능이 우수하다. 그러나 RLS의 방법은 상당히 편이된 해를 구하게 되므로 수렴특성이 좋다고 하더라도 S/N 비가 낮을 때에는 사용할 수 없게 된다. 결국 RLS와 RGTLS의 성능을 비교하여 볼 때 RGTLS의 성능이 우수함을 알 수 있다.

Ⅵ. 결 론

본 논문에서는 GTLS 문제가 ARMA 시스템의 계수를 구하기 위한 방법으로 새로이 적용되었으며, 적응 GTLS 알고리듬이 ARMA 시스템의 파라미터를 구하기 위한 적응 알고리듬으로서 제안되었다. 또한 GTLS 및 적응 GTLS알고리듬에 대한 섭동분석이 이루어졌다. GTLS 및 적응 GTLS 알고리듬의 성능이 컴퓨터 모의실험을 통하여 평가되었다. 실험결과 GTLS 방법이 잡음이 섞인 ARMA시스템에 대해서 시스템 계수를 구하는데 계수의 편이와 표준편차의 측면에서 LS나 TLS 방법보다 우수함을 보였다. 그리고 적응 GTLS알고리듬이 매우작은 계수편이와 항상 최적의 해에 수렴하며, 수렴속도가 빠름을 보였다.

부록 A

본 부록에서는 참고문헌 [6]의 섭동 formulation과 참고문헌 [9]의 통계학적 결과를 이용하여 GTLS방법에 의하여 추정된 AR 계수들의 $O(N^{-1})$의 첫번째 및 두번째의 모멘트(first and second moment)를 유도해낸다. 다음의 행렬 $\boldsymbol{R}$을 정의하자.

$$\boldsymbol{R}=\boldsymbol{R}_{yy}-\boldsymbol{R}_{yx}\,\boldsymbol{R}_{xx}^{-1}\,\boldsymbol{R}_{xy} \tag{A.1}$$

Wilkinson's 방법[6]에 따르면 (p+1) by (p+1) 행렬 $\boldsymbol{R}$을 섭동계수 a를 가진 불규칙섭동(random perturbation)으로 표현할 수 있다. 다음과 같다. 단, α는 $0<\alpha<1$ 이다.

$$\hat{\boldsymbol{R}}=\boldsymbol{R}+(\hat{\boldsymbol{R}}-\boldsymbol{R})=\boldsymbol{R}+q\,\boldsymbol{B} \tag{A.2}$$

여기서 $\boldsymbol{B}$는 Hermitian이며, 행렬의 각원소가 근사적으로 샘플행렬 $\hat{\boldsymbol{R}}_{yy}$, $\hat{\boldsymbol{R}}_{xx}$, $\hat{\boldsymbol{R}}_{xy}$와 서로 독립인 평균이 0인 불규칙 행렬(random matrix)이다. 행렬 $\hat{\boldsymbol{R}}$는 (p+1) 개의 고유치와 그에 상응하는 고유벡터를 갖는다. 그러면 행렬 $\hat{\boldsymbol{R}}$의 가장 작은 고유치에 상응하는 고유벡터 $\hat{\boldsymbol{v}}_p$는 다음의 식으로 표시된다[9].

$$\hat{\boldsymbol{v}}_p\cong(1-\frac{1}{2}\sum_{j=0}^{p-1}|t_{1j}^{(p)}|^2q^2)\,\boldsymbol{v}_p+\sum_{j=0}^{p-1}t_{1j}^{(p)}q\boldsymbol{v}_j+\sum_{j=0}^{p-1}t_{2j}^{(p)}q^2\boldsymbol{v}_j \tag{A.3}$$

여기서 $t_{kj}^{(i)}$, $k=1, 2, \ldots$ 은 섭동계수들이다. 그러면 l번째 AR 계수의 근사값은 다음과 같이 표현된다.

