Border Basis of an Ideal of Points and its Application in Experimental Design and Regression

Samira Pourkhajouei, Sareh Goli, Amir Hashemi*,
Department of Mathematical Sciences,Isfahan University of Technology Isfahan,

Received: 2018/11/18 Accepted: 2019/03/02

Extended Abstract

Paper pages (549-562)

Introduction

Border bases are a generalization of Gröbner bases for zero-dimensional ideals which have attracted the interest of many researchers recently. More precisely, border bases provide a new method to find a structurally stable monomial basis for the residue class ring of a polynomial ideal and this yields a special generating set for the ideal possessing many nice properties.

Given a finite set of points, finding the set of all polynomials vanishing on it (so-called either ideal of points or vanishing ideal of the set of points) has numerous applications in several fields in Mathematics and other sciences. In 1982, Buchberger and Möller proposed an algorithm to compute a Gröbner basis for an ideal of points. This algorithm proceeds by performing Gaussian elimination on a generalized Vandermonde matrix. In 2006, Farr and Gao presented an incremental algorithm to compute a Gröbner basis for an ideal of points. The main goal of their paper is to calculate a Gröbner basis for the vanishing ideal of any finite set of points under any monomial ordering, and for points with nontrivial multiplicities they adapt their algorithm to compute the vanishing ideal via Taylor expansions.

The method of border bases is a beneficial tool to obtain a set of polynomial models identified by experimental design and regression. The utilization of Gröbner bases theory in experimental design was introduced by Pistone and Wynn. However, using Gröbner bases we cannot find all possible models which form structure of an order ideal for an experiment. For example, if we consider the design {(-1,1),(1,1),(0,0),(1,0),(0,-1)}, the model $\{1, x, y, x^2, y^2\}$ cannot be computed by Gröbner bases method. This fact is expected this method relies on monomial orderings.

Material and methods

In this paper, we first present the Buchberger-Möller and Farr-Gao algorithms and then by applying these algorithms, we describe an algorithm which computes a border basis for the ideal of points corresponding to the input set of points with nontrivial multiplicity. In addition, we focus on presenting different models related to an experiment by using the concept of monomial bases for the residue class ring of a polynomial ideal.

Results and discussion

As we mentioned earlier, Buchberger-Möller algorithm is an efficient algorithm to compute a Gröbner basis for an ideal of points. We describe a simpler presentation of this algorithm in which we use the function NormalForm which receives as input a linear polynomial p and a Gröbner basis G = $\{g_1, \ldots, g_m\}$ of linear polynomials in $y_1, \ldots, y_s$ and returns f and q=$[q_1, \ldots, q_m]$ where f is the remainder of the division of p by G and p=$q_1g_1+\cdots+q_mg_m$+r.

Furthermore, we compare the efficiency of this algorithm with the function VanishingIdeal of Maple.

Given a finite set of points, we consider the case in which some points in the set have nontrivial multiplicity. Based on the Farr-Gao algorithm, we prepare an algorithm that computes a border basis for the vanishing ideal of the finite set of points by using Taylor expansions.

Suppose that n is the number of factors in an experiment. An experimental design is a finite set $X \subset K^n$ of points. The set of all polynomials vanishing at the design is called a design ideal. Regression analysis is a useful statistical process for the investigation of relationships between a response (or dependent) variable and one or more predictor (or independent) variables. When there is more than one predictor variable in a regression model, the model is a multiple linear regression model which we can call polynomial model. Suppose a random sample of size n is given (then we have exactly n data points are observed from (Y,$\underline{X}$). The expession

$$\beta_0 + \beta_1 x_1 + \beta_2 x_1 + \cdots + \beta_k x_k + \text{error?}$$

is the model for multiple linear regression where β_i's are called slopes or regression coefficients. Also, representing the merged effects of the predictor variables on the response variable is called interaction effect. By using multiple linear regression, we can analyze models containing interaction effects. For example, let us consider the following model

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 {x_1}^2 + \beta_4 x_1 x_2 + \text{error}.$$

By substituting ${x_1}^2 = x_3$ and $x_1 x_2 = x_4$, we have a multiple linear regression as follows

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \text{error}.$$

In addition, multiple R-squared or R^2 is a statistical measure that states the square of the relationship between the predicted response value and response value. It should be noted that multiple R-squared is always any value between 0 and 1, where a value closer to 1 informing that a greater proportion of variance is computed for the model. Statistically, a high multiple R-squared shows a well-fitting regression model. Also in multiple regression, tolerance is used as an indicator of multicollinearity. Tolerance may be said to be the opposite of the coefficient of determination and is obtained as $1 - R^2$. All other things equal, researchers desire higher levels of tolerance, as low levels of tolerance are known to affect adversely the results associated with a multiple regression analysis. The smaller the tolerance of a variable, the more redundant is its contribution to the regression (i.e., it is redundant with the contribution of other independent variables). In the regression equation, if the tolerance of any of the variables is equal to zero (or very close to zero), the regression equation cannot be evaluated (the matrix is said to be ill-conditioned in this case, and it cannot be inverted).

Conclusion

The following conclusions were drawn from this research.

- We present a simpler variant of Buchberger- Möller algorithm (which seems to be easier for the implementation issue) for computing a border basis for an ideal of points.
- We present an algorithm that incrementally computes a border basis for the vanishing ideal of any finite set of points in which some points have multiplicity.
- We provide good statistical polynomial models which are more suitable for practical applications due to the stability of border bases models compared with Gröbner bases models.

Keywords: Border basis, Ideal of points, Buchberger-Möller algorithm, Experimental design, Regression.

*Corresponding author: amir.hashemi@cc.iut.ac.ir