Alternate Kalman Filter

Kalman Filter를 다시 정리해보면 다음과 같다.
우선 다음과 같은 Discrete-Time Linear System이 주어진다.

$$x_k = F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1}$$
$$y_k = H_k x_k + v_k$$

$$w_k \sim (0, Q_k)$$
$$v_k \sim (0, R_k)$$
$$E(w_k w_j^T) = Q_k \delta_{k-j}$$
$$E(v_k v_j^T) = R_k \delta_{k-j}$$
$$E(v_k w_j^T) = 0$$

이 때, 칼만 필터는 다음과 같다.

$$\hat{x}_k^- = F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1}$$
$$P_k^- = F_{k-1}P_{k-1}^+ F_{k-1}^T + Q_{k-1}$$
$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1}$$
$$\hat{x}_k^+ = \hat{x}_k^- + K_k(y_k - H_k \hat{x}_k^-)$$
$$P_k^+ = P_k^- - K_k H_k P_k^-$$

1. Sequential Kalman Filter

위와 같은 칼만 필터에서 y_k가 i개의 요소로 이루어진 벡터라고 생각해보자. 그에 따라 y_k의 노이즈인 v_k의 분산은 $r \times r$의 diagonal 행렬로 다음과 같다.

$$R_k = \begin{bmatrix} R_{1k} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & R_{rk} \end{bmatrix}$$

따라서 measurement y_k는 다음과 같이 구성된다.

$$y_{ik} = H_{ik}x_k + v_{ik}$$
$$v_{ik} \sim (0, R_{ik})$$

해당 문제에 대해서 원래 칼만 필터를 적용하듯 순차적으로 r개의 관측값에 따라 각각 칼만 필터를 적용시켜준다.
우선 첫 번째 상태를 다음과 같이 정의한다.

$$\hat{x}_{0k}^+ = \hat{x}_k^-$$
$$P_{0k}^+ = P_k^-$$

이에 대해 $i = 1, \dots, r$에 대해 순차적으로 필터를 적용한다.

$$K_{ik} = P_{i-1}^- H_{ik}^T (H_{ik} P_{i-1}^- H_{ik}^T + R_{ik})^{-1}$$
$$\hat{x}_{ik}^+ = \hat{x}_{i-1,k}^- + K_{ik}(y_{ik} - H_{ik}\hat{x}_{i-1,k}^-)$$
$$P_{ik}^+ = P_{i-1,k}^- - K_{ik}H_{ik}P_{i-1,k}^-$$

그리고 최종적인 상태는 다음과 같이 된다.

$$\hat{x}_k^+ = \hat{x}_{rk}^+$$
$$P_k^+ = P_{rk}^+$$

만약 R_k가 diagonal matrix가 아니라면 다음과 같이 치환하여 해결할 수 있다.

$R_k = R$이 symmetric하며 positive definite하다면, eigenvalue decomposition을 통해 $R = S\hat{R}S^{-1}$로 분해할 수 있다.

그리고 measurement를 다음과 같이 변환하여 생각할 수 있다.

$$\tilde{y}_k = S^{-1}y_k = S^{-1}(H_k x_k + v_k) = \tilde{H}_k x_k + \tilde{v}_k$$

다음과 같이 변환된 measurement의 noise는 아래와 같이 diagonal한 분산을 가지게 된다.

$$\begin{aligned} E(\tilde{v}_k \tilde{v}_k^T) &= E(S^{-1}v_k v_k^T S^{-T}) \\ &= E(S^{-1}v_k v_k^T S) \\ &= S^{-1}RS \\ &= \hat{R}\,(diagonal) \end{aligned}$$

위 과정에 따라 diagonal한 R이 되었으니 앞의 식을 동일하게 적용할 수 있게 된다.

2. Information Filter (중요)

Information Filter는 Kalman Filter의 변형으로 계산 상의 편의와 안정성을 가져다 주는 방법이다.

다음과 같이 Information matrix를 정의한다.

$$I = P^{-1}$$

해당 matrix는 정보가 많으면 커지고 정보가 없으면 작아지는 metric으로 $I \to 0, P \to \infty$이면 정보가 없다는 뜻이고 $I \to \infty, P \to 0$이면 완벽한 정보를 가진다는 뜻이다.

그에 따라 칼만 필터 과정을 Information matrix를 이용해 표현해보면 다음과 같다.

여기에는 다음과 같은 Matrix Inversion Lemma가 사용된다.

