

Summary of “Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures”

Matt Dickenson mcd31
STA571/CS590.01

Due: 24 March, 2014

Orbanz and Roy provide an introduction to exchangeability in Bayesian models of arrays of random variables, such as matrices and graphs. Their results apply generally to d -dimensional arrays but are more accessible when understood in reference to sequences (1-arrays) and matrices (2-arrays). Exchangeability is an important assumption in Bayesian modeling, because if it holds then the empirical distributions of observations from experiments converge to the random probability measure Θ as $n \rightarrow \infty$ with probability 1 for every measurable set A :

$$\hat{S}_n(\cdot) := \frac{1}{n} \sum_{i=1}^n \delta_{X_i}(\cdot)$$

$$\lim_{n \rightarrow \infty} \hat{S}_n(A) \rightarrow \Theta(A)$$

The Chinese Restaurant Process is a well-known example of a random partition.

In many types of data (e.g. time series) the observations themselves are not exchangeable. This obstacle is not insurmountable, as components of the model other than the variables representing the observations can be assumed to be exchangeable. Graphs and matrices are other examples where it seems difficult to assume exchangeability, because the rows and columns of a matrix are not disjoint as in de Finetti's theorem.

The authors rely on the Aldous-Hoover theorem for two notions of exchangeability to address this: jointly exchangeable and separately exchangeable matrices. A matrix is jointly exchangeable iff there is a random measurable function $F : [0, 1]^3 \rightarrow X$ such that

$$(X_{ij}) := (F(U_i, U_j, U_{\{\{i,j\}\}})) \quad (1)$$

where $(U_i)_{i \in N}$ and $(U_j)_{j \in N}$ are sequences of iid Uniform[0,1] random variables and $(U_{\{\{i,j\}\}})_{i,j \in N}$ is an array of the same. The array accounts for randomness in the row-column interactions. An example of a jointly exchangeable matrix is the adjacency matrix for an undirected, unweighted graph without self loops. This allows us to ignore the diagonal variables $U_{\{\{i,j\}\}}$. For fixed values of U_i and U_j , $F(U_i, U_j, \cdot)$ is defined on $[0,1]$.

A matrix is separately exchangeable iff there is a random measurable function $F : [0, 1]^3 \rightarrow X$ such that

$$(X_{ij}) := (F(U_i^{row}, U_j^{col}, U_{ij})) \quad (2)$$

where $(U_i^{row})_{i \in N}$ and $(U_j^{col})_{j \in N}$ are sequences of iid Uniform[0,1] random variables and $(U_{i,j})_{i,j \in N}$ is an array of the same. The only difference between this case and the jointly exchangeable one is the index structure of the uniform random variables. An example of this case is collaborative filtering, e.g. for users and movies. Seeing any particular realization of the (movie ratings) matrix does not depend on the ordering of the units (users and movies).

Summary of “Bayesian Models of Graphs, Arrays
and Other Exchangeable Random Structures”

Matt Dickenson mcd31
STA571/CS590.01

Due: 24 March, 2014

The theorems developed in this paper show that any Bayesian model on an exchangeable array can be specified with a prior on functions. These random functions often come in one of two types. The first type is random piecewise-constant functions, with priors such as the Chinese Restaurant Process and the Indian Buffet Process. Many clustering applications, including collaborative filtering, fall into this category. The second type includes random continuous functions, with priors such as Gaussian processes.

Orbanz and Roy show that the theory of exchangeable arrays can be extended to higher dimensions, to include data structures other than sequences, matrices, and graphs. However, exchangeable random structures are never sparse: an exchangeable random graph is either dense or empty. Network data often exhibit sparse properties (e.g. power law phenomena, or small world networks). This presents a challenge for Bayesian modeling of random networks.