Skoltech

Skolkovo Institute of Science and Technology

Experimental Data Processing

MA060238

Determining and removing drawbacks of exponential and running mean. Task 1

Authors:

Luca Breggion

Irina Yareshko

October 12, 2022

1 Introduction

The aim of this work is to determine conditions for which running and exponential mean methods provide effective solution and conditions under which they break down.

2 Backward exponential smoothing

2.1 Apply backward exponential smoothing to forward exponential estimates to further smooth measurement errors

Working on Assignment 2, it was possible to observe that the exponential smoothing results showed a delay of estimations. The optimal smoothing coefficient in exponential smoothing obtained was $\alpha = 0.25$ and the window size M that provides equality of the running mean variance and the exponential smoothing was $M = 7$, considering a set of 300 points.

Therefore, in order to smooth the results even more, a backward exponential smoothing, **Equation 2**, is applied to the forward exponential estimates, **Equation 1**.

$$X_i^f = X_{i-1}^f + \alpha(z_i - X_{i-1}^f), \quad for\ i = 2, ..., N \tag{1}$$

$$X_i^b = X_{i+1}^b + \alpha(X_i^f - X_{i+1}^b), \quad for\ i = N-1, ..., 1 \tag{2}$$

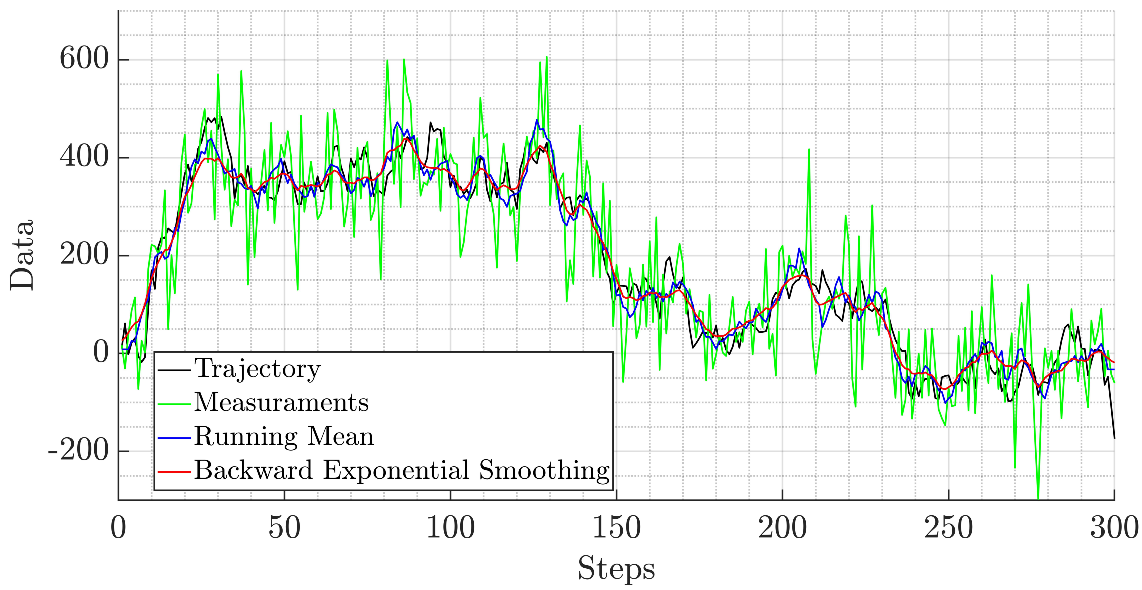

Figure 1: Comparison of the Running mean and Backward exponential smoothing

As shown in **Figure 1**, the backward exponential smoothing provides better smoothing results than running mean.

