Identifiability of the Trans-dimensional Bradley-Terry Model*

1 Identification of the Likelihood

Consider the sampling distribution for a response π 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, rotation invariance, and label switching.

Location and Scale Invariance

For illustration, consider the TDBT model with the linear predictor $z_{ij} = w^\top (f_i - f_j)$. Adding a constant vector $\delta = (\delta_1, \ldots, \delta_K)^\top$ to both f_i and f_j does not change the like-

*April 26, 2025

lihood. Furthermore, if each element of w is divided by the corresponding constant in $c = (c_1, \dots, c_K)^\top$ with $c_k \in \mathbb{R}\backslash\{0\}$, and each element of f_i and f_j is multiplied by the corresponding constant in c, the likelihood also remains unchanged:

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

where $\tilde{w} = \left(\frac{w_1}{c_1}, \dots, \frac{w_K}{c_K}\right)^\top$, $\tilde{F} = \{\tilde{f}_1, \dots, \tilde{f}_N\}$ and $\tilde{f}_n = (c_1(f_{n1} + \delta_1), \dots, c_K(f_{nK} + \delta_K))^\top$ for $n = 1, \dots, N$; $k = 1, \dots, K$.

Reflective Invariance

Even after additive and multiplicative invariance have been removed, the likelihood of the TDBT model still exhibits a sign ambiguity. Specifically, multiplying the k-th coordinate of w and the corresponding coordinate of every f_i by -1 leaves the likelihood unchanged, so there are 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:

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

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.

Rotation Invariance

When $K \geq 2$, the linear predictor $z_{ij} = w^\top(f_i - f_j)$ is also invariant under orthogonal rotations of the latent space. Specifically, for any orthogonal matrix $R \in \mathbb{R}^{K\times K}$ satisfying $R^\top R = RR^\top = I_K$, the transformation $\bar{f}_i = Rf_i \quad \text{and} \quad \bar{w} = Rw$ leaves the likelihood unchanged:

$$\bar{w}^\top(\bar{f}_i - \bar{f}_j) = (Rw)^\top(Rf_i - Rf_j) = w^\top(f_i - f_j).$$

Label Switching

When $K \geq 2$, there are $K!$ permutations of parameter labels that yield the same likelihood.

References

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.