Final Project

STAT 447

2024-04-11

Introduction and Review:

Before we discuss Dirichlet Processes, it is crucial to establish a groundwork in probability measure theory. We will briefly revisit the concepts of σ-algebra, probability measures and Dirichlet Distributions. The relevant literature for probability measure theory are (Billingsley 2012) and (Rudin 1986). In the following, we present generalized definitions, which are discussed in rigour in these foundational measure theory works.

Let $\mathbb{X}$ be a well-defined sample space. A σ-algebra $\mathcal{F} \subseteq P(\mathbb{X})$ is a set satisfying the following:

1. The entire sample space $\mathbb{X}$ is in $\mathcal{F}$ and $\emptyset$ is in $\mathcal{F}$. This is referred to in the literature as "non-emptiness and universality."
2. For all sets $A \in \mathcal{F}$, the complement $A^c \in \mathcal{F}$. This property is referred to as "closure under complementation."
3. For any countable collection of sets $\{A_i\}_{i \in I}$, where I is a countable index set, if $\forall i \in I, A_i \in \mathcal{F}$ then $\bigcup_{i \in I} A_i \in \mathcal{F}$. This is referred to as "closure under countable unions."

For the sake of this work, we are more interested in *probability measures*, which are built on σ-algebra. A probability measure $\mu : \mathcal{F} \mapsto [0, 1]$ satisfies the following familiar axioms of probability:

1. All mappings are non-negative: $\forall A \in \mathcal{F}, \ \mu(A) \geq 0$. This is referred to as "non-negativity."
2. $\mu(\mathbb{X}) = 1$, and $\mu(\emptyset) = 0$.
3. For any countable set $\{A_i\} \subseteq \mathcal{F}$ where $\forall i \neq j, A_i \cap A_j = \emptyset$, we have that $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$, this is referred to as "countable additivity."

Before discussing Dirichlet Distributions and Processes, we acknowledge that the above definitions might be abstract for those new to measure theory. For clarity, the Appendix includes examples to illustrate σ-algebra properties and verify a probability measure μ on a simple finite set.

(Teh 2006)

Appendix

Acknowlededgements: Miscellaneous information such as knowledge on sets, power sets, subsets and countability from (Demirbas and Rechnitzer 2023).

Example of a σ-Algebra

For additional clarity, we provide an example of a σ-algebra $\mathcal{F}$ to demonstrate the properties mentioned in the literature review.

Let's consider the finitely countable and simple set $\mathbb{X} = \{a, b, c\}$. Directly, $P(\mathbb{X}) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$ is the power set of $\mathbb{X}$. Consider $\mathcal{F} = \{\emptyset, \{a\}, \{b, c\}, \{a, b, c\}\} \subseteq P(\mathbb{X})$.

We will apply the properties discussed in the literature review section to prove that $\mathcal{F}$ is a σ-algebra.

To verify Property 1 (universality and non-emptiness), we note that we can also write $\mathcal{F}$ as $\{\emptyset, \{a\}, \{b, c\}, \mathbb{X}\}$. From this definition, it is direct to see that $\mathbb{X} \in \mathcal{F}$ and $\emptyset \in \mathcal{F}$, verifying universality and non-emptiness.

To verify Property 2 (closure under complementation), we note that $A^{c^c} = A$. Hence, if we show $A \in \mathcal{F} \implies A^c \in \mathcal{F}$ this is equivalent to showing that $A \in \mathcal{F} \iff A^c \in \mathcal{F}$. Since $\mathcal{F}$ is finitely countable, we can consider a case-wise basis for verification. Firstly, we have $A = \emptyset$. By definition, $A^c = \mathbb{X}$. We see that $\mathbb{X} \in \mathcal{F}$. As mentioned before, this implies that the case where $A = \mathbb{X}$ also holds. Now, we can proceed to verify the case where $A = \{a\}$. We note that $\{a\}^c = \{b, c\}$, and that $\{b, c\} \in \mathcal{F}$. Therefore, this case holds and thus so does the case where $A = \{b, c\}$ by biconditionality. Hence, we can conclude that $\mathcal{F}$ is closed under complementation.

To verify Property 3 (closure under countable unions), we first consider the concrete case where $\{A_i\}_{i=1}^2 = \{\{a\}, \{b, c\}\}$ where $A_1 = \{a\}$ and $A_2 = \{b, c\}$. We note that $A_1, A_2 \in \mathcal{F}$, so we would expect that $A_1 \cup A_2 \in \mathcal{F}$. Directly, $A_1 \cup A_2 = \{a\} \cup \{b, c\} = \{a, b, c\} = \mathbb{X}$, and we see that $\mathbb{X} \in \mathcal{F}$. For the remaining cases, $\forall A \in \mathcal{F}, A \cup \emptyset = A$ by fundamental set properties, and directly $A \in \mathcal{F}$ by construction. Similarly, $\forall A \in \mathcal{F}, A \cup \mathbb{X} = \mathbb{X}$ and we know that $\mathbb{X} \in \mathcal{F}$. Hence, for all cases, $\mathcal{F}$ is closed under countable unions.

From all of these properties, we can conclude that $\mathcal{F}$ is a σ-algebra, as required $\square$.

Example of a Probability Measure

Using the σ-algebra $\mathcal{F} = \{\emptyset, \{a\}, \{b, c\}, \{a, b, c\}\}$ we will define $\mu : \mathcal{F} \mapsto [0, 1]$ and prove that μ is a probability measure using the properties discussed in the literature review.

