

EM 알고리즘을 통한 칼만 필터의 성능 개선

강지혜, 김성수
충북대학교

Improved Kalman filter performance via EM algorithm.

Kang Jee-Hye, Kim Sung-Soo
Chungbuk National University

Abstract - The Kalman filter is a recursive Linear Estimator for the linear dynamic systems(LDS) affected by two different noises called process noise and measurement noise both of which are uncorrelated white. The Expectation Maximization(EM) algorithm is employed in this paper as a preprocessor to reinforce the effectiveness of Kalman estimator. Particularly, we focus on the relation between Kalman filter and EM algorithm in the LDS.

In this paper, we propose a new algorithm to improve the performance on the parameter estimation via EM algorithm, which improves the overall process of Kalman filtering. Since Kalman filter algorithm not only needs the system parameters but also is very sensitive the initial state conditions, the initial conditions decided through EM turns out to be very effective. In experiments, the computer simulation results are provided to demonstrate the superiority of the proposed algorithm.

1. 서 론

칼만 필터는 우수한 순환 선형 추정기로 다양한 분야에 걸쳐 그 응용력이 뛰어나다. 칼만 필터 알고리즘은 선형 필터링과 예측 문제의 전반에 걸쳐 응용된다. 특히 랜덤 신호들의 예측이나 랜덤 잡음으로부터 랜덤한 신호들의 분리, 랜덤 잡음이 포함된 원하는 신호의 검출에 적용된다. 이것은 모두 시스템의 파라미터를 알고 있다는 가정에서 출발하게 된다. 또한 관측된 데이터가 없는 초기의 추정 상태값과 추정오차 분산값은 전혀 알 수 없으므로 임의로 주어지게 된다. 초기치 선정에 있어서, EM 알고리즘을 이용하여 최대 가능성(Maximum Likelihood:ML)의 원리에 따라 반복적으로 그 값을 갱신하여 대가함수(cost function)가 국부 최소점에 도달하는 방향으로 순환하게 된다. 본 논문에서는 EM 알고리즘을 이용한 시스템 파라미터 추정을 새롭게 칼만 필터에 적용시켜 위에서 언급한 문제점을 보완하고 그 성능을 향상시키기 위한 알고리즘을 제시한다.

2. 본 론

2.1 칼만 필터 알고리즘

칼만 알고리즘의 핵심은 추정 상태의 순환 최소평균 자승오차를 최소화 하는 것에 있다. 이것은, 현재 각 상태에서 갱신된 추정값은 이전의 추정값과 새로 들어온 입력 데이터로부터 계산되고, 이 때의 순환 최소 평균 자승 추정 문제의 해결을 위해 특별한 스토캐스틱 프로세스의 상관 특성을 이용한 이노베이션 접근 방법이다.

시간이 n 일 때, 관측된 데이터 $y(n)$ 은 $n=1$ 에서부터 $n-1$ 까지의 관측된 과거 값들로 나타내고, 이때 관측된 데이터 $y(n)$ 의 최소 평균 자승 추정값을 $\hat{y}(n|y_{n-1})$ 과 같이 표현한다[1]. $y(n)$ 을 갖는 이노베이션 프로세스는 식(1)과 같이 측정된 데이터의 새로운 정보로 표현된다.

$$\alpha(n) = y(n) - \hat{y}(n|y_{n-1}) \quad (1)$$

이러한 이노베이션 프로세스는 관측된 데이터 $y(n)$ 의 선형 조합으로 표현된 식(2)와 같이 과거의 관측된 값들 $y(1), y(2), \dots, y(n-1)$ 과 직교를 이룬다.

$$E[\alpha(n)y(k)] = 0, \quad 1 \leq k \leq n-1 \quad (2)$$

이것은 Gram-Schimidet 직교화를 이용하여 얻을 수 있는데, 관측된 데이터 시퀀스들이 조건으로 주어졌을 때 $x(n)$ 의 최소평균자승추정 $\hat{x}(n|y_n)$ 은 등가적으로 이노베이션들이 조건으로 주어졌을 때의 최소 평균 자승 추정과 같기 때문이다. 그러므로 식(3)과 같은 선형조합의 관계로 놓을 수 있다.

