9th International Conference on Computational Social Science IC^{2}S^2
July 17-20, 2023, Copenhagen, Denmark

Escalatory Dynamics in International Relations

Keywords: Latent Variable Models, International Relations, Bayesian Time Series, Ordinal Measurement, Political Event Data

Introduction: Observed Country Interactions and Latent Relationships

Being part of a global diplomatic community, countries shape and develop their relationships with other countries. The state of a relationship is typically only evidenced through observed country interactions and can thus be considered latent. Different political interactions have been categorized and ranked on a conflict–cooperation intensity scale, ranging from "fight" to "provide aid" (Goldstein, 1992). It is natural to assume that conflictual actions like "fight" are more likely observed if countries are in an "enemy" than in an "ally" state (Schrodt, 2006; Stoehr et al., 2022). This means, the ordering of action types should be reflected in an ordering of latent relationship states. Another long-standing assumption is that relationships escalate and de-escalate through a ladder of escalatory stages (Davis and Stan, 1984; Schrodt, 2006; Randahl and Vegelius, 2022). We expect a smooth transitioning from enemies to allies and vice versa.

Model: Ordinal Dynamical System

Emission and Transition Matrices. There exist various statistical models of dyadic country interactions and their latent relationships (O'Connor et al., 2013; Schein et al., 2016). Many models rely on the notion of a state-to-action emission matrix and a state-to-state transition matrix such as the prominent Hidden Markov Model (HMM). The emission matrix (Fig. 1 (left)) describes how likely we observe (emit) action type a when being it the latent relationship state k. The transition matrix (Fig. 1 (right)) models the probability of transitioning from the latent state k to all other states. The rows of these stochastic matrices describe discrete distributions that are typically sampled independently from a Dirichlet distribution. This leads to the states and actions being categorical, without particular ordering, a problem known as label switching. We refer to this kind of stochastic matrices as *Standard Matrix Dirichlet* (SMD).

Ordered Matrix Dirichlet (OMD). So how can we reflect the assumption of ordered action types and ordered latent states in our model? We propose the *Ordered Matrix Dirichlet* (OMD) which ensures that the probability mass of a stochastic matrix is, on average, shifted to the right

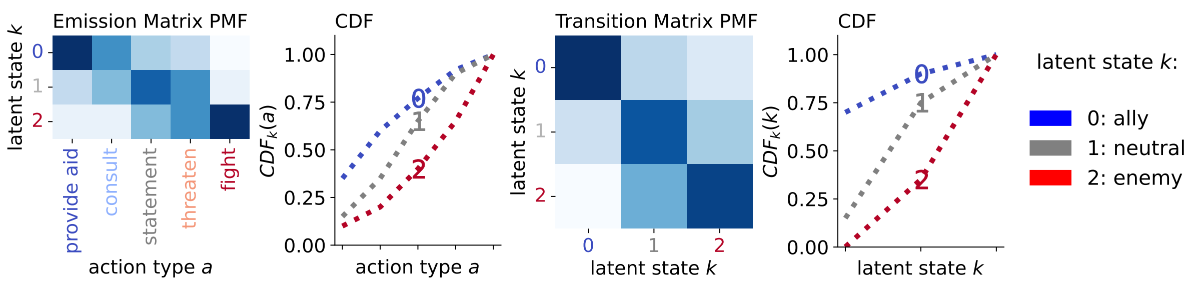

Figure 1: Observed action types a are naturally ordered from cooperative to conflictual. Emission matrix: we expect cooperative actions (e.g. "provide aid") to be associated with more cooperative latent states k (e.g. "ally"). Transition matrix: we expect latent states to transition to adjacent states.

as we move down rows which can be observed in Fig. 1. More formally, the distribution of row k has First-order Stochastic Dominance (FSD) over row $k+1$ if its cumulative distribution function $\text{CDF}_k(a)$ at each point a is higher than or equal to $\text{CDF}_{k+1}(a)$. We bake this FSD constraint into a stick-breaking algorithm (Gelman et al., 2013, p. 585) for sampling from a Dirichlet random variable. The algorithm, provided in Algorithm 1, returns an OMD variate, a stochastic matrix with uniquely ordered action types and latent states.

