

Linear Mixed Models for Genome and Epigenome-Wide Association Studies

Understanding the genetic underpinnings of disease is important for screening, treatment, drug development, and basic biological insight. Genome-wide associations, wherein individual or sets of genetic markers are systematically scanned for association with disease are one window into disease processes. Naively, these associations can be found by use of a simple statistical test. However, a wide variety of confounders lie hidden in the data, leading to both spurious associations and missed associations if not properly addressed. These confounders include population structure, family relatedness, cell type heterogeneity, and environmental confounders. I will discuss the state-of-the art approaches (based on linear mixed models) for conducting these analyses, in which the confounders are automatically deduced, and then corrected for, by the data and model.