(5) We know that Kalman filter gives us the MMSE estimate.

We can consider the given scenario as,

$$x(k+1) = x(k+1) = \mathcal{E}(k+1) \quad \text{—— (1)}$$

$(\mathcal{E} \sim WN(\mu, \sigma_x^2)$ assuming the observations are i.i.d.)

and $y(k) = x(k) + v(k)$ —— (2)

So we have,

$A = 0 = B = q$, $C = 1$, $q = \sigma_x^2$, $r = \sigma_v^2$

Substituting these values in the Kalman filter expression for a first order system,

$$K_{f,k} = \frac{(a^2 P_{k-1} + q)C}{(c^2(a^2 P_{k-1} + q) + r)} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2} = \text{constant} = K_f$$

Now $\hat{x}(k|k-1) = \hat{x}(k-1|k-1)$ —— (3)

and $\hat{x}(k|k) = \hat{x}(k-1|k-1) + K_{f,k}(y(k) - \hat{x}(k-1))$

$\Rightarrow \hat{x}(k|k) = (1 - K_f)\hat{x}(k-1|k-1) + K_{f,k}\, y(k)$

$= (1-K_f)\hat{x}(k-1|k-1) + K_{f,k}\, y(k)$ —— (4)

and $\hat{x}(0|-1) = \bar{x}$ —— (5)

$\Rightarrow \hat{x}(0|0) = \bar{x} + K_f(y(k) - \bar{x})$ —— (6)

By recursive eliminating terms until $\hat{x}(0|0)$

$$\hat{x}(N|N) = \sum_{k=0}^{N} K_f (1-K_f)^k (y(N-k)) + (1-K_f)^{N+1}\bar{x}$$

$$\Rightarrow \hat{x}(N|N) = \sum K_f (1-K_f)^k (y(N-k)) +$$
$$+ (1-K_f)^{N+1} \bar{x}$$

So $\beta(N) = (1-K_f)^{N+1}$

~~&~~ we don't have ~~a~~ a clear $\alpha(N)$. $\sum y(k)$ is actually a weighted ~~some~~ sum, with less weight given to older data

In the limit $N \to \infty$,

~~The sum~~ $\beta(N)$ vanishes, since,

$$1-K_f = \frac{\sigma^2 v}{\sigma^2 x + \sigma^2 v} < 1.$$

For $\alpha(N)$, consider $y(k) = \mu + \underbrace{\varepsilon(k)}_{\sim WN(0,\sigma_x^2)} + \underbrace{v(k)}_{\sim WN(0,\sigma^2 v)}$

So $\lim_{N\to\infty} E\left(\sum (K_f (1-K_f)^k y(N-k))\right)$

$$\Rightarrow \lim_{N\to\infty} E(\alpha(N)) = \sum_{k=0}^{\infty} K_f (1-K_f)^k \mu .$$
$$= \mu \qquad (\because E(\varepsilon) = E(v) = 0)$$

So as $\lim_{N\to\infty}$, we completely rely on the data, and the average value of the estimate is the mean of the R.V. x.

METHOD-2

Since the ε used in 1 has non-zero mean, I tried using a different formulation having 0 mean ε

$$\begin{bmatrix} x(k+1) \\ \mu \end{bmatrix} = \begin{bmatrix} 1\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x(k) \\ \mu \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \varepsilon(k)$$

and $y(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x(k) \\ \mu \end{bmatrix} + v(k)$

However $O_n = \begin{pmatrix} C \\ CA \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\therefore$ system is observable rank$(O_n) = 2$

We note that both $x(k)$ and μ are unknowns.

(If μ is known, $\hat{x}_{MMSE} = E(x) = \mu$, irrespective of N)

We can jointly estimate both in an MMSE fashion using Kalman filtering.

($\varepsilon \sim WN(0, \sigma_x^2)$ and $v \sim WN(0, \sigma_v^2)$ and in addition both are Gaussian).

$A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$

$Q = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & 0 \end{bmatrix}$ ($\because$ μ is a fixed value without an error)

$R = \sigma^2 v$

~~Riccatti eqns~~: $\bar{P_k} = AP_{k-1}A^T + Q$ —— ①

$K_{f,k} = (\bar{P_k})(C^T)(R + C\bar{P_k}C^T)^{-1}$ —— ②

$P_k = (I - K_{f,k}C)(\bar{P_k})$ —— ③

Let P_{k-1} be a 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$\bar{P_k} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} \sigma x^2 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} c & d \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} \sigma x^2 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \sigma x^2 + d & d \\ d & d \end{pmatrix}$$

So, $K_{f,k} = \begin{pmatrix} \sigma^2 x + d & d \\ d & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}(R + C\bar{P_k}C^T)^{-1}$ —— ④

Consider $R + C\bar{P_k}C^T$

$$= \sigma v^2 + \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} \sigma^2 x + d & d \\ d & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$= \sigma v^2 + \begin{pmatrix} \sigma x^2 + d & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$\Rightarrow R + CP_k^- C^T = \sigma_v^2 + \sigma^2 x + d$ —— ⑤

