# Simulation assignment 2
# Variation of MSE with step-size using the LMS algorithm

Balasubramaniam M C
Roll- EE18B155

November 25th 2020

## Abstract:

This assignment is about finding the variation of excess Mean Squared Error with step size by estimating a filter of given no. of taps using the LMS algorithm for independent and correlated Gaussian( with shift structure) input data in presence of noise. Then, it's verified with theoretical expressions for excess MSE derived for small and reasonably high step-sizes ( using the separation principle).
The theoretical expressions for excess MSE for the LMS algorithm is given by the following:
Small step sizes:

$$\zeta^{LMS} = \mu {\sigma_v}^2 Tr(R_u)/2 \tag{1}$$

Under reasonably high step sizes, using the separation principle-

$$\zeta^{LMS} = \mu {\sigma_v}^2 Tr(R_u)/[2 - \mu Tr(R_u)] \tag{2}$$

where,
$\mu$ is the step size
$R_u$ is the covariance matrix
${\sigma_v}^2$ is the noise variance

## Introduction

The MSE for each step size is obtained by running the LMS algorithm for $4*10^5$ iterations, averaging the errors for each step size for 100 experiments. Finally to obtain the MSE in steady state for each step size, we average the errors $|e(i)|^2$ for the last 5000 iterations. This assignment can be divided into two parts based on kind of input data.

- Independent Gaussian Regressors:For each iteration of the LMS algorithm we generate a zero mean gaussian input regressor $u_i$ with a covariance matrix $R_u$ having an eigenvalue spread $\rho$ of 5. The desired output is generated as per:

$$d(i) = u_i w_o + v(i) \tag{3}$$

  where,
  d(i) is the desired output
  $u_i$ is the input regressor for that iteration
  $w_o$ is the optimal weight vector for the filer
  v(i) is the white gaussian noise with given variance

- Correlated Gaussian Regressors : First, we generate zero mean, unit variance data {s(i)} and pass it through a first order auto-regressive filter with the following transfer function

$$\sqrt{1-a^2}/(1-az^{-1}) \tag{4}$$

  where,
  a = 0.8
  After we obtain correlated data {u(i)} using this, we use a tapped delay line to generate $u_i$ (of the size of the LMS weight vector) . The desired output d(i) is generated similar to the first case.

We then compare the experimental results for MSE with the theoretical values for both cases.

# Procedure:

## Independent Gaussian Regressors:

We intialize an optimal weight vector. Then, we make the $R_u$ matrix (here, for simplicity it's a diagonal matrix) with given eigenvalue spread of 5 as shown below.

```
%generating Ru
val = 10*[0.1 0.12 0.2 0.5 0.25 0.18 0.16 0.14 0.22 0.4];
%with eigenvalue spread 5
Ru = diag(val); %covariance matrix
```

We then iterate through L=100 experiments. In each experiment, we generate input and desired output data according to what was discussed in the introduction.

The main code being:

```
%generating input regressors and stacking below
%mean is a 0 vector
ui = mvnrnd( mean, Ru, 1);
u(i,1:M)= ui;
%generating 1 sample of 0 mean gaussian noise of given variance
vi = sqrt(var_noise)*randn(1);
d(i)= ui*wo + vi;  %desired o/p
```

After this, we iterate through step sizes( from $10^{-4}$ to $10^{-2}$) and execute the lms algorithm for N = $4 * 10^5$ iterations for each step size. We use the 'ui' generated earlier in every ith iteration.
The error values $|e(i)|^2$ are stored in a matrix for every iteration and step size for the given experiment. At the end of the experiment iteration loop we accumulate this mse matrix over experiments. Finally, we average this over L = 100 experiments.
The last error values of last 5000 iterations are taken for each step size and averaged to produce MSE for that step size.
Using the theoretical values of EMSE in (1) and (2) we find the theoretical MSE by adding ${\sigma_v}^2$. This is because in steady state the MSE(using random variables) is given by :

$$J(i) = E(|\mathbf{d} - \mathbf{u}w_{i-1}|^2) = {\sigma_v}^2 \tag{5}$$

Since, $w_{i-1}$ reaches a steady state weight vector (close to the optimal) and data is generated according to eqn(3), the $J_{min}$ tends to the variance of the noise itself. The plots for MSE are shown in fig. 1 and 2.

