EE6110 Adaptive Signal Processing

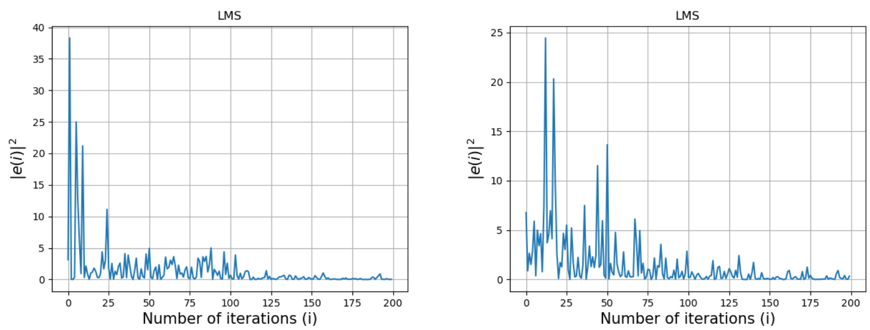
Subhankar Chakraborty EE17B031

December 4, 2020

1 Convergence plots for different algorithms

In this section we compare the convergence speeds of LMS, NLMS and RLS. The experimental details can be found in Example 10.1 of Adaptive Filters by Ali Sayed and are being skipped for the sake of brevity.