$$\hat{a}_l=\frac{\hat{v}_{pl}}{\hat{v}_{p0}}\cong\frac{(1-\frac{1}{2}\sum_{j=0}^{p-1}|t_{1j}^{(p)}|^2q^2)\,v_{pl}+\sum_{j=0}^{p-1}t_{1j}^{(p)}qv_{jl}+\sum_{j=0}^{p-1}t_{2j}^{(p)}q^2v_{jl}}{(1-\frac{1}{2}\sum_{j=0}^{p-1}|t_{1j}^{(p)}|^2q^2)\,v_{p0}+\sum_{j=0}^{p-1}t_{1j}^{(p)}qv_{j0}+\sum_{j=0}^{p-1}t_{2j}^{(p)}q^2v_{j0}} \tag{A.4}$$

위의 식을 Taylor series를 사용하여 전개한 후 q^2의 항까지로 근사시키면 다음과 같이 나타낼 수 있다.

$$\hat{a}_l\cong\frac{\boldsymbol{v}_{pl}}{\boldsymbol{v}_{p0}}-\sum_{j=0}^{p-1}t_{1j}^{(p)}\,q\,\frac{\boldsymbol{v}_{jl}\boldsymbol{v}_{pl}}{\boldsymbol{v}_{p0}^2}+\sum_{j=0}^{p-1}t_{1j}^{(p)}\,q\,\frac{\boldsymbol{v}_{pl}}{\boldsymbol{v}_{p0}}-\sum_{j=0}^{p-1}t_{2j}^{(p)}\,q^2\,\frac{\boldsymbol{v}_{jl}\boldsymbol{v}_{pl}}{\boldsymbol{v}_{p0}^2}-\sum_{j_1=0}^{p-1}\sum_{j_2=0}^{p-1}t_{1j_1}^{(p)}t_{1j_2}^{(p)}\,q^2\,\frac{\boldsymbol{v}_{j_2l}\boldsymbol{v}_{j_10}}{\boldsymbol{v}_{p0}^2}+\sum_{j_1=0}^{p-1}\sum_{j_2=0}^{p-1}t_{1j_1}^{(p)}t_{1j_2}^{(p)}\,q^2\,\frac{\boldsymbol{v}_{j_10}\boldsymbol{v}_{j_20}\boldsymbol{v}_{pl}}{\boldsymbol{v}_{p0}^3} \tag{A.4}$$

$t_{ij}^{(p)}$의 식은 식(A.4)을 고유치 문제에 있어 섭동등식(perturbation equation)에 대입하여 구할 수 있는데, 문헌 [10]에 따르면 아래와 같다.

$$t_{1j}^{(p)}=\frac{\boldsymbol{v}_j^H\,\boldsymbol{B}\,\boldsymbol{v}_p}{\lambda_j-\lambda_p} \tag{A.5}$$

그리고

$$t_{2j}^{(p)}=\frac{\boldsymbol{v}_p^H\,\boldsymbol{B}\,\boldsymbol{v}_p\,\boldsymbol{v}_j^H\,\boldsymbol{B}\,\boldsymbol{v}_p}{(\lambda_j-\lambda_p)^2}-\sum_{k=0}^{p-1}\frac{\boldsymbol{v}_j^H\,\boldsymbol{B}\,\boldsymbol{v}_k\,\boldsymbol{v}_k^H\,\boldsymbol{B}\,\boldsymbol{v}_p}{(\lambda_j-\lambda_p)(\lambda_k-\lambda_p)} \tag{A.6}$$

추정된 AR 계수의 모멘트를 구하기 위하여 또한 다음의 결과를 이용하자[10].

$$E[(\boldsymbol{A}_1^H\,\boldsymbol{B}\,\boldsymbol{A}_2)(\boldsymbol{A}_3^H\,\boldsymbol{B}\,\boldsymbol{A}_4)]=\frac{(\boldsymbol{A}_1^H\,\boldsymbol{R}\,\boldsymbol{A}_4)(\boldsymbol{A}_3^H\,\boldsymbol{R}\,\boldsymbol{A}_2)}{Nq^2} \tag{A.7}$$

그리고

$$\boldsymbol{v}_i^H\,\boldsymbol{R}\,\boldsymbol{v}_j=\lambda_i\,\delta_{ij} \tag{A.8}$$

그러면 다음의 식과 같이 l 번째 AR 계수의 편이값을 구할 수 있다.

