Genome-Wide Association Studies Using Haplotype Clustering with A New Haplotype Similarity

Jianhua Guo
Key Laboratory for Applied Statistics of MOE and School of Mathematics and Statistics
Northeast Normal University, Changchun 130024, China
Email: jhguo@nenu.edu.cn Joint work with Lina Jin and Wensheng Zhu

Abstract

Association analysis, with the aim of investigating genetic variations are designed to detect genetic associations with observable traits, which have played an increasing part in understanding the genetic basis of diseases. Among these methods, haplotype-based association studies are believed to possess prominent advantages, especially for the rare diseases in case-control studies. However, when modeling these haplotypes, they are subjected to statistical problems caused by rare haplotypes. Fortunately, haplotype clustering offers an appealing solution. In this research, we have developed a new befitting haplotype similarity for "affinity propagation" clustering algorithm, which can account for the rare haplotypes primely, so as to control for the issue on degrees of freedom. The new similarity can incorporate haplotype structure information, which is believed to enhance the power and provide high resolution for identifying associations between genetic variants and disease. Our simulation studies show that the proposed approach offers merits in detecting disease-marker associations in comparison with the cladistic haplotype clustering method CLADHC. We also illustrate an application of our method to cystic fibrosis, which shows quite accurate estimates during fine mapping.