Matrix Inversion Lemma

$$\begin{aligned} (A + BC^{-1}D)^{-1} &= A^{-1} - A^{-1}B(C + DA^{-1}B)^{-1}DA^{-1} \\ (A + BC^{-1}D)^{-1}BC^{-1} &= A^{-1}B(C + DA^{-1}B)^{-1} \end{aligned}$$

$$\begin{aligned} I_k^- &= (P_k^-)^{-1} \\ &= (F_{k-1}(I_{k-1}^+)^{-1}F_{k-1}^T + Q_{k-1})^{-1} \\ &= Q_{k-1}^{-1} - Q_{k-1}^{-1}F_{k-1}(I_{k-1}^+ + F_{k-1}^T Q_{k-1}^{-1} F_{k-1})^{-1} F_{k-1}^T Q_{k-1}^{-1} \\ I_k^+ &= (P_k^+)^{-1} \\ &= (P_k^-)^{-1} + H_k^T R_k^{-1} H_k \\ &= I_k^- + H_k^T R_k^{-1} H_k \end{aligned}$$

정리하면 다음과 같다.

Information Filter (1) $\mathcal{I} = P^{-1}$

$$\begin{aligned} \mathcal{I}_k^- &= Q_{k-1}^{-1} - Q_{k-1}^{-1}F_{k-1}(\mathcal{I}_{k-1}^+ + F_{k-1}^T Q_{k-1}^{-1}F_{k-1})^{-1}F_{k-1}^T Q_{k-1}^{-1} \\ \mathcal{I}_k^+ &= \mathcal{I}_k^- + H_k^T R_k^{-1} H_k \\ K_k &= (\mathcal{I}_k^+)^{-1} H_k^T R_k^{-1} \\ \hat{x}_k^- &= F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1} \\ \hat{x}_k^+ &= \hat{x}_k^- + K_k(y_k - H_k\hat{x}_k^-) \end{aligned}$$

여기에 추가적으로 상태값인 $\hat{x}$도 $\hat{z}=P^{-1}\hat{x}$로 치환할 수 있다. 예시로 가우시안 pdf를 치환해보면 다음과 같다. 우선 원래의 함수는 다음과 같다.

$$p(x|\mu,\Sigma)=\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$

여기서 $I=\Sigma^{-1}$, $z=\Sigma^{-1}\mu$로 치환하면 다음과 같다.

$$p(x|z,I)=\exp\left(-\frac{1}{2}\left(n\log(2\pi)-\log|I|+z^TIz\right)+z^Tx-\frac{1}{2}x^TIx\right)$$

이와 같이 칼만 필터에서 $\hat{x}$를 $\hat{z}$로 치환하면 다음과 같다. 여기서도 위의 Matrix Inversion Lemma가 사용된다.

$$\begin{aligned}
\hat{z}_k^- &= (P_k^-)^{-1}\hat{x}_k^- \\
&= (P_k^-)^{-1}F_{k-1}\hat{x}_{k-1}^+ \\
&= \left(F_{k-1}P_{k-1}^+F_{k-1}^T+Q_{k-1}\right)^{-1}F_{k-1}P_{k-1}^+\hat{z}_{k-1}^+ \\
&= Q_{k-1}^{-1}F_{k-1}\left((P_{k-1}^+)^{-1}+F_{k-1}^TQ_{k-1}^{-1}F_{k-1}\right)^{-1}\hat{z}_{k-1}^+ \\
&= Q_{k-1}^{-1}F_{k-1}\left(I_{k-1}^++F_{k-1}^TQ_{k-1}^{-1}F_{k-1}\right)^{-1}\hat{z}_{k-1}^+ \\
\hat{z}_k^+ &= (P_k^+)^{-1}\hat{x}_k^+ \\
&= (P_k^+)^{-1}\left(\hat{x}_k^-+P_k^+H_k^TR_k^{-1}(y_k-H_k\hat{x}_k^-)\right) \\
&= \left((P_k^+)^{-1}-H_k^TR_k^{-1}H_k\right)\hat{x}_k^-+H_k^TR_k^{-1}y_k \\
&= \left((P_k^+)^{-1}-H_k^TR_k^{-1}H_k\right)P_k^-\hat{z}_k^-+H_k^TR_k^{-1}y_k \\
&= \left((P_k^-)^{-1}+H_k^TR_k^{-1}H_k-H_k^TR_k^{-1}H_k\right)P_k^-\hat{z}_k^-+H_k^TR_k^{-1}y_k \\
&= \hat{z}_k^-+H_k^TR_k^{-1}y_k
\end{aligned}$$

정리하면 다음과 같다.

