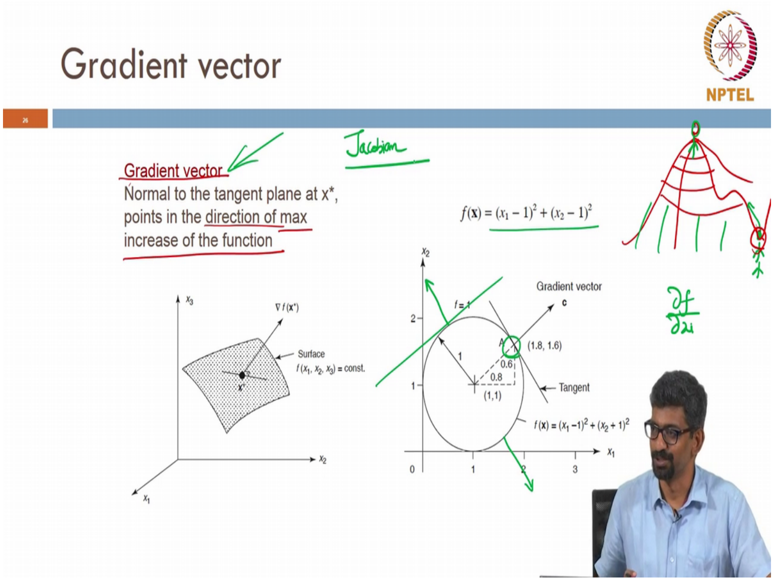
Surrogates and Approximations in Engineering Design
Prof. Palaniappan Ramu
Department of Engineering Design
Indian Institute of Technology, Madras

Lecture - 04
"Calculus related to Optimization"

(Refer Slide Time: 00:18)



So, some quick calculus stuff. So, basically we are always interested in something the words tangent gradient. We are interested in the gradient information, why? The example that we discussed now should tell you why I will be interested in the gradient information.

Student: Normal (Refer Time: 00:50) where it in increase and that is decrease.

Because gradient tells you, what information that gradient says?

Student: Whether it is we have it is maxima.

So, it gives me this information, the direction of maximum increase in a 2-D, problem it was very clear, ok. Now imagine you have a hill like that. An optimization problem can be described like that, ok. This is how an algorithm starts to work in this is you, who is trying to climb this. So, you can start here when I say there it could be anywhere here, ok. And assume that you cannot see the you cannot see, meaning you cannot see means I am blind folding you. And then you have a stick in your hand that is a usual example that we give, ok. So, what you will do is this? You look for the maximum increase because; you also want to reach there.

Student: (Refer Time: 02:10).

As fast as you can, ok; so, what you will do is, this you will take your stick and you will do like this, and you will find out whether the maximum increases here, here, here, here, which direction it is. Let us assume that this is where the maximum is. So, you can you will start going in this direction. Instead of going step by step you can also jump. So, will you jump for one feet, will you jump assuming that you can jump, ok, like how many our feet, it is will you jump for half of feet will you jump for one feet, will you jump for 10 feet. Because you do not know right I have put a value here. Instead you could take a smaller slope, smaller gradient and you might be able to reach faster, ok. So, that is called line search so, this is a direction.

The direction in which you need to start, because I blindfold you I rotate you and I leave you use the stick and then you have 360 degrees you can climb if you can climb. So, and then you find out, ok. So, let us even put this slightly differently. So, let us say it is actually a valley kind of a stuff. So, you are inside the valley, ok. So, you can check all sides and you are you are free to get on whatever side it is. So, which one will you make a decision, where there is a increase number one and if you have multiple increases which one will you choose.

Student: (Refer Time: 03:40).

If I did not put that valley on one half of are like a semicircle area, it will be going down. So, that tells you like I should not go that said because I want to go to the top of the hill. So, you will select only the other half, in that other half if you have multiple slopes that are going up which one will you choose, normally the one do not ask me is a 90 degree can I climb, that is a different thing the assumption is you can climb, ok.

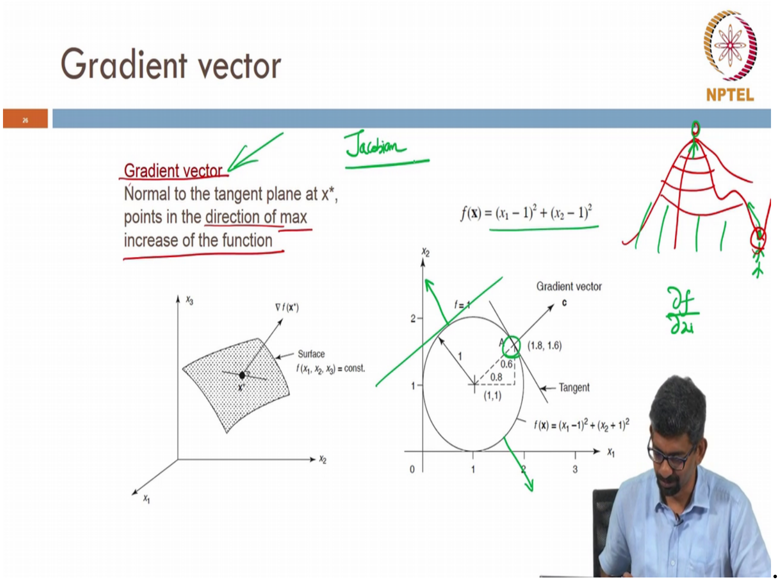
So, you will take the one that can take you the fastest that is one. The second is how far do you go in that? That is called line size that will come, the direction is given by your gradient, but what information does it say? It says it is increasing; now just take your hill example outside represent it as a function. So, as a design point, this green guy who you wear is a design point, ok. The algorithm looks for the design point which side it should go. So, it will choose the direction of the gradient, ok. So, that is important the direction of the, sorry the grade in the direction of the gradient means it will it tells you the direction in which the function is increasing if you are minimizing you will actually go in the.

Student: Opposite.

Opposite direction, imagine that you are here you are a top here, ok. So, then you want to go to the valley, ok. So, what will it will say that the function is increasing in this direction? So, you will take the opposite direction, because it is supposed to decrease that side it is a very simple. So, hence you need to have an understanding about a gradient vector. So, and you know all this, right it is normal to the tangent plane. So, if I have a surface there is a tangent it is normal to the tangent plane that is one thing. So, just to give you an idea if this is a circle that we are talking about, ok. So, at this particular point this is the tangency, the circle can increase in all sides. So, you are talking only specific to this point here, ok. If I had another constraint that is going like this, then it is increasing in this direction.

