Assignment 2

Comparison of the exponential and running mean for random walk model

By

Group 19

Ahmed Baza, Skoltech, 2020

Ahmed Elgohary, Skoltech, 2020

Mohamed Abbas, Skoltech, 2020

Ziad Baraka, Skoltech, 2020

Part I: Determination of optimal smoothing constant in exponential mean.

1.1-firts we generate two trajectories: X_i using the random walk model and noise Z_i with $\sigma_w^2, \sigma_\eta^2$ = 13, 8 accordingly. one with size 3000 and the other with size 300 to compaire which one will estimate better results of the error variances.

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

$Z_i = X_i + \eta_i$ (2)

also gentrating the residuals ρ_i, ν_i used in identification method.

$\nu_i = Z_i - Z_{i-1}$ (3)

$\rho_i = Z_i - Z_{i-2}$ (4)

```
clc
clear
Axis=1:3000; Axis2=1:300; % Arrays to plot the data aganist.
X=1:3000; Z=1:3000; V=1:2999; Rho=1:2998; X(1)=10; Z(1)=10+normrnd(0,sqrt(8));   % initializing
X2=1:300; Z2=1:300; V2=1:299; Rho2=1:298; X2(1)=10; Z2(1)=10+normrnd(0,sqrt(8)); % initializing
N=length(X); N2=length(X2);
    for i=2:3000
    X(i)=X(i-1)+normrnd(0,sqrt(13));
    Z(i)=X(i)+normrnd(0,sqrt(8));
    V(i-1)=Z(i)-Z(i-1);   %Residual
    end
    for i=3:length(Z)
    Rho(i-2)=Z(i)-Z(i-2); % Residual
    end
    for i=2:300
    X2(i)=X2(i-1)+normrnd(0,sqrt(13));
    Z2(i)=X2(i)+normrnd(0,sqrt(8));
    V2(i-1)=Z2(i)-Z2(i-1);   %Residual
    end
    for i=3:length(Z2)
    Rho2(i-2)=Z2(i)-Z2(i-2); % Residual
    end
```

1.2 Identify σ_w^2 and σ_η^2 using identification method to compaire them and find which size approximate them closer to the true values (13,8):

$$E\left[v_i^2\right] = \sigma_w^2 + 2\sigma_\eta^2 \qquad (5)$$

$$E\left[\rho_i^2\right] = 2\sigma_w^2 + 2\sigma_\eta^2 \qquad (6)$$

```
E_V=(1/(N-1))*sum(V.^2);       % Expectance of Residual V
E_Rho=(1/(N-2))*sum(Rho.^2);   % Expectance of Residual Rho
Sigma2_W= E_Rho-E_V           % from Eq6 - Eq5 we get this relation
```

```
Sigma2_W = 12.7995
```

```
Sigma2_eta=0.5*(E_V-Sigma2_W) % direct sub in Eq5.
```

```
Sigma2_eta = 7.7933
```

```
E_V2=(1/(N2-1))*sum(V2.^2);    % Expectance of Residual v
E_Rho2=(1/(N2-2))*sum(Rho2.^2);% Expectance of Residual Rho
Sigma2_W2= E_Rho2-E_V2        % from Eq6 - Eq5 we get this relation
```

```
Sigma2_W2 = 15.6151
```

```
Sigma2_eta2=0.5*(E_V2-Sigma2_W2) % direct sub in Eq5.
```

```
Sigma2_eta2 = 4.9453
```

As we can see the trajectory with the bigger size got a closer approximation of the true error variances of our random walk trajectory $\sigma_w^2, \sigma_\eta^2$ = (13, 8) . so, more data we have the better we can estimate the error and hence the better we can estimatet he optimal smoothing coefficient.