$$\hat{x}(n|y_n) = \sum_{k=1}^n b_k \alpha(k), \quad b_k = \frac{E[x(n)\alpha(k)]}{E[\alpha(n)\alpha(k)]} \quad (3)$$

$$1 \leq k \leq n$$

$$\hat{x}(n|y_n) = \hat{x}(n-1|y_{n-1}) + b_n \alpha(n) \quad (4)$$

위 식(3)은 아래 식(4)와 같이 두 항으로 분리하여, 순환 추정 알고리즘의 표현으로 이전의 추정값 $\hat{x}(n-1|y_{n-1})$ 와의 관계로 나타낼 수 있다. 그러므로 이노베이션의 선형적 비율의 조합으로 새로운 추정치를 얻을 수 있다.

본 논문에서 전개하는 선형 이산 시간 동적 시스템은 식(5)로 표현되고, 여기서는 상태변수를 제어하는 또 다른 시스템 파라미터의 입력값 $u(n)$ 을 배제하였다. 프로세스 잡음 $v_1(n)$ 과 관측잡음 $v_2(n)$ 이 존재할 경우로 이들은 서로 통계적으로 독립인 AWGN으로 가정한다.

· 프로세스 방정식 :

$$x(n+1) = F(n+1, n)x(n) + v_1(n)$$

· 측정 방정식 : $y(n) = C(n)x(n) + v_2(n) \quad (5)$

$F(n+1, n)$ 은 상태전이 행렬, $C(n)$ 은 측정 행렬이고, 잡음의 통계적 특성은 다음의 식(6)으로 표현된다.

$$P(v_1) \approx N(0, Q_1), \quad P(v_2) \approx N(0, Q_2) \quad (6)$$

$$E[v_1(n)v_1^H(k)] = \begin{cases} Q_1(n) & n=k \\ 0 & n \neq k \end{cases}$$

$$E[v_2(n)v_2^H(k)] = \begin{cases} Q_2(n) & n=k \\ 0 & n \neq k \end{cases}$$

최소자승평균 추정 $\hat{y}(n|y_{n-1})$ 와 이노베이션 프로세스와의 관계를 살펴보면, 그 정의에 의해 아래의 식(7)과 같이 나타낼 수 있다.

$$\alpha(n) = C(n)\epsilon(n, n-1) + v_2(n) \quad (7)$$

여기서 $\epsilon(n, n-1) = x(n) - \hat{x}(n|y_{n-1})$ 은 시간 n 일 때, 상태벡터와 1단계 예측벡터와의 차로써 예상되는 상태-오차 벡터(state-error vector)이다. 구하고자하는 최소평균자승 추정치 $\hat{x}(n+1|y_n)$ 의 식(8)을 다음과 같이 쓸 수 있다.

$$\hat{x}(n+1|y_n) = F(n+1, n)\hat{x}(n|y_{n-1}) + G(n)\alpha(n) \quad (8)$$

이것은 선형능동시스템 상태의 최소 평균자승 추정값을 구하는데, 전처리 추정값 $\hat{x}(n|y_{n-1})$ 에 상태전이 행렬 $F(n+1, n)$ 이 곱해지고, 더해지는 보정항은 $G(n)\alpha(n)$ 이다. 이제 계산된 칼만 이득을 가지고 최소 평균자승 추정값 $\hat{x}(n+1|y_n)$ 의 식을 사용하여 1단계 (one-step) 예측을 할 수 있고, 또한 주어진 $\hat{x}(n|y_{n-1})$ 과거값을 가지고 $\hat{x}(n+1|y_n)$ 을 계산한다. 그러나, 칼만이득을 계산하는 식은 특별히 유용하지는 않다. 그것은 예측되는 상태오차 상관행렬 $K(n, n-1)$ 을 알아야 하기 때문에 이를 위한 반복적인 계산식이 유도된다. 예측되는 상태오차 벡터 $\epsilon(n+1, n)$ 는 상태 $x(n+1)$ 과 1단계 예측값 $\hat{x}(n+1|y_n)$ 의 차이이므로, 예측되는 상태-오차 상관행렬의 반복적인 계산을 위한 Riccati difference equation 식(9)을 얻는다.

