A Unified Framework for Ideal Point Estimation

Max Goplerud
Harvard University

Overview

This paper creates a unified framework for ideal point estimation (`uIRT`) by using recent developments in Bayesian statistics (Polson et al. 2013; Linderman et al. 2015). The core model derived in this paper estimates ideal points from multinomial and ordinal data with arbitrarily many categories quickly and without need for variational (or other) approximations. Extensions such as 'dynamic' smoothing (Martin and Quinn 2002) are easily incorporated. As an empirical example, I focus in detail on the recent literature on abstentions in the UN General Assembly and show that the `uIRT` can implement two state-of-the-art models (Bailey et al. 2015; Rosas et al. 2015) in a fraction of the time, whilst also allowing them to be generalised to recover new and interesting results.

Developments in Ideal Point Estimation

Recent Developments

- Imai et al. (2016) – Use EM to make IRT models scale easily.
- New sophisticated models to bridge different types of data.

Remaining Limitations

- Proliferation of 'bespoke' models.
- No (Bayesian) model for multinomial ideal points
- A reliance on variational approximations for fast estimation.
- Fast estimation limited to particular types of models.

The Unified IRT Framework: `uIRT`

The Model

- Stick-breaking representation to create a unified framework.
- Pólya-Gamma data augmentation makes logistic links tractable.

Benefits

- Fast estimation of multinomial, ordinal (many-valued), and/or binary data.
- Permits scaling all types of survey questions in one model.
- Easily integrates 'bespoke' models into the same framework.
- Estimation via EM (without approximations) or MCMC.

Stick-Breaking Representation

Assume that $y \sim Multi(1, \boldsymbol{p})$ and that we have imposed some (arbitrary!) ordering on the outcomes from $k = 1, \cdots, K$. Define y^k as indicators for whether $y = k$.

$$\forall k \in \{1, \cdots, K-1\}, \quad y^k = Bin(C_k, \sigma^k); \quad C_k = \mathbb{I}(y > k); \quad \sigma^k = \frac{\Pr(y = k)}{\Pr(y \geq k)}$$

Any multinomial can be expressed by the above representation.

σ^k are 'stick-breaks': Given that y is *at least* k, what is the probability of k?

Actual probabilities of $y = k$ are recovered by:

$$\sum_{n=1}^{K} p^n = 1; \quad p^1 = \sigma^1; \quad p^k = \sigma^k \prod_{n=1}^{k-1} 1 - \sigma^n, \quad \forall k < K$$

The `uIRT` Model

I individuals; J votes with K_j outcomes. Impose some arbitrary ordering on the K_j outcomes. Define the probability of i's vote on j in terms of σ_{ij}^k:

$$\sigma_{ij}^k \equiv \Pr(y_{ij} = k |\, y_{ij} \geq k) = \frac{\exp(\psi_{ij}^k)}{1 + \exp(\psi_{ij}^k)}; \quad \psi_{ij}^k = \kappa_j^k + \beta_j^k x_i$$

- 'Stick-breaking' or 'continuation' logit (Linderman et al. 2015; Mare 1980)
- Two Parameter IRT model with κ_j^k and β_j^k

Model Specification:

- Ordinal: Set all β_j^k equal to some common value β_j.
- Multinomial: Allow β_j^k to be unconstrained.

Pólya-Gamma Augmentation

Polson et al. (2013), drawing on Biane et al. (2001), define a Pólya-Gamma random variable as an infinite sum of Gamma random variables:

$$\omega \sim PG(b, c) \equiv \omega = \frac{1}{2\pi^2} \sum_{n=1}^{\infty} \frac{Z_n}{(n - 1/2)^2 + c^2/(4\pi^2)}; \quad Z_n \sim^{i.i.d.} Gamma(b, 1)$$

They note that:

$$\frac{\exp(\psi)}{1 + \exp(\psi)} = \frac{1}{2} \exp(1/2\psi) \int_0^\infty \exp\left(-\omega/2 \cdot \psi^2\right) f(\omega) d\omega; \quad \omega \sim PG(1, 0)$$

