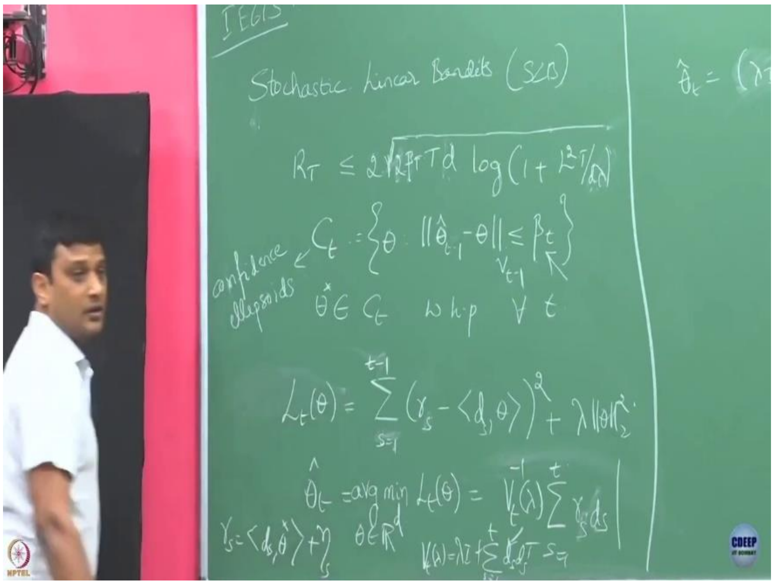
Bandit Algorithm (Online Machine Learning)
Prof. Manjesh Hanawal
Industrial Engineering and Operations Research
Indian Institute of Technology, Bombay

Lecture – 48
Construction of Confidence
Ellipsoid – I

So, today we will continue our discussion on our Stochastic Linear Bandits. So, in the last class, we just argued that if you make certain assumption about my confidence set and about the mean rewards and also about the norm of my context or norm of my decision element, we came up with a method to give a regret bound, right.

(Refer Slide Time: 00:54)



So, today what we will do is we just try to understand if at all it is possible to come up with a confidence set which satisfies our assumption. What was our assumption? We said that there exists some set C_t which we could defined it as $C_t = \{\theta: \left|\left|\hat{\theta}_{t-1} - \theta\right|\right|_{V_{t-1}} \leq \beta_t\}$. We say that $\theta^* \in C_t \; w.h.p \quad \forall\, t$.

So, I wanted that at every time you have this confidence set. We henceforth we will call it as confidence ellipsoid, such because now it is like a ball around, centered around this point $\hat{\theta}$. So, our question is such a thing exists, if it all I can define some β_t for which this holds.

So, today we can again before we going to state it formally like such a β_t exist, we are again going to make further more assumptions and now try to see that if it all it is possible under a bit more restricted cases. Then, we will see that how to relax that restricted case, to a more general case where we can get such a β.

So, the more general case requires us to get into something a bit more sophisticated like martingales and mixture methods. It will may be we will not just go in to that, but we will just talk about a some basic ideas and just point out where the method of mixtures and martingales may have to be used, ok.

But just we will try to get a sense of how this β_t can exist, ok. To understand this we are further going to make a couple of assumptions before we see how it looks like we were going to first set $\lambda = 0$ that is there is no regularization.

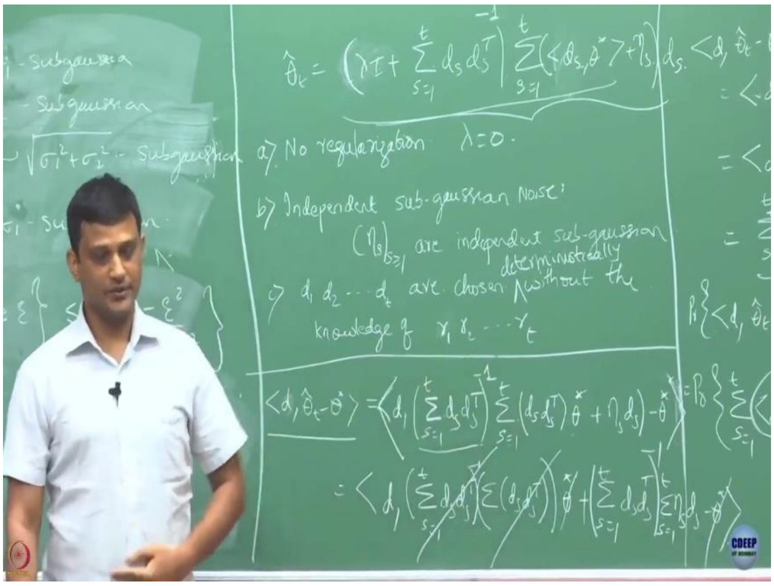
Before I write this. So, how did I find my $\hat{\theta}$? $\hat{\theta}$ we find by minimizing a least square, regularized least square function, right. So, what was that function? We defined a function which was, so this was our last function we defined and we try to minimize it over θ.

