Adaptive Echo Cancellation

Koushik Nadakuduti
Dept. of Electrical and Computer Engineering,
Herbert Wertheim College of Engineering
University of Florida
Gainesville, Florida, USA
k.nadakuduti@ufl.edu

***Abstract*— In this paper, different methods of implementation of adaptive filter (Weiner method, LMS method) using FIR, IIR gamma filters for echo cancellation are discussed with sample values, results and plots.**

Keywords—adaptive;filter;echo;LMS;FIR;IIR;gamma;

I. Introduction

Adaptive filters are the linear filters whose transfer function can be modified by varying the parameters and adjusting them using an optimizing algorithm [1]. Each delay line in the filter holds the memory of one previous input component. So as the number of delay lines (i.e., Model order) increases, the memory goes that deep into past. By using the knowledge of previous input, estimating the future value is the most important application of adaptive filter.

Adaptive filters have a wide range of applications in the modern world of engineering. Some of them are prediction, prediction for modelling, model based spectral analysis, System identification, Inverse modelling, Interference cancelling, echo cancellation, adaptive inverse control etc., (Principe et al. 1993) They are very powerful and efficient to use. Now, the application we are taking for our discussion is various methods of Echo cancellation using adaptive filters.

II. Methodology

The concept behind using Adaptive filters for echo cancellation is that, the adaptive filters, work in a way that, the output of the filter will be the correlated data of the input and desired output, whereas the uncorrelated data will be the error signal. In our application if we give the music signal as an input to the adaptive filter, and corrupted signal which has both music and speech mixed, as the desired response, because of the correlation between the music component in the corrupted signal and the original music signal, music will be separated from corrupted signal and will be given as the filter output whereas the speech signal separated from corrupted signal will be the error. So, by collecting the error signal, we're able to recover the speech signal.

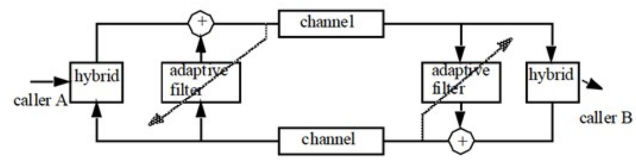


Fig: Block diagram of adaptive echo cancellation [4]

We can get a clear understanding of our application by looking into the above block diagram. The hybrid at the caller A is imperfect which makes the speech signal corrupted by adding noise to it which is the music from caller B in this case. It makes, caller B hear the corrupted speech which includes noise mixed with the speech of A. Let us assume hybrid at far end (caller B) is working fine for simplicity.

Standardization of a dataset is a common requirement for many machine learning estimators, they might behave badly if the individual feature do not more or less look like standard normally distributed data. In general, many elements used in the objective function of learning algorithm assume that all features are centered around zero and have variance in the same order. If a feature has a variance that is orders of magnitude larger than others, it might dominate the objective function and make the estimator unable to learn from other features correctly as expected [2] Hence, it is important to keep in mind that, the data has to be transformed before working on it as explained above for our algorithms to work perfectly

A. Using FIR Filter:

To introduce delay in the system, one of the best methods is using FIR (Finite Impulse Response) filter, while implementing an adaptive filter using FIR filter, the memory of system is proportional to number of delay taps. There are two different methods of implementation using FIR filters, they are a) Wiener solution and
b) LMS solution.

Wiener method:

This method gives the analytical solution to the problem. It gives a solution by finding the weights such that the output of the system is the orthogonal projection of desired response on to the input vector space. Since, the orthogonal projection gives the shortest distance vector, the error will be minimum. The length of the orthogonal projection is equal to the minimum error and this condition gives best performance of the filter.

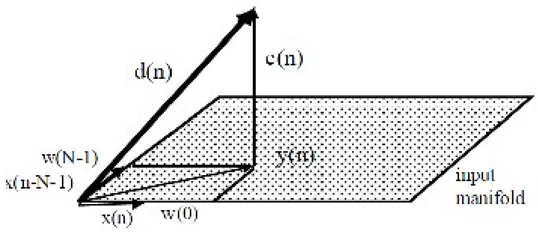


Fig: Wiener solution in vector space [4]

Wiener method is designed to work well with stationary signals. But, coming to our application, since the music wave is not stationary, the wiener solution which is an analytical won't work well. The equation used to update weights in this method is $w^* = (R+\lambda I)^{-1}P$, where R is the correlation matrix, λ is the regularization term and P is X^Td, where d is the desired response. The ERLE of wiener method will be as below

