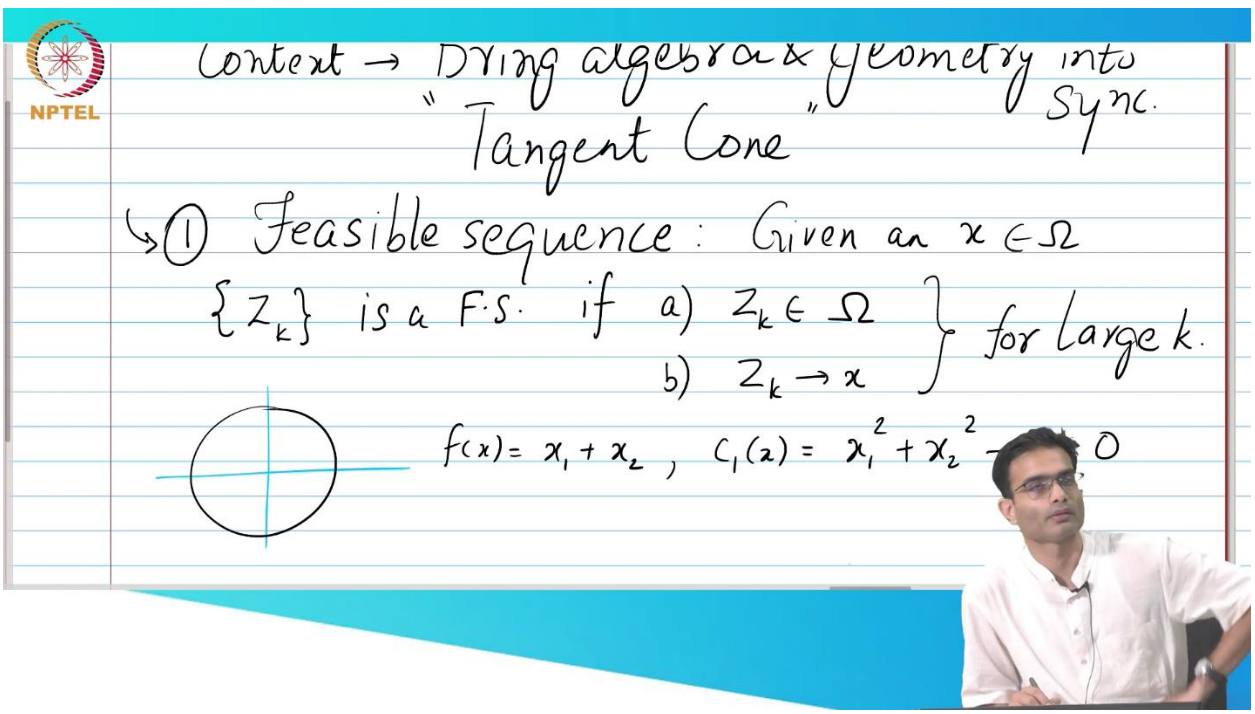
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 10
Lecture - 66

Feasible sequences and tangent cone

So, let us get started. We have been looking at constraint optimization and in particular we are looking at what all did we look at? We looked at an equality constraint to start with, then we looked at an inequality constraint and then we looked at two inequality constraints and based on that intuition we have built some kind of idea of what or how to solve not solve, but how to identify whether or not we have reached a useful point in the algorithm. Then the last thing that we looked at was this kind of a discrepancy between algebra and geometry. right. So, we saw for example that algebra tells us something which gives us a set of what did we call them linearized feasible directions. And depending on how I wrote the specifications for example, in here we took the constraint was specified as square of something and when I took square of something and I tried to calculate the feasible directions I landed up into trouble I found out that any direction was feasible whereas actually that is not the case.

So, we said that we ended this previous section by saying that we need something extra to bring algebra and geometry into sync and that something extra is called constraint qualification ok. So, now let us go a little bit deeper into it. maybe we can just continue it here itself. So, what is the context over here that we It’s very simple bring algebra and geometry into sync.

