Summary of "The Indian Buffet Process: An Introduction and Review"
Matt Dickenson `mcd31`
STA571/CS590.01   Due: 10 February, 2014

Griffiths and Ghahramani (2011) present several derivations of the Indian Buffet Process (IBP), and describe a number of applications and adaptations. By analogy to the Chinese Restaurant Process, the IBP describes customers visiting a buffet and filling their plates with its seemingly infinite number of dishes. N customers enter the restaurant in sequence, and the first fills her plate with a Poisson(α) number of dishes. Then, the ith customer samples dishes in proportion to their popularity, with probability $\frac{m_k}{i}$, where m_k is the number of previous customers who sampled dish k. The ith customer then samples $K_1^{(i)} \sim$ Poisson($\frac{\alpha}{i}$) number of new dishes. The customers represent objects and their chosen dishes represent features; in this way, the IBP serves as a prior for sparse binary matrices.

Of the potentially infinite number of columns in the binary matrix $\mathbf{Z}$ ($N \times K$), only a finite number will have non-zero entries. This number is the effective dimension of the matrix, K_+. Because of the way that the Poisson parameter is reduced with each new customer, $K_+ \sim$ Poisson(αH_N) (where H_N is the Nth harmonic number). By exchangeability, the number of dishes on each customer's plate is distributed Poisson(α). $\mathbf{Z}$ remains sparse as $K \to \infty$: effective dimensions of $\mathbf{Z}$ are $N \times K_+$, and the expected number of entries is $N\alpha$. In some applications this coupling of the average number of features per object and the total number of features may be undesirable, in which case the two-parameter generalization of the IBP can be applied.

The IBP and the CRP are related in more than just their analogies to cuisine. The Beta process is the de Finetti mixing distribution underlying the Indian buffet process, just as the Dirichlet process is for the CRP. This parallel helps to develop analogous derivations of the IBP, including the stick-breaking construction. The IBP can be computationally time-consuming for inference, but ongoing research has developed alternative inference strategies to the Gibbs sampler presented by Griffiths and Ghahramani that are more appropriate to other applications (such as variational inference).