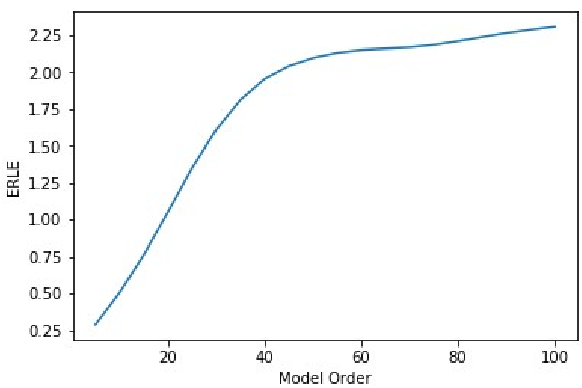


FIG: ERLE CURVE

ERLE means Echo Return Loss Enhancement. It is the measure of echo cancellation in the output. It is the Signal to Nosie ratio (SNR) which gives the part of signal component and noise component in the signal. More the value of ERLE means more SNR implies more signal and less noise which means better the echo cancellation in the output. So, from the above graph we can see that, around model order 40, the value of ERLE increased till that point and then getting flat and not varying very much. That means, the best output should be around model order 50. Let us consider the weight tracks for different samples of same size from our data. Weight track gives the variation of weights over the number of iterations. If the variation in weights is bouncy, it means that, weights are unable to get closer to the optimal value and varying drastically over it. By varying step size, we can make weight curves simple and smooth so that, the weights are not varying much around the optimal value.

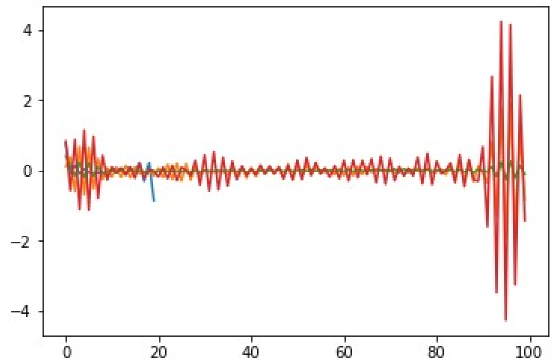


FIG: Weight tracks over different samples

We can see from the above graph that, for same sample size but continuous windows of input data, the weight coefficients are varying drastically so that there is no similarity between weights of any continuous samples. It is because, the input wave is a non-stationary signal and weights are not able represent different statistics of the data. That's the reason why, Wiener solution cannot be used with non-stationary signals.

LMS Method:

LMS means 'Least Mean Squares'. As the name itself suggests, the weights are updated in a fashion that, it tends to the minimum absolute value of the error for each sample. The main and important difference between wiener method and LMS method is that, weights are updates locally for each sample in this method. Because of this reason, this method works even with non-stationary signals unlike wiener method which works only with stationary signals. One interesting thing about LMS method is that, for stationary signals, it approaches the functionality of wiener method. By maintaining proper and optimal step size, weights will be able to track the dynamics of the signal gradually. Now let us see what model order is, and how is it selected.

Model order is basically the order of FIR filter. It is equal to the number of taps. To create any output sample, some number of past inputs, past outputs are used. They are nothing but delayed inputs, delayed outputs. So, the maximum amount of delay used in creating output sample is called filter order or model order [3]. As the order increases, the number of computations involved also increases. So, we need to be careful while selecting model order for the system performance to be optimal.

Let us see the ERLE curve obtained in this method

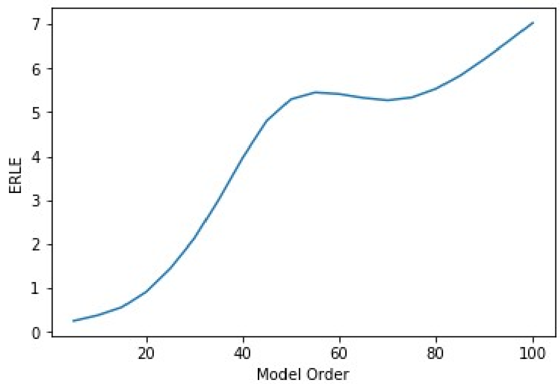


FIG: ERLE CURVE

We can see it in the curve that, as the model order increases, the value of ERLE is also increased. But we need to make sure that, all weights are changing with the iterations. If we take model order more than the optimal value, some of the weights will only be at zero and do not vary. So, we also need to consider this factor while finalising model order. Let us see the weight tracks for model order 100.