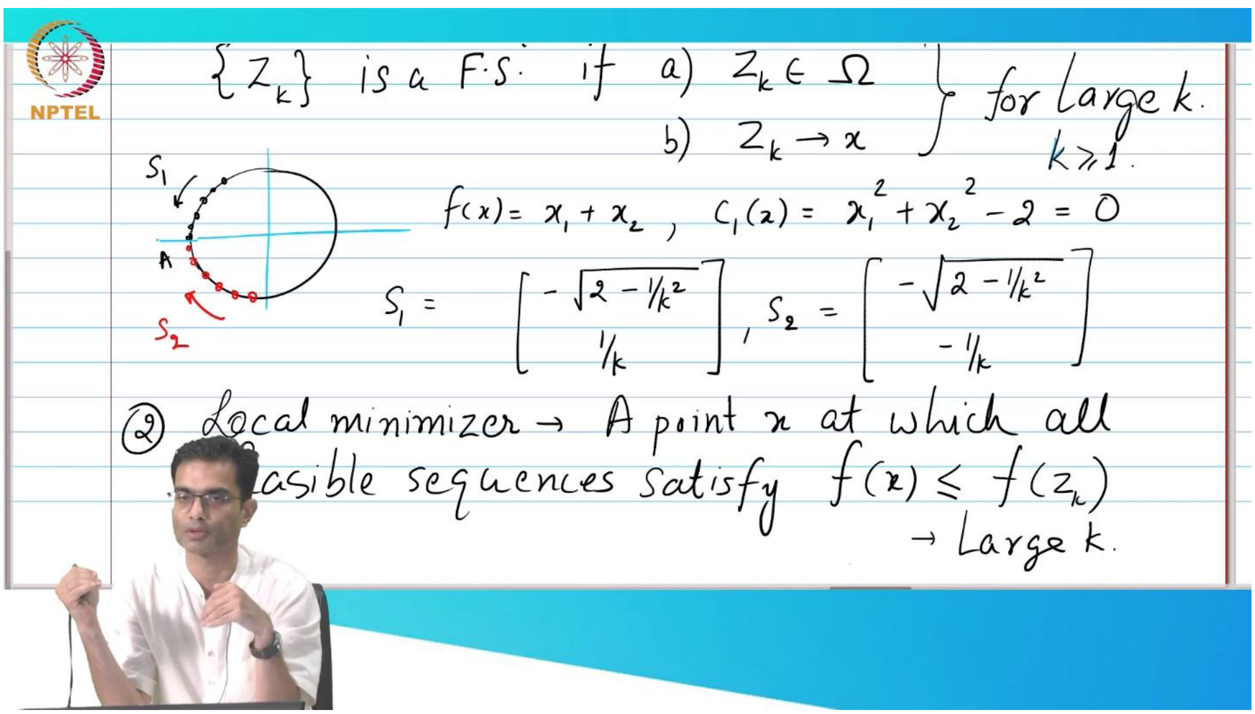


In order to do this now we have spent a lot of time in building up algebra. For example, we had those linearized feasible directions, but so far if you notice we have not spent any effort in bringing any new tools to understand geometry which is sort of going with drawing two dimensional things. So, that is not going to help us when I try to think of a general n dimensional problem right. So, now is the time that we need to build a few tools to help us form up geometry ok. So, in order to do this I have to define a few new objects which will help us to get to the final geometric object that we are wanting to use.

That final geometric object is called a tangent cone that is going to be the geometric object that I am going to use to compare it with the algebraic object. The algebraic object was linearized feasible directions. So, the geometry object is going to be called the tangent cone ok. So, let us just mention this so that you know where we are going ok. But of course, it is not possible to jump into defining a tangent cone in one step.

So, we have to make a few intermediate steps. So, let us have a look at those. So first is what I define as a feasible sequence. The words are all very common sense kind of words. Feasible.

The words themselves are suggestive. Sequence means a sequence of points. Feasible means something to do with feasibility. So what the way I define it is, so I take a point x which belongs to the feasible set to start with and the feasible sequence, remember a sequence I denote with these curly brackets. This is a feasible sequence fs for short.



If two things happen, first is that z_k belongs to the feasible set The second is z_k tends to this point x, but there is a small relaxation we say this for So, this is the idea of a feasible sequence. So, what is it saying? To define a feasible sequence I first need you to tell me what is the point x. So, let us look back at our example that we were discussing yesterday. This is my same example which we have been using objective function is $x_1 + x_2$ and I had a constraint $C_1(x)$ which was equality constraint $x_1^2 + x_2^2 - 2 = 0$. So, what is the feasible set in this case? Boundary, boundary of the circle.