$$E[\hat{a}_l]\cong\frac{\boldsymbol{v}_{pl}}{\boldsymbol{v}_{p0}}+\sum_{j=0}^{p-1}\frac{\lambda_p\lambda_j}{(\lambda_p-\lambda_j)^2N}\left(\frac{v_{pl}v_{j0}^2}{v_{p0}^3}-\frac{v_{jl}v_{j0}}{v_{p0}^2}\right) \tag{A.9}$$

한편 l 번째 AR 계수의 두번째 모멘트에 대한 N^{-1} 차수의 근사식은 다음과 같다.

$$(\hat{a}_l-E[\hat{a}_l])(\hat{a}_m-E[\hat{a}_m])\cong$$

$$\left(-\sum_{j=0}^{p-1}t_{1j}^{(p)}\,q\,\frac{v_{jl}v_{pl}}{v_{p0}^2}+\sum_{j=0}^{p-1}t_{1j}^{(p)}\,q\,\frac{v_{pl}}{v_{p0}}\right.$$

$$\left.-\sum_{j=0}^{p-1}t_{2j}^{(p)}\,q^2\,\frac{v_{jl}v_{pl}}{v_{p0}^2}-\sum_{j_1=0}^{p-1}\sum_{j_2=0}^{p-1}t_{1j_1}^{(p)}t_{1j_2}^{(p)}\,q^2\,\frac{v_{j_2l}v_{j_10}}{v_{p0}^2}\right.$$

$$+\sum_{j_1=0}^{p-1}\sum_{j_2=0}^{p-1} t_{1j_1}^{(p)}\, t_{1j_2}^{(p)}\, q^2\,\frac{v_{j_1 0}\, v_{j_2 0}\, v_{pl}}{v_{p0}^3} - \Delta a_l)\cdot$$

$$(-\sum_{j=0}^{p-1} t_{1j}^{(p)}\, q\,\frac{v_{jm}\, v_{pm}}{v_{p0}^2} + \sum_{j=0}^{p-1} t_{1j}^{(p)}\, q\,\frac{v_{pm}}{v_{p0}}$$

$$-\sum_{j=0}^{p-1} t_{2j}^{(p)}\, q^2\,\frac{v_{jm}\, v_{pm}}{v_{p0}^2} - \sum_{j_1=0}^{p-1}\sum_{j_2=0}^{p-1} t_{1j_1}^{(p)}\, t_{1j_2}^{(p)}\, q^2\,\frac{v_{j_1 m}\, v_{j_2 0}}{v_{p0}^2}$$

$$+\sum_{j_1=0}^{p-1}\sum_{j_2=0}^{p-1} t_{1j_1}^{(p)}\, t_{1j_2}^{(p)}\, q^2\,\frac{v_{j_1 0}\, v_{j_2 0}\, v_{pm}}{v_{p0}^3} - \Delta a_m) \quad \text{(A.10)}$$

여기서 $\Delta a_l = a_l - E[\hat{a}_l]$이다. 따라서 위의 편이를 구할 때 사용한 것과 유사한 방법을 사용하여 다음과 같은 공분산 행렬을 얻을 수 있다.

$$COV[\hat{a}, \hat{a}]_{lm} = E[(\hat{a}_l - E[\hat{a}_l])(\hat{a}_m - E[\hat{a}_m])$$

$$\cong \sum_{j=0}^{p-1} \frac{\lambda_p \lambda_j}{(\lambda_p - \lambda_j)^2 N}$$

$$\left(\frac{v_{pl}\, v_{pm}\, v_{j0}^2}{v_{p0}^4} + \frac{v_{jl}\, v_{jm}}{v_{p0}^2} - \frac{v_{j0}(v_{pl}\, v_{jm} + v_{pm}\, v_{jl})}{v_{p0}^3}\right) \quad \text{(A.11)}$$

위의 식 (A.11)은 차수 N^{-2} 이상의 항을 무시함으로써 구하여 진다.

