

適應 ARMA 모형의 적용에 관한 고찰

이 대 호
전기및 전자공학과
(1983. 4. 30 접수)

<요 약>

ARMA 모형에 의한 음성 분석에 있어서 가장 곤란점의 하나인 입력추정에 대해서 검토하고 보조수단으로서 임펄스 추정 알고리즘을 제안하였다. 제한된 조건하에서 제안된 알고리즘은 극히 우수한 특성을 보였으나 조건의 변화에 따라 그 특성이 급격히 열화되는 것이 관측되었으며 일반적 적용은 타당하지 않음이 판명되었다. 아울러 파라미터 추적과정에서 수렴에 적절한 신호의 진폭과 수렴판정의 척도에 대해서 평가와 제안을 실었다.

On the Application of the Adaptive ARMA Model

Lee, Tai-ho

Dept. of Electrical and Electronic Engineering

(Received April 30, 1983)

<Abstract>

The major difficulty in analyzing the voice signal by ARMA model lies on the fact that the excitation signal can not be separated from the estimation error. An algorithm appropriate to estimate isolated impulse type excitation signal is presented along with the brief discussion of recent work related to the topic.

The proposed algorithm exhibits highly reliable performance under the favorable conditions. However it has been found the algorithm is so sensitive to the change of condition that its applicability is limited to the special purposes. Also, the effect of the signal amplitude on the convergence, and the measure of convergence are discussed, and suggestions are given for better use of ARMA model for voice signal analysis.

I. 서 론

음성신호의 처리에 있어서 파라미터를 정확히 추출한다는 일은 일차적 중요성을 가진다. 신호의 전송을 위한 符號化, 인식용 위한 판단 기준 및 합성을 위한 원료가 모두 정확한 파라미터를 요구하기 때문이다. 여러가지 수단 중에서도, 단순하고 비교적 믿을만한 자료를 얻을 수 있는 선형예측기술(linear prediction technic)이 주로 많이 이용된

다. [1], [2] 선형예측기법은窓(window)의 이용이 불가피하기 때문에窓의 효과가 대상신호의 특성에 중첩되는 것을 피할 수 없고 또 분석구간을 좁힐수록 이 효과는 증대되므로 비교적 넓은 구간을 요청한다. [3] 따라서 얻어지는 결과는 평균적 특성이 되어 포맷트의 내역폭을 알기 어렵게 된다. 최근에 ARMA(autoressive and moving average) 모형을 사용하는 파라미터의 추정이 활발히 논의되는 것은 이와 같은 문제를 해소하려는 한 노력이라고 볼 수 있다. [4], [5] 적응(adaptive) ARMA 모

형의 수단은 선형예측에서와 마찬가지로 추정값과 실제값 사이의 자승오차를 최소화하려는 것으로

$$\frac{\partial E}{\partial \phi_i} = 0 \quad (1)$$

을 만족하는 한 組의 파라메터 $\Phi = \{\phi_i\}$ 를 구하는 과정이며 여기서, $E = \sum v_k^2$ 이고 v_k' 은 추정오차로서 다음과 같이 표현된다.

$$v_k' = s_k + s_{k/k-1} \quad (2)$$

또한,

$$s_{k/k-1} = S_k^T \Phi_{k-1}$$

$$S_k^T = (s_{k-1} \ s_{k-2} \ \dots \ s_{k-N} \ \hat{u}_{k-1} \ \dots \ \hat{u}_{k-M})$$