So, let us take this point over here. Now, I want to give you an example of what a feasible sequence is. So, remember I am saying that z_k which is the sequence member should belong to the feasible set, but for large values of k. So, you can assume without any loss of generality that k begins from 1. So, let us just note that over here.

$k \geq 1$ just a matter of numbering ok. So, for large values of k all of these z_k's begin to live on the feasible set Ω and should finally approach this value x ok. So, let us take supposing I think of points like this. So, this is one sequence So, sequence I start from somewhere you know wherever else on the top of the circle and as the sequence index k increases this z_k begins to tend towards my point a right. So, without first defining it geometrically it makes sense what is a feasible sequence, it lives on the feasible set and it is approaching the point A as I increase my index right.

Another sequence could be for example, if I start from let us say from somewhere here and then I start as I increase my k this could be another feasible sequence. You can there are there are in fact infinite possible sequences that I can define that live on Ω and approach my point A ok. So, this was the geometric picture let us try to define it ok. So, I am going to call this as S_1, I am going to call this give you two examples of a sequence ok.

So, S_1. So, I need to define a x coordinate, I need to define a y coordinate ok. How would I construct this? the clue is that these points of the sequence must live on the circle. So, they should satisfy $x_1^2 + x_2^2 = 2$, right. So, I could define S_1 like this. Does that work? Is $x_1^2 + x_2^2 = 2$? it works right.

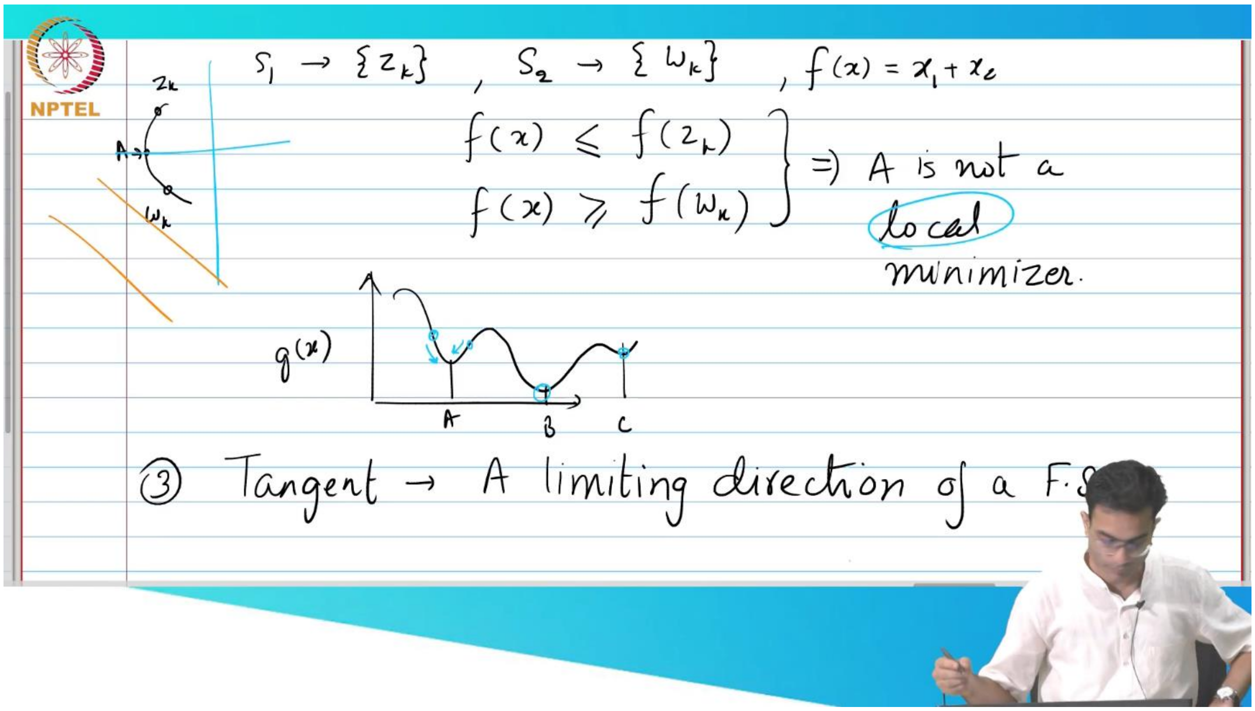
So, it is on the circle and for example, if I take k equal to 1 where do I start from $(-1,1)$ where is that? second quadrant right, second quadrant. So, it is starting from in between over here and then as I increase k where is it going? It is following the arrow and falling down into a right. When k tends to infinity what does this coordinate become? $\left(-\sqrt{2}, 0\right)$ which is my point a ok. Similarly, to get S_2 there is not much different that I would do, what would I do? Just flip that $1/k$ to $-1/k$.

