Skoltech

Skolkovo Institute of Science and Technology

Experimental Data Processing

MA060238

Comparison of the exponential and running mean for random walk model

Authors:

Luca Breggion
Irina Yareshko
Ivan Kozhin

October 11, 2022

1 Introduction

The objective of this laboratory work is to compare the errors of exponential and running mean to choose the most effective quasi-optimal estimation method in conditions of uncertainty.

2 Determination of optimal smoothing constant in exponential mean

2.1 Generate a true trajectory and the measurements of this true trajectory

Starting from the initial condition $X_1 = 10$ and the given variances $\sigma_w^2 = 12$, $\sigma_\eta^2 = 9$ it is possible to determine the respectively normally distributed random noise with zero mathematical expectations, as follows:

$$w_i = \sqrt{\sigma_w^2} \cdot randn(n, 1) \quad \eta_i = \sqrt{\sigma_\eta^2} \cdot randn(n, 1) \tag{1}$$

where the built-in *Matlab* function *randn* creates a vector of n random values. During this study two cases are considered: $n = 300$ points and $n = 3000$ points.
Furthermore, the true trajectory X_i using the random walk model is determined as:

$$X_i = X_{i-1} + w_i \tag{2}$$

and the generated measurements z_i of the process X_i are obtained:

$$z_i = X_i + \eta_i \tag{3}$$

2.2 Identify σ_w^2 and σ_η^2 using identification method

In order to determine the noise statistics σ_w^2 and σ_η^2, the identification method is used. As a first step, the residuals ρ_i and v_i were calculated:

$$\rho_i = w_i + w_{i-1} + \eta_i - \eta_{i-2} \qquad v_i = w_i + \eta_i - \eta_{i-1} \tag{4}$$

subsequently, the Math. expectations of the respective residual are:

$$E_{v_i^2} \approx \frac{1}{N-1} \sum_{k=2}^{N} v_k^2 \qquad E_{\rho_i^2} \approx \frac{1}{N-2} \sum_{k=3}^{N} \rho_k^2 \tag{5}$$

and solving the following system, **Equation 6**, consisting of two unknowns and two equations, it is possible to retrieve σ_w^2 and σ_η^2 for both trajectory sizes.

$$\begin{cases} \sigma_w^2 + 2\sigma_\eta^2 - E_{v_i^2} = 0 \\ 2\sigma_w^2 + 2\sigma_\eta^2 - E_{\rho_i^2} = 0 \end{cases} \tag{6}$$

The results obtained are listed in the **Table 1**, and it is observed that as the number of steps increase, the variance values get closer and closer to those given.

Steps number	300	3000
σ_w^2	11.9738	13.9473
σ_η^2	7.6658	7.4674

Table 1: Noise statistics σ_w^2 and σ_η^2 for 300 and 3000 steps

2.3 Determine optimal smoothing coefficient in exponential smoothing

Determined the coefficient χ as:

$$\chi = \frac{\sigma_w^2}{\sigma_\eta^2} \tag{7}$$

it is feasible to calculate the optimal α for random walk model with **Equation 8** and get the final results listed in **Table 2**.

$$\alpha = \frac{-\chi + \sqrt{\chi^2 + 4\chi}}{2} \tag{8}$$

Steps number	300	3000
χ	0.6928	0.7214

Table 2: Values of χ for two cases

2.4 Perform exponential smoothing with the determined smoothing coefficient

Exponential smoothing is a technique used for forecasting where the forecast is made through the exponentially weighted average of prior observations. It can be written as:

$$\widehat{X}_i = \alpha z_i + (1-\alpha)\widehat{X}_{i-1}, \tag{9}$$

where $\widehat{X_i}$ is a smoothed estimate at time i, α is a smoothing coefficient that takes values from 0 to 1 and z_i is a measurement at time i. The results of exponential smoothing performed in Matlab are shown in **Figure 1** and **Figure 2** for two cases with different number of steps.

