Summary 5: The Indian Buffet Process: An Introduction and Review

Susan Cherry

This paper introduces the Indian Buffet Process (IBP), which applies nonparametrics to stochastic models in which objects are represented with an unknown number of latent features. It specifies how the IBP can be used to specify a prior distribution in latent feature models. The IBP can also be used to define a prior distribution in any setting where the latent structure can be expressed in the form of a binary matrix with a finite number of rows and infinite number of columns.

The paper begins by discussing latent class models. First, it describes finite mixture models and then extends to infinite mixture models. The paper's review of infinite mixture models shows that infinite statistical models can be defined by specifying priors over infinite combinatorial objects and it shows how to derive these priors by taking the limit of priors of finite models. Next, the paper discusses latent feature models, in which each object is represented by a vector of latent feature values f_i and the properties x_i are generated from a distribution that is determined by the latent feature values. Again, the paper first describes the finite feature model and then takes the infinite limit.

This brings us to the IBP. Intuitively, the process works as follows: N customers enter a restaurant containing infinite dishes, one after another. The first customer starts at the buffet and takes a serving of each dish, stopping after a Poisson(α) number of dishes. The ith customer samples dishes in proportion to their popularity, serving with a probability of $\frac{m_k}{i}$, where m_k is the number of previous customers who have sampled a dish. At the end of the previously sampled dishes, the customer tries a Poisson($\frac{\alpha_i}{i}$) number of new dishes. The paper then shows an example of a linear-gaussian latent feature model with binary features and illustrates how inference can be conducted using Gibbs Sampling.

Finally, the paper discusses how the IBP is related to the Dirichlet Process. A simple way to think about the Dirichlet process is to think of DP as a probability measure over probability measures. A straightforward way to think about this is the stick-breaking construction that we have discussed multiple times in class. A similar stick breaking construction can be defined for the IBP. Sort the π_k that represent the probability of each feature being possessed by an object from largest to smallest. The distribution of the sequence of stick lengths corresponds to the distribution of these ordered probabilities. The stick breaking construction was used in defining the variational inference algorithm that is discussed earlier in the paper and can be used to derive other inference algorithms for the iBP, as well.