동적시스템의 상태변수 추정기법 (I)
– 이산칼만필터의 이해와 적용 –

■ 황 익 호
(국방과학연구소)

1. 개 요

동적 시스템이란 입력에 대한 시스템의 반응이 시간에 관련되어 표현되는 시스템이다. 이러한 동적 시스템의 거동과 성질을 잘 표현하기 위한 방법으로 상태변수(state)를 이용하는 방법이 많이 이용되어 왔다. 즉, 어떤 동적 시스템의 특성과 거동은 그 시스템의 상태변수들이 시간에 따라 어떻게 변화하는 가를 살펴봄으로써 이해할 수 있으며, 시스템을 제어한다는 것은 시스템의 상태변수들이 원하는 값을 가지도록 하는 것이라 할 수 있다. 그러므로 시스템의 상태변수를 추정해내는 것은 시스템의 현재 상황을 정확히 이해하고 이에 근거하여 적절한 판단을 수행하는데 있어서 매우 중요하다.

동적 시스템의 상태변수 추정문제는 시스템의 운동특성에 대한 정보와 시스템의 출력 측정치가 시스템과 어떤 관계를 가지는가에 대한 정보, 매시간 측정되는 측정치들을 통하여 획득되는 정보를 종합하여 시스템의 현재 상태변수를 알아내는 문제이다. 시스템의 운동 특성에 대한 정보는 시스템 상태변수들의 운동특성을 미분방정식 또는 차분방정식으로 표현한 시스템 모델을 통하여 표현하며, 측정치와 시스템간의 관계는 측정치와 상태변수간의 관계를 기술한 측정모델을 이용하여 표현한다. 그러나 이러한 모델은 실제 시스템의 개략적인 운동 특성을 수학적인 방법으로 나타낸 것이므로 실제 시스템의 운동과 정확히 일치하지 않는다. 이 뿐 아니라, 측정치에는 여러가지 원인에 기인한 측정오차가 포함된다. 따라서 시스템의 추정문제는 이러한 시스템 모델 불확실성 및 측정오차 등을 적절히 고려하여 최적의 상태변수 값을 추정하는 문제라 할 수 있다.

시스템 모델과 측정모델에 기반하여 측정치로부터 최적의 추정치를 얻는 문제에 대한 해법은 그 동안 많은 연구자들에 의하여 다양한 방법으로 시도되어 왔다. 이러한 방법들 중에서 전통적으로 많이 이용되는 것으로 최소자승 추정기법과 베이시안 추정기법(Bayesian estimation technique)을 들 수 있다[1,2]. 특히 시스템이 선형이고 시스템 모델 불확실성 및 측정오차가 정규분포를 가지는 확률프로세스(random process)의 합으로 표현되는 경우에 대한 결과들이 많이 연구되고 적용되었으며, 이 경우에 대한 대표적인 최적해가 칼만필터(Kalman filter)이다[1-3].

근래에 들어 여러가지 센서들을 저렴한 가격으로 이용할 수 있게 됨에 따라, 칼만필터를 제어시스템에 응용하려는 요구가 증대되고 있다. 이에 부응하여 이번 '알기쉬운 제어이론'에서는 2회에 걸쳐 베이시안 추정기법의 관점에서 이산칼만필터를 유도하고 그 의미를 설명하고자 한다. 구체적으로 이번 호에서는 Gauss-Markov모델로 모델링되는 동적 시스템의 상태변수 정보를 확률론적 방법으로 표현하는 방법

을 도입하고, 이러한 개념에서 상태변수를 추정하는 것은 어떤 의미인지를 설명하였으며, 다음 호에서는 이러한 설명을 바탕으로 이산칼만필터를 유도하고 그 특징을 소개하기로 한다.

2. 선형 이산시스템에 대한 추정문제

많은 경우에 제어 대상인 동적시스템의 운동은 공칭궤적(nominal trajectory) 주변에서 선형화한 미분방정식으로 표현할 수 있고, 이산화(discretization)과정을 통하여 다음과 같은 선형 차분방정식으로 표현할 수 있다[4].

$$\begin{cases} x_{k+1} = F_k x_k + \Gamma_k u_k, \ \ x_{k=0} = x_0 \\ z_k = H_k x_k \end{cases} \tag{2.1}$$

