

[All the problems I discussed with Spandan, Soumyadeep]

Problem 1

Let X, Y_1, Y_2 be three random variables with joint density f_{X,Y_1,Y_2} . For a fixed y_1 , consider two random variables \tilde{X}, \tilde{Y}_2 with joint distribution $g_{\tilde{X},\tilde{Y}_2}$ defined as $g_{\tilde{X},\tilde{Y}_2}(x, y_2) = f_{X,Y_2|Y_1}(x, y_2 | y_1)$. Show that

$$\mathbb{E}[\tilde{X} | \tilde{Y} = y_2] = \mathbb{E}[X | Y_1 = y_1, Y_2 = y_2]$$

What is the relevance of this fact in our derivation of recursive estimation in the lecture?

Solution: We have

$$g_{\tilde{X}|\tilde{Y}_2}(x | y_2) = \frac{g_{\tilde{X},\tilde{Y}_2}(x, y_2)}{g_{\tilde{Y}_2}(y_2)} = \frac{f_{X,Y_2|Y_1}(x, y_2 | y_1)}{f_{Y_2|Y_1}(y_2 | y_1)} = \frac{\frac{f_{X,Y_1,Y_2}(x, y_1, y_2)}{f_{Y_1}(y_1)}}{\frac{f_{Y_1,Y_2}(y_1, y_2)}{f_{Y_1}(y_1)}} = f_{X|Y_1,Y_2}(x | y_1, y_2)$$

Therefore $\mathbb{E}[\tilde{X} | \tilde{Y} = y_2] = \mathbb{E}[X | Y_1 = y_1, Y_2 = y_2]$. This is used to derive the iterative estimator which is used in the recurrence relation for Kalman Filter. ■

Problem 2

Consider the Kalman filtering problem for the scalar system:

$$X_k = \alpha X_{k-1} + W_k \quad Y_k = hX_k + Z_k$$

as described in class (i.e., $W_k \sim N(0, \sigma_W^2)$ i.i.d, $Z_k \sim N(0, \sigma_Z^2)$, and X_1 are independent). The initial condition is $X_1 \sim N(0, \sigma_{X_1}^2)$. For the numerical exercises below you can assume $\sigma_{X_1}^2 = \sigma_Z^2 = \sigma_W^2 = h = 1$.

- Plot sample paths of the process $\{X_k\}$ for different values of α . Pick a representative set of values of α to show the effect of α on how the sample paths look like. Can you explain qualitatively the effect?
- Let $\hat{X}_k = E[X_k | Y_1, \dots, Y_k]$. For those sample paths of $\{X_k\}$ plotted in part (a), plot in the same figure the sample paths of the estimates $\{\hat{X}_k\}$. What is the qualitative effect of α on the estimation errors?
- Let $\tilde{X}_k = E[X_k | Y_k]$. This is the state estimate based only on the current observation. For the sample paths in (a) and (b), plot the sample paths of $\{\tilde{X}_k\}$ in the same figure as well. How does the difference in the accuracy of the estimators \hat{X}_k and \tilde{X}_k depend on the value of α ? Explain qualitatively.
- Let f_k be the conditional distribution of X_k given the observations up to time k . For your favorite value of α , plot f_k for several values of k to get a feel of how the distribution evolves in time. Do these distributions depend on the random outcome of the experiment? How?
- What happens to the distribution of X_k as $k \rightarrow \infty$? Give a quantitative answer. Does your answer depend on α ? Does your answer depend on $\sigma_{X_1}^2$?
- What happens to the MMSE estimation error σ_k^2 of \hat{X}_k as $k \rightarrow \infty$? Does it converge to zero, a finite non-zero value or infinity? How does your answer depend on α ? An answer supported by numerical evidence together with some analysis would be fine; it doesn't have to be totally rigorous.

Solution:

- (a) Here we have taken the values of α to be $\{-1, 0.8, 1, 1.2\}$ in Figure 1. Here we can see that when the value of α is 1.2 then the sample value increases. And when the value of α is -1 it oscillates around 0. But for $\alpha = 0.8$ the sample values remains close to 0. Therefore the sample values converges when $|\alpha| < 1$ and otherwise diverges.