s_k : k 번째 음성표본

\hat{u}_k : k 번째 추정입력

$$\Phi_k^T = (\phi_{k1}, \dots, \phi_{k, N+M})$$

이며 N 과 M 은 각각 극점과 영점의 수, 그리고 Φ_k 가 구하려는 파라메터 벡터이다. 분석의 과정은 음성표본의 순서를 따라 순차적으로 추정을 진화시키므로 파라메터 값이 충분히 수렴되면 선형예측에서와는 달리 구간의 평균적 특성이 아니라 최종 표본 부근의 좁은 구간의 정확한 특성에 좀더 접근할 것이 기대된다. 그러나 실제 적용의 결과는 복잡한 보조수단은 동일한 경우에도 현저한 개선은 보이지 않는다. [5] 중요한 이유는 추정입력 \hat{u}_k 의 추출이 곤란하다는 점, 실제적 음성이 완전한 定常상태 (steady state)를 유지하기 어렵다는 점, 그리고 또한 예측자수를 정확히 알기 어렵다는 점 들이며, 이들은 선형예측에서도 문제가 되는 것이지만, ARMA 모형의 경우에는 파라메터 수렴을 방해하므로 근본적인 문제를 야기한다. 본 연구에서는 입력 추출의 한가지 보조수단을 제시하여 그 특성을 평가하고, 그밖에 신호의 크기의 영향, 수렴판정의 수단에 대해서 검토하여 ARMA 모형을 적용하는데 필요한 자료를 제공한다.

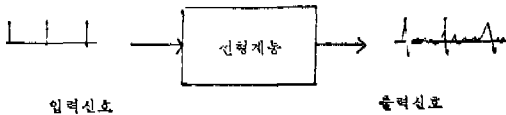


그림 1. 신호발생계통

II. 입력신호의 추정

1. 입력신호의 영향

식(1)로부터 얻어지는 파라메터 추정공식은 다음과 같다. [6]

$$\Phi_k = \Phi_{k-1} + P_k S_k (1 + S_k^T P_k S_k)^{-1} v_k \quad (3)$$

여기서 v_k 는 교정입력이며 P_k 는 陽行列로서 다음식을 따른다.

$$P_k = P_{k-1} - P_k S_k (1 + S_k^T P_k S_k)^{-1} S_k^T P_k \quad (4)$$

식(3)과(4)에 의한 추정은 150~200 표본정도에서 수렴이 되고 [4], 또한 안정된 구간에서는 60~70 표본정도면 수렴되는 것을 볼 수 있다. 결과적으로 얻어지는 파라메터 Φ_k 는 매 분석구간의 최종 부분에 관한 한 오차가 작은 좋은 추정값이 될 것이다. 그러나 실제 음성에 적용된 결과는 현저한 파라메터에 대한 개선이 보이지 않는다. [5] 가장 큰 곤란은 그림 1에서 대장 시스템의 입력신호를 정확히 추적할 수 없다는 점이다. ARMA 모형은 대장 시스템의 입력을 알고 있을 것으로 간주되어 이루어지는 것이며 [6] 그렇지 못한 때는 입력추정을 도달하는 모형을 이용할 수 밖에 없는데 추진오가의 입력을 분리해 내는 것은 불가능한 것이기 때문에 입력신호가 있는 순간 추정과정에는 교란이 발생하며 다시 안정하는데 상당기간이 걸리게 된다. 편제 시도되고 있는 두가지 모형을 일별하면 다음과 같다.

Morikawa와 Fujisahi [4]에 의한 모형은 $v_k = v_k'$ 으로 하여 식(3)에 적용하고 새로이 얻어진 Φ_k 를 이용하여

$$\hat{u}_k = s_k - S_k^T \Phi_k \quad (5)$$

에 의하여 추정입력을 얻어내고 있다.

한편 Miyaraga 등 [5]은 '수정 MRAS' 방식이라는 이름으로 2개모형의 병렬운전에 의한 모음 파라메터 추정을 시도하고 있다. 여기서는 $v_k' = v_k + \hat{u}_k$ 의 형태로 추정입력성분과 교정입력, 즉 순수한 오차성분을 분리하려는 의도를 나타낸 것이지만 실제로는 단순히, 어느 한계를 초과하는 v_k' 에 대해서는 이를 피치입력으로 보고 파라메터의 수정은 약 2ms 간 보류하며 그 밖의 경우에만 적절한 인자를 곱해서 추정입력으로 받아들이도록 하고 있다.