부록 B

부록 B에서는 MA계수들에 대한 첫번째 그리고 두번째 모멘트에 대하여 기술한다. 식(14)로부터 MA 계수들은 다음과 같다.

$$\boldsymbol{b} = -\boldsymbol{R}_{xx}^{-1}\,\boldsymbol{R}_{xy}\,\boldsymbol{a} \quad \text{(B.1)}$$

$\boldsymbol{U} = -\boldsymbol{R}_{xx}^{-1}\,\boldsymbol{R}_{xy}$ 그리고 $\boldsymbol{b} = \boldsymbol{U}\,\boldsymbol{a}$라고 정의하자. 그러면 섭동된 MA 계수들은 다음의 식으로 표현되어 진다.

$$\hat{\boldsymbol{b}} = \hat{\boldsymbol{U}}\hat{\boldsymbol{a}} = (\boldsymbol{U} + \delta\boldsymbol{U})(\boldsymbol{a} + \delta\boldsymbol{a})$$
$$= \boldsymbol{U}\boldsymbol{a} + \boldsymbol{U}\delta\boldsymbol{a} + \delta\boldsymbol{U}\boldsymbol{a} + \delta\boldsymbol{U}\delta\boldsymbol{a} \quad \text{(B.2)}$$

여기서 $\delta\boldsymbol{U}$는 행렬 $\boldsymbol{U}$의 섭동행렬이고 $\delta\boldsymbol{a}$는 벡터 $\boldsymbol{a}$의 섭동벡터이다. 그런데 행렬 $\delta\boldsymbol{U}$의 평균은 영행렬이고, 벡터 $\delta\boldsymbol{a}$가 부록 A에서 정의된 $\boldsymbol{B}$의 함수이고 $\boldsymbol{B}$의 확률밀도함수가 $\delta\boldsymbol{U}$의 함수와 다르므로[11, pp319], $\delta\boldsymbol{a}$와 독립이다. 따라서 추정된 MA 계수들의 기대값은 다음과 같이 표현될 수 있다.

$$E[\hat{\boldsymbol{b}}] = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{U}\Delta\boldsymbol{a}, \quad \text{(B.3)}$$

$$\Delta\boldsymbol{b} = \boldsymbol{b} - E[\hat{\boldsymbol{b}}] = \boldsymbol{U}\Delta\boldsymbol{a} \quad \text{(B.4)}$$

여기서 $\Delta\boldsymbol{a}$는 AR계수의 편이값이며 $\Delta\boldsymbol{b}$는 MA 계수의 편이값이다. 벡터 $\hat{\boldsymbol{b}} - E[\hat{\boldsymbol{b}}]$의 공분산은 다음과 같이 구하여진다.

$$COV[\hat{\boldsymbol{b}}, \hat{\boldsymbol{b}}] = E[(\hat{\boldsymbol{b}} - E[\hat{\boldsymbol{b}}])(\hat{\boldsymbol{b}} - E[\hat{\boldsymbol{b}}])]$$
$$= E[(\boldsymbol{U}\delta\boldsymbol{a} + \delta\boldsymbol{U}\boldsymbol{a} + \delta\boldsymbol{U}\delta\boldsymbol{a} - \Delta\boldsymbol{b})$$
$$(\boldsymbol{U}\delta\boldsymbol{a} + \delta\boldsymbol{U}\boldsymbol{a} + \delta\boldsymbol{U}\delta\boldsymbol{a} - \Delta\boldsymbol{b})^T]$$
$$= E[(\boldsymbol{U}(\delta\boldsymbol{a} - \Delta\boldsymbol{a}) + \delta\boldsymbol{U}\boldsymbol{a} + \delta\boldsymbol{U}\delta\boldsymbol{a})$$
$$(\boldsymbol{U}(\delta\boldsymbol{a} - \Delta\boldsymbol{a}) + \delta\boldsymbol{U}\boldsymbol{a} + \delta\boldsymbol{U}\delta\boldsymbol{a})^T] \quad \text{(B.5a)}$$

위 식에서 $\delta\boldsymbol{a}$와 $\delta\boldsymbol{U}$가 서로 독립임과 섭동에 대한 3차 이상의 항을 무시하면 다음과 같은 식으로 정리할 수 있다.