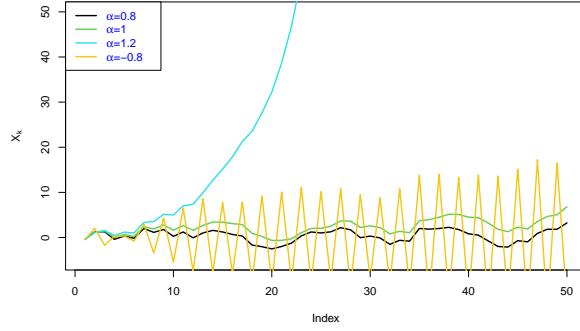
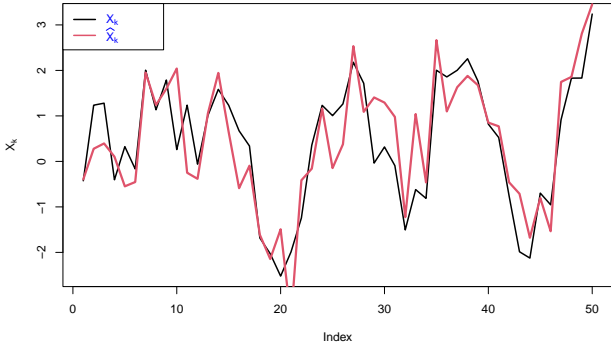
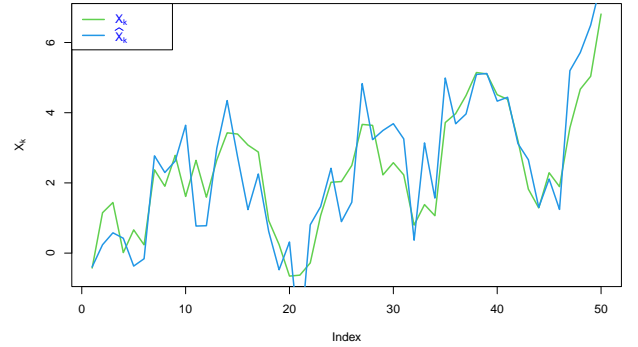


Figure 1: Plot of X_k for different $\alpha \in \{-1, 0.8, 1, 1.2\}$

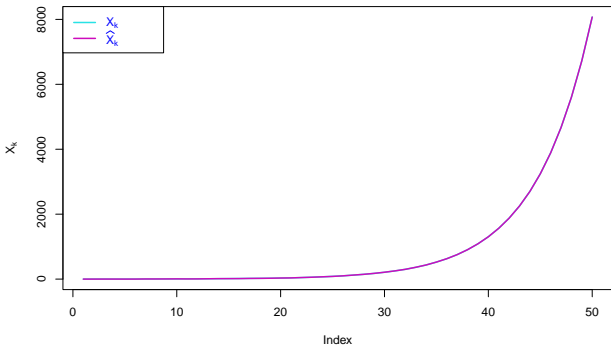
- (b) In the following plot we can see that the predicted values $\mathbb{E}[X | Y_1, \dots, Y_k]$ matches almost correctly with the sample values X_k . From the plots we conclude that as $|\alpha|$ becomes larger it has lesser effect on the estimation which we also showed in part (f) where we showed if $|\alpha|$ becomes larger then the MMSE estimation is independent of α .



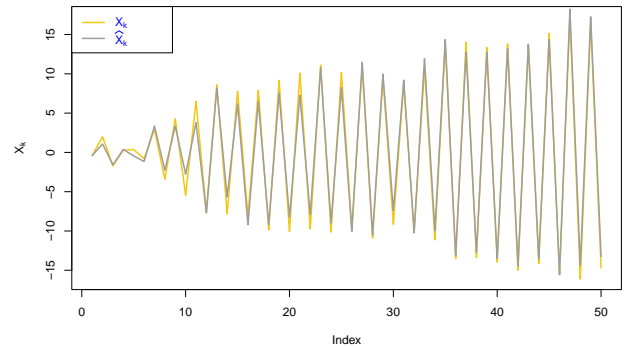
(a) Plot of X_k vs \hat{X}_k for $\alpha = 0.8$



(b) Plot of X_k vs \hat{X}_k for $\alpha = 1$



(c) Plot of X_k vs \hat{X}_k for $\alpha = 1.2$



(d) Plot of X_k vs \hat{X}_k for $\alpha = -1$

Figure 2: Compared X_k and predicted \hat{X}_k

- (c) Here we compare X_k , \hat{S}_k and $\tilde{X}_k = \mathbb{E}[X_k | Y_k]$ for all values of α . Now we have $\text{Cov}(X_k, Y_k) = h\rho_k^2$ and