In order to verify the validity of the smoothing procedure and to compare the running and backward exponential smoothing, two indicators are used:

- Deviation indicator:

$$I_d = \sum_{i=1}^{N} (z_i - \widehat{X}_i)^2 \tag{3}$$

- Variability indicator:

$$I_v = \sum_{i=1}^{N-2} (\widehat{X}_{j+2} - 2\widehat{X}_{j+1} + \widehat{X})^2 \tag{4}$$

the results obtained are listed in the **Table 1**.

Method	Running Mean	Backward Exp. Smoothing
I_d	$2.1541 \cdot 10^6$	$2.0599 \cdot 10^6$
I_v	$1.9650 \cdot 10^5$	$1.4289 \cdot 10^4$

Table 1: Deviation and Variability indicators for Running Mean and Backward Exponential Smoothing

An additional evidence is given by the deviation and variability indicators: since the backward exponential values are lower better than the running mean ones.

3 Drawbacks of running mean. First trajectory

3.1 Generate a true trajectory X_i of an object motion disturbed by normally distributed random acceleration

As a first step, the actual trajectory X_i was determined (**Equation 5**) considering:

- Size of the trajectory 300 *points*
- Initial conditions: $X_1 = 5$, $V_1 = 0$, $T = 0.1$

$$X_i = X_{i-1} + V_{i-1}T + \frac{a_{i-1}T^2}{2} \tag{5}$$

where V_i is evaluated as:

$$V_i = V_{i-1} + a_{i-1}T \tag{6}$$

and a_i is a random vector from normal distribution with zero mean and given variance $\sigma_a^2 = 10$. Similarly, to generate the measurements z_i of the process X_i, it is considered a variance of $\sigma_n^2 = 500$.

$$z_i = X_i + \eta_i \tag{7}$$

3.2 Determine empirically the window size M and smoothing coefficient α that provide the best estimation of the process X_i using measurements z_i

It is known that the true trajectory is close to a line, however the measurements have a huge noise component. It gives us a hint that we should look for quite a small value of coefficient α to obtain acceptable smoothing and it is reasonable to have a bigger window width M for running mean.
In order to choose the right window size M of the running mean and the optimal coefficient α of exponential smoothing, the plots were built for different values of these coefficients.