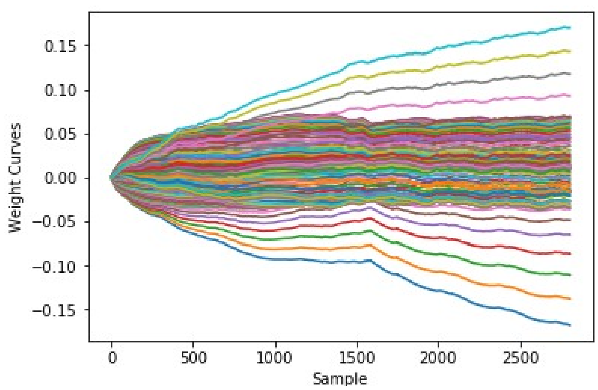


Fig: Weight tracks

We can observe that, most of the weights are centered around origin in the above plot. So, the model order 100 is more than needed. Now, let us consider model order=50

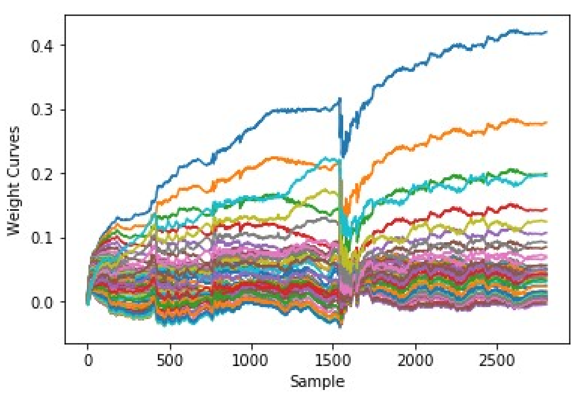


Fig: Weight tracks

From the above figure, we can see that almost all the weights are varying with time. Though some of the weights are around the center, the quality of the output is better when model order is around 50. If we consider model order around 20, almost all weights will be perfectly varying with samples. So, for model orders around 50 we can get the optimal output.

Step size is another important parameter in getting better output. Range of step size is dependent upon the eigen value distribution of the input data. If the step size we selected is more than optimal value, the speed of converging of algorithm will be more but there will be a lot of misadjustment whereas, if the step size is smaller than optimal, the algorithm converges slower with less misadjustment. So, the choice of step size should be such that, the speed of algorithm and misadjustment are properly balanced. It is a tradeoff between speed and misadjustment of data.

We can estimate step size from the weight tracks. If the tracks are bouncy and jumping very much, it means that step size is large and we may need to reduce it. For an optimal value of step size, the curve will be smooth and less bouncy. From the above weight tracks, we can see that the weight tracks are not very bouncy and are smooth enough. So, the step size we used for it (~0.0001) seems to be optimal in this case.

Normalized LMS (NLMS):

The better and improved version of LMS algorithm is NLMS algorithm which means Normalized LMS. In this method we divide the sample by the local input power of that sample. The advantage of this method is that, it will make sure that outliers will not deviate the weight tracks far from its correct trajectory.

The weight tracks obtained when using NLMS method when parameters M=50 and step size = 0.001 taken are as follows:

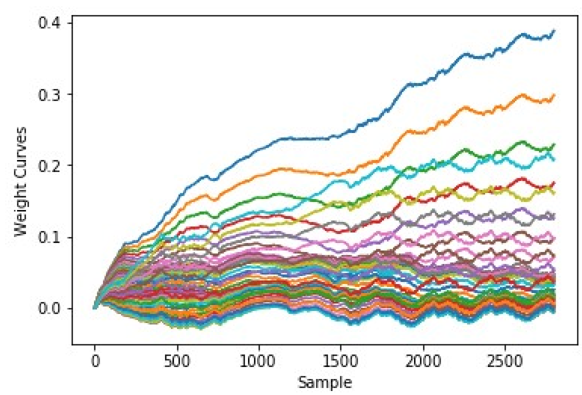


Fig: Weight tracks

We can see the comparison that, in the normalized LMS method, the weights are not being drastically varied at points because of outliers. They have been smooth throughout

If we listen to the audio cleared from noise and other disturbances, in the recovered speech, we can identify that the speech is not very clear but can understand that someone is speaking in the background. Clear observation can help us identify some words at the end of the audio.

A. *Using Gamma filter*

Another way of introducing delay is using Gamma filter. Gamma filter is essentially an IIR filter with restricted feedback architecture. The gamma filter borrows desirable features from both IIR and FIR systems - trivial stability, easy adaptation and yet the decoupling between the region of support of the impulse response and the filter order (Principe et al. 1993) [3].