So, these are all more like local quantities. So, oftentimes when you are talking about optimization, you are talking only in the neighborhood of the particular point. So, sometimes global optimization (Refer Time: 05:52) I mean, because it is only in the neighborhood that is where you get local optimas. So, at this point this is a tangent and this is my gradient vector, ok. So, it is increasing in this point it is increasing in this direction, if you go to this point it will increase in this direction, ok. That is something that you need to appreciate. Now this is the first order of business; obviously, what is the second order, it is d squared f, ok.

(Refer Slide Time: 06:27)



This is.

Student: (Refer Time: 06:24).

The other one is.

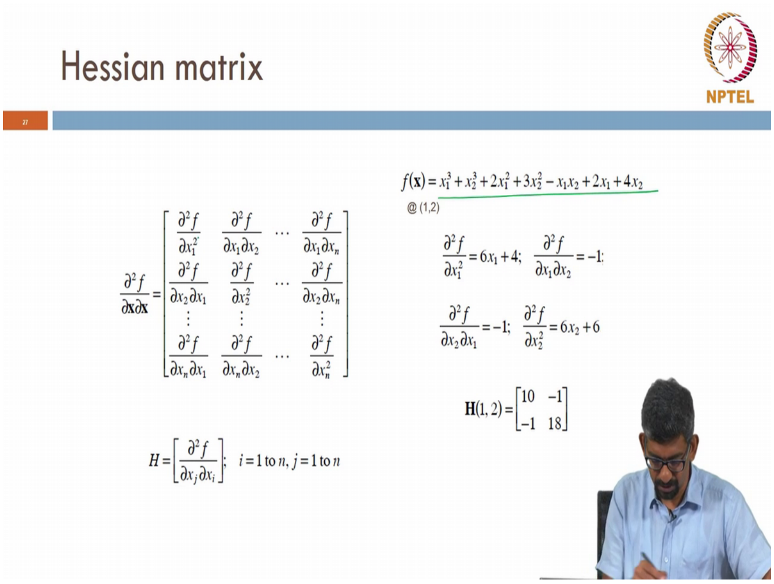
Student: (Refer Time: 06:28).

Hessian what is this guy called, gradient is also called as.

Student: (Refer Time: 06:32).

Do not use a word slope anymore, you are engineers that to graduate engineers. So, use the word gradient, it is more more n dimensional, ok.

(Refer Slide Time: 06:50)



So, the second order partial differentiation is your hessian matrix, ok. This information is also required. So, what does notionally, what does hessian vector give you? Because, this guy we know gradient means what gradient means what.

Student: (Refer Time: 07:07).

What does a gradient mean? Simplest notional understanding of what a gradient is.

Student: Direction.

Direction is actually we derive it, ok. So, when I say gradient see people use sometimes while constructing this, you know, what you call inclusive floors and all that, they say the gradient of this wall is this much.

Student: Slope.

So, it is slope, so, in n dimensional; that means, slope. So, basically it tells you what the slope is, by how much it is increasing the rate at which the function is increasing that is what your gradient information is telling. So, what should the hessian give? The hessian should give something more than that, what is that?

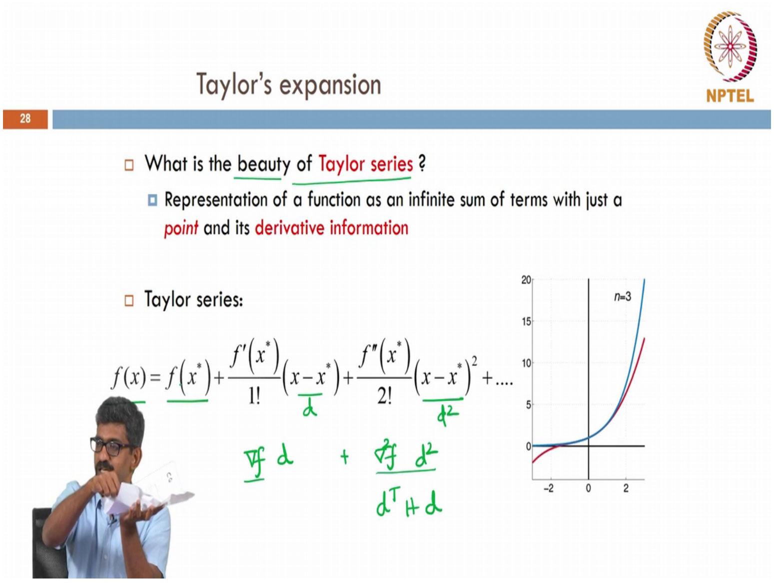
Student: Concavity.

So, it captures the convexity. So, this guy is the tangent not tangential information, the tangent based information, right so, that is slope. So, what the hessian says? Is it gives you the second order, the same information, but it gives you a higher order information, ok. So, someone comes to the once to apply for your department, comes and asks you what the entry criteria is, that is the first order information.

If they go and ask your professor, sir what is the entry criteria? That is the second order information; they go and ask your HOD that is the third level information, ok. So, similar, they all are the same thing, but they give more and more information as you go. So, hessian gives for the same problem and gives you a second order information that order is important, right. So, I want to do a second order. So, this gives as he pointed out it is the curvature information that you look at, I guess you guys remember your dou y by dou x equal to 0, from your 12 standard 10th standard whenever you studied that, ok.

Whenever you want to do minimum maximum what do you do? You take a function you take dou f by dou x equal to 0, that is because that is what the slope is, and then you do dou squared f by dou x squared to find out because both for minima and maxima that is the same thing, let me come back to that. So, this is just tells you like if you take a function like this. You want to construct dou f dou squared f. So, this is, this is this is here written right.

(Refer Slide Time: 09:24)



Now, we will just step, one step in the side we will try to talk about something called the Taylor series. And then we will go back and see why this is important in optimization. As the question here what is the beauty of Taylor series? Maybe you can answer yeah you do you two cannot answer, yeah.

Student: Can approximate (Refer Time: 09:51).

Sorry?

Student: (Refer Time: 09:56) approximate.

Taylor series is an approximation; approximation of what?

Student: Sir (Refer Time: 10:00) it can approximate the neighborhood values (Refer Time: 10:03).

Fine, very good, it is an approximation of the function that is a very important point that he said in the neighborhood, ok. What information does it require? That is where there.

Student: (Refer Time: 10:19).

Because there are a lot of function approximations, but Taylor series is very interesting for optimization, for the very fact that he pointed out neighborhood. It is in the immediate neighborhood I do not know I cannot tell you where what is the function value? There in the immediate neighborhood of where I am standing I will tell you what the function value is, but I need some information what is that.

Student: Ok.

You need the.

Student: (Refer Time: 10:42).

Of course the function value of where I am standing, and then.

Student: Gradient.

The gradient information, you just need to give me these 2 values, and I will give you the function value in the immediate neighborhood, please understand. It is a very complex function I do not know how it looks like, or it could be a linear function also I do not know that. But wherever I am standing you just need to tell me what that function value is, and you need to give me the gradient information. And I will tell you in the immediate neighborhood how my function looks like.