## Correlated Gaussian Regressors:

After initializing an optimal weight vector, based on the auto-correlation property of {u(i)}, the $R_u$ matrix is shown to have a toeplitz structure:

$$R_u = a^{|i-j|}, 0 <= i, j <= M - 1 \tag{6}$$

This is generated by:

```
for i= 1:M
    for j= 1:M
        Ru(i,j) = a^(abs(i-j));
    end
end
```

The procedure for this case would be exactly the same as the previous one, except for the tapped delay line construction to generate correlated input regressors. As said earlier, we generate zero mean, iid gaussian data and pass it through a filter to get correlated data as shown below:

```
%generating 0 mean unit variance gaussian iid data
s = randn(1,N);
%zero padding for the starting samples
u = zeros(1,N + M-1);
%passing through filter to get correlated data
u(M:end) = filter([sqrt(1-a^2)],[1 -a],s);
```

The following mimics the sliding window

```
%Taking a sliding window block of M samples every iteration
ui = u(i:i+M-1);
```

The rest of the procedure is identical to the previous case. The plots for MSE are shown in fig. 3 and 4.

## Observations and Conclusions on Graphs:

- In general, it can be seen that the MSE in steady state increases with step size for both independent and correlated gaussian input. This is because intuitively, it can be visualized as jumps at many points of the MSE function ( function of the weight vector). In the steady state (or after many iterations), if the step size is high the LMS algorithm causes big jumps in this MSE function due to high changes in the weight vector 'wi'. Hence, for lower step sizes it converges better. Additionally to reduce the excess MSE, one can look at the theoretical EMSE expression (eqn 1) and approximating the order of the step size from this based on $Tr(R_u)$ and noise variance to give low EMSE (which in this case turns out to be approx. $10^{-4}$ or $10^{-3}$).

- We assume wide-sense stationary data for both input regressors (independent and correlated). Hence, the variation of the MSE is similar in both cases depending only on $Tr(R_u)$. This also implies that in certain applications where we can choose the input data (like channel estimation or filter prediction), the $R_u$ matrix can be set appropriately to give minimum Excess MSE as possible. This also answers the question why the variation in the independent case is upto -29.4dB, while in the correlated case its upto a lower value. It's because the $Tr(R_u)$ is different in both cases, hence the slope of MSE with step size also varies. In the independent case, if the eigenvalue spread of $R_u$ is increased, it's trace and hence the EMSE also increases. However, in the correlated case if the value of "a" changes we see that the trace of the toeplitz matrix $R_u$ remains unchanged (depends on the value of M- filter tap length) thereby not changing the EMSE.

- It's seen that in case of both independent and correlated gaussian input, that the theoretical expression for excess MSE for small step sizes doesn't adhere to the simulation results at higher step sizes as shown in fig. 1 and

3. This is because in deriving this expression, the following was assumed:

$$E(||u_i||^2|e_a(i)|^2) << {\sigma_v}^2 Tr(R_u) \tag{7}$$

where $e_a(i)$ is the a-priori error.
However, we know that with increasing step size this error increases, hence the approximation may no longer be valid.

- The theoretical expression for MSE using separation principle seems to adhere to simulation results for both correlated and independent gaussian data. This is because it uses the separation principle which is:

  At steady state, $||u_i||^2$ is independent of $e_a(i)$.

  This is fairly reasonable since, in steady state the error is less affected by the input regressor. It's better than the assumption used earlier and hence, it can be seen from fig. 2 and 4 that the theoretical expression for a wide range of step sizes (eqn 2) adheres to the simulation results for higher and lower step sizes.

- Infact, from eqn. 2 ,it can be seen that for small step sizes, it can be approximated to eqn 1.

- From the plots, the minimum MSE- $J_{min}$ is around -30dB which is about the value of noise variance( in dB) as seen earlier.

- Finally, another reason why MSE from simulation may not adhere to the theoretical MSE in the plots might be that we are using an ensemble average of $\{|e(i)|^2\}$ in the last 5000 iterations for the MSE value, while theoretical expressions are derived considering random variables.

