

## 적응예측기를 이용하여 잡음 섞인 음성신호로부터 autoregressive 계수를 추산하는 방법

### An Autoregressive Parameter Estimation from Noisy Speech Using the Adaptive Predictor

구 본 응\*  
(Boneung Koo\*)

#### 요 약

잡음 섞인 관측데이터로부터 AR 모수를 추정하는 방법을 제안하였다. AP방법이라고 이름붙인 이 방법은 단순하고도 신뢰성있는 적응예측기를 이용하려는 시도의 산물이다. 잡음 섞인 입력수열로부터 계산된 AR 모수의 추정치보다 예측수열로부터 계산된 AR 모수의 추정치가 원래의 모수에 스펙트럼상의 거리가 더 가깝다는 것을 이론적으로 증명하였다. 실제 음성 신호와 칼만필터를 사용한 실험결과도 이론과 일치함을 보였다. 대략적으로, AP방법으로 계산된 추정치를 사용하였을때의 잡음감쇠 성능은 잡음 섞인 입력수열로부터 계산된 AR 모수의 추정치를 사용하였을때보다는 우수하였고, EM반복법에 의한 추정치를 사용하였을때보다는 약간 못한 것으로 나타났다. 그러나, 제안된 방법은 그 단순성으로 인하여 경우에 따라 더 복잡한 다른 방법의 대안으로 사용될 수 있을 것이다.

#### Abstract

A new method for autoregressive parameter estimation from noisy observation sequence is presented. This method, termed the AP method, is a result of an attempt to make use of the adaptive predictor which is a simple and reliable way of parameter estimation. It is shown theoretically that, for noisy input, the parameter vector computed from the prediction sequence is closer to that of the original sequence than the noisy input sequence is, under the spectral distortion criterion. Simulation results with the Kalman filter as a noise reduction filter and real speech data supported the theory. Roughly speaking, the performance of the parameter set obtained by the AP method is better than the noisy one but worse than the EM iteration results. When the simplicity is considered, it could provide a useful alternative to more complicated parameter estimation methods in some applications.

#### I. 서 론

음성신호의 선형모델로서는 계산의 편의상 보통

AR (autoregressive) 모델을 사용한다. 부호화, 잡음 제거, 인식, 피치검출 등의 분야에서 AR 계수는 매우 중요한 역할을 한다. 잡음이 없을 때에는 Yule-Walker equation을 풀어서 AR 계수를 구하면 되지만, 잡음이 섞인 경우에도 같은 방법을 사용하면 그 해는 bias가

\*경기대학교 전자공학과  
접수일자: 1995년 4월 24일

큰 추산치가 되어, 가령 이 解를 잡음제거 필터에 직접 사용했을 경우 왜곡이 심한 출력을 얻게 된다. 本稿에서는 잡음 섞인 WSS(wide-sense stationary) AR sequence로부터 잡음의 영향을 덜 받는 AR 계수의 추정 방법을 제시하였다. 이 방법은 ADPCM(Adaptive Differential Pulse Code Modulation) 등에서 사용되는 적용예측기(Adaptive Predictor)를 이용하는 방법으로서, 구조가 간단하고 계산속도가 빠르다는 장점을 갖는다.

잡음 섞인 음성신호 또는 AR sequence로부터 source model 계수의 estimates를 구하는 문제는 오래전부터 주요 연구대상이 되어왔다[1]-[6]. 잡음이 白色(white)인 경우에 대해서는 여러 연구결과가 있으나 有色(colored)인 경우에 대해서는 별로 발표된 것이 없다. 백색잡음의 경우, 최적해를 구하려면 batch mode에서만 계산가능한 비선형 최적화기법을 사용해야하는 한편[1], [5], 선형화된 준최적해는 음성신호와 같은 대역신호에 적용하기에는 미흡한 경우가 많다[3], [4], [6]. 本稿에서는 적용예측기를 이용하여 백색 및 유색잡음에 적용할 수 있는 새로운 알고리즘을 제안하였다.

이 방법은 음성신호에 부가된 잡음을 감소시키는 칼만필터[7]에 적용할 목적으로 고안되었다. 최근에, 같은 용도로 사용할 수 있는 계수추정 알고리즘으로서 EM 반복법이 발표된 바 있다[8]. EM 반복법으로 추정된 계수는 필터계수로서의 성능은 최적해에 근접하게 우수하지만 2회 내지 3회의 반복계산을 필요로 하는데 반하여, 본 연구에서 고안된 방법은 성능은 다소 미흡하지만 반복법이 아니라는 장점을 갖는다.

