Kalman Filter to estimate Pandemic

Learning objectives:

- System modeling
- Simulating system with noise
- Designing and implementing a simple Kalman Filter

We will aim to implement a Kalman filter on a simplified pandemic model.

Disclaimer: [This project does not aim to make light of ongoing pandemic. The Pandemic prediction is extremely complicated and there has significant loss of life and a great deal of suffering. You may choose any other system, so long as you can develop the model, have a noise and disturbance term and implement either a discrete or continuous time Kalman Filter]

"Real" Advancements:

There has been considerable recent effort. For instance Singh et al. "Kalman filter based short term prediction model for COVID-19 spread" in Applied Intelligence, Published: 03 November 2020 [https://link.springer.com/article/10.1007/s10489-020-01948-1] presents an interesting and promising results. A full discussion is well beyond the scope of this course.

Model:

We have a simple discrete time model with 20 states [x_1 to x_{20}], where the value of state x_i corresponds the number of persons who have been infected for $i - days$. (So $x_4(k) = 20$ implies that there are 20 people in the 4[th] day of infection). [After 20 days the person is deemed to not be infectious and is not tracked by the model]. With each passing day we shift the number of infections (this ignores adverse outcomes)

$$\begin{aligned} x_{20}(k+1) &= x_{19}(k) \\ x_{19}\,(k+1) &= x_{18}(k) \\ &\vdots \\ x_2(k+1) &= x_1(k) \end{aligned} \qquad \text{Eq. 1}$$

The real number of new infection is given by

$$x_1(k+1) = a_1x_1(k) + a_2x_2(k) + \cdots + a_{19}x_{19}(k) + w \qquad \text{Eq. 2}$$

Here $w\sim(0, Q)$ is a zero mean Gaussian process to account for any uncertainties (assume Q=2).

If it given that

$$a_1, a_2, \cdots, a_5 = 0 \qquad \text{Eq. 3}$$

$$a_6, a_7, \cdots, a_{14} = 0.125$$
$$a_{15}, a_{16}, \cdots, a_{19} = 0.0025$$

It is assumed that we only measure the number of people who have been infected between 14-15 days

$$y_k = 0.5 \times x_{14}(k) + 0.5 \times x_{15}(k) + v \qquad \text{Eq. 4}$$

Here $v \sim (0, R)$ is a zero mean Gaussian process to account for any uncertainties (assume R=1).

The provided m-file

You are given a matlab file (panmodel.m) that has a function “panmodel”. Each time you call the function it will report y_k, the measured number of infections. [Please use “clear all” to reset the simulation time to 0]. The simulations start at an unknown state

Project task

1. Create a discrete time model for Eq. 1-Eq. 4
2. Assuming $\hat{x}(k = 0) = 0$, $P = I$, simulate the Kalman filter for 100 days
    a. You can implement this as a simple loop in Matlab (k=1:100) [use the equations for the discrete time Kalman Filter (in the handout) to estimate P_k^-, K_k, P_k^+, $\hat{x}_k^-$ and $\hat{x}_k^+$ at each time step]
    b. At each step store y_k, $H_k\hat{x}_k^-$ and $\hat{x}_k^+(1)$ in separate variables and plot the evolution of these variables over time.
3. You have been provided data from BC CDC . (This is only for education purpose).
    a. Describe how you might make a forecast using the Kalman Filter
    b. See how well your data is able to make a 1 day forecast starting from day 150.
    c. How can you improve your prediction?

    [Note: This is an open ended problem & there isn’t 1 correct answer]