따라서 '수정 MRAS'는 적용대상신호의 성격을 상당히 알고 있는 상황에서 가능하며 특히 잡음요인이 강하게 작용하는 子음성의 경우나 二重피치현상이 발견되는 여성음의 경우에는 적용이 불가능하거나 효과가 거의 없게 된다. 위의 두 수단중 수렴을 촉진한다는 점에서는 식(5)의 目的의 수단이 더 우수하다. 그림 2c는 식(5)에 의한 추정오차와 추정입력을 보인 것으로 입력신호 이후 상당한 기간의 교란 상태가 있음을 보여준다.

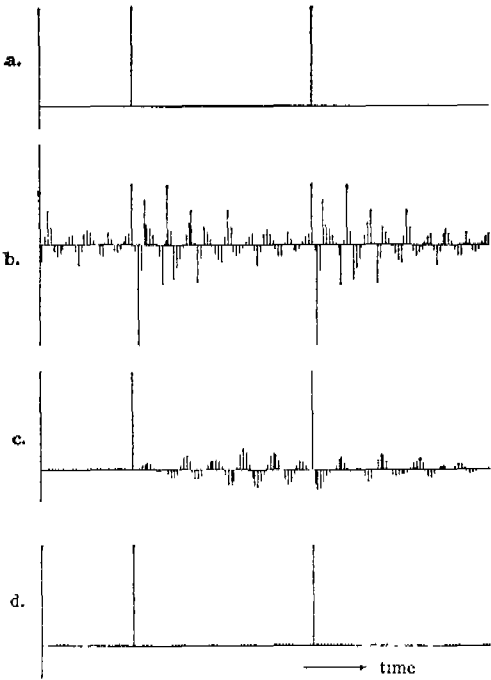


그림 2. a. 입력신호, b. 합성음성, c. 추정입력 : [4]에 의함, d. 추정입력 : 임펄스추정 알고리즘에 의한.

2. 임펄스 입력 추정 알고리즘

본연구에서는 입력추정의 한 보조수단으로서 다음과 같은 단순한 적용기법을 제안한다.

1) 파라미터 추정은 식 3)과 4)에 따른다.

2) k 번째 교정입력 v_k 의 후보로서 식(2)의 v_k' 다중의 두 값을 선택하고 그 중 v_{k+1}' 값은 작게 하는 쪽을 v_k 로 정한다.

3) \hat{u}_k 는 식(5)를 따르되 2)항의 결과로 얻어진 Φ_k 를 이용한다.

만일 $v_k=0$ 이 선택되는 경우라면 3)항의 내용은 $\hat{u}_k=v_k'$ 의 해당 표본에 관한 파라미터 수정은 보류된다. 이 성질은 입력이 고립된 임펄스인 경우에는 추정에 큰 도움을 주게 된다. 그림 1에서 입력이 고립 임펄스일 때는 해당 표본에서 추정오차는 순전히 입력 값일 가능성이 크다. 만일 단순히 [4]의 방법을 이용하면 이 추정오차를 이용하여 파라미터를 수정하려는 노력을 하게 되므로 추정에 큰 교란을 가져오고 그 영향이 그다음의 여러 표본에 걸쳐 나타나게 된다. 임펄스입력 추정 수단의 성격은 다

음에 논의한다.

1) 이상적 조건, 즉, ① 추정차수가 대상과 정합되고, ② 입력신호가 고립 임펄스이며, ③ 잡음교란이 작은 경우에는 위의 1)~3)에 의한 임펄스입력 추정 알고리즘은 매우 우수한 추정능력을 보이는 것은 그림 2로 알 수 있다.

2) 추정 차수가 어긋나는 경우는 파라미터의 수렴이 매우 늦어지며 경우에 따라서는 식(5)의 수단보다 열등한 결과를 얻었다. 이것은 $v_k=0$ 가 되는 표본마다 추정이 보류되는 까닭에 수렴이 더욱 늦어지게 되는 때문이다.

3) 입력의 형태가 임펄스가 아니거나 가깝게 배치된 연속 임펄스인 배에도 파라미터 수렴에 혼란이 발생한다. 이 경우에는 식(5)의 수단도 마찬가지이며(그림 3) 시음 ARMA 모형의 큰 결함의 하나이다.

