# Identifiability of the Trans-dimensional Bradley-Terry Model*

## 1 Identification of the Likelihood

Consider the sampling distribution for a response $\pi$ in a trans-dimensional Bradley-Terry (TDBT) model given by

$$p(Y \mid w, F) = \prod_{i<j} \frac{\exp\{w^\top (f_i - f_j)\}}{1 + \exp\{w^\top (f_i - f_j)\}}. \tag{1}$$

where $Y = \{y_{12}, \ldots, y_{N-1,N}\}$, $w = (w_1, \ldots, w_N)^\top$ and $F = \{f_1, \ldots, f_N\}$. When $K = 1$ and $w = 1$, the model (1) reduces to the conventional Bradley-Terry model (Bradley and Terry, 1952). The multiple unobservable parameters in the equation (1) induce the problem of identification in the general data model $\prod_{i<j} \pi_{ij}$.

Identification can be dealt with by specifying additional assumptions and constraints that render the model estimable. In general, we say that a statistical model is identifiable if distinct parameter values generate different joint probability distributions. The likelihood of a TDBT model can be underidentified for the following five reasons: location invariance, scale invariance, reflective invariance, label switching, and rotation invariance.

### Location and Scale Invariance

For illustration, consider the TDBT model with the linear predictor $z_{ij} = w(f_i - f_j)$. If we add a constant vector $\delta = (\delta_1, \ldots, \delta_K)^\top$ to both $f_i$ and $f_j$, the likelihood still remains

*April 24, 2025

the same. In addition, if we divide $w$ by a constant vector $c = (c_1, \ldots, c_K)^\top$ ($c_k \in \mathbb{R}\backslash\{0\}$) and multiply both $f_i$ and $f_j$ by the same constant, the likelihood remains unchanged:

$$p(Y \mid w, F) = p\Big(Y \mid \tilde{w}, \tilde{F}\Big)$$

where $\tilde{w} = \left(\frac{w_1}{c_1}, \ldots, \frac{w_K}{c_K}\right)^\top$, $\tilde{F} = \{\tilde{f}_1, \ldots, \tilde{f}_N\}$ and $\tilde{f}_n = (c_1(f_{n1} + \delta_1), \ldots, c_K(f_{nK} + \delta_K))^\top$ for $n = 1, \ldots, N$; $k = 1, \ldots, K$. To address this issue, one method is to assign a very strong prior to $f_i$ (e.g., Bafumi et al., 2005), a fixed population distribution that can determine the location and scale of the parameter space for the entire likelihood function $p(Y \mid w, F)$. The constraint is simply

$$f_i \sim \mathcal{N}_K(\mathbf{0}, \boldsymbol{I}) \quad \text{for } i = 1, 2, \ldots, N.$$

## Reflective Invariance

Even after eliminating additive and multiplicative aliasing, the likelihood of the TDBT model remains invariant when every element of the $k$-th dimension is multiplied by $-1$ simultaneously in the weight vector $w$ and in all worth vectors $f_i$, so there exists $2^K$ equivalent parameterizations. A common remedy to remove this ambiguity is to impose a sign-identification rule, for example, to restrict the support of $w$ to the positive orthant as follows:

$$w_k > 0 \quad \text{for } k = 1, \ldots, K.$$

either by imposing $w$ a truncated prior or by re-parameterising $w$ (e.g., $\log w_k \in \mathbb{R}$). Since the sign of each element of $w$ can no longer change, the sign of $f_i$ is identifiable and still capture both positive and negative latent traits. Therefore, this positivity constraint resolves reflective invariance without affecting model flexibility.

## Label Switching

In multidimensional setting, another issue arises. For $K$ dimensions, there are $K!$ permutations of parameter labels that yield the same likelihood. Combined with sign reversals, the likelihood has $2^K \times K!$ symmetric modes.

## Rotation Invariance

## References

Bafumi, J., A. Gelman, D. K. Park, and N. Kaplan (2005). Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation. *Political Analysis 13*(2), 171–187.

Bradley, R. A. and M. E. Terry (1952). Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika 39*(3/4), 324–345.