Deriving the Iterative Kalman Filter Scheme

Assume we are able to "guess" an estimate of a system at "time k" denoted as $\hat{x}_k'$ (the prime denotes prediction)

Goal: $\underset{K}{\text{minimize}} \quad J = \mathbb{E}\left[(x_k-\hat{x}_k)^T(x_k-\hat{x}_k)\right]$ (1)

x_k = actual state
$\hat{x}_k$ = optimal estimate

subject to:

(C1) $x_{k+1} = Fx_k + u_k + w_k$
(C2) $y_k = Hx_k + v_k$

F, H, u_k, y_k known

(C3) $Q_k = \mathbb{E}[w_k w_k^T]$
(C4) $R_k = \mathbb{E}[v_k v_k^T]$

covariances
Q_k, R_k given

(C5) $\hat{x}_k = \hat{x}_k' + K_k(\underbrace{y_k - H\hat{x}_k'}_{})$

often called "innovation" or "measurement residual"

can rewrite as $\hat{x}_k = (I-K_kH)\hat{x}_k' + K_k y_k$

Starting with the objective function (1) we can rewrite J as follows

$$J = \mathbb{E}\left[(x_k-\hat{x}_k)^T(x_k-\hat{x}_k)\right]$$

fun fact: $\langle \vec{a}, \vec{a}\rangle = \vec{a}^T\vec{a}$
$= \text{trace}(\vec{a}\vec{a}^T)$ ← check for yourself

$$= \mathbb{E}\left[\text{trace}\left((x_k-\hat{x}_k)(x_k-\hat{x}_k)^T\right)\right]$$

$$= \text{trace}\left[\mathbb{E}\left[(x_k-\hat{x}_k)(x_k-\hat{x}_k)^T\right]\right]$$

→ J = trace of the error covariance!

So, let's replace $\mathbb{E}\left[(x_k-\hat{x}_k)(x_k-\hat{x}_k)^T\right]$ with the covariance matrix "P_k"

This means that our objective function (1) becomes

$$\underset{K}{\text{minimize}} \quad J = \text{trace}(P_k) \quad (2)$$

Let's rewrite P_k using $P_k = \mathbb{E}\left[(x_k-\hat{x}_k)(x_k-\hat{x}_k)^T\right]$ and (C1)-(C5)

recall: $\hat{x}_k = (I-K_kH)\hat{x}_k' + K_k y_k$ (C5)
$y_k = Hx_k + v_k$ (C2)

plug in (C2) → (C5)

$$\hat{x}_k = (I - K_kH)\hat{x}_k' + K_k(Hx_k + v_k)$$

$$\hat{x}_k = \hat{x}_k' + K_kH(x_k - \hat{x}_k') + K_k v_k$$

Now, $x_k - \hat{x}_k = x_k - \hat{x}_k' - K_kH(x_k-\hat{x}_k') - K_k v_k$

$$x_k - \hat{x}_k = (I-K_kH)(x_k-\hat{x}_k') - K_k v_k$$

Finally: $P_k = \mathbb{E}\left[(x_k-\hat{x}_k)(x_k-\hat{x}_k)^T\right]$

$$= \mathbb{E}\left[\left((I-K_kH)(x_k-\hat{x}_k') - K_kv_k\right)\cdot\left((I-K_kH)(x_k-\hat{x}_k) - K_kv_k\right)^T\right]$$

$$= \mathbb{E}\left[(I-K_kH)(x_k-\hat{x}_k')(x_k-\hat{x}_k')^T(I-K_kH)^T\right] - K_k\underbrace{\mathbb{E}\left[v_k(x_k-\hat{x}_k')^T\right]}_{*}(I-K_kH)^T$$

$$- (I-K_kH)\underbrace{\mathbb{E}\left[(x_k-\hat{x}_k')v_k^T\right]}_{*}K_k^T + K_k\underbrace{\mathbb{E}\left[v_kv_k^T\right]}_{=R \text{ via } C_4}K_k^T$$

(*) these terms go to zero because we assume that v_k and $x_k-\hat{x}_k'$ are not correlated..

So... $$P_k = (I-K_kH)\underbrace{\mathbb{E}\left[(x_k-\hat{x}_k')(x_k-\hat{x}_k')^T\right]}_{\text{looks a lot like "P" but with } \hat{x}_k'}(I-K_kH)^T + K_kRK_k^T$$

Call this P_k' !

