Kalman Filter
Problem Set-up:

Reality:

1. Process model: $x_k = A x_{k-1} + w_p$
2. Measurement model: $y_k = H x_k + n_m$

State estimator: Combine process & measurement

$\hat{x}_k = x_k^P + K_k (y_k - \hat{y}_k)$

If I have $x_{k-1}$, then I can model both $x_k^P$ and $\hat{y}_k = H x_k^P$
What do we need to do this?

- Estimate $K_k$
- time index

- If process noise $n_p = 0$,
  - $K_k$ should be 0
- if measurement noise $n_m = 0$
  - $K_k$ should be $H^{-1}$

because,  
$$\hat{x}_k = x_k^p + H^{-1}(Hx_k + n_m - Hx_k^p)$$
$$= x_k + H^{-1}n_m \approx x_k$$

Main intuition: seems like possible to use $K_k$ as a knob that combines the process and measurement.

If I turn $K_k$ the wrong way, the prediction should diverge from the true $x_k$, which will also manifest in gap in true and modeled measurement ($y_k - \hat{y}_k$).
434: Kalman Filters

1. Start with an initial $\hat{x}_0$

2. That gives us a process based estimate $x_1^p$

3. If this was correct, then I expect the measurement $y_1$ to match with my modeled measurement $Hx_1^p$

   If it does not match, then the error is composed of both process error and measurement error.

4. My goal is to modulate the process estimate $x_1^p$ with some linear function of this error.

   This gives me an estimate of the state variable as: $\hat{x}_1 = x_1^p + K_1(y_1 - Hx_1^p)$

5. What should $K_1$ be?

   - Well, design it such that it minimizes the MSE of state estimate error, defined as: $e_1 = x_1 - \hat{x}_1$

   - So we want to minimize $E[e_1^2]$

6. Let's model $e_1$ ... $e_1 = x_1 - (x_1^p + K_1(y_1 - Hx_1^p)) = x_1 - (x_1^p + K_1(Hx_1 + n_m - Hx_1^p))$

   $(1 - K_1H)x_1 - (1 - K_1H)x_1^p + K_1n_m$

   $(1 - K_1H)(x_1 - x_1^p) + K_1n_m$

   $\text{e}_1^p = (1 - K_1H)e_1^p + K_1n_m$

   Not surprising that this error has both un-modeled components – the process error $e_1^p$ and the measurement error $n_m$

7. For MSE, we compute $E[e_1e_1^T]$ ... since $e_1$ can be a vector ... or you can stack up $[e_1, e_2, e_3 ...]$ to make a vector

   - This expectation then becomes: $P_1 = E[e_1e_1^T] = (1 - K_1H)(e_1 - x_1^p)(e_1 - x_1^p)^T(1 - K_1H)^T + K_1n_m n_m^T K_1^T$

   Cross terms aren’t present because $e_1^p$ and $n_m$ are uncorrelated. Why? Because the process and measurement errors are independent. Thus,

   $E[e_1e_1^T] = (1 - K_1H)P_1(1 - K_1H)^T + K_1R_m K_1^T$

   Not surprising that this error covariance has the process error covariance $P_1^p$ and the measurement error covariance $R_m$

8. Find $K_1$ that minimizes this covariance, i.e.,\[
\min_K (1 - K_1H)P_1(1 - K_1H)^T + K_1R_m K_1^T
\]

   - This gives: $K_1 = \frac{P_1^p H^T}{(H P_1^p H^T + R_m)}$

   [ See equations 11.21 to 11.24 in this article for minimization ]

9. Perhaps we can assume we know covariance for the measurement error $R_m$ ... but we don't have $P_1^p$

   Since $K_1 = f(P_1^p)$ ... we need $P_1^p$

10. Ok, so let's define $P_1^p = E[e_1^p e_1^p]^T$ where $e_1^p = x_1 - x_1^p = Ax_0 + n_p - Ax_0 = A e_0 + n_p$

    So, $P_1^p = E[(A e_0 + n_p)(A e_0 + n_p)^T] = A P_0 A^T + R_p$

    ... again, assuming $e_0$ and $n_p$ are uncorrelated.

    So we have $P_1^p = A P_0 A^T + R_p$

    Of course, this $P_1^p$ depends on the previous $P_0$

    Assuming we have resolved the previous states well, we now have everything we need for $K_1$

11. Now, let's resolve this $P_1^p$ from last step ... which is the same as resolving $P_1$ since it would be used in the next step.

    $P_1^p = f(K_1, P_1)$ ... so we have resolved both $K_1$ and $P_1$ now

    In fact, plugging $K_1$ into $P_1$ ... and then simplifying, we get:

    $P_1 = (1 - K_1H)P_1^p$

    ▼ Derivation
\[ P_1 = (1 - k_i H) P_1^P (1 - k_i H)^T + k_i R_m K_i^T \]
\[ = (P_1^P - k_i H P_1^P) (1 - H^T K_i) \]
\[ = P_1^P - P_1^P H^T K_i^T - k_i H P_1^P + k_i (H P_1^P H^T + R_m) K_i^T \]

Substitute \( K_i = \frac{P_1^P H^T}{H P_1^P H^T + R_m} \)

\[ = P_1^P - P_1^P H^T \left( \frac{P_1^P H^T}{H P_1^P H^T + R_m} \right) - \left( \frac{P_1^P H^T}{H P_1^P H^T + R_m} \right) P_1^P + \frac{P_1^P H^T}{H P_1^P H^T + R_m} \]

\[ = P_1^P - k_i H P_1^P \]

\[ = (1 - k_i H) P_1^P \]

Thus, \( P_1 = (1 - k_i H) P_1^P \)
\[
\begin{align*}
\hat{x}_k &= A \hat{x}_{k-1} \\
P_k^P &= A P_{k-1} A^T + R_p
\end{align*}
\]

\[
\begin{align*}
\hat{x}_k &= x_k^p + K_k (y_k - H x_k^p) \\
K_k &= P_k^p H^T \\
&= \frac{P_k^p H^T}{H P_k^p H^T + R_m} \\
P_k &= (1 - K_k H) P_k^p
\end{align*}
\]
Questions