

Fibers of multi-way contingency tables given conditionals: relation to marginals, cell bounds and Markov bases

Aleksandra Slavković · Xiaotian Zhu ·
Sonja Petrović

Received: 6 November 2012 / Revised: 6 January 2014 / Published online: 25 June 2014
© The Institute of Statistical Mathematics, Tokyo 2014

Abstract A fiber of a contingency table is the space of all realizations of the table under a given set of constraints such as marginal totals. Understanding the geometry of this space is a key problem in algebraic statistics, important for conducting exact conditional inference, calculating cell bounds, imputing missing cell values, and assessing the risk of disclosure of sensitive information. Motivated by disclosure problems, in this paper we study the space of all possible tables for a given sample size and set of observed conditional frequencies. We show that this space can be decomposed according to different possible marginals, which, in turn, are encoded by the solution set of a linear Diophantine equation. Our decomposition has two important consequences: (1) we derive new cell bounds, some including connections to directed acyclic graphs, and (2) we describe a structure for the Markov bases for the given space that leads to a simplified calculation of Markov bases in this particular setting.

Keywords Conditional tables · Contingency tables · Diophantine equations · Disclosure limitation · Directed acyclic graphs · Marginal tables · Markov bases · Optimization for cell entries