And what did it give us? So, this will, this give us $\hat{\theta}_t = argmin_{\theta \in R^d} L_t(\theta) = V_t^{-1}(\lambda) \sum_{s=1}^{t} r_s d_s$ .I already observe that this is going to be the solution.

So, earlier I just defined V_t. What was the V_t matrix? So, λ plus some outer products some of outer products, right, but that was a function of λ that depends on which λ I chose, so that is why I made this a function of λ here. So, this $V_t(\lambda)$ is nothing but $V_t(\lambda) = (\lambda I + sum_{\{s=1\}}^{t} d_s d_s^T)$, and this s is going to be 1 to T, ok.

Now, let us try to slightly simplify this, what is this estimate we have gotten here. This quantity here, so this is going to do it only to t - 1 and then at time t this is what my loss is going to be look like, ok.

(Refer Slide Time: 06:58)



Now, my $\hat{\theta}_t$ just, so I am just replacing my $V_t(\lambda)$ function here. $V_t(\lambda) = (\lambda I + sum_{\{s=1\}}^{t} d_s d_s^T)$.

$r_s = < d_s, \theta^* > + \eta_t$. So, this is the noisy version of the reward. I got in round s. This is simply the reward sample I have observed in round s. So, let me replace that quantity over there. So, what this gives me here if I just replace r_s by that quantity that is ds and this is multiplied by my quantities ds, there is inverse.

So, now, before I start further simplifying these things and see how my confidence ellipsoids are going to look like, I am going to make a couple of assumptions now, simplifying assumptions, ok. First thing I am going to assume the case that there is, I am going to assume λ = 0 that is the case corresponding to no regularization.

Second, independent, so I am going to assume that these noises here are actually independent sub Gaussian noises. So, earlier in the original set up, how this noise term is assumed to be?

Student: Conditionally.

It is conditionally sub Gaussian, but now we are just assuming that there are sub Gaussian and they are independent. So, these are my simplifier. Third point they are assuming that this d_1, d_2 all the way up to the d_t are chosen without the knowledge of r_1, r_2 all the way up to r_t.

So, what is that we are assuming? First ok, we are assuming that, there is no regularization. That means, if there is no regularization here the thing is this inversion maybe not well defined, ok. But if we had this regularization λ greater than 0, this inversion was always well defined. But let us ignore that fact. So, let us set this λ to be 0. And then we are going to say that my η is are all sub Gaussian then their independent.

The third point, this is the very bad assumptions as we are making we are just saying, the arm played in till round t their selection is made without the knowledge of r_1, r_2, r_t. That means we are, right now saying my algorithm is kind of not adoptive, right. So, every time I chose my arm it dependent on actually what I have played previously which arm an played and what is the corresponding reward I got for that.

But now we are saying when I making this assumption like basically I am ignoring the reward I have observed so far and then making a choice of r, ok. That means, the kind of decoupling the dependency of the arm I choice in round t on the observations have made so far. This is and this is actually not true, but let us say we are making this assumption.

So, when we had our set up, if you have to learn you have to make sure that we I am going to choose the arm I am going to play in the next round based on what observations I have made, right. But I as I said that makes things complicated to analyze that is why we will make the simplifying assumption, ok. Let us try to analyze.

So, now, just try to let us try to follow of the sequence of analysis how it goes, then we will see that what we actually wanted. We just now, as of now our focus is to see that will I be able to come up with such a set which contains my θ* with high probability for every t, ok. So, the steps we are going to go through it is going to be certain sequence of manipulation. Just follow this manipulation, we will arrive at this step at some point, ok.

