AFRICAN INSTITUTE FOR MATHEMATICAL SCIENCES

(AIMS RWANDA, KIGALI)

Name: Akor Stanley
Course: DA1

Assignment Number: 1
Date: March 20, 2021

Question 1

Given the real-valued state variables;

$$X_i, i = 1, 2, 3, 4.$$

Where;

$$X_{i+1} = 4X_i(1 - X_i) \tag{1}$$

The data model is given by;

$$Y_i = X_i + \epsilon_i$$

(a) Given that the observations are y_2 and y_3, the cost function is;

$$J(x_2, x_3; y_2, y_3) = \epsilon_2^2 + \epsilon_3^2 \tag{2}$$

$$= (x_2 - y_2)^2 + (x_3 - y_3)^2 \tag{3}$$

From equation1, we obtain $x_3 = 4x_2(1 - x_2)$, subtituting the obtained result for x_3 in equation3, we shall have;

$$J(x_2; y_2, y_3) = (x_2 - y_2)^2 + (4x_2(1 - x_2) - y_3)^2 \tag{4}$$

To minimise equation4, we set $\frac{\partial J}{\partial x_2} = 0$, so;

$$\frac{\partial J}{\partial x_2} = 2(x_2 - y_2) + 2(4 - 8x_2)(4x_2(1 - x_2) - y_3) = 0 \tag{5}$$

$$(x_2 - y_2) + (4 - 8x_2)(4x_2(1 - x_2) - y_3) = 0 \tag{6}$$

(b) Let $y_2 = 0.5$ and $y_3 = 0.01$, and substituting in equation6, and upon simplifying, we shall obtain;

$$32x_2^3 - 48x_2^2 + 17.08x - 0.54 = 0 \tag{7}$$

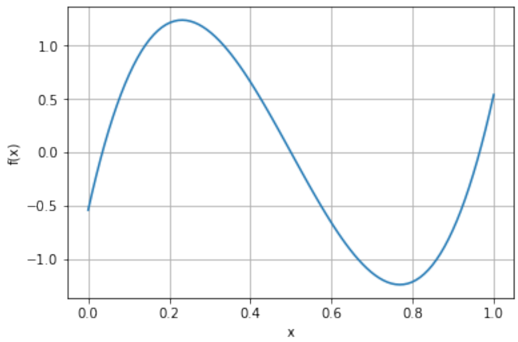

Figure 1: Graphical Solution

Figure1 shows the graphical solution to equation7, it follows that the exact values for which the function is minimum are;

$$x_2 = 0.03497311190565431, \quad x_2 = 0.500000000000000, \quad x_2 = 0.965026880943457$$

Finding the roots of equation7 is a neccessary condition that the function is a minimum at some values of x_2, however, this is not a sufficient condition, we still need extra information in order to get the filtered estimate of x_2.

(c) The cost funtion is given by;

$$J(x_2; y_2, y_3) = (x_2 - y_2)^2 + (4x_2 (1 - x_2) - y_3)^2$$

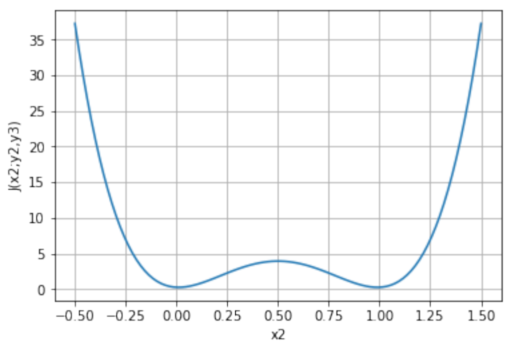

Figure 2: A plot of the cost function

Even with the extra given information, we still cannot get the filtered extimate of x_2 because from our plot2, we observe that they are two local minimum and no global minimum, also we notice that the point at which $x_2 = 0.50000$ is not a local minimum, but rather, it is a local maximum.

