

동적시스템의 상태변수 추정기법 (II)

- 이산칼만필터의 이해와 적용 -

■ 황 익 호

(국방과학연구소)

1. 개요

지난 호에서는 동적 시스템의 확률적 모델링 방법에 대하여 설명하고, 확률론적 관점에서 추정이론의 개념을 소개하였다. 많은 동적 시스템은 Gauss-Markov 모델로 모델링 할 수 있으며, 이 동적시스템의 상태변수 추정문제는 측정치에 대한 상태변수의 조건부확률을 순차적으로 갱신하여 나아가는 과정이라 할 수 있음을 설명하였다. 전보에서 기술하였듯이 이러한 추정문제는 다음과 같은 이산칼만필터 문제로 귀결되므로, 본 고에서는 정규분포(normal probability distribution, or Gaussian probability distribution)를 가지는 확률변수들의 성질을 이용하여 이산칼만필터(discrete Kalman filter) 문제의 해를 구하고 그 특성을 간단히 살펴보도록 하겠다.

우선 이산칼만필터 문제를 다시 살펴보자.

[이산칼만필터 문제]

다음과 같은 모델에 따라 운동하는 동적 시스템을 고려하자.

$$x_{k+1} = F_k x_k + w_k \quad (1.1.a)$$

$$z_k = H_k x_k + v_k \quad (1.1.b)$$

여기서 x_k , z_k 는 각각 시각 k 에서의 상태변수, 측정치이며, F_k , H_k 는 각각 시스템 행렬, 측정행렬이다. 상태변수의 초기치 $x_{k=0}$ 와 공정잡음 w_k 및 측정잡음 v_k 는 서로 독립이며

그 확률분포는 다음과 같이 정규분포를 이룬다.

$$x_{k=0} \sim N(\hat{x}_0; P_0), \quad w_k \sim N(0; Q_k), \quad v_k \sim N(0; R_k)$$

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

위의 문제를 살펴보면 이산칼만필터문제를 이루는 확률벡터(random vector) 또는 확률프로세스(random process)는 모두 정규분포(Normal distribution, 또는 Gaussian distribution)와 관련이 되어 있으며, 시스템 모델 및 측정 모델을 이루는 차분방정식은 모두 선형방정식(linear equation)으로 이루어져 있음을 알 수 있다. 이러한 특성을 보다 잘 활용하여 해를 쉽게 구하기 위하여 정규분포를 가지는 확률벡터 또는 확률프로세스의 성질을 재확인하는 것으로 다음 장을 시작한다.

2. 정규확률변수의 특성

본 장에서는 정규분포를 가지는 확률변수의 몇 가지 특징을 알아보고 다음 장에서는 이를 이용하여 이산칼만필터의 해를 구하고자 한다. 다음은 정규확률변수 또는 정규확률프로세스의 대표적인 성질들이다.

[NRV1] 정규분포를 가지는 확률변수 또는 벡터는 평균과 분산만으로 전체 확률밀도함수를 표현할 수 있다.

평균 \bar{x} , 분산 P 인 정규확률분포를 가지는 확률벡터 (random vector) x 의 확률밀도함수(pdf, probability density function) $p(x)$ 는 다음 식과 같이 주어지며, 이를 $x \sim N(\bar{x}; P)$ 와 같이 표현한다.

$$p(x) = \frac{1}{(2\pi)^{n/2} |P|^{1/2}} \exp \left\{ -\frac{1}{2} (x - \bar{x})^T P^{-1} (x - \bar{x}) \right\} \quad (2.1)$$

윗 식에서 $|P|$ 는 P 의 determinant를 의미한다. 식에서 알 수 있듯이 확률밀도함수는 평균과 분산만으로 표현되어 있다. 정규분포를 가지는 이산확률프로세스인 경우는 위의 식에서 \bar{x} 와 P 가 시간에 대하여 변화하는 경우로 생각하면 된다. 이와 같은 특성은 어떤 확률프로세스가 정규분포라는 것을 알 경우, 전체 확률밀도함수의 변화를 추적할 필요 없이 단지 그 평균과 분산만을 추적하더라도 그 확률프로세스의 모든 정보를 손실 없이 유지하기에 충분하다는 사실을 알려 준다.

[NRV2] 정규분포를 가지는 두 개의 확률변수 또는 벡터의 선형조합도 정규분포를 가진다.

