Convex Optimization

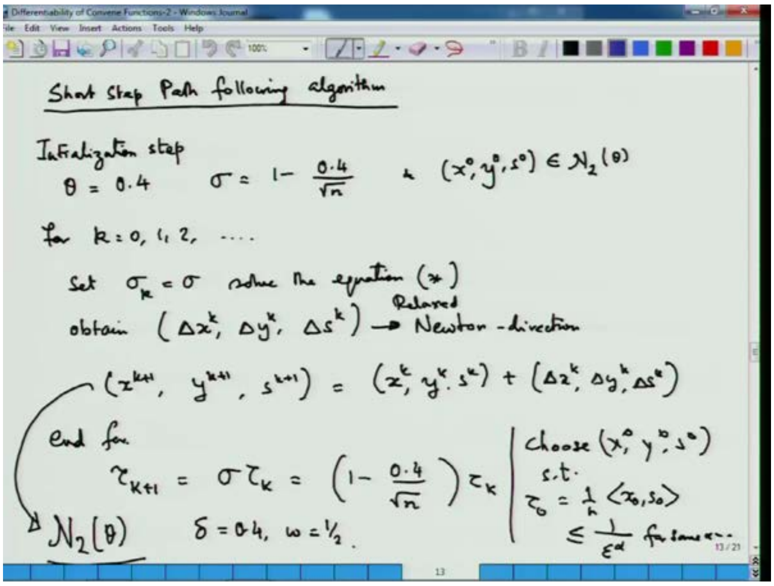
Prof. Joydeep Dutta

Department of Mathematics and Statistics

Indian Institute of Technology, Kanpur

Lecture No. # 32

(Refer Slide Time: 00:26)



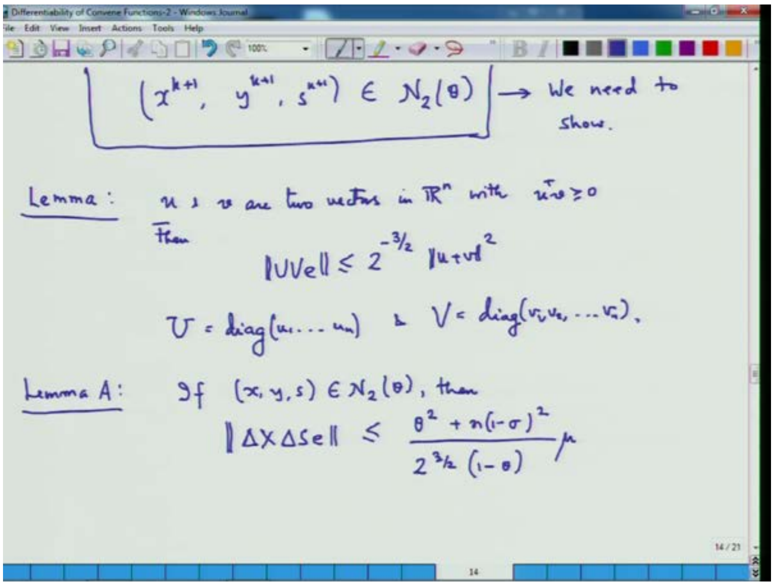
In the last class, we studied at the end the shorts type path following algorithm. We would just like to recollect it once more, that you have been initialization step. Where, your parameter theta which creates the never would end to theta, along with the sigma. The centering parameter is given, and you start with the point which is already in the neighborhood. Then you keep on doing these steps. We have shown that, we can actually show that tau k plus 1; the duality measure is related to tau k in this fashion. And once we have chosen delta to be this and omega to be half, we can show that we can and here naturally here delta is 0.4, and omega is half. Then we can easily show that this short step path for following algorithm is in polynomial time, because of the previous result, where we had to choose tau k plus 1 to be related to tau k in this fashion, and that is exactly what we had done.