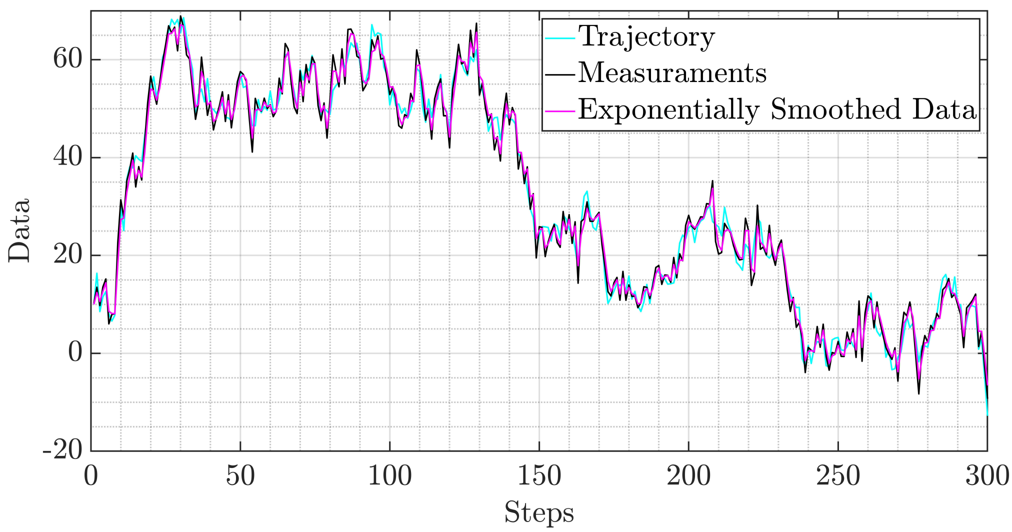

Figure 1: Exponential Smoothing with 300 steps

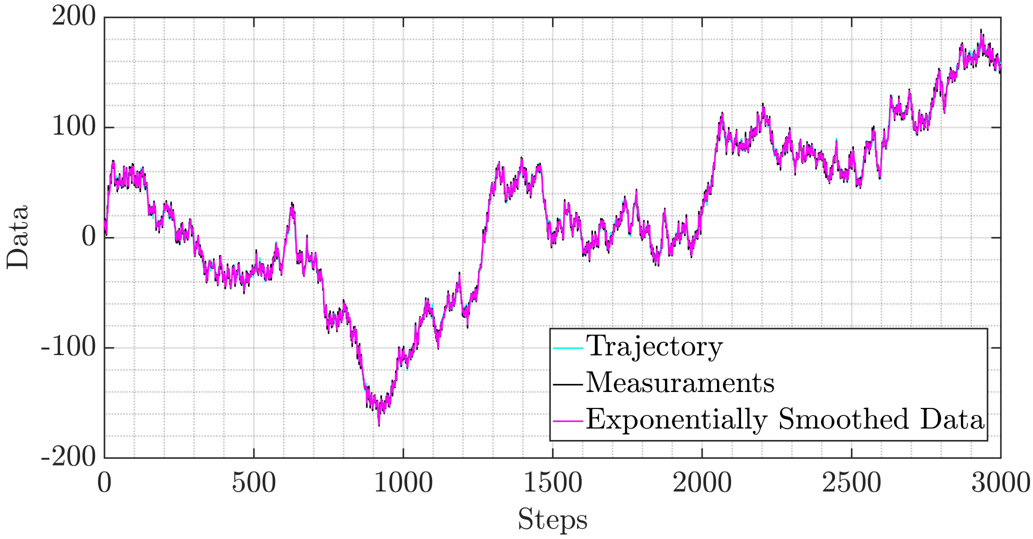

Figure 2: Exponential Smoothing with 3000 steps

To compare the accuracy of the method for two cases **Figure 3** and **Figure 4** are presented with identical scales.

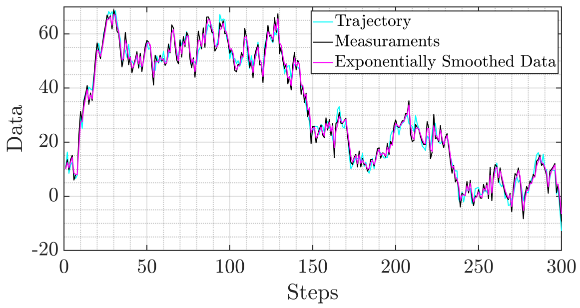

Figure 3: Exponential Smoothing with 300 steps

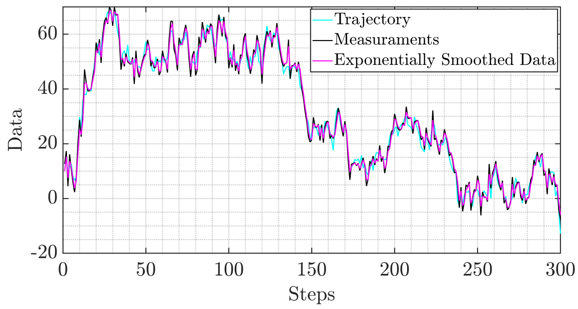

Figure 4: Exponential Smoothing with 3000 steps

In addition to the figures, the Root Mean Squared Error ($RMSE$) was calculated:

$$RMSE = \sqrt{\frac{(\sum_{i=1}^{N}(X_i - \widehat{X}_i)^2}{N}} \quad (10)$$

For two cases with different number of steps values of RMSE are presented in the **Table 3**.

