Bayesian inference for partial orders

Kate Lee

University of Auckland

This is a joint work with

Geoff K. Nicholls[1], Rukuang Huang[1], Alexis Muir-Watt[3], Jessie Jiang[1]

1 University of Oxford, 2 University of Southampton, 3 Private researcher

Biometrics in the Bay of Islands 2023

Regional Rankings	North America	Latin America	Western Europe	Eastern Europe	China	India	APAC Ex China, India	Middle East and Africa
1st	Apple	Apple	Apple	Apple	Apple	Apple	Apple	Apple
2nd	SAMSUNG	SAMSUNG	SAMSUNG	SAMSUNG	HUAWEI	OnePlus	SAMSUNG	SAMSUNG
3rd	Google	Motorola	Xiaomi	Xiaomi	oppo	SAMSUNG	Xiaomi	ASUS
4th	OnePlus	Xiaomi	oppo	ASUS	vivo	vivo	vivo	Xiaomi
5th	Motorola	HUAWEI	OnePlus	oppo	Xiaomi	oppo	oppo	HUAWEI

New entrant to top 5 in 2021 as compared to 2020

Models for rank list data

$x = (x_1, \ldots, x_n)$ is a permutation of $[n] = \{1, 2, \ldots, n\}$, and think of $x_1 \to x_2 \to \cdots \to x_n$ as a simple DAG representing a "ranking".

Ranking models for the central total order[1] to the list (data). i.e., Mallows model, Plackett-Luce model

We allow the centering order to be a partial order. Mannila et al. [1, 2, 3] treats subclasses (bucket orders & VSP's) of PO's. Beerenwinkel at al. [4, 5] includes Bayes/MCMC for PO's.

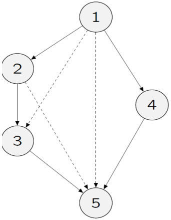
Similar but different to network models.

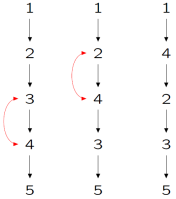
Bayesian inference for partial order.

[1] totally ordered set

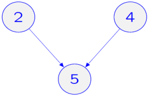
Queues and partial orders (PO) [6, 7]

Queue of n actors $[n] = \{1, \ldots, n\}$ constrained by PO $h \in \mathcal{H}_{[n]}$.



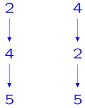


Sub-order $h[o]$



$o = \{2, 4, 5\}$

LE's $\mathcal{L}(h[o])$



Partial order h Linear extensions $\mathcal{L}(h)$

If just $o \subseteq [n]$ are queuing then the constraining suborder is $h[o]$.
For $X = (X_1, \ldots, X_n)$ a random queue, $\Pr(X = x|h) = |\mathcal{L}(h)|^{-1}$.

For $i \in [N]$, actors $o_i = \{o_{i,1}, \ldots, o_{i,n_i}\}$, list data $y_i = (y_{i,1}, \ldots, y_{i,n_i})$,

$$p(y|h) = \prod_{i=1}^{N} |\mathcal{L}(h[o_i])|^{-1}.$$

Noisy lists: Realised lists not perfect LE's.

$$p(x|h) = p(X_1 = x_1|h) \times \Pr(X_2 = x_2|h_{-x_1}) \ldots \Pr(X_n = x_n|h_{-(1:n-1)})$$

$$= \frac{|\mathcal{L}(h[x_{2:n}])|}{|\mathcal{L}(h[x_{1:n}])|} \times \frac{|\mathcal{L}(h[x_{3:n}])|}{|\mathcal{L}(h[x_{2:n}])|} \times \cdots \times \frac{|\mathcal{L}(h[x_n])|}{|\mathcal{L}(h[x_{n-1:n}])|}$$

Queue jumping (QJ) up queue: with p select next one at random,

$$p(x|h,p) = \prod_{j=1}^{n-1} \left(\frac{p}{n-j+1} + (1-p) \frac{|\mathcal{L}(h[x_{j+1:n}])|}{|\mathcal{L}(h[x_{j:n}])|} \right)$$

Posterior for h

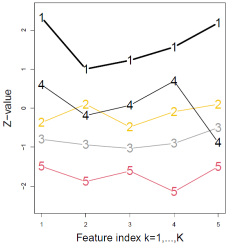
Give a prior over $h \in \mathcal{H}_{[n]}$ then $\pi_{[n]}(h,p|y) \propto \pi_{[n]}(h)\,\pi(p)\,p(y|h,p)$.