Now, once you have an idea of this, the most important thing is that when you make this algorithmic moment from s k, y k, s k to s k plus 1, y k plus 1, s k plus 1 where alpha in this case 1. It is important that you should keep a note that this also has to go to N 2 theta. That is…

(No audio from 01:41 to 01:59)

Should belong to N 2 theta.

(Refer Slide Time: 01:12)



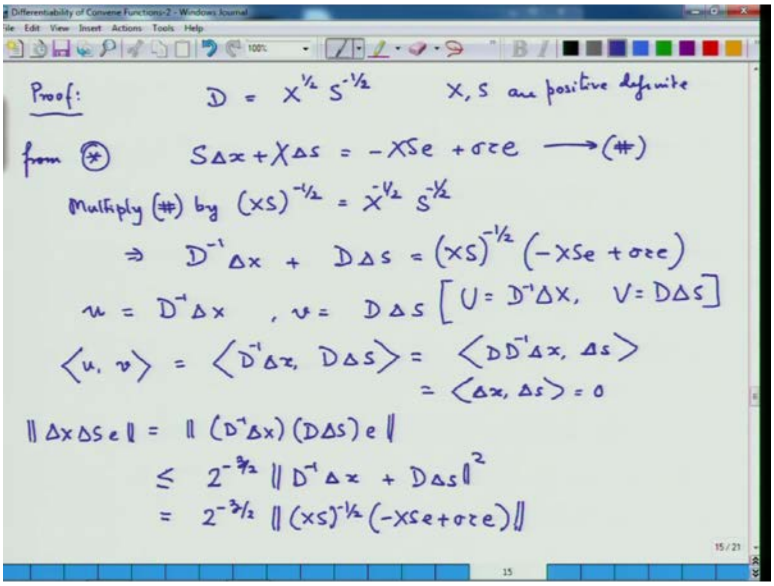
So, that. So, this is something we need to show. (No audio from 02:05 to 02:13) And this is where lies order optimization algorithms. The final algorithms slicing showing; that whatever you are doing is actually making sense. So, how do you show, whatever you are doing is an actually making sense? You do it by showing the basic requirements needed for your algorithm. And this is the basic requirement. So, what do we; do we here.

In order to do this we take some we had to prove certain things step by step. So, in order to do it, we will first come to prove a lemma not a prove, I will just mention this lemma. That if u and v are two vectors in R n (No audio from 03z:01 to 03:08) with u transpose v greater than equal to 0. Then U V e so, you I am writing capital U, capital V e. You know this is nothing but the diagonal, and this is nothing but the diagonal matrix; consisting of the components of v 1 and v in the diagonal. Inequalities look very strange, but these sorts of inequalities are helpful. Of course, one would say how do I figure out such in inequalities. The fact is that; you do not from the air develop this inequality. You try to prove that this is in this. And I will prove that you will see that you would require some estimate like this. And then when you then you try to such a thing is true. As before U is the diagonal matrix, and V is the diagonal matrix.

Now, what we are going to show is the following another lemma. So, lemma A: I am not being very particular about the numbering of lemmas and theorems, where it is idea which is more essential that knowing which theorem mono theorem to. So, what we are trying to now prove is the following that. If x, y and s is in N 2 theta, then change that is the grad x, grad y, grad s that you compute out of the Newton's step. The Newton step itself, then you take the diagonal matrixes consisting of the components of grad x. Diagonal matrix component consisting of the component of grad s. And these are these things.

This is less than theta square plus n into 1 minus sigma whole square 2 to the power 3 by 2… Of course, you can ask me, how should I prove this?