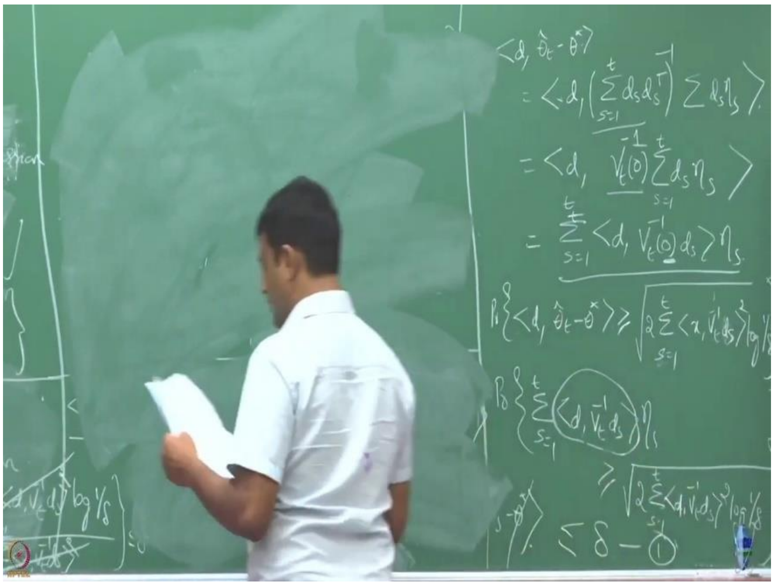
So, now, first we are going to look at this quantity here. So, what is this quantity basically? $\hat{\theta}$ is your estimate in round t, θ*t your true parameter. So, $\widehat{\theta_t}$ minus θ* is the error you are making in that round. But what this is basically telling in the projection of this arm x on this difference, ok. So, instead of x let me just call it d because we are denoting my decision point has d. These are all vectors.

Now, what is this quantity? This quantity is nothing but d and now I am going to replace my $\widehat{\theta}$ by whatever I have by making this λ to be 0. So, this will give me what? So, it will give me summation s equals to 1 to t, d_s, d_s transpose inverse of this, summation, this quantity over here s = t.

So, what I will now do is, so I am just substituting this quantity for $\widehat{\theta_t}$ here. What I will do is I will do this I will pull this d s inside, so that will give me, I have just replace the value of $\widehat{\theta^*}$ from this equation here, and just brought in this d_s just inside this.

So, now, if you see this if I just simplify this d, if I pull this matrix inside this, θ* is constant that is fixed. Now, if I multiply this quantity here this first term with this that will give me an identity. So, let me write this. So, I just expanded this inner product here, ok. So, now, you see that this quantity gets matrix and I am multiplying by inverse. So, they get canceled. And now what remains is θ* which θ* will get knocked up with this single θ* I had.

(Refer Slide Time: 18:07)



So, finally, what remains for me is this projection term here equals to d times. This quantity here is equals $< d,\ \left(\sum_{\{s=1\}}^{t} d_s d_s^T\right)^{-1} \sum d_s \eta_s >$.

But now I am forcing let my λ equals to 0 and trying to assume that holds, ok. And now just let me just to make that more explicit let me write it as V t of 0. So, this V t 0 is nothing but this entire quantity. Just simplifying this further. What I had is this is the inner product between these two.

So, this is between s 1 to t, now if I just take this equals to 1 to t. I can further write it as $\sum_{\{s=1\}}^{t} < d,\ V_t^{-1}(0)d_s > \eta_s$. I have just simply manipulated this like, ok.

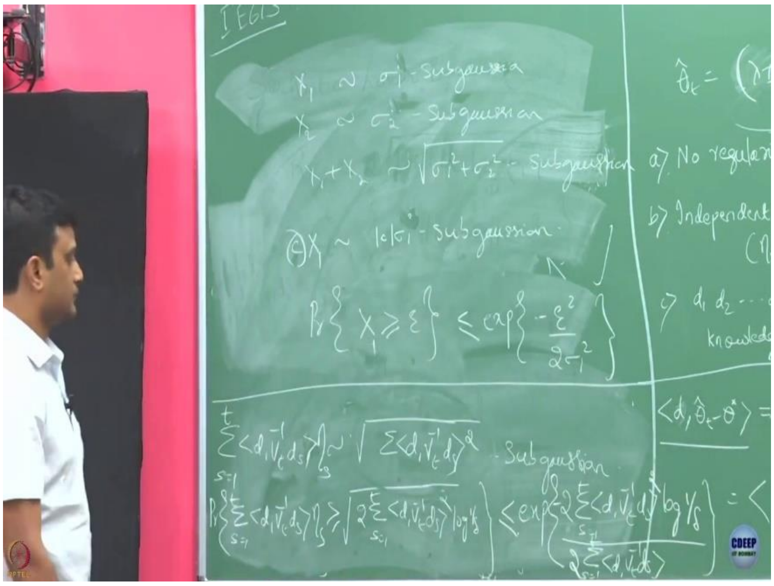
Just notify it everything is fine like. If you just take the inner product of this, right this is inner product of d with this matrix and d_s, and now that is summed over or s equals to 1, but I have just pulled the sum out and I have I am just writing it like this, ok. By doing what I have basically done is this projection of my arm d in the on this error term I have written it in nothing but a linear combination of the noise terms; n_s is the noise term, right, n s is the one sub Gaussian noise. And what is this?