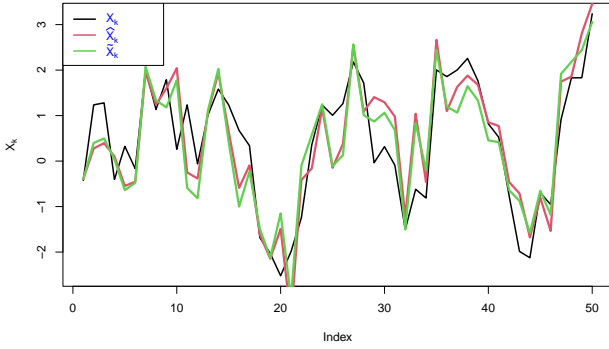
$\text{Var}[Y_k] = h^2 \rho_k^2 + \sigma_Z^2$. Hence we have

$$\mathbb{E}[X_k | Y_k] = \frac{h \rho_k^2 Y_k}{h^2 \rho_k^2 + \sigma_Z^2}$$

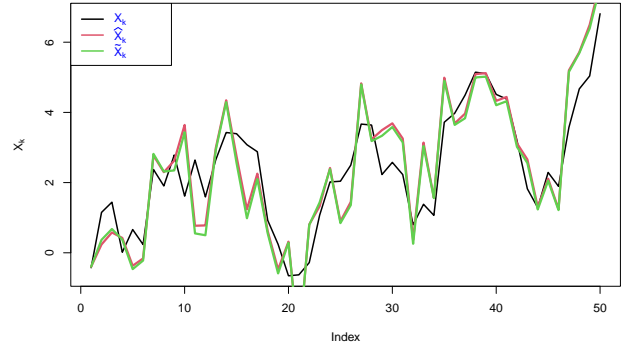
So we have

$$\begin{aligned} \mathbb{E}[X_k - \tilde{X}_k]^2 &= \mathbb{E} \left[X_k - \frac{h \rho_k^2 Y_k}{h^2 \rho_k^2 + \sigma_Z^2} \right]^2 \\ &= \frac{1}{(h^2 \rho_k^2 + \sigma_Z^2)^2} \mathbb{E} [(h^2 \rho_k^2 + \sigma_Z^2) X_k - h \rho_k^2 (h X_k + Z_k)]^2 \\ &= \frac{1}{(h^2 \rho_k^2 + \sigma_Z^2)^2} \mathbb{E} [\sigma_Z^2 X_k - h \rho_k^2 Z_k]^2 \\ &= \frac{\sigma_Z^4 \rho_k^2 + h^2 \rho_k^4 \sigma_Z^2}{(h^2 \rho_k^2 + \sigma_Z^2)^2} = \frac{\sigma_Z^2 \rho_k^2}{h^2 \rho_k^2 + \sigma_Z^2} \end{aligned}$$

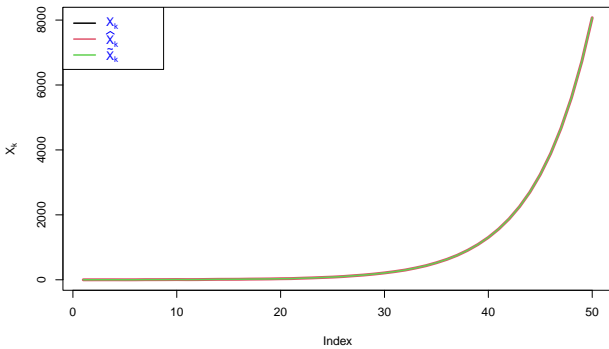
Now this is the MMSE estimation of σ_k^2 of X_k which comparing with part (f) we can see that we obtained the same estimation value. Therefore both \hat{X}_k and \tilde{X}_k are equally good estimating sample values.



