~~Pythatman~~ Kalman Filter

Hidden states ~~xx~~ x_t ← ~~but~~ not observed
measurements z_t

~~$x_0 \sim$ Gaussian (μ_0, Σ_0)~~

$x_{t+1} \stackrel{?}{=} A_t x_t + b_t + \epsilon_{t+1}$

$x_k = A x_{k-1} + B u_{k-1} + w_{k-1}$ X - dynamics

$z_k = H x_k + v_k$ ~~Z - dynamics~~ Z - measurement

x hidden state with dynamics known.
H the measurement ~~deten~~ noisy mapping of the hidden state x.

w, v are assumed to be independent

$p(w_k) \sim N(0, Q)$

$p(v_k) \sim N(0, R)$

Ideally, $Q \perp R$ and constant.
while they may change over time

We can observe z_k, which is a known mapping of hidden status x_k. We have knowledge of x_k dynamics
~~while~~

$\hat{x}_k^-$: prior of x_k without z_k
$\hat{x}_k$: posterior of x_k given z_k.

$e_k^- = x_k - \hat{x}_k^-$ prior error

$e_k^+ = x_k - \hat{x}_k$ posterior error.

$P_k^- = E[e_k^- e_k^{-T}]$ matrix

$P_k = E[e_k \cdot e_k^T]$

$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-)$

$e_k = x_k - \hat{x}_k = x_k - \hat{x}_k^- - K(z_k - H\hat{x}_k^-)$.

min $P_k = E[e_k \cdot e_k^T]$

$\tilde{P}_k = E[e_k^T e_k]$

$= E[(x_k^T - \hat{x}_k^{-T} - (z_k^T - \hat{x}_k^{-T} H^T) K^T)$

$(x_k - \hat{x}_k^- - K(z_k - H\hat{x}_k^-))]$

$$\varepsilon_k = z_k - H\hat{x}_k^-$$

$$= E\left[(e_k^{-T} - \varepsilon_k^{-T}K^T)(e_k^- - K\cdot\varepsilon_k^-)\right]$$

$$= E\left[e_k^{-T}e_k^- + \varepsilon_k^{-T}K^TK\varepsilon_k^- - 2\varepsilon_k^{-T}K^Te_k^-\right]$$

$$= E[e_k^{-T}e_k^-] + E[\varepsilon_k^{-T}K^TK\varepsilon_k^-]$$

$$\underline{-2E[\varepsilon_k^{-T}K^Te_k^-]}$$

$$\shortparallel$$

$$\underline{-2\, e_k^{-T}H^TK^Te_k^-}$$

$$e_k^- = x_k - \hat{x}_k^-$$

$$\varepsilon_k^- = H(x_k - \hat{x}_k^-) + V_k$$

$$E[\varepsilon_k^{-T}K^TK\varepsilon_k^-] = E[(V_k^T + (x_k - \hat{x}_k^-)^TH^T)K^TK$$

$$(V_k + H(x_k - \hat{x}_k^-))]$$

$$= E[V_k^TK^TKV_k]$$

$$= e_k^{-T}e_k^- - 2e_k^{-T}H^TK^Te_k^-$$

$$+ E[\varepsilon_k^{-T}K^TK\varepsilon_k^-]$$

$$\epsilon_k^- = \cancel{\theta}\, H x_k + v_k - H\hat{x}_k^-$$

$$= \cancel{H \ldots}\ \cancel{\theta}\, H e_k^- + v_k$$

$$\epsilon_k^{-T} K^T K \epsilon_k^-$$

$$= (v_k^T + e_k^{-T} H^T) K^T K (\cancel{\theta} H e_k^- + v_k)$$

$$E[\epsilon_k^{-T} K^T K \epsilon_k^-] = E[v_k^T K^T K v_k] + E[e_k^{-T} H^T K^T K H e_k^-]$$

$$= \overset{\text{trace}}{\cancel{\text{diag}}}(KRK^T)$$

$$+ e_k^{-T} H^T K^T K H e_k^-$$

$v_k \sim N(0, R)$

$K v_k \sim N(0, KRK^T)$

$E[v_k v_k^T] = R.$ $\qquad e_{k\ n\times 1}^- \qquad H_{m\times n} \qquad K_{n\times m}$

