Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 01
Lecture - 07

Summary of background material - Analysis II

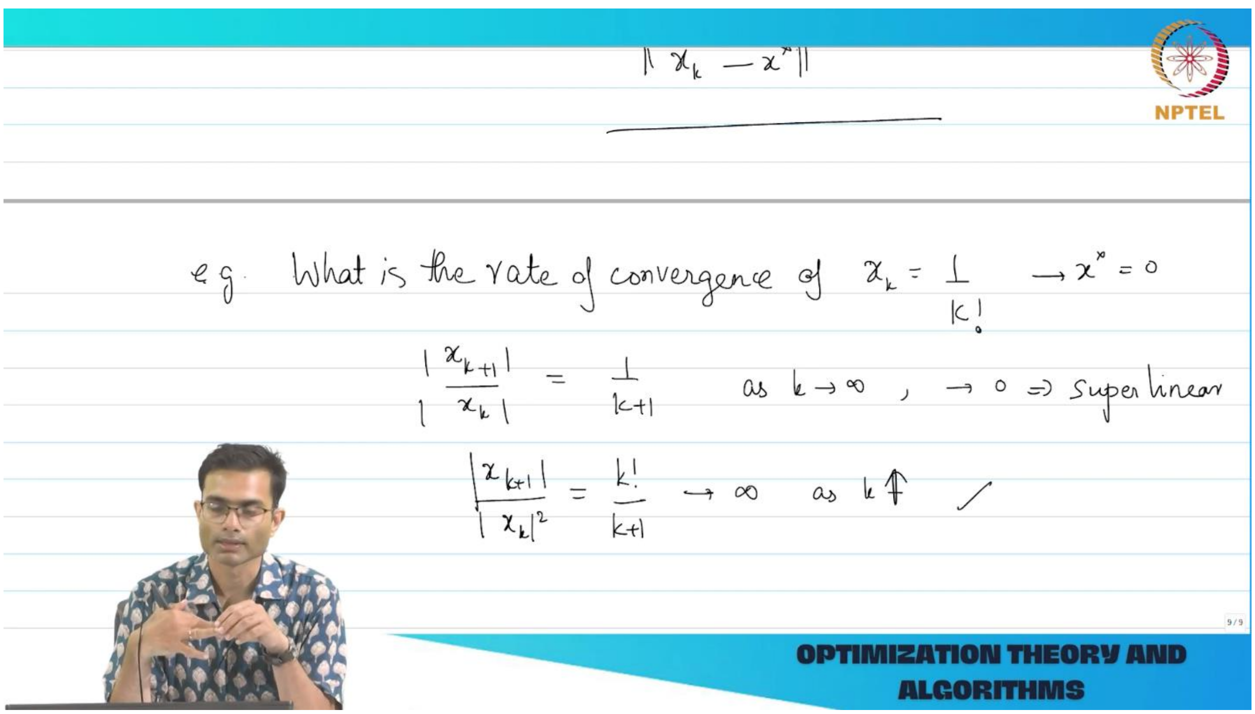
So, again this will sound a little abstract. We will just take one numerical example—well, not numerical, but one example to narrow down and understand this better, okay? So, for example: what is the rate of convergence? Let us work this out. Consider the sequence:

$$x_k = \frac{1}{k!}$$

So, I am giving you a sequence. The first term is 1, then $\frac{1}{2}$, and then $\frac{1}{6}$, and so on, where $k!$ denotes the factorial. So, what is x^* for this sequence?

$$x^* = 0$$

Very intuitively, it is clear that $x^* = 0$, right?



Now, in this case, it is a scalar sequence, not a vector sequence, so we do not need the norm. So, if we take, for example,

$$\frac{x_{k+1}}{x_k}$$

what do we get? We get:

$$\frac{1}{k+1}$$

Now, what can we say about $\frac{1}{k+1}$? As $k \to \infty$, it tends to 0, right? So, this is superlinear, right? Okay. Can it also be quadratic? I mean, we found that it satisfies superlinear convergence, but is it also quadratic? Let us just check. So, we take:

$$\frac{x_{k+1}}{x_k^2}$$

What do we get?

$$\frac{k!}{(k+1)!}$$

What is this limit as $k \to \infty$? Infinity. It is not going to a positive number, nor is it being upper-bounded by some constant number, right? So, this is actually infinity as k increases, right? Therefore, this sequence does not have a quadratic rate of convergence; it has a superlinear rate of convergence. You will find that only one type of convergence holds at a time—it is not going to be both.

Now, we have spent time talking about linear algebra and analysis. The next thing I want to talk about is convexity, which is an important part of analysis. Let us see… How are we doing on time? I think we are out of time, so I will pause here.

What about Linear Convergence?

What about linear convergence? When I take the ratio:

$$\frac{1}{k+1}$$

Is it linear or not? For large k, notice the subtlety. Is there a restriction on r? It is between 0 and 1, yes, but is it a closed interval $(0,1)$ or an open interval? It is an open interval, meaning 0 is not included. So, as k becomes arbitrarily large, r tends to 0. It slips out of linear and becomes superlinear. For linear convergence, r needs to be strictly positive, and $\frac{1}{k!}$ is not. This is why it is crucial that it is an open interval $(0,1)$.

$$r < 0.5$$

That is one way to look at it, but the comprehensive way is that if a sequence satisfies the higher-order definition, we declare it to have that convergence. If it is quadratic, we say it has quadratic convergence.

Key Concepts Reviewed

We have reviewed linear algebra topics, including vector norms, matrix norms, fundamental subspaces, eigenvalue decomposition, and singular value decomposition. We also discussed the convergence of sequences, definitions of upper bounds, and rate of convergence. All of this is in the class notes on the website, so please do have a look.

Practical Example

In one case, I gave a student a number that was 3 times the age of the universe. This was a combinatorial problem. The variable of optimization had length 100, and each component had 2 or 3 options (0, 1, 2). Thus, there were 3^{100} possibilities. Using an Intel Core i9 processor, where one calculation takes 1 microsecond, you can estimate the total time by multiplying it all out. Many times, back-of-the-envelope calculations can help, though this may not always be possible.

Schur Decomposition

For those interested, the proof for the SVD decomposition can be found on the course website. It starts with forming $A^T A$, AA^T, and factorizing them. The Schur decomposition, an important theorem, is used to convert the eigenvalue decomposition (EVD) into a singular value decomposition (SVD).

Condition Numbers

The condition number tells you the amplification factor. If b has a small amount of noise, the amplification will result in a reasonable amount of noise, but if b has more noise, the amplification factor remains the same. This depends only on A, and the condition number captures these effects. Numerical analysts are very particular about condition numbers.