**P0004** – Kalman Filter

간단히 예를 들기 위해서, 일차원에서 움직이는 물체에 대한 운동방정식을 모델링 해보기로 하자.

$X\_{t}=\left[\genfrac{}{}{0pt}{}{x\_{t}}{\dot{x}\_{t}}\right]$

물체를 움직이게 하는 제어 입력에 대한 정의는 아래와 같이 할 수 있다.

$u\_{t}=\frac{f\_{t}}{m}$

물체의 모멘트(혹은 질량) $m$에 미치는 힘 $f\_{t}$의 관계로 설명할 수 있다. 이것으로부터 시간 $t$에 해당하는 물체의 위치와 속도에 대한 관계식은 아래와 같이 정의할 수 있다.

 $x\_{t}=x\_{t-1}+\left(\dot{x}\_{t-1}×∆t\right)+\frac{f\_{t}\left(∆t\right)^{2}}{2m}$

 $\dot{x}\_{t}=\dot{x}\_{t-1}+\frac{f\_{t}∆t}{m}$

위의 식을 시스템의 상태 방정식 형태(state space variable form)로 표현하면 아래와 같다.

 $\left[\begin{matrix}x\_{t}\\\dot{x}\_{t}\end{matrix}\right]=\left[\begin{matrix}1&∆t\\0&1\end{matrix}\right]\left[\begin{matrix}x\_{t-1}\\\dot{x}\_{t-1}\end{matrix}\right]+\left[\begin{matrix}\frac{∆t^{2}}{2}\\∆t\end{matrix}\right]u\_{t}$ (1)

 🡺 $X\_{t}=A∙X\_{t-1}+B∙u\_{t}$ (2)

여기서 $X\_{t}$는 우리가 관심을 가지는 시스템의 state vector이고, $u\_{t}$는 control input이다. $A$는 시스템의 state space transition matrix이며, $B$는 control input matrix이다.

그런데, 우리가 $x\_{t}$에 대한 정확한 정보를 알 수 있다면 좋겠지만 경우에 따라서는 정확한 정보를 알 수가 없어서 센서 등의 계측기로부터 정보를 얻어 $x\_{t}$에 대한 추정estimation을 해야 한다. $x\_{t}$에 대한 추정 값(estimation value) $\hat{X}\_{t}$는 $X\_{t}$와 마찬가지의 시스템 상태 방정식을 가진다.

$\hat{X}\_{t}=A∙\hat{X}\_{t-1}+B∙u\_{t}$

예측 정보인 $\hat{x}\_{t}$는 $x\_{t}$와 오차를 가질 수 밖에 없다. 이 오차는 매 순간 t마다 일정 분포 안에 포함되는 값으로 표현할 수 있다. 이 오차는 확률 모형으로 정의될 수 있는데, 가우시안 분포를 가지는 이 확률 모형은 평균값과 분산을 가지게 된다. 추정 오차에 대한 공분산 기대 값을 아래와 같이 정의하자.

 $P\_{t}^{-}=E\left[\left(x\_{t}-\hat{x}\_{t}^{-}\right)\left(x\_{t}-\hat{x}\_{t}^{-}\right)^{T}\right]$

$P\_{t}^{-}$는 시간 t에서의 추정 오차 공분산에 대한 기대 값 matrix이다. $\hat{x}\_{t}^{-}$는 시간 t에 추정 오차에 대한 보상을 하기 전의 값이다. 시스템의 상태 값을 추정하는 데 있어 일정 수준의 오차가 존재함을 알았다. 이제 이를 기반으로 시스템 상태 방정식을 다시 정의하면 아래와 같다.

$X\_{t}=A∙X\_{t-1}+B∙u\_{t}+w\_{t}$

여기서 $w\_{t}$는 시스템 모델에 대한 에러이며, 프로세스 노이즈process noise라고 한다. 시간 t-1와 t 사이에 발생하는 오차는 아래와 같이 정의된다.

$w\_{t}\~N\left(0, σ\_{w}\right)$

위의 뜻은 평균이 0이며 분산이 $σ\_{w}$인 확률 밀도 함수Probability Density Function의 값이다.