윗 식에서 x_k, u_k, z_k 는 각각 시각 k 에서의 상태변수, 입력, 측정치이며, F_k, Γ_k, H_k 는 각각 시스템 행렬(system matrix), 입력 행렬(input matrix), 측정행렬(measurement matrix)이다. 그러나 위의 선형화 및 이산화 과정을 거치는 과정에서 시스템의 운동모델은 오차를 포함하게 되며, 측정치 또한 측정 중 발생하는 다양한 요인의 오차를 수반한다. 이때, 이 오차들을 정규분포를 가지는 확률프로세스로 모델링 할 수 있다고 가정하면 다음과 같은 모델을 얻을 수 있다.

$$\begin{cases} x_{k+1} = F_k x_k + \Gamma_k u_k, + w_k \\ z_k = H_k x_k + v_k \end{cases} \tag{2.2}$$

식에서 w_k 는 시스템의 모델링시 발생한 오차를 의미하는 확률프로세스로서 평균이 0이고 공분산이 Q_k 인 정규분포를 가지며, 이를 $w_k \sim N(0;Q_k)$ 라 표시한다. 마찬가지로 v_k 는 측정오차를 나타내며 $v_k \sim N(0;R_k)$ 이다. 이들의 명칭으로, w_k 는 각각 공정잡음(process noise)이라 하며, v_k 는 측정잡음(measurement noise)라 한다. 마지막으로 (2.2)의 초기치도 $x_{k=0} \sim N(x_0;P_o)$ 와 같이 정규분포를 가지며, 그 평균과 분산을 안다고 가정한다.

그런데 (2.2)에서 $\Gamma_k u_k$ 의 크기는 불확실성이 없고 우리가 입력으로 넣어주는 이미 아는 값이므로 이 것에 의하여 (2.2)에 포함된 정보의 불확실성이 변화하지는 않는다. 따라서 일반적인 추정문제를 고려할 때에는 (2.3)과 같이 $\Gamma_k u_k$ 가 없는 모델을 고려한다. (2.3)과 같은 모델을 Gauss-Markov 시스템 모델이라 하며, (2.3.a)는 시스템 모델, (2.3.b)는 측정모델이라 한다.

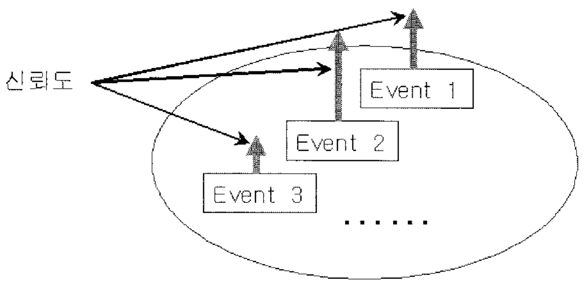


그림 2.1 정보 표현 모델

$$x_{k=1} = F_k x_k + w_k \tag{2.3.a}$$
$$z_k = H_k x_k + v_k \tag{2.3.b}$$

만약 $\Gamma_k u_k$ 에 의한 영향을 고려하여야 하는 경우에는 (2.1)의 해에서 $\Gamma_k u_k$ 에 의한 영향이 고려된 부분만 추출하여 상태변수 추정치에 추가하여 필터(상태변수 추정기)를 구성하면 된다.

이제 정보를 표현하는 방법에 대하여 생각해 보자. 일반적으로 어떤 정보를 이용하기 위해서는 그 정보의 내용과 신뢰도를 동시에 고려하여야 한다. 따라서 정보를 표현하는 방법으로 그림 2.1과 같이 해당 정보와 관련된 사건들을 모두 모아 놓고 그 사건들이 사실일 신뢰 확률들을 지정한다면 적절한 정보 표현 모델이 될 수 있다[3].

그런데, 확률공간(probability space)이란 관련된 사건들을 모두 모아놓고 이들에 대하여 적절한 확률측도(probability measure)를 지정한 것이므로, 기본적으로 그림 2.1과 같은 구조이다. 그러므로 확률공간은 정보를 표현하고 이용하기 위한 적절한 수학적 모델이라 할 수 있다. 확률공간을 구성하기 위해서는 어떤 정보와 관련된 모든 사건들과 이들 사건들의 교집합 및 합집합을 모두 포함하는 전체 집합 Ω 를 생각하고, 이 전체집합에 확률밀도함수(probability density function) 등을 통하여 합리적인 신뢰확률을 지정하면 된다. 이를 상태변수의 경우로 맞추어 생각해 보면, 상태변수가 가질 수 있는 값 들의 집합(보통은 상태변수의 차원에 해당하는 크기의 벡터공간)을 생각하고, 이 집합에 대하여 확률밀도함수를 적절히 결정함으로써 상태변수에 관련된 정보를 표현할 수 있음을 알 수 있다.