이상의 검토에서 보는 바에 의하면 임펄스추정 알고리즘을 파라미터추정의 기본수단으로 사용하는 것은 무리가 있으며 조건이 불리한 때에는 식(5)의 수단에도 미치지 못할 수가 있으므로 이 두 모형은 병행 운전하여 유리한 측의 파라미터를 서로 교환하는 방법을 사용하는 것이 단독 운전보다 좋은 결과를 얻을 수 있을 것이다.

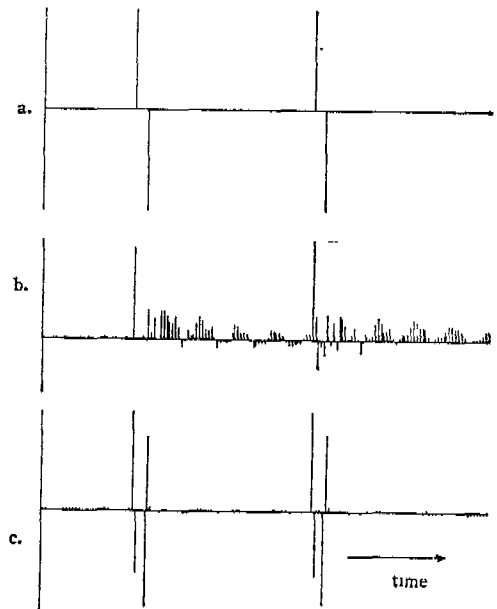


그림 3. 입력추정 실패의 예

a. 입력신호, b. 추정입력 : [4]에 의함.

c. 추정입력 : 임펄스추정 알고리즘에 의한.

Ⅲ. 신호의 진폭과 수렴판정

1. 신호진폭의 영향

ARMA 모형을 음성해석에 적용하는데 있어서 교정행렬 P_k 의 초기치가 수렴에 미치는 영향에 대해서는 알려져 있으나 신호 자체의 크기에 대해서는 고려되어 있지 않다. 실제로 신호 s_k 가 어느정도 크기 이상이면 큰 영향은 나타나지 않는것 같다. 그러나 신호의 평균 진폭이 작아질수록 수렴속도가 느려지는 경향이 관측된다. 파라미터 교정은 식(3)의 우변 제 2항에서 이루어지는 것이므로 $1 < \langle S_k^T P_k S_k \rangle$ 가 될 정도로 작은 크기의 입력이라던 교정값은 P_k, R_k, v_k 가 되는데 v_k 역시 신호크기의 범위에 있으므로 \langle 교정량 \rangle 의 관계가 성립한다. 따라서 신호가 작은 경우에는 적절한 수치로 스케일링할 필요가 있다. 최대 진폭이 10부근 이면 안정한 것으로 관측되었다.(그림 4)

2. 수렴판정의 척도

추정과정에서 한 분석구간을 종료시키고 파라미터를 추출한 후 새로운 분석구간으로 옮겨가게 되는 기준은 수렴판정에 의하여 결정하게 되는데 [4]의 경우 이 척도로서 $\|\Phi_k - \Phi_{k-1}\|$ 을 사용하고 있다. 일견 타당한 접근으로 보여지지만 오차의 여지가 있다. 그림 (6)의 예에서 추정오차는 매우 큰 값을 가

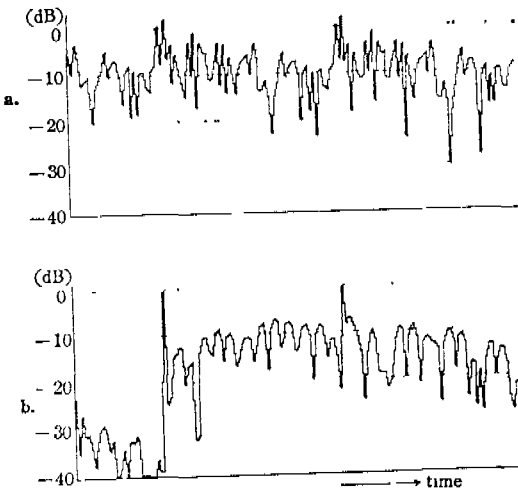


그림 4. 신호진폭이 입력추정에 미치는 영향

(입력신호는 그림 2, a와 같음) a. 최대진폭 0.01의 경우, b. 최대진폭 10의 경우.