The advantage of using gamma filter is that, it theoretically has infinite memory. Its memory consists of components of input components way deep into the past, gradually decreasing. It means that, nth sample has some memory of 1st component, some more memory of 2nd element and its memory of the components increases as we reach through (n-1)th component. So nth sample remembers more about n-1th sample than n-2th element and so on.

Let us have a look at Gamma filter before going into its functionality.

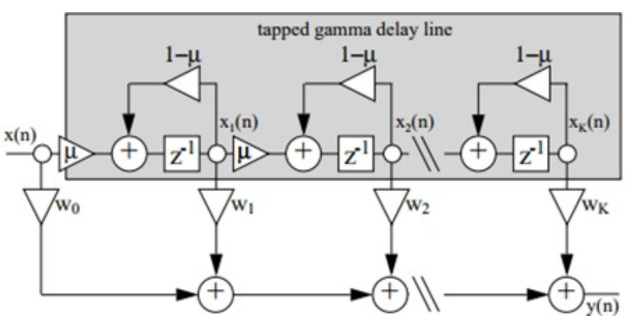


Fig: Gamma – IIR filter [4]

As we can see, each stage of gamma filter has a feedback component and the output of first stage is fed in as the input of the second stage. Because of this structure, gamma filter has a memory very deep into the past.

The equation used to find the output of the system is

$$y(n) = \sum_{k=0}^{M} w_k x_k(n) \text{ [4]}$$

Where,

$$x_k(n) = (1-\mu)x_k(n-1) + \mu x_{k-1}(n-1),\ k = 1, \ldots, M$$

In the above equation [4], the first term is the input part and the second half is the feedback part of the equation.

To make sure that the filter is stable the value of μ should be positive and less than 1 i.e., $0 < \mu < 1$

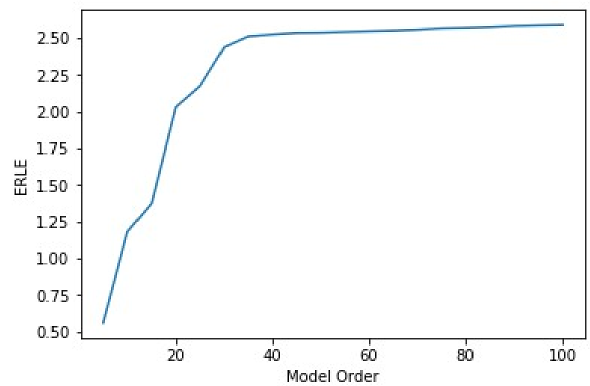
By calculating the gradient and minimum value of error-function, we will reach the below equation for weight updating.

$$w = w + 2\eta_1 e(i) x_k(i) \text{ [4]}$$

Where, η_1 is the step size.

In this method we can also update the value of μ on online basis as we are doing it for weights. But for simplicity, we are taking a fixed value of μ.

Let us look at the ERLE plot as a function of number of taps (model order):



. Fig: ERLE Curve

By following the same conceptual analysis in selecting different parameters like number of taps and step size, we can find out that, best output will be around M=50 and step size=0.001. For simplicity taking a constant value of feedback parameter as μ=0.2. Now, let us look at the weight curves with parameters model order=50, step size 0.001 and μ=0.2

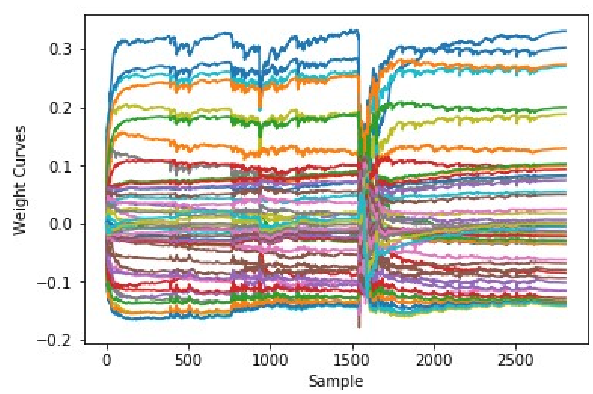


Fig: Weight Tracks

If we listen to the audio cleared from noise and other disturbances, in the recovered speech, we can identify the speech clearly. It is very much clear and better than that obtained using FIR filter.

III. Results

A. Using Gamma filter

So, we are now able to compare the FIR and Gamma-IIR filters qualitatively as well to find out IIR works better for us. Now, let us see some numerical values of performance measuring parameters and the comparison of various methods discussed.