두 개의 확률벡터 x, y 가 상호정규분포(jointly normal probability distribution)를 가지며, 그 평균과 분산이 각각 $\bar{x}, \bar{y}, P_x, P_y$ 이고, 공분산(covariance)이 P_{xy} 라 하자. 이 경우, 두 확률벡터의 선형조합으로 표현되는 확률벡터 $Z = Ax + By$ 도 역시 정규분포를 가지며 그 평균 \bar{z} 와 분산 P_z 는 다음과 같다[4].

$$\bar{z} = A\bar{x} + B\bar{y} \quad (2.2)$$

$$P_z = AP_x A^T + AP_{xy} B^T + BP_{yx} A^T + BP_y B^T \quad (2.3)$$

앞에서와 마찬가지로 이 성질은 확률프로세스인 경우에도 성립한다.

[NRV3] 상호정규분포확률벡터의 조건부 확률벡터도 역시 정규분포를 가진다.

상호정규분포확률벡터 x, y 로 이루어진 정규확률벡터 $z = \begin{bmatrix} x \\ y \end{bmatrix}$ 를 생각해 보자. 이 벡터의 평균과 분산을 각각

$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}, \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix}$ 라 하면, $\{y=b\}$ 라는 조건에 대한 x 의 조건부확률분포도 역시 정규분포이며 그 평균과 분산은 다음과 같다[4,5].

$$E\{x | y=b\} = \bar{x} + P_{xy} P_y^{-1} (b - \bar{y}) \quad (2.4)$$

$$P_{x|y=b} = P_x - P_{xy} P_y^{-1} P_{yx} \quad (2.5)$$

[NRV4] 정규분포확률벡터의 확률밀도함수가 최대인 점은 하나 뿐이며, 그 위치는 그 확률벡터의 평균 지점이다.

이 성질은 (2.1)에서 명백하다.

다음 장에서는 이와 같은 성질을 이용하여 (1.1)의 Gauss-Markov 모델을 통하여 결정되는 상태변수 벡터의 통계적 특성이 어떻게 전이되는지를 살펴보기로 하자.

3. Gauss-Markov 모델의 통계적 특성과 이산칼만필터

본 장에서는 2장의 지식을 이용하여 x_k 의 확률밀도함수를 추적함으로써 이산칼만필터의 해를 유도한다. 지난 호에서 기술하였듯이 추정문제는 측정치에 대한 상태변수의 조건부 확률분포를 구하는 것이다. 측정치에 대한 조건부확률은 시간 k 에서의 측정치 z_k 의 인가여부에 따라 구별되는 사전 확률밀도함수(a priori probability density) $p(x_k | Z^{k-1})$ 와 사후 확률밀도함수(a posteriori probability density) $p(x_k | Z^k)$ 의 두 가지가 있으며, 이들을 구별하기 위하여 시간에 대한 첨자를 $k | k-1$ 과 $k | k$ 의 두 가지로 구분하여 사용하기로 한다. 이제 $k=0$ 에서 시작하여 순차적으로 상태변수의 확률밀도함수를 계산하자.

[k=0] 이산칼만필터의 가정에서 상태변수는 정규분포를 이루며, 그 평균과 분산은 각각 \hat{x}_0, P_0 이다.

[k=1] x_0 와 w_0 가 정규확률벡터이므로 [NRV2]에 의하여 x_1 의 사전 확률분포도 정규분포 $x_{10} \sim N(\hat{x}_{10}; P_{10})$ 를 이루며, 그 평균과 분산은 (2.2)-(2.3)을 이용하여 다음과 같이 계산할 수 있다.

$$\hat{x}_{10} = F_0 \hat{x}_0 \quad (3.1.a)$$

$$P_{10} = F_0 P_0 F_0^T + Q_0 \quad (3.1.b)$$

한편, (1.1.b)를 살펴보면, 측정치 z_1 은 두 개의 확률벡터 x_1 과 v_1 의 선형조합으로 주어지므로, z_1 도 역시 정규분포를 가지며, 평균과 분산은 다음과 같음을 알 수 있다.

$$E\{z_1\} = H_1 \hat{x}_{1|0} \quad (3.2.a)$$

$$P_z(1) = E\{(z_1 - E\{z_1\})(z_1 - E\{z_1\})^T\} \\ = E\{(H_1 x_1 + v_1 - H_1 \hat{x}_{1|0})(H_1 x_1 + v_1 - H_1 \hat{x}_{1|0})^T\} \quad (3.2.b) \\ = H_1 P_{1|0} H_1^T + R_1$$