So, this is an it is multiplied by some weight which is nothing, but the inner product of these two quantities and that is some weight. And now this is noise. So, what I have basically done is this quantity nothing but weighted some of the sub Gaussian noise and if I want to bound this quantity and now has to bound this weighted some of the sub Gaussian noise, ok.

So, now, suppose let us say I am interested in how big this quantity is, So, suppose let us say I wanted to ask the question, probability that $\Pr\left\{< d,\ \widehat{\theta}_t - \theta^* >\right\} \geq \sqrt{2}\sum_{\{s=1\}}^{t} < x,$ $V_t^{-1}d_s >^2 \log\frac{1}{\delta}$. I will we will see why I am using this number, but let us say I am interested to know whether it is larger than this number.

So, now what I have that this quantity is weighted sum of the sub Gaussian noise and I am now asking the question what is the probability that this quantity exceeds this value. How you are what result you are going to apply to found find this? So, we have already bounded how the tails of sub Gaussian looks like, right. So, let us try to apply that here, ok. Is this quantity entire quantity here is a sub Gaussian?

(Refer Slide Time: 25:03)



So, if it is sub Gaussian with that parameter. So, we had if X_1 is σ_1^2 sub Gaussian and X_2 is let us say σ_2^2 sub Gaussian. What is $X_1 + X_2$ are like? This is like $\sqrt{\{\sigma_1^2 + \sigma_2^2\}}$

Sub Gaussian, right. So, now, η_s is what? One sub Gaussian. But now it is multiplied by some number. If X 1 is sigma 1 sub Gaussian if I am going to multiply it by constant C, what is CX_1 is like? $|C|\sigma_1^2$sub gaussian.

Sub Gaussian now what is this quantity is like? So, this is going to be sigma Gaussian sorry, this is going to be sub Gaussian with what parameter?

Right. So, this weights are going to, so initially it was one sigma Gaussian. Now, each term in the summation is going to be this much time multiplied sub Gaussian and I if I want to find the sub Gaussian it have the sum it is nothing, but square root of sum of each one of them, ok, fine. Now, I know that this quantity is sigma sub Gaussian with some sigma.

Now, I want to bound this quantity to be upper is at least this much. How to find that? So, let us say, so if some X is, let us say X is or let us say X_1 is σ_1^2 sub Gaussian, what is this probability? We have discussed this, right. We have this probability. What is this quantity is like? Upper bounded as exp $\left\{\frac{\epsilon^2}{2\,\sigma^2}\right\}$.

So, we had already discussed this, right, when we talked about set of concentration equality. Now, let us say, so let us say this X_1 corresponds to this quantity over here and this ϵ corresponds to this quantity over here, ok.

Notice that I can treat this quantity here as a constant because whether I am making my assumption, ok. So, let us say whatever I have observed so far, right, these are all the random quantities, but I am not making my decision based on this random quantities. I am just playing some arms in every round according to some and I has also have to be specific that they are chosen deterministically.

Yes, they are chosen without knowing what are this random quantity, and also that when I choose them I am not choosing randomizing my choices. In every time I am going to choose something deterministically.

For example, one strategy could be, ok, I am going to choose my arms in a round robin fashion in the first round I choose d_1, in the second round I chose d_2, the third round I chose d_3 and after I exhaust all of them and then I will come back and choose. Let us assume that is one such deterministic case where we have just ignore what has been happened and you have been just simply choosing.

Let us see if I applied this what is that I am going to get. So, what is σ_1^2 for me? σ_1^2 for me is nothing, but $\sum_{\{s=1\}}^{t} < d, V_t^{-1} d_s > \eta_s \sim \sqrt{\sum < d, V_t^{-1} d_s >^2} -$ Sub gaussian.