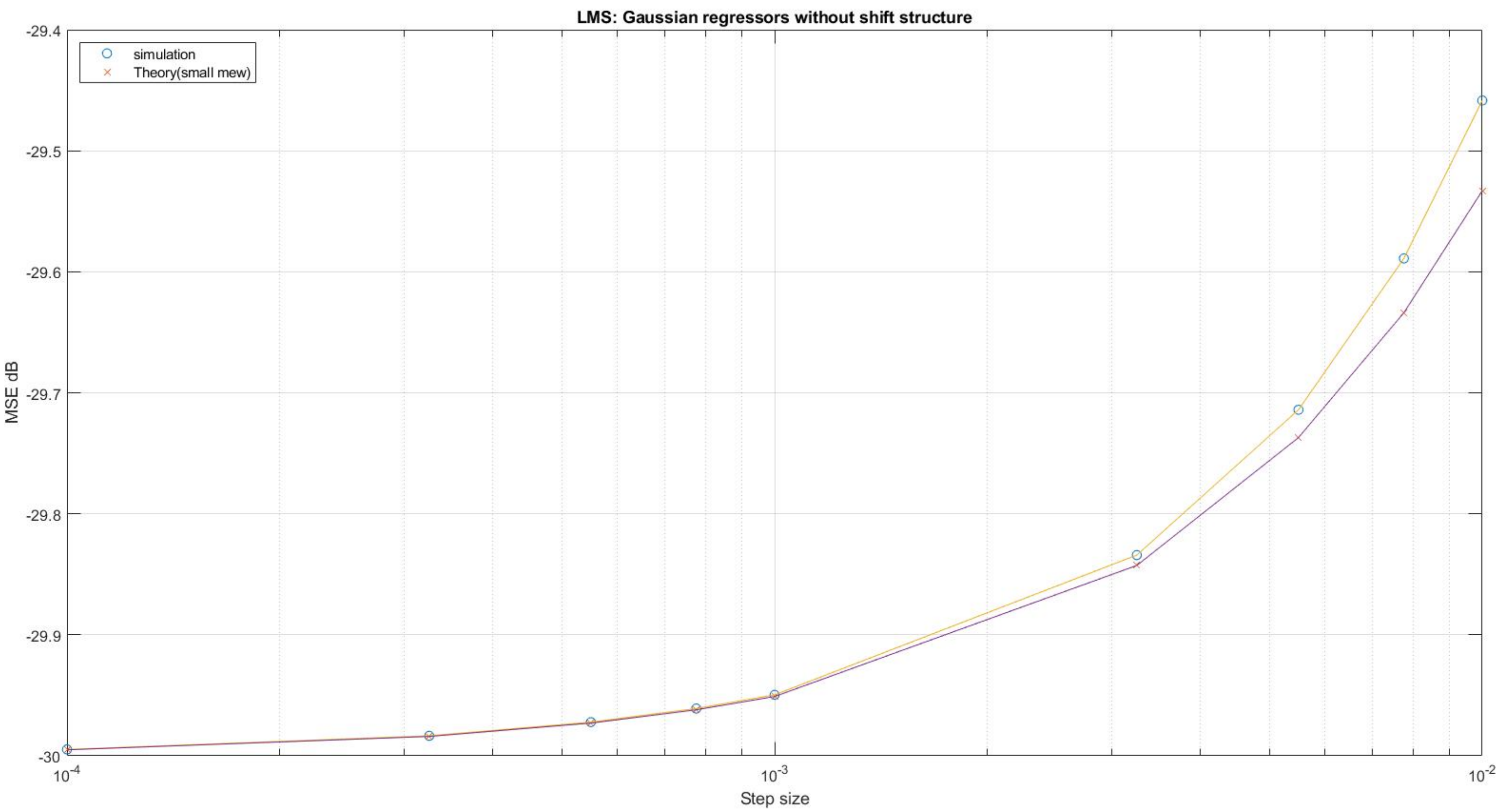


Figure 1: Plot for MSE vs iterations with independent gaussian regressors (with noise variance of 0.001) and comparing with it's theoretical expression for small step sizes.