또한, 상태변수 x_1 과의 공분산도 다음과 같이 계산할 수 있다.

$$P_{zx}(1) = E\{(x_1 - \hat{x}_{1|0})(z_1 - E\{z_1\})^T\} \\ = E\{(x_1 - \hat{x}_{1|0})(H_1 x_1 + v_1 - H_1 \hat{x}_{1|0})^T\} \quad (3.2.c) \\ = P_{1|0} H_1^T$$

$$P_{zx}(1) = [P_{zx}(1)]^T = H_1 P_{1|0} \quad (3.2.d)$$

그러므로 [NRV3]의 성질에 따라 측정치 z_1 이 인가된 조건에 대한 조건부 확률 $p(x_1 | Z^1 = \{z_1\})$ 도 역시 정규확률밀도 함수를 이루며, 그 평균과 분산은 (2.4)-(2.5)에 의하여 다음과 같이 주어짐을 알 수 있다.

$$\hat{x}_{1|1} = \hat{x}_{1|0} + P_{1|0} H_1^T [H_1 P_{1|0} H_1^T + R_1]^{-1} (z_1 - H_1 \hat{x}_{1|0}) \quad (3.3.a)$$

$$P_{1|1} = P_{1|0} - P_{1|0} H_1^T [H_1 P_{1|0} H_1^T + R_1]^{-1} H_1 P_{1|0} \quad (3.3.b)$$

이상의 결과를 종합하면 $k=1$ 의 시간이 종료된 후 상태변수의 사후확률밀도함수도 역시 정규확률분포를 가지는 것을 알 수 있으며, 이는 $k=0$ 에서 상태변수가 정규분포를 가지는 상태로 출발한 것과 동일한 상황임을 의미한다. 따라서 $k=2, k=3, \dots$ 에서도 동일한 방법으로 상태변수의 확률밀도함수를 계산할 수 있다. 이러한 맥락에서 (3.1)과 (3.3)을 일반화하면 다음과 같은 칼만필터 관계식을 얻을 수 있다.

Kalman Filter System Propagation

$$\hat{x}_{k+1|k} = F_k \hat{x}_{k|k} \quad (3.4.a)$$

$$P_{k+1|k} = F_k P_{k|k} F_k^T + Q_k \quad (3.4.b)$$

Kalman Filter Measurement Update

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1} (z_{k+1} - H_{k+1} \hat{x}_{k+1|k}) \quad (3.5.a)$$

$$P_{k+1|k+1} = (I - K_{k+1} H_{k+1}) P_{k+1|k} \quad (3.5.b)$$

여기서

$$K_{k+1} = P_{k+1|k} H_{k+1}^T [H_{k+1} P_{k+1|k} H_{k+1}^T + R_{k+1}]^{-1} \quad (3.5.c)$$

칼만필터 식에서 $\hat{x}_{k+1|k}$ 와 $P_{k+1|k}$ 는 시간 $k+1$ 에서 상태변수의 사전 추정치(a priori estimate) 및 그 오차공분산(a priori error covariance)이라 하고 $\hat{x}_{k+1|k+1}$ 과 $P_{k+1|k+1}$ 은 상태변수의 사후 추정치(a posteriori estimate) 및 그 오차공분산(a posteriori error covariance)이라 하며, 각각 측정치 z_{k+1} 이 인가된 전/후의 상태변수 추정치와 그 오차 공분산을 의미한다.

4. 이산칼만필터의 성질

본 장에서는 앞에서 유도된 이산칼만필터의 기본적인 특징들을 몇 가지 살펴보겠다.

첫째, 칼만필터로 구해지는 것은 상태변수의 추정치와 오차공분산 뿐만 아니라, 상태변수의 확률밀도함수 전체이다. 그 이유는 Gauss-Markov 모델과 칼만필터를 통하여 구해지는 상태변수의 확률분포는 모두 정규분포이며, 정규분포는 그 평균과 분산만으로 전체 확률밀도함수가 표현될 수 있기 때문이다.