Finally: $P_k = (I-K_kH)P_k'(I-K_kH)^T + K_kRK_k^T$ (3)

$$= P_k' - P_k'H^TK_k^T - K_kHP_k' + K_kHP_k'H^TK_k^T + K_kRK_k^T$$

$$= P_k' - \underline{P_k'H^TK_k^T} - \underline{K_kHP_k'} + K_k(HP_k'H^T+R)K_k^T$$

↳ note: $K_kHP_k' = (P_k'H^TK_k^T)^T$

$$J = \text{trace}(P_k) = \text{trace}(P_k') - 2\,\text{trace}(K_kHP_k') + K_k(HP_k'H^T+R)K_k^T$$

Now... $\frac{dJ}{dK_k} = -2(HP_k')^T + 2K_k(HP_k'H^T+R) = 0$ for critical values

$$(HP_k')^T = K_k(HP_k'H^T+R)$$

So... $\boxed{K_k = (HP_k')^T(HP_k'H^T+R)^{-1} \quad (4)}$ * We have found the gain!

Now, note from earlier in (3) that $P_k = (I-KH)P_k' - (I-KH)P_k'(KH)^T + KRK$

Substituting (4) into (3),

$$P_k = (I-KH)P_k'(I-KH)^T + KRK^T$$

$$= (I-KH)P_k'I^T - (I-KH)P_k'(KH)^T + KRK^T$$

$$= (I-KH)P_k' - P_k'H^TK^T + KHP_k'H^TK + KRK^T$$

$$= (I-KH)P_k' - P_k'H^TK^T + K(HP_k'H^T+R)K^T$$

$$= (I-KH)P_k' - P_k'H^TK^T + (HP_k')^T\cancel{(HP_k'H^T+R)^{-1}(HP_k'H^T+R)}K^T$$

$$= (I - KH)P_k' - \cancel{P_k' H^T K^T} + \cancel{P_k'^T H^T K^T}$$

since $P_k' = (P_k')^T$

So, we now have that $\boxed{P_k = (I - KH)P_k'}$ $\leftarrow$ corrected covariance calculation.

We should pause: here is where we are at:

Given some "guess" for the state and covariance at iteration k (denoted $\hat{x}_k'$ and P_k' respectively) we have

"iteration k"

$\rightarrow$ given $\hat{x}_k'$ P_k' predictions

① find the Kalman Gain

$$K_k = (HP_k')^T (HP_k' H^T + R)^{-1}$$

② Correct the state and covariance

$$\hat{x}_k = \hat{x}_k' + K_k(y_k - H\hat{x}_k')$$

$$P_k = (I - K_k H)P_k'$$

But what happens at the next iteration? How does one find the "next" prediction $\hat{x}_{k+1}'$ and P_{k+1}'?

We haven't used c_1 and c_3!

Since $x_{k+1} = Fx_k + u_k + w_k$ $\longrightarrow$ guess that $\boxed{\hat{x}_{k+1} = F\hat{x}_k + u_k}$

Now, we know that $P_{k+1} = \mathbb{E}\left[(x_{k+1} - \hat{x}_{k+1})(x_{k+1} - \hat{x}_{k+1})^T\right]$

and $(x_{k+1} - \hat{x}_{k+1}) = (Ax_k + u_k + w_k - A\hat{x}_k - u_k) = A(x_k - \hat{x}_k) + w_k$

So $P_{k+1} = \mathbb{E}\left[\left[A(x_k - \hat{x}_k) + w_k\right]\left[(x_k - \hat{x}_k)^T + w_k^T\right]\right]$

$$= F\,\mathbb{E}\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T\right]F^T + \cancelto{0}{F\,\mathbb{E}\left[(x_k - \hat{x}_k)w_k^T\right]} + \cancelto{0}{\mathbb{E}\left[w_k(x_k - \hat{x}_k)^T\right]F^T} + \underbrace{\mathbb{E}\left[w_k w_k^T\right]}_{Q}$$

$$\boxed{P_{k+1} = FP_kF^T + Q}$$

why? b/c the error and noise are not correlated in a single iteration.

$x_{k+1} - \hat{x}_{k+1}$ and w_k are correlated

We assume that w_{k+1} w_k ... independent

So, we will use this info to predict at the next iteration

So, we will use this info to predict at the next iteration

$$\hat{x}'_{k+1} = F\hat{x}_k + u_k$$

$$P'_{k+1} = FP_kF^T + Q$$

we assume that w_{k+1} w_k are indedendent

$\boxed{\text{Process summary}}$ ✱

At some $k = k_0$, predict an initial state and covariance $\hat{x}'_{k_0}$ P'_{k_0}

for $k = k_0, k_0+1, \ldots$

① get observed values y_k

② compute Kalman gain

$$K_k = (HP'_k)^T (HP_kH^T + R)^{-1}$$

③ correct state and covariance estimates

$$\hat{x}_k = \hat{x}'_k + K(y_k - H\hat{x}'_k)$$

$$P_k = (I - KH)P'_k$$

④ update predictions for the state and covariance in next iteration

$$\hat{x}'_{k+1} = F\hat{x}_k + u_k$$

$$\hat{P}_{k+1} = FP_kF^T + Q$$

⑤ increment k and repeat

$$k = k + 1$$