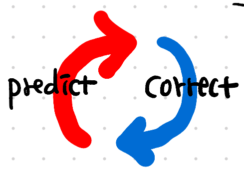
2024.09.29

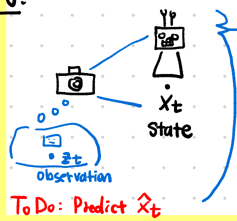
⟨Kalman Filter⟩

special case of a Bayes' filter w/ dynamics and sensory models

: optimal recursive estimator. (If all noise is Gaussian, KF minimizes MSE of estimated parameters)

: one way for tracking.





→ $t+1$
→ state changes
→ new observation is obtained.
→ Prediction also should be updated.

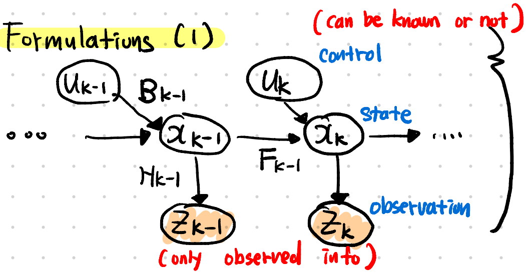
- **Key assumptions**: (1) Only the immediate past matters

$$P(X_i \mid X_1, \dots X_{i-1}) = P(X_i \mid X_{i-1})$$

(2) Measurement depends only on current state.

$$P(z_i, z_j, \dots z_k \mid X_i) = P(z_i \mid X_i)\, P(z_j, \dots, z_k \mid X_i)$$

- **Formulations (1)**



F_k : State-transition model
H_k : observation model
Q_k : Covariance of process noise
R_k : covariance of observation noise.

"True state" at time k evolved from $k-1$ is : $x_k = F_k x_{k-1} + B_k U_k + \underline{W_k}$ process noise $W_k \sim N(0, Q_k)$

"Observation" (or measurement) is made according to : $z_k = H_k x_k + \underline{V_k}$ observation noise $V_k \sim N(0, R_k)$

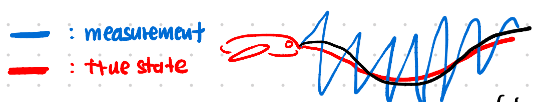
- **Simple example** : airplane altitude.

$$\text{altitude}_k = 0.98 \times \text{altitude}_{k-1} + \text{Control}_k + \text{turbulence}_k$$

$$\text{measured altitude}_k = \text{altitude}_k + \text{noise}_k.$$

Possible Question : why don't we just use measurement as it is?

⇒



goal : predict state from measurement

(+ z_k can have different dimension from x_k)

- Details.

→ $\hat{x}_{n|m}$: estimate of x at time n given observation upto time m. $(m \le n)$

$\hat{x}_{k|k}$: posteriori state estimate mean at k given k observation.

$P_{k|k}$: posteriori estimate covariance matrix
(a measure of estimated accuracy of state estimate)

... and we have two phases:

available from prev. step

(1) Predict : $\hat{x}_{k|k-1} = F_k \hat{x}_{k-1|k-1} + B_k u_k$

(estimate how state would change from k-1 to k)

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k$$

$$E\left((x_{k|k-1} - \hat{x}_{k|k-1})(x_{k|k-1} - \hat{x}_{k|k-1})^T\right)$$

$$= E\left[\left((F_k x_{k-1|k-1} + \underline{B_k u_k} + w_k) - (F_k \hat{x}_{k-1|k-1} + \underline{B_k u_k})\right)\left((F_k x_{k-1|k-1} + \underline{B_k u_k} + w_k) - (F_k \hat{x}_{k-1|k-1} + \underline{B_k u_k})\right)^T\right]$$

$$= E\left[\left(F_k (x_{k-1|k-1} - \hat{x}_{k-1|k-1}) + w_k\right)\left(F_k (x_{k-1|k-1} - \hat{x}_{k-1|k-1}) + w_k\right)^T\right]$$

$$= E\left[F_k (x_{k-1|k-1} - \hat{x}_{k-1|k-1})(x_{k-1|k-1} - \hat{x}_{k-1|k-1})^T F_k^T\right]$$
$$+ E\left[\underline{w_k}\left(F_k (x_{k-1|k-1} - \hat{x}_{k-1|k-1})\right)^T\right]$$
$$+ E\left[\left(F_k (x_{k-1|k-1} - \hat{x}_{k-1|k-1})\right)\underline{w_k^T}\right] + E\left(w_k w_k^T\right)$$

noise are independent.

$$\Rightarrow F_k P_{k-1|k-1} F_k^T + Q_k$$

(2) Update

I:

→ $\tilde{y}_k = z_k - H_k \hat{x}_{k|k-1}$ [Innovation or pre-fit residual]

→ $S_k = H P_{k|k-1} H^T + R_k$ [Covariance of Innovation] If high, less believe observation

↳ $Cov(\tilde{y}_k) = Cov(z_k) + Cov(-H_k \hat{x}_{k|k-1})$

$= R_k + H_k \cdot \underbrace{Cov(\hat{x}_{k|k-1})}_{P_{k|k-1}} H_k^T$

→ $K_k = P_{k|k-1} H_k^T S_k^{-1}$ [Kalman Gain]

If) $P_{k|k-1} \Uparrow$, $S_k^{-1} \Downarrow$: Pred : Uncertain, Obs : Certain. : $K_k \Uparrow$

$P_{k|k-1} \Downarrow$, $S_k^{-1} \Uparrow$: Pred : Certain, Obs : Uncertain : $K_k \Downarrow$

→ $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{y}_k$ [Update (a posteriori) state estimate]

$= \hat{x}_{k|k-1} + K_k (z_k - H_k \hat{x}_{k|k-1})$

$= (I - K_k H_k) \hat{x}_{k|k-1} + K_k z_k$

If) $K_k H_k \Uparrow$, $K_k \Uparrow$: weight more to observation

$\Downarrow$, $\Downarrow$: weight more to prediction

→ $P_{k|k} = (I - K_k H_k) P_{k|k-1}$.

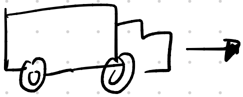
→ $\tilde{y}_{k|k} = z_k - H_k \hat{x}_{k|k}$ [Post-fit residual]

Cov : $H_k P_{k|k} H_k^T + R_k$.

↳ If High, R or Q should be re-defined.

or F / H can be fixed.

• Example.



$$x_k = \begin{bmatrix} p \\ \dot{p} \end{bmatrix} = \overset{\begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}}{F} x_{k-1} + w_k$$

[no Bu_k since no known control]

$$z_k = \underset{[1 \quad 0]}{H} x_k + v_k$$

[position is observed with noise]

$$\Rightarrow \hat{x}_{0|0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$ [We know the initial state]

$$P_{0|0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$ [We let KF know that $\hat{x}_{0|0}$ is certain]

Q) How R_k, Q_k are made?

⇒ Should be designed by engineer.

ex) R_k : depend on sensor's quality!

& check whether $z_k = \begin{bmatrix} z_k^1 \\ \vdots \\ z_k^m \end{bmatrix}$ are independent or not.

Q) Is linear dynamic the best?

→ We can set : $x_{k+1} = f_k(x_k, u_k) + w_k$, $w_k \sim N(0, Q_k)$

$z_k = h_k(x_k) + v_k$, $v_k \sim N(0, R_k)$

[Extended Kalman Filter].