Dynamic Poisson Tucker Model (DPT). Equipped with the OMD, we can now develop an highly expressive Dynamic Poisson Tucker (DPT) model tailored to political event data. We consider ICEWS, one of the largest collections of country-level "who-did-what-to-whom-at-what-time" quadruplets. The data can be organized into a 4-mode count tensor $Y \in \mathbb{Z}_0^{I \times J \times A \times T}$ where each entry describes the number of times country i took action a towards country j during time step t. Importantly, the 20 action types (listed on the left in Fig. 3) are ordered based on the Goldstein Scale (Goldstein, 1992). Our DPT model combines motifs behind Tucker decomposition and Poisson–Gamma Dynamical Systems Schein et al. (2016):

$$y^{(t)}_{i \xrightarrow{a} j} \sim \text{Pois}\left(\sum_{c_1=1}^{C}\sum_{c_2=1}^{C}\sum_{k=1}^{K} \psi_{c_1 i}\psi_{c_2 j} \underbrace{\phi_{ka}}_{\text{emission}} \lambda^{(t)}_{c_1 \xrightarrow{k} c_2}\right) \quad \lambda^{(t)}_{c_1 \xrightarrow{k} c_2} \sim \text{Gam}\left(\sum_{k_1=1}^{K} \lambda^{(t-1)}_{c_1 \xrightarrow{k_1} c_2} \underbrace{\pi_{kk_1}}_{\text{transition}}, \tau_0\right)$$

The emission matrix $\Phi \in [0,1]^{K \times A}$ and transition matrix $\Pi \in [0,1]^{K \times K}$ are modeled using the OMD. $\lambda^{(t)}_{c_1 \xrightarrow{k} c_2}$ forms the *core tensor* $\Lambda^{(t)} \in \mathbb{R}_+^{C \times C \times K}$ that models the rate at which countries in community c_1 take actions in state k towards countries in community c_2 during time step t The parameters $\psi_{c_1 i}$ and $\psi_{c_2 j}$ represent the rate at which the countries i / j act as source / target in community c_1 / c_2 and describe a community-country matrix $\Psi \in \mathbb{R}_+^{C \times V}$ (see Fig. 2 and Fig. 4).

Evaluation: Data Imputation and Forecasting

To evaluate our DPT model, we conduct imputation and forecasting tasks on held-out data. We find that the OMD does not significantly forfeit predictive fit while bringing the benefit of easily interpretable latent representations corresponding to country relationships. In forecasting and settings of data scarcity, the inductive bias of the OMD even improves performance. The OMD as a modeling motif has wide applicability to a broad family of dynamical systems with Dirichlet-sampled matrices and application domains exhibiting (de-)escalatory dynamics.

References

Paul K. Davis and Peter J. E. Stan. 1984. *Concepts and models of escalation: A report from the Rand Strategy Assessment Center*. Rand.

Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2013. *Bayesian data analysis*. Chapman and Hall/CRC.

Joshua Goldstein. 1992. A conflict-cooperation scale for WEIS events data. *The Journal of Conflict Resolution*, 36(2):369–385.

Brendan O'Connor, Brandon Stewart, and Noah Smith. 2013. Learning to extract international relations from political context. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 1094–1104.

David Randahl and Johan Vegelius. 2022. Predicting escalating and de-escalating violence in Africa using Markov models. *International Interactions*, pages 1–17.

Aaron Schein, Mingyuan Zhou, David M. Blei, and Hanna Wallach. 2016. Bayesian poisson Tucker decomposition for learning the structure of international relations. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48*, pages 2810–2819.

Philip A. Schrodt. 2006. Forecasting conflict in the Balkans using hidden Markov models. In *Programming for Peace*, pages 161–184. Springer Netherlands.

Niklas Stoehr, Lucas Torroba Hennigen, Josef Valvoda, Robert West, Ryan Cotterell, and Aaron Schein. 2022. An ordinal latent variable model of conflict intensity. In *arXiv*, volume 2210.03971.