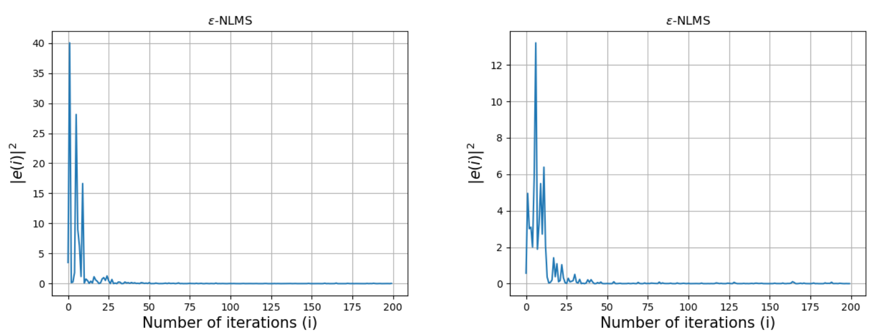
1.1 LMS



(a) Experiment 1 (b) Experiment 2

Figure 1: Convergence plot for two realizations of LMS

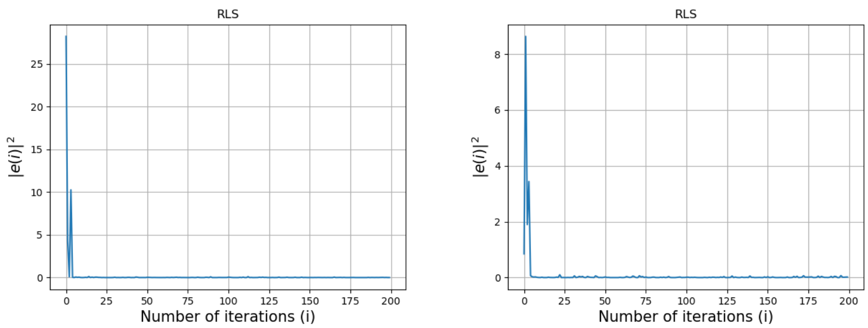
1.2 ϵ-NLMS



(a) Experiment 1 (b) Experiment 2

Figure 2: Convergence plot for two realizations of ϵ-NLMS

1.3 RLS



(a) Experiment 1 (b) Experiment 2

Figure 3: Convergence plot for two realizations of RLS

1.4 Ensemble Average Learning Curve

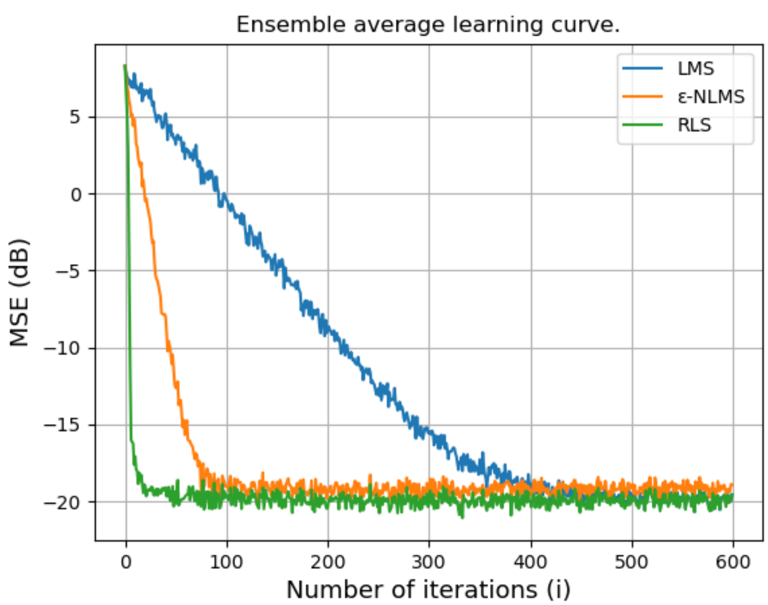


Figure 4: Ensemble average learning curves for the three algorithms obtained by averaging over 300 experiments.

1.5 Observation and Inference

The following observations can be drawn from the plots.

- Even though the LMS algorithm is a very crude instantaneous approximation of the steepest gradient method, the error eventually converges to zero irrespective of the realization of the data. This shows that the approximation works well in practice.

- ϵ-NLMS and RLS also eventually converge irrespective of the realization of the data showing that these too work well in practice.

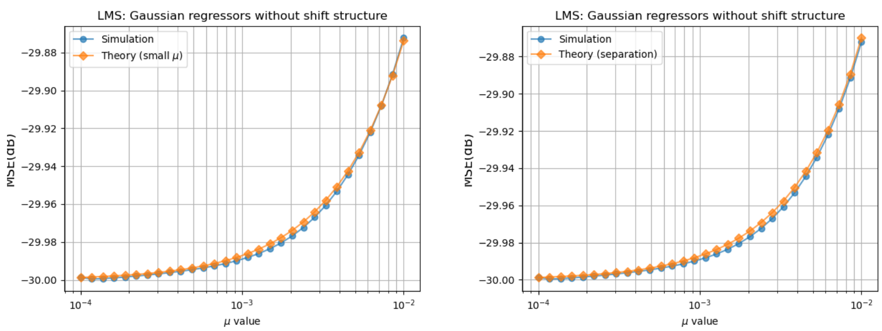
- We saw in class that the computation complexity follows the order $RLS \geq \epsilon - NLMS \geq LMS$. However, the added computational complexity comes with the benefit of faster convergence. The convergence rates follow the trend $RLS \geq \epsilon - NLMS \geq LMS$ showing the fact that there is a tradeoff between convergence speed and computational complexity.

- Although not apparent from the plots here, the steady state MSE for the given parameters follows the trend $MSE_{\epsilon-NLMS} > MSE_{LMS} > MSE_{RLS}$, i.e., RLS performed the best in terms of MSE and $\epsilon-$NLMS performed the worst. This was inferred by generating data as mentioned in the textbook example but running each algorithm for 60000 iterations instead of 600 and taking the average error of the last 50 iterations. This was repeated 20 times with independent realizations of the data and MSE was obtained by averaging over the 20 experiments.

2 Performance of LMS

In this section we compare the steady state MSE for two LMS experiments to theoretical. We look at independent Gaussian regressors without shift structure and correlated Gaussian input with shift structure. The experimental details can be found in Section 16.6 of Adaptive Filters by Ali Sayed and are being skipped for the sake of brevity.

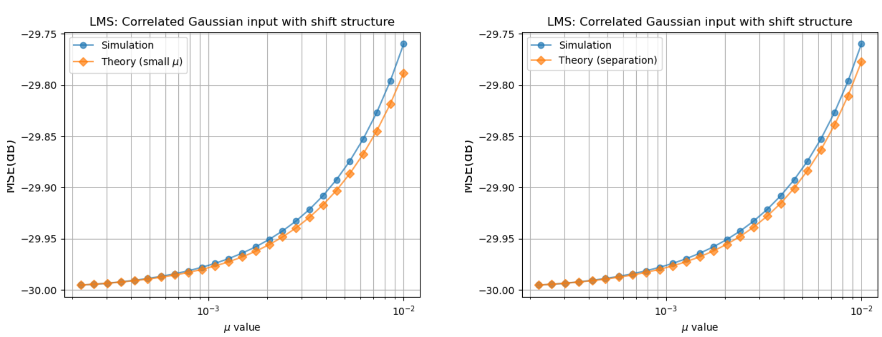
2.1 Gaussian regressors without shift structure



(a) Small step-sizes assumption (b) Separation assumption

Figure 5: Theoretical and simulated MSE for a 10-tap LMS filter with $\sigma_v^2 = 0.001$ and Gaussian regressors without shift structure.