(Refer Slide Time: 05:55)



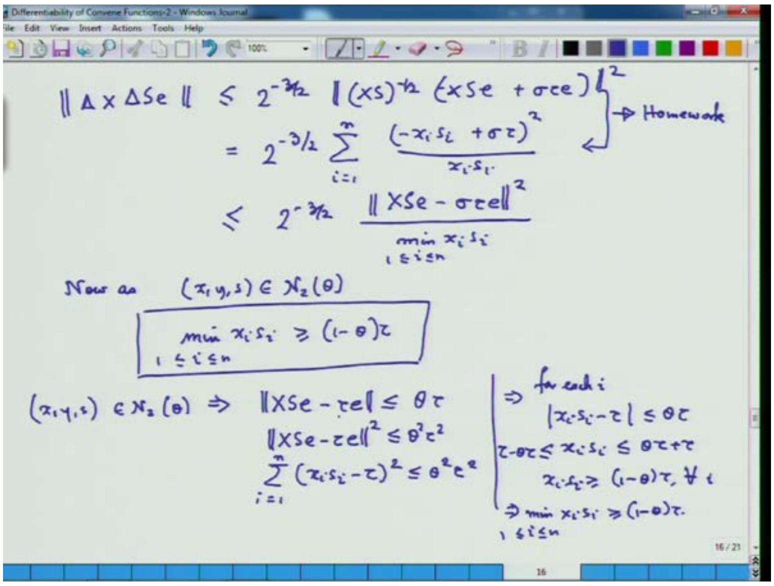
So, I have to prove, what I have just mentioned, in order to prove this. I take this matrix D as X half and S minus half. So, you know this X and S both are positive definite. So, there is a matrix whose square is nothing but a X. So, that is X half, there is the positive definite matrix square is X half. So, if I do that, then what we do is look at the last thing that, you get from the Newton equation. The last expression that from star, we have to go back to earlier notes; that when you go back to the last expression of the relax Newton system, that you have solving is this is minus X S e plus sigma tau e.

Now, once you know this so, I will multiply… So, this equation I can call it as hash. So, multiply hash by X S to the power minus half, that is basically X minus half S minus half. Because S half and X half of both positive definite, and hence invariable. If that happen, then what you will get? This would imply that taking D as this expression. The D inverse delta x plus D of delta s is X S minus half into minus X S e plus sigma tau e. Now you observe your two vectors.

So, you take u to be D inverse delta x, and v to be a vector D inverse delta s. So, what is u transpose v? That is something you have to observe. So, observe the fact u in a product of v. (no audio from 08:26 to 08:38) so, this is because D transpose is D here, because a symmetric matrixes. So, this is identity. So, this is nothing but delta s delta x delta s, and you know that this in a product is already we have proved earlier, that this is nothing but 0. We are studying the general format of the IP algorithm; we had proved that is estimate 0. See where this estimated actually helping. Now, once you know these; you can now write delta s e. This can be written as D inverse delta x… These are diagonal matrixes and all are invariable.

Now, all are commuting. So, D of delta s into e; this is the capital U, capital V basically. So, of course, from we will immediately know the capital U is D inverse delta x, and capital V is D of delta s. So, once you know this you apply; what we have just learned, because the product is equal to 0. So, this is nothing but less than 2 to the power minus 3 by 2 times u v whole square. So, u is my D inverse delta x plus D of delta s whole square. So, this is same as writing 2 to the power minus 3 by 2 into… Now this thing is nothing but this; so again write this as X S minus half into minus X e plus sigma tau e.

(Refer Slide Time: 10:39)



Once I do this; I can now write again, I will repeat what I have return in the last page. So, that you have do not loose the link. We have show that; this is less than 2 to the power minus 3 by 2 into norm of X S minus half into X S minus X S e plus sigma tau e whole square. So, this can be now returning as 2 to the power minus 3 by 2. I am opening the I am taking that this to be the standard Euclidean norm, and I am opening up the Euclidean norm. So, you can open up write the whole thing in the Euclidean norm, this is what I will get finally. (No audio from 11:27 to 11:37)

So, from here to here, this part you figure it out as homework. Now, once you have done this; this actually means the following. This is less than… Because this meant thing will come out, so, this is less than…

(No audio from 12:03 to 12:19)

