

A General Framework for Variable Selection in Linear Mixed Models with Applications to Genetic Studies with Structured Populations

Sahir R Bhatnagar^{1,2}, Karim Oualkacha³, Yi Yang⁴, Marie Forest², and
Celia MT Greenwood^{1,2,5}

¹Department of Epidemiology, Biostatistics and Occupational Health, McGill
University

²Lady Davis Institute, Jewish General Hospital, Montréal, QC

³Département de Mathématiques, Université de Québec À Montréal

⁴Department of Mathematics and Statistics, McGill University

⁵Departments of Oncology and Human Genetics, McGill University

July 4, 2018

Abstract

Complex traits are ~~thought~~ ^{known} to be influenced by a combination of environmental factors and rare and common genetic variants. However, detection of such multivariate associations can be compromised by low statistical power and confounding by population structure. Linear mixed effect models (LMM) can account for correlations due to relatedness but ^{are} not applicable in high-dimensional (HD) settings where the number

*should you say
"have not been" ?* 1