이러한 관점에서 생각해 볼 때, 상태변수 추정문제는 시간에 따라 측정치에 의하여 새로이 인가되는 정보를 잘 고려하여 상태변수의 확률밀도함수(또는 확률)를 갱신하는 문제라 할 수 있다. 새 측정치 정보를 확률공간에 집어 넣어서 갱신된 확률밀도함수(또는 확률)를 얻는 방법은 일반적으로 조건부

확률 밀도함수(또는 조건부 확률)를 계산하는 것이다. 조건부 확률 $P(X|Z)$란 어떤 사건 Z가 일어난 조건에서 사건 X가 일어날 확률을 의미한다. 즉, 사건 X가 발생하는 것에 관련된 모든 사건을 포함하는 전체집합을 Ω라 한다면 $P(X)$는 Ω에 속하는 모든 경우들에 대하여 정의되는 확률인 반면에, $P(X|Z)$는 $\Omega \cap Z$로 축소된 집합에 대하여 정의되는 확률이다. 따라서 조건부 확률 $P(X|Z)$를 계산한다는 것은 사건 Z가 발생하지 않는 경우는 모두 무시함으로써 X에 대한 불확실성을 감소시키는 과정이라 할 수 있다. 이러한 조건부 확률은 다음의 Bayesian Rule에 의하여 계산할 수 있다.

$$P(X|Z) = \frac{P(XZ)}{P(Z)} \tag{2.4}$$

윗 식에서 $P(XZ)$는 사건 X와 Z가 동시에 발생하는 확률이다.

조건부 확률을 이용한 정보의 갱신 개념에 대한 이해를 돕기 위하여 다음의 예를 생각해 보자.

〔예 2.1〕 조건부 확률에 의한 정보 갱신

A, B의 두 개의 동전을 던져서 앞면이 나오는 횟수를 X라 하고, A동전의 앞면이 나오는 횟수를 Z라 하자. 이 경우 X가 가질 수 있는 경우는 0, 1, 2이고, Z는 0, 1이다. 모든 동전의 앞면과 뒷면이 나올 확률이 같다고 가정하고 $P(X)$와 $P(X|Z)$를 계산해 보자.

$P(X=0)$은 동전 A와 B가 모두 뒷면이 나오는 한가지 경우에 대한 확률이며, $P(X=1)$은 A만 앞면이 나오는 경우와 B만 앞면이 나오는 경우의 두 가지, $P(X=2)$는 동전 A와 B가 모두 앞면이 나오는 한 가지 경우에 대한 확률이므로 다음과 같은 확률을 계산할 수 있다.

$$P(X=0) = \frac{1}{4},$$
$$P(X=1) = \frac{1}{2},$$
$$P(X=2) = \frac{1}{4}$$

이번에는 동전 A의 앞면이 나왔다는 것을 알고 있는 경우, 즉 $\{Z=1\}$인 경우에 대한 X의 조건부 확률을 구해보자. 이 경우 이미 앞면이 하나 나왔으므로 X가 가질 수 있는 범위는 1, 2 두 가지뿐이며 동전 B가 앞면이냐 뒷면이냐에 의하여 결정된다. 따라서, 조건부 확률은 다음과 같이 계산된다.

$$P(X=0|Z=1) = 0,$$
$$P(X=1|Z=1) = \frac{1}{2},$$
$$P(X=2|Z=1) = \frac{1}{2}$$

이상의 결과를 살펴보면, $\{Z=1\}$이라는 정보가 추가됨에 따라 X가 가질 수 있는 값의 범위와 그 확률이 동시에 변화하였음을 알 수 있다. 예를 들어 $\{Z=1\}$이라는 정보가 없다면 우리는 앞면이 하나도 나오지 않을 수 있다는 생각을 할 수도 있지만, $\{Z=1\}$을 안다면 동전들이 모두 뒷면이 나올 수도 있을 것이라는 생각은 버려야 하며, 따라서 $P(X=0|Z=1) = 0$ 이다. 이 예는 조건부 확률을 통하여 X의 값이 얼마인지에 대한 정보에 $\{Z=1\}$이라는 정보를 성공적으로 부가하였음을 의미한다. ■

예에서 확인하였듯이 확률론적 관점에서 추정문제의 본질은 측정치에 대한 상태변수의 조건부 확률을 시간에 따라 갱신하여 나가는 과정이라 할 수 있다. 마코프 이산 시스템의 경우, 일반적으로 이 과정은 시스템 전파(system propagation)와 측정치 갱신(measurement update)의 두 단계를 통하여 수행할 수 있음이 알려져 있다[1,5].