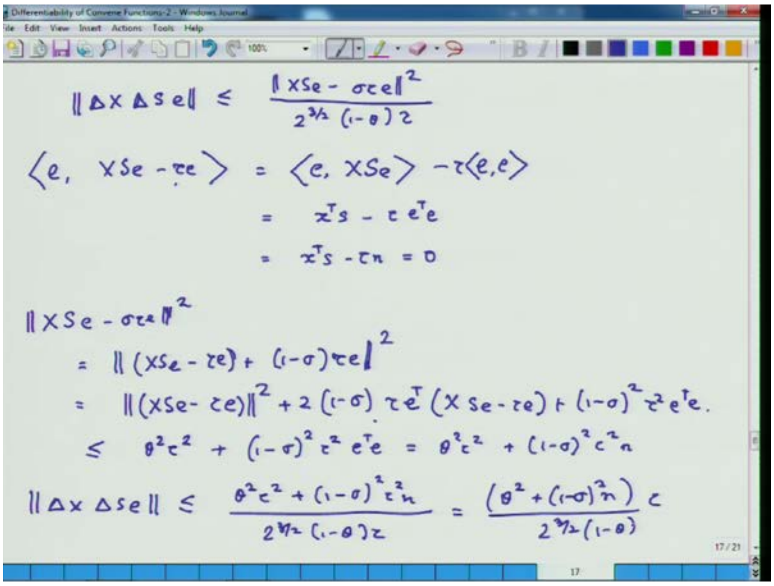
This is something you mean easily prove.

Now, as what you have is that x, y, s; is element of N 2 theta. So, the next y, s is element of N 2 theta, this will give us the following. (No audio from 12:43 to 12:55) How does it give me so, so; if you observe, what will happen? x, y, s element of N 2 theta is implying that not mu sorry tau. This is tau sorry. So, once this is there, it this simply tells X S e minus tau e, the norm 2; that is standard Euclidean norm. I am not writing, because will not using any other norm is less than equal to theta times tau.

So, what is what will happen here is the following? I can now write this whole thing as; norm X S e minus tau e whole square less then equal to theta square tau square. So, this means, summation x i s i minus tau whole square. So, which would imply, these are all non-negative quantities this implies that for each I. x i s i minus tau is less than equal to theta term; which means x i s i, it is less than theta tau plus tau less than tau minus theta tau.

So, x i s i is bigger than equal to 1 minus theta tau for all I, between 1 to n. So, this would imply that the mean of x i s i with i learning for 1 to n is bigger than equal to 1 minus theta tau. And this is exactly, what we intended to proof, for this what. So, now once I know this fact. So, my expression will become very simple.

(Refer Slide Time: 15:17)



So, once I know this fact. So, what does it mean? It means that non delta X delta S e is less than equal to I can put 2 to the power minus 3 by 2 here into 1 minus theta tau. And then we have to really compute out this part, X S e minus sigma tau e whole square. This is what, we essentially have to compute. So, also observed the fact; that if I take the inner product of minus tau e, then sorry I make a mistake, then what will get is e into X S e minus tau into e, e. What will this give me? This will simply give me x transposes s minus tau e transpose e, which is n.

So, this is nothing but x transpose s minus tau n, and this is nothing but 0, because x transpose s phi n is equal to tau the duality measure. So, which means that now, if I compute this one, let us, we can use this fact inside it. (No audio from 16:57 to 17:05) So, this is equal to X S e. I am a making little mistake and confusing, sigma was used for difference purpose. Here sigma x X S e…

(No audio from 17:20 to 17:38)

1 minus sigma… (No audio from 17:41 to 17:48) I make mistake. It is tau e 1 minus sigma tau, were we just arrange the things slightly. So, we have added a tau e subtracted tau e, and then we… Once have this; I can now write this as the standards. We have compute norms 1 square of norms…

(No audio from 18:18 to 18:37)