This is all is required oftentimes; you do not need to have a overall understanding of what your function is in. The neighborhood is what I you need to know what my function values. That is what your Taylor series does. It is a representation of a function as an infinite sum of terms the interesting point is, you say why you are stopping only with gradient information I will give you hessian good I can approximate my function even better. So, it is a it is an infinite sum of terms with higher order terms.

So, this is your function value at the new point. This is your old point, gradient information divided by the difference the second order information divided by d squared, ok. So, if you would write you would write it like this. Del f t plus, or you can write it as how can you write this one.

Student: (Refer Time: 12:05).

That is a quadratic form of writing it. You understood what I have written, right. So, del f means gradient information times d x minus x star, it is the distance between the point, that you want and the actual point where you know the function that is d, and the square of that is d squared. So, this information f double prime is nothing but your hessian second order information. So, I can write it as del squared f d squared, but in quadratic sense this del squared f is nothing but your Hessian h I am just giving the representation h, ok. And this may be what is the usual.

Student: (Refer Time: 12:41).

G.

Student: (Refer Time: 12:44).

Jacobian.

Student: J j (Refer Time: 12:47).

You can give, yeah you can give g or something, but it will confuse here. So, you just keep it as del f only. So, hessian is h in this particular case. That is all it is quadratic form d transpose, because this d is a vector, please understand, because I am just writing it as x, but x could be x 1, x 2, x 3 it is in space. There is a point in space. Here is a function value, x star what is the value. So, you have a plain paper someone, all my you just give me a plain paper. Because I have papers, but it is all than just one second that is,, that is. This is a non-linear function; imagine that this is a non-linear function.

Now, ok, do not worry, this way you do not need to know, because you cannot see projection right. So, I am interested in approximating my function value here, ok. So, this is x 1, this is x 2, ok. And this guy is x 2 and this guy is x 1, meaning for you it is x 1. This is x 1 for you and this is x 2, ok. What I am saying is any function; any value in this surface is described by f of x 1 comma x 2.

So, if I want to approximate my function value here, where I know the function value here, what I am asking is give me that function value plus what is the gradient information. So, what will be the; I am not asking the elements, but I am asking the number of elements in the gradient vector for this one right. So, the point that you need to understand why I took this example is because, I wrote this one, and I told it is a vector. Because when I have a point here, it is described by x 1-star x 2 star, and I want the value at x which is x 1 x 2. I want the function value at.

Student: (Refer Time: 15:06).

Please understand, in this whole point you do not have this function. That is something that you will have to appreciate. For the sake of discussion, I am showing you the function. In reality you do not have this function. So, when god comes what will you ask?

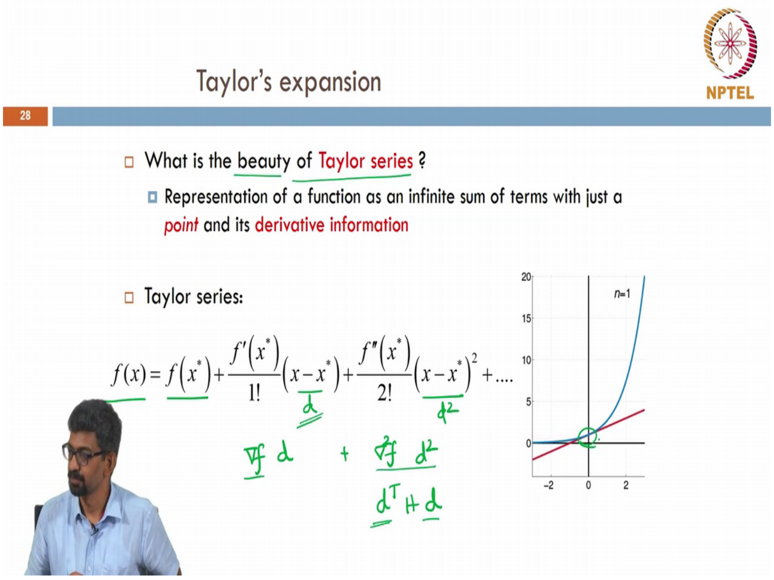
Student: (Refer Time: 15:23).

You say boss, please give me the f, this is what you told me right. Earlier you told me give me the f, you know the f then it is very easy. You do not even need Taylor series. If you do not have the function now I told, f values here, can you tell me what is the f value here? It could be linear, it could be quadratic, it could be non-linear like this, it could be anything, how will you know? You cannot interpolate just like that. You will get sucked, if you interpolate in this just like that. I gave you 4 values here interpolate you are gone your interpolation will go like this the actual function is in the other side, ok.

So, it essentially we are looking at interpolation on limit. Do not worry, will correlation there is a little bit of a correlation with your output function, and then you load is still interpolation only do not worry, it is not directly interpolation that is all, ok. You will you imagine whenever I show the function something that you need to record on your mind is you do not have the function. That is the whole point. Now if you want to go from point x star to x, ok, then it is a it is basically a direction or it is a distance between the 2 points. So, it is x 1 minus x 1 star x 2 minus x 2 star, that is why it is, ok.

So, it is a vector that is why I told that d transpose times d. You can write it like that when it is x minus x star squared in a matrix sense when you write it is called the quadratic form. You do not write x minus x star squared that is for single variable when it is a multidimension you write it like this. So, it is called the quadratic nature quadratic form, ok. So, this figure tells you what happens to the approximation when you use higher order terms.

(Refer Slide Time: 17:18)



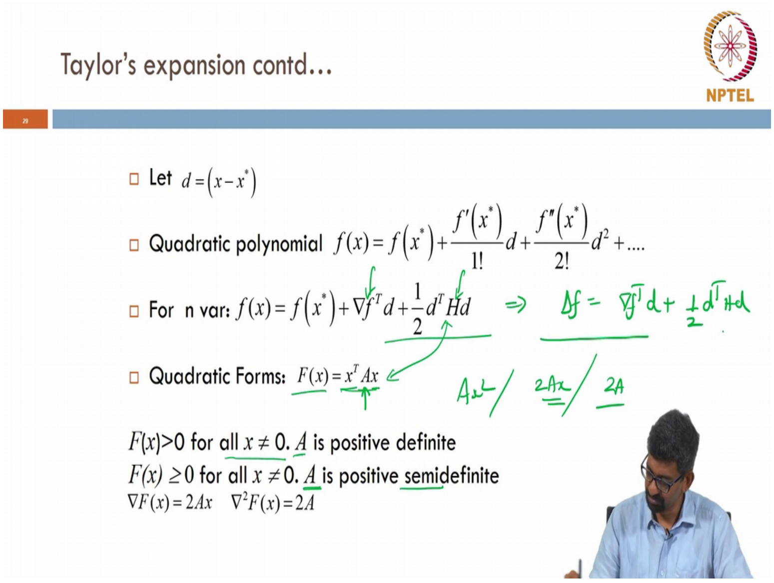
So, let us wait when you see n equal to 1, you saw that point right, at that neighborhood it will be correct.