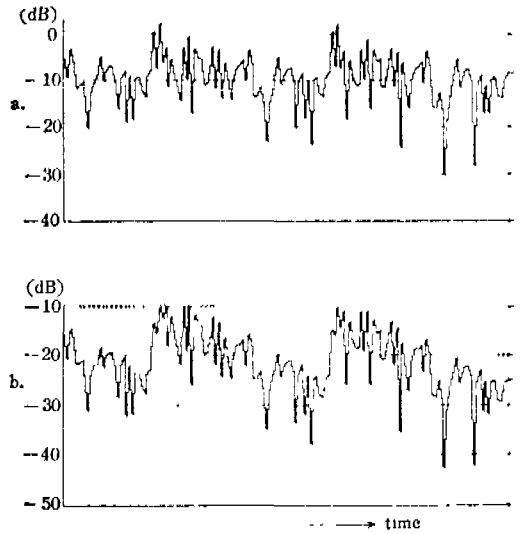


그림 5. 수렴판정 실패의 예(입력신호는 그림 2, a)

a. 추정오차 b. $\|\Phi_k - \Phi_{k-1}\|$

지고 있어 추정이 식철히 이루어지지 못했음을 보이지만 $\|\Phi_k - \Phi_{k-1}\|$ 값은 한계를 정하기에 따라서는 충분히 수렴된 것으로 보여진다. 이와 같은 착오는 식(3)의 교정행렬 P_k 의 원소값이 k 의 증가에 따라 작아져 가며 여기에 대상신호 s_k 값도 작으면 그 효과가 증폭되어 식(3)의 우변 제 2항이 매우 작아진다. 따라서 Φ_k 가 어느 값에 수렴하게 되기는 하지만 이것은 부정확한 값이 되는 것이다.

이 오류를 피하기 위해서는 추정오차 자체를 이용하는 것이 유리하며 $|v_k'/\sigma_{s,k}|$ 를 척도로 사용하는 것을 제안한다. 여기서 $\sigma_{s,k}$ 는 대상신호의 k 번째 표본까지의 표준 편차이다.

Ⅳ. 결 론

음성분석의 한 수단인 ARMA 모형을 적용하는데 있어서의 곤란점, 즉 입력추정과 수렴에 관한 문제를 토의하였다.

입력 추정에 있어서는 임펄스 추정 알고리즘을 제안하였고 진폭이 수렴에 끼치는 영향에 대해서 검토하였으며 수렴판정의 척도에 대한 문제점을 논의하고 가능한 개선책을 제안하였다.

백잡음 형태의 자음 여기신호, 짧은 간격으로 반복되는 모음의 피치임력 그리고 연속적으로 변화를 보이는 파라미터 등은 넓은 폭의 개인차와 함께 음

성의 본질적 성격으로 파악되는데 이 모든 것들이 ARMA 모형에 의한 음성신호분석을 간단하게 하고 때로는 불가능하게 까지 하는 요인이 된다. 그러나 비교적 단순한 수법으로 연속적인 추정이 가능한 ARMA 모형은 좋은 조건하에서는 우수한 자료를 제공할 수 있는 까닭에 선형예측 방법과 병행운영 하는 상보적 체계를 구성하면 우수한 계통운 언을 것이다.

참 고 문 헌

1. J. Makhoul, Proc. IEEE, 63, 561(1975)
2. J.D. Markel, IEEE Trans. on Audio and Electro Acoustics, Au-21, 149(1973)
3. F. J. Harris, Proc. IEEE, 66, 51(1978)
4. H. Morikawa and H. Fujisaki, IEEE Trans. on Acoustics, Speech, and Signal Processing, ASSP-30, 77(1982)
5. Y. Miganaga, et al, *ibid*, 88(1982)
6. R. Isermann, Digital Control Systems, Springer-Verlag, Berlin(1981) pp. 364-386.