그림 4.1은 이 내용을 이해하기 쉽게 설명한 그림이다. 시간 $k-1$ 에서 정규분포를 이루고 있는 상태변수는 시간이 k 로 시스템 전이하면서 공정잡음 등으로 인하여 상태변수의 불확실성이 커지게 된다. 이는 그림에서 $p(x_{k-1} | Z^{k-1})$ 이 $p(x_k | Z^{k-1})$ 보다 더 넓적한 형태가 되는 것으로 표현되었다. 실제로 (3.4.b)를 보면 시스템 전이 후의 공분산 $P_{k+1|k}$ 의 크기는 이전 시간에서의 공분산 전이량 $F_k P_{k|k} F_k^T$ 보다 공정잡음 공분산 Q_k 만큼 커진 것을 알 수 있다.

이와 같이 불확실성이 커진 상태변수의 분포는 측정치가 인가됨으로써 그 양이 축소되게 되는데, 그림에서 측정치 갱신을 통하여 $p(x_k | Z^{k-1})$ 보다 $p(x_k | Z^k)$ 의 확률분포가 측정치 주변으로 이동하고 동시에 평균 부위에 집중 분포하는 형태로 변화하는 것으로 표현되었다. 이는 (3.5.a)에 의하여 확률분포의 평균이 측정치 주변으로 이동하고, (3.5.b)에 의하여 측정치 갱신 후 공분산 $P_{k+1|k+1}$ 가 측정치 갱신 전 공분산 $P_{k+1|k}$ 보다 $K_{k+1} H_{k+1} P_{k+1|k}$ 만큼 축소되는 것을 통하여 확인할 수 있다.

둘째, 칼만필터를 통하여 얻어지는 상태변수 추정치 $\hat{x}_{k|k}$ 는 최소오차분산추정치(Minimum Variance estimate)인 동시에 최대사후확률 추정치(Maximum A Posteriori estimate)이다. 지

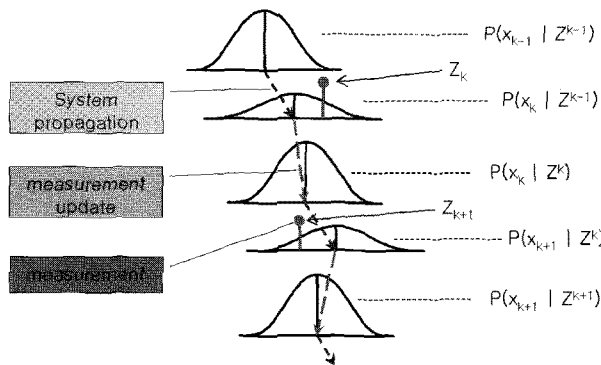


그림 4.1 칼만필터로 구해지는 확률밀도함수 변화

난 호에서 설명하였듯이 최소오차분산 추정치는 조건부 평균값이므로 당연히 칼만필터의 추정치는 최소오차 분산 추정치이다[1,4,5]. 또한 [NRV4]의 성질에 따라 조건부 평균에서 정규확률밀도함수의 최대확률이 나타난다. 이 뿐 아니라 다른 연구결과에 의하면 칼만필터는 정규분포라는 가정이 없는 경우에도 최소자승오차를 제공하는 선형필터임이 알려져 있다[5].

마지막으로 칼만필터가 정상적으로 동작할 때의 잔류오차(residual) $r_{k+1} \triangleq z_{k+1} - H_{k+1} \hat{x}_{k+1|k}$ 는 평균이 0이고 분산이 $H_{k+1} P_{k+1|k} H_{k+1}^T + R_{k+1}$ 인 정규분포를 가짐이 알려져 있다. 따라서 칼만필터를 구동하였을 때에 잔류오차의 평균이 0이 아니고 편향오차(bias error)를 가진다면, 칼만필터를 설계할 때 이용한 모델의 오차가 상당히 크다고 생각할 수 있다. 특히 칼만필터는 모델에 근거하여 동작하는 필터이므로 필터 설계 시 고려한 모델과 실제 모델의 차이가 큰 경우에는 오동작할 수 있으며, 이를 확인하는 방법으로 칼만필터의 잔류오차를 관찰하는 것이 일반적인 방법이다. 만약 모델링 오차가 매우 큰 시스템에 대한 필터를 설계하여야 한다면 본 고에서 소개된 기본적인 칼만필터는 적합하지 않을 수 있으며, 다른 강인 필터기법이나 적응필터기법들을 고려해 보아야 한다.