$$\text{Target} = E[e_k^{-T} e_k^-] - 2 e_k^{-T} H^T K^T e_k^-$$

$$+ \mathrm{tr}(KRK^T) + e_k^{-T} H^T K^T K H e_k^-$$

$$\frac{\partial \, Target}{\partial K} = -2E[e_k^- e_k^{-T}]H^T + 2KR + 2KHE[e_k^- e_k^{-T}]H^T$$

$$= 0$$

$$\Rightarrow \underset{n \times m}{K} = P_k^- H^T (H P_k^- H^T + R)^{-1}$$

Interesting Fact:

As $R \to 0$ $K \to H^+$ (pseudo inverse)

As $P_k^- \to 0$ $K \to 0$

i.e. if measurement is flawless, weight concentrates over measurement ~~(Kalman~~

if prior is perfect weight concentrates over forecasting.

K is called Kalman Gain.

Now turn back and check Kalman Model.

The scenario:

A hidden process $\{X_k\}$ not observable

The evolution of X_k

$X_k = A X_{k-1} + B u_{k-1} + w_{k-1}$

$w_{k-1} \sim N(0, Q)$

The observation / measurement.

$Z_k = H X_k + V_k$ $V_k \sim N(0, R)$

The initial state x_0 is known. How to estimate $(x_1, x_2, \cdots, x_n)$?

Note: the ~~evolution~~ forecasting/measurement process both have errors. The goal is to aggregate the two noisy estimation.

$\hat{x}_k^-$ prior estimate without z_k

$\hat{x}_k$ posterior estimate with z_k.

$e_k^- = x_k - \hat{x}_k^-$ prior error

$e_k = x_k - \hat{x}_k$ posterior error.

The known: $x_0, u_0, (z_1, z_2, \cdots, z_n), A, B, Q, R$.

unknown/estimate: $\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_n$

optimization suggests

$$\underbrace{\hat{x}_k}_{n\times 1} = \underbrace{\hat{x}_k^-}_{n\times 1} + \underbrace{K}_{n\times m}(\underbrace{z_k}_{m\times 1} - \underbrace{H}_{m\times n}\underbrace{\hat{x}_k^-}_{n\times 1})$$

$$K = P_k^- H^T (H P_k^- H^T + R)^{-1}$$

$$P_k^- = E[e_k^- {e_k^-}^T]_{n\times n}$$

$$P_k = E[e_k \cdot e_k^T]_{n\times n}$$

It can be proved that

$$E[X_k] = \hat{X}_k$$

$$E[(X_k - \hat{X}_k)(X_k - \hat{X}_k)^T] = P_k$$

And if as stated, the forecasting / measurement has normal error (mean 0), then the ~~posterior distribution follows~~ the true / hidden state of X_k $\sim N(\hat{X}_k, P_k)$

H known

Kalman Filter Algo.

A, B, Q, R known.

$\hat{X}_0 = X_0$

P_0 known.

1) $\hat{X}_k^- = A\hat{X}_{k-1} + Bu_{k-1}$

$P_k^- = AP_{k-1}A^T + Q$

Forecasting.

2) $K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$

Measurement correction

$\hat{X}_k = \hat{X}_k^- + K_k(Z_k - H\hat{X}_k^-)$

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

Following this way, one can find $(\hat{X}_1, \hat{X}_2, \cdots, \hat{X}_n)$ and the post cov matrix. $(P_1, P_2, \cdots, P_n)$ where $X_k \sim N(\hat{X}_k, P_k)$

Note: In each step, it is easy to assume R_k, Q_k changes. Even A_k, B_k is allow to change. ~~The only~~ They merely play the role to evolve ~~the change~~

Extended Kalman Filter (EKF)

It's an approximate version of Nonlinear process Kalman
(Ad-hoc, the ~~~~ problem is nonlinearly model
does not keep normality)