Steps number	300	3000
$RMSE$	1.543	1.441

Table 3: $RMSE$ values for two cases

It is observed from the figures and from calculations of the Root Mean Squared Error that with the increase in the number of steps, the accuracy of exponential smoothing grows respectively. We can conclude that exponential smoothing method is a good technique for forecasts. However, as it assigns exponentially decreasing weights for newest to oldest observations, exponential smoothing does not always handle trends well, which makes it applicable best to short-term forecasts in the absence of seasonal or cyclic variations.

3 Comparison of methodical errors of exponential and running mean

3.1 Generate a true trajectory, measurements and determine smoothing coefficients

The true trajectory and measurements are generated similarly as in the previous section with accordance to **Equations 1 - 3**, with initial condition $X_1 = 10$ and variance of noises $\sigma_w^2 = 28^2$, $\sigma_\eta^2 = 97^2$. Smoothing coefficient $\alpha = 0.25$ is obtained from **Equations 7 - 8**.

3.2 Determine the window size M

The component of full error that is related to measurement errors for running mean is determined as:

$$\sigma_{RM}^2 = \frac{\sigma_\eta^2}{M} \tag{11}$$

The component of full error that is related to measurement errors for exponential smoothing is determined as:

$$\sigma_{ES}^2 = \sigma_\eta^2 \frac{\alpha}{2-\alpha} \tag{12}$$

The window size M that provides equality of the running mean variance, σ_{RM}^2, and the exponential smoothing, variance σ_{ES}^2, is:

$$M = \frac{2-\alpha}{\alpha} = 7 \tag{13}$$

4 Apply running mean using determined window size M and exponential mean using determined smoothing constant α to measurements z_i

A running mean is a number that shows a middle or normal value for a set of data and continually changes as more data points are collected. The expression for running mean can be written as follows:

$$\widehat{X}_i = \frac{1}{M} \sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_k \tag{14}$$

Exponential mean is performed with accordance to Equation (9)

Figures below show the result of applying running mean and exponential smoothing to the data.