II절에서는 알고리즘의 내용을 소개하고, III절에서는 백색 및 유색잡음에 대하여 각각 스펙트럼상의 거리를 계산하여 이 방법으로 구한 AR 계수가 noisy sequence로부터 직접 구한 것보다 잡음의 영향을 덜 받는다는 것을 보였다. IV절에는 실험결과를, V절에는 검토 및 결론을 제시하였다.

## II. 적용예측기의 응용

일반적으로 차수가  $p$ 인  $AR(p)$  sequence  $\{x(n)\}$  과, 잡음이 섞인 sequence  $\{s(n)\}$ 은 각각 다음의 식(1)과 (2)로 표시된다.

$$x(n) = \sum_{i=1}^p a_i x(n-i) + w(n) \quad (1)$$

$$s(n) = x(n) + v(n) \quad (2)$$

여기서,  $w(n)$ 은 mean이 0이고 variance가  $\sigma_w^2$ 인 white Gaussian sequence이고,  $v(n)$ 은 백색 또는 유색 부가잡음이다. 본고에서 다루고자하는 문제는, 길이가  $L$ 인 관측 데이터  $\mathbf{s} = \{s(1), \dots, s(L)\}$ 가 주어졌을 때  $\mathbf{a} = [a_1, \dots, a_p]^T$ 와  $\sigma_w^2$ 의 추산치를 구하는 것이다. 잡음  $v(n)$ 은  $x(n)$ 과는 무관(uncorrelated)하다고 가정한다.

식(1)을 다시 쓰면,

$$x(n) = x_p(n) + w(n) \quad (3)$$

$$x_p(n) = \sum_{i=1}^p a_i x(n-i) \quad (4)$$

가 된다.  $\{x(n)\}$ 을 알면 자기상관계수를 구한 다음 Yule-Walker 식으로부터 선형 예측계수  $\mathbf{a}$ 와  $\sigma_w^2$ 을 구할 수 있다. 적용예측기는 새로운 입력 샘플  $x(n)$ 이 들어올 때마다 예측계수를 새로 계산하여  $x_p(n)$ 을 출력하는 기능을 가지므로, 예측계수 추산 기능도 갖고 있고, 본 연구에서는 바로 이 점에 착안하여  $x_p(n)$ 을 계산하는 과정에서 구해지는 예측계수를 이용하고자 하였다.

적용예측기는 음성코딩 분야에서 널리 쓰이고 있고, 그 종류도 다양한데[7], 본 논문에서 사용한 구조는 그림1과 같다.  $\{x(n)\}$ 의 모수(parameters)는  $(\mathbf{a}, g)$ 로 표시하였는데, 여기서  $g \equiv \sigma_w$ 이다. 예측신호  $\hat{x}_p(n)$ 는

$$\hat{x}_p(n) = \sum_{i=1}^p \hat{a}_i(n-i) x(n-i) \quad (5)$$

이고,  $\hat{a}_i(n)$ 은 시각지수  $n$ 에서 계산된  $a_i$ 의 추산치이다. 매 입력  $x(n)$ 이 들어 올 때마다 오차신호  $e(n) \equiv x(n) - \hat{x}_p(n)$ 을 계산하고 적용알고리즘은 새로운 예측계수  $\hat{a}_i(n+1)$ 을 구한다.

적용알고리즘은 매 시각에서의 오차 에너지가 감소하도록 계수값을 갱신한다. 이러한 알고리즘에는 LMS 알고리즘, adaptive gradient 알고리즘, 적응칼만 알고리즘 등이 있다[9]. 잘 조정된 알고리즘을 사용하면, 기준 입력이 잡음없는  $\{x(n)\}$ 일 때 시간이 경과함에 따라  $\{\hat{a}_i(n)\}$ 은  $\{a_i\}$ 로 수렴하고, 따라서  $\hat{x}_p(n)$ 은  $x_p(n)$ 으로 수렴하게 된다.