1.3 Determine optimal smoothing coefficient in exponential smoothing:

$$\chi = \frac{\sigma_w^2}{\sigma_\eta^2} \qquad (7)$$

$$\alpha = \frac{-\chi + \sqrt{\chi^2 + 4\chi}}{2} \qquad (8)$$

```
chi=Sigma2_W/Sigma2_eta;                    % Xai
alpha=0.5*(-chi+sqrt(chi^2+4*chi))          % Calculating alpha for trajectory 1.
```

```
alpha = 0.7009
```

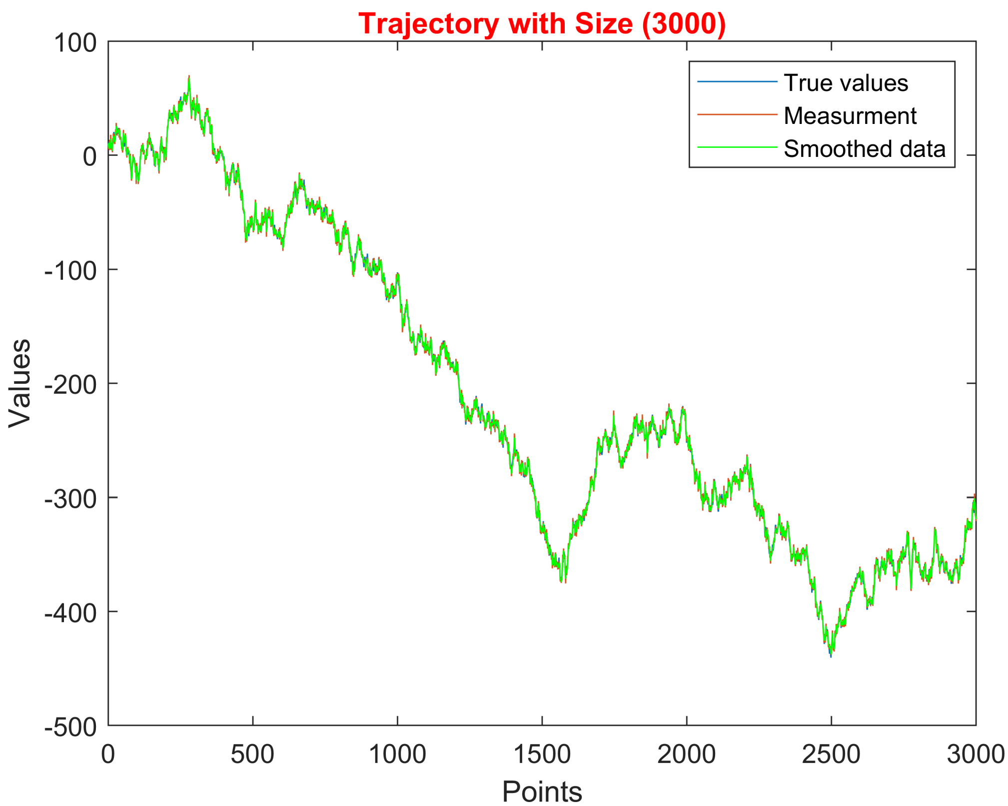
```
chi2=Sigma2_W2/Sigma2_eta2;                 % Xai
alpha2=0.5*(-chi2+sqrt(chi2^2+4*chi2))      % Calculating alpha for trajectory 2.
```

```
alpha2 = 0.7982
```