$$K(n+1, n) = F(n+1, n)K(n)F^H(n+1, n) + Q_1(n) \quad (9)$$

앞에서 제시된 칼만 필터의 이론은 시스템 파라미터 $F(n+1, n), C(n), Q_1(n), Q_2(n)$ 를 모두 알고 있는 것으로 가정한다. 하지만 그 이론은 입력 $y(n)$ 에 의존하는 더 많은 행렬이 있는 상황을 포함해서 일반화 될 수 있다. 그런 상황에서, 칼만 이득과 예측되는 상태오차 상관행렬 $K(n+1, n)$ 미리 계산되지 않는다. 오히려, 그것들은 입력 $y(n)$ 에 의존한다. 이것은 $K(n+1, n)$ 이 조건부 오차 상관행렬이며, 입력 $y(n)$ 에 조건적으로 의존함을 의미한다. 또한, 고려할 다른 중요한 하나는, 앞에서 설명한 한 단계 예측과 필터링 알고리즘을 사용하기 위한 초기조건들이 필요하다는 것이다. 그러나 프로세스 방정식의 초기조건은 전혀 모르는 것이고, 또한 그것은 일반적으로 기대값과 상관행렬로 표현되어있다. 시간 $n=0$ 때에 어떤 관측된 데이터도 없으므로, 초기에 예측되는 추정값을 다음의 식(10)과 같이 선택한다.

$$\hat{x}(1|y_0) = E[x(1)] \quad (10)$$

그리고 그것의 상관행렬을 아래의 식(11)과 같이 쓴다.

$$K(1, 0) = E[(x(1) - E[x(1)])(x(1) - E[x(1)])^H] = \Pi_0 \quad (11)$$

초기조건에 대한 이 선택은 직관적으로 만족하고, 치우치지 않은 필터된 추정값 $\hat{x}(n|y_n)$ 을 전개하는 이점들 갖는다.

2.2 EM 알고리즘을 통한 시스템 파라미터 추정

선형 동적 시스템(LDS)에서의 파라미터 추정에 EM 알고리즘을 사용하는 것은 이미 알려진 사실이다. 시스템의 모델은 앞에서 제시된 칼만의 경우와 동일하고, 가우시안 분포 특성을 지닌 오차를 포함하고 있는 관측된 데이터를 가지고 있다. 다음의 식(12), 식(13)은 시스템 모델을 기반으로 하여 상태와 출력의 조건부 밀도함수를 구하는 것이다.

$$P(y|x) = \exp\left\{-\frac{1}{2}[y - Cx]'Q^{-1}[y - Cx]\right\}(2\pi)^{-p/2}|Q|^{-1/2} \quad (12)$$

$$P(x|x_{-1}) = \exp\left\{-\frac{1}{2}[x - Fx_{-1}]'Q^{-1}[x - Fx_{-1}]\right\}(2\pi)^{-k/2}|Q|^{-1/2} \quad (13)$$

종종 시스템의 입력 또한 관측되는 경우도 있다. 이러한 경우는 시스템의 입력력 반응을 모델링 하는 것이 중요하지만, 만약 시스템의 출력만 관측된다면 무감독(Unsupervised problem)문제로서 지금과 같이 입력이 배제된 시스템의 파라미터 추정에 EM 알고리즘을 적용하게 된다. 이러한 무감독 모델의 경우는 각 차원에서의 출력잡음이 서로 무상관하다는 가정아래, 최대 가능성(ML) 요소분석을 통하여 반복적으로 수행하게 된다.

선형 능동 시스템은 HMM(Hidden Markov model)의 연속적인 상태로써 볼 수도 있는데 HMM로부터의 전방향은 이미 LDS의 칼만 필터에 의해서 계산되고, 또한 Rauch의 순환을 사용하여 후방향 역시 계산할 수 있다[2]. 이러한 두 가지 순환들은 주어진 관측 데이터 시리로부터 상태의 확률을 추정하는 문제의 해결책으로 사용된다. 이것으로써 EM 알고리즘의 E 단계(Expectation step)를 기반으로 하여 후행확률을 구하게 된다.