5. 이산칼만필터의 설계 예

본 장에서는 간단한 시스템에 대하여 칼만필터를 구동한 예를 통해 칼만필터의 수행 결과에 대한 이해를 돕고자 한다. 다음과 같이 위치 p 와 속도 s 를 상태변수 $x_k \triangleq \begin{bmatrix} p_k \\ s_k \end{bmatrix}$ 로 하는 시스템이 있다고 가정하자.

실제 시스템

$$x_{k+1} = Fx_k + w_k \tag{5.1.a}$$

$$z_k = Hx_k + v_k \tag{5.1.b}$$

여기서 $x_{k=0} \sim N(\hat{x}_0; P_0)$, $w_k \sim N(0; Q)$, $v_k \sim N(0; R)$ 는 각각 서로 독립인 정규확률프로세스로서 그 평균과 분산은 각각

$$\hat{x}_0 = \begin{bmatrix} 0 \\ 30 \end{bmatrix}, P_0 = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}, R = 100$$

을 가진다. 또, $F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $H = [1 \ 0]$ 이다.

이 시스템에 대한 칼만필터는 단지 (3.4)-(3.5)를 그대로 적용하는 것만으로 손쉽게 구현할 수 있다. 특히 이 경우는 측정치가 스칼라이므로 (3.5.c)의 칼만 이득은 역행렬을 구할 필요 없이 다음과 같이 나눗셈으로 계산할 수 있다.

$$K_{k+1} = \frac{P_{k+1|k} H^T}{HP_{k+1|k} H^T + R} \tag{5.2}$$

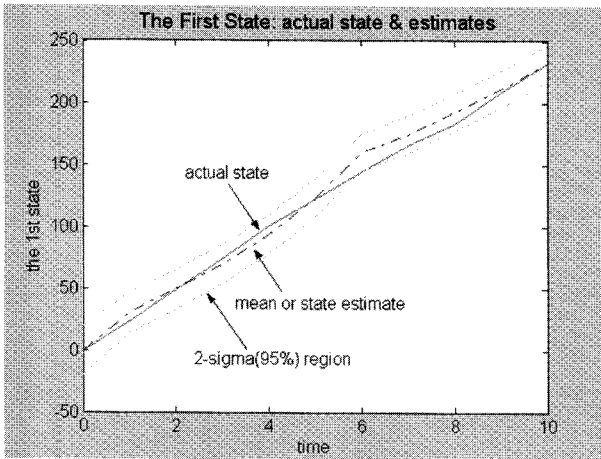
그림 5.1은 이에 대한 시뮬레이션 결과를 보인 것이다. 여기서 유의할 것은 시스템이 확률프로세스인 w_k 에 의하여 구동되고 그 초기값도 역시 확률변수이므로 시뮬레이션 프로그램의 random seed에 따라 그림과 다른 결과가 나올 수 있다는 것이다. 따라서 칼만필터의 출력은 실제 시스템 상태변수가 시간에 따라 어떤 범위 안에 존재하는지를 추정한다는 관점에서 이해되어야 한다.

전술하였듯이 칼만필터의 상태변수 추정치와 공분산은 매 순간 실제 시스템의 상태변수가 존재할 수 있는 분포의 평균과 분산을 의미하므로, 만약 칼만필터 추정치 주변에서 2배의 표준편차 구간을 그린다면 대략 95%의 확률로 실제 상태변수가 그 구간안에 존재할 것이라 판단할 수 있다. 다시 말하면,

$$\text{칼만필터의 출력을 } \hat{x}_{k|k} \triangleq \begin{bmatrix} \hat{p}_{k|k} \\ \hat{s}_{k|k} \end{bmatrix}, P_{k|k} \triangleq \begin{bmatrix} P_{k|k}^{11} & P_{k|k}^{12} \\ P_{k|k}^{21} & P_{k|k}^{22} \end{bmatrix}$$

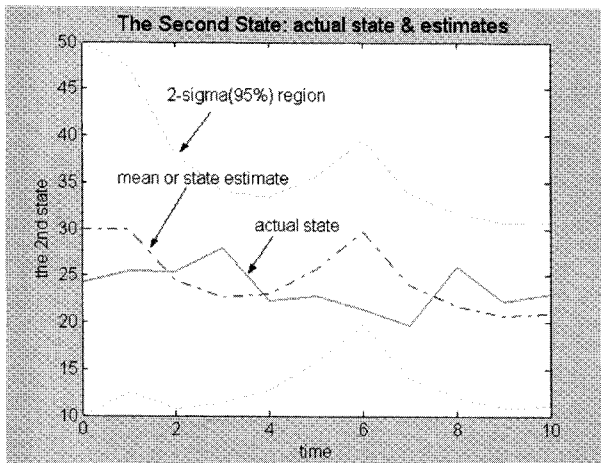
라 할 때, 실제 상태변수는 95%의 확률로 $\hat{p}_{k|k} \pm 2\sqrt{P_{k|k}^{11}}$ 및 $\hat{s}_{k|k} \pm 2\sqrt{P_{k|k}^{22}}$ 의 구간에 있다고 추측할 수 있다.