시스템 전파:

$$\begin{aligned} p(x_k|Z^{k-1}) &= \int p(x_k|x_{k-1}, Z^{k-1}) p(x_{k-1}|Z^{k-1}) dx_{k-1} \\ &= \int p(x_k|x_{k-1}) p(x_{k-1}|Z^{k-1}) dx_{k-1} \end{aligned}$$

측정치 갱신:

$$\begin{aligned} p(x_k|Z^k) &= \frac{p(z_k|x_k, Z^{k-1}) p(x_k|Z^{k-1})}{p(z_k|Z^{k-1})} \\ &= \frac{p(z_k|x_x) p(x_k|Z^{k-1})}{\int p(z_k|x_k, Z^{k-1}) p(x_k|Z^{k-1}) dx_k} \end{aligned}$$

윗 식에서 x_k는 시간 k에서의 상태변수 값이고, z_k는 k에서 측정된 측정치이며, Z^k는 시간 k까지 누적된 측정치 전체의 집합 $Z^k = \{z_1, z_2, \cdots, z_k\}$이다. 또한 $p(x|z)$는 x의 조건부 확률밀도함수이다.

시스템 전파란 시스템의 동적 특성을 고려하여 시간 $k-1$에서의 상태변수에 대한 정보를 시간 k에서의 정보로 갱신

하는 것을 말하며, 이 과정을 통하여 생성된 조건부 확률밀도함수 $p(x_k \mid Z^{k-1})$를 사전 확률밀도함수(*a priori* probability density)라 한다. 측정치 갱신이란 시간 k에서 입력된 측정치 z_k의 정보를 추가하여 상태변수 x_k에 대한 불확실성을 줄이는 과정이며, 이 때 만들어지는 상태변수의 조건부 확률밀도함수 $p(x_k \mid Z^k)$를 사후 확률밀도함수(*a posteriori* probability density)라 한다. (2.5)와 (2.6)은 확률적분(stochastic integral)을 포함하고 있으므로, 일반적으로 이를 계산하는 것은 매우 어려운 일이며 특별한 경우들에 한하여 그 해가 연구되어 있다.

3. 추정치 선정 기준

앞장에서 설명하였듯이 추정문제는 상태변수로 이루어지는 확률공간의 조건부확률밀도함수를 적절히 결정하는 문제로 정의할 수 있다. 그러나 조건부 확률밀도함수를 구하였다고 하더라도 실제로 추정 결과를 적용하기 위해서는 특정한 추정치가 요구되는 경우가 일반적이다. 본 장에서는 상태변수의 조건부 확률밀도함수가 주어진 경우에 어떤 기준으로 최종 추정치를 얻어낼 것인가 하는 문제에 대하여 설명하였다.

확률밀도함수가 가용한 상황에서는 사실상 상태변수의 모든 정보가 있는 것이므로 어떠한 기준의 최종추정치든지 모두 계산할 수 있다. 추정치 선정문제는 그 집단의 통계적 특성을 모두 아는 경우에 그 집단의 대표값을 선정하는 문제와 같은 문제이므로 흔히 사용하는 평균값, 중간값, 최빈값 등의 개념을 모두 적용할 수 있으며, 적용하고자 하는 문제의 성질에 적합한 추정치 기준을 설정하여 이용하여야 한다. 공학 문제에서 일반적으로 많이 이용하는 개념은 최소오차분산(minimum error variance) 추정치이다.

최소 오차분산 추정치는 최소자승(least squares) 추정치 또는 최소자승평균(minimum mean squares) 추정치 등과 같은 개념의 추정치로서, 추정 오차의 조건부 분산을 최소화하는 추정치이다. 다시 말하면, 최소오차분산 추정치 $\hat{x}$는 상태변수 x가 가질 수 있는 모든 값에 대하여 $J(\hat{x}) = E\{(x-\hat{x})^T(x-\hat{x}) \mid Z\}$을 최소화하는 추정치이다. 이 값을 구하기 위하여 임의의 추정치 y에 대한 $J(y)$를 계산해 보자.

