Z117a Analysis of Integral Field Spectroscopic Data as a High-Dimensional Low-Sample Size Data Problem

竹内　努 (1,2), 河野 海 (1), 中西 康一郎 (3), 矢田 和善, 青嶋 誠 (4), 石井 晶 (5), 江頭 健斗 (4) (1 名古屋大, 2 統数研, 3 国立天文台, 4 筑波大, 5 東京理科大)

Current astronomical instruments provide us with overwhelmingly large data on atoms, molecules, ions, and dust in the interstellar medium (ISM) of observed objects, providing quantitative information of their physical condition. However, detailed observations are generally very time-consuming, and it is not easy to map objects to obtain many independently sampled measurements. If we denote the dimension in the wavelength (or frequency) with d and the number of samples with n, we often find that $n \ll d$. Traditionally in astrophysics, such a situation is regarded as an ill-posed problem, and there was no choice but to throw away most of the information in wavelength direction to let $d < n$. Naturally, it is desirable to avoid such a waste of data and make full use of the amount of information. The data with $n \ll d$ is referred to as high-dimensional low sample size (HDLSS), often found in the field of genome analysis, etc. In order to deal with HDLSS problems, a method called high-dimensional statistics has been developed rapidly in the last decade (e.g., Aoshima 2018). Focusing on the evolution of the interstellar medium of the galaxy, we performed a completely new analysis of integral field spectroscopic data of ALMA with high-dimensional statistics. This is indeed a typical example of the HDLSS data in astrophysics. We discuss the current performance of the method and report the first result on this analysis, as well as some potential problems specific to astronomical data.