본 연구에서는 기준입력이 식(2)로 주어진 잡음이 부가된 신호이다. 그림2에 각 수열과 계수를 명시하

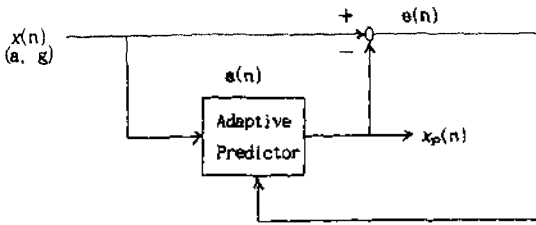


그림 1. 잡음없는 입력에 대한 적응예측기  
Fig. 1. Adaptive Predictor with Noise-free Input

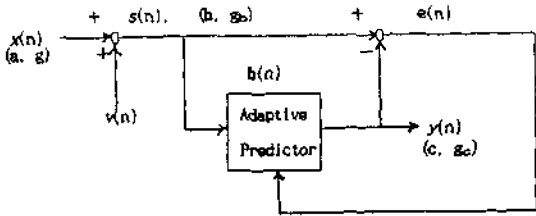


그림 2. 잡음섞인 입력에 대한 적응예측기  
Fig. 2. Adaptive Predictor with Noisy Input

였다. noise-free sequence  $x = \{x(n)\}$ 의 AR( $p$ ) 모델의 모수는  $(a, g)$ 로 표시하였는데,  $a$ 는 예측계수 벡타이고,  $g$ 는 예측오차에너지의 평방근, 즉,  $\sigma_w$ 이다.  $(b, g_b)$ 는 noisy sequence  $s = \{s(n)\}$ 의 모수이고,  $b(n)$ 은  $b$ 를 추적하는 계수벡타이다. 수열  $y = \{y(n)\}$ 은 예측기의 출력으로서

$$y(n) = \sum_{i=1}^p \hat{b}_i(n) s(n-i) \quad (6)$$

이고,  $(\hat{b}, g_b)$ 는  $y$ 의 AR( $p$ )의 모델계수들이다.

그림2에서와 같이 기준입력이 잡음이 부가된 신호  $\{s(n)\}$ 인 경우, 당연히 적응예측기는 원래의 신호  $\{x(n)\}$ 이 아닌  $\{s(n)\}$ 을 추적할 것이고, 따라서, 적응 알고리즘이 추산하는 예측계수는  $\{s(n)\}$ 의 것이 되는데, 이것은 본래의 목적과는 명백히 다른 것이다. 따라서, 본 연구에서는 적응예측기가 생성하는 또하나의 수열  $\{y(n)\}$ 에 주목하게 되었다.

다음 절에서, 어떤 조건이 만족되면,  $\{y(n)\}$ 으로부터 계산된 AR( $p$ ) 예측계수가  $\{s(n)\}$ 으로부터 계산된 것보다 잡음의 영향을 덜 받는다는 것을 입증하였다.

### III. 스펙트럼상의 거리 비교

앞에서 언급한 두 종류의 예측계수들과 원래의 신

호  $\{x(n)\}$ 의 예측계수와의 스펙트럼상의 거리를 비교하기 위하여 LPC 등에서 많이 사용되는 rms log 스펙트럼 거리 측정기준[10]을 사용하였다. 두개의 수열 벡타  $x$ 와  $y$ 의 스펙트럼을 각각  $X(e^{j\omega})$  및  $Y(e^{j\omega})$ 라고 하면, 두 수열사이의 rms log 스펙트럼 거리의 제곱  $d^2$ 는

$$d^2 = \int_{-\pi}^{\pi} \{\log |X(e^{j\omega})|^2 - \log |Y(e^{j\omega})|^2\}^2 d\omega \quad (7)$$

로 정의되고, 이것은  $L_2$  norm의 일종으로서 수학적인 metric의 모든 성질들, 예컨대, 비음(nonnegative)이고, 대칭성 및 삼각부등식을 만족시킨다는 특성을 갖고 있다. 식(7)로 정의된  $d$ 는 Itakura-Saito distortion measure 및 이것의 변형들을 포함하여 다른 distortion measure들을 비교하는 기준으로 사용된다[11].

다음에, 백색잡음과 유색잡음의 각 경우에 대하여 어떤 조건이 만족되면 그림2에 표시된 수열  $y$ 가  $s$ 보다  $x$ 에 더 가깝다는 것, 즉,

$$d(x, y) < d(x, s) \quad (8)$$

이 됨을 보였다.