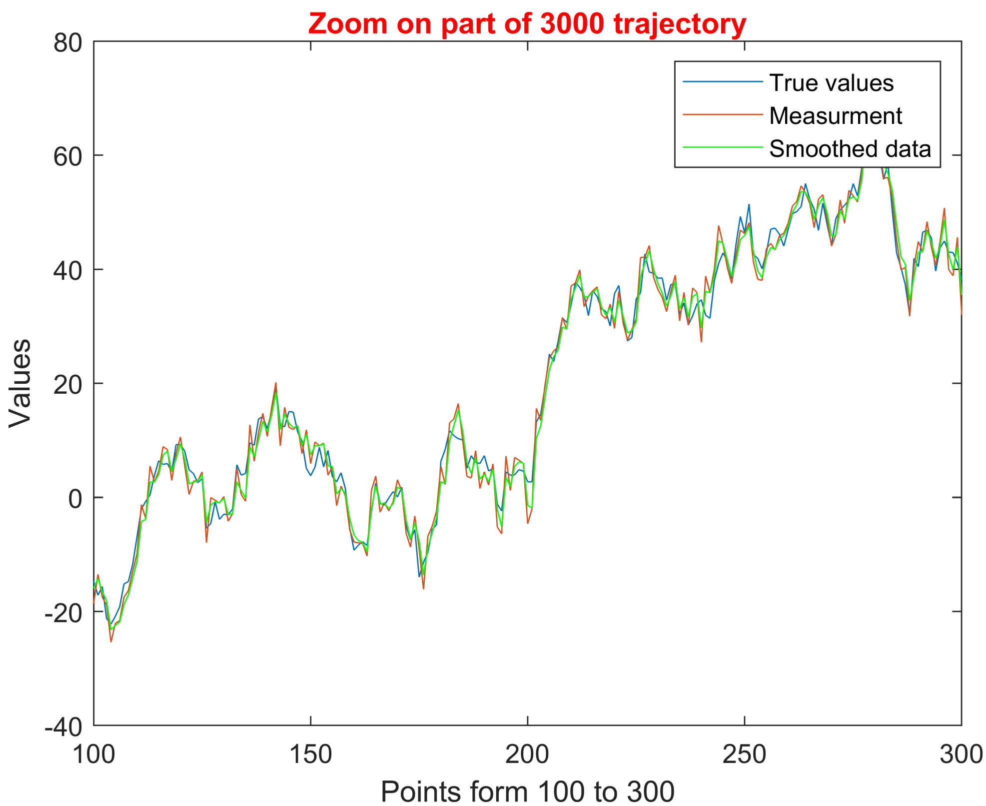
As we see here bothe sizes estimated α greather than 0.5, so we will not have a great smothed curve or the results will not have good filiteration for the measurement noise.

1.4 Perform exponential smoothing with the determined smoothing coefficient. Plot results.

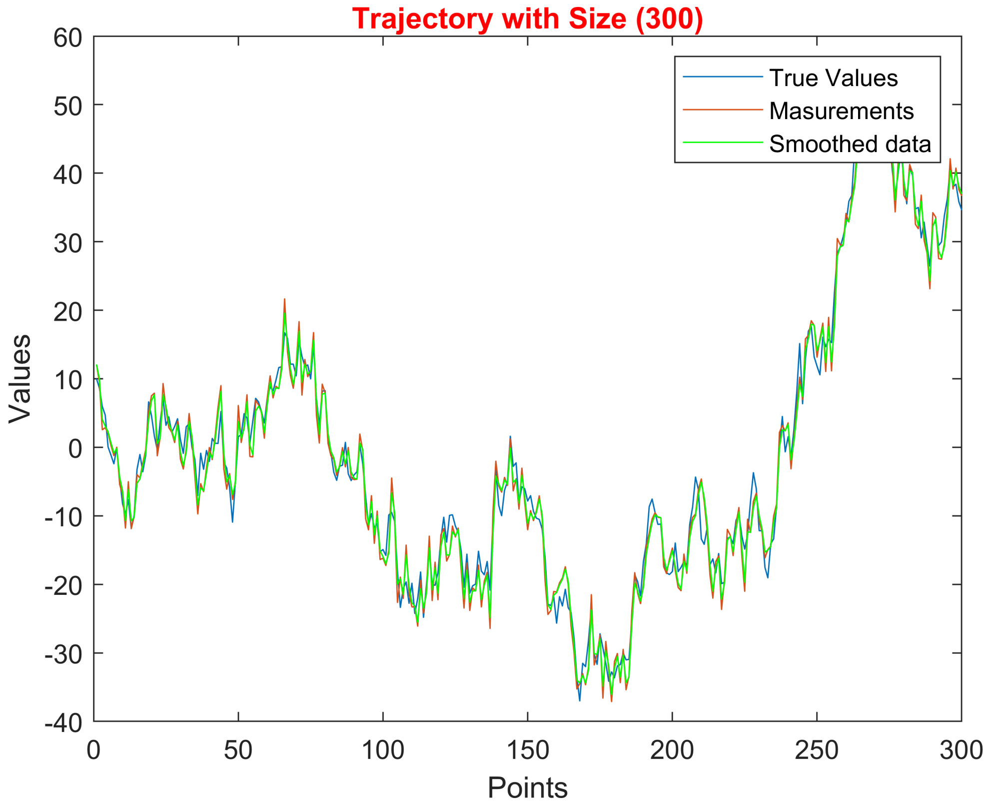
```
XS=Z;   %smoothed X intiation for trajectory 1.
XS2=Z2; %smoothed X intiation for trajectory 2.
for i=2:N
    XS(i)=alpha*Z(i)+(1-alpha)*XS(i-1); % Smoothing
end
for i=2:N2
    XS2(i)=alpha2*Z2(i)+(1-alpha2)*XS2(i-1); % Smoothing
end
plot(Axis,X,Axis,Z,Axis,XS,'g')
legend('True values','Measurment','Smoothed data')
title('Trajectory with Size (3000)','color','r')
xlabel('Points')
ylabel('Values')
```




```
plot(Axis,X,Axis,Z,Axis,XS,'g')
legend('True values','Measurment','Smoothed data')
title('Zoom on part of 3000 trajectory','color','r')
xlabel('Points form 100 to 300')
ylabel('Values')
xlim([100 300])
```




```
plot(Axis2,X2,Axis2,Z2,Axis2,XS2,'g')
legend('True Values','Masurements','Smoothed data')
title('Trajectory with Size (300)','color','r')
xlabel('Points')
ylabel('Values')
```



