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

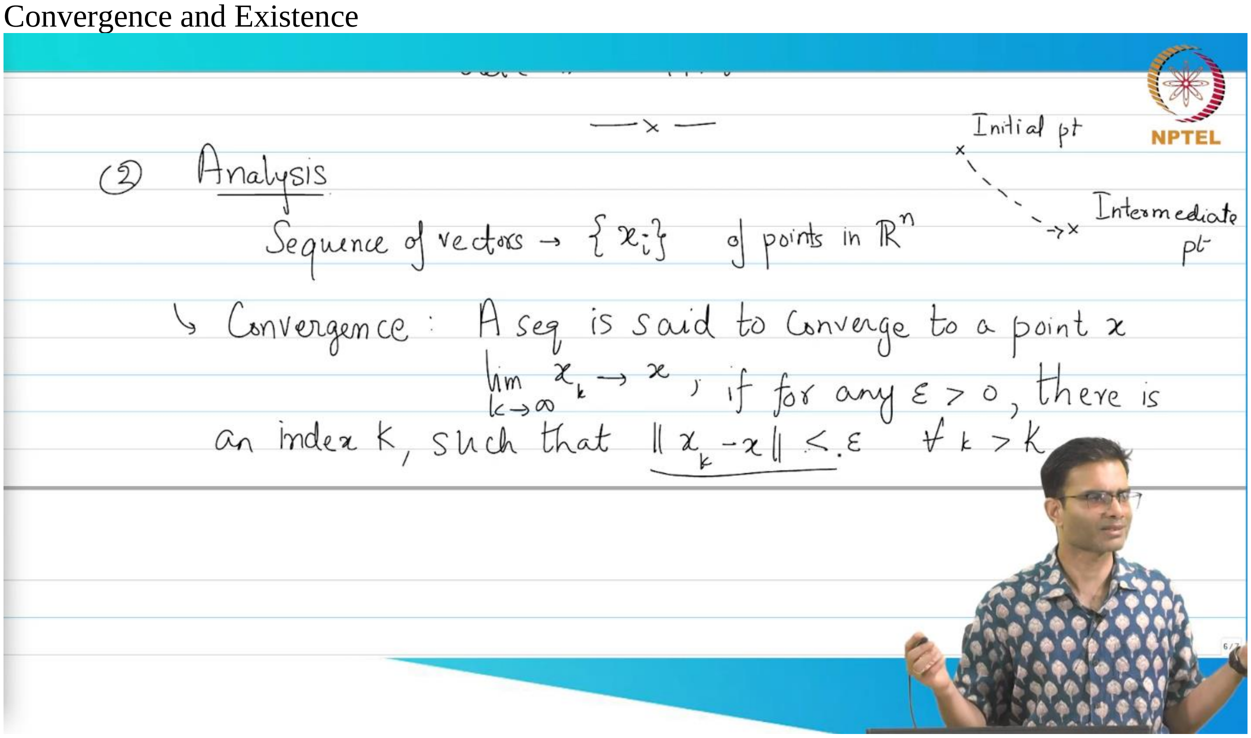
Lecture - 06

Summary of Background Material - Analysis I

The next recap focuses on analysis, which is essential in optimization algorithms due to the need for understanding convergence. In any optimization algorithm, we start from an initial point and aim to reach an optimal solution. Before coding and implementing the algorithm, it's crucial to have confidence that it will work, which we quantify using analysis tools that assess convergence. The key questions are whether the algorithm converges and, if so, how fast.

Many algorithms, even in research and industry, lack a formal proof of convergence, making this property significant when it can be established. However, even if an algorithm converges, a slow rate of convergence can render it impractical. Knowing the convergence rate beforehand prevents wasting time on simulations that might never complete within a reasonable timeframe. Therefore, understanding both convergence and the rate of convergence is crucial.