We can augment Pólya-Gamma variables to a logistic link to get:

$$f(\psi, \omega) \propto \exp\left(1/2\psi - \omega/2\psi^2\right) f(\omega)$$

UN General Assembly Voting

Estimation

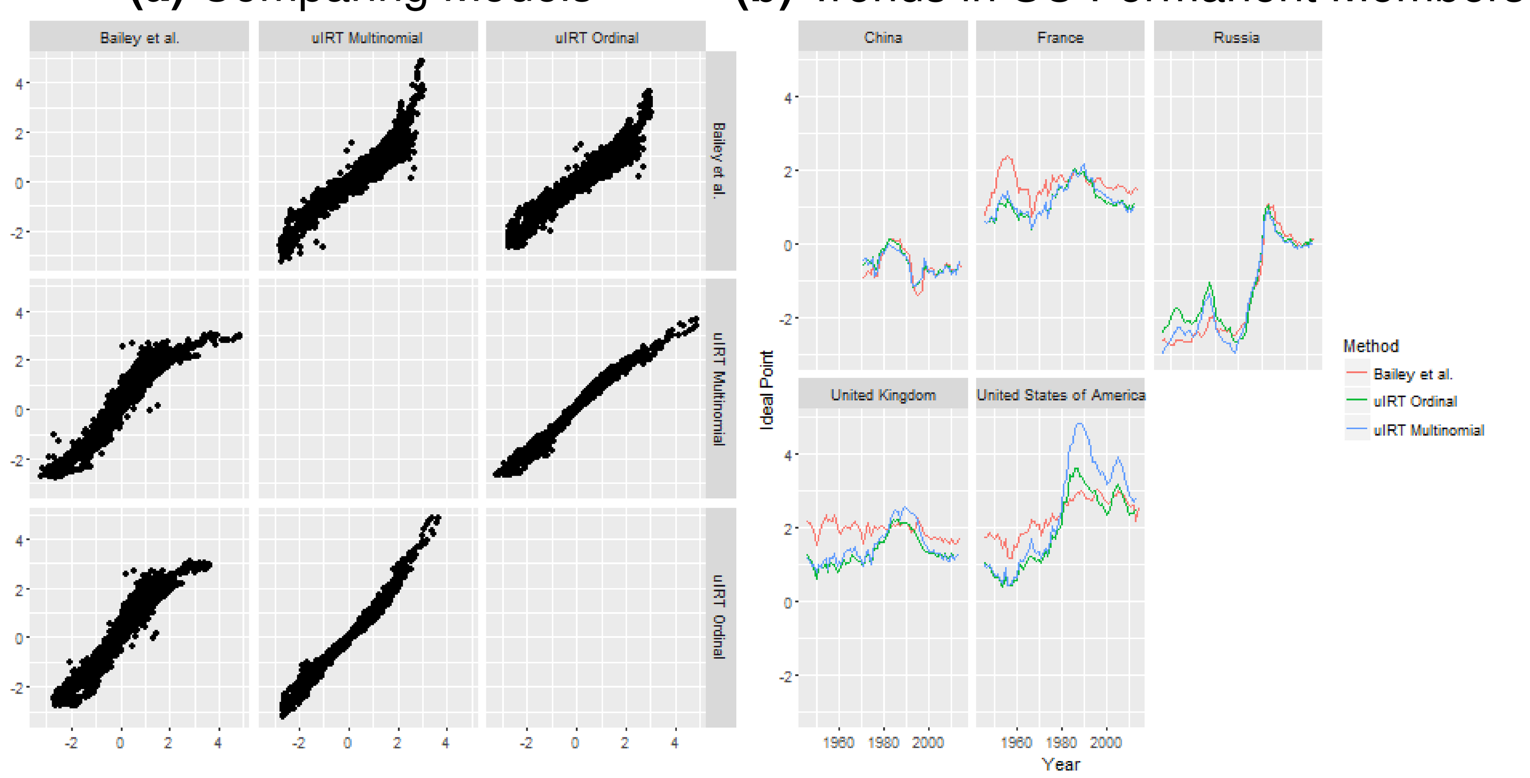
- Each σ_{ij}^k is augmented with an independent ω_{ij}^k
- Gives a quadratic complete data log-likelihood.
- Add priors for the parameters κ_j^k, β_j^k, x_i.
- Use an exact EM algorithm (no variational approximations needed!)
  - E-step: The augmented ω_{ij}^k are conditionally Pólya-Gamma: $PG(1, \psi_{ij}^k)$
  - M-step: The complete data loglikelihood is quadratic; simple updates.
- MCMC Approach:
  - Gibbs Sampler with normal and Pólya-Gamma updates.
  - Much slower but recovers full posterior.

Ordinal and Multinomial Dynamic Ideal Points

- Bailey et al. (2015) analyse UN voting data as three-leveled ordinal (yes, abstain, no) with Martin-Quinn 'dynamic smoothing'.
- MCMC methods take a long time (many hours).
- Assumes ordinality for all items.
- Not possible to implement in `emIRT` (Imai et al. 2016)
- Existing EM approaches would required complicated variational updates.

`uIRT` runs in minutes, uses exact EM algorithm, and relaxes ordinality.

(a) Comparing Models **(b)** Trends in SC Permanent Members



Prepared for the Sixth Annual Government Department Poster Session