<u>After plotting the 3 curves together (the smoothed, the measurements and the true values curve):</u>

It can be clearly seen that the smoothed curve is closer to the measurements rather than the true values. Rerunning this little experiment multiple times gives the same kind of behavior, and that was clearly due to the value of alpha as it's been between 0.5 and 1, and as we studied, the more the value approaches 1 the more it follows the behavior of the measurements with no filtration of the noise. Hence with great values like 0.7 and 0.8 it should be expected that the smoothed curve be more similar to the measurements. And the value of the alpha was higher surely due to the values of our error variance And this behavior was similarly seen in both trajectories, the 300 points and the 3000 points, and in each run they both had close values of alphas.

Part II: Comparison of methodical errors of exponential and running mean.

2.1 generate trajectory X_i using the random walk model and noise Z_i with $\sigma_w^2, \sigma_\eta^2$ = $28^2, 97^2$ accordingly with size 300 .

$X_i = X_{i-1} + \omega_i$ (7)

$Z_i = X_i + \eta_i$ (8)

```
X3=1:300; Z3=1:300;  % New trajectory
eta=normrnd(0,97,1,300); W=normrnd(0,28,1,300);
X3(1)=10+W(i)  ;Z3(1)=10+eta(1);
for i=2:300
    X3(i)=X3(i-1)+W(i);
    Z3(i)=X3(i)+eta(i);
```

```
end
```

2.2 Determine optimal smoothing coefficient α using equations (7,8):

```
SW= 28^2;              %Sigma_W
Seta=97^2;             %Sigma_eta
Chi3=SW/Seta;          %chi
alpha3=0.5*(-Chi3+sqrt(Chi3^2+4*Chi3))
```

```
alpha3 = 0.2500
```

2.3 Determine the window size M:

by equating the componant of full error of both the running mean method and the exponential mean method then solve for M:

```
SES = Seta*(alpha3/(2-alpha3));            % SigmaES Esponinitial Smoothing
M= floor(Seta/SES)                        % M: the Running mean window
```

```
M = 7
```

