A Minimum Bias Latin Hypercube Design

Deterministic engineering design simulators can be too complex to be amenable to direct optimization. An indirect route involves data collection from the simulator and fitting of less complex surrogates: metamodels, which are more readily optimized. However, common statistical experiment plans are not appropriate for data collection from deterministic simulators due to their poor projection properties. Data collection plans based upon number-theoretic methods are also inappropriate because they tend to require large sample sizes in order to achieve their desirable properties. We develop a new class of data collection plan, the Minimum Bias Latin Hypercube Design (MBLHD), for sampling from deterministic process simulators. The class represents a compromise between empirical model bias reduction and dispersion of the points within the input variable space. We compare the MBLHD class to previously known classes by several model independent measures selected from three general families: discrepancies, maximin distance measures, and minimax distance measures. In each case, the MBLHD class is at least competitive with the other classes; and, in several cases the MBLHD class demonstrates superior performance. We also make a comparison of the empirical squared bias of fitted metamodels. We approximate a mechanistic model for water flow through a borehole, using both kriging and polynomial metamodels. Here again, the performance of the MBLHD class is encouraging.