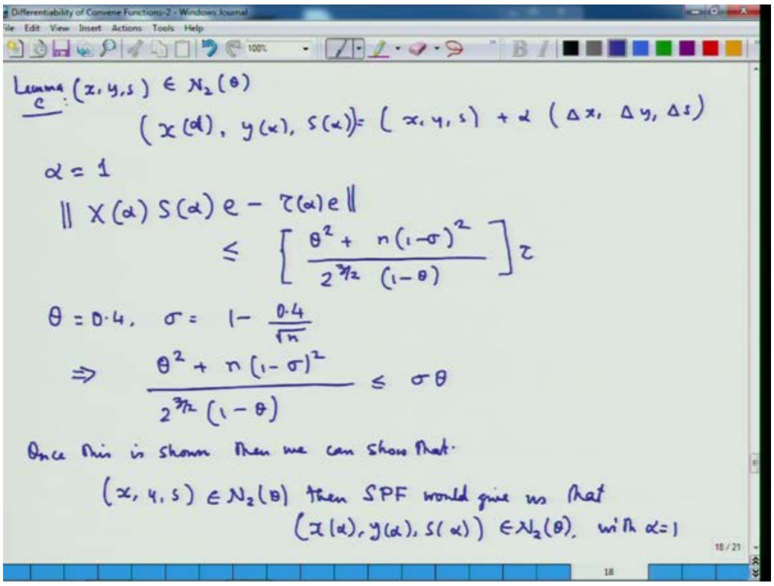
Now, e transpose tau e transpose X S e minus e tau e plus 1 minus sigma whole square mu square e transpose e… (No audio from 18:54 to 19:02) Now this part is 0, because of this fact. Now, we have this part which by into theta, because x x, y, s is into theta. This is nothing but less than theta square tau square, plus I have 1 minus sigma whole square tau square e transpose e. e transpose is nothing but n. So, this is nothing but theta square tau square plus 1 minus sigma square tau square n.

Now, once this is done, we are almost true, because one this is done; it will show that, I can put this whole thing that take the tau square common (( )) with the tau. So, and that n is our result.

(No audio from 20:01 to 20:19)

Tau square is taken now would so, I will just our tau. So, enough theta square plus 1 minus sigma square n time's tau by 2 to the power 3 by 2 into 1 minus theta. So, this is the first estimate that is required to prove our go to our goal. Our second estimates; which will say the following the fact. That is let me know for a current set up. So here, what I am trying to do, is instead of doing the whole thing for N 2 theta, I can actually take x, y, s in N 2 theta. I can actually show it for a general case go do. I will just do it for our own story here.

(Refer Slide Time: 21:08)



So, if, so this is my again I have a lemma. So, it is possible lemma C. So, if x, y, s is in N 2 theta, and your running the short step algorithm. You will have the following thing x beyond quality alpha, will call it X k plus 1, I would guess. So, I am just calling it alpha, but alpha is 1. May be you will run confuse. So, now basically we will have… Since, We have we always write instead of X k plus 1 just short an our efforts. We always write… (No audio from 21:54 to 22:03) plus the Newton step, original i direction plus alpha times in a Newton step. Now what we have going to do, our alpha here is 1?

So, with the alpha equal to 1, we will have… minus tau alpha e. Tau alpha is of course, you will you understands what. Inner product of these two is y n. This is actually less than equal to 1 alpha is equal to 1. It is again less than…

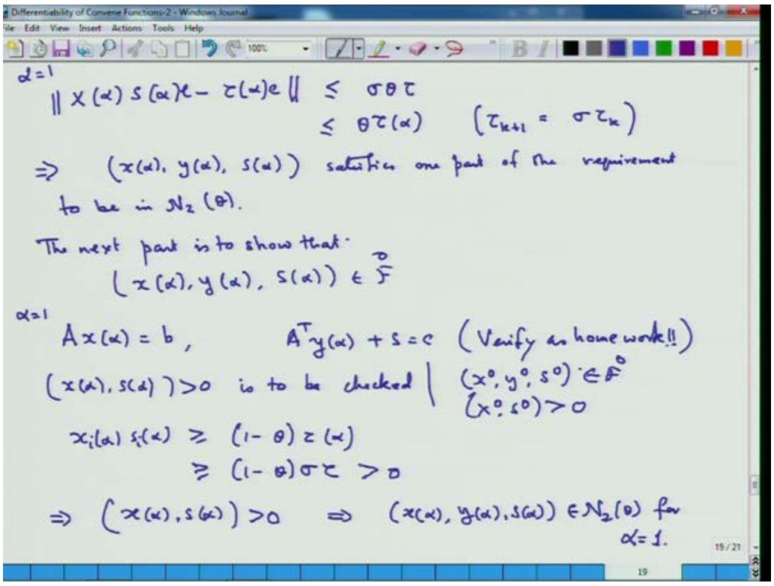
(No audio from 22:54 to 23:12) Times it.