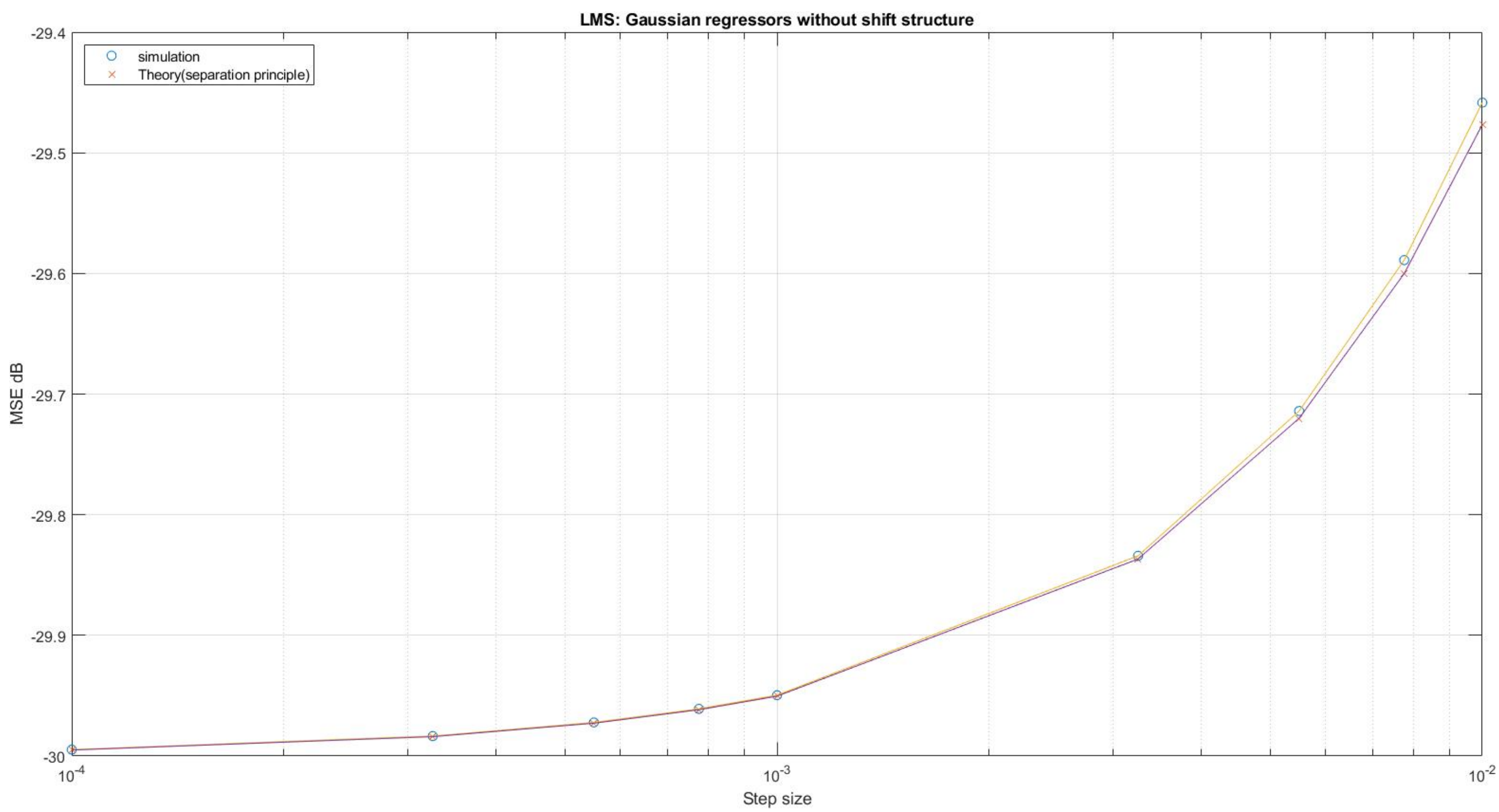


Figure 2: Plot for MSE vs iterations with independent gaussian regressors (with noise variance of 0.001) and comparing with it's theoretical expression for relatively bigger step sizes( using the separation principle ).