EM 알고리즘의 E step은 다음의 식(14)과 같이 대수가능성의 기대값을 구하고자 한다.

$$Q = E[\log P(\{x\}, \{y\}) | \{y\}] \quad (14)$$

위 식의 값은 다음의 세 가지 기대값들(15)에 좌우된다.

$$\hat{x}_t = E[x_t | y]$$

$$P_t = E[x_t x_t^H | \{y\}] \quad (15)$$

$$P_{t, t-1} = E[x_t x_{t-1}^H | \{y\}]$$

여기서 상태 추정값 \hat{x}_t 은 칼만 필터에서 언급한 과거의 관측된 값과 예측된 값에 의존하는 상태 추정값 $\hat{x}_t = E[x_t | \{y\}]$ 과는 다른 것이다.

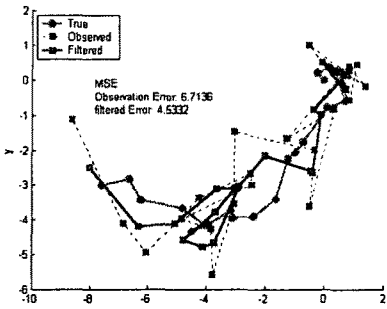
본 논문에서 언급한 시스템 파라미터는 F, C, Q_1, Q_2 과 초기 상태 추정값 $\hat{x}(0|y_0)$, 초기 상태 추정분산 $K(1, 0)$ 이다. 시스템 파라미터 추정에 사용되는 EM 알고리즘의 E-step은 임의로 주어진 시스템의 초기 파라미터로부터 칼만 필터의 순방향 순환과 후방향 순환을 이용하여 상태 추정값과 상태 오차 추정 분산값, 예측된 상태 추정값 $\hat{x}(n|y_{n-1})$, 예측된 상태 추정 오차 분산 $K(n, n-1)$ 을 구하게 된다. 이 값들과 관측된 데이터를 이용하여 앞서 언급했던 대수가능성의 기대값을 좌우하는 세 가지 기대값들을 계산하게 된다. 이렇게 해서 얻은 위의 세 파라미터들은 다음의 M-step에서 이것

을 최대로 만드는 시스템 파라미터를 추정하게 된다.

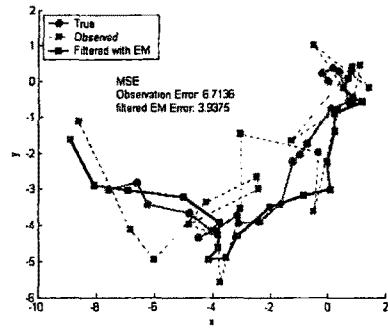
바로 M-step은 앞의 E step에서 구해진 대수 가능성의 기대값 Q 를 각각의 시스템 파라미터로 편미분을 하였을 때 0이 되는 값으로 구해지는데, 이렇게 시스템의 파라미터들은 갱신되는 과정을 반복적으로 수행한다. 이러한 과정을 통해서 최종적으로 가장 작은 대가함수(cost function)의 값으로 도달하는 즉, 최대의 가능성을 지닌 시스템 파라미터 값을 구하게 된다. 여기서도 반드시 고려해야 할 요소는 역시 초기상태의 값에 민감하게 반응하게 된다는 점이다. 이것은 칼만 필터나 EM 알고리즘과 같이 순환적으로 값을 갱신해주는 과정에서는 초기값에 따라 도달되는 오차분산의 극부 최소점이 달라진다. 본 논문에서 제안된 알고리즘의 핵심은 바로 칼만 필터를 이용한 위치 추적과정에서 시스템의 파라미터를 모를 경우나 초기의 추정값을 전혀 알 수 없는 대부분의 경우, EM 알고리즘을 통해서 최대의 가능성을 지닌 파라미터를 이용하는데 있다. 이것은 순환적인 과정을 수행함으로써 오차의 값을 점차 감소시켜 나가고, 최종적으로 최소의 추정오차 분산을 얻게 된다. 즉, 최대 가능성을 지닌 파라미터로 갱신됨에 따라 최소의 오차를 갖는 방향으로 진행하게 된다. 그 결과 오차를 포함한 관측된 데이터만을 가지고 추정한 것과 비교할 때 향상된 성능을 얻게 된다.

