하계학술대회
논문86-5-5

고속이득 연산식의 수리특성
Numerical Properties of Fast Gain Algorithm

양 홍 석 서울대학교
정 찬 수* 숭전대학교
이 상 철 대유공업전문대학
이 석 원 "

I. 서 론

시스템 식별 방법은 그 응용분야가 넓은 만큼 다양하게 연구·발전되어 왔으며 각각 서로 다른 장점들을 갖고 있다.[1] 최근에는 대용량 계산기와 값싸고 다양한 집적회로의 보급으로 실시간 신호처리가 실용화되기 시작하면서, 보다 수렴속도가 빠르고 계산량이 적은 시스템 식별 연산식이 요구되는 추세이다.

P개의 미지수(식별해야 할 시스템계수)를 가진 선형방정식을 풀기 위해서는 일반적으로 $O(p^3)$의 곱셈이 필요하며, 시스템 방정식에 나오는 행렬의 특수한 형태를 이용하여 Levinson 과 Durbin은 $O(p^2)$의 곱셈만 필요한 고속 연산식을 개발하였다. (Levinson-Durbin 연산식)[2]

그후 1974년 Morf 에 의해 "이동 특성 (shifting property)" 이 발표되고 이를 이용한 고속 격자형 필터의 연구가 활발히 진행되었으며 70년대 후반에 Viera, Lee 등에 의해 계속 연구 발전되어 오고 있다.[3,4]

또 한편으로는 Ljung, Robins 등에 의해 고속으로 Kalman 이득을 계산하는 연산식이 연구되었다.[5,6] 위의 두 흐름은 근본적으로 비슷한 개념에서 출발하였으며 최소자승법에 근거하고 있으므로 수렴속도가 빠르고, 계산량은 O(p)에 불과하여 신호처리에서와 같이 P가 20~30 되는 경우에도 엄청난 계산량의 감소로 실시간 처리를 가능케 하고 있다. 그러나 이들 연산식의 안정성의 증명, 수렴속도에 관한 해석등 해결되어야 할 과제가 있으며 본 논문에서는 연산식을 소개하고 이의 수리적 안정성을 증명하고자 한다.

II. 이동 특성

설명을 간단히 하기 위하여 다음과 같은 3차 AR 모형을 생각하자.

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \eta_t \qquad (1)$$

y_t : 시간 t 에서의 측정값
η_t : 추정 오차(residuals)

시스템 상수 $\theta = (a_1, a_2, a_3)^T$를 추정하기 위한 측정값들의 공분산행렬은 $t = N-1$ 에서

$$R_{N-1} = \left[\begin{array}{cc|c} \sum_{t=0}^{N-1} y_t y_t & \sum_{t=1}^{N-1} y_t y_{t-1} & \sum_{t=2}^{N-1} y_t y_{t-2} \\ \sum_{t=1}^{N-1} y_{t-1} y_t & \sum_{t=1}^{N-1} y_{t-1} y_{t-1} & \sum_{t=2}^{N-1} y_{t-1} y_{t-2} \\ \hline \sum_{t=2}^{N-1} y_{t-2} y_t & \sum_{t=2}^{N-1} y_{t-2} y_{t-1} & \sum_{t=2}^{N-1} y_{t-2} y_{t-2} \end{array}\right] \quad (2)$$

다음의 측정값이 들어올때 즉 $t=N$ 에서는

$$R_{N} = \left[\begin{array}{c|cc} \sum_{t=0}^{N} y_t y_t & \sum_{t=1}^{N} y_t y_{t-1} & \sum_{t=2}^{N} y_t y_{t-2} \\ \hline \sum_{t=1}^{N} y_{t-1} y_t & \sum_{t=1}^{N} y_{t-1} y_{t-1} & \sum_{t=2}^{N} y_{t-1} y_{t-2} \\ \sum_{t=2}^{N} y_{t-2} y_t & \sum_{t=2}^{N} y_{t-2} y_{t-1} & \sum_{t=2}^{N} y_{t-2} y_{t-2} \end{array}\right] \quad (3)$$

위의 두 식을 비교해보면 R_{N-1}의 왼쪽 윗부분과 R_N의 오른쪽 아랫부분이 완전히 일치함을 알 수 있다.