Student: Ok.

You cannot you cannot say sir; at minus 2 it is a very bad upper 2 it will be a very bad approximation because, you have used only one value. The second it will use your gradient information, and it will still be only like this. Then n equal to 3 gets slightly better, when you go keep on going in this direction, for n equal to 7 and 8 it gets a very good approximation, ok. But do I really need an overall approximation? Maybe not, because that is what that is how optimization works.

What it says is, it looks for the stuff and then it goes to the new point, and then it has to evaluate that. So, it is only in the immediate neighborhood that I need to know. Because that is what you also did right you searched. Only in the immediate neighborhood you cannot have an infinite stick and search, there it does not make sense also.

(Refer Slide Time: 18:14)



So, Taylor's expansion is used widely in optimization. Because sometimes people say that oh this Taylor series allows you to cheat. Because most complex problem can be simply approximated function approximation using Taylor series, but of course, it has some disadvantages also, ok. Ok, this is this is what I wrote in the previous slide.

Student: (Refer Time: 18:39).

This understanding is important because we are going to define stuff based on gradient and hessian. Yeah, the definition comes here, at a later point we will come up with something called necessary condition and sufficient condition. They will say what your gradient and hessian should take, ok. For that we will say your hessian should be positive definite, this is a matrix terminology, does any of you know what positive definiteness means ?

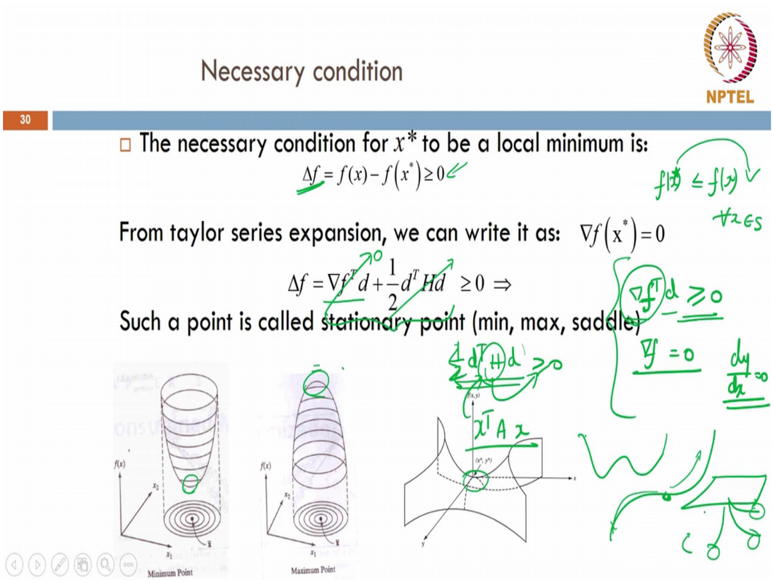
Student: It is transpose A x should be positive.

Yeah it is given here also. So, if you take something like x x is a vector, A A matrix, and then x is a vector here. And you call this entire resultant as f, ok. If f is greater than 0, for this is important, for all x not equal to 0, ok. You cannot have x all the x should be positive greater than 0, then this A is called positive definite. If the x condition is again not equal to 0, but if your f of x is greater than or equal to 0, greater than are equal to 0 then a is called positive.

Student: Semi.

Semi definite, ok; the same thing also holds good for hessian, ok. Of course, it does not matter this is not also it holds good for hessian. Because this is how we are writing the hessian these please relate these two that is important. So, gradient of, this is just writing it that way, del squared f you can if you take this it is x squared. So, it will be 2A x that is for the gradient and if you do del squared f it is just 2A. You understood what I am saying, right? This is nothing but A x squared. So, if you take a first order differentiation it will be 2A x that is what we write here if you take a second order differentiation it will be 2A that is all. So, we will see where this positive definiteness and all that is come into picture.

(Refer Slide Time: 10:53)



There is something called necessary condition and sufficient condition, very simply and quickly put what it means is, if someone who wants to get into your department, ok. You set up some kind of an entry criteria. That is necessary condition, ok, just because you came to the interview is there a guarantee that you will get selected you are only called a.

Student: (Refer Time: 21:16) Candidate.

You are only called a interview candidate.

Student: (Refer Time: 21:20).

You are not a successful candidate; you are only an interview candidate, ok. But that is the entry criteria so, but anyway there are lot of people who did not make the entry criteria. So, the necessary condition is the entry criteria. But there are multiple people, 100 people came, we selected only 12 people. One of you is you are one of those 12; I am just giving some number. So, 100 people came they were all candidates, and we selected only 12 of you. Because you are the ones who satisfied all and, sorry all the 100s satisfy those constraints first, but you only qualified to be the optimal combinations right.

So, necessary condition means it selects candidate points, ok. You have already come up with this necessary condition and sufficiency condition, but do you just do not know what you didn’t relate them as necessary insufficient. Because when I told have you, you have been introduced to minima and maxima earlier, by saying you know f of x. Then if you want to find the gradient you take dou f by dou x, if you want to find the minima what do you do?

Student: (Refer Time: 22:37).

You do.

Student: (Refer Time: 22:41).

No man for finding minima, I give you a function and I ask you to find the minima, what will you do?

Student: (Refer Time: 22:49).

What have you done in 12 standard.

Student: (Refer Time: 22:50) positive.

Positive definiteness? When positive definiteness does not all it is engineering, right you do not learn that in 10th or 12th. But in 10th or 12th itself we have done.

Student: (Refer Time: 23:01).

Minima maxima.

Student: (Refer Time: 23:03).

He is the honest guy, he says dou f by dou x equal to 0, that is what I will do, right?

Student: (Refer Time: 23:12).

But then, why do you do the second order?

Student: Find and check (Refer Time: 23:16), which point it is maxima (Refer Time: 23:17).

So, dou f by dou x equal to 0 is only a candidate, ok. I have a function like this, here also dou f by dou x equal to 0, here also dou f by dou x, what does it say? The slope does not change, that is what it says, ok. I give you a function like this. Take our roads potholed road as, wherever it is flat it is dou f by dou x equal to 0. So, all of them are candidate, but I do not know whether it is a hill or it is a valley. Whether it is a peak or a valley, I do not know depending on whether you are maximizing or minimizing I am interested in knowing whether it is a hill or a.

Student: Valley.

That is when I am interested to see whether it is going this way or it is coming that way. So, for that and curvature of course, that is when I need to know my d squared information. So, your first order information is what your necessary condition is. So, you are a candidate the peak is also a candidate the valley is also a candidate, ok. The curvature will tell me whether you qualified as a peak or a valley that is all this is. But we will see how it is extended to n dimensions that is all.