Some point right the feasible sequence can be defined at any point on the feasible set including the solution point. So, this is a feasible, this is example of several feasible sequences. Now, related to the question you asked, we can also using the concept of feasible sequences, we can have a different alternate definition of what is a minimizer, or what we are calling a stationary point. So, let us call that as point number 2.

So, local minimizer. So, you have got the idea of a feasible sequence. Now, using the idea of feasible sequences if I ask you to come up with a definition of a minimizer to your problem, can you think of a clever way? What would distinguish the minimizer from all other points in the feasible set and you do not have any calculus to work with you only have feasible sequences which you can construct at will. What would distinguish the minimizer from any other point in the set right that is exactly right. If I find a point and at that point I find that no matter what feasible sequence I construct the value of the function at that point is always less than or equal to the feasible sequence. That is another signature of right.

So, no matter in which direction and what way I approach this point the function value at that point is always less than what is there on the feasible sequence that is a perfect signature for a feasible point. So, you see how we are using geometry to arrive at minimizer now right. So, it is a point at which all feasible sequences, move this down a bit satisfy that $f(x) \leq f(z_k)$ and again we are going to make this statement for large k. ok. So, based on these two simple concepts which you have got, we have got a feasible sequence and we have got a local minimizer.

Can we conclude that the point A is not a local minimizer? So, let us take sequence 1 S_1 ok. So, and everyone knows the figure. So, sequence S_1 has let us say z_k and let us say sequence S_2 let us call the points as w_k. This was our oops. So, this is my z_k, this is my w_k, this is my point A over here.



Right. What was to check increasing or decreasing the objective function I need to take into account is $x_1 + x_2$, $x_1 + x_2$ we already said has what are the contours of constant f I think right. So, they are like this. So, $f(z_k)$ let us fill this $f(x)$ $f(z_k)$ $f(x)$ and $f(w_k)$. So, the first one what is the which way is the inequality $f(x)$ is always greater than or less than or equal to $f(z_k)$? It is less than let us say equal to and the other one it is always greater than or equal to. So, this together implies what? A is not a local minimizer. Yeah, that is quite straightforward and it is a non-calculus way of arriving at a local minimizer.

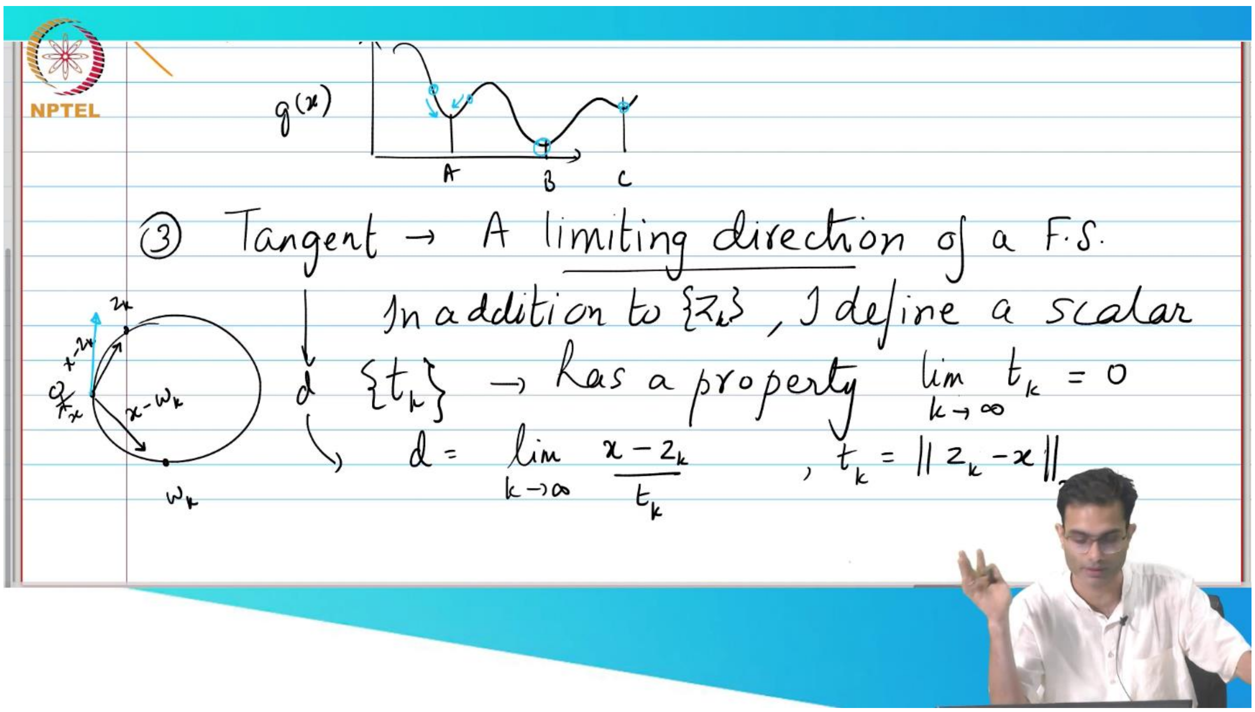
Remember this signature will only work for a local minima. Will it give us a global minima also? For convex functions there is only one minima. So, there is no question, but in general supposing I had my let us take another example this is my $g(x)$ which looks like this right. So, supposing I take two points A, B, C and I construct a feasible sequence about point A, it is clear what we are doing. So, let us say I start from here or I start from here and the feasible sequence should tend to A that is the meaning of a feasible sequence.