즉 R_N에는 R_{N-1}에 있던 많은 정보가 그대로 있으며, 일 부분만 새로 추가하고 잘라버리면 되므로 이 성질을 잘 이용하면 매우 효율적으로 계산할 수 있게 된다.

Ⅲ. 고속이득 연산식(회귀 모형식별 연산식)

(1)식에 대응되는 p차의 AR모형을 생각하자. [0,t] 의 시간 영역에서의 식을 벡터로 표시하면,

$$Y_t = X_t \hat{\theta}_t + H_t \quad (4)$$

단, $Y_t = (y_t, y_{t-1}, \cdots y_1)^T$ $t \times 1$ 벡터

$X_t = (Y_{t-1}, Y_{t-2}, \cdots\cdots, Y_{t-p})$ $t \times p$ 행렬

$\hat{\theta}_t = (\hat{a}_1, \hat{a}_2, \cdots\cdots, \hat{a}_p)_t^T$ $p \times 1$ 벡터

$H_t = (\eta_t, \eta_{t-1}, \cdots\cdots, \eta_1)^T$ $t \times 1$ 벡터

$\hat{\theta}_{t-1}$ 에서 $\hat{\theta}_t$ 를 계산한 순차식은

$$\hat{\theta}_t = \hat{\theta}_{t-1} + K_t \eta_t \quad (5)$$

이며, K_t 는 Kalman 이득으로 새로운 측정값이 얻어질 때 마다 새로 계산하게 되며 다음과 같이 정의 된다.

$$[X_t^T X_t] K_t = x_t \quad (6)$$

단, $x_t = (y_{t-1}, y_{t-2}, \cdots y_{t-p})^T$

K_t 및 K_{t+1} 의 정의식을 확장하면

$$\begin{bmatrix} Y_t^T Y_t & Y_t^T X_t \\ X_t^T Y_t & R_t^- \end{bmatrix} \begin{bmatrix} c \\ K_t \end{bmatrix} = \begin{bmatrix} \hat{y}_t \\ x_t \end{bmatrix} \quad (7)$$

$$\begin{bmatrix} R_{t+1}^- & X_{t+1}^T Y_{t-p} \\ Y_{t-p}^T X_{t+1} & Y_{t-p}^T Y_{t-p} \end{bmatrix} \begin{bmatrix} K_{t+1} \\ c \end{bmatrix} = \begin{bmatrix} x_{t+1} \\ \hat{y}_{t-p} \end{bmatrix} \quad (8)$$

위의 (7), (8) 식을 관찰해 보면 좌변의 첫번째 행렬은 서로 같은 것이며 우변도 첫항과 마지막항이 조금 다를 뿐 나머지는 서로 같다. 이러한 유사성을 이용하고 새로운 변수들을 도입하여 K_t 로 부터 K_{t+1}을 얻는 순차 연산식을 구하면 다음과 같다.[7]

$$\eta_t = y_t - x_t^T \hat{\theta}_{t-1} \quad (9)$$

$$r_t = y_{t-p} - x_{t+1}^T \hat{\phi}_{t-1} \quad (10)$$

$$\hat{\theta}_t = \theta_{t-1} + K_t \eta_t \quad (11)$$

$$e_t = y_t - x_t^T \hat{\theta}_t \quad (12)$$

$$E_t = E_{t-1} + e_t \eta_t \quad (13)$$

$$\begin{bmatrix} k_t^* \\ k_t^{**} \end{bmatrix} = \begin{bmatrix} E_t^{-1} e_t \\ K_t - \hat{\theta}_t E_t^{-1} e_t \end{bmatrix} \quad (14)$$

$$K_{t+1} = (k_t^* + \hat{\phi}_{t-1} k_t^{**}) / (1 - r_t k_t^{**}) \quad (15)$$

$$\hat{\phi}_t = \hat{\phi}_{t-1} + K_{t+1} r_t \quad (16)$$

여기서 $\hat{\phi}_t$는 후향예측 계수이며, r_t는 그 예측오차이고 E_t는 추정오차의 제곱합이다.