So, the necessary condition for x to be a local minima we already know this, I have written this for you. Do not look at this, we have we have written, right? I have said I would have written it in a slightly different way, I would have said f of x.

Student: X star equal to than x.

Yeah sorry, f of x star.

Student: Was greater than.

Should be minimum means should be less than or equal to f of x.

Student: For all

For all.

Student: X

For all x belonging to s, this is what I would have written, right. Now if I write it this way, if I take this f x star to the other side I can write it like this, correct? I am just calling this difference as delta f that is all. Now what I am saying is, from Taylor series, because I do not know these values, these are all theoretically any function you take this will hold good. This is only theorem, now what I am saying is from Taylor series. I can approximate this delta f, correct? Taylor series was that only, if you take this f x to this side.

Student: (Refer Time: 25:33).

That is delta f correct? So, this can be written as delta f equal to plus, ok. Half is not required, fine. That is what we are doing here, ok. So, I am just replacing this guy by a Taylor series expansion and I am saying it should be greater than 0, clear? This should be greater than 0, this is the idea that is when it is minima, I am replacing this delta f by the Taylor series expansion and I am saying it should be greater than 0. What does this mean? What is the meaning of this? Ok.

Student: (Refer Time: 26:17).

No, no not global minima and all. What is the meaning of this is, see what happens is, usually your d is very small, correct? You will not going to stand here and ask me what is on the outside of MSB what is the function value, you will ask me within this within this building, I mean within this room what is the function value that. So, d is usually small in terms of unit, in terms of percentage right. So, d square is going to be.

Student: Even smaller.

Even smaller, ok so, usually this term does not dominate, ok. So, it reduces the order of domination reduces assuming keep increasing your number of terms. So, we are not considering that at this point in time. So, we are only taking this guy. So, what does this say? There are only 2 quantities in the product, one is gradient of f, the other one is d.

Student: (Refer Time: 27:21).

Can d be negative or positive?

Student: (Refer Time: 27:30).

No, there is a point; you cannot say that you can always ask only on the positive side. I can also ask in the negative side.

Student: (Refer Time: 27:41).

So, there is no way that I can say that, correct? Ok, that is one point.

Student: (Refer Time: 27:49).

The other one is gradient of f.

Student: (Refer Time: 27:53) should be 0.

No, no gradient of f just do not jump into known solution. Is there a control on the gradient of f values?

Student: At x star (Refer Time: 28:09) at the candidate point x star (Refer Time: 28:11).

At the candidate point x star and that is derived from this. How do you know at the candidate point you will be 0?

Student: (Refer Time: 28:22) sir, if it is a minima or maxima.

Did I say that it is a minima or maxima yet.

Student: Sir, for yet to be a maxima or minima.

Correct, that is derived from this, ok. So, can d be equal to 0.

Student: (Refer Time: 28:37).

Does not make sense.

Student: (Refer Time: 28:43).

It can be 0, you can say yeah mathematically it can be 0, correct? But does it make sense you are evaluating at the same point. So, you are going to drop means, you are not dropping out of the equation, you are not considering d for the discussion anymore. So, the value that gradient of f takes becomes important, ok. So, it can be positive negative that is. So, the only thing that it can do for this equation to hold good. You cannot say gradient of f should be positive always can also to be negative, right. So, in order for this inequality to hold good gradient of f should be.

Student: 0.

This is what you already knew. You just didn’t know this in terms of gradient of f. You knew it in terms of.

Student: (Refer Time: 29:29).

Dy by dx equal to 0 so, if you understand what dy by dx physically means is it is flatness. Now it is just flatness in multidimension that is all. This is your necessary condition. Naturally sufficiency condition is what? Second order business in this equation what is the second order business.

Student: (Refer Time: 29:51).

You know this is to be equal to 0, that is what sorry, that is what you just saw this guy is equal to 0. So, for this equation to hold good what you should do?

Student: Half (Refer Time: 30:02).

Half d transpose d should be greater than or equal to 0. Do not worry about half.

Student: (Refer Time: 30:08).

D can be same story, d can be positive or negative you do not have a control on that then what does it mean ? H should be.

Student: Positive definiteness.

Positive definite or semi definite, it will take appropriate values that is all. So, that is what positive definiteness came into picture. Now you understand we just use the terminology x transpose A x. Now just replace that by this that is all. So, then if your f is meaning the outcome of this is greater than or equal to 0, then h is positive semi definite or definite if it is greater than 0. It is definite if it is a greater than or equal to 0 it is positive semi definite that is all. So, what this one says is, this is your I do not know some chips are coming these days, right.

Earlier we used to give the example of Pringles chips. You have seen the Pringles chips, today I mean yesterday or day before I saw on TV some chips are now coming in the same shape, but it is a packet. Have you seen the Pringles chips?

Student: Yes.

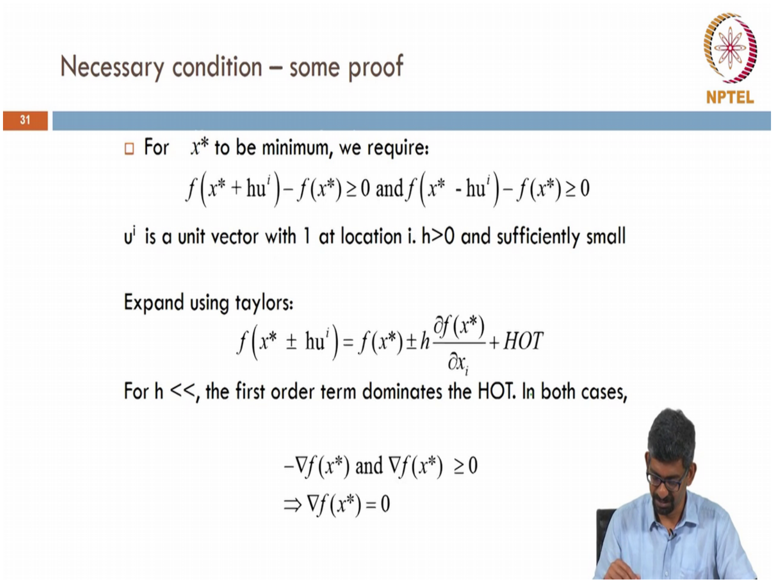
Pringles chips is the classical example of saddle point.

Student: (Refer Time: 31:18).

Used in control theory optimization, any function approximation, ok? It will be like this just like a horse saddle. The point is for this quantity for this quantity and for this quantity, it is gradient equal to 0 the 3 points, that I marked here the gradient is equal to 0. So, that is only candidate, it could be your minima it could be your maxima, it could be a inflection point or a what is an inflection point this is an inflection point you are here. You cannot say anything you move this side you will fall down; you go this side you will go up. So, it is an inflection point that is what this saddle point says, ok, that is all.

