DP Notes, 10/8

1 Kolmogorov Extension Theorem

Before we get into the lecture notes, let's look at the Kolmogorov Extension Theorem. This section is a lightly modified version of Erhan Çinlar's treatment of the subject (Ch.4, Sec. 4).

We would like to show the existence of a probability space $(\Omega, \mathcal{H}, \mathbb{P})$ that can carry a process $\{X_t : t \in I\}$, where I is an index set that is arbitrary (and possibly continuous). It is helpful to think of I as time. All the X_t must take values in the same measurable set $(E, \mathcal{E})$.

Let I be an arbitrary index set and $(E, \mathcal{E})$ a measurable space. We are given the probability measures π_J, one for each finite subset J of I on the product space $(E^J, \mathcal{E}^J)$. The goal is to construct a probability space $(\Omega, \mathcal{H}, \mathbb{P})$ and a stochastic process $X = (X_t)_{t \in I}$ over it such that π_J is the distribution of the random variable $X_J = (X_t)_{t \in J}$ for each finite $J \subset I$.

We start by letting $(\Omega, \mathcal{H}) = (E, \mathcal{E})^I$. That is, Ω is the collection of all functions $t \mapsto \omega(t)$ from I into E, and $\mathcal{H}$ is the σ-algebra generated by the finite-dimensional measurable rectangles. Define X_t to be the coordinate variables: $X_t(\omega) = \omega(t)$ for all t and ω. Each X_t is measurable with respect to $\mathcal{H}$ and $\mathcal{E}$. Moreover, $\mathcal{H} = \sigma\{X_t : t \in I\}$.

For $I \supset J \supset K$, let p_{JK} be the natural projection from E^J onto E^K. For instance, if $J = (s, t, u)$ and $K = (t, u)$, then $p_{IJ}(\omega) = (\omega(s), \omega(t), \omega(u))$, $p_{JK}(x, y, z) = (y, z)$ and $p_{IK}(\omega) = (\omega(u), \omega(t))$. Let $\mathcal{J}_f$ denote the collection of all finite sequences of elements of I, and $\mathcal{J}_c$ the collection of all infinite (countable) sequences.

The probability measure $\mathbb{P}$ we are seeking will be the probability law of X; accordingly, we want

$$\mathbb{P}\{X_J \in A\} = \pi_J(A), \qquad A \in \mathcal{E}^J, \quad J \in \mathcal{J}_f. \tag{1}$$

This requires that the finite dimensional distributions be consistent:

$$\pi_K = \pi_J \circ p_{JK}^{-1}, \qquad K \subset J \in \mathcal{J}_f, \tag{2}$$

since $X_K = p_{JK} \circ X_J$ for $K \subset J$. The following is the Kolmogorov Extension Theorem:

Theorem 1.1. *Suppose that $(E, \mathcal{E})$ is a standard measure space and that $\{\pi_J : J \in \mathcal{J}_f\}$ satisfies*

$$\mathbb{P}\{X_J \in A\} = \pi_J(A), \qquad A \in \mathcal{E}^J, \quad J \in \mathcal{J}_f. \tag{3}$$

Then, there exists a unique probability measure $\mathbb{P}$ on $(\Omega, \mathcal{H})$ such that

$$\pi_K = \pi_J \circ p_{JK}^{-1}, \qquad K \subset J \in \mathcal{J}_f \tag{4}$$

holds.

2 The Mathematical Underpinnings of Dirichlet Processes

2.1 Assumptions

Let $\mathcal{X}$ be a discrete space and α be a finite measure on $\mathcal{X}$. Let $\bar{\alpha} = \frac{\alpha}{\alpha(\mathcal{X})}$. Suppose that

$$\begin{aligned} G &\sim Dirichlet(\alpha) \\ X_n &\sim G \end{aligned} \tag{5}$$

We want a predictive distribution for the next X where

$$\begin{aligned} \mathbb{P}\{X = i | X_1, \ldots, X_N\} &= \frac{\alpha(\mathcal{X})}{\alpha(\mathcal{X}) + N} \bar{\alpha}(i) + \frac{N}{\alpha(\mathcal{X}) + N} \frac{n_i}{N} \\ &= \frac{\alpha(i) + n_i}{\alpha(\mathcal{X}) + N} \end{aligned} \tag{6}$$

where n_i is the number of times that i has been seen in the data.

2.2 Does a Dirichlet Process Actually Exist?

This is a discussion of Sections 3.1-2 of (1) in the case of a finite space. Let $A_1, \ldots, A_k$ be a partition of $\mathcal{X}$. Let $M(\mathcal{X})$ be the set of all probability measures on $\mathcal{X}$.

Theorem 2.1. *Let α be a finite measure on $\mathcal{X}$. Then there exists a unique probability measure D_α on $M(\mathcal{X})$ called the Dirichlet process with parameter α satisfying*

$$(G(A_1), G(A_2), \ldots, G(A_k)) \sim \mathit{Dirichlet}(\alpha(A_1), \alpha(A_2), \ldots, \alpha(A_k)) \tag{7}$$

for every partition $A_1, A_2, \ldots, A_k$ of $\mathcal{X}$.

This follows from the Kolomogorov Extension Theorem. The consistency requirement of Equation 4 comes from the Tail-Free property of Equation 7:

$$G(\cdot | A_i) \sim \mathit{Dirichlet}(\alpha_{A_i}), \tag{8}$$

where $G(A_i) > 0$ and α_{A_i} is the restriction of α into A_i. (1) extends this construction to Dirichlet processes on the real line.

References

[1] Ghosh, J. and Ramamoorthi, R., *Bayesian Nonparametrics*, Springer, 2003.