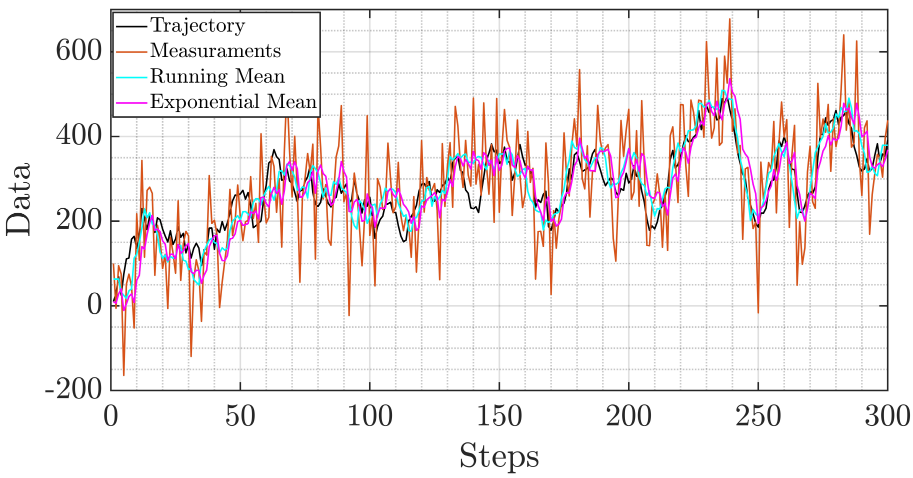

Figure 5: Comparison of Exponential Smoothing and Running Mean

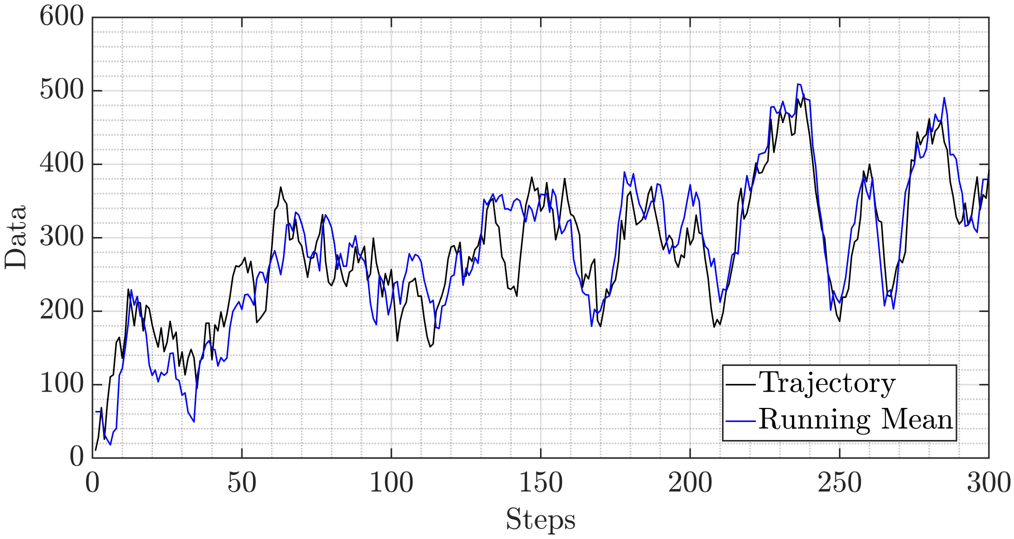

Figure 6: Running Mean

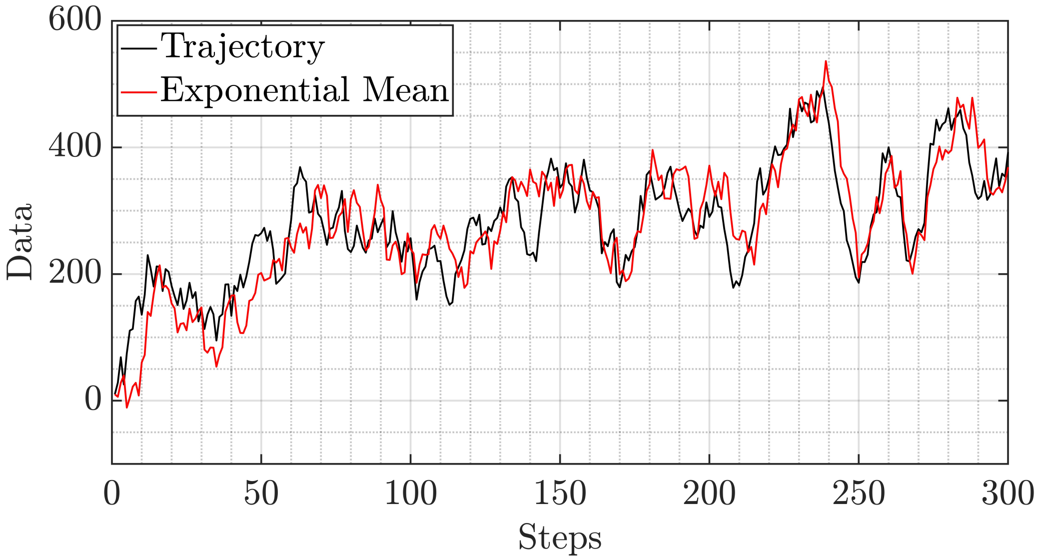

Figure 7: Exponential Smoothing

It is observed from the figures above that running mean provides a better smoothing as the offset and deviation is much less for running mean than for exponential smoothing.

5 Conclusion

In this assignment we studied and compared two different smoothing data techniques, which are running mean method and exponential smoothing, and learnt the methods to determine their optimal smoothing constants. We learnt that exponential smoothing method is a good technique for forecasts, that assigns exponentially decreasing weights for newest to oldest observations. This makes it applicable best to short-term forecasts in the absence of seasonal or cyclic variations.
Running mean performs forecasts by capturing the average changes in the data. The main difference between these methods is that running mean method considers all the point having equal weights, while exponential smoothing method gives more weight to the newest points for forecasting.
We learnt that it is important to choose the smoothing coefficients correctly to get both an acceptable level of filtration errors and reaction to changes in a system dynamics. For example, for exponential smoothing

coefficient α and for running mean coefficient M no filtration of error is performed when $\alpha = 1$ and $M = 1$, and very effective error filtration but no reaction to changes in dynamics when $\alpha \to 1$ and $M \to \infty$.
For the data case studied in this assignment the running mean showed better performance having lower deviation and offset compared to exponential smoothing that are presented in **Figure 6** and **Figure 7**.

Individual contribution:

- Luca Breggion: Matlab code, report
- Irina Yareshko: Matlab code, report
- Ivan Kozhin: Report