MAT377 Lecture Notes

Arky!! :3c

'24 Fall Semester

Contents

§1 Day 1: Introduction to Probability (Sep. 4, 2024)

Link to textbook.

We start with a sampler problem that on the surface, seems unrelated to probability. Let $v_1, \dots, v_n \in \mathbb{R}^n$ be unit vectors on the unit sphere, i.e.t $\|v_i\| = 1$. If we are to pick $\varepsilon_i = \{-1, 1\}$ at random, what is our expectation on how large will

$$\sum_{i=1}^{n} \varepsilon_i v_i$$

be? We could brute force and average out over all probabilities as follows,

$$\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \left(\sum_{i=1}^{n} \varepsilon_i v_i \right) = \sum_{i=1}^{n} \left(\underbrace{\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \varepsilon_i}_{=0} \right) v_i = 0.$$

Now, consider that

$$\begin{aligned}\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \left| \sum_{i=1}^{n} \varepsilon_i v_i \right|^2 &= \frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{i_1, i_2 = 1}^{n} \varepsilon_{i_1} \varepsilon_{i_2} \langle v_{i_1}, v_{i_2} \rangle \\ &= \sum_{i_1, i_2 = 1}^{n} \left(\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \varepsilon_{i_1} \varepsilon_{i_2} \right) \langle v_{i_1}, v_{i_2} \rangle\end{aligned}$$

To simplify the bracketed summation, we could consider the following two cases:

- If $i_1 \neq i_2$, we would have that

$$\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \varepsilon_{i_1} \varepsilon_{i_2} = \frac{2^{n-2}}{2^n} \sum_{\substack{\varepsilon_{i_1} \in \{-1,1\} \\ \varepsilon_{i_2} \in \{-1,1\}}} \varepsilon_{i_1} \varepsilon_{i_2} = 0.$$

- If $i_1 = i_2$, we would have

$$\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \varepsilon_{i_1} \varepsilon_{i_2} = \frac{2^{n-1}}{2^n} \sum_{\varepsilon_i \in \{-1,1\}} \varepsilon_{i_1} \varepsilon_{i_2} = 1.$$

By linearity of expectation, we obtain

$$\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \left| \sum_{i=1}^{n} \varepsilon_i v_i \right|^2 = n,$$

and

$$\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \left| \sum_{i=1}^{n} \varepsilon_i v_i \right| = c\sqrt{n},$$

where c is a positive real constant.

We now abstract a few probability terms;

- Ω is a sample space, i.e. the set of possible outcomes.
- Let P denote probability, i.e. a mapping of subsets of Ω to $[0,1]$ (read: probability of getting these subsets of Ω); the probability of an event ε out of S occuring is given by $P(\varepsilon \in S, S \subset \{-1,1\}^n) = \frac{1}{|S|}$, assuming that each event in S is equally likely. With this, we have three important properties of P to define:
  1. $P(\Omega) = 1$; the chance of an event in the probability space happening is 1.
  2. Let $\mathcal{F}$ be a collection of subsets $A_1, \ldots, A_n$. Then
  $$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i),$$
  given that $A_i \cap A_j = \emptyset$ for all $1 \leq i, j \leq n$. This is linearity of expectation.
  3. $P(A^C) = 1 - P(A)$, which is a property of set complement.
- When our collection $\mathcal{F}$ of subsets of Ω satisfy the following properties, we call it a σ-algebra:
  1. $\emptyset \in \mathcal{F}$,
  2. Closed under countable union: $A_n \in \mathcal{F} \implies \bigcup_{i=1}^{n} A_i \in \mathcal{F}$,
  3. Closed under complement: $A_i \in \mathcal{F} \implies A_i^C \in \mathcal{F}$.

  In a finite sample space, the power set $\mathcal{F} = \mathcal{P}(\Omega)$ is one such example of a σ-algebra. As an example, let $\Omega = [0,1)$; then $P([a,b]) = b - a$ (wlog, let $a < b$). Now, let $\mathcal{F} = \mathcal{P}([0,1))$. Define the equivalence $x \sim y$ if $x - y \in \mathbb{Q}$.
- We now introduce the axiom of choice; Let A be a set containing one element of each equivalence class from the above defined equivalence. Consider $\tau_q A := \{A\} + q$; let us claim that
  $$\bigcup_{q \in \mathbb{Q}} \tau_q A = [0,1),$$
  which is a countable union of $[0,1)$, since $\mathbb{Q}$ is countable. We have that $P(A) = P(\tau_q A)$ because intervals don't change size under shifting by q. However, observe that
  $$P([0,1)) = P\left(\bigcup_{q \in \mathbb{Q}} \tau_q A\right) = \sum_{q \in \mathbb{Q}} P(\tau_q A).$$
  Then either
  $$\begin{aligned} P(A) = 0 &\implies P(\tau_q A) = 0 \implies P([0,1)) = 0, \text{or} \\ P(A) \neq 0 &\implies P(\tau_q A) \to \infty \implies P([0,1)) \to \infty, \end{aligned}$$
  which doesn't make sense (for now). This shows that we need to pick our σ-algebra properly; observing that the intersection of two σ-algebras is also a σ-algebra, it is appropriate to let $\mathcal{F}$ be the smallest σ-algebra containing $[a,b)$. This is called a *Borel Set*.