# Lecture 1d. Limits Review

COMP90051 Statistical Machine Learning

Sem2 2020
Lecturer: Ben Rubinstein

THE UNIVERSITY OF MELBOURNE

# This lecture

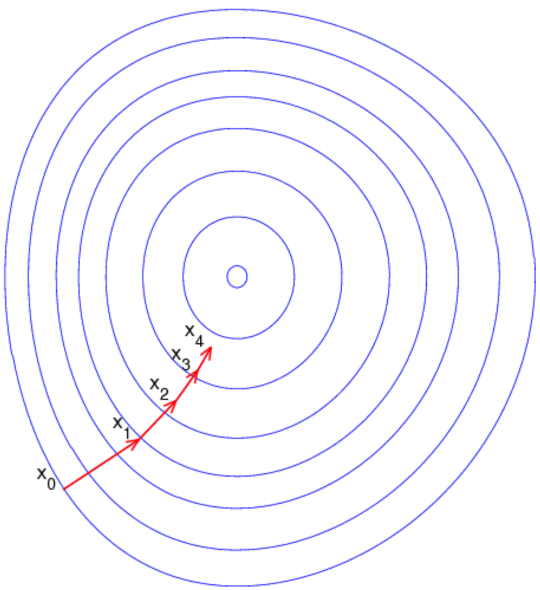
- About COMP90051
- Review: Probability theory
- Review: Linear algebra
- **Review: Sequences and limits**

# Sequences and Limits

Sequences arise whenever we have iterations (e.g. training loops, growing data sample size). Limits tell us about where sequences tend towards.

# Infinite Sequences

- Written like $x_1, x_2, \ldots$ or $\{x_i\}_{i \in \mathbb{N}}$
- Formally: a function from the positive (from 1) or non-negative (from 0) integers
- Index set: subscript set e.g. $\mathbb{N}$
- Sequences allow us to reason about test error when training data grows indefinitely, or training error (or a stopping criterion) when training runs arbitrarily long



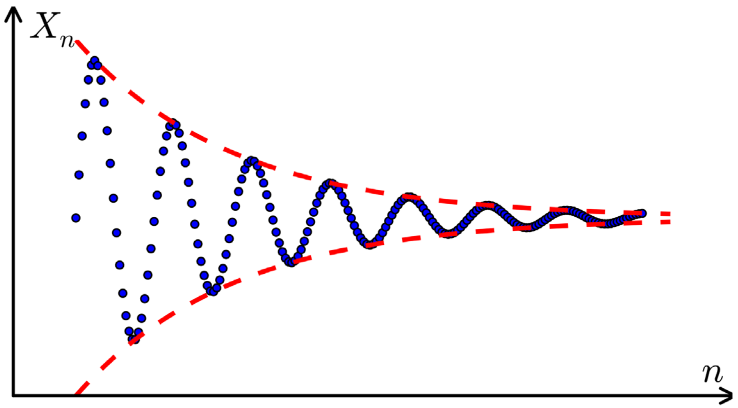

Wikipedia public domain

# Limits and Convergence

- A sequence $\{x_i\}_{i\in\mathbb{N}}$ converges if its elements become and remain arbitrarily close to a fixed limit point $L$.
- Formally: $x_i \to L$ if, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $\|x_n - L\| < \varepsilon$

Notes:

- Epsilon $\varepsilon$ represents distance of sequence to limit point
- Distance can be arbitrarily small
- Definition says we eventually get that close (at some finite $N$) and we stay *at least* that close for ever more




Wikipedia public domain

# Supremum

Generalising the maximum: When a sequence never quite peaks.

# When does the Maximum Exist?

- Can you always take a max of a set?
- Finite sets: what's the max of $\{1, 7, 3, 2, 9\}$?
- Closed, bounded intervals: what's the max of $[0,1]$?
- Open, bounded intervals: what's the max of $[0,1)$?
- Open, unbounded intervals: what's the max of $[0,\infty)$?