#### 백색잡음의 경우

부가잡음  $v(n)$ 은 평균값이 0이고 분산이  $\sigma_v^2$ 인 백색잡음으로서 잡음이 더해지기 전의 신호  $x(n)$ 과는 무관(uncorrelated)하다고 가정한다. 식(1)의 AR( $p$ ) 신호  $x(n)$ 은 백색잡음  $w(n)$ 을 입력으로 하는 선형시스템의 출력이므로,  $x(n)$ 의 PSD(Power Spectral Density)는

$$|X(e^{j\omega})|^2 = \frac{g^2}{|A(e^{j\omega})|^2} \quad (9)$$

인데, 여기서,  $g = \sigma_w$ 이고,

$$A(e^{j\omega}) = 1 - \sum_{i=1}^p a_i e^{-j\omega i} \quad (10)$$

이다. 또,  $x = \{x(n)\}$ 와  $v = \{v(n)\}$ 는 무관하므로,  $s = x + v$ 의 PSD는

$$|S(e^{j\omega})|^2 = |X(e^{j\omega})|^2 + \sigma_v^2 \quad (11)$$

이다.

적용예측기가 입력  $\mathbf{s}$ 를 완벽하게 추적하여  $\hat{\mathbf{b}} = \mathbf{b}$ 가 된다면, 그림2로부터,

$$v(n) = \sum_{i=1}^p \hat{b}_i s(n-i) \quad (12)$$

가 되고, 따라서,

$$|Y(e^{j\omega})|^2 = |B(e^{j\omega})|^2 \cdot |S(e^{j\omega})|^2 \quad (13)$$

$$B(e^{j\omega}) = \sum_{i=1}^p \hat{b}_i e^{-j\omega i} \quad (14)$$

인데, 식(7)의 정의에 의하여,

$$d^2(\mathbf{x}, \mathbf{s}) = \int_{-\pi}^{\pi} \{\log |X(e^{j\omega})|^2 - \log |S(e^{j\omega})|^2\}^2 d\omega$$

이고,

$$d^2(\mathbf{x}, \mathbf{y}) = \int_{-\pi}^{\pi} \{\log |X(e^{j\omega})|^2 - \log |Y(e^{j\omega})|^2\}^2 d\omega$$

이므로,

$$J(\omega) = \{\log |X(e^{j\omega})|^2 - \log |S(e^{j\omega})|^2\}^2 - \{\log |X(e^{j\omega})|^2 - \log |Y(e^{j\omega})|^2\}^2 \quad (15)$$

라 하면,

$$d^2(\mathbf{x}, \mathbf{y}) - d^2(\mathbf{x}, \mathbf{s}) = \int_{-\pi}^{\pi} J(\omega) d\omega \quad (16)$$

와 같이 표시할 수 있다. 식(9), (13), (14)를 (15)에 대입하여 정리하면,

$$J(\omega) = \left\{ \log \frac{g^2}{g^2 + \sigma_v^2 |A(e^{j\omega})|^2 \cdot |B(e^{j\omega})|^2} \right\}^2 - \left\{ \log \frac{g^2}{g^2 + \sigma_v^2 |A(e^{j\omega})|^2} \right\}^2$$

이 되는데, 여기서,  $\sigma_v^2 |A(e^{j\omega})|^2 > 0$  이므로, 뒷식 우변의 두번째 항은 0보다 작다. 따라서,

$$|B(e^{j\omega})|^2 < 1, \forall \omega \in [-\pi, \pi] \quad (17)$$

이면, 식(16)에서,  $J(\omega) < 0$ 이 만족되므로, 식(8)이 성립한다.

즉, 식(17)의 조건이 만족되면, 백색잡음이 섞인 신호  $\mathbf{s}$ 보다 예측신호  $\mathbf{y}$ 가 원래의 신호  $\mathbf{x}$ 에 스펙트럼

저리상으로 더 가깝다는 것을 알 수 있다.

### 유색잡음의 경우

이 경우에는 부가잡음  $v(n)$ 이 유색이고, 다음 식과 같이 AR( $q$ ) process로 모델링할 수 있다고 가정한다.

$$v(n) = \sum_{i=1}^q h_i v(n-i) + \xi(n) \quad (18)$$

여기서,  $\xi(n)$ 은 mean과 variance가 각각 0,  $\sigma_\xi^2$ 인 백색 Gaussian 잡음이다.