2.4 Apply running mean using determined window size α and exponential mean:

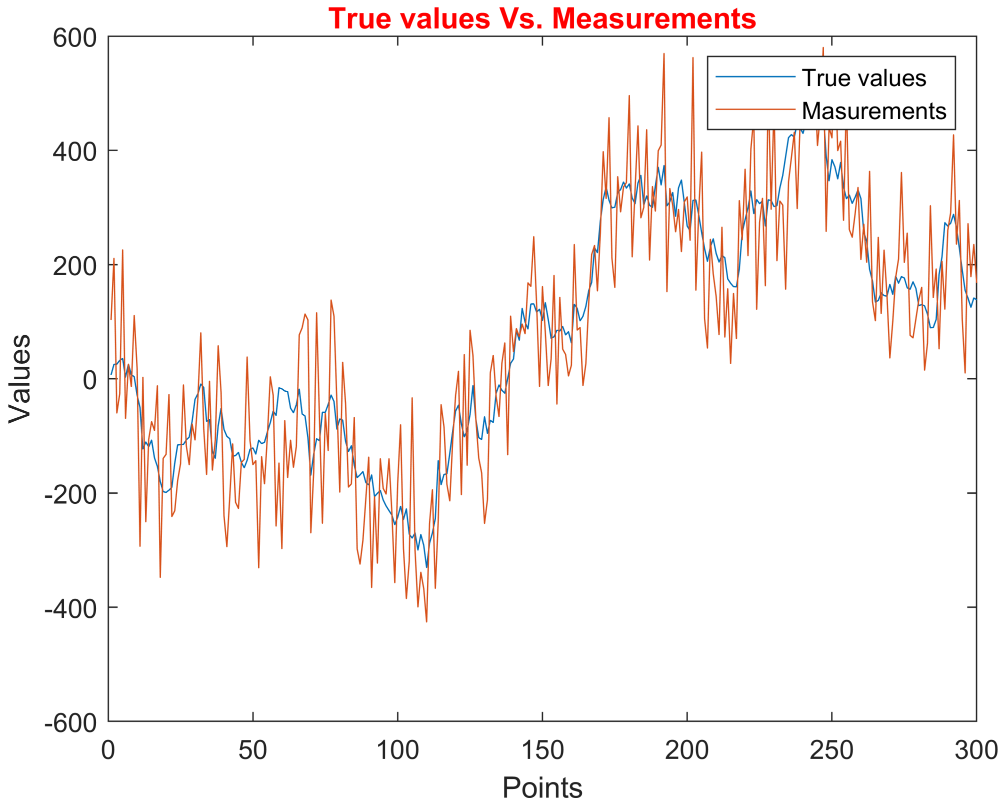
```
XRM =1:N2;                           % Runnig mean smoothed trajectory
XFS=1:N2; XFS(1)=Z3(1);              % Exponential smoothe trajectory.

for i=(M+1)/2:N2-(M+1)/2
    XRM(i)=(1/M)*sum(Z3(i-((M-1)/2):i+((M-1)/2))); % runnig mean smoothing
end
for i=2:N2
    XFS(i)=alpha3*Z3(i)+(1-alpha3)*XFS(i-1);       % Exponential smoothing
end
XRM(1:(M-1)/2)=mean(Z3(1:(M-1)/2));
XRM(N2-((M-1)/2):N2)=mean(Z3(N2-(M-1)/2:N2));
```

