Partial Order Models for Episcopal Social Status in 12th Century England

Geoff K Nicholls[1,2] (corresponding) and Alexis Muir Watt[2]

[1] nicholls@stats.ox.ac.uk
[2] Statistics Department, University of Oxford, 1 South Parks Road, Oxford OX13TG, UK

Abstract: Our data are lists of bishops signed in the 12th Century, in an order which respects the relative importance of the individual bishops. We model the underlying social order as a partial order, and the list data as a random complete order which respects this underlying partial order. We give static and dynamical models for the partial order. We summarize the posterior distribution using MCMC samples and a particle filter. We fit the models and find evidence for significant order, and for significant change in the order, over time.

Keywords: Social Hierarchy, Partial Orders, Bayesian

1 Data and Questions

Witness lists are ordered lists of the signatories to historical legal documents called acta. We have a large collection of "Royal Acta" from twelfth century England (these were provided by Dr David Johnson of St Peter's College, University of Oxford and Dr Nicholas E Karn, History, School of Humanities, University of Southampton). Witnesses generally signed in order of importance. The different social classes signed in groups in an order which is very obvious. What order relations existed between the bishops who appear in the lists? How did they change? Changes in this hierarchy reflect political events of the time. In the period 1070AD to 1150AD there are $m = 511$ lists with two or more bishops. The time at which a list was signed is known to within an interval (mean interval length 4 years, 90% less than 11 years). Approximately one half of the lists have length just two bishops, and the mean length is 3.5. Each bishop is given a numerical index. For $i = 1, 2, ..., m$, let $y_i = (y_{i,1}, y_{i,2}, ..., y_{i,n_i}), i = 1, 2, ..., m$ give the ordered set of indices of the n_i bishops who witnessed the ith list.

2 Models and Inference

We represent the unknown true order relation in the data as a partial order, that is, as a transitively closed directed acyclic graph $h = h[1 : n]$ with n

nodes, one node for each bishop in the analysis. We can think of h as a binary matrix with $h_{a,b} = 1$ (or 0) if bishop a ranks above b (or not). The model reflects a rigid social hierarchy, which is respected by all, but subject to occasional upheaval. Any particular witness list y_i is modeled as a random total order (a linear extension) respecting the suborder $h[o_i]$ for the bishops who attended that signing. This observation model arises if the signing order is a snapshot of a rapidly evolving total order on the individuals present. However, individuals may jump the queue. Before the j'th person signs in the ith list, there are $n_i - j + 1$ individuals remaining. With probability p the next to sign is chosen at random, ignoring any order constraints, and otherwise, the next person is the first person in a random linear extension of the suborder for the remaining individuals. Let $C(h)$ be the number of total orders consistent with partial order h, and let $C_i(h) = C(h[-i])$ be the number of linear extensions headed by bishop i. The likelihood $L(h, p; y_i) = \Pr(Y_i = y_i | H = h, O = o_i, P = p)$ for partial order h is

$$L(h, p; y_i) = \prod_{j=1}^{n_i - 1} \left(\frac{p}{n_i - j + 1} + (1-p) \frac{C_{y_j}(h[y_{j:n_i}])}{C(h[y_{j:n_i}])} \right)$$

so that $L(h, p; y_i) = 1/C(h[y_i])$ at $p = 0$. We can compute the count $C[h]$ quickly for partial orders on up to about $n = 15$ bishops. There were in the period of interest just over twenty bishops at any given time, just a subset of whom are active, so our algorithms are just adequate. There is recent work in Beerenwinkel *et al.* (2007) on maximum likelihood partial orders for conjunctive Bayesian networks, applications of Bayesian inference for 'bucket' orders can be found in Mannila (2008), and on Bayesian inference for generalized Bradley-Terry models in Caron *et al.* (2010). However, we know of no well-developed Bayesian framework for partial orders.

We describe a family of prior distributions for partial orders. These prior models for partial orders are derived from k-dimensional random orders, reviewed in Brightwell (1993). They are marginally consistent for suborders. The prior probability for a suborder is the marginal probability for that order in the prior for any superset of its nodes. Latent variables $Z = (Z_1, Z_2, ...Z_n)$ determine the partial order h on n bishops. The i'th bishop has K real-valued traits $Z_i = (Z_{i,1}, ..., Z_{i,K})$. These traits are not physical, but act as measures of status. Bishop a beats Bishop b ($h_{a,b} = 1 = 1 - h_{b,a}$ if $Z_{a,j} > Z_{b,j}$ for all $j = 1, 2, ..., K$ so that $h = h(Z)$. If the variables overlap then $h_{a,b} = h_{b,a} = 0$. Prior elicitation informed the partial order depth, so we have parameterized the prior to control depth by correlating the latent variables for a given Bishop. Let $Z_i \sim MVN(0, \Sigma)$ with $\Sigma_{i,j} = \rho$ for $i \neq j$ and $\Sigma_{i,i} = 1$. The hyperprior for $R = \rho$ is $R \sim \text{Beta}(1, 1/6)$. The joint latent variable prior, $f(z, \rho)$ say, gives a prior on partial orders (through $h = h(Z)$) which is roughly uniform on depth.

We have extended $h(Z), R$ to a process $h(Z(\tau)), R(\tau)$ in time. At the event times $\phi = (\phi_1, \phi_2, ...)$ of a Poisson process (the catastrophe process) of rate λ_C, there is a change point where $(Z(\phi_j), R(\phi_j)) \sim f$ independent of all history. At the event times $\psi = (\psi_1, \psi_2, ...)$ of a Poisson process (the singleton process) of rate λ_S the latent variables of a single Bishop $i \sim U\{1, 2, ..., n\}$ are (independently) renewed $Z_i(\psi_j) \sim MVN(0, \Sigma)$ at fixed R. This process has equilibrium $f(z, \rho)$.

