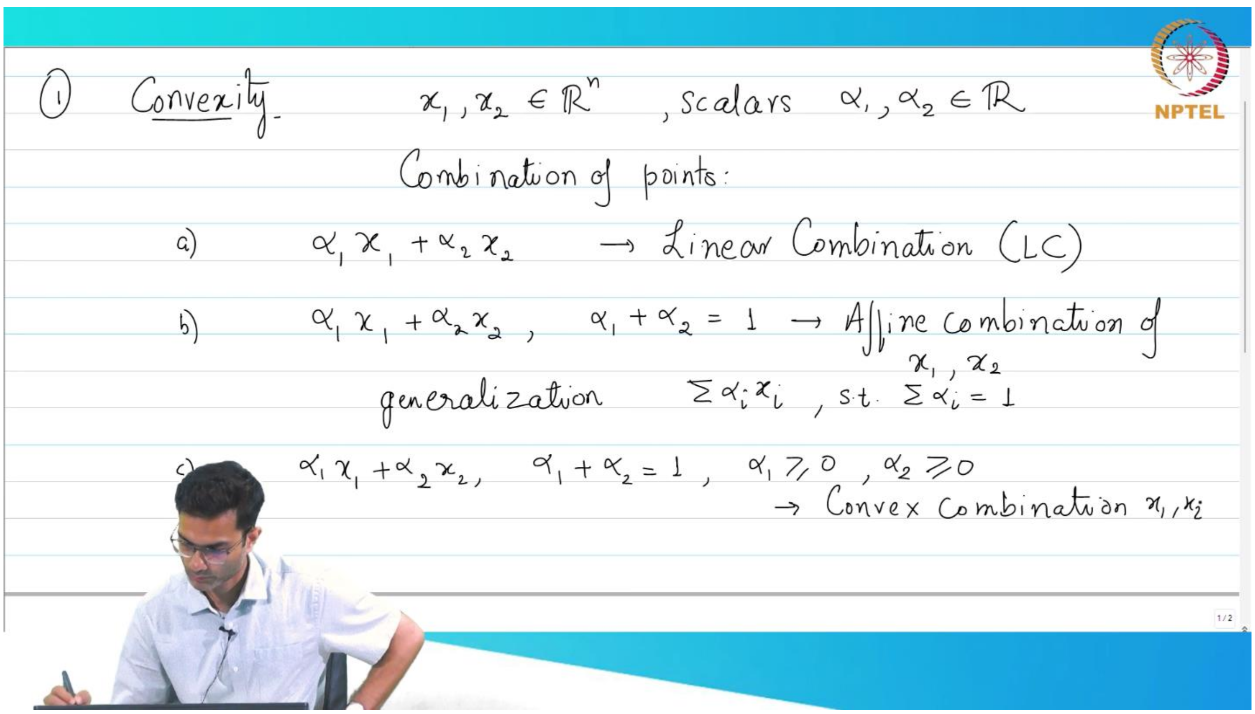
wCourse Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 02
Lecture - 08

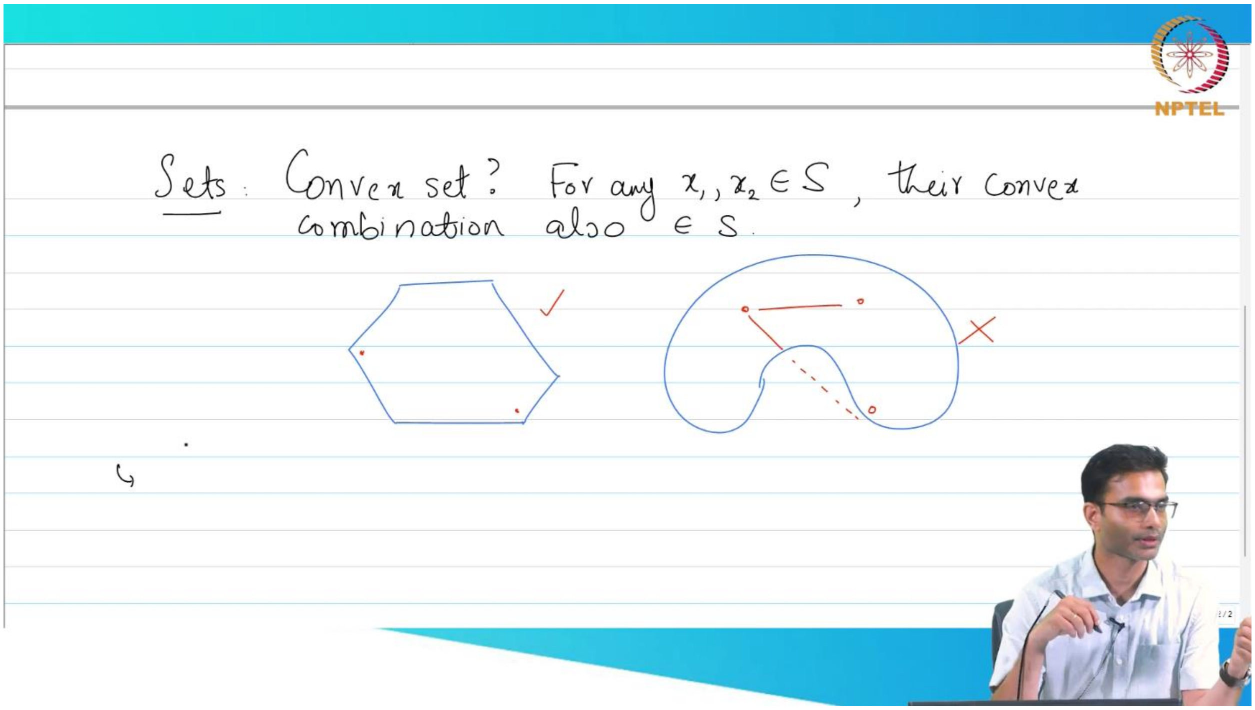
Summary of background material - Analysis III