Kalman Filter

$$\begin{aligned}
\hat{x}_k^- &= F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1} \\
P_k^- &= F_{k-1}P_{k-1}^+F_{k-1}^T + Q_{k-1} \\
\hat{x}_k^+ &= \hat{x}_k^- + P_k^-H_k^T(H_kP_k^-H_k^T+R_k)^{-1}(y_k-H_k\hat{x}_k^-) \\
P_k^+ &= P_k^- - P_k^-H_k^T(H_kP_k^-H_k^T+R_k)^{-1}H_kP_k^-
\end{aligned}$$

r x r matrix inversion

Information Filter (2) $\mathcal{I}=P^{-1}$, $\hat{z}=P^{-1}\hat{x}$, and $G_k=0$

$$\begin{aligned}
\hat{z}_k^- &= Q_{k-1}^{-1}F_{k-1}(\mathcal{I}_{k-1}^+ + F_{k-1}^TQ_{k-1}^{-1}F_{k-1})^{-1}\hat{z}_{k-1}^+ \\
\mathcal{I}_k^- &= Q_{k-1}^{-1} - Q_{k-1}^{-1}F_{k-1}(\mathcal{I}_{k-1}^+ + F_{k-1}^TQ_{k-1}^{-1}F_{k-1})^{-1}F_{k-1}^TQ_{k-1}^{-1} \\
\hat{z}_k^+ &= \hat{z}_k^- + H_k^TR_k^{-1}y_k \\
\mathcal{I}_k^+ &= \mathcal{I}_k^- + H_k^TR_k^{-1}H_k
\end{aligned}$$

n x n matrix inversion

위와 같이 measurement가 state보다 많을 때, 즉 n<<r일 때에는 Information Filter가 계산상 더 효율적이다. 하지만 여전히 R_k의 역함수를 계산해야 한다는 점에서 이의 계산량을 고려해야 한다.

만약 $P_0^+=\infty$일 때 Kalman filter는 돌리지 못하지만 $I_0^+=0$이므로 information filter는 동작할 수 있다. 반대로 $I_0^+=\infty$일 때는 information filter는 돌릴 수 없지만 $P_0^+=0$이기 때문에 Kalman filter는 동작한다.

3. Square Root Filter

만약 P가 symmetric하고 positive definite하면 Cholesky decomposition을 통해 $P = SS^T$로 분해할 수 있다. 여기서 S를 P의 square root라고 한다.

1) Dynamic Update

우선 dynamic update단계부터 살펴보자. 먼저 다음 조건을 만족하는 orthogonal한 $2n \times 2n$의 matrix T가 있다고 가정하자.

$$\begin{bmatrix} (S_k^-)^T \\ 0 \end{bmatrix} = T \begin{bmatrix} (S_{k-1}^+)^T F_{k-1}^T \\ Q_{k-1}^{T/2} \end{bmatrix} = \begin{bmatrix} T_1 & T_2 \end{bmatrix} \begin{bmatrix} (S_{k-1}^+)^T F_{k-1}^T \\ Q_{k-1}^{T/2} \end{bmatrix}$$

$$T_1^T T_2 = T_2^T T_1 = 0$$
$$T_1^T T_1 = T_2^T T_2 = I$$

그리고 양변에 해당 식을 transpose한 것을 각각 곱하면 다음과 같다.

$$\begin{bmatrix} S_k^- & 0 \end{bmatrix} \begin{bmatrix} (S_k^-)^T \\ 0 \end{bmatrix} = \begin{bmatrix} T_1 (S_{k-1}^+)^T F_{k-1}^T + T_2 Q_{k-1}^{T/2} \end{bmatrix}^T [\ldots]$$

$$\begin{aligned} S_k^- (S_k^-)^T &= F_{k-1} S_{k-1}^+ \, T_1^T T_1 (S_{k-1}^+)^T F_{k-1}^T + Q_{k-1}^{1/2} \, T_2^T T_2 Q_{k-1}^{T/2} \\ &= F_{k-1} S_{k-1}^+ (S_{k-1}^+)^T F_{k-1}^T + Q_{k-1}^{1/2} \, Q_{k-1}^{T/2} \\ P_k^- &= F_{k-1} P_{k-1}^+ F_{k-1}^T + Q_{k-1} \end{aligned}$$