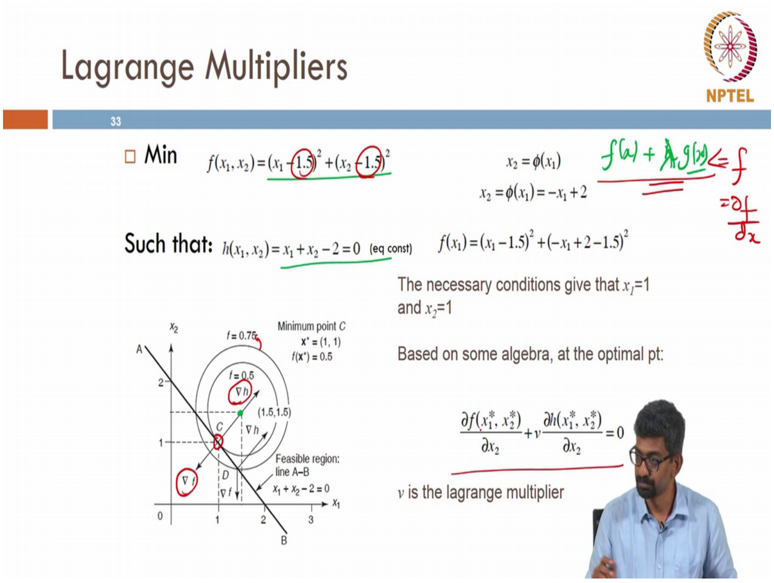
This is very interesting because in what you call like a double well systems. Classical example of double well system is if you go to these roadside dabas, ok. You are even no not in hours, if you see they will have table foot like this. On top of this the table will be there, ok. And either because of some issues in this one or on the surface in which it is landed, it will do like this, ok. So, it is stable in this end it is stable in this end. It will not topple. So, that is a double well system. It is stable here it is stable here, but in between it is a unstable. So, it is a classical example of a double well system. So, these are classically discussed in multiple cases not only in optimization theory. That is that is the reason that I showed that.

(Refer Slide Time: 32:58)



So, you do not have to worry about these necessary condition discussion, I am just gonna skip.

(Refer Slide Time: 33:02)



This it is the same idea, that is that is all I have done here, ok. So, there is a concept of Lagrangian multipliers which we might not use lot, but I will quickly go through it. So, what it says is there is, whatever we discussed no we have discussed constraint problems, yeah. So, usually we discuss only unconstrained problems first, and how do you solve constraint problem is you can actually reformulate the problem as your objective function plus something times your sorry lambda times that is what your Lagrangian multiplier is times your constraint, ok. Then it becomes an unconstrained problem.

So, this is this information whatever I am going to tell you is something that you already know, which we discussed from our circle example, ok. That is the same stuff here; they are using a circle as an objective function, and some straight line as a constraint. So, if you see this might be much better than my pictures that I drew. So, this is my center, this point here 1.5 comma 1.5 that is what the center is here, right. So, this is my center, and as I told the circle is expanding, ok. So, at this point the function value is 0.75, for this guy for this inner circle it is 0.5, at this dot it was 0. So, it is expanding in that direction.

Now, if you see this is your equation, but this is a slightly different one, it is an equality constraint. It is not an inequality constraint. So, what is the feasible domain in an equality constraint.

Student: On the line.

On the line.

Student: On the line

So, what it means? It says.

Student: (Refer Time: 34:57).

You do not have buffer time; 8 o' clock means the 8 o' clock. That is what it says; class starts equal to 8 o' clock, ok. So, you have to be there at 8. You can you do not have a grace time. It does not say less than or equal to 8.02 or 8.2. So, it is difficult sometimes, not most times to satisfy an equality constraint because, in inequality constraint it is a domain. Here I mean if it is 2-D, it is a line there it is an area. If it is 3-D, it is a surface there it is a volume. So, it is much easier volume means you have more space, right. So, now where will my optima lie, obviously, it should lie on the.

Student: (Refer Time: 35:48).

On the line, and then it touches. So, just keep on expanding and finally touched here. Now they are taking a simple feature here and saying. What they are saying is you are in this particular direction, your function is increasing in this direction gradient of f. And your objective sorry your constraint is going in that direction which is.

Student: (Refer Time: 36:14).

Gradient of h, gradient of f they both are in the.

Student: (Refer Time: 36:21) opposite direction.

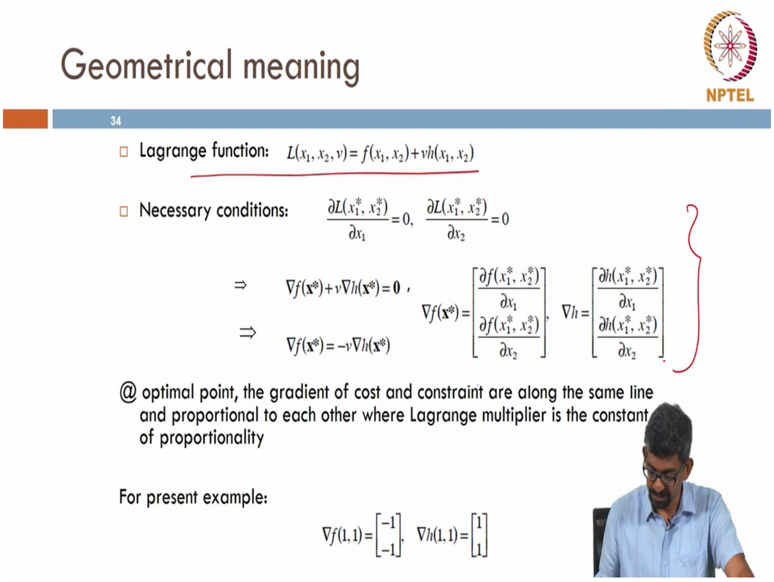
So, that is a feature at the optimal point. So, there you can use it 2 ways. Here this is similar to what Kiran was trying to tell me right like del f equal to 0. You know that you can use it in the reverse fashion. Or while trying to seek an optimal point, you just need to see where my gradient and this guy are in the opposite direction. Or if I have got in an approximate manner if I have got my optimum point I just need to check whether this condition holds good. So, the del f and del h will be in the opposite direction proportional they are not equal. They will be proportional that is what this equation says. The constant of proportionality is your.

Student: Lagrangian.

Lagrangian multiply, ok, this is a way to understand in optimization theory. But Lagrangian multiplier is used in different areas. One of my colleagues teachers parallel manipulator system dynamics, where they use Lagrangian multipliers, ok. Computational geometry they use Lagrangian multiplier. In one sense this is geometry not computational geometry, but then in computational geometry also Lagrangian multiplier. So, in different contexts it is used, in the context of optimization, you can just imagine that they will be not imagined you just imagine visualize it that you are constrained and your objective functions are in opposite directions, but it need not be equal, ok. It is only proportional, ok. So, yeah so that is what it says, dou f because that is your condition, right.