그림 5.1은 이러한 관점에서 $\hat{p}_{k|k} \pm 2\sqrt{P_{k|k}^{11}}$ 와



(a) 첫 번째 상태변수

($\hat{p}_{k|k}$ -일점쇄선, $\hat{p}_{k|k} \pm 2\sqrt{p_{k|k}^{11}}$ -점선, p_k -실선)



(b) 두 번째 상태변수

($\hat{s}_{k|k}$ -일점쇄선, $\hat{s}_{k|k} \pm 2\sqrt{p_{k|k}^{22}}$ -점선, s_k -실선)

그림 5.1 상태변수 실제값(실선)과 칼만필터 추정치(일점쇄선) 및 95% 존재영역(점선)

$\hat{s}_{k|k} \pm 2\sqrt{p_{k|k}^{22}}$ 의 구간(점선)과 실제 상태변수(실선)를 동시에 도시한 것이다. 여기서 $\hat{p}_{k|k}$, $\hat{s}_{k|k}$, $p_{k|k}^{11}$, $p_{k|k}^{22}$ 는

$$\hat{x}_{k|k} \triangleq \begin{bmatrix} \hat{p}_{k|k} \\ \hat{s}_{k|k} \end{bmatrix}, P_{k|k} \triangleq \begin{bmatrix} p_{k|k}^{11} & p_{k|k}^{12} \\ p_{k|k}^{21} & p_{k|k}^{22} \end{bmatrix}$$

로 정의되는 칼만 필터 추정치 벡터와 오차 공분산 행렬의 원소이다. 그림 5.1을 살펴보면 실선으로 표현된 실제 상태변수는 점선으로 표현된 $\hat{p}_{k|k} \pm 2\sqrt{p_{k|k}^{11}}$ 및 $\hat{s}_{k|k} \pm 2\sqrt{p_{k|k}^{22}}$ 의 구간 안에 존재함을 볼 수 있다.

6. 결론

지난 호에 이어서 본 고에서는 기본적인 정규확률프로세스의 성질을 이용하여 이산 칼만필터를 유도하고, 간단한 예제를 통하여 칼만필터의 추정치의 의미를 살펴보았다. 이산칼만필터는 Gauss-Markov 프로세스로 주어지는 선형확률 시스템 상태변수의 분포를 알려주는 통계적 추정기이며, 그 추정치는 최소오차공분산 또는 최대 사후확률의 관점에서 최적값을 제공한다. 칼만필터는 그 구조가 순환적(recursive)이므로 컴퓨터를 이용한 프로그램이 쉬운 장점이 있으며, 선형시스템으로 모델링되는 많은 문제에 손쉽게 적용하여 우수한 상태변수 추정성능을 얻을 수 있다. 특히 원하는 상태변수를 직접 측정하지 못하는 경우 칼만필터를 적용함으로써 손쉽게 상태변수 관측기(state observer)로 이용할 수 있다. 최근에 팔복할 만한 발전을 이루어 가는 센서 기술과 더불어, 쉽고 편리하게 디지털 시스템에 적용할 수 있는 칼만필터는 보다 많은 실제 문제에 적용될 것으로 기대된다.

참고문헌

- [1] 황익호, 알기쉬운 제어이론 '동적시스템의 상태변수 추정 기법[1]', 제어·로봇·시스템학회지, vol. 15, no. 2, 여름호(6월호), 2009.
- [2] 황익호, 나원상, '베이지안 상태변수 추정기와 다중가설 필터기법', 제어·로봇·시스템학회지, vol. 14, no. 3, 가을호(9월호), 2008.
- [3] C.-T. Chen, *Linear System Theory and Design*, Holt, Rinehart and Winston, 1984.
- [4] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, Academic Press, 1979.
- [5] F. L. Lewis, *Optimal Estimation*, John Wiley & Sons, 1986.

○ 저자 약력



황 익 호

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