Statistica Sinica **35** (2025), 1-17
doi:https://doi.org/10.5705/ss.202022.0236

EFFICIENT ESTIMATION AND INFERENCE FOR THE SIGNED β-MODEL IN DIRECTED SIGNED NETWORKS

Haoran Zhang and Junhui Wang*

Southern University of Science and Technology and
The Chinese University of Hong Kong

Abstract: This paper proposes a novel signed β-model for directed signed network, which is frequently encountered in application domains but largely neglected in literature. The proposed signed β-model decomposes a directed signed network as the difference of two unsigned networks and embeds each node with two latent factors for in-status and out-status. The presence of negative edges leads to a non-concave log-likelihood, and a one-step estimation algorithm is developed to facilitate parameter estimation, which is efficient both theoretically and computationally. We also develop an inferential procedure for pairwise and multiple node comparisons under the signed β-model, which fills the void of lacking uncertainty quantification for node ranking. Theoretical results are established for the coverage probability of confidence interval, as well as the false discovery rate (FDR) control for multiple node comparison. The finite sample performance of the signed β-model is also examined through extensive numerical experiments on both synthetic and real-life networks.

Key words and phrases: Directed network, estimating equation, false discovery rate, node ranking, one-step estimation, status theory.

1. Introduction

Network data has attracted increasing attention from different scientific communities, due to its flexibility in describing various pairwise relations among multiple objects of interest. In literature, various network models have been developed, such as the Erdös-Rényi model (Erdös and Rényi, 1960), the stochastic block model (Holland, Laskey and Leinhardt, 1983; Zhao, Levina and Zhu, 2012), the β-model (Chatterjee, Diaconis and Sly, 2011), the latent space model (Hoff, Raftery and Handcock, 2002), and the network embedding model (Zhang, He and Wang, 2022). Among them, the β-model is one of the most popular models (Rinaldo, Petrović and Fienberg, 2013; Karwa and Slavković, 2016; Graham, 2017; Chen, Kato and Leng, 2021), which explicitly represents each node i with a numeric factor β_i to accommodate degree heterogeneity. Yet, most existing development of the β-model focuses on undirected and unsigned networks, and it is only recently that the directed β-model (Yan, Leng and Zhu, 2016; Yan et al.,

*Corresponding author. E-mail: junhuiwang@cuhk.edu.hk