We will define a concrete example as follows, and show it is a probability measure on $\mathcal{F}$.

$$\mu(A) = \begin{cases} 2/3, & |A| = 1 \\ 1/3, & |A| = 2 \\ 1, & |A| = 3 \\ 0, & \text{otherwise} \end{cases}$$

Let's evaluate the properties discussed in the literature review to verify that μ is in fact a probability measure.

First, we evaluate Property 1 (non-negativity.) In effect, we wish to verify that $\forall A \in \mathcal{F}, \mu(A) \geq 0$. We can easily evaluate the universal in a case-wise basis.

    a. $\emptyset \in \mathcal{F}$ and $\mu(\emptyset) = 0 \geq 0$.
    b. $\{a\} \in \mathcal{F}$ and $\mu(\{a\}) = 2/3 \geq 0$.
    c. $\{b, c\} \in \mathcal{F}$ and $\mu(\{b, c\}) = 1/3 \geq 0$.
    d. $\mathbb{X} \in \mathcal{F}$ and $\mu(\mathbb{X}) = 1 \geq 0$.

Hence, $\forall A \in \mathcal{F}, \mu(A) \geq 0$, verifying the non-negativity clause. Further, we see that $\forall A \in \mathcal{F}, 0 \leq \mu(A) \leq 1$.

In addition, we can utilize the evaluations above to verify Property 2. Directly, we see that $\mu(\emptyset) = 0$ and $\mu(\mathbb{X}) = 1$, verifying Property 2 that a value of 1 is assigned to the sample space $\mathbb{X}$.

Finally, we verify Property 3 (countable additivity.) Again, because the σ-algebra is finitely countable, we verify all pairwise disjoint intersections on a case-wise basis.

    a. First, we consider the set of pairwise disjoint sets $\{\emptyset, A\}$ for $A \in \mathcal{F}$. We see that $\forall A \in \mathcal{F}, \emptyset \cap A = \emptyset$ and $\emptyset \cup A = A$. Using, $\mu(\emptyset) = 0$, we observe that $\mu\left(\bigcup_i A_i\right) = \mu(A) = \mu(A) + \mu(\emptyset) = \sum_{i=1} \mu(A_i)$, as required.

b. Similarly, we consider $\mathbb{X} \in \mathcal{F}$, noting that $\forall A \in \mathcal{F}, \mathbb{X} \cap A = A$ is non-disjoint, so the case holds vacuously by falsity of the antecedent. A similar argument can be applied for $\{A, A\}, A \in \mathcal{F}$ which is evidently a non-disjoint pair.

c. The other pairwise disjoint set in $\mathcal{F}$ is $\{A_i\} = \{\{a\}, \{b, c\}\}$, since $\{a\} \cap \{b, c\} = \emptyset$. Hence, this pair should be countably additive. We can see directly that $\bigcup_i A_i = \{a\} \cup \{b, c\} = \mathbb{X}$, Hence, $\mu\left(\bigcup_i A_i\right) = \mu(\mathbb{X}) = 1$, so we anticipate $\sum_i \mu(A_i) = 1$. To verify, we see that $\sum_i \mu(A_i) = \mu(\{a\}) + \mu(\{b, c\}) = 2/3 + 1/3 = 1$, so countable additivity holds.

From all of these properties, we can conclude that μ is a probability measure on σ-algebra $\mathcal{F}$, as required $\square$.

$$\mathbb{G}_0 : \left\langle \text{dir}(\alpha_{j,1:K(j)}), N(m_{j,k}, s_{j,k}^2) \right\rangle = \left\langle \text{dir}(1_{1:4}), N(0, 10) \right\rangle$$

Infinite Dirichlet Process Mixture Model
$$\{\pi_k\}_{k=1}^{\infty} \sim \text{GEM}(\alpha)$$
$$\{\langle \alpha_k, \beta_k \rangle\}_{k=1}^{\infty} \sim \mathbb{G}_0$$
$$y_i \mid \{\pi_k\}, \{\langle \alpha_k, \beta_k \rangle\} \sim \sum_{k=1}^{\infty} \pi_k \text{beta}(y_i \mid \langle \alpha_k, \beta_k \rangle)$$

So, we would... Define prior $\mathcal{P} \sim \text{DP}(\alpha \mathbb{G}_0)$ The implemented code for the Gibbs Sampler is modified from the DPMM code by (Komodakis 2023)

Sources

Billingsley, Patrick. 2012. *Probability and Measure, Anniversary Edition.* Wiley.

Demirbas, Seckin, and Andrew Rechnitzer. 2023. "An Introduction to Mathematical Proof : MATH 220." Free web and pdf textbook. https://personal.math.ubc.ca/~PLP/.

Komodakis, Manolis. 2023. "Dirichlet Process Mixture Models." *GitHub Repository.* https://github.com/mkomod/dpmm/tree/master/mixture_models; GitHub.

Rudin, Walter. 1986. *Real and Complex Analysis.* 3rd ed. McGraw-Hill.

Teh, Yee Whye. 2006. "Dirichlet Process." Course Notes for Gatsby Computational Neuroscience Unit Tutorial. https://mlg.eng.cam.ac.uk/zoubin/tut06/ywt.pdf.