Lecture 2 and 3
(3 via video)

10 min

First - Regular problem recap.
At the end of these next two lectures, you should have a good grasp of the basic notation we'll use. We will also cover sequences and how they converge.
This will include information on Cauchy sequences as well as
- linear convergence
- exponential convergence
- geometric convergence
- arithmetic convergence
- quadratic convergence

As well as some simple algorithms that illustrate these convergence rates.

Logistics
Survey due Friday
TA Office hours on Friday before class
I will email the current list of enrolled students about a few questions regarding our course policies.

I'll post to Brightspace

Tom T. Huang

Preliminary

My handwriting is not always optimal for reading! I'm sorry

Sequences

$\{x_k\}$ is a sequence in $\mathbb{R}^n$

x_k understood to denote a vector not a scalar.

x_k k^{th} element in vector x

$\{x^{(k)}\}$ is also a sequence

so $x_j^{(k)}$ is the jth element in $x^{(k)}$

$\alpha, \beta, \gamma, \ldots$ are scalars

$A, B, C, \ldots$ are matrices.

The goal in optimization is to use an Algorithm / Procedure / Code / method / subroutine / etc to produce a sequence $\{x_k\}$ that converges to an optimum point of f

Ex:

$f(x) = x^2$

$f(0) = 0$ is the optimum pt.

Usually x_k goes on forever!

We write: Converge to follow

$\lim_{k \to \infty} x_k = x^*$

or $\lim_{k \to \infty} x_k = x^*$

What does $x_k \to x^*$ mean?

Formally...

If for any $\varepsilon > 0$

There exists a $J \geq 1$

$\|x^* - x_k\| \leq \varepsilon$

for all $k \geq J$, we say $x_k \to x^*$

What about $\|x^* - x_k\|$?

$\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$

$= \sqrt{x^T x}$

Hard to tell. Some times closed converge

Code = difficult

tricky bot

Slow for all x^*

then it's an ε

st. the sequence never gets closer than ε!

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

$x_k = 2 x_{k-1}$

$x_k \to \infty$,

$x_1 = 1$

$x_k = 2 x_{k-1}$

$f_k = \cos(x_k)$

Ex:

$f(x) = x^2$

$f(0) = 0$ is the Extrema pt.

Usually x_k goes on forever!

Why is this slow

Sometimes, we like to evaluate things that close smith converge...

$x_k = \begin{bmatrix} \cos\left(\frac{2\pi k}{10}\right) - 2^{-k} \\ \sin\left(\frac{2\pi k}{10}\right) - 2^{-k} \end{bmatrix}$

A sequence has a limit point or accumulation point when there is a subsequence that converge to x^*

If $\{x_k\}$ is the sequence then if $y_j = x_{k_j}$ is the subsequence and $y_j \to y$ then y is a limit pt of $\{x_k\}$

Ex:

$\{(x_k)\} = \left\{ \begin{bmatrix} (1-2^{-k}) \\ (-1)^k \end{bmatrix} \right\}$

the index = column vector!

$= \left\{ \begin{bmatrix} 1-2^{-k} \\ (-1)^k \end{bmatrix} \right\}$

$\Rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$1 - 2^{-k} \to 1$

$\log(x) \to \infty$

$\sqrt{\log(x)} \to \infty$

Ex: $\left\{ \begin{bmatrix} 2^{-k} \cos(\pi k) + 1 \\ 2^{-k} \sin(\pi k) \end{bmatrix} \right\}$

Does this converge?

Close guess!

Yes! $\Rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Why? $2^{-k} \to 0$ $\boxed{|\cos(\pi k)| \leq 1}$

A sequence is called Cauchy if for any $\varepsilon > 0$ there is a J s.t.

$\|x_k - x_l\| \leq \varepsilon$

for all $k, l \geq J$

Cauchy $\Leftrightarrow$ Converge!

A fundamental question in optimization

- How fast does my algorithm converge?

$\Leftrightarrow$

- How fast does my sequence converge? (cauchy)

Why? $x_k \to x_{k+1}$

We want $x_k \to x^*$ fast!

CS 590 - Purdue University - Computer Science - David F. Gleich - 2022

First notion of speed

if $\dfrac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \le \rho < 1$

holds, we say x_k converges linearly (Q-linear) at rate ρ.

for all $k \to \infty$

$\|x_{k+1} - x^*\| \le \rho^k \underbrace{\|x_0 - x^*\|}_{\delta}$

$\log(\rho^k \delta) = k \log(\rho) + \log(\delta)$

if $\rho = 0.5 = 2^{-1}$

if $\rho = 0.9999$

(0.999)$^{2^{12}}$ ≈ 0.6

for the class linear is slow convergence!

linear = geometric = exponential

What is slow than linear!

Attitude Convergence.

$\|x_{k+1} - x^*\| \le \boxed{\dfrac{1}{k^\sigma}} \delta_0$

if $\sigma = 1, 2, 0.5$.

Faster than linear Super linear!

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

Faster: Quadratic

$\dfrac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} \le \mu$

if $\|x_0 - x^*\| = 0.001$

$\mu = 10$ $(10) \cdot (0.001)^2 = (0.001) \cdot (0.001)$

$\|x_{k+1} - x^*\| \le \mu \cdot \|x_k - x^*\|^2$

$\le (0.0000 1)$

Faster than quadratic? Cubic!

$\dfrac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^3} \le \mu$

$10^{-3} \to 10^{-9}$

CS 590 - Purdue University - Computer Science - David F. Gleich - 2022