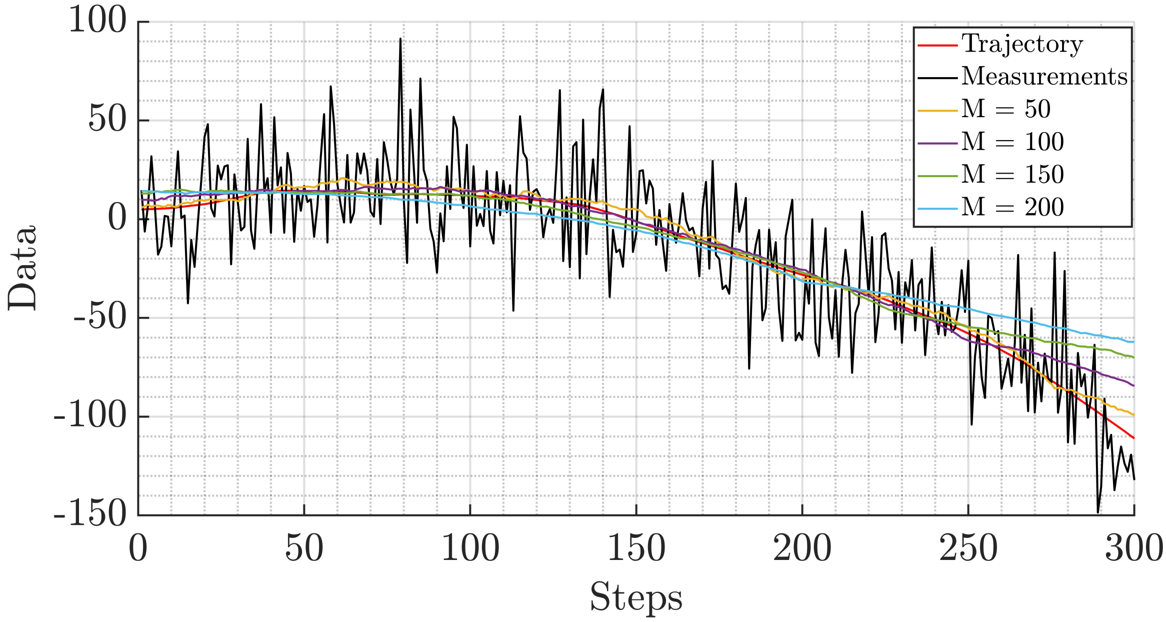

Figure 2: Running Mean with different values of M

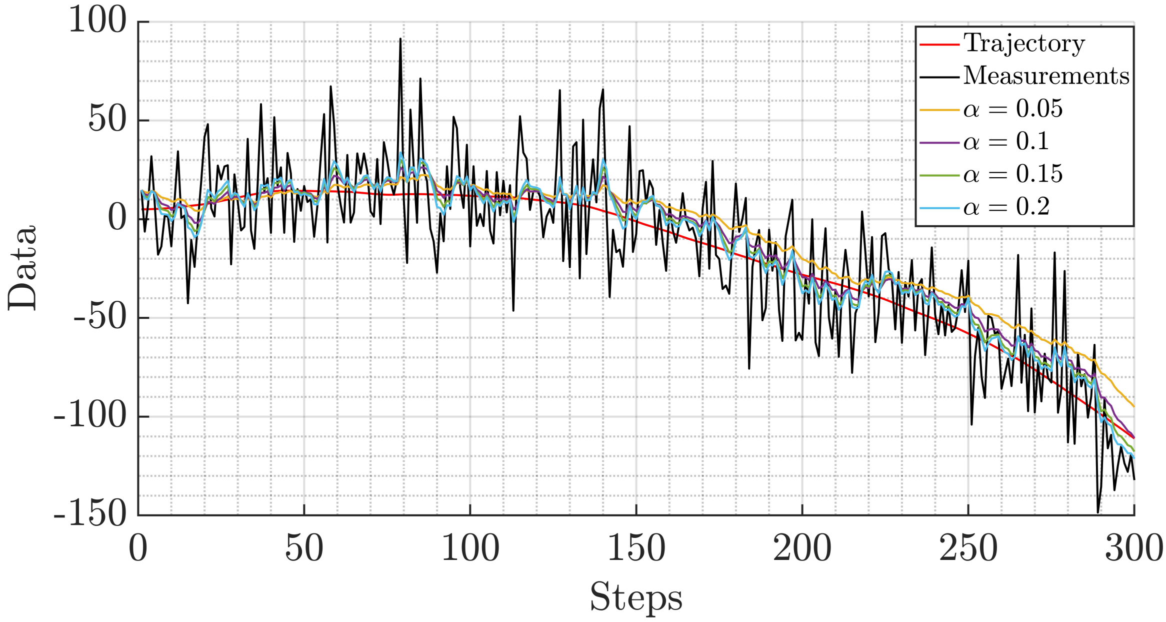

Figure 3: Exponential smoothing with different values of α

It is seen from the **Figure 2** and **Figure 3** that the optimal coefficients are: $M = 50$ for the running mean and $\alpha = 0.05$ for the exponential smoothing.

3.3 Chose better smoothing method using deviation and variability indicators

Similarly to **Section 2.1**, Deviation and Variation indicators were obtained, using **Equation 3** and **Equation 4** respectively. The results of the calculations are presented in the **Table 2**.

Method	Running mean	Backward exp. smoothing
I_d	$1.600 \cdot 10^5$	$1.799 \cdot 10^5$
I_v	264.2	822.1

Table 2: Deviation and Variability indicators for Running Mean and Backward Exponential Smoothing

According to the indicators value, Running Mean method demonstrates a better smoothing compared to Exponential Smoothing, as both Deviation and Variation indicators are significantly larger for the latter one.