2.2 Correlated Gaussian input with shift structure



(a) Small step-sizes assumption (b) Separation assumption

Figure 6: Theoretical and simulated MSE for a 10-tap LMS filter with $\sigma_v^2 = 0.001$ and regressors with shift structure. The regressors are generated by feeding correlated data into a tapped delay line.

2.3 Observation and Inference

The following observations can be drawn from the plots.

- MSE approximated using the separation assumption is in general valid over much larger range of μ than that approximated under the assumption of small step sizes. This is especially apparent in Figure 6.

- The theoretical and simulation values of the MSE for both approximations of ζ^{MSE} are much more closer to each other in the case of Gaussian regressors without a shift structure than in the case of correlated Gaussian input with shift structure. It implies the following.

  - The assumption that in *steady-state*

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

    is much more valid for the case of Gaussian regressors without shift structure than that of correlated Gaussian input with shift structure.

  - The separation principle, i.e., in *steady-state* $(i \to \infty)$

    $$\mathbb{E}\| \|u_i\| \|^2 \|e_a(i)\|^2 = (\mathbb{E}\| \|u_i\| \|^2)(\mathbb{E}\|e_a(i)\|^2) = \mathrm{Tr}(R_u)\zeta^{LMS}$$

    is again much more valid for the case of Gaussian regressors without shift structure than that of correlated Gaussian input with shift structure.

- Although not apparent from the plots here, the simulation MSE for the correlated Gaussian input with shift structure case was really high for the initial few μ values. It could probably be resolved if I took 4×10^6 or higher iterations instead of 4×10^5 but that was computationally limiting for me and my code kept getting killed. Those points have been removed to make the plots cleaner.

3 Other Details

3.1 Choice of parameters

The 10-tap filter is chosen to have values $\{1, -2, 3, -4, 5, -5, 4, -3, 2, -1\}$ for both Figure 5 and Figure 6. For Figure 5, R_u is chosen to be a diagonal matrix with entries $\{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 0.5\}$. This satisfies the given condition that the eigenspread of R_u is 5. To generate the ensemble average curves, averaging is taken over 80 experiments. μ is chosen as a geometric series, i.e, of the form $a, ar, ..., ar^{29}$ with $a = 10^{-4}$ and $r = 10^{\frac{2}{29}}$. Rest parameters are the same as those mentioned in section 16.6 of the book.

3.2 Calculation of J_{min} for theoretical values of MSE

We used two formulae for the excess MSE (ζ^{LMS}).

$$\zeta^{LMS} = \frac{\mu\sigma_v^2 Tr(R_u)}{2}$$

under the assumption of small values of μ, and

$$\zeta^{LMS} = \frac{\mu\sigma_v^2 Tr(R_u)}{2 - \mu Tr(R_u)}$$

under the assumption of separability. For the total MSE achieved, we need the optimal theoretical MSE, J_{min} as well.

$$J_{Total} = J_{min} + \zeta^{LMS}$$

J_{min} is given by

$$J_{min} = \sigma_d^2 - R_{ud}R_u^{-1}R_{du}$$

with the definitions of the parameters as used in the book in Chapter 10. However, for this case it can be calculated much more easily. The given model is

$$d = uw_0 + v$$

with u being the input and d being its corresponding output, w_0 being the filter coefficients and v being the noise. v is assumed independent of u. J_{min} is defined as

$$\min_w \quad \mathbb{E}[\|d - uw\|^2]$$

It can be easily shown that the above minima occurs at $w = w_0$, i.e,

$$\begin{aligned} J_{min} &= \mathbb{E}[\|d - uw_0\|^2] \\ &= \mathbb{E}[v^2] \\ &= \sigma_v^2 \end{aligned}$$

So, J_{Total} is given by

$$J_{Total} = \zeta^{LMS} + \sigma_v^2$$

3.3 Initial condition to generate $u(i)$ from $s(i)$ in the case of correlated Gaussian input with shift structure

For the plots of correlated Gaussian input with shift structure, the transfer function mapping $s(i)$, a unit variance i.i.d. Gaussian random process, to $u(i)$ is given by

$$\frac{\sqrt{1 - a^2}}{1 - az^{-1}}$$

For an input sequence $x(i)$, the output sequence $y(i)$ is given by

$$y(i) = ay(i-1) + \sqrt{1 - a^2}x(i)$$

To solve for y exactly, we need some initial conditions on y. Since nothing such is mentioned in the example for the sequence $u(i)$, we assume it to be a causal sequence. Hence, the initial condition becomes

$$u(0) = au(-1) + \sqrt{1-a^2}s(0)$$

Using the causality assumption, we get

$$u(0) = \sqrt{1-a^2}s(0)$$

The sequence $u(i)$ has been generated from $s(i)$ using this initial condition.