So, you see the Newton step and the one which is when alpha is 1, the one which is the higher one. So, this is having the same relationship. Here what is happening? Here this is same. So, here I am taking an estimate of the product of the Newton step, and this is x n s component. I am having the same sort of estimate, when alpha is equal to 1. What we can once we know this estimate, we are not giving a no proof of this.

Now, if we what we have to do? It is that; we have to choose. So, our theta is chosen to be in the algorithm, our theta is chosen to be 0.4, theta is chosen to be 0.4 sigma is chosen to be 1 minus… were just come sigma is chosen to be 1 minus 0.4 by root n, which is fixed. Once I choose these I can show, then it would imply that theta square plus n 1 minus sigma whole square 2 to the power 3 by 2 into 1 minus theta. So, these are the quantities, which are not known here. If you put on these quantities, then this is a less than sigma time's theta.

Once this is shown, then we can show that; whenever x, y, s is element to N 2 theta, then the short step path following algorithm. So, SPF would give us that x alpha, y alpha, s alpha is also an element of N 2 theta, with alpha is equal to 1. Of course, you can prove it with any other alpha, also what then this estimation will changes a lot. So, this is how you continue to do it? So, how you do this fact? Let us just try to do the proof.

(Refer Slide Time: 26:21)



So, let me look into the fact this fact. Now with alpha equal to 1, I have x alpha s alpha minus tau alpha x alpha s alpha e minus tau alpha e. e is the vector 1 1 1. I want to remain you is less than equal to… If you good go by this estimate, this whole estimates less than sigma theta. So, it is less than equal to sigma theta into tau. So, if you look at the relation between tau k plus 1 and tau k, in the short step path following algorithm, we know that tau k plus 1 is equal to sigma into tau k. that was our.

So, here sigma into tau k is would give me tau alpha, because we just k plus 1 here is alpha is being that true. So, that is what you get? So, what you have proved? You have proved that x alpha, y alpha, s alpha satisfies one express satisfies one part of the requirement to be N 2 theta. (No audio from 28:00 to 28:10) Now, other part is to see that the next part…

(No audio from 28:16 top 28:29)

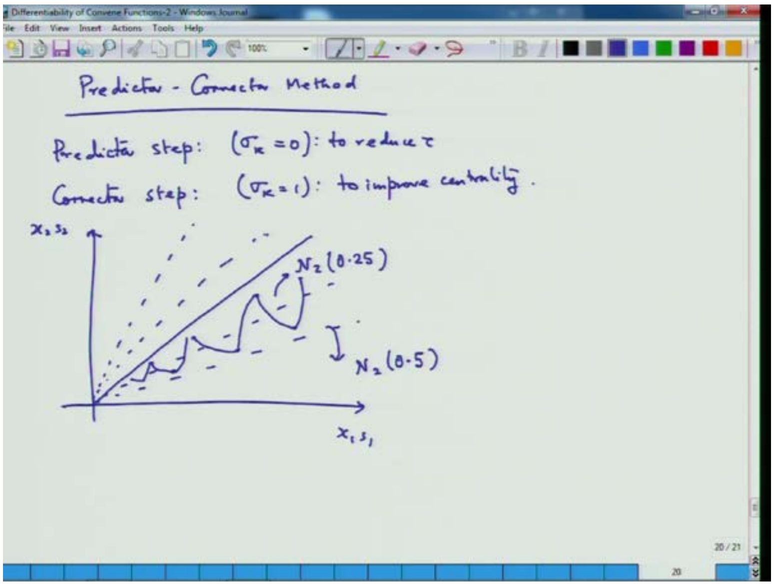
So, that this one... So, this is what you have to show? Now how do I show that? I leave it you to verify that A of x alpha is equal to b, alpha is equal to 1. And A transpose y alpha plus s is equal to C, Verify as homework. (No audio from 28:56 to 29:04) What we have to now check? It is the positivity of… is to be checked. Now what you will get? What we will get is the following.