Ⅳ. 고속 연산식의 안정성

앞절에 소개한 연산식의 안정성을 증명하기 위해 다음과 같이 Likeryhood Variable 을 정의 한다.

$$\gamma_t = x_t^T K_t \quad (17)$$

이 값은 y_t 가 Gauss 공정이 아니면 1에 가까와 지는 성질이 있으며 고속 연산식에서 스스로 최적 K_t 를 찾아가도록 하는 중요한 역할을 한다.[3]

이를 이용하면 Innovation e_t는

$$\begin{aligned} e_t &= y_t - x_t^T \hat{\theta}_t \\ &= y_t - x_t^T(\hat{\theta}_{t-1} + k_t \eta_t) \\ &= (1-\gamma_t)\eta_t \qquad \text{——— (18)} \end{aligned}$$

따라서 (13)식은

$$E_t = E_{t-1} + (1-\gamma_t)\eta_t^2 \qquad \text{——— (19)}$$

정리1) γ_t는 항상 1보다 작은 값을 갖는다.

증명) 식(15)의 양변에 x_{t+1}^T을 곱하면

$$\gamma_{t+1} = x_{t+1}^T(k_t^* + \hat{\varphi}_{t-1} k_t^{**})/(1-\gamma_t k_t^{**}) \quad \text{— (20)}$$

좌변분자를 다시 정리하면

좌변의 분자

$$= x_{t+1}^T \begin{bmatrix} E_t^{-1} e_t \\ \underline{k}_t - \underline{\hat{\theta}}_t E_t^{-1} e_t \end{bmatrix} + \underline{x_{t+1}^T \hat{\varphi}_{t-1}}(k_t^p - \hat{\theta}_t^p E_t^{-1} e_t)$$

$$= y_t E_t^{-1} e_t + \underline{x}_t^T \underline{k}_t - \underline{x_t^T \hat{\theta}_t E_t^{-1} e_t} + (y_{t-p} - \gamma_t)\cdot$$

$$(k_t^p - \hat{\theta}_t^p E_t^{-1} e_t) = y_t E_t^{-1} e_t - (\underline{x}_t^T \underline{\hat{\theta}}_t + E_t^{-1} e_t +$$

$$y_{t-p} \hat{\theta}_t^p E_t^{-1} e_t) + (\underline{x}_t^T \underline{k}_t + y_{t-p} k_t^p) - \gamma_t(k_t^p - \hat{\theta}_t^p E_t^{-1} e_t)$$

$$= (y_t - x_t^T \hat{\theta}_t) E_t^{-1} e_t + \underline{x_t^T k_t} - \underline{\gamma_t(k_t^p - \hat{\theta}_t^p E_t^{-1} e_t)}$$

$$= \gamma_t + E_t^{-1} e_t^2 - u_t \qquad \text{——— (21)}$$

여기서 벡터에 밑줄을 그은 것은 그 벡터의 마지박 항을 제외한 벡터이고, 오른편 위에 p를 쓴것은 그 벡터의 마지막 항을 의미하며 $u_t = \gamma_t(k_t^p - \hat{\theta}_t^p E_t^{-1} e_t)$ ——— (22) 이다.

이제(20)식은

$$\gamma_{t+1} = \frac{\gamma_t + E_t^{-1} e_t^2 - u_t}{1 - u_t} \qquad \text{——— (23)}$$

그런데 $E_t = E_{t-1} + e_t^2/(1-\gamma_t)$ ——— (24)

이므로

$$\gamma_t + E_t^{-1} e_t^2 = 1 - (1-\gamma_t)(1 - E_{t-1}/E_t) \quad \text{— (25)}$$

이다. (25)식과 (24)식을 비교하면 (25)의 우변은 $\gamma_t > 1$이면 1보다 크고 $\gamma_t < 1$ 이면 1보다 작으며 그 역도 성립한다. 그런데 $\gamma_{t+1} > 1$이면 $\gamma_t > 1$ 이고 이는 $\gamma_0 = 0$에 모순이다. 즉 $\gamma_t < 1$ 이다.