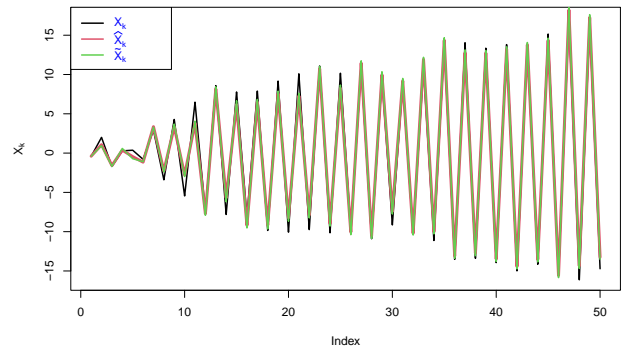
(a) Plot of X_k vs \hat{X}_k vs \tilde{X}_k for $\alpha = 0.8$



(b) Plot of X_k vs \hat{X}_k vs \tilde{X}_k for $\alpha = 1$



(c) Plot of X_k vs \hat{X}_k vs \tilde{X}_k for $\alpha = 1.2$



(d) Plot of X_k vs \hat{X}_k vs \tilde{X}_k for $\alpha = -1$

Figure 3: Compared X_k and predicted \hat{X}_k and \tilde{X}_k

(d) Here we plot f_k for values $k \in [10]$ with $\alpha = 0.8$. Now the conditional distribution $X | Y_1, \dots, Y_k$ approaches the distribution $N\left(0, \frac{\sigma^2}{1-\alpha^2}\right)$ for large k and also we notice from the plot that this doesn't depend on the Y_k .

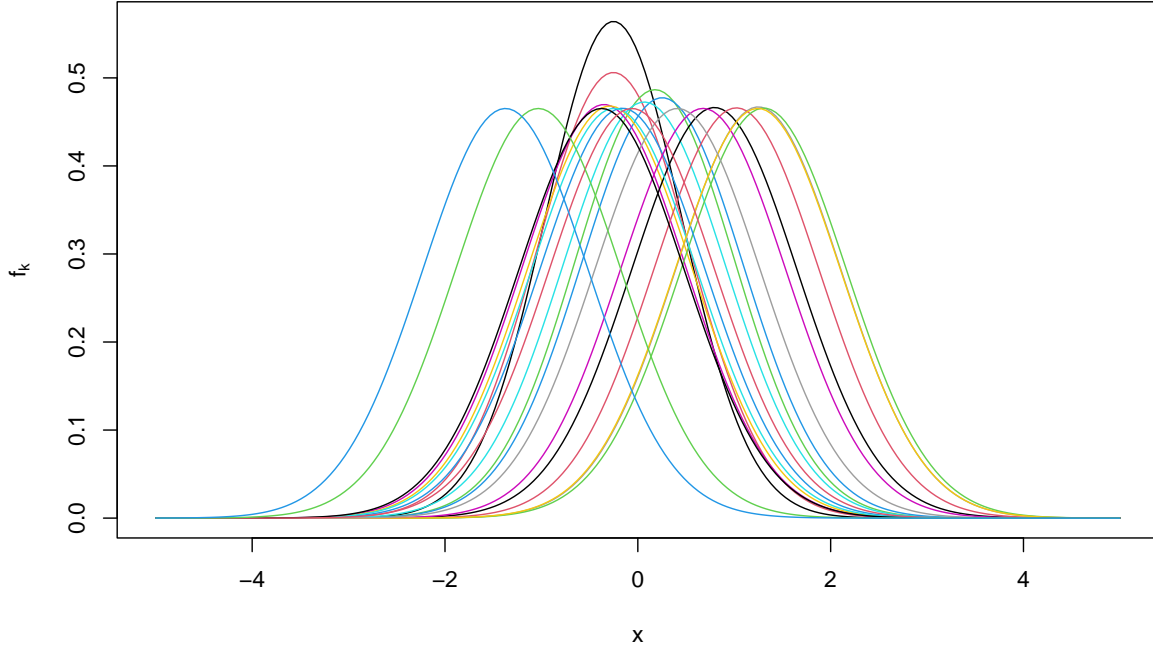


Figure 4: Plot of density f_k of $X_k \mid Y_1, \dots, Y_k$ for $\alpha = 0.8, k \in [10]$

- (e) We will induct on k . Since X_1, W_2 are independent and we have $X_2 = \alpha X_1 + W_2$ hence $X_2 \sim N\left(0, \alpha^2 \sigma_{X_1}^2 + \sigma_W^2\right)$. Now X_{k-1} and W_k are independent and we have $X_k = \alpha X_{k-1} + W_k$. By inductive hypothesis X_{k-1} follows Gaussian Distribution. Hence X_k also follows Gaussian Distribution. Hence $\mathbb{E}[X_k] = 0$. Now we have to calculate $\text{Var}[X_k]$.