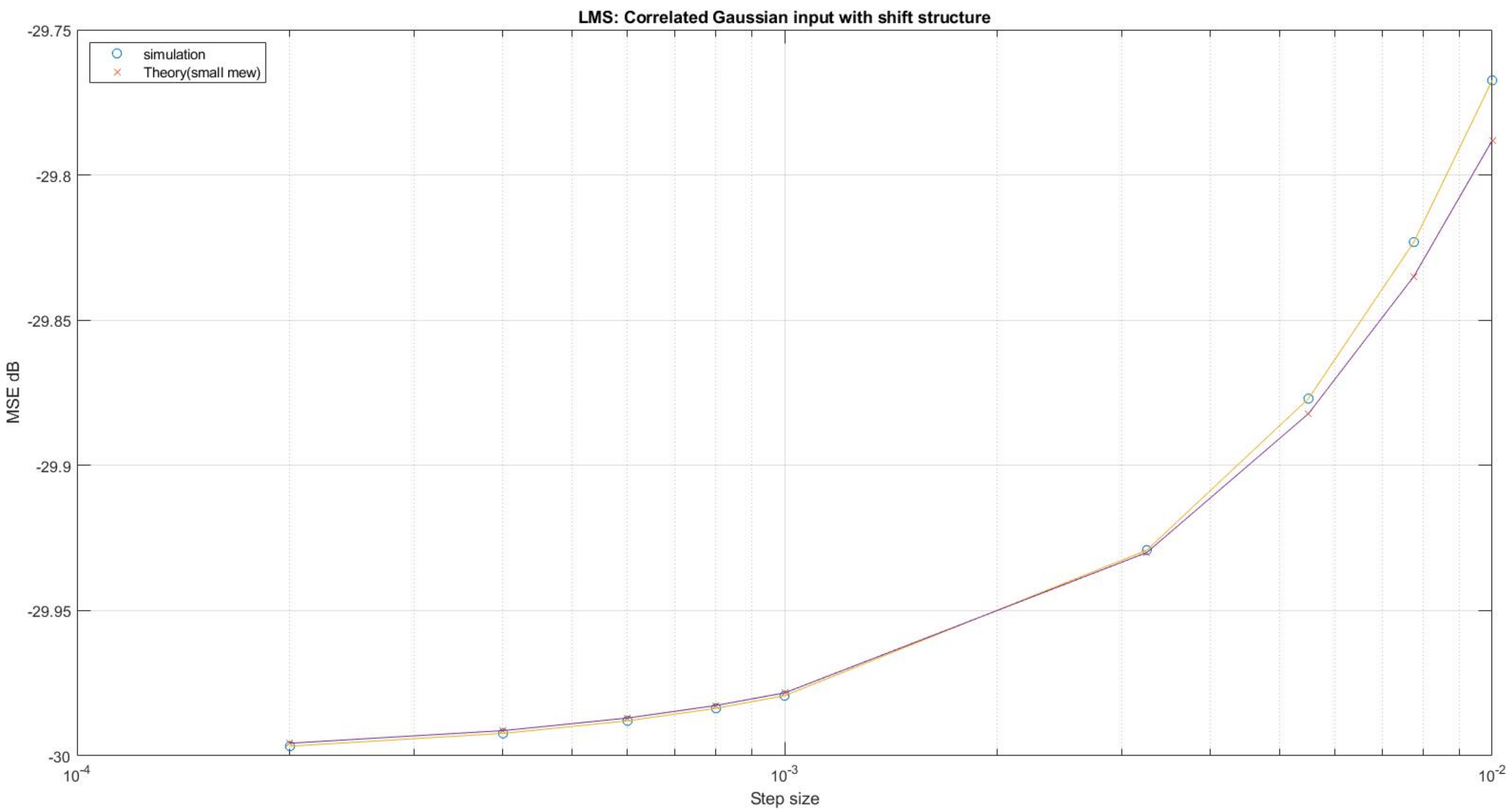


Figure 3: Plot for MSE vs iterations using correlated gaussian regressors with shift structure (with noise variance of 0.001) and comparing with it's theoretical expression for small step sizes.