$$|V(e^{j\omega})|^2 = \frac{\sigma_\xi^2}{|H(e^{j\omega})|^2} \quad (19)$$

이고, 여기서,

$$H(e^{j\omega}) = 1 - \sum_{i=1}^q h_i e^{-j\omega i} \quad (20)$$

이다. 또,  $\mathbf{x} = \{x(n)\}$ 와  $\mathbf{v} = \{v(n)\}$ 는 무관하므로,  $\mathbf{s} = \mathbf{x} + \mathbf{v}$ 의 PSD는

$$|S(e^{j\omega})|^2 = |X(e^{j\omega})|^2 + |V(e^{j\omega})|^2 \quad (21)$$

이다.  $|X(e^{j\omega})|^2$ 는 식(9)와 (10)으로 주어졌다. 적용예측기가 입력  $\mathbf{s}$ 를 완벽하게 추적하여  $\hat{\mathbf{b}} = \mathbf{b}$ 가 된다면, 그림2의  $y(n)$ 은 식(12),  $|Y(e^{j\omega})|^2$ 은 식(13),  $B(e^{j\omega})$ 는 식(14)로 각각 표시된다.

이제, 백색잡음의 경우에서와 같은 방법으로, 식(15)와 (16)을 사용하여  $d^2(\mathbf{x}, \mathbf{y})$ 와  $d^2(\mathbf{x}, \mathbf{s})$ 의 크기를 비교해 보자. 식 (13)을 (15)에 대입하면,

$$J(\omega) = \{\log |X(e^{j\omega})|^2 - \log |S(e^{j\omega})|^2 - \log |B(e^{j\omega})|^2\}^2 - \{\log |X(e^{j\omega})|^2 - \log |S(e^{j\omega})|^2\}^2$$

이 되는데, 식(21)에 의하여,

$$\log |X(e^{j\omega})|^2 - \log |S(e^{j\omega})|^2 < 0, \forall \omega \in [-\pi, \pi]$$

이므로,

$$2(\log |X(e^{j\omega})|^2 - \log |S(e^{j\omega})|^2) < \log |B(e^{j\omega})|^2 < 0,$$

또는,

$$\left\{ \frac{|X(e^{j\omega})|^2}{|S(e^{j\omega})|^2} \right\}^2 < |B(e^{j\omega})|^2 < 1, \forall \omega \in [-\pi, \pi] \quad (22)$$

이며, 식 (16)에서,  $f(\omega) < 0$ 이 만족되므로, 식(8)이 성립한다.

즉, 식(22)의 조건이 만족되면, 유색잡음이 섞인 신호  $\mathbf{y}$ 보다 예측신호  $\mathbf{y}$ 가 원래의 신호  $\mathbf{x}$ 에 스펙트럼 거리상으로 더 가깝다는 것을 알 수 있다.

백색잡음에 대한 조건 식(17)과 유색잡음에 대한 조건 식(22)를 비교해보면, 식(22)에는 식(17)의 조건이 포함되어 있고, 왼쪽에 조건이 하나 더 추가되어 있음을 알 수 있다. 이 조건은 다양한 모든 종류의 AR신호와 유색잡음에 대하여 일반적으로 증명할 수 없으므로 항상 성립한다고는 할 수 없다. 백색잡음에 대한 조건 식(17)도 마찬가지로 일반적으로 성립한다고는 할 수 없다. 다만, 입력되는 신호와 잡음의 표본에 따라 성립할 수도 있고, 안할 수도 있다.

#### IV. 실험 결과

본 논문에서 제안한 방법(이하, AP 방법)의 성능을 실험하기 위하여 실제 음성신호에 백색 및 유색잡음을 더한 것을 입력신호로 하고 칼만 필터의 출력과 원래 잡음이 없을때의 음성신호의 SNR을 계산하였다. 음성신호는 남녀 화자가 발음한 각기 다른 2초 가량의 문장을 8kHz, 12bits로 A/D 변환한 것이다. 백색잡음은 컴퓨터로 생성한 Gaussian 잡음을 사용하였고, 유색잡음은 헬기의 조종석에서 녹음한 것으로서 2초 동안은 거의 정제적(stationary)인 것을 사용하였다. 잡음을 제거하기 위하여 사용한 칼만필터[7]에는 음성신호의 AR계수와 잡음의 계수가 사용되는데, 음성신호의 AR계수는 본 논문에서 제안된 AP 방법으로 계산된 것을 사용하였고, 잡음의 계수는 잡음신호로부터 직접 계산하여 사용하였다.