Let us see the comparison of SNR (ERLE) for same parameters for different methods

Model order, Step size	LMS	NLMS	Gamma-IIR
M=10, 0.001	0.5594	0.1343	3.4849
M=25,0.001	1.9404	1.1951	9.8261
M=50,0.001	8.1536	5.6956	18.5512

Table 1: Comparison of ERLE for different methods

From the values in the above tabular diagram we can compare the performance of three methods working with same parameters. Though the values of ERLE in NLMS are less than LMS, it gives the advantage of smooth changes in weights even when there are outliers

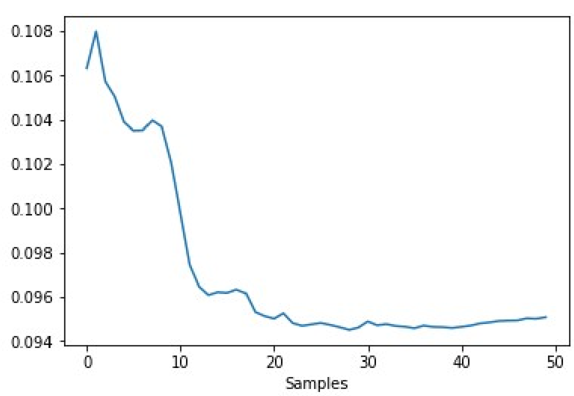
The number of computations required to generate output in IIR filter is more than the FIR filter, which makes its working slow.

We can also see during the execution that, as the model order increases the computation time also increases proportionally.

By making an interesting comparison of IIR and FIR filter by setting the feedback parameter μ=1 in the IIR method which makes it theoretically equivalent to FIR method we can see that ERLE of both values are close to each other. Also, the output recovered speech sound seems similar

Now, let us plot the learning curves of different algorithms to see their progress with number of samples.

LMS Method:



As the samples increases, the cost function is decreasing as we can see from the above graph. To get the better understanding of variation of the cost function with samples, less number of samples are considered

Now, let us look at the NLMS method's learning curve:

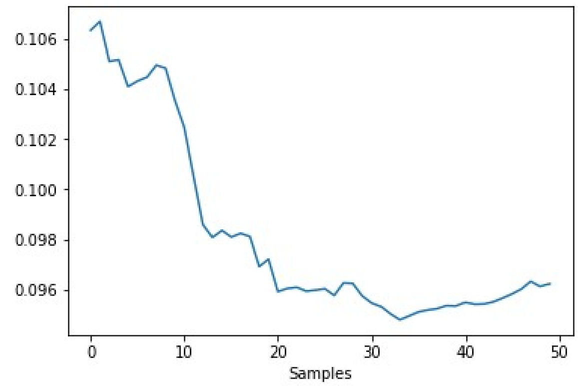


Fig: Learning Curve

We can notice that, the learning curve of LMS and NLMS are similar.

Now, let us look at Gamma filter method's learning curve

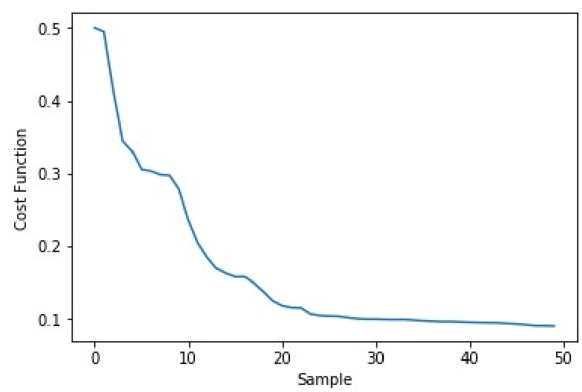


Fig: Learning Curve

IV. Discussion

We can see, the performance of IIR filter is better than that of an FIR filter. But since we are confined to linear model of regression, the recovery is not optimal as the operation of hybrid is not linear. With some higher order regression technique, we would achieve a better recovery of the speech.

V. Conclusion

So, from the above all results and discussions, we can conclude that performance wise, IIR filter works better than FIR filter but the number of computations required to implement IIR filter is more than that of an FIR filter. Because of IIR filter's large memory of past inputs and outputs, it performs well in recovering speech whereas FIR filter's performance is not up to mark because of its limited memory. So, it's a tradeoff between efficiency and resources in selection of a filter.

REFERENCES:

[1] https://en.wikipedia.org/wiki/Adaptive_filter

[2] http://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

[3] https://www.quora.com/What-does-order-of-a-filter-mean

[4] Concepts discussed during info session on project

Declaration:

KOUSHIK NADAKUDUTI

I confirm that this assignment is my own work, is not copied from any other person's work (published or unpublished), and has not been previously submitted for assessment either at University of Florida or elsewhere.