So, this will satisfy the condition that A is a local minimizer. But does it tell me anything about what is happening at B? Tells me absolutely nothing. So that is why this is a signature only for local minimizers. There is absolutely no way that you can get non-local information over here.

And that is the case with almost, not almost, all the algorithms that we are talking about in this course, they are all local optimization tools. You need some other tricks to figure out what is happening in other valleys.

So, we have got feasible sequence, we have got a minimizer, let us go further. The third thing that I need to define having used a, the idea for feasible sequence is a tangent. We have all looked at we all know what tangents have we have been doing drawing tangents since class 9. Let us formalize this in a way derived from feasible sequences ok. So, let us I have our diagram is above over here ok.

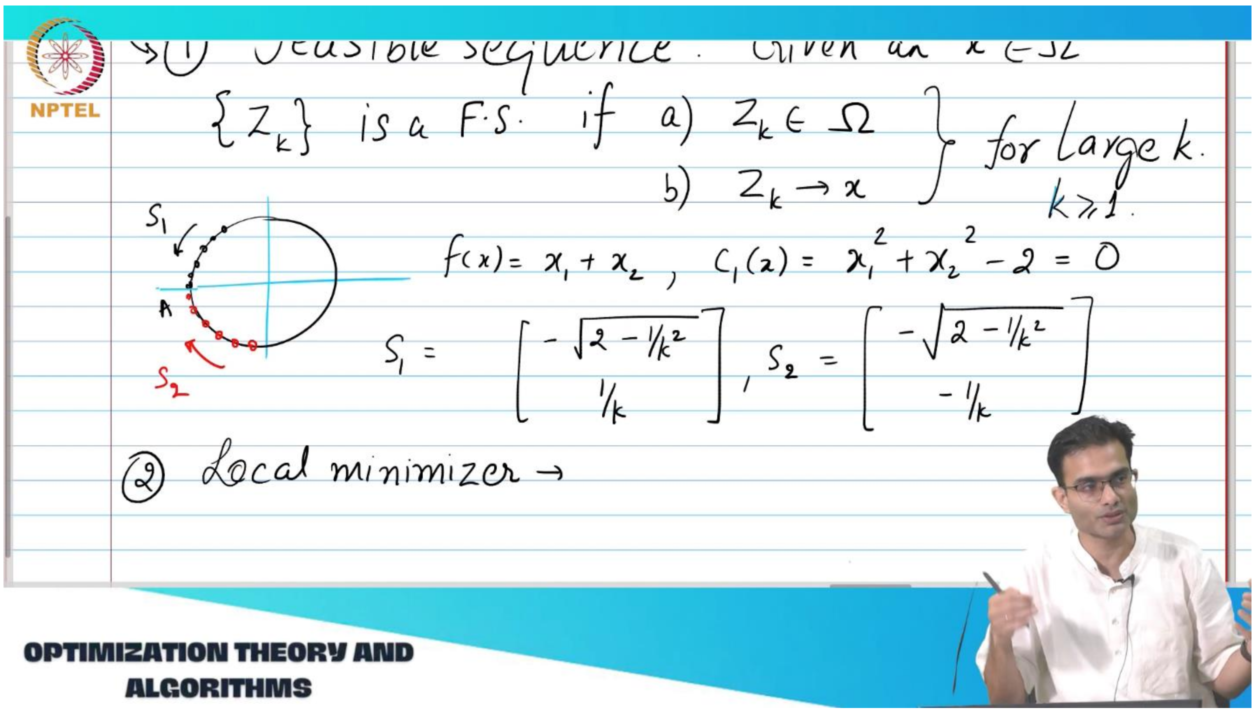
So, a tangent is simply a limiting direction of a feasible sequence. So, keep this diagram in mind over here this is my z_k this is my A. So, as I am beginning to suppose I. So, this point x over here I need to draw this again let us draw this again over this is my point z_k, this is my point w_k. So, what is $x - z_k$, which way does it look? I draw a line from x to z_k.