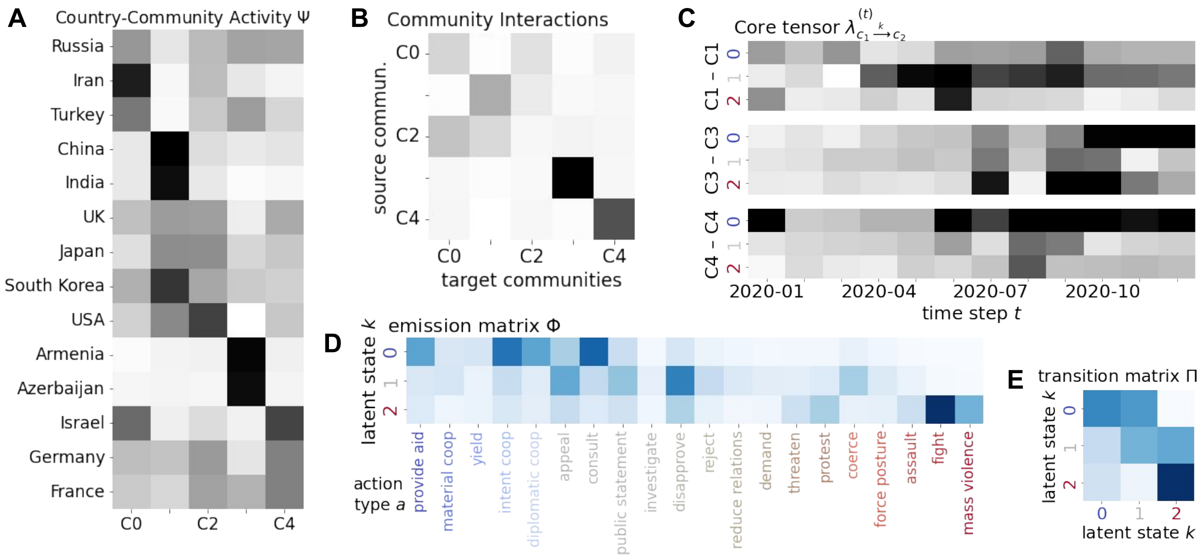

Figure 2: Posterior parameters of DPT model fitted to ICEWS subset of 2020. (A) The latent matrix Ψ indicates country-community activity. Armenia and Azerbaijan are standing out as they are involved in one community only; (B) Interactions between latent communities. We find that communities C3 and C4 predominantly interact between themselves. (C) Selected community interactions over time: C3-C3 are in conflict, while C4-C4 are mostly friendly; (D) Latent emission matrix Φ representing global state-to-action probabilities. Thanks to the ordering, we know that state $k = 2$ represents an "enemy" relationships; (E) Latent transition matrix Π show smooth transition probabilities.

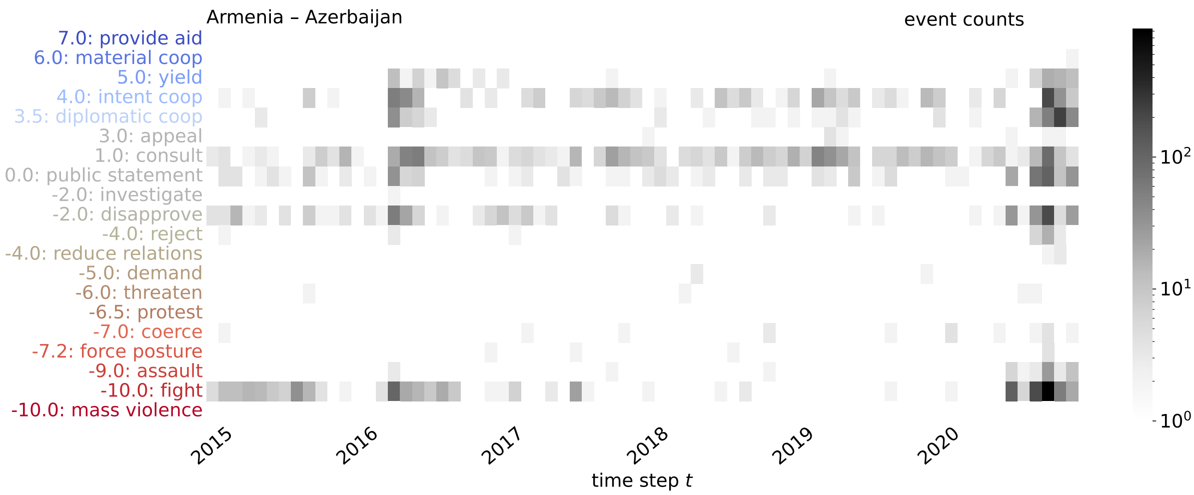

Figure 3: Subset of ICEWS event data showing how country interactions are changing smoothly over time. The full event tensor is of shape $I \times J \times A \times T$ and represents actions a exchanged between countries i and j at time step t. Here, we display Armenia and Azerbaijan, aggregated by months.

