

## Summary of “Hierarchical Beta Processes and the Indian Buffet Process”

Matt Dickenson mcd31

STA571/CS590.01

Due: 12 February, 2014

---

Thibaux and Jordan (2007) build on the development of the Indian Buffet Process (IBP) as a prior for binary matrices as introduced by Griffiths and Ghahramani (2005). Griffiths and Ghahramani show that a reordering of feature matrices, as generated by the IBP, into their “left-ordered form” (i.e. treating the columns as binary numbers and sorting in descending order) produces a class of equivalent matrices. Thus, the IBP is an exchangeable distribution over binary matrices. By de Finetti’s theorem, any infinitely exchangeable sequence can be written as  $P(Z_1, \dots, Z_n) = \int \{\prod_{i=1}^n P(Z_i|B)\} dP(B)$ . This means that there exists a random element  $B$  that renders the  $Z_i$  conditionally independent.

Thibaux and Jordan show that the beta process serves as  $P(B)$  for matrices generated by the IBP. That is, 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 aids in the development of the two-parameter generalization of the IBP, in which  $\pi_k|\alpha, \beta \sim \text{Beta}(\frac{\alpha\beta}{K}, \beta)$ . As in the one-parameter IBP, the first customer fills her plate with a  $\text{Poisson}(\alpha)$  number of dishes. Then, the  $i^{\text{th}}$  customer tastes previously sampled dishes, with probability  $\frac{m_k}{\beta+i-1}$ , where  $m_k$  is the number of previous customers who sampled dish  $k$ . The  $i^{\text{th}}$  customer then samples an additional  $\text{Poisson}(\alpha\beta/(\beta+i-1))$  new dishes. The average number of dishes per customer (i.e. features per object) is still  $\alpha$ , but the expected number of overall features and non-zero columns is  $\alpha \sum_{i=1}^n \frac{\beta}{\beta+i-1}$ .  $\beta$  gives us control over the size and sparsity of the matrix, decoupling the average number of features from the total. Understanding the beta process as the basis for the IBP also shows how a hierarchical version can be developed, allowing sets of features to be shared across nested data. Thibaux and Jordan’s contribution helps to deepen our understanding of the IBP and extend it to new applications.