이와 같이 전개하면 기존의 dynamic update 식이 나오게 된다. 따라서 처음에 가정했던 다음의 식에서 $T^T T = I$를 만족하는 T를 찾으면 그에 따라 dynamic update가 가능하다.

$$T \begin{bmatrix} (S_{k-1}^+)^T F_{k-1}^T \\ Q_{k-1}^{T/2} \end{bmatrix} = \begin{bmatrix} n \times nmatrix \\ 0 \end{bmatrix} = \begin{bmatrix} (S_k^-)^T \\ 0 \end{bmatrix}$$

2) Measurement Update

앞서 진행했던 것처럼 $i = 1, \ldots, r$인 sequential kalman filter를 가정하자.

그리고 초기 값은 $P_{0k}^+ = P_k^-$로 설정한다. 그 이후 다음과 같은 sequential 과정을 거친다.

$$P_{i-1,k}^+ = S_{i-1,k}^+ S_{i-1,k}^{+\,T}$$

$$\begin{aligned} K_{ik} &= \frac{P_{i-1,k}^+ H_{ik}^T}{H_{ik} P_{i-1,k}^+ H_{ik}^T + R_{ik}} = \frac{S_{i-1,k}^+ S_{i-1,k}^{+\,T} H_{ik}^T}{H_{ik} S_{i-1,k}^+ S_{i-1,k}^{+\,T} H_{ik}^T + R_{ik}} \\ P_{ik}^+ &= (I - K_{ik} H_{ik}) P_{i-1,k}^+ \\ &= \left(I - \frac{S_{i-1,k}^+ S_{i-1,k}^{+\,T} H_{ik}^T H_{ik}}{H_{ik} S_{i-1,k}^+ S_{i-1,k}^{+\,T} H_{ik}^T + R_{ik}} \right) S_{i-1,k}^+ S_{i-1,k}^{+\,T} \\ &= S_{i-1,k}^+ (I - a\phi\phi^T) S_{i-1,k}^{+\,T} \end{aligned}$$

위 식에서 ϕ와 a는 다음과 같은 값을 가진다.

$$\phi = S_{i-1,k}^{+\,T} H_{ik}^T$$

$$a = \frac{1}{\phi^T \phi + R_{ik}}$$

또한 다음의 식이 성립한다.

$$I - a\phi\phi^T = (I - a\gamma\phi\phi^T)^2, \; where \;\; \gamma = \frac{1}{1 \pm \sqrt{aR_{ik}}}$$

이를 위 식에 다시 대입하면 다음과 같다.

$$
\begin{aligned}
P_{ik}^{+} &= S_{ik}^{+} S_{ik}^{+\,T} \\
&= S_{i-1,k}^{+} (I - a\phi\phi^{T}) S_{i-1,k}^{+\,T} \\
&= S_{i-1,k}^{+} (I - a\gamma\phi\phi^{T})^{2} S_{i-1,k}^{+\,T}
\end{aligned}
$$

$$
S_{ik}^{+} = S_{i-1,k}^{+} (I - a\gamma\phi\phi^{T})
$$

따라서 다음과 같이 square root값을 update할 수 있다.

정리하면 다음과 같은 순서로 measurement update를 진행할 수 있다.

1. After the *a priori* covariance square root S_k^- and the *a priori* state estimate $\hat{x}_k^-$ have been computed, initialize

$$
\begin{aligned}
\hat{x}_{0k}^{+} &= \hat{x}_k^{-} \\
S_{0k}^{+} &= S_k^{-}
\end{aligned}
\tag{6.73}
$$

2. For $i = 1, \cdots, r$ (where r is the number of measurements), perform the following.
   (a) Define H_{ik} as the ith row of H_k, y_{ik} as the ith element of y_k, and R_{ik} as the variance of the ith measurement (assuming that R_k is diagonal).
   (b) Perform the following to find the square root of the covariance after the ith measurement has been processed:

$$
\begin{aligned}
\phi_i &= S_{i-1,k}^{+T} H_{ik}^{T} \\
a_i &= \frac{1}{\phi_i^T \phi_i + R_{ik}} \\
\gamma_i &= \frac{1}{1 \pm \sqrt{a_i R_{ik}}} \\
S_{ik}^{+} &= S_{i-1,k}^{+} (I - a_i \gamma_i \phi_i \phi_i^T)
\end{aligned}
\tag{6.74}
$$

