**Computational Neuroscience**
**Dr. Sharba Bandyopadhyay**
**Department of Electronics and Electrical Communication Engineering**
**Indian Institute of Technology Kharagpur**
**Week – 06**
**Lecture – 27**

Lecture 27: Response to Stimulus Mapping (Decoding)

Welcome. So, we have been discussing the spike triggered average method to estimate receptive fields of neurons. We showed that simply by doing the spike-triggered average of the stimulus, we can come up with a kernel which is the impulse response of a linear time invariant system which can depict the overall pathway from the stimulus to the response transformation albeit with further elements like static non-linearity, then driving a point process to generate the spikes and so on. But ultimately it is the the main idea is that in order to go from the stimulus to response, the main part or the filtering done by the pathway or the neuron can be obtained through the spike triggered average. And these kind of methods have various extensions as we have discussed a little bit , but it can be extended to multiple dimensions and so on. However, there is one point that is the deserves your attention and that is that in the whole derivation of the spike triggered average as a model for the neuron, we have assumed a white noise Gaussian stimulus that is there is zero mean and variance is sigma square and the the main thing is that the auto correlation function is a delta function at with no dependence across time.

However, that is not generally true in the sense that we will be using many different kind of stimuli in estimating the receptive fields of neurons using the spike triggered average. And so we need a more general or rather an alternative to estimate those receptive fields, it is again using the spike triggered average, but with different kinds of stimuli. So, that is one aspect that we need to cover and another point that requires your attention is that so far we have been treating all this analytically and we may I mean in in order to implement all this we may have to go into discrete time and discrete representation of the stimulus space and so on. So, we need to develop the same thing for discrete time based analysis.

So, in order to do that let us consider the model let us say we are trying to estimate the rate response of a neuron at the time point t or this is basically the center of a time being let us say and that is defined by basically some kind of weightage of the stimulus preceding the that particular time. So, we will have this weight vector here and this is the stimulus. So, let us say that we are recording from a neuron this is time and this is the rate response. So, let us say this is 0 time and we are this is the probability of spike or rate or in terms of number of spikes per second occurring over time. I mean I have drawn it in a continuous manner, but imagine this to be discrete and parallely just like in the spike triggered average case we have a stimulus in this case not necessarily a Gaussian white noise, but can be a stimulus that whose autocorrelation function is not a delta function.

So, let this $r_t$. So, this is our $s$ as a function of time, this is $r$ as a function of time, this is our time axis, this $r_t$ let us say is the rate in this time being this is our time t and this is the rate value that is this particular rate this this is our $r_t$ at particular time t. So, the response in this particular time is dependent on from causality is dependent on some past history of the stimulus is what we will assume. So, at this time t and the preceding stimulus. So, since we have discretized the time let us say this is $s$ at the time point t the stimulus preceding that time point t is represented by the vector $s_t$ ok which is which contains the elements of the stimulus over time preceding that time point t.

So, let us say this $s_t$ is simply a vector that is $s_0$ $s_1$ or $s_{t0}$ $s_{t1}$ and so on up to $s_t$ some particular memory length capital M capital M time bits. So, this is capital M minus 1 time wins we can make this 1 make this 2 that is some significant factor. So, let us say this is the first bin and going backwards is the second bin and this is the mth bin. And the corresponding values of the stimulus from the time point t backwards are in this vector $s_t$. And we are saying that the rate produced at this time point t let us say we are considering this time point t here the rate produced at this time point t is dependent on this stimulus $s_t$ by some linear weighting.

So, as we have been treating this as a linear time invariant system. So, some linear weighting in the sense that this w or we we can call it let us say h let us say. So, this $h^T$ it is another vector times $s_t$ this is another vector. So, what is h is essentially how the way how the stimuli are weighted in order to obtain the response. So, we have note that we have removed the non-linearity in this case because we know that that we can treat that you know I mean in order to actually implement it and I mean it still the results would hold and in the from the from our previous discussions.

So, we are now talking about the system without the non-linearity and simply writing the rate as weighted sum of the stimulus preceding that particular time point. So, in other words so let us say this is our going to be $h_1$ $h_2$ up to $h_n$. So, the first stimulus is weighted by $h_1$ the second preceding stimulus time point is weighted by $h_2$ and so on. So, we have this $h^T$s as the estimate of the rate. So, what we now have is this $rhat$ at time point t is simply the $h^T$.

So, this is fixed for the neuron this h is fixed for the neuron in the sense this is what we are trying to estimate and s at time point t is this vector. So, overall this thing is a scalar. So, in order to find out h that best describes the rate based on the stimuli and this particular linear model we will consider the mean squared error that is our actual rate that we have measured which is r let us say we have estimated rate as r hat the you subtract this out and the square of that and average this over all the time points that we are using that is all the t this is an average over all the time base. So, now we can simply replace our r hat with our model that is $h^T$s. So, we will drop the suffix t and assume it implicitly and the averaging is also over time which we will assume implicitly.