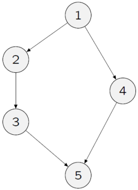
Marginally Consistent prior for h [9]

Fix $K \geq 1$ and $0 < \rho < 1$. Latent variables for actor $j = 1, \ldots n$,

$$(Z_{j,1}, \ldots, Z_{j,K}) \sim N(0_K, \Sigma^{(\rho)})$$

with $\Sigma^{(\rho)}$ a $K \times K$ covariance with $\Sigma^{(\rho)}_{k,k} = 1$ and $\Sigma^{(\rho)}_{k,k'} = \rho$, $k \neq k'$.





$h(Z) = \{\langle i,j \rangle : Z_{i,k} > Z_{j,k}, \text{ all } k\}$

- $\Pr(h(Z) = h|\rho)$ is MC.
- Prior $\pi_{[n]}(h|\rho)$ has parameter ρ controls depth $d(h)$ distribution.
- If $K \geq \lceil n/2 \rceil$ then $\pi_{[n]}(h|\rho) > 0$ for all $h \in \mathcal{H}_{[n]}$.[8]

Covariates[10]

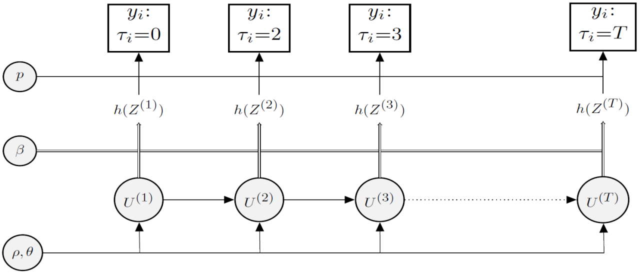
In list y_i actor j has covariates $x_{i,j} \in R^p$ so take

$$(U_{j,1},\ldots,U_{j,K}) \sim N(0_K,\Sigma^{(\rho)})$$

and $Z^{(i)}_{j,:} = U_{j,:} + x^T_{i,j}\beta 1_K$. Effects β push path $Z_{j,:}$ up or down.

Timeseries[10]

List data y_i come with time stamp $\tau_i \in \{1,\ldots,T\}$. Extend to HMM using $U = (U^{(t)}\ t = 1,\ldots,T)$ and $U \sim VAR_{\rho,\theta}(1)$.



Posterior (with $Z = Z(U,\beta)$)

$$\pi(U,\beta,\theta,\rho,p|y) \propto p(y|h(Z),p)\pi(U|\theta,\rho)\pi(\beta,\theta,\rho,p).$$

Case study - Witness list

Witness lists of legal documents from 20 Anglo-Norman dioceses (yr 1080-1155)[11]

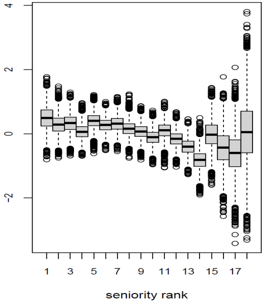
12th Century Acts: witness lists example (1127)

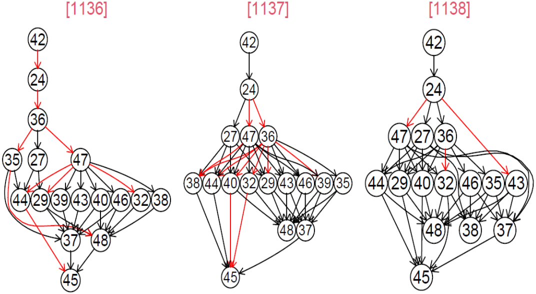
```
[1] William, Archbishop of Canterbury
[2] Roger, Bishop of Salisbury
[3] William, Giffard, bishop of Winchester, 1100-1129
[4] Bernard, Bishop of St David's
[5] William, de Warelwast, bishop of Exeter
[6] Urban, bishop of Llandaff
[7] Geoffrey, Rufus, Bishop of Durham
[8] Robert, de Sigillo, Bishop of London
   (Richard, de Belmeis I, bishop of London, 1108-1126)
[9] Miles, of Gloucester, earl of Hereford 1141-1143
[10] Henry, de Port
[11] Walter, de Amfreville
[12] William, de Folis
[13] Roger, de Port
[14] William, de Port
```

Historians want evolving status hierarchy of bishops: we have $N = 371$ lists over $T = 76$ years; estimate $h = (h^{(1)}, \ldots, h^{(T)})$.