$$
\begin{aligned}
J(y) &= E\{(x-y)^T(x-y) \mid Z\} \\
&= \int (x-y)^T(x-y) p_{X|Z}(x \mid Z)dx \\
&= \int x^T x p_{X|Z}(x \mid Z)dx - 2y^T \int x p_{X|Z}(x \mid Z)dx + y^T y \\
&= \left[y^T - \int x^T p_{X|Z}(x \mid Z)dx\right]\left[y - \int x p_{X|Z}(x \mid Z)dx\right] \\
&\quad + \int x^T x p_{X|Z}(x \mid Z)dx - \left|\int x p_{X|Z}(x \mid Z)dx\right|^2
\end{aligned}
$$

(3.1)은 y에 대한 2차식 이므로, $J(y)$의 최소값을 주는 최적의 y^*를 $\hat{x}$이라 하면, $\hat{x}$는 다음과 같이 조건부 평균으로 주어짐을 알 수 있다[6].

$$\hat{x} = \int x p_{X|Z}(x \mid Z)dx = E_{X|Z}\{x \mid Z\} \tag{3.2}$$

이상을 종합하면 측정치에 의하여 매시간 갱신되는 조건부확률밀도함수를 이용하여 상태변수의 조건부 평균을 구하면 그 값은 최소오차분산의 개념에서 최적의 추정치가 될 수 있다.

4. 소 결론

이번 호에서는 확률론적 관점에서의 베이시안 추정이론 개념을 설명하였다. 지금까지 언급한 내용을 정리하면 다음과 같다.

- 많은 동적 시스템은 공칭궤적 주변에서 (2.3)과 같은 Gauss-Markov 모델로 모델링 할 수 있다.
- 상태변수 추정문제의 본질은 측정치에 대한 상태변수의 조건부 확률을 시간에 따라 갱신하여 나가는 과정이라 할 수 있다.
- 조건부 확률이 주어지면 조건부 평균을 구함으로써 최소오차분산 추정치를 계산할 수 있다.

이러한 관점에서 이산칼만필터 문제를 정리하면 다음과 같다.

[이산칼만필터 문제]

다음과 같은 모델에 따라 운동하는 동적 시스템을 고려하자.

$$x_{k+1} = F_k x_k + w_k \tag{2.3.a}$$
$$z_k = H_k x_k + v_k \tag{2.3.b}$$

여기서 x_k, u_k, z_k는 각각 시각 k에서의 상태변수, 입력, 측정치이며, F_k, Γ_k, H_k는 각각 시스템 행렬, 입력 행렬, 측정

행렬이다. 상태변수의 초기치 $x_{k=0}$와 공정잡음 w_k 및 측정잡음 v_k 는 서로 독립이며 그 확률분포는 다음과 같이 정규분포를 이룬다.

$$x_{k=0} \sim N(x_0;P_0),\ w_k \sim N(0;Q_k),\ v_k \sim N(0;R_k)$$

또, 공정잡음과 측정잡음은 시간에 대하여 서로 독립인 백색잡음(white noise)이다. 이와 같은 시스템에 대하여 상태변수의 조건부 확률밀도함수 $p(x_k \mid Z^k)$를 계산하고, 상태변수의 최소오차분산 추정치을 구하여라. ■

다음 호에서는 정규확률변수의 특성을 이용하여 이와 같은 이산칼만필터 문제의 해를 구하고, 그 특성을 설명하도록 하겠다.

참고문헌

[1] 황익호, 나원상, '베이시안 상태변수 추정기와 다중가설 필터기법,' 제어 · 로봇 · 시스템학회지, 제14권 제3호, 2008.

[2] 나원상, 황익호 '최소자승관점에서의 선형상태추정기법,' 제어 · 로봇 · 시스템학회지, 제14권 제3호, 2008.

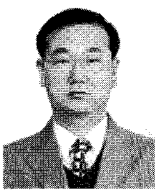
[3] 황익호, '고급상태변수 추정문제에 대한 확률론적 접근-다중가설필터 기법을 중심으로,' 중앙대학교 정보통신연구원 세미나 자료, 2007.

[4] C.-T. Chen, Linear System Theory and Design, Holt, Rinehart and Winston, 1984.

[5] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, Academic Press, 1979.

[6] F. L. Lewis, *Optimal Estimation*, John Wiley & Sons, 1986.

저자약력



황 익 호

- 1988년, 1990년, 1995년 서울 대학교 공과대학 제어계측공학과 공학사, 공학석사, 공학박사.
- 1995년~현재 국방과학연구소 유도조종부 책임연구원.
- 관심분야 : 유도조종기법, 추정론, 표적추적필터 등