Question 2

Given the process model;

$$X_{i+1} = \alpha X_i + \delta_i \qquad i = 1, 2, 3$$

Where X_1 is given, and the data model

$$Y_i = X_i + \epsilon_i, \qquad i = 2, 3$$

Assuming $\epsilon_i \sim N(0, \tau)$ or $\forall_i$, and
$Y_2 \sim N(\alpha x_1, 1 + \tau^2)$, $\quad Y_3 \sim N(\alpha^2 x_1, 1 + \tau^2 + \alpha^2)$ and $Cov(Y_2, Y_3) = \alpha$

(a) To find the inverse of the covariance matrix $Cov(Y)^{-1}$, we shall first determine the $Cov(Y)$

$$\begin{aligned} Cov(Y) &= \begin{pmatrix} var(Y_2) & Cov(Y_2, Y_3) \\ Cov(Y_3, Y_2) & var(Y_3) \end{pmatrix} \\ &= \begin{pmatrix} 1 + \tau^2 & \alpha \\ \alpha & 1 + \tau^2 + \alpha^2 \end{pmatrix} \end{aligned}$$

Thus, $Cov(Y)^{-1}$ is given by;

$$Cov(Y)^{-1} = \frac{1}{|Cov(Y)|} \begin{pmatrix} 1 + \tau^2 + \alpha^2 & -\alpha \\ -\alpha & 1 + \tau^2 \end{pmatrix}$$

Where;

$$\begin{aligned} |Cov(Y)| &= (1 + \tau^2)(1 + \tau^2 + \alpha^2) - \alpha^2 \\ &= (1 + \tau^2) + \tau^2(1 + \tau^2 + \alpha^2) \end{aligned}$$

Therefore,

$$Cov(Y)^{-1} = \frac{1}{(1 + \tau^2) + \tau^2(1 + \tau^2 + \alpha^2)} \begin{pmatrix} 1 + \tau^2 + \alpha^2 & -\alpha \\ -\alpha & 1 + \tau^2 \end{pmatrix}$$

$$f_{Y|x_1} \propto \exp\left(-\frac{1}{2((1 + \tau^2) + \tau^2(1 + \tau^2 + \alpha^2))} (y_2 - \alpha x_1, y_3 - \alpha^2 x_1) Cov(Y)^{-1} (y_2 - \alpha x_1, y_3 - \alpha^2 x_1)^T \right)$$

Extracting the exponent term, and let;

$$\begin{aligned} R &= 2((1 + \tau^2) + \tau^2(1 + \tau^2 + \alpha^2)) \\ S &= (y_2 - \alpha x_1, y_3 - \alpha^2 x_1) \begin{pmatrix} 1 + \tau^2 + \alpha^2 & -\alpha \\ -\alpha & 1 + \tau^2 \end{pmatrix} \begin{pmatrix} y_2 - \alpha x_1 \\ y_3 - \alpha^2 x_1 \end{pmatrix} \end{aligned}$$

Now let us resolve the matrix S

$$\begin{aligned} S &= (y_2 - \alpha x_1, y_3 - \alpha^2 x_1) \begin{pmatrix} 1 + \tau^2 + \alpha^2 & -\alpha \\ -\alpha & 1 + \tau^2 \end{pmatrix} \begin{pmatrix} y_2 - \alpha x_1 \\ y_3 - \alpha^2 x_1 \end{pmatrix} \\ &= (y_2 - \alpha x_1, y_3 - \alpha^2 x_1) \begin{pmatrix} (1 + \tau^2 + \alpha^2)(y_2 - \alpha x_1) - \alpha(y_3 - \alpha^2 x_1) \\ -\alpha(y_2 - \alpha x_1) + (1 + \tau^2)(y_3 - \alpha x_1) \end{pmatrix} \\ &= (1 + \tau^2 + \alpha^2)(y_2 - \alpha x_1)^2 - 2\alpha(y_2 - \alpha x_1)(y_3 - \alpha^2 x_1) + (1 + \tau^2)(y_3 - \alpha^2 x_1)^2 \end{aligned}$$

Thus the value of the exponent is given by;

$$-\left[\frac{(1+\tau^2+\alpha^2)(y_2-\alpha x_1)^2-2\alpha(y_2-\alpha x_1)(y_3-\alpha^2 x_1)+(1+\tau^2)(y_3-\alpha^2 x_1)^2}{2((1+\tau^2)+\tau^2(1+\tau^2+\alpha^2))}\right]$$

The dimensionn of the exponent is 1 because it is a scalar.