So, this is my $x - z_k$. Similarly, so let us write this, this is $x - z_k$, then I can also construct a vector like this, this is $x - w_k$ fine. Now, what am I defining the tangent as? A limiting direction of a feasible sequence. Now, imagine what will happen? Supposing I am standing over here, I am standing at the point x and I am watching the feasible sequence come and collide into me. Can I think of a limiting direction? As the points get closer and closer, so right now you see that $x - z_k$ has an angle like this, but as I start approaching z_k coming closer and closer to x which way do you think this will start approaching? A tangent, right? As long as I do not exactly become equal to x I am some small distance away, this limiting direction that is why this limiting direction over here this direction is going to become tangent to the circle you can see that right. So, that is basically the that is basically how I come up with the definition of a tangent, ok.

So, now to mathematically define a tangent I have got the difference of two vectors. z_k and x, but there is a problem with that right if z_k begins to approach x what happens to the magnitude of this vector? It becomes smaller and smaller, finally, in the limit what will happen to the magnitude of this vector? It becomes 0. So, how do I rescue? Divide by the distance for example, right? So, that is why I need a feasible sequence and I also need a sequence of scalars which will divide this, right. So, that is the common sense way of arriving at it.

So, in addition to z_k which I had I define some length. So, a scalar let us call it t_k ok which has the property that what do you think is the most critical property of this t_k which we said is going to be like the length that as k tends to a very large number what should t_k tend towards? 0, right? So, it should go as small and at the same rate at what $z - x_k$ is happening, right? That is, I will just come to that. So, t_k let us write that over here ok.



So, give me a minute over there. So, the tangent now becomes I am going to call it d. So, d becomes simply now I can take the limit nicely. So, k tends to infinity $\frac{x-z_k}{t_k}$ ok. So, this comes this as you can see that if t_k has this property then this $\frac{x-z_k}{t_k}$ is going to give me a well-defined direction, right? It is going to give me. So, if I want to draw it is going to give me this vector in blue this is going to be using z_k I get this vector.

Now, question is what is a good example or a candidate for t_k? So, the first idea that comes to our mind is the two-norm $\| x - z_k \|$. This is good enough for our purposes ok, but as it turns out in the literature it is not necessary. Supposing I had defined this as the one-norm or some other Euclidean, some other p-norm of it, it would still work. Why? The only property that it needs to satisfy is this that for large k it tends to 0, ok.

So, this is actually left to us. So, you can choose it in a way that makes the math of the problem as straightforward and convenient as possible ok. And this is usually you do not have to stray away from this choice, this works. So, as this length shortens and shortens you can see this gives me a well-defined direction, ok. So, we have got a tangent. If I were to choose w_k as my sequence will I get what will I get? So, $x - w_k$ is kind of pointing bottom right, right?

Now, as w_k begins to approach this way which way is this vector going to look? Downwards, right? So, what is that telling us? It is fine, it is telling us that associated with every feasible sequence I am going to come up with a particular tangent. There is no reason why we expected everything to give me the same tangent, right? You can see both the blue arrow and the red arrow are legitimate tangent directions to the circle at point A, that is fine, right.