So, this can be written as our average of $h^T s$ minus r square average which is now if we multiply this let us say we have $h^T s$ minus r into $h^T$ this is the scalar. So, we can write either one before the other one or the other does not matter and this is being averaged over all the time points. So, this can be written as $h^T$ of ss$^T$h minus $h^T s$ times r$^T$r is a scalar at that point. So, we can simply write r then minus r h sorry this is $h^T$r $h^T$r s$^T$ h plus r square the average of this whole thing. So, this boils down to the average of $h^T$ss$^T$h minus the average of $h^T s$ r minus r s$^T$ h average of that plus average of r square.

So, this is simply the variance or rather I mean average of r square if we treat it as if we convert it into a zero mean rate without any loss of generality this is the variance of that r . So, but this this is already available to us and it is a constant. So, this is the mean squared error if we use a linear model like $h^T$s in order to estimate r and we need to minimize this obviously. So, we take and we have to take the derivative with respect to h in order to obtain the h for which this is minimum and before that we would like to bring our attention to these terms ss$^T$ sr rs$^T$. So, sr and rs$^T$ they are essentially essentially the same because r is a scalar and so it is simply an average of the different r s with the previous stimulus.

So, it is essentially the cross correlation between s and r estimated from the data and ss$^T$ is the auto covariance matrix of s and s being ah I mean s being 0 if we can make a zero mean then this is simply the auto correlation matrix of s. So, by representing that the average of ss$^T$ we write it as $C_{ss}$ that is the auto correlation matrix and similarly the average of sr is simply $C_{rs}$ or the cross correlation between the response rate r and the stimulus and so these two terms both provide the $C_{rs}$ term and so we can write this as $h^T$ $C_{ss}$ h minus $C_{rs}$ ah times h ah or rather twice that plus the average of r square. So, now if we take the derivative of this mean squared error with respect to h what we obtain here is simply ah 2 $C_{ss}$ h ah. So, this is if we consider this as e then what we are doing is de/dh that comes out to be $C_{ss}$ h minus 2 $C_{rs}$ h and ah. So, for this we equate this to be 0 in order for ah this to be the minimum the mean squared error to be at the minimum.

So, that implies that our h is given by $C_{ss}^{-1}C_{rs}$. Now, remember this is exactly the same of what we had with the spike triggered average barring the fact that this $C_{ss}^{-1}$ is an identity matrix for a stimulus that is white Gaussian noise identity matrix and scaled by the variance of the white Gaussian noise because the auto correlation matrix of the white Gaussian noise would be a diagonal matrix which each diagonal element as the sigma square of that particular Gaussian. So, what we are doing getting here is that this is the $C_{ss}^{-1}$ is providing a way to get rid of the correlations in the stimulus and allowing us to use any kind of stimulus per say to obtain ah the linear ah kernel estimate for the neuron. So, it must be noted that we use using the $C_{ss}^{-1}$ is a little tricky because due to because we have finite amount of data we will always have finite amount of data and even though we may be using a sample of say white Gaussian noise for our experiment the $C_{ss}^{-1}$ might have very small elements that are non zero which actually should be zero had we had an infinite amount of a long infinitely long stimuli. So, in that case if we use the $C_{ss}^{-1}$ ah we blow up those errors those non zero values which should have been zero those small values and actually amplify the error quite a bit way by taking the $C_{ss}^{-1}$ and get very ah ridiculous results.

So, we have to be careful in using this $C_{ss}^{-1}$ when we are applying this to correct for the correlations within the stimulus. We have to be absolutely sure about the strength of the correlations in order to apply this. So, we ah we have introduced the idea of ah doing this whole thing in discrete time ah and also using the stimulus as in discrete time and ah values ah within those time bins ah and we will use this set up only to now ask the question that if I have ah the given that I am observing the rate response of a stimulus where in response to a stimulus given some new rate responses can I reconstruct the stimulus. So, so far what we have been looking at has been the forward problem that is we have the stimulus going into this box ah and then the output is our rate let us say after number of steps. So, now, ah if the question is that ok this we have a set of observations like this where we have ah stimulus and rate like we do, but ah now can we if given some new rates or new ah number of observations of responses can we find out what the stimulus which is essentially the decoding problem or reconstruction problems.