Into 1, right because each η_s is 1 sigma Gaussian, so because of that I only have to worry about it, ok. So, what I did? I have messed up. This is only d here. So, now, let us go back to that. Now, because of that let me write it here. Probability that I am just repeating that quantity there, this being greater than 2 times. So, this probability should be what?

This should be upper bounded by I am now just going to apply this quantity here. What is ϵ for me? ϵ for me is this quantity over here. So, I am going to get exp, ϵ is, so after than the square I am going to get 2 times and there is of course, log there is a minus here because of this minus log(1/δ). And what is this quantity here? 2. What is σ_1^2? σ_1^2 is this entire quantity for me that is; so and this is a square term here.

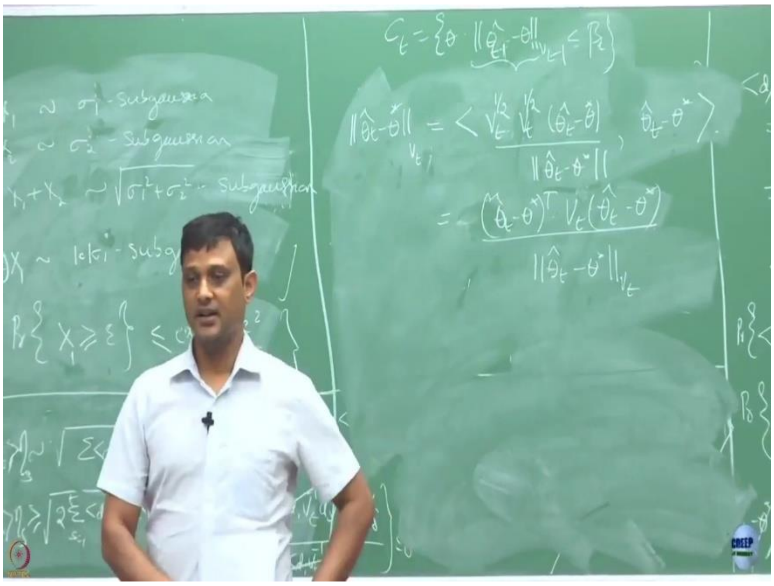
(Refer Slide Time: 33:12)

So, now, if you see that this cancels, this cancels. What remains here for me is only exp minus log(1/ δ)

So, now let us try to understand this. What we are able to argue is if you are going to look at this quantity which is nothing, but the projection of my decision and the error term being larger than; this quantity is very small that is it is upper bounded by δ. So, what we are just showing is this projection will not be larger than this quantity, and that and that depends on this whatever that δ parameter we have chosen, ok.

Now, we want to now convert it into whatever this result we have. Now, we want to convert it to into the case where how to come up with this decision set, right. So, let us let us see ah. So, let us say, so this is one part we have shown. So, what have shown now equals to δ. So, we have this part. Now, we want to translate. So, in my confidence set if you recall what are the C_t I had, I had the terms like.

(Refer Slide Time: 34:47)



So, I need to first, now try to get how does this term look like, ok. So, what is the probability? So, what do you want? We want this happening for if it happens, if I consider a set which satisfies this condition, I want to show that my θ* contains this with high probability, right.

So, now whatever we have shown here now try to connect that with this quantity, $\widehat{\theta_t}$ - θ*. I now instead of θ I want to specifically want to show that if I take θ* this will also satisfy this condition that is why that θ is contained in that. So, now, let us take my θ* here. And now let us try to connect whatever we have with this quantity, ok. Now, just let me take this θ here. I think this is θ - 1.

So, this can be written as θ_t half sum Y, I will tell you what is this Y and, or maybe just we can directly write it as did not you d_t half into $\widehat{\theta_t}$ minus θ* this whole divided by $\widehat{\theta_t}$ minus θ* comma $\widehat{\theta_t}$ minus θ*. So, do you agree that if I take this in the product this is the exactly equals to this. So, what is this? This is simply V_t matrix, right. Just take. So, and denominator is a constant, ok.

So, now if you are going to transpose of this quantity this is this quantity transpose into V t multiplied by this quantity. So, this is nothing but what? So, the numerator is transpose V_t transpose, but that is symmetric matrix. So, I do not care about this. Then, $\widehat{\theta_t}$ minus θ* this whole quantity divided by $\widehat{\theta_t}$ minus θ*, right. But numerator is nothing, but the squared quantity of the denominator.

Student: (Refer Time: 38:23).

Where?

Student: (Refer Time: 38:24).

Numerator nothing but the square quantity of the denominator that is why we get it back.