시간 t에서의 상태와 추정 오차가 포함된 상태 사이의 오차는

 $x\_{t}-\hat{x}\_{t}^{-}$ $=\left(A∙x\_{t}+B∙u\_{t}+w\_{t}\right)-\left(A∙\hat{x}\_{t}+B∙u\_{t}\right)$

 $=A∙\left(x\_{t}-\hat{x}\_{t}\right)+w\_{t}$

이므로,

 $P\_{t}^{-}$ $=E\left[\left(x\_{t}-\hat{x}\_{t}^{-}\right)\left(x\_{t}-\hat{x}\_{t}^{-}\right)^{T}\right]$

 $=E\left[\left(A∙\left(x\_{t-1}-\hat{x}\_{t-1}\right)+w\_{t}\right)\left(A∙\left(x\_{t-1}-\hat{x}\_{t-1}\right)+w\_{t}\right)^{T}\right]$

$=A∙E\left[\left(\left(x\_{t-1}-\hat{x}\_{t-1}\right)\left(x\_{t-1}-\hat{x}\_{t-1}\right)^{T}\right)\right]∙A^{T}$

$+A∙E\left[\left(\left(x\_{t-1}-\hat{x}\_{t-1}\right)∙w\_{t}^{T}\right)\right]+E\left[\left(w\_{t}∙\left(x\_{t-1}-\hat{x}\_{t-1}\right)^{T}\right)\right]∙A^{T}+E\left[w\_{t}∙w\_{t}^{T}\right]$

여기서 추정 에러인 $x\_{t-1}-\hat{x}\_{t-1}$와 프로세스 노이즈인 $w\_{t}$은 uncorrelated 하므로

$A∙E\left[\left(\left(x\_{t-1}-\hat{x}\_{t-1}\right)∙w\_{t}^{T}\right)\right]=E\left[\left(w\_{t}∙\left(x\_{t-1}-\hat{x}\_{t-1}\right)^{T}\right)\right]∙A^{T}=0$

이 된다. 따라서,

$P\_{t}^{-}=A∙E\left[\left(\left(x\_{t-1}-\hat{x}\_{t-1}\right)\left(x\_{t-1}-\hat{x}\_{t-1}\right)^{T}\right)\right]∙A^{T}+E\left[w\_{t}∙w\_{t}^{T}\right]$

의 식으로 정리할 수 있다. 위의 식을 간단하게 표현하면 아래와 같다.

$P\_{t}^{-}=A∙P\_{t-1}∙A^{T}+Q\_{t}$ (3)

이상으로 시스템 모델에 대한 추정으로부터 오차 매트릭스를 정리했다. 그런데 이 오차는 어떤 식으로 표현을 해야 할까? 상태 $x\_{t}$에 대한 추정에 대한 불확실성을 표현하는 방법으로 가장 간단하면서 강력한 방법 중 하나는 Gaussian Function으로 정의하는 것이다. 상태 추정 값 $\hat{x}\_{t}$는 평균 값mean과 분산variance를 가지는 확률 밀도 함수Probability Density Function로 정의할 수 있다. 바꾸어 말하면 상태 $x\_{t}$에 대한 추정은 평균 값 $\hat{x}\_{t}$를 가지고 이를 중심으로 일정 분산의 모델링 오차가 합해진 형태로 표현할 수 있다는 것이다.

추정 오차에 대해 좀 더 알아보도록 하자.

실제 상태인 $x\_{t}$가 존재하는 위치의 추정은 Gaussian PDF를 사용하여 표현할 수 있다고 했다. 이럴 경우 $x\_{t}$의 위치가 $r$에 위치할 확률을 표현하면 아래와 같다.

 $y\left(r; μ, σ\right)=\frac{1}{\sqrt{2πσ^{2}}}e^{-\frac{\left(r-μ\right)^{2}}{2σ^{2}}}$

여기서 함수 $y$는 $μ$인 평균Mean을 가지고 $σ$인 분산Variance을 가질 때, $x\_{t}$의 위치가 $r$에 위치할 확률이다. 이런 식으로 Gaussian PDF를 활용하여 모델 오차뿐만이 아니라 센서가 가지고 있는 측정 오차를 표현할 수가 있다. 따라서 칼만 필터의 연산 단계에서 필요한 모델 오차와 측정 오차를 위의 개념을 이용하여 Gaussian PDF의 합성을 통해 최종 추정 식을 유도하게 된다.