Substituting ~~in~~ ~~from~~ ⑤ in ④

$$\Rightarrow K_{f|k} = \begin{pmatrix} \dfrac{\sigma^2_x + d}{\sigma^2_v + \sigma^2_x + d} \\ \\ \dfrac{d}{\sigma^2_v + \sigma^2_x + d} \end{pmatrix} \quad \text{——} ⑥$$

$I - K_{f,k} C =$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{\sigma^2_v + \sigma^2_x + d} \begin{pmatrix} \sigma^2_x + d \\ d \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix}$$

$$\therefore ③ \Rightarrow P_k = \begin{bmatrix} \dfrac{\sigma^2_v}{\sigma^2_v + \sigma^2_x + d} & 0 \\ -\dfrac{d}{\sigma^2_v + \sigma^2_x + d} & 1 \end{bmatrix} \begin{bmatrix} \sigma^2_x + d & d \\ d & d \end{bmatrix}$$

Since the 2nd diagonal term $(2,2)$ is the only ~~important~~ relevant term,

$$d_k = \left(\frac{-d_{k-1}}{\sigma^2_v + \sigma^2_x + d_{k-1}} \right) d_{k-1} + d_{k-1}$$

$$\Rightarrow d_k = \frac{(\sigma^2_v + \sigma^2_x)\, d_{k-1}}{\sigma^2_v + \sigma^2_x + d_{k-1}} \quad \text{——} ⑦$$

We know: $\hat{x}(k|k) = \hat{x}(k-1|k-1) + K_{f,k}\left(y(k) - C\hat{x}(k|k-1)\right)$

where $\hat{s}(k) = \begin{bmatrix} \hat{x}(k) \\ \hat{\mu}(k) \end{bmatrix}$

$$\begin{bmatrix} \hat{x}(k|k-1) \\ \hat{\mu}(k|k-1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}(k-1|k-1) \\ \hat{\mu}(k-1|k-1) \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \hat{x}(k|k-1) \\ \hat{\mu}(k|k-1) \end{bmatrix} = \begin{bmatrix} \hat{\mu}(k-1|k-1) \\ \hat{\mu}(k-1|k-1) \end{bmatrix}$$

$$\therefore \begin{pmatrix} \hat{x}(k|k) \\ \hat{\mu}(k|k) \end{pmatrix} = \begin{pmatrix} \hat{\mu}(k-1|k-1) \\ \hat{\mu}(k-1|k-1) \end{pmatrix} + \begin{pmatrix} \dfrac{\sigma^2_x + d}{\sigma^2_v + \sigma^2_x + d} \\ \dfrac{d}{\sigma^2_v + \sigma^2_x + d} \end{pmatrix} \left(y(k) - (1 \;\; 0) \begin{pmatrix} \hat{\mu} \\ \hat{\mu} \end{pmatrix} \right)$$

$$= \begin{pmatrix} \hat{\mu}(k-1|k-1) + \dfrac{\sigma^2_x + d}{\sigma^2_v + \sigma^2_x + d} \left(y(k) - \hat{\mu}(k-1|k-1) \right) \\ \hat{\mu}(k-1|k-1) + \dfrac{d}{\sigma^2_v + \sigma^2_x + d} \left(y(k) - \hat{\mu}(k-1|k-1) \right) \end{pmatrix}$$

$$\cancel{= \begin{pmatrix} \hat{\mu}(k-1|k-1) \left(\dfrac{\sigma^2_v}{\sigma^2_x + \sigma^2_v + d} \right) - \left(\dfrac{\sigma^2_x + d}{\sigma^2_v + \sigma^2_x + d} \right) y(k) \\ \hat{\mu}(k-1|k-1) \left(\dfrac{\sigma^2_v}{\ } \right. \end{pmatrix}}$$

$$\Rightarrow \hat{x}(k|k) = \hat{\mu}(k-1|k-1)\left(\frac{\sigma_v^2}{\sigma_v^2+\sigma_x^2+d}\right) + \left(\frac{\sigma_x^2+d}{\sigma_v^2+\sigma_x^2+d}\right) y(k)$$

$$\hat{\mu}(k|k) = \hat{\mu}(k-1|k-1)\left(\frac{\sigma_v^2+\sigma_x^2}{\sigma_v^2+\sigma_x^2+d_{k-1}}\right) + \frac{d_{k-1}}{\sigma_v^2+\sigma_x^2+d_{k-1}} y(k) \quad \text{—— (8)}$$

$$\lim_{k\to\infty} \hat{\mu}(k|k) = \hat{\mu}(k-1|k-1)$$

$$\left(\because (7) \Rightarrow d_k \to 0 \text{ because } \frac{\sigma_v^2+\sigma_x^2}{\sigma_x^2+\sigma_v^2+d_{k-1}} < 1\right)$$

→ To conclude, since d_k is itself in a non-linear recursion, I am not able to write it in the expected form

→ Nonetheless by performing the recursion in (7) and (8), we can obtain the MMSE estimate of x.

→ And as $N \to \infty$, our estimate from the data converges.