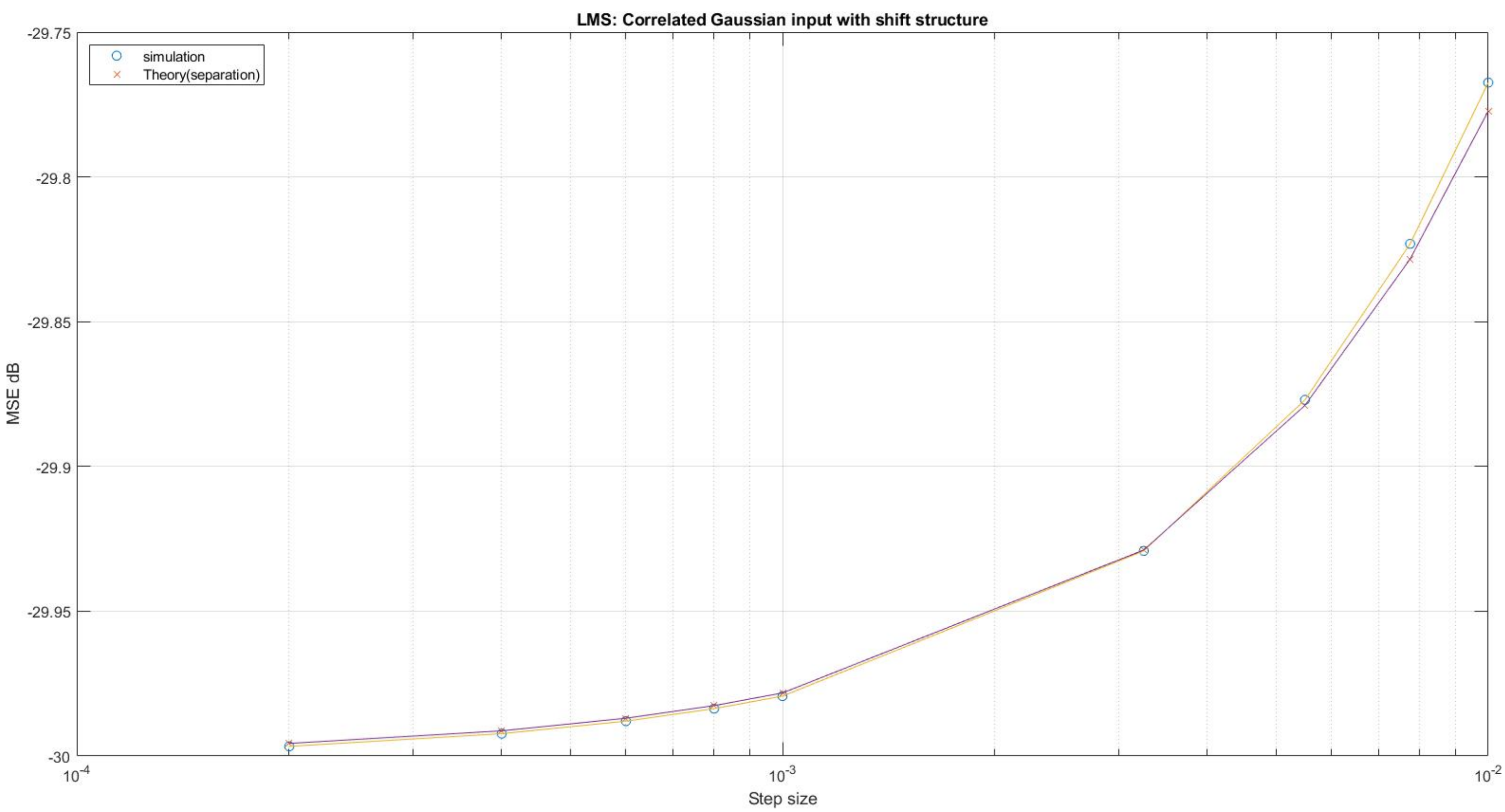


Figure 4: Plot for MSE vs iterations using correlated gaussian regressors (with noise variance of 0.001) and comparing with it's theoretical expression for relatively bigger step sizes (using separation principle).