(b) By substituting the expression of our exponent into our distribution function, we shall obtain;

$$f_{Y|x_1} \propto \exp\left(-\left[\frac{(1+\tau^2+\alpha^2)(y_2-\alpha x_1)^2-2\alpha(y_2-\alpha x_1)(y_3-\alpha^2 x_1)+(1+\tau^2)(y_3-\alpha^2 x_1)^2}{2((1+\tau^2)+\tau^2(1+\tau^2+\alpha^2))}\right]\right) \tag{8}$$

From the above equation, we can obtain $f_{Y|x_2}$ using the linear process model, that is;

$$x_{i+1} = \alpha x_i + \delta_i$$

We assume the $\delta_i = 0$

$$x_2 = \alpha x_1$$

Substituing $x_2 = \alpha x_1$ in equation(8), we obtain;

$$f_{Y|x_2} \propto exp\left(-\left[\frac{(1+\tau^2+\alpha^2)(y_2-x_2)^2-2\alpha(y_2-x_2)(y_3-\alpha x_2)+(1+\tau^2)(y_3-\alpha x_2)^2}{2((1+\tau^2)+\tau^2(1+\tau^2+\alpha^2))}\right]\right) \tag{9}$$

From 9, we want to maximize our joint distribution function, thus we extract the numerator of the exponent and set it as J,we ignore the denominator because it does not depend on x_2. Therefore;

$$J = (1+\tau^2+\alpha^2)(y_2-x_2)^2-2\alpha(y_2-x_2)(y_3-\alpha x_2)+(1+\tau^2)(y_3-\alpha x_2)^2$$

To maximize $f_{Y|x_2}$, we have to minimize J.

$$\frac{\partial J}{\partial x_2} = 0$$

So,

$$-2(1+\tau^2+\alpha^2)(y_2-x_2)-2\alpha\left[-(y_3-\alpha x_2)-\alpha(y_2-x_2)\right]-2\alpha(1+\tau^2)(y_3-\alpha x_2)=0$$

Upon simplification, we obtain;

$$x_2(1+\tau^2+\alpha^2\tau^2)-y_2(1+\tau^2)-y_3\alpha\tau^2=0$$

$$x_2 = \frac{y_2(1+\tau^2)+y_3\alpha\tau^2}{1+\tau^2+\alpha^2\tau^2} \tag{10}$$

(c) **case 1:** $0 < \alpha < 1$**:** For α close to zero, then equation10, becomes;

$$x_2 \approx \frac{y_2(1+\tau^2)}{(1+\tau^2)} = y_2$$

This implies that data y_2 will have an effect on the filtering while y_3 will have no impact. Similary, for α close to 1, equation10 becomes;

$$x_2 \approx \frac{y_2(1+\tau^2)+y_3\tau^2}{1+2\tau^2} = y_2\left(\frac{(1+\tau^2)}{1+2\tau^2}\right) + y_3\left(\frac{\tau^2}{1+2\tau^2}\right)$$

Comparing the magnitude of y_2 and y_3 from above, we observe that y_2 will have a greater magnitude, thus even though both y_3 and y_2 would affect the filtering, y_2 will have a greater effect.

case 2: $\tau \ll 1$**:** For τ close to zero, then equation10 becomes;

$$x_2 \approx y_2$$

This implies that only data y_2 will have an effect on the filtering.