정리2) $u_t < 1$ 이다.

증명) (23)식을 변형하여 u_t에 관해 정리하면

$$u_t = \frac{(\gamma_t + E_t^{-1} e_t^2) - \gamma_{t+1}}{1 - \gamma_{t+1}} \qquad \text{——— (26)}$$

그런데 정리1) 에서

$$\gamma_{t+1} < 1 \qquad \text{——— (27)}$$

$$\gamma_t + E_t^{-1} e_t^2 < 1 \qquad \text{——— (28)}$$

이므로 $u_t < 1$ 이다.

정리1) 에서 $\gamma_t < 1$이므로 E_t 는 단조증가 함수 이고 그 초기 값은 작은 양의 실수로 잡으므로 항상 그 역수는 유한하며, 또한 $u_t < 1$ 이므로 유수를 줄이려는 방향으로 K_t 를 갱신시켜 나간다.

4. 결론

이동 특성을 이용한 고속 이득 연산식은 식이 간단하고 계산량이 적으며 최소 자승법에 근거한 연산식으로 수렴 속도가 빠른 특징을 갖고 있다. 수리적 안정성이 증명 되므로서 그 효용성을 한층 높였다.

앞으로 수렴속도의 해석, 잡음의 크기와 추정오차의 하한선과의 함수 관계의 해석 등의 문제가 있으며, 특히 기구변수를 사용할 수 있는 안정하고 견실한 연산식의 개발이 과제로 남아있다.

참고문헌

1) 양흥석, 남현도, 김진기, "스펙트럼 추정을 위한 공분산 기구 변수 격자알고리즘" 대한전기학회 논문지 Vol.35, 1986. 4

2) G.Carayannis et al.,"Fast Recursive Algorithms for a Class of Linear Equations", IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-30,NO.2,1982.4.

3) A.Vieira et al.,"A Classification of Algorithms for ARMA Models and Ladder Realizations", Appendix A of the Report to the Defense Communication Agency, Contact NO.DCA 100-77-C-0005, 1978.

4) D.T.Lee,"Canonical Ladder Form Realizations and Fast Estimation Algorithms", The PH.D. Assertation Paper, Stanford University,1980.

5) L.Ljung et al.,"Fast Calculation of Gain Matrices for Recursive Estimation Schemes", Int.J.Control, Vol.27, NO.1,1978.1.

6) A.J. Robins et al.,"Recursive System Identification Using Fast Algorithms", Int.J. Control, Vol.33, NO.3, 1981.

표) 고속 이득 연산식의 계산기 모의 실행 결과

k	a_1	a_2	a_3	a_4	$(\|\hat{\varepsilon}\|/\|\varepsilon\|)\times 100$
20	2.5[illegible]± 0.1[illegible]	-2.410± 0.371	2.251± 0.412	-0.759± 0.209	11.25
40	2.642± 0.099	-3.472± 0.301	2.285± 0.342	-0.737± 0.185	9.98
60	2.720± 0.062	-3.689± 0.147	2.525± 0.145	-0.859± 0.057	3.52
80	2.696± 0.057	-3.640± 0.125	2.477± 0.117	-0.841± 0.042	4.88
100	2.747± 0.035	-3.762± 0.065	2.597± 0.060	-0.889± 0.027	1.52
120	2.760± 0.014	-3.788± 0.043	2.620± 0.045	-0.912± 0.035	0.77
128	2.767± 0.012	-3.800± 0.034	2.637± 0.033	-0.901± 0.013	0.63
참값	2.7607	-3.8106	2.6535	-0.9238	–