   (c) Compute the Kalman gain for the ith measurement as

$$
\boxed{K_{ik} = a_i S_{i-1,k}^{+} \phi_i} \tag{6.75}
$$

   (d) Compute the state estimate update due to the ith measurement as

$$
\hat{x}_{ik}^{+} = \hat{x}_{i-1,k}^{+} + K_{ik}(y_{ik} - H_{ik}\hat{x}_{i-1,k}^{+}) \tag{6.76}
$$

3. Set the *a posteriori* covariance square root and the *a posteriori* state estimate as

$$
\begin{aligned}
S_k^{+} &= S_{rk}^{+} \\
\hat{x}_k^{+} &= \hat{x}_{rk}^{+}
\end{aligned}
\tag{6.77}
$$

4. U-D Factorization

symmetric하고 positive definite한 $n \times n$ matrix P는 다음과 같이 $P = UDU^T$로 U-D Factorization이 가능하다.

$$
\begin{aligned}
\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{bmatrix}
&= \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 \\ u_{12} & 1 & 0 \\ u_{13} & u_{23} & 1 \end{bmatrix} \\
&= \begin{bmatrix} d_{11} + d_{22}u_{12}^2 + d_{33}u_{13}^2 & d_{22}u_{12} + d_{33}u_{13}u_{23} & d_{33}u_{13} \\ d_{22}u_{12} + d_{33}u_{13}u_{23} & d_{22} + d_{33}u_{23}^2 & d_{33}u_{23} \\ d_{33}u_{13} & d_{33}u_{23} & d_{33} \end{bmatrix}
\end{aligned}
$$

위 식에서 u_{ij}와 d_{ij}만 구하면 된다.

1) Dynamic Update

우선 P^-는 다음과 표현이 가능하다.

$$
\begin{aligned}
P^- &= FP^+F^T + Q \\
&= \begin{bmatrix} FU^+ & I \end{bmatrix} \begin{bmatrix} D^+ & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} U^{+T}F^T \\ I \end{bmatrix} \\
&= W\hat{D}W^T
\end{aligned}
$$

우리는 여기서 $P^- = U^-D^-U^{-T} = W\hat{D}W^T$와 같은 대각화 표현이 필요하다.

이를 위해 $v_k\hat{D}v_j^T = 0 \ (k \neq j)$가 되도록 Gram-schmit를 이용해 $W = U^-V$로 나누어준다.

그러면 다음과 같이 계산이 가능하다.

$$
\begin{aligned}
W\hat{D}W^T &= (U^-V)\hat{D}(U^-V)^T \\
&= U^-(V\hat{D}V)U^{-T} \\
&= U^-D^-U^{-T}
\end{aligned}
$$

2) Measurement Update

P는 Kalman Filter에 의해 다음과 같이 계산된다.

$$
P_i = P_{i-1} - P_{i-1}H_i^T(H_iP_{i-1}H_i^T + R_i)^{-1}H_iP_{i-1}
$$

여기서 각각 $P_{i-1} = U_{i-1}D_{i-1}U_{i-1}^T$, $P_i = U_iD_iU_i^T$로 대각화하고 $\alpha_i = H_iP_{i-1}H_i^T + R_i$로 치환하면 다음과 같다.

$$
\begin{aligned}
U_iD_iU_i^T &= U_{i-1}D_{i-1}U_{i-1}^T - \frac{1}{\alpha_i}U_{i-1}D_{i-1}U_{i-1}^TH_i^TH_iU_{i-1}D_{i-1}U_{i-1}^T \\
&= U_{i-1}\left(D_{i-1} - \frac{1}{\alpha_i}\left(D_{i-1}U_{i-1}^TH_i^T\right)\left(D_{i-1}U_{i-1}^TH_i^T\right)^T\right)U_{i-1}^T
\end{aligned}
$$

여기서

$$
\overline{U}\overline{D}\overline{U}^T = \left(D_{i-1} - \frac{1}{\alpha_i}\left(D_{i-1}U_{i-1}^TH_i^T\right)\left(D_{i-1}U_{i-1}^TH_i^T\right)^T\right)
$$

로 치환하게 된다면, 다음과 같이 U_i와 D_i를 구할 수 있다.

$$
\begin{aligned}
U_iD_iU_i^T &= U_{i-1}\left(\overline{U}\overline{D}\overline{U}^T\right)U_{i-1}^T \\
&= \left(U_{i-1}\overline{U}\right)\overline{D}\left(U_{i-1}\overline{U}\right)^T
\end{aligned}
$$