Convergence and Existence

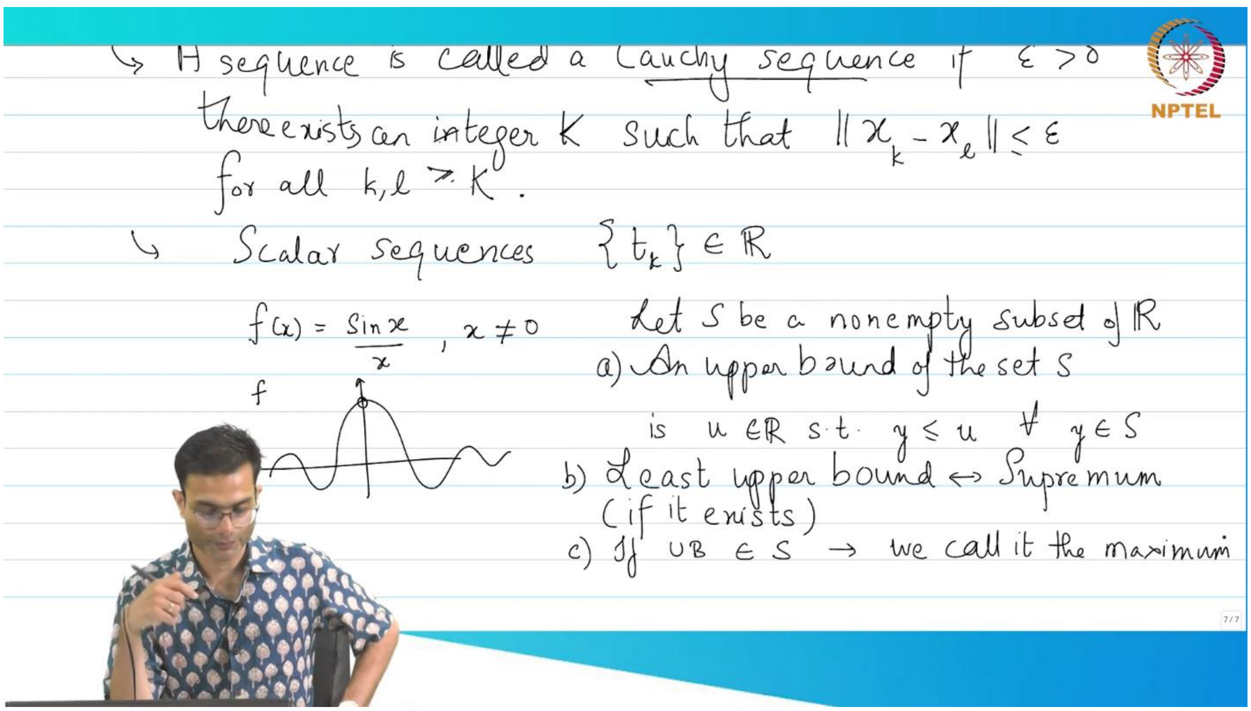


Convergence is closely related to the existence of a solution. If a sequence converges, it suggests that there is a point where the sequence stabilizes. This is not always easy to prove, but where possible, it's studied. A sequence is often denoted as a set of vectors, which is important since most engineering problems exist in n-dimensional spaces, not just scalar sequences.

The convergence of a sequence is intuitively defined as the sequence approaching a specific point as the sequence number increases. Formally, a sequence

$$\{x_k\}$$

converges to a point x if, for any small positive number ϵ, there exists an index K such that the distance between x_k and x is less than ϵ for all $k \geq K$. This distance is measured using norms, typically the 2-norm unless otherwise specified.