And decoding can be done in many ways and reconstruction is one of the ways in order to get back the original stimulus which is not necessarily what the brain will be doing with wherever this rate information is going ah into probably the next state neuron, but this may be what we may be required to do in case of ah let us say brain computer interfaces where we need to know what was the stimulus given a set of responses from a population of neurons and then decide ah the stimulus and then based on that make a decision about what is to be done with a with a ah machine ah that has to be driven by the brain signals. So, in those kind of problems these ah methodologies would be useful albeit with more sophisticated methods ah. Here we will stick to the ah the first step which is a linear method for reconstructing the stimulus exactly as we did ah previously we set up the same problem where now we have ah set of rates just like the last time and we have a stimulus ah in again in ah discrete time all of them. So, this is our rate as a function of time and this is our stimulus as a function of time ah we this time is obviously as I said ah discrete and that is what we will be treating here. So, we set up the same problem in the same way ah.

So, let us consider that ah this is the time being t ah let us say this value here this value here in that particular bin is the stimulus and this is the rate. Now a little difference ah in the problem here is that we observe the rate at a particular time t it is not possible to estimate the stimulus at that particular time point. We because we have a causal system there is a finite amount of gap ah based on which ah the responses are created based on stimuli remember earlier we were using a vector of length m bins ah of stimulus to obtain the rate at time point t. So, this whole stimulus ah this whole period of stimulus was being used to determine this rate and now ah from this rate ah we obviously cannot determine the entire stimulus. So, for the stimulus at this particular time point to be determined let us say now we are representing the stimulus in green here in order to calculate or estimate the stimulus we need the rates in future because now the problem is the other way round.

So, having observed ah the rate and stimulus relationship let us say this is not online, but offline that is we have ah corresponding pair of stimulus and responses then in order to estimate the reconstruction filter that we will use to a new set of observed rates ah to find out what the stimulus was ah that problem is such that ok we will look into the future set of rates and from that we will be estimating ah stimulus in the past in the beginning ok. So, that is essentially taking care of the causality I mean there are number of ways of setting this up, but this is probably the simplest way that ah we have a future set of rates and we now want to get a kernel that will allow us to use these rates to calculate the stimulus. So, we will assume that we will go ah forward let us say in this case capital M bins ah from 1 and likewise we essentially set up the same problem ah where we are saying that the estimate of the stimulus at the time bin t is given by some rate of the time bin H this is now a kernel that will be operating on the rate responses to estimate the stimulus $H^T$R at ah particular ah t plus M ah that particular vector. So, what we mean by $R_{t+M}$ is a vector that contains the rates at time point 1 right after t up to time point t plus M after ah $R_M$ which is at time point t plus M $R_1$ to $R_M$ and this is a suffix by t or we can introduce the t here itself ah ah the rate at which the rate is given by the rate at which the rate is given by the rate at which the rate is given and in fact write it out simply as ah $R_{t+1}$ $R_{t+2}$ up to $R_{t+M}$. So, now the the problem is ah exactly the same ah that we were solving earlier except now we have to ah use one particular ah constraint and that is we require ah the stimulus to be 0 mean for using of the auto correlation function.

So, we need to correct for that and that here we will be ah providing as ah basically ah $H^T$ $H^T$ $H^T$ times the R bar at over the entire stimulus. So, ah essentially what we have is our S hat t ah is $H^T$R we are dropping the suffix ah and minus R bar. And so by subtracting out the mean rate we are ensuring that the estimated stimulus is also 0 mean and so we can call this we can call ah delta or a D or something and then proceed in the same way as we did ah for the forward problem. So, $H^T D$. So, we have the exact same problem now and ah the ensuing solution is also going to be the same now we will have the auto correlation function $C_{RR}^{-1}$ times $C_{RS}$ and ah.

So, S is here a scalar quantity for each of the elements and this is going to provide us our H ah if you go on or I am sorry we need to do this D D and this is D S ah after having subtracted out ah the ah mean rate we are getting D. So, by this ah now we can now after we have obtained H if we observe ah rate responses for a period M and more after observing the first M ah points of rate we can use the H in order to estimate the first element of the stimulus and in a sliding manner go forward and calculate the rate. So, ah here after we observe let us say the M bins of rate we can now see what the first or estimate the first point of the stimulus then going one step forward we can estimate the next point of the stimulus and carry on with it in order to fully reconstruct the entire stimulus ah given a new set of rates. So, ah with this we will ah stop our discussions on the linear models ah and later on we will take up a little bit more of second order and beyond models ah small discussion, but we will now change gears in order to go on ah into information theory as we have said that it is the way to go forward ah because of the problems of using non-linear models because of the huge amount of data needed because of the huge expansion of the number of parameters that are required to define a non-linear system ah in terms of polynomial function of the stimulus space and as the stimulus space the parameters in the stimulus space keep growing a second order model itself where there are a square number of ah parameters that have to be estimated. So, ah that is why we will ah now introduce you to ah information theory techniques or ideas from information theory and then show one particular methodology ah that can be used in order to use the advantages of information theory and still be able to model these kind of non-linear systems. Thank you.