Dimitrova, Elena; Hu, Jingzhen; Liang, Qingzhong; Stigler, Brandilyn; Zhang, Anyu

Algebraic model selection and experimental design in biological data science. (English)
Zbl 07430815

Summary: Design of experiments and model selection are essential steps in biological data-science pipelines and coupling them has the potential to improve computational efficiency. In particular approaches which can handle data from unspecified distributions and nonlinear models are of interest. We propose an algebra-based computational framework to unify experimental design and model selection for discrete data sets and minimal polynomial models. We use an affine transformation, called a linear shift, to identify data sets and bases of polynomial models simultaneously. We present the theoretical foundation for a web-accessible database. This framework enables us to address two illustrative questions that may arise when studying biological systems: finding the data which identifies a set of known interactions and finding identifiable interactions given a set of data. As an example, we apply the methodology to a previously constructed pharmacodynamic model of epidermal derived growth factor receptor (EGFR) signaling.

MSC:
13P10 Gröbner bases; other bases for ideals and modules (e.g., Janet and border bases)
13P25 Applications of commutative algebra (e.g., to statistics, control theory, optimization, etc.)
62P10 Applications of statistics to biology and medical sciences; meta analysis
11Txx Finite fields and commutative rings (number-theoretic aspects)

Keywords:
model selection; experimental design; biological data science; Gröbner bases; standard monomials; finite fields

Full Text: DOI

References:


This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.