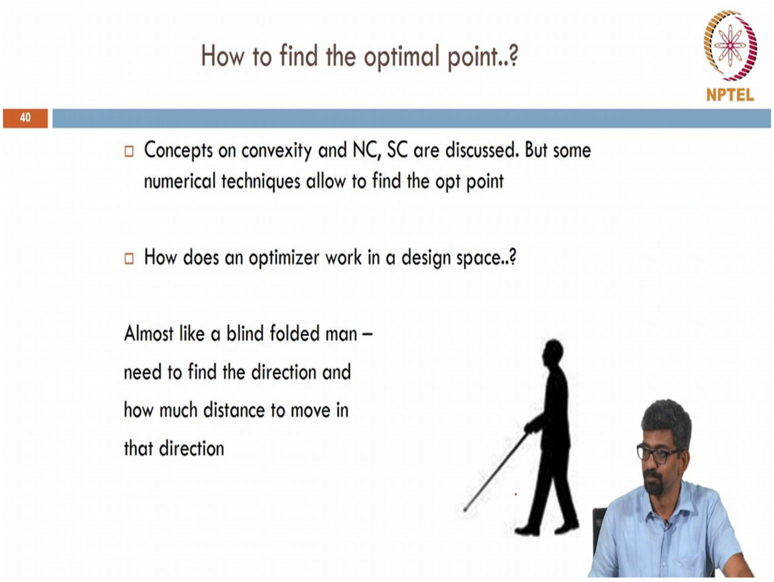
So, if I rewrite your equation like this if I call this guy as f, then if I want to find the minima then I should say dou f by dou x equal to 0 that is what we are trying to do here, ok. And that we are whatever lambda is your Lagrangian multiplier.

(Refer Slide Time: 38:24)



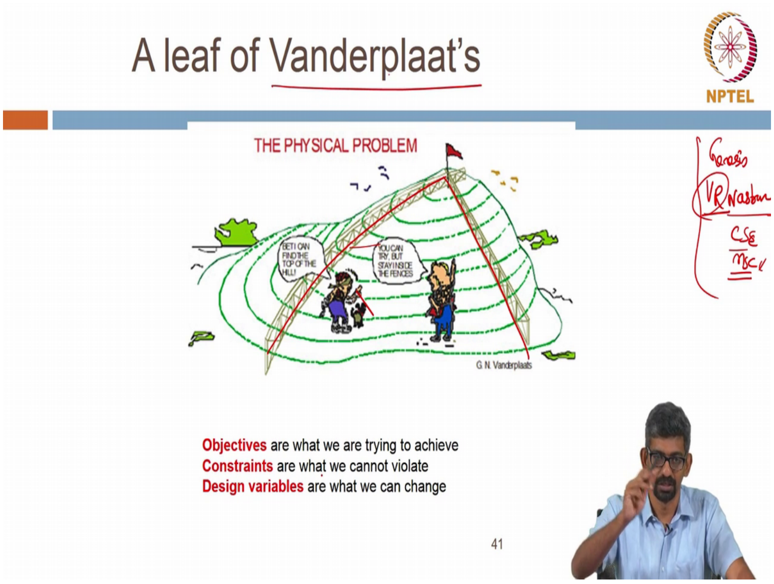
So, you recast your problem using a Lagrangian function. Then you come up with your necessary conditions, yeah.

(Refer Slide Time: 38:31)



This is the blindfold guy's example, ok, yeah.

(Refer Slide Time: 38:35)



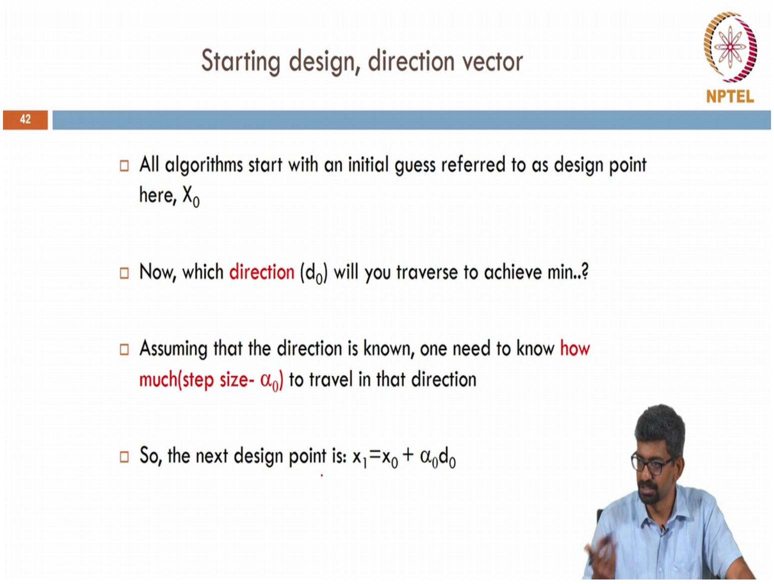
This is something that I will tell you this will wrap up the optimization part; we will probably start a little bit on design of experiments.

So, this is from a famous guy called Vanderplaats he has a book on engineering optimization. He also runs this software company called genesis, ok. So, it is also called VR Nastran VR stands for Vanderplaats, ok. So, you can read this, this guy says bet I can find the top of the hill. You can see that this guy is blindfolded. What this guy says is you can try, but stay inside the fences, ok. So, these are your objective function, these are ISO contours, ok, the green lines here are they your ISO contours. And this guy has a dog, but we can give him a stick in his hand, ok. So, that he can find and then you can go up.

So, what they are saying is this is your constraint that is what he is saying, these are your constraint this is your feasible domain, make sure that you do not touch your fences because they are electrified, you will get a shock. So, you will have to go in that. So, if you touch your fence you will have to come back that is what it says. So, how will he go about? Objectives are what you are trying to achieve, what is that I am trying to achieve? I am trying to go to the top of the hill, constraints are what you cannot violate, what is that you cannot violate? The fences, I cannot go beyond the fences.

The design variables are something that you can change. What is it that I can change? The direction in which I am traveling and the distance I am traveling in that particular direction, that is all I can do.

(Refer Slide Time: 40:13)



So, any search based algorithms, they start with the initial guess what we refer as a design point, here you can call it as x naught. Now you need to make 2 decisions. One is which direction the direction is given by the.