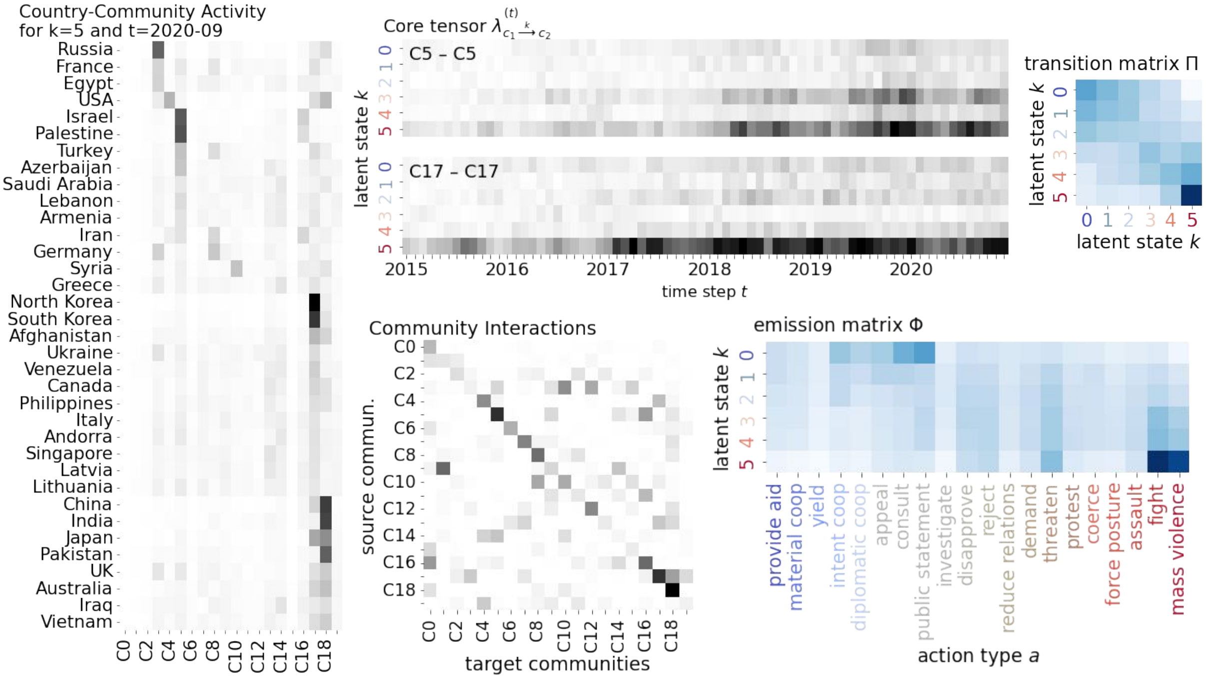

Figure 4: Posterior parameters of Dynamic Poisson Tucker Model, with $K = 6$ latent states and $C = 12$ latent communities, fitted to full temporal range (2000-2020) of ICEWS data. We find that the probability mass of the transition matrix is centered along the diagonal revealing step-wise (de-)escalatory dynamics. Particularly state 5 has high probability representing an "absorbing state" of conflict that is hard to escape. The country-community affiliation matrix Ψ provides no information on whether communities represent allies or enemies per se. To obtain this information, we interact the country-community matrix with the core tensor $\psi^{(\rightarrow)}_{c_1 i} \sum_{c2=1}^{C} \sum_{j=1}^{J} \psi^{(\leftarrow)}_{c_2 j} \lambda^{(t)}_{c_1 \xrightarrow{k} c_2}$ for specific choice of k and t.

Algorithm 1 Ordered Matrix Dirichlet

Input: alpha concentration prior $\alpha \in \mathbb{R}_+^A$
for k = 1, …, K **do**
  $\tilde{\phi}_{k1} \overset{\text{iid}}{\sim} \text{Beta}\left(\alpha_1, \sum_{a=2}^{A} \alpha_a\right)$
end for
$(\phi_{11}, \ldots, \phi_{K1}) \leftarrow \text{Sort}\Big((\tilde{\phi}_{11}, \ldots, \tilde{\phi}_{K1})\Big)$
for a = 2, …, A-1 **do**
  for k = 1, …, K **do**
    $\tilde{\beta}_{ka} \overset{\text{iid}}{\sim} \text{Beta}\left(\alpha_a, \sum_{a'=a+1}^{A} \alpha_{a'}\right)$
  end for
  $(\beta_{1a}, \ldots, \beta_{Ka}) \leftarrow \text{Sort}\Big((\tilde{\beta}_{1a}, \ldots, \tilde{\beta}_{Ka})\Big)$
  for k = 1, …, K **do**
    $\phi_{ka} \leftarrow \left(1 - \sum_{a'=1}^{a-1} \phi_{ka'}\right) \beta_{ka}$
  end for
end for
for k = 1, …, K **do**
  $\phi_{kA} \leftarrow 1 - \sum_{a'=1}^{A-1} \phi_{ka'}$
end for
Output: OMD variate $\Phi \in \mathbb{R}_+^{K \times A}$