적용예측알고리즘에는 여러가지가 있는데, 그중에서 LMS(Least Mean-Square) 알고리즘, Gradient 알고리즘, 칼만 알고리즘 등이 많이 쓰이고 있는데[9], 본 연구에서는 추적능력이 가장 우수한 칼만 알고리즘[12]을 사용하였다.

백색잡음의 경우, 실험에 사용된 신호들에 대하여 식(17)의 조건을 검증해 본 결과, 두가지 음성신호에 대한  $|B(e^{j\omega})|^2$ 의 최대값은 각각 0.891, 0.631로서 모두 1보다 작아서 식(8)의 부등식이 성립함을 확인하

였다. 즉, 예측신호  $\mathbf{y}$ 로부터 계산된 AR계수가 잡음 섞인 신호  $\mathbf{y}$ 의 것보다 원래의 신호  $\mathbf{x}$ 의 것에 더 가깝고, 따라서, 이 계수를 잡음감쇄필터에 사용했을 때 더 큰 SNR값을 얻어야 할 것이다.

시스템입력의 SNR이 5dB인 경우, 백색 및 유색잡음에 대한 실험 결과를 각각 표1 및 표2에 보였다. Scalar형 및 벡터형 칼만필터[7]를 모두 적용하여 필터출력의 SNR을 계산하였다. 비교를 위하여 EM 알고리즘의 실험결과[8]도 함께 수록하였다. 표1과 표2에서 "Noisy"항은 5dB의 잡음이 섞인 입력신호로부터 계산한 AR(p) 계수를 칼만필터계수로 사용한 경우이고, "AP"항은 본 논문에서 제안한 적용예측기를 사용한 경우이고, "EM-1"항은 "Noisy" 경우의 KF 출력으로부터 계산한 AR(p) 계수를 다시 KF 계수로 사용한 경우이다. "Ideal"항은 잡음이 없을 때의 음성신호로부터 계산한 AR(p) 계수를 KF 계수로 사용한 경우, 즉 이상적인 경우의 KF 출력의 SNR이다.

표 1. 칼만필터 출력의 SNR(백색잡음의 경우)

Table 1. SNR at the Kalman Filter Output(for white noise)

KF Type	문 장	A		B	
		4	10	4	10
Scalar	p				
	Noisy	8.073	8.188	7.798	7.928
	AP	9.157	9.475	8.591	8.843
	EM-1	9.237	9.400	8.701	8.922
	Ideal	9.925	10.094	9.171	9.509
Vector	p				
	Noisy	8.576	8.770	8.187	8.397
	AP	9.495	9.627	8.821	8.795
	EM-1	10.252	10.655	9.423	9.849
	Ideal	11.039	11.750	9.956	10.725

표 2. 칼만필터 출력의 SNR(유색잡음의 경우)

Table 2. SNR at the Kalman Filter Output(for colored noise)

KF Type	문 장	A		B	
		(4, 10)	(10, 10)	(4, 10)	(10, 10)
Scalar	(p, q)				
	Noisy	8.461	8.240	8.032	8.079
	AP	9.119	8.923	8.548	8.666
	EM-1	9.541	9.409	8.833	9.008
	Ideal	9.983	10.153	9.227	9.562
Vector	(p, q)				
	Noisy	8.078	8.490	7.698	8.272
	AP	8.543	9.290	8.084	8.950
	EM-1	8.986	9.877	8.409	9.321
	Ideal	9.535	10.805	8.923	9.974

백색잡음의 경우, 필터의 차수는 AR모델차수  $p$ 와 일치한다. 표1에서, Scalar KF(Kalman Filter)를 사용한 경우, 필터의 차수  $p$ 값에 따라 약간의 편차는 있지만, AP에 의한 SNR은 Noisy한 경우의 SNR보다 대략 1dB정도 증가했고, EM-1에 의한 SNR에 0.1dB 이내로 접근하고 있다. Vector KF를 사용한 경우에는 AP에 의한 SNR은 Noisy한 경우의 SNR과 EM-1에 의한 SNR의 중간쯤 되는 값을 갖는 것으로 나타났다. 어느 경우에도 AP에 의한 SNR은 Noisy한 경우의 SNR보다 필터의 종류와 입력신호에 따라 0.5dB 내지 1.3dB 증가했음을 알 수 있다.

유색잡음의 경우, 필터의 차수는 음성신호의 AR모델차수  $p$ 와 잡음의 AR모델차수  $q$ 의 합과 같다. 표2를 보면, AP에 의한 SNR은 Noisy한 경우의 SNR보다 필터의 종류와 입력신호에 따라 대략 0.5dB 내지 0.8dB 증가했음을 알 수 있다.