We fit the static model to m witness lists from short intervals of time (it does not allow for evolution in the order). This analysis treats the uncertain list-dates as fixed (some are dated, and the uncertainty is often small). We use MCMC to simulate the posterior distribution $\pi(z, p, \rho|y) \propto L(h(z), p; y)f(z, r)$. The prior for p is uniform in $[0, 1]$. We use a hybrid MCMC/particle filtering approach as in Andrieu *et al.* (2010) to simulate the posterior distribution for the dynamical model. Let $t = (t_1, t_2, ..., t_m)$ parameterize the unknown true dates associated with the m lists, and let $t_{[i]}$ be the i'th date in an ordered list of the dates. We carry out MCMC for t, λ_S, λ_C. We estimate $p(y|t, \lambda_S, \lambda_C)$ using paths from a particle filter. The filter integrates the $Z(\tau), R(\tau)$ process using a discrete time HMM with hidden states $(Z(t_{[i]}), R(t_{[i]}), p)$ and emitted states $y_{[i]}$ (index ordered on t), so the HMM states are maintained at the m list times only.

3 Results and Conclusions

We present results for both the static and dynamical models. We illustrate the static model using lists taken from two windows of time. The lists are given below. What partial orders on status constrain the lists at each time? Do the orders differ from one window to the next?

Witness lists 1119-1121

[1119]	5	6	4	7		
[1120]	3	4				
[1121]	1	2				
[1121]	10	1	2	5	6	8
[1121]	1	10	2	5	6	9
[1121]	1	2				
[1121]	10	1	2			

Witness lists 1127-1129

[1127]	9	10				
[1127]	2	9	10			
[1127]	2	1	6	5	8	10
[1127]	2	6				
[1127]	2	9	6	10		
[1127]	2	9	6	10		
[1129]	7	10				
[1129]	6	7	4	10		
[1129]	3	4	2			

We make a separate static Bayesian analysis for each window. Figure 1 displays a graphical summary of the posterior distribution on partial orders in each time window. The posterior probability for each edge is estimated via MCMC, and thresholded at one half. We illustrate the dynamical model on the 1119-21 data, conditioning on $\lambda_S = 1$ and $\lambda_C = 0.1$, just to show consistency. The consensus order for the year 1120 is show in Figure 1. It agrees well with the adjacent result for the corresponding static analysis.

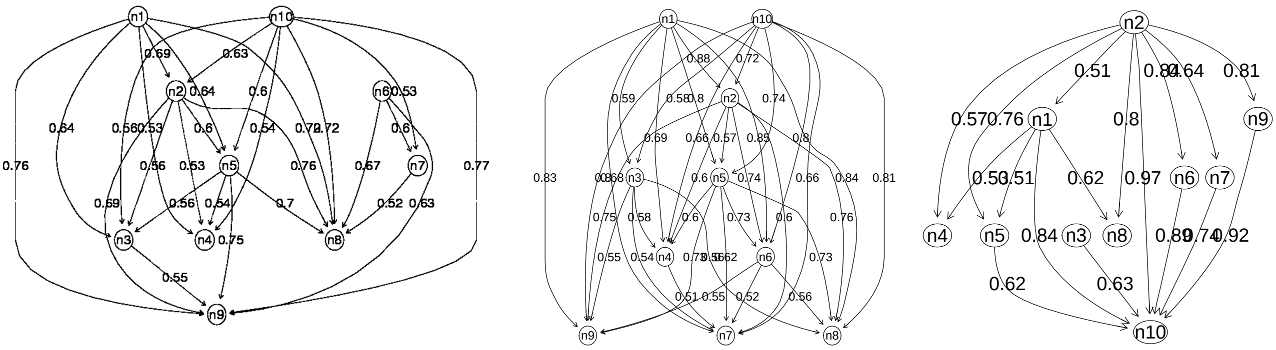

FIGURE 1. Consensus partial orders show marginal posterior support for each directed edge. Edge labels are marginal posterior probabilities. (Left) Dynamical model/MCMC-Particle filter, 1120. (Mid) Static model/MCMC 1119-21 (Right) 1127-29

We see from Figure 1 that there is evidence for significant order (edges supported with high marginal posterior probability). There is evidence for change. The postholder of the position of Bishop of London changes from the left to right graph, and node 10 (Bishop of London) correspondingly moves from the top to the bottom of the figure with some strongly supported edges changing direction. The probability for a catastrophe in the short interval 1119-21 was low (10%). The linear extensions in 1119-21 are well explained by the higher rate singleton change process.

Acknowledgments: Thanks to Prof Bernard Silverman, Dr David Johnson and Dr Nicholas E Karn for their help.

References

Andrieu, C., Doucet, A., and Holenstein, R. (2010) Particle Markov chain Monte Carlo methods. J. Roy. Statist. Soc. B, **72**, 269-342.

Beerenwinkel, N., Eriksson, N., and Sturmfels, B. (2007) Conjunctive Bayesian networks. *Bernoulli*, **13**, 893-909.

Brightwell, G. (1993) Models of random partial orders. In: *Surveys in Combinatorics* 5383, ed. K. Walker, Cambridge Univ. Press

Caron, F., and Doucet, A. (2010). Efficient Bayesian Inference for Generalized Bradley-Terry Model. *J. Comp. Graphical Stat* to appear.

Mannila, H. (2008) Finding Total and Partial Orders from Data for Seriation. *Lecture Notes in Comp. Sci.*, **5255**, 16-25.