CPSC 550: Machine Learning II	2008/9 Term 2

Lecture 6 — Feb 10, 2009

Lecturer: Nando de Freitas *Scribe: Matt Hoffman*

6.1 Background

Consider some probability measure ν and some event E whose complement $\overline{E}$ has null-measure, i.e. $\nu(\overline{E}) = 0$. Then for some statement that holds for all $w \in E$ we say that the statement holds *almost surely.* In other words, for some probability triple $(\Omega, \mathfrak{F}, P)$ an event $E \in \Omega$ occurs a.s. (almost surely) if $P(E) = 1$. We can now define the following limits:

1. *Weak* or *in probability convergence* holds if

$$\lim_{n\to\infty} P\big(\|X_n - X\| > \epsilon\big) = 0 \quad \forall\epsilon$$

   and we write this as

$$X_n \xrightarrow{P} X \quad \text{or} \quad \operatorname*{plim}_{n\to\infty} X_n = X.$$

2. *Strong* or *almost sure convergence* (also known as *convergence with probability 1*) holds if

$$P\Big(\lim_{n\to\infty} X_n = X\Big) = 1$$

   and we write this as $X_n \xrightarrow{\text{a.s.}} X$. Using the definition of an almost sure event above we can expand this as

$$P\Big(\{\omega \in \Omega | \lim_{n\to\infty} X_n(\omega) = X(\omega)\}\Big) = 1$$

   i.e. every ω for which this does not occur will have null-measure. If the measurable space in question is the Borel σ-algebra this will include even uncountable sets such as the Cantor set.

3. Finally, *sure convergence* holds if

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \quad \forall\omega,$$

   i.e. if X_n and X assign the same measure to every event in the limit.

Given these definitions we can write the following theorem:

Theorem 6.1 (Measure continuity). *Given an infinite sequence of sets $\{A_i\}$ where each set is a subset of the following set, i.e. $A_i \subset A_{i+1}$, the following limit holds:*

$$P\Big(\lim_{n\to\infty} A_n\Big) = \lim_{n\to\infty} P(A_n).$$

Corollary 6.2. *A consequence of this theorem is that the probability of unions of these sets is also given by the limit $P(\cup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} P(A_n)$.*

Proof: We first note that each set can be written as a union of disjoint sets

$$A_n = A_1 \cup (A_2 - A_1) \cup (A_3 - A_2) \cup \cdots \cup (A_n - A_{n-1}).$$

Also we can trivially see by induction that each A_i for $i < n$ is a subset of A_n and as a result

$$\bigcup_{i=1}^{\infty} A_i = \lim_{n\to\infty} \bigcup_{i=1}^{n} A_i = \lim_{n\to\infty} A_n.$$

Finally we can see that the probability of these unions is

$$P(\cup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} \left[P(A_1) + \sum_{i=2}^{\infty} P(A_j - A_{j-1}) \right] = \lim_{n\to\infty} P(A_n).$$

□

6.2 The limit inferior and superior

Given a sequence $\{X_n\}$ we will define the limit inferior and limit superior respectively as

$$\liminf_{n\to\infty} X_n = \sup_{n\geq 0} \inf_{m\geq n} X_m = \lim_{n\to\infty} \inf_{m\geq n} X_m \text{ and}$$
$$\limsup_{n\to\infty} X_n = \inf_{n\geq 0} \sup_{m\geq n} X_m = \lim_{n\to\infty} \sup_{m\geq n} X_m.$$

An example of this behavior is shown in Figure 6.1 for a sequence of real numbers.

The infimum and supremum are also well-defined for sets, and thus the liminf and limsup exist where $\{X_n\}$ is a sequence of sets. Here $\liminf X_n$ and $\limsup X_n$ are sets. Intuitively, for some element a

- $a \in \liminf X_n$ if and only if there exists some integer n_0 such that $a \in X_n$ for all $n > n_0$. This means that the liminf consists of all elements that are in X_n for all but finitely many n.
- $a \in \limsup X_n$ if and only if for all n_0 there exists $n > n_0$ such that $a \in X_n$. This means that the limsup consists of all elements that are in X_n for infinitely many n. We say that these elements occur *infinitely often* (or i.o.).

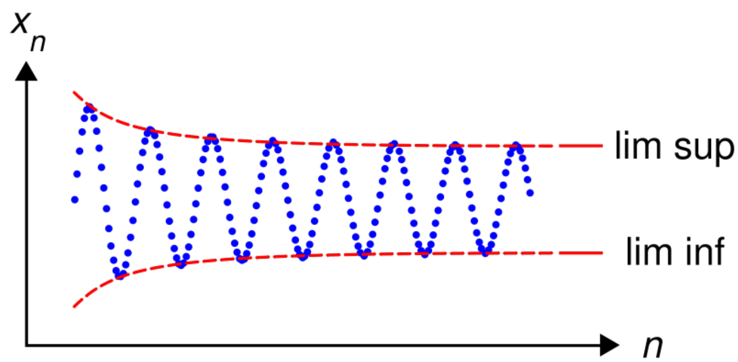

Figure 6.1. A depiction of the liminf and limsup of some sequence of real numbers $\{X_n\}$. This figure is taken from Wikipedia, `http://commons.wikimedia.org/wiki/Image:LimSup.svg`.

Example. Consider a sequence of sets $\{X_n\}$ corresponding to

$$\{0\}, \{1\}, \{0\}, \{1\}, \ldots$$

The liminf and limsup of this sequence of sets is given by

$$\begin{aligned} \operatorname{limsup} X_n &= \{0, 1\} \text{ and} \\ \operatorname{liminf} X_n &= \emptyset. \end{aligned}$$

The definitions given before weren't entirely precise, in that we never introduced the exact definition of the infimum and supremum of a sequence of sets. We will do so now and then use this to give an alternative definition of the liminf and limsup. For some sequence of sets $\{X_n\}$ the infimum and supremum can be directly related to countable unions and intersections, i.e.

$$\inf_{m \geq n} \{X_m\} = \bigcap_{m=n}^{\infty} X_m \quad \text{and} \quad \sup_{m \geq n} \{X_m\} = \bigcup_{m=n}^{\infty} X_m.$$

Since the infimum are defined in terms of these intersections we can define a sequence of infimum $I_n = \inf_{m \geq n} X_m$ and it is easy to see that $I_n \subset I_{n+1}$ for all n. We can then write the liminf as

$$\operatorname{liminf} X_n = \lim_{n \to \infty} \bigcap_{m=n}^{\infty} X_m = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} I_m.$$

Similarly we can see that a sequence of supremum $S_n = \sup m \geq n X_m$ are disjoint and as a result write the limsup as

$$\operatorname{limsup} X_n = \lim_{n \to \infty} \bigcup_{m=n}^{\infty} X_m = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} S_m.$$