2.3 실험 및 고찰

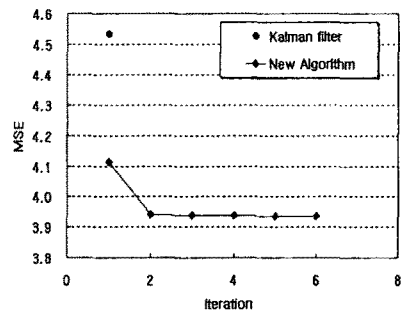
본 논문이 제안한 알고리즘의 타당성을 보이기 위해 프로세스 오차와 관측오차를 가지는 위치추적 시뮬레이션을 통해서 기존의 칼만 필터 알고리즘과 새롭게 제안된 알고리즘을 동일한 조건하에 실행하였다. 초기값과 초기 추정오차분산은 각각 제로입력과 단일분산으로 설정하였고, 관측된 데이터 주기는 20샘플로 주었다. 그림1은 기존 칼만 필터 알고리즘을 적용하였을 때의 추적과정과 관측오차, 추정오차(4.5332)를 보여준다. 그림에서 보는 바와 같이 추정된 값들이 오차를 포함하는 관측된 데이터에 민감하게 반응하고 있음을 알 수 있다. 이와 비교하여 그림2에서는 본 논문이 제안한 EM 알고리즘을 통한 시스템 파라미터 추정을 적용한 칼만 필터의 위치추적 과정과 그때의 추정오차(3.9773)를 제시하였다. 그림1과 비교해 볼 때 관측된 데이터의 변화에 덜 민감하며 실제의 위치데이터와의 분산이 작음을 알 수 있다. 두 그림에 제시된 값은 앞의 식에서 언급한 추정 오차분산 Mean Square Error(MSE)를 계산한 값이다. 그림3은 기존의 칼만 알고리즘을 적용하여 얻은 MSE값과 EM 알고리즘을 통한 시스템 파라미터를 적용한 칼만 필터과정을 반복적으로 수행하면서 추정오차의 분산이 수렴하는 과정을 반복 횟수와 MSE값을 통해 나타내었다.



< 그림 1. 기존의 칼만 필터를 통한 위치추정 >



< 그림 2. 제안된 알고리즘을 통한 결과 >



< 그림 3. 제안된 알고리즘의 반복과정을 통한 MSE비교 >

기존의 칼만 필터를 적용한 MSE와 시스템 추정을 통한 추정오차분산은 그 차이가 크다는 것을 쉽게 알 수 있다. 무엇보다도 제안된 알고리즘의 추정오차가 극부 최소점에 도달하기까지의 허용오차한계를 10^{-4} 로 설정하여 순환적으로 적용 했을 때, 두 번째 반복 과정에서 오차분산이 많이 감소되고 곧 수렴된다. 따라서 비교적 적은 반복횟수를 통해 향상된 추정결과를 얻을 수 있다. 뿐만 아니라, 민감한 초기값 설정 및 초기 오차분산의 값을 EM 알고리즘을 통해 적용하기 때문에 칼만 필터의 최종 추정오차가 감소하게 된다.

3. 결 론

본 논문에서는 순환 선형 추정기로 가장 많이 응용되고 있는 칼만 필터의 성능을 개선하고자 시스템의 파라미터 추정에 이용되는 EM 알고리즘을 적용시켰다. EM 알고리즘은 시스템의 파라미터를 모를 경우에만 이용되었는데 이것을 칼만 필터의 순환 과정에서 적용하여 초기의 상태 추정값과 초기 추정 오차의 분산 값을 설정할 수 있게 한다. 이것은 순환 알고리즘이 그 초기값에 민감하게 반응한다는 사실을 고려할 때 최종적으로 추정 오차 분산의 최소화를 통해 성능을 개선시켜준다. 아울러 동일한 조건하에 위치를 추적하는 시뮬레이션을 가지고 기존의 칼만 필터와 제안된 알고리즘을 적용한 결과를 비교하여 제시하였다.

[참고 문헌]

- [1] Simon Haykin, "Adaptive Filter Theory", Prentice Hall.
- [2] Zoubin Ghahramani, Geoffrey E. Hinton, "Parameter Estimation for Linear Dynamical System", U.Toronto tech. report, 1996