So, when we start with x naught and s naught, they are strictly bigger than 0. So, the mu is strictly bigger than 0, and so you start on. So, Mu alpha mu is strictly bigger than 0. So, what you get here is sorry not mu tau is strictly, because you start with us points. So, when you start, you start with x 0, y 0, s 0. This is in f naught. So, x naught f naught is x naught f naught pair x naught s naught pair is already a vector with positive components.

So, now we will use this fact that we have just proved, if you basis back. It is that x alpha, s alpha for every component of that vector. That this product is bigger than 1 minus theta into tau or alpha, but what is tau alpha? Tau alpha is nothing but in this case sigma times tau. Now theta is positive, in our case theta is 0.4, and sigma is positive and tau is positive. So, this is strictly bigger than 0. So, none of them can be here 0, because all of them are greater than equal to 0. Any one of them any one of them is 0, if either if one of them is 0, then the product is 0. So, it cannot be so.

So, which means x alpha, s alpha is strictly bigger than 0. And this would finally, imply that x alpha, y alpha, and x alpha is element of N 2 theta, for alpha is equal to 1. So, with this is finish or understanding of the short step path following module. But the problem with the short step path following model is not really implementable model. It is the toy model, allowing you to understand how to do an analysis after that. But another model which is slightly better than this which can be which latter on leads to amount practical algorithm; it is the predictor corrector method.

(Refer Slide Time: 31:47)



(No audio from 31:44 to 31:55) So, in the predictor corrector method; of course that a two-step predictor step and the corrector step; so in the predictor step so, here my sigma values; the centering parameter values will change. Sigma k at an every predicted step is 0, is done the idea of this step is to reduce tau, and then this is the corrector step. Where it sigma k equal to 1; idea is to improve the centrality, because there is a centering parameter. And take it force the thing to a more towards the central path.

Now, this is something we will discuss tomorrow, because here what we do is that we used two values of thetas also. One is for the predictors step, another is for the corrector step; the fact is that we start with when we start, we initialize we start with a point in one neighborhood, and then when we improve our values, we go to another neighborhood. So, there starting neighborhood, and there is the predictor neighborhood, and there is a corrector neighborhood.

The fact is that even if you predicts one in means the sense that you start from one neighborhood; go to another neighborhood, start from other neighborhood; go to another neighborhood and so on; that is the way it goes. Basically, if you look at the x s space so, here is my central path. So, here I will have two neighborhoods. Slightly bigger neighborhood which is N 2(0.5); theta is 0.5, and slightly us smaller neighborhood at this is N 2 (0.25), slightly more smaller. So, you start from something here, start from us a smaller neighborhood come to a bigger neighborhood. Then again go back to a smaller neighborhood, again come back to the bigger neighborhood, and again go back to the smaller neighborhood and so on. You are able to takes steps, which is slightly bigger than the short step path following algorithm.

So, instead of writing this thing in detail, this is just a very basic expansion that we have given. So, tomorrow we will describe the predictor corrector algorithm, and with that and also the long step path following algorithm. And with this we will end the fact that we will end our story of the path following interior point methods. Once we have done that from the class on words next class, after the next one there is day after tomorrow class, we will talk about semi definite programming.

We will take two three class to talk about a very exiting and important an area of convex optimization call semi definite programming. And then we will give some miscellaneous resent developments of convex optimization, in the last two hour two classes. So, we I think with this we would have some six classes left, and with it tomorrows class they will be one class gone. So, they are five classes three semi definite programming, and two certain miscellaneous stories about optimization, and that would end the course. So, thank you very much.