Let us start with convergence and rate of convergence, which is what we spoke about at the end of the lecture last time. What and how do we choose allowable rates of convergence? We spoke about how many different types of convergence rates there are: linear, quadratic, and something in between called super-linear. In fact, that quadratic, if you remember, had a square in the denominator, which is why we called it quadratic. You could replace that by any other number; you could make it 3, 4, and it will be a fourth-order rate of convergence. So, the question is: how do we choose? The bad news is we cannot choose. We come up with an algorithm, and do its analysis; it will turn out to be this or that.



It is typically quite difficult to plan to have a certain rate of convergence; you get what you get. At least the first time when you are doing this, maybe if you get much more experience, you can design for a certain rate of convergence, but typically, no. Is it okay to attend this class without any proper concepts of norm, singular value decomposition, linear algebra? Well, as you saw in last time's class, it may be difficult, but you can take your chances; you can drop it or not. It will be useful if the handwritten notes could be shared.

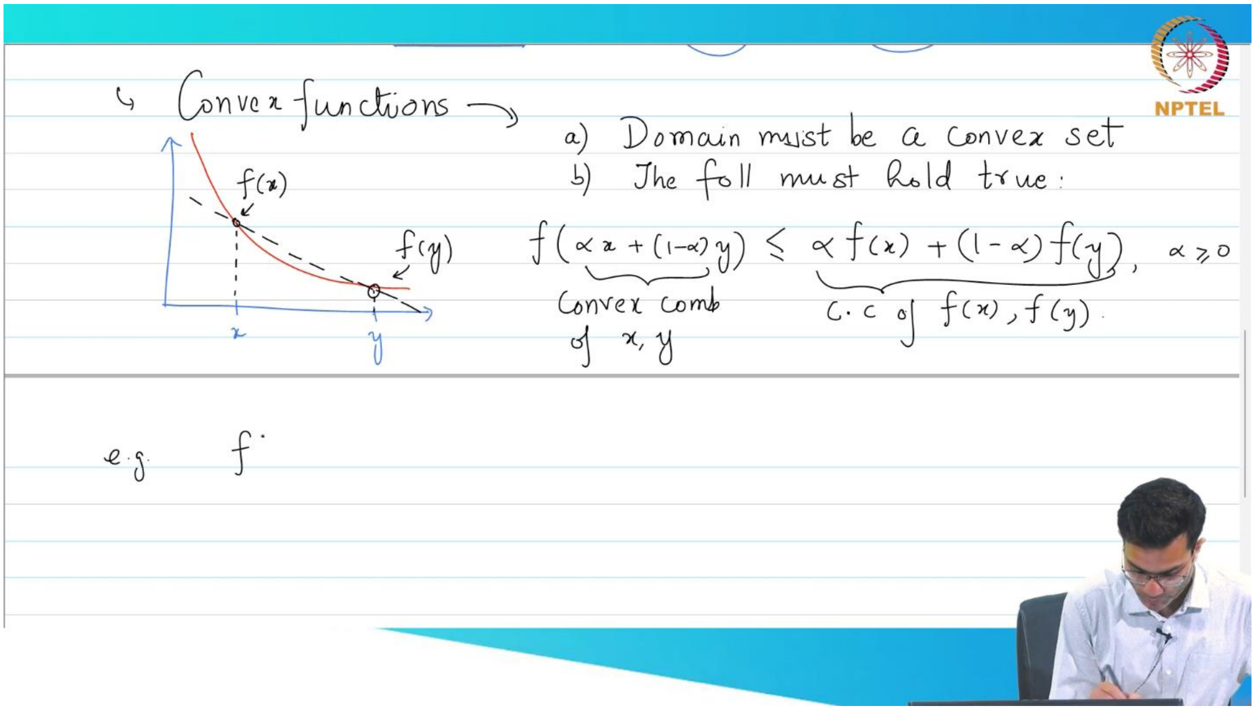
I am not too keen on that. I mean, everything that I am writing is already there in notes form on the class website. Has anyone here seen the notes? A fair number of you have seen them. So, everything that I am talking about is covered there. If you want, the course, I mean, the lecture content is there on the website; go and read it.