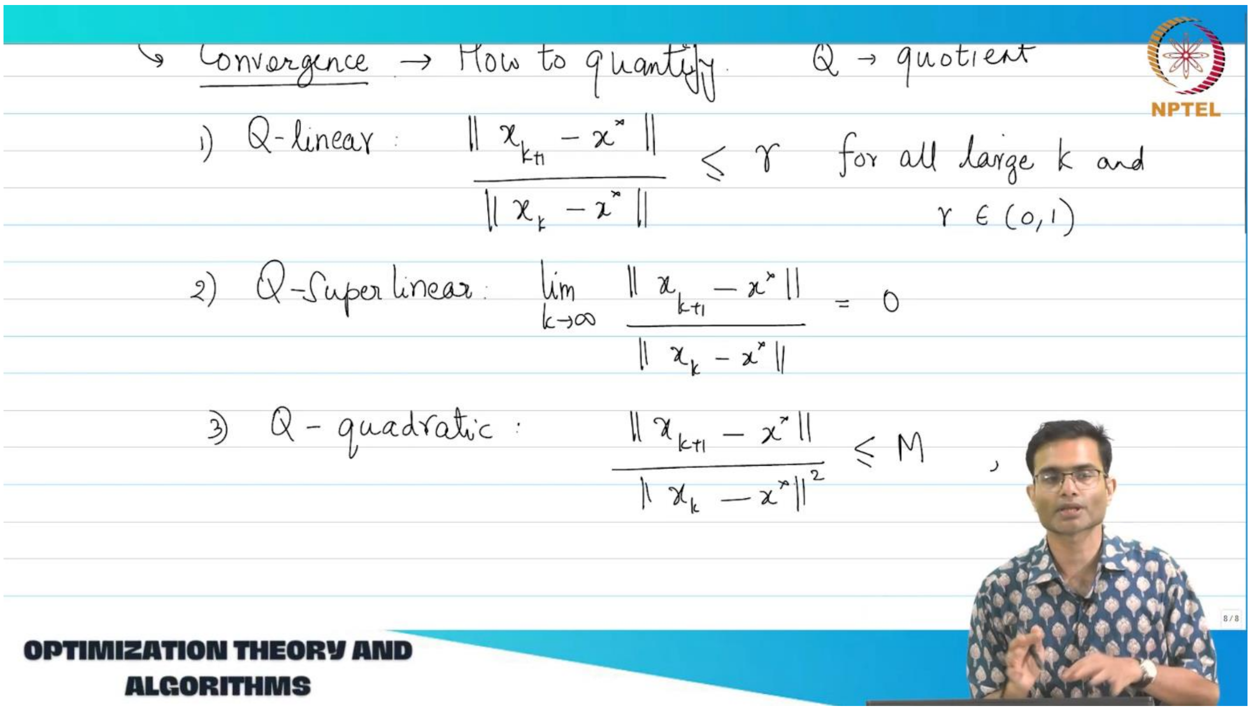


Cauchy Sequences

A Cauchy sequence is a special type of sequence where, beyond a certain point, the distance between any two points in the sequence becomes arbitrarily small. While not all sequences that converge are Cauchy, this concept is useful in understanding algorithm proofs and their convergence properties.

Scalar Sequences and Bounds

Moving to scalar sequences, one key concept is boundedness. A set is upper-bounded if there exists a number such that all elements of the set are less than or equal to this number. Among all upper bounds, the least upper bound, or supremum, is the smallest one. If this upper bound is part of the set itself, it is called the maximum. Similarly, on the lower side, we have the infimum and the minimum.



Rate of Convergence

The rate of convergence is essential, as it affects whether a simulation will complete within a practical timeframe. Three types of convergence rates are important:

1. Q-linear convergence: The ratio of the difference between successive points in a sequence is a constant r where $0 < r < 1$. This indicates that the sequence is converging linearly.

$$\frac{\| x_{k+1} - x^* \|}{\| x_k - x^* \|} \leq \gamma \quad \text{for all large } k \text{ and } \gamma \in (0,1)$$

2. Q-superlinear convergence: The ratio of successive differences approaches zero as the sequence progresses, indicating faster-than-linear convergence.

$$\lim_{k \to \infty} \frac{\| x_{k+1} - x^* \|}{\| x_k - x^* \|} = 0$$

3. Q-quadratic convergence: The ratio involves the square of the differences, indicating an even faster convergence rate.

$$\frac{\| x_{k+1} - x^* \|}{\| x_k - x^* \|^2} \leq M$$

These rates help categorize the performance of most algorithms. Understanding and quantifying these rates is crucial for ensuring the efficiency and practicality of optimization algorithms.