$$\text{Var}[X_k] = \alpha^2 \text{Var}[X_{k-1}] + \sigma_W^2 = \alpha^2 (\alpha^2 \text{Var}[X_{k-2}] + \sigma_W^2) + \sigma_W^2 = \dots = \alpha^{2k-2} \sigma_{X_1}^2 + \sigma_W^2 \sum_{i=0}^{k-1} \alpha^{2i}$$

Therefore $X_k \sim N\left(0, \alpha^{2k-2} \sigma_{X_1}^2 + \sigma_W^2 \sum_{i=0}^{k-1} \alpha^{2i}\right)$. Hence if $|\alpha| < 1$, $\lim_{k \rightarrow \infty} \text{Var}[X_k] = \frac{\sigma_W^2}{1-\alpha^2}$. Hence as $k \rightarrow \infty$, $X_k \rightarrow N\left(0, \frac{\sigma_W^2}{1-\alpha^2}\right)$. Now if $|\alpha| \geq 1$, then as $k \rightarrow \infty$, α^{2k-2} diverges. Therefore $\text{Var}[X_k]$ diverges to $+\infty$.

If $|\alpha| < 1$ then $\text{Var}[X_k] = \frac{\sigma_W^2}{1-\alpha^2}$. Hence it doesn't depend on $\sigma_{X_1}^2$.

- (f) Let ρ_k denote the variance of X_k . Then we have the formula

$$\rho_n = \alpha^2 \rho_{n-1}^2 + \sigma_W^2 \quad \text{for } n \geq 2$$

Hence we know the behavior of the conditional variance σ_k^2 of X_k . Hence we know the MMSE of $X_k \mid Y_1, \dots, Y_k$ as $k \rightarrow \infty$. Now we have

$$\sigma_k^2 = \frac{\rho_k^2 \sigma_Z^2}{h^2 \rho_k^2 + \sigma_Z^2} = \frac{\sigma_Z^2}{h^2 + \frac{\sigma_Z^2}{\rho_k^2}}$$

From the previous part if $|\alpha| < 1$ then $\lim_{k \rightarrow \infty} \rho_k = \frac{\sigma_W^2}{1-\alpha^2}$ and if $|\alpha| \geq 1$ then as $k \rightarrow \infty$, ρ_k diverges to $+\infty$.

Therefore when $|\alpha| < 1$, $\lim_{k \rightarrow \infty} \sigma_k^2 = \frac{\sigma_Z^2 \sigma_W^2}{h^2 \sigma_W^2 + (1-\alpha^2) \sigma_Z^2}$ and when $|\alpha| \geq 1$ we have $\lim_{k \rightarrow \infty} \sigma_k^2 = \frac{\sigma_Z^2}{h^2}$.

■

Problem 3

For the system in [Problem 2](#) derive a recursive algorithm for computing the one-step ahead estimator: $\mathbb{E}[X_k | Y_1, Y_2, \dots, Y_k, Y_{k+1}]$. This means we can look ahead one step to estimate the state.

Solution: We have $Y_{k+1} = hX_{k+1} + W_{k+1}$. and $X_{k+1} = \alpha X_k + Z_k$. Therefore combining these two we have

$$Y_{k+1} = \alpha h X_k + (hW_k + Z_{k+1})$$

Now take $T_k = hW_k + Z_{k+1}$. Then $T_k \sim N(0, h^2\sigma_W^2 + \sigma_Z^2)$. Now denote $Y_1^k = (Y_1, \dots, Y_k)$ and denote $y_1^k = (y_1, \dots, y_k)$. Then we have

$$f_{X_k | Y_1^{k+1}}(x | y_1^{k+1}) = \frac{f_{Y_{k+1} | X_k, Y_1^k}(y_{k+1} | x_k, y_1^k)}{f_{Y_{k+1} | Y_1^k}(y_{k+1} | y_1^k)} f_{X_k | Y_1^k}(x_k | y_1^k)$$

Combining this with $\alpha h X_k + (hW_k + Z_{k+1})$ we can write

$$\mathbb{E}[X_k | Y_1^{k+1}] = h\alpha \mathbb{E}[X_{k-1} | Y_1^k]$$

■