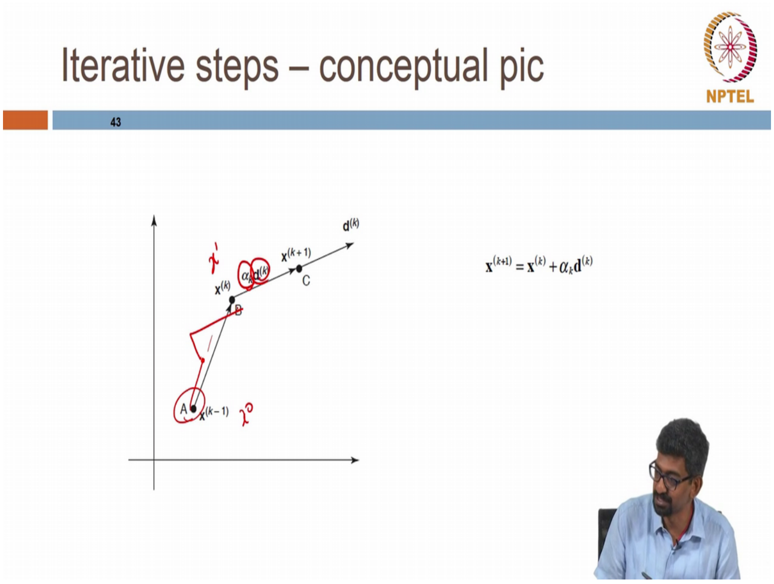
Student: Gradient.

Gradient information so, if you are going into a valley, you will go in the negative gradient and it is called the steepest descent. Now, if you are going to the hill it will be called the steepest.

Student: Ascent.

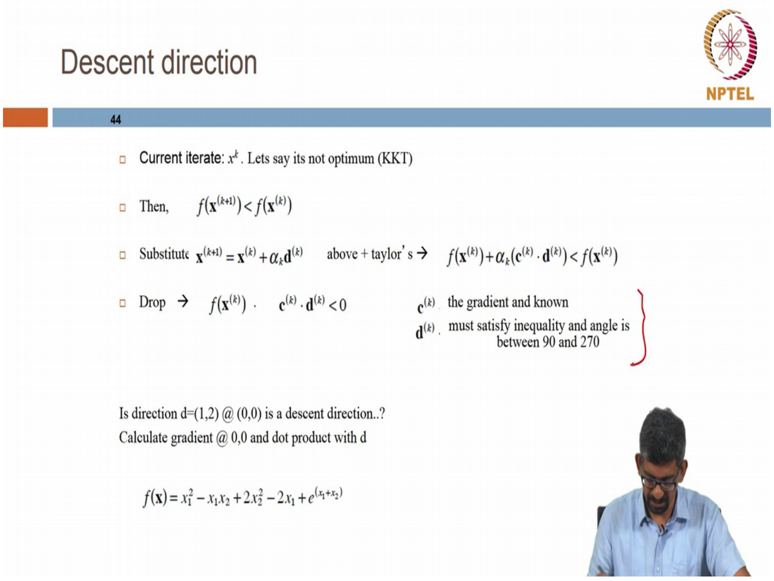
Ascent that is all positive gradient you need to look, you will transversed achieve. Assuming the direction is known, I told you right will you jump one feet or will you jump 10 feet, ok. There is known one need to know how much step size that is your alpha naught to travel in that direction. This alpha naught is kind of solved as a line search problem. It is very convenient because irrespective of n dimension, it will be solved into a one dimensional problem. So, your next design point is simple it will be x naught plus alpha times d naught. I need to go this much feet in that direction, that is all. So, this is achieved by gradient information, and this is given by your step size calculation. It is an it is a minimization problem it.

(Refer Slide Time: 41:23)



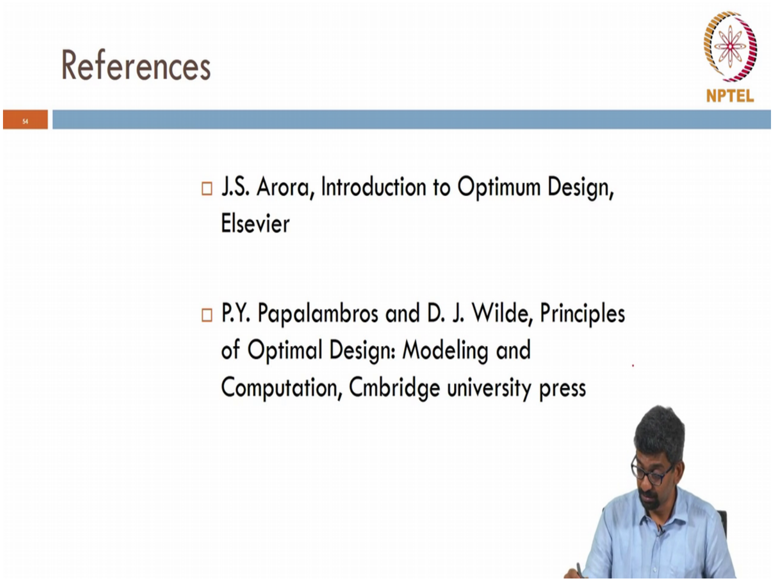
So, this is how it works. It is just pictorial representation. You start from k minus 1 you can call this x naught, ok. Then this will be your x 1. Then this is alpha k d k then you go to k plus 1, you go to k plus 2 you keep continuing like that. So, this alpha is your line search, this d is your direction. So, please understand only from here to here it is in this direction. When you went to that point the direction is something else. So, let us say that you had a function something different here, ok, your line search should have allowed you to go only up to here. Then you might have to go in this direction and then come in this direction, may be, yeah.

(Refer Slide Time: 42:05)



Let us not worry these are interesting derivations, but I am going to skip all this is not required for us. This is about the steepest descent and yeah, yeah.

(Refer Slide Time: 42:23)



So, for this one which took much longer than what I thought, is if you want any references, you want to read a little bit more on these things. Arora is the best book that you can pick up and read. But the second book is a very interesting book. It is by Papalambros he is from university of Michigan, at Ann Arbor.

He gives it in a very nice fashion, ok, for instance Arora's book is very comprehensivem it is this big, I think it is about 650 rupees or 750 rupees cheap additions are available. So, is the case for Papalambro's also. Yeah, Papalambro's gives it to introduces the course in an design sense, ok. So, he does not he does give details about the mathematics, but he introduces everything in a practical sense. So, it is a good book to hold for now and for later also; when you go work or go for a goes talk or something. It might be a good book for you to rely upon. So, it is a good idea to buy that book also not too expensive.

But these two give you a very good introduction, and also there are lot of other books like a Chandrupatla and Belegundu, ok, Ashok Belegundu from Penn state is a very good book. That is also a very good book, and of course, our evergreen.

Student: Kalyan.

Kalyanmoy Deb is there, but that the again that is also he uses a different set of example problem. So, you can buy anything. I suggest these 2, ok, now there are numeric numerous books on introduction to optimization, ok. But these are the standard books that you might want to depend upon if you want to read a little bit more.