이러한 실험결과를 앞의 III절에서 유도했던 이론적인 결과, 즉, 잡음쉬인 신호보다는 적용예측기의 예측신호로부터 구한 모델계수가 원래의 것에 더 가깝다는 추론과 부합되는 것이다.

## V. 검토 및 결론

본 논문에서는 잡음쉬인 음성신호로부터 원래 음성신호의 AR 계수를 추출하는 새로운 방법을 제시하였다. 이 방법은 성능이 입증된 간단한 적용예측기 계수추정에 이용하려는 시도이다. 입력에 잡음이 없는 경우에는 예측알고리즘이 추적하는 계수가 원래의 계수에 수렴하므로 아무 문제가 없지만, 잡음이 섞인 입력의 경우에는 예측알고리즘이 추적하는 계수보다 예측신호로부터 구한 계수가 원래 신호에 스펙트럼상에서 더 가깝다는 것을 이론적으로 유도하였다. 이 계수를 잡음감쇠용 칼만필터에 적용하여 실험결과가 이론과 일치함을 보였다.

이 방법으로 구한 계수의 잡음감쇠성능은 잡음신호에서 직접 구한 계수의 성능보다는 좋고, EM반복법보다는 떨어진다고 대략적으로 말할 수 있다. 그러나, 종래의 비선형이나 EM반복법보다는 계산시간이 짧으므로 경우에 따라 실시간 음성신호의 잡음제거 필터로 사용될 수 있을 것이다.

## 참고 문헌

1. M. Pagano, "Estimation of models of autoregressive signals plus white noise," *Ann. Stat.*, vol. 2, pp. 99-108, 1974.
2. J. S. Lim and A. V. Oppenheim, "All-pole modeling of degraded speech," *IEEE Trans. Acoust., Speech and Signal Proc.*, vol. ASSP-26, pp. 197-210, June 1978.
3. W. J. Done and Rushforth, "Estimating the parameters of noisy all-pole process," *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc.*, pp. 228-231, April 1979.
4. K. Bry and J. LeRoux, "Comparison of some algorithms for identifying autoregressive signals in the presence of noise," *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc.*, pp. 224-227, April 1982.
5. B. Friedlander and B. Porat, "A spectral matching technique for ARMA parameter estimation," *IEEE Trans. Acoust., Speech and Signal Proc.*, vol. ASSP-32, pp. 338-343, April 1984.
6. K. K. Paliwal, "A noise-compensated long correlation matching method for AR spectral estimation of noisy signals," *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc.*, pp. 1369-1372, April 1986.
7. J. D. Gibson, B. Koo, and S. D. Gray, "Filtering of colored noise for speech enhancement and coding," *IEEE Trans. Signal Proc.*, vol. 39, no. 8, pp. 1732-1742, Aug. 1991.
8. 구본용, "EM 알고리즘을 이용한 Noisy All-pole Process의 Autoregressive 계수추정 및 잡음제거 방법", 경기대학교 산업기술종합연구소 논문집, 제8집, pp. 19-26, 1992. 2.
9. J. D. Gibson, "Adaptive prediction in ADPCM speech coding," *IEEE Proceedings*, vol. 68, pp. 488-525, Apr. 1980.
10. L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Englewood Cliffs, N. J. : Prentice-Hall, 1978.
11. A. H. Gray, Jr. and J. D. Markel, "Distance measures for speech processing," *IEEE Trans. Acoust., Speech and Signal Proc.*, vol. ASSP-24, pp. 380-391, Oct. 1976.
12. J. D. Gibson, "Sequentially adaptive backward prediction in ADPCM speech coding," *IEEE Trans. Commun.*, vol. COM-26, pp. 145-150, Jan. 1978.

▲具 本 應(Boneung Koo) 1953년 8월 2일생



1975년 2월 : 서울대학교 공과대  
학 공업교육학과  
전자전공(학사)

1984년 12월 : Texas A&M Univ.  
전기공학과(석사)

1988년 12월 : Texas A&M Univ.  
전기공학과(박사)

1977년 1월~1982년 7월 : 한국원  
자력연구소(연구원)

1989년 3월~현재 : 경기대학교 전자공학과(부교수)

※주관심분야 : Speech Coding, Speech Enhancement