Gaussian PDF의 특성 중 하나는 서로 다른 두 개의 Gaussian PDF의 곱은 새로운 하나의 Gaussian PDF로 만들어진다는 것이다. 우리는 이런 특성을 이용해 복잡도를 증가시키지 않고 매 상태State에 대한 최종 추정 값을 계산할 수가 있다.

[Gaussian PDF 합성에 관한 그림]

상태 모델에 대한 Gaussian PDF의 식을 아래와 같이 정의하자.

$y\_{1}\left(r; μ\_{1}, σ\_{1}\right)=\frac{1}{\sqrt{2πσ\_{1}^{2}}}e^{-\frac{\left(r-μ\_{1}\right)^{2}}{2σ\_{1}^{2}}}$

마찬가지로 상태 측정에 대한 Gaussian PDF의 식을 아래와 같이 정의하자.

$y\_{2}\left(r; μ\_{2}, σ\_{2}\right)=\frac{1}{\sqrt{2πσ\_{2}^{2}}}e^{-\frac{\left(r-μ\_{2}\right)^{2}}{2σ\_{2}^{2}}}$

위의 두 개의 Gaussian PDF의 곱을 통해 PDF가 합성됨으로써 새로운 Gaussian PDF의 식을 얻을 수 있다.

$y\_{fused}\left(r; μ\_{1}, σ\_{1}, μ\_{2}, σ\_{2}\right)$ $=\frac{1}{\sqrt{2πσ\_{1}^{2}}}e^{-\frac{\left(r-μ\_{1}\right)^{2}}{2σ\_{1}^{2}}}×\frac{1}{\sqrt{2πσ\_{2}^{2}}}e^{-\frac{\left(r-μ\_{2}\right)^{2}}{2σ\_{2}^{2}}}$

 $=\frac{1}{2π\sqrt{σ\_{1}^{2}σ\_{2}^{2}}}e^{-\left\{\frac{\left(r-μ\_{1}\right)^{2}}{2σ\_{1}^{2}}+\frac{\left(r-μ\_{2}\right)^{2}}{2σ\_{2}^{2}}\right\}}$

위의 식을 다시 정리하면,

$y\_{fused}\left(r; μ\_{fused}, σ\_{fused}\right)=\frac{1}{\sqrt{2πσ\_{fused}^{2}}}e^{-\frac{\left(r-μ\_{fused}\right)^{2}}{2σ\_{fused}^{2}}}$

여기서 $μ\_{fused}$와 $σ\_{fused}^{2}$는 다음과 같이 정리할 수 있다.

 $μ\_{fused}$ $=\frac{μ\_{1}σ\_{1}^{2}+μ\_{2}σ\_{2}^{2}}{σ\_{1}^{2}+σ\_{2}^{2}}$

 $=μ\_{1}+\frac{σ\_{1}^{2}\left(μ\_{2}-μ\_{1}\right)}{σ\_{1}^{2}+σ\_{2}^{2}}$

$σ\_{fused}^{2}$ $=\frac{σ\_{1}^{2}σ\_{2}^{2}}{σ\_{1}^{2}+σ\_{2}^{2}}$

 $=σ\_{1}^{2}-\frac{σ\_{1}^{4}}{σ\_{1}^{2}+σ\_{2}^{2}}$

위의 두 식은 칼만 필터 알고리즘 상에서 업데이트 부분에 있는 수식이다. 그러나 위의 식은 프로세스 노이즈 및 측정 노이즈가 같은 스케일에 있다고 가정한 전개이다. 하지만 현실에서는 꼭 그러지만은 않을 수도 있기 때문에 스케일링을 해줄 필요가 있다. 이를 위해 프로세스 노이즈 쪽에 임의의 상수 c로 나누어 측정 노이즈에 맞추어 주도록 한다.