2.5 Make visual comparison of results.

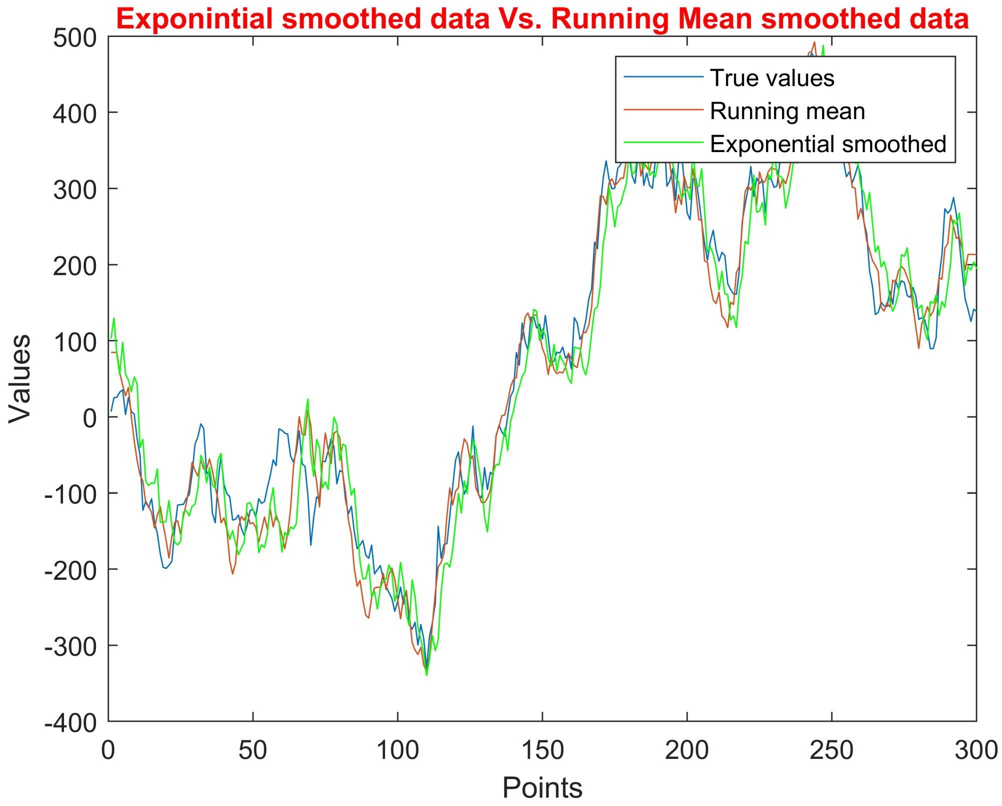
ploting the results to make a comparison between the two methods

```
plot(Axis2,X3,Axis2,Z3)
legend('True values','Masurements')
title('True values Vs. Measurements','Color','r')
xlabel('Points')
ylabel('Values')
```



The graph above shows both the measurements along with the true trajectory and we can see that the measurements extremely noisy due to the high error variance $\sigma_\eta^2 = 97^2$.

```
plot(Axis2,X3,Axis2,XRM,Axis2,XFS,'g')
legend('True values','Running mean','Exponential smoothed')
title('Exponintial smoothed data Vs. Running Mean smoothed data','Color','r')
xlabel('Points')
ylabel('Values')
```



Plotting both the running mean and the exponential against the true values:

We can see both their pros and cons as discussed in the class, so the exponential has an inherit shift in values from the true trajectory, but what's beautiful about it is that it has the same trend with the true one, unlike running mean where some of the values can be seen increasing while the true values were decreasing and the other way around.

2.5 **Make conclusions which methods give greater methodical error in conditions of equal errors conditioned by measurement errors for this particular generated trajectory.**

as we can see in the above grapgh, the Exponitial mean smoothing curve has a shift to the right from the true trajectory on the other hand the runnig mean curve follows the main trinds of the true trajectory however it fails to follow the fast changes (for example the sudden jumps got smoothed). So generally both methods have a their errors but the error in the exponitial mean smoothing curve is more significant and given that both methods has the same measurement error, We thinke the constant shift will add up to get greater methodical error than the runnig mean.

Learning log:

In this assignment we have dealt with the relation between number of data points and how we can rely on them for an estimation of true values, and also compared the running mean with the exponential method and to what extent we can depend on for our calculations as smoothing methods and how their behavior is affected by their parameters, and if we can conclude bullet points to rely on in future analysis they would be:

- The more the data points you can gather from an experiment, the more you can rely on your analysis to be close to the true physical model.

- How the exponential smoothing coefficient affects the smoothing process of measurement data, and the practical sense of the number as it approaches both 0 and 1.

- How the running mean can be effective as well as the exponential in smoothing data, although it accounts for the last point as the most important one. Also, how the behavior of the smoothing changes from a model to another

- Smoothing methods don’t always follow the trend of the true trajectory of data, so they cannot be fully trusted, but they are the best approximation we have. Like the shift in the exponential and also the change of concave and convex curves in the running mean.

  - we also tried to run the α estimation for 1000 runs then get the avrage of the α array as a trial get better estimation for the smoothing coeffecient but, we couldn't get it any better the average was close to our estimation from the 3000 size trajectory.