Case study - Witness list

Some analysis result [10]



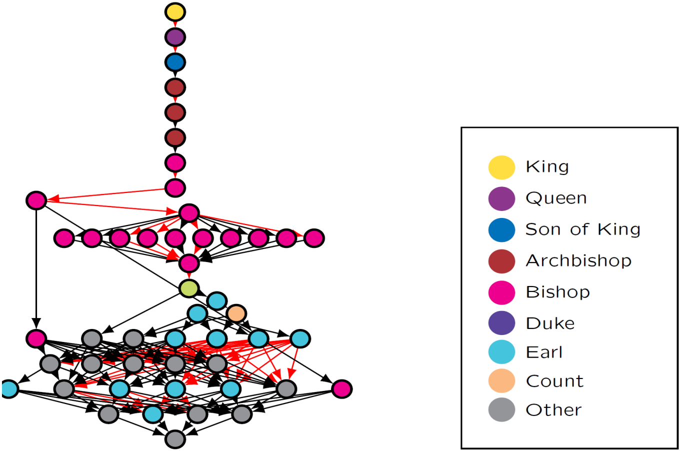


Seniority effect (β)

Consensus PO in 1136-1138
(significant (black) and strong (red) orders.)

Case study - Witness list

Social hierarchy of $n = 49$ witnesses for 1134-1138 [12]



VSP[2]/QJ-U model. Consensus order. Significant/strong order relations are indicated by black/red edges respectively.

[2] Vertex-series-parallel partial orders (VSP)

Conclusion

- PO offers the largest class of rank relationships, including the total order, bucket order, VSP.
- Marginal consistent prior and properties.
- Temporal PO model with covariate effects for noise-free/noisy data.
- Scalable VSP-PO for large PO's.

Future work - Less restrictive noise model, hierarchical PO, clustering, computation.

References I

[1] H. Mannila and C. Meek, “Global partial orders from sequential data,” in *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD ’00, Boston, Massachusetts, USA: Association for Computing Machinery, 2000, pp. 161–168, ISBN: 1581132336. DOI: 10.1145/347090.347122. [Online]. Available: https://doi.org/10.1145/347090.347122.

[2] A. Gionis, H. Mannila, K. Puolamäki, and A. Ukkonen, “Algorithms for discovering bucket orders from data,” in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2006, pp. 561–566.

[3] H. Mannila, “Finding total and partial orders from data for seriation,” in *Discovery Science*, J.-F. Boucault, Ed., ser. LNAI, vol. 5255, Berlin Heidelberg: Springer-Verlag, 2008, pp. 16–25.

[4] N. Beerenwinkel, N. Eriksson, and B. Sturmfels, “Conjunctive bayesian networks,” *Bernoulli*, vol. 13, no. 4, pp. 893–909, 2007.

[5] T. Sakoparnig and N. Beerenwinkel, “Efficient sampling for Bayesian inference of conjunctive Bayesian networks,” *Bioinformatics*, vol. 28, no. 18, pp. 2318–2324, Jul. 2012, ISSN: 1367-4803. DOI: 10.1093/bioinformatics/bts433. eprint: https://academic.oup.com/bioinformatics/article-pdf/28/18/2318/698477/bts433.pdf. [Online]. Available: https://doi.org/10.1093/bioinformatics/bts433.

[6] G. Brightwell and P. Winkler, “Counting linear extensions,” *Order*, vol. 8, no. 3, pp. 225–242, 1991.

[7] G. Brightwell, “Surveys in combinatorics,” in (London Mathematical Society Lecture Note Series), K. Walker, Ed., London Mathematical Society Lecture Note Series. Cambridge Univeristy Press, 1993, vol. 187, ch. Models of random partial orders, pp. 53–83.

[8] T. Hiraguchi, “On the dimension of partially ordered sets,” *The science reports of the Kanazawa University*, vol. 1, pp. 77–94, 2 1951.

[9] P. Winkler, “Random orders,” *Order*, vol. 1, no. 4, pp. 317–331, 1985.

References II

[10] G. K. Nicholls, J. E. Lee, N. Karn, D. Johnson, R. Huang, and A. Muir-Watt, *Bayesian inference for partial orders from random linear extensions: Power relations from 12th century royal acta*, 2022. arXiv: 2212.05524 [stat.ME].

[11] R. Sharpe, D. Carpenter, H. Doherty, M. Hagger, and N. Karn, "The charters of William II and Henry I," Online: Last accessed 27 October 2022, 2014.

[12] C. Jiang, G. K. Nicholls, and J. E. Lee, *Bayesian inference for vertex-series-parallel partial orders*, (To appear UAI2023), 2023.

Thank you.

Any question?