It is in fully written English sentences. So, you can understand it better rather than this hodgepodge which I am writing over here. By the way, is someone finding it difficult to follow my handwriting? No one will say yes, but if you do, then write it in the feedback chat at the end of the day. Sir, you should have been a doctor instead, or something like that. The condition number still seems abstract; it will be helpful if it is covered in the tutorial.

So, we will spend a lot of time talking about the condition number later on in the course. So, we will get to it. This question, I think, got asked at the end of the class regarding the example we took: $x_k = \frac{1}{k!}$ which we took in the end. Why is it not linear? Because I can find an r such that $\frac{x_{k+1}}{x_k} < r$, and we said that if it satisfies a higher rate of convergence, we just declare it as super-linear even though the question is correct. Then there are some questions about the definition of the p norm.

So, why is the ∞ norm the maximum value? We derive this in a linear algebra course; it is a simple matter of playing around with the definition of the p norm. Remember there is a $\frac{1}{p}$ at the end of the expression, you can take that to the left-hand side and then take the limit as $p \to \infty$; only one term survives, which is the maximum term. So, that is why you are left with it. Similarly, there is a question about why the 2 norm of a matrix equals the singular value. We would derive this in a linear algebra course, not here; the derivation is not that difficult. So, I will not spend time over here, but if you still find it difficult, I can post a link on Google Classroom for it.

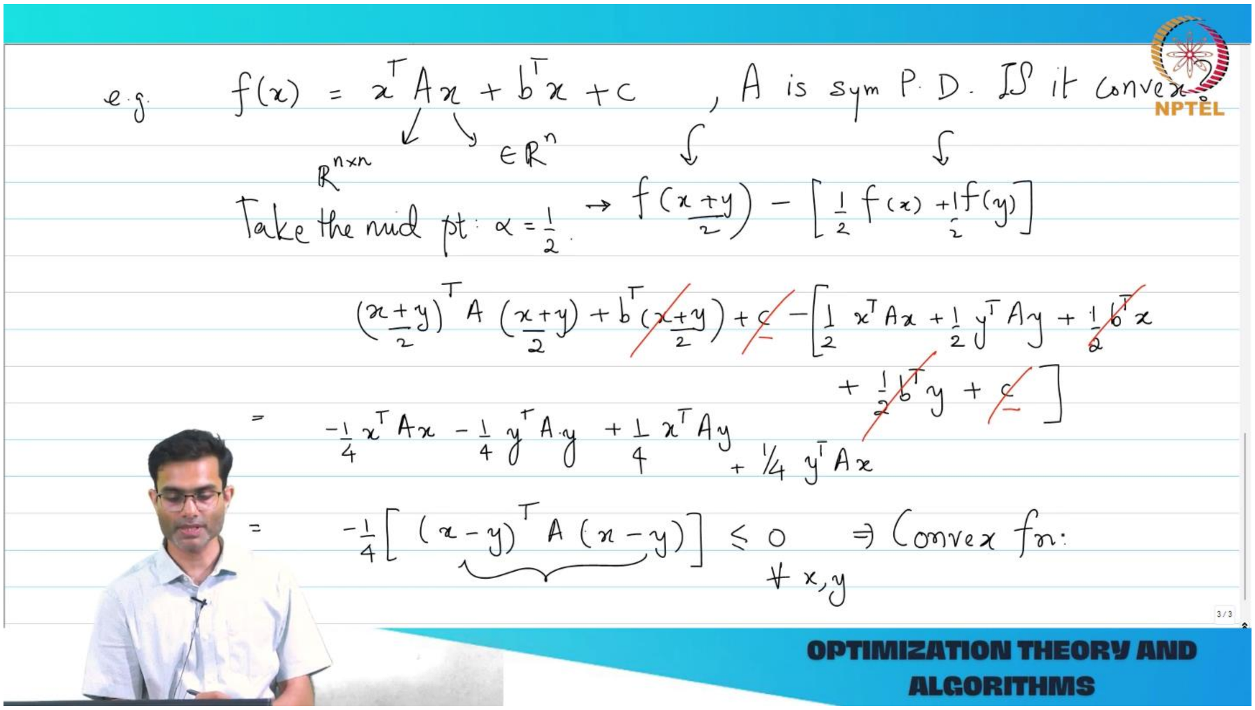


After speaking about rates of convergence, which is kind of a concept of linear algebra, or sorry, analysis, when we are talking about sequences and what rate they converge to, the next thing that we want to talk about is a very important concept of convexity. Now, before we talk about convexity, I want to talk about something a little bit broader than convexity, which is simply how do I combine several points in space? Sounds like a very simple idea. So, let us start with as simple as possible; let us take two points. So, I am going to take x_1 and x_2 as two points, and these live in an n-dimensional space.