4 Drawbacks of running mean. Second trajectory

4.1 Generate a cyclic trajectory X_i

The cyclic trajectory was generated according to the following **Equation 8** and **Equation 9** with initial condition $A_1 = 1$.

$$X_i = A_i sin(\omega i + 3) \tag{8}$$

$$A_i = A_{i-1} + w_i \tag{9}$$

Using the given variance $\sigma_w^2 = 0.08^2$ it is possible to determine the normally distributed random noise with zero mathematical expectations, as follows:

$$w_i = \sqrt{\sigma_w^2} \cdot randn(n, 1), \tag{10}$$

where the built-in *Matlab* function *randn* creates a vector of n random values. Size of the trajectory $n = 200$ points is given.
Given the periods of oscillations $T = 32$ steps, it is possible to evaluate the angular frequency ω as:

$$\omega = \frac{2\pi}{T} = 0.1963\ rad/s \tag{11}$$

4.2 Generate measurements z_i of the process X_i

Similarly to the previous section, following the **Equation 10**, η_i is obtained considering the variance $\sigma_\eta^2 = 0.05$. Therefore, the measurements z_i are generated:

$$z_i = X_i + \eta_i \tag{12}$$

Next, the Running Mean, **Equation 13**, with window size $M = 13$ is applied to the measurements z_i

$$\widehat{X}_i = \frac{1}{M} \sum_{k=i-\frac{M-1}{2}}^{i+\frac{M-1}{2}} z_k \tag{13}$$

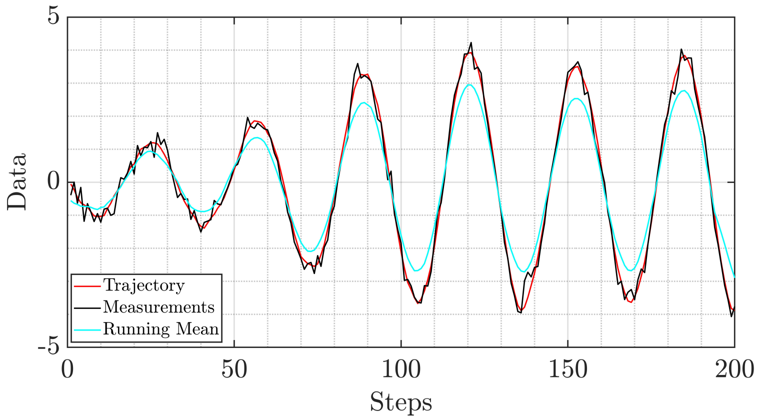

Figure 4: Running Mean with $M = 13$ for a cyclic trajectory

4.3 Determine the period of oscillations for different cases

Determine the periods of oscillations for which running mean with a given window size $M = 19$

- a) produces inverse oscillations

It was obtained empirically that running mean produces inverse oscillations when the value of period of oscillations T is smaller that the value of window size M. It is illustrated by **Figure 5** for $T = 15$ and $M = 19$