$$COV[\hat{\boldsymbol{b}}, \hat{\boldsymbol{b}}]$$

$$\cong E[\boldsymbol{U}(\delta\boldsymbol{a} - \Delta\boldsymbol{a})(\delta\boldsymbol{a} - \Delta\boldsymbol{a})^T\boldsymbol{U}^T] + E[\delta\boldsymbol{U}\boldsymbol{a}\,\boldsymbol{a}^T\,\delta\,\boldsymbol{U}^T]$$

$$= \boldsymbol{U}COV[\hat{\boldsymbol{a}}, \hat{\boldsymbol{a}}]\,\boldsymbol{U}^T + E[\delta\boldsymbol{U}\boldsymbol{a}\,\boldsymbol{a}^T\,\delta\,\boldsymbol{U}^T] \quad \text{(B.5b)}$$

그리고

$$E[\delta\,\boldsymbol{U}\boldsymbol{a}\boldsymbol{a}^T\,\delta\,\boldsymbol{U}^T]$$
$$= \int\int \delta\,\boldsymbol{U}\boldsymbol{a}\boldsymbol{a}^T\,\delta\,\boldsymbol{U}^T\, f_{\delta U|T}(\delta U|T)\, f_T(T)\, dT\, dU \quad \text{(B.6)}$$

여기서

$$f_{\delta U|T} =$$

$$\frac{\exp(\dfrac{-1}{2}\, tr\, T\delta\boldsymbol{U}(\boldsymbol{R}_{yy} - \boldsymbol{R}_{yx}\,\boldsymbol{R}_{xx}^{-1}\,\boldsymbol{R}_{xy})^{-1}\,\delta\,\boldsymbol{U}^T)}{(2\pi)^{(q+1)(p+1)/2}\,|T^{-1}|^{(p+1)/2}\,|\boldsymbol{R}_{yy} - \boldsymbol{R}_{yx}\,\boldsymbol{R}_{xx}^{-1}\,\boldsymbol{R}_{xy}|^{(q+1)/2}} \quad \text{(B.7)}$$

$$f_T = \frac{|T|^{(N-q-2)/2}\exp(-\frac{1}{2}\,tr\,R_{xx}^{-1}T)}{2^{N(q+1)/2}\pi^{q(q+1)/4}|R_{xx}|^{N/2}\prod_{i=1}^{q+1}\Gamma[\frac{1}{2}(N+1-i)]} \quad (B.8)$$

이다[11]. 위의 식 (B. 5)에서 MA 계수의 공분산식도 N^{-1}의 근사식임을 알 수 있다.

참 고 문 헌

1. S. L. Marple, *Digital Spectral Analysis with Application*, Prentice Hall, 1983.
2. T. Soderstorm and P. Stoica, *System identification*, Prentice-Hall, 1989.
3. M. A. Rahman and K. Yu, "Total least square approach for frequency estimation using linear prediction," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, No. 10, pp. 1440-1454, 1987.
4. G. H. Golub and C. F. Van Loan, "Analysis of the total least square problem," *SIAM J. Numer. Anal.*, vol. 17, pp. 883-893, 1980.
5. S. V. Huffel and J. Vandelwalle, "Analysis and properties of the generalized total least squares problem when some or all columns in A are subject to error," *SIAM J. Matrix Anal.*, vol. 10, No. 3, pp. 294-315, 1989.
6. S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 1986.
7. J. J. Shynk, "Adaptive IIR filter," *IEEE Acoustics, Speech, and Signal Processing Magazine*, pp. 2-21, April, 1989.
8. Basilevsky, *Applied Matrix Algebra in the Statistical Science*, North-Holland, 1983.
9. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, 1965.
10. M. Kaveh and A. J. Barbell, "The statistical performance of the MUSIC and the Minimum-norm algorithms in resolving plane waves," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-34, No. 2, pp. 331-341, 1986.
11. S. F. Aranold, *The Theory of Linear Models and Multivariate Analysis*, John Wiley & Sons, 1981.

▲김 재 인

14권 1호 참조

현재 : 한국통신 S/W 연구소 음성언어연구팀 선임연구원

▲김 진 영

14권 1호 참조

현재 : 전남대학교 공과대학 전자공학과 전임강사

▲이 태 원

14권 1호 참조

현재 : 고려대학교 공과대학 전자공학과 교수