# What about "Least Upper Bound"?

- Can you always take a least-upper-bound of a set? (much more often!)
- Finite sets: what's the max of {1, 7, 3, 2, 9}?

    max=9 LUB=9

- Closed, bounded intervals: what's the max of [0,1]?

    max=1 LUB=1

- Open, bounded intervals: what's the max of [0,1)?

    max=N/A LUB=1

- Open, unbounded intervals: what's the max of $[0,\infty)$?

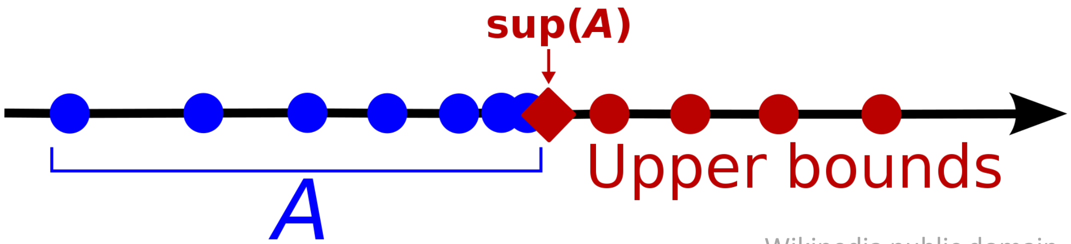
    max=N/A LUB=$\infty$

# The Supremum

- Consider any subset $S$ of the reals
- Upper bound $u \in \mathbb{R}^+$ of set $S$ has: $u \geq x$ for all $x \in S$
- If $u$ is no bigger than any other upper bound of $S$ then it's called a least upper bound or supremum of $S$, written as $\sup(S)$ and pronounced "soup":
  - $z \geq u$ for all upper bounds $z \in \mathbb{R}^+$ of $S$

FreeSVG public domain

- When we don't know, or can't guarantee, that a set or sequence has a max, it is better to use its sup




Wikipedia public domain

# Infimum

- The greatest lower bound or infimum is generalisation of the minimum
- Written inf($S$) pronounced "inf"
- Useful if we're minimising training error but don't know if the minimum is ever attained.

# Stochastic Convergence

When random events or quantities can sometimes be expected to converge (e.g. test error likely drops to a minimal value)

# Why Simple Limits Aren't Enough

- Consider running your favourite learner on varying numbers of $n$ training examples giving classifier $c_n$
- If your learner minimises training error, you'd wish its test error wasn't much bigger than its training error
- If $R_n = err_{test}(c_n) - err_{train}(c_n)$, you'd wish for $R_n \to 0$ as this would mean eventually tiny test error
- But both training data and test data are random!
- Even if $R_n \to 0$ *usually* happens, it won't *always*!!

# Stochastic Convergence

- A sequence $\{X_n\}$ of random variables (CDFs $F_n$) converges in distribution to random variable $X$ (CDF $F$) if $F_n(x) \to F(x)$ for all constants $x$
- A sequence $\{X_n\}$ of random variables converges in probability to random variable $X$ if for all $\varepsilon > 0$: $\Pr(|X_n - X| > \varepsilon) \to 0$
- A sequence $\{X_n\}$ of random variables converges almost surely to random variable $X$ if:$\Pr(X_n \to X) = 1$
- Chain of implications:

almost sure (strongest) $\Longrightarrow$ in probability $\Longrightarrow$ in distribution (weakest)

# But **don't** worry…

- We're not going to do *any* calculations with stochastic convergence
- Close understanding of it won't be necessary in this subject
- But it's good to be aware that its "out there" and we **may refer to it** (v briefly) within StatML theory

CCA4.0 Vincent Le Moign

# Summary

- Sequences
- Limits of sequences
- Supremum is the new maximum
- Stochastic convergence

Next time: L02 Statistical schools

Homework week #1: Watch all week 1 recordings.
Jupyter notebooks setup and launch (at home)