Constrained Sampling of Full Probability Density Input-Output Table Model

Topic: Input-Output Theory and Methodology (5)
Author: Oleg LUGOVOY
Co-Authors: Nikita Lugovoy

Input output tables have multiple data uncertainty issues routing from surveys, they coverage and frequency, compilation of national accounts methods, including averaging and aggregation, to updating tables from former years to the most recent economic data (see Temursho, 2017 for detailed overview of uncertainty sources and treatment in IO analysis). The paper discusses a theoretical framework for marginal probability density functions of input-output (io-) coefficients and explicit probability model of the IOT matrix as a set of truncated Dirichlet distributions. The full probability density model is derived directly from the input-output model rows and columns constraints and is consistent with proposed by Jackson (1986) framework for individual io-coefficients who modeled io-coefficients of individual firms using Gamma distribution, and with Raa and Steelâ€™s (1994) assumption of Beta distribution of technical coefficients. In the paper we reiterate that both Gamma and Beta distributions are natural fit for the matrices and lay out the framework for modeling dependencies in the matrix via a set of Dirichlet distributions for individual rows and columns. Dirichlet distribution and process have been used to model IOTs before. The key difficulty of sampling is the constraint on the columns which are hard to integrate into the sampling algorithm especially for large matrices. We show that the IOT matrix can be represented as a set of truncated Dirichlet distributions where one distribution describes one row of the matrix. This enforces rows-sum constraints. Column constraints can be satisfied by truncating Dirichlet distribution to the feasible interval of every io-cell. The problem has an equivalent dual representation of column-wise Dirichlet distribution with truncation imposed by row-sums.
Still, sampling from truncated Dirichlet distribution with row and column constraints is not as straightforward as from the original Dirichlet distribution. Alternative methods include MCMC to collect samples that satisfy the truncation requirements. MCMC methods are feasible but generally computationally intensive and may not meet both types of constraints. An efficient sampling algorithm would generate random matrices with exact match to all sets of the constraints. Here we propose a generalization of one-dimensional â€œstick braking algorithmâ€• (also known as â€œstring cuttingâ€•) used to sample from Dirichlet distribution to two dimensions and consider several of its variations. We call this family of algorithms Plate Breaking since they share the same idea that: at each level of recursion, the current submatrix is partitioned into subregions, where some form of prior is used to propose how the parent mass should be divided among them, and the process recurses. Variations of constructions of Plate Breaking differ in how they partition, what prior they place over the proposed division, or how feasibility is enforced, but all share the property that mass is conserved exactly at every level and constraints aim to be satisfied by construction rather than by rejection.
The implementation of the sampling algorithms is tested on small synthetic tables, as well as Scottish IOT (2007) 11x11, WIOD-Australia table (34x34), the US IOT (402+ industries), and synthetic matrix of 10 thousand rows and columns. All algorithms allow control of concentration parameters of the matrix and particular rows, columns, show unbiased estimates for all concentrations, and computationally light. The algorithms have potential broad implications for input output analyses and updating, for national accounts, general equilibrium modeling, and optimization.
Temursho U. (2017), Uncertainty treatment in input-output analysis, in: Ten Raa, T., Handbook of Input-Output Analysis, Edward Elgar Publishing, Chapter 12, pp. 407-463.
Jackson, R.W. (1986), â€˜The full-distribution approach to aggregate representation in the input-output modeling frameworkâ€™, Journal of Regional Science, 26 (3), 515-531.
Thijs ten Raa, Mark F.J. Steel, Revised stochastic analysis of an input-output model, Regional Science and Urban Economics, Volume 24, Issue 3, 1994, Pages 361-371.