Furthermore, I am going to take a couple of scalars which will hang out with these x_1, x_2; I am going to call them α_1 and α_2; they also live in the real space. Okay, now, the question is how can I combine these two points? In general, how do I combine n points? Okay, so if I just write this as $\alpha_1 x_1 + \alpha_2 x_2$, what would you call this? You are giving me the geometry of the combination, which is a line, but in plain English, what will I call this? A linear combination of the points. A linear combination of the points, exactly. Also abbreviated as LC.

Okay, when I made a linear combination of these two points, did I have to make any specifications on the alphas? No. No, alphas can be anything in the real world, right? So, this is the most general combination of two points. Now, let us talk about one small restriction. So, I am going to say $\alpha_1 x_1 + \alpha_2 x_2$, and I am going to put one constraint: $\alpha_1 + \alpha_2 = 1$.

So, there is now a constraint on α. Anyone wants to guess what this is called? A line joining... again, I am getting geometry; is there a name that comes to mind for this? Linear combination; this is still a linear combination, but there is something a little more special about it. Affine. Correct. So, this is called an affine combination of x_1, x_2.



So, how would I generalize this to n points or m points? The generalization would simply be $\sum_i \alpha_i x_i$ such that $\sum_i \alpha_i = 1$, however many points there are. Now, let us look at one more restriction over here: $\alpha_1 x_1 + \alpha_2 x_2$. I am going to retain this affine constraint, and I am going to add one more constraint which is $\alpha_1 \geq 0, \alpha_2 \geq 0$. Any guesses what this is called? The hint is the first word on the page. This is called a convex combination of x_1 and x_2.

So, if I were to draw this, this is my x_1; this is my x_2, and this is a line running through these points. Okay, so now it is easy to visualize this when I talk about an affine combination; where can the points be? So, an affine combination means their sum should be 1, but α_1 and α_2 can be negative also, right? So, what do you think is the locus of points over here of the affine combination of x_1? Any point or what? Any quadrant. What is the locus of points that is traced by going through different values of α in an affine combination? An ellipse. Any point? What do you think if I tell you consider the red line? So, $\alpha_1 + \alpha_2$ has to be 1. So, let us take for example $\alpha_1 = 0$ and $\alpha_2 = 1$ and vice versa, right?

So, the two points x_1, x_2 are obviously on this, right? If I take $\alpha_1 = \frac{1}{2}$ and $\alpha_2 = \frac{1}{2}$, where am I? Middle, right? If I take $\alpha_1 = 2$, is it possible? Yes, it is possible for an affine combination. Where will I be? So, this... do you agree that all of these points will still be on the line, right? They will still be on the line. So, it looks like I have drawn this line to only be in the first quadrant, but it need not be; I mean, I can extend it like this, right?

So, an affine combination basically traces out the entire line connecting x_1 and x_2, whereas the convex combination, now you can appreciate, is the line segment between x_1 and x_2; that is the restriction, right? So, this is just convexity of points. Let us talk about now sets. So, how do you think I would define a convex set? Everything is building on the previous step. So, I have told you what a convex combination of two points is.

So, what do you think is a convex set? If you pick any two points in that set, if the line segment connecting those two points is entirely within the set, then that set is convex. Correct. So, a set C is convex if and only if for all $x_1, x_2 \in C$, and for all $\alpha \in [0,1]$, it holds that $\alpha x_1 + (1-\alpha)x_2 \in C$. So, that is the mathematical definition of a convex set.

Now, let us go back and see what it means in terms of 2D and 3D objects; you will see that it is easier to visualize in two dimensions. So, a few examples of convex sets: the first one is a line. I am going to take the entire line. Do you agree that this is a convex set? Yes. So, any point you take on the line, will the entire line segment lie within it? Yes. So, the line is a convex set.

The next one is a polygon; so, a triangle or a square is a convex set, right? Now, if you take any points inside this, say x_1 and x_2, and you take the line segment connecting x_1 and x_2, it should also lie within that set. So, polygons are a classic example of convex sets.

What about this? A circle. Any points you take on the circle; do you think that the line segment connecting those points is still in that circle? Yes, it is. However, this one: a half circle. Is it a convex set? No, it is not a convex set because if I take x_1 on one side of the half circle and x_2 on the other side, the line connecting them will go outside of the half circle.

So, you see that a simple idea like a convex set means that you should be able to draw a line connecting two points, and that line should be contained within the set; that is essentially what convexity means.