$y\_{1}\left(r; μ\_{1}, σ\_{1}, c\right)=\frac{1}{\sqrt{2π\left(\frac{σ\_{1}}{c}\right)^{2}}}e^{-\frac{\left(r-\frac{μ\_{1}}{c}\right)^{2}}{2\left(\frac{σ\_{1}}{c}\right)^{2}}}$

$y\_{2}\left(r; μ\_{2}, σ\_{2}\right)=\frac{1}{\sqrt{2πσ\_{2}^{2}}}e^{-\frac{\left(r-μ\_{2}\right)^{2}}{2σ\_{2}^{2}}}$

위의 수식을 통해 다시

$\frac{μ\_{fused}}{c}$ $=\frac{μ\_{1}}{c}+\frac{\left(\frac{σ\_{1}}{c}\right)^{2}\left(μ\_{2}-\frac{μ\_{1}}{c}\right)}{\left(\frac{σ\_{1}}{c}\right)^{2}+σ\_{2}^{2}}$

🡺$μ\_{fused}$ $=μ\_{1}+\left\{\frac{\left(\frac{σ\_{1}^{2}}{c}\right)}{\left(\frac{σ\_{1}}{c}\right)^{2}+σ\_{2}^{2}}\right\}∙\left(μ\_{2}-\frac{μ\_{1}}{c}\right)$

위의 식을 단순화 하기 위해 $H={1}/{c}$ 그리고 $K={\left(Hσ\_{1}^{2}\right)}/{\left(H^{2}σ\_{1}^{2}+σ\_{2}^{2}\right)}$으로 치환을 하여 아래와 같이 정리한다.

 $μ\_{fused}=μ\_{1}+K∙\left(μ\_{2}-Hμ\_{1}\right)$

마찬가지로 분산의 합성에 대해서도 정리를 하면,

$\frac{σ\_{fused}^{2}}{c^{2}}$ $=\left(\frac{σ\_{1}}{c}\right)^{2}-\frac{\left(\frac{σ\_{1}}{c}\right)^{4}}{\left(\frac{σ\_{1}}{c}\right)^{2}+σ\_{2}^{2}}$

🡺 $σ\_{fused}^{2}$ $=σ\_{1}^{2}-\left\{\frac{\frac{σ\_{1}^{2}}{c}}{\left(\frac{σ\_{1}}{c}\right)^{2}+σ\_{2}^{2}}\right\}∙\frac{σ\_{1}^{2}}{c}$

 $σ\_{fused}^{2}$ $=σ\_{1}^{2}-KHσ\_{1}^{2}$

이제까지의 수식 정리를 바탕으로 칼만 필터의 오차 공분산 업데이트에 관한 정리를 이해할 수 있다.

칼만 필터에서는 업데이트와 보정의 과정을 순차적으로 거쳐 우리가 알고 싶어하는 상태의 추정 값을 계산해 낸다.

칼만 필터를 실제 적용하는 환경은 discrete domain에서 구현을 하므로 수식은 discrete domain 기준으로 시간 $t$를 시점 $k$로 변경해서 정리하면 다음과 같다.

Prediction (모델 업데이트)

1. Project the state ahead

$\hat{x}\_{k}^{-}=A\hat{x}\_{k-1}^{-}+Bu\_{k}$

1. Project the error covariance ahead

$P\_{k}^{-}=AP\_{k-1}A^{T}+Q$

Correction (측정 업데이트)

1. Compute the kalman gain

$K\_{k}=P\_{k}^{-}H^{T}\left(HP\_{k}^{-}H^{T}+R\right)^{-1}$

1. Update estimate with measurement $z\_{k}$

$\hat{x}\_{k}=\hat{x}\_{k}^{-}+K\_{k}\left(z\_{k}-H\hat{x}\_{k}^{-}\right)$

1. Update the error covariance

$P\_{k}=\left(1-K\_{k}H\right)P\_{k}^{-}$

위의 Prediction과 Correction 과정이 순차적으로 연산이 되면서 우리가 알고 싶어하는 상태에 대한 추정 값 $\hat{x}\_{k}$를 얻을 수 있다. 이 때 $k=1, 2, 3, …, N$ 이다.

칼만 필터의 내용을 좀 더 깊게 이해하기 위해 다음과 같은 내용을 참고하는 것을 추천한다.

* 공분산Covariance
* 공분산 행렬Covariance Matrix