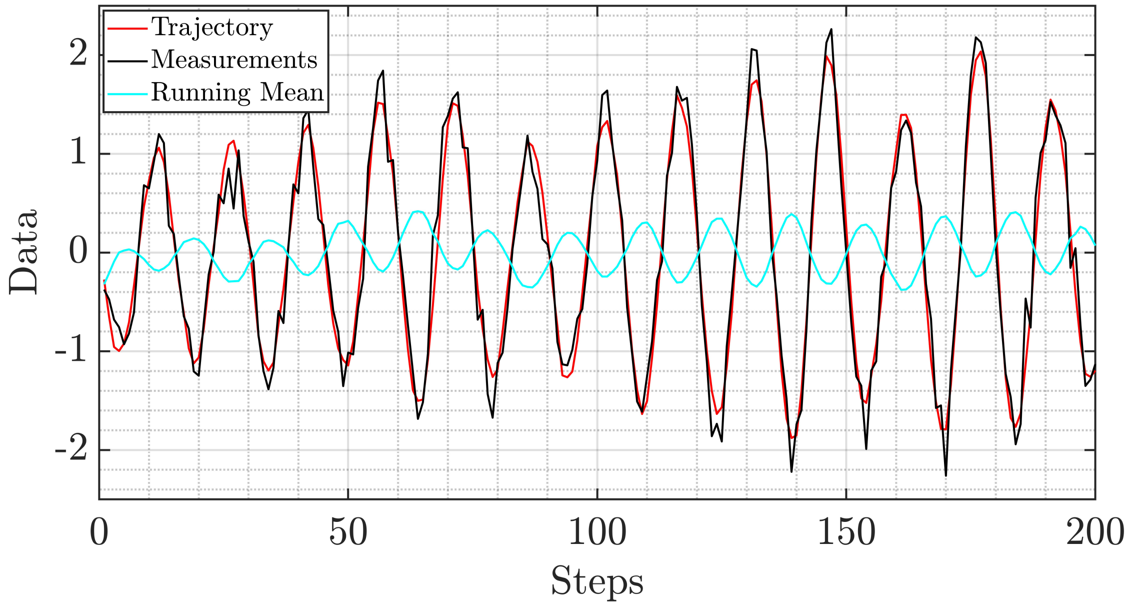

Figure 5: Running Mean with T = 15, M = 19

- b) leads to the loss of oscillations (zero oscillations)

It was obtained empirically that running mean leads to the loss of oscillations (or produce zero oscillations) when the value of period of oscillations T is equal to the value of window size M. It is illustrated by **Figure 6** for $T = 19$ and $M = 19$

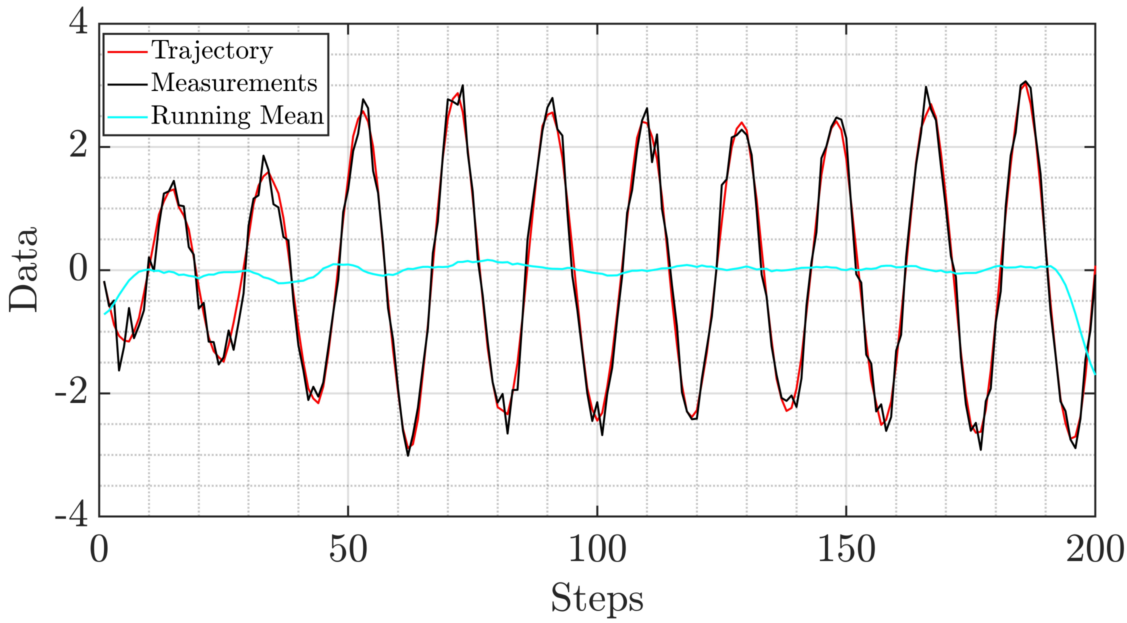

Figure 6: Running Mean with T = 19, M = 19

- c) changes the oscillations insignificantly

It was obtained empirically that running mean changes the oscillations insignificantly when the value of period of oscillations T is bigger than the value of window size M. It is illustrated by **Figure 7** for $T = 25$ and $M = 19$

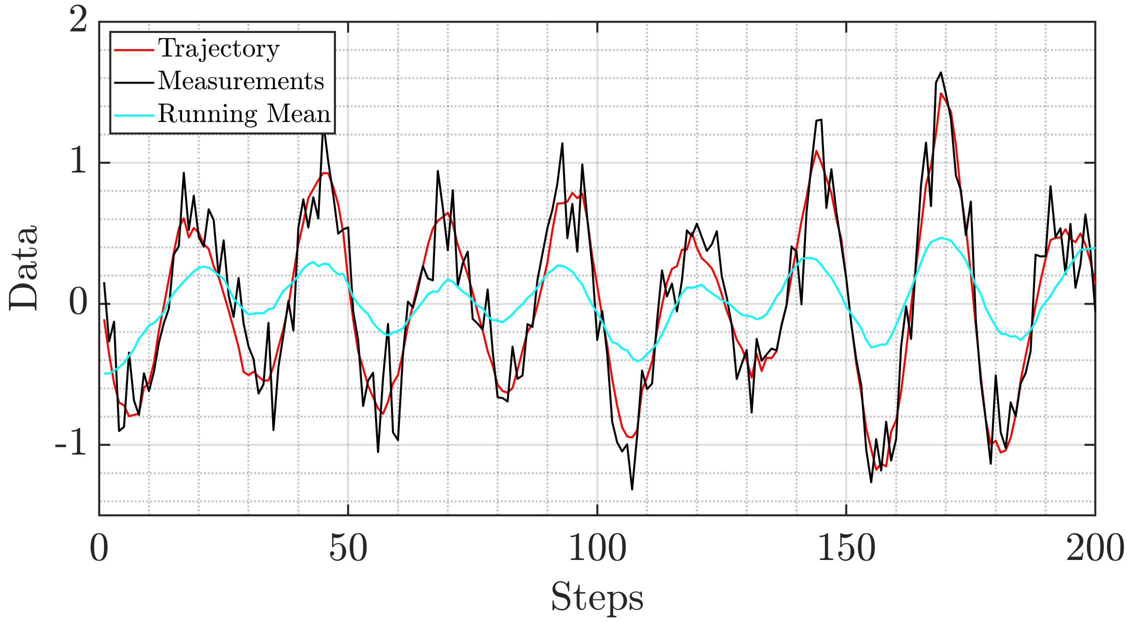

Figure 7: Running Mean with T = 25, M = 19

5 Conclusion

By the end of this assignment, we learnt how to determine conditions for which Running and Exponential Mean methods provide effective solution and conditions under which they break down.

We found out by experience, that the offset and deviation of forward exponential can be reduced by application of backward exponential smoothing on top of the result. This allows for forward-backward exponential smoothing to exceed the running mean in accuracy.

We learnt that for different systems different methods fit best and how to choose the most appropriate technique by calculation deviation and variation coefficients. As well as that, we gained as understanding of generation of cyclic functions and reflected upon the correlation between smoothing coefficient and the period for such type of systems.

For the running mean method with a given window width M the period of oscillations T produces inverse oscillations if $M > T$, leads to the loss of oscillations if $M = T$, and changes the oscillations insignificantly if $M < T$.

Individual contribution:

- Luca Breggion: Matlab code, report
- Irina Yareshko: Matlab code, report