Journal of Computational and Applied Mathematics 346 (2019) 609–619

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Multi-variable regression methods using modified Chebyshev polynomials of class 2

Zibo Li [a], Guangmin Sun [a,*], Cunfu He [b], Xiucheng Liu [b,*], Ruihuan Zhang [b], Yu Li [a], Dequn Zhao [a], Hao Liu [a], Fan Zhang [a]

[a] Faculty of Information Technology, Beijing University of Technology, Beijing, 100124, China
[b] College of Mechanical Engineering and Applied Electronics Technology, Beijing University of Technology, Beijing, 100124, China

ARTICLE INFO

Article history:
Received 22 November 2016
Received in revised form 14 November 2017

Keywords:
Polynomial regression
Chebyshev polynomials of class 2
Feature selection
Cascade regression

ABSTRACT

So far, many regression works have been implemented by using linear regression methods. Although more accurate predictions results could be obtained, polynomial regression is not used as much as compared to linear regression in real applications due to occurrence of coefficient explosion. To overcome this problem, two regression algorithms using Chebyshev polynomials of class 2 based on cascade regression and feature selection are proposed in this paper. In the experimental part, three separate experiments including function interpolation and real-case regression were conducted on three datasets to test the proposed algorithms. As shown by the experimental results, the proposed algorithms performed better than other regression methods in terms of both accuracy and processing time.

1. Introduction

Among various quantitative prediction methods, regression is one of the commonly used methods. With the contributions brought by mathematicians, a series of popular regression tools such as linear regression, polynomial regression and rigid regression etc. have been adopted and among them polynomials regression could be treated as one of the most accurate method [1]. In all the polynomials regression methods, Chebyshev polynomials regression is becoming an outstanding one for its fitting capabilities. However, polynomials (e.g. Chebyshev polynomials) are difficult to be employed in multi-variable regression case in the current stage. In this paper, two techniques including the feature selection and cascade regression are considered to improve the performance of multi-variable Chebyshev polynomials regression.

According to Hoel P.G. [1], Chebyshev polynomials regression was firstly adopted for single-variable function interpolation in 1966. Based on the theory of Chebyshev polynomials of class 2 (CP-2), Chebyshev polynomials regression has some eminent properties [1–4], such as high approximation capability with low polynomials and weighted orthogonality. Especially, in the case of analytical understanding, Chebyshev polynomials could be adopted to ease the Runge phenomenon. Based on this advantage, we could try to use the high order of Chebyshev polynomials to deal with a regression problem. And

* Corresponding authors.
E-mail addresses: glory0117@163.com (Z. Li), gmsun@bjut.edu.cn (G. Sun), xiuchliu@bjut.edu.cn (X. Liu).

Abbreviations: CP-2, Chebyshev polynomials of class 2; MIFS-CR, Mutual Information-based Feature Selection with Class-dependent Redundancy; PR, Polynomial regression; PLSR, partial least squares regression; PCPR, partial Chebyshev polynomials regression; MCPR, Multi-variable Chebyshev Polynomials Regression; AQD, Air Quality Dataset; CCPPD, Combined Cycle Power Plant Dataset; CV, cross-validation; BP, Back propagation; PGPR, Partial Gegenbauer polynomial regression; PLPR, Partial Legendre polynomial regression; PHPR, Partial Hermit polynomial regression; PPPR, Partial Pollaczek polynomial regression

https://doi.org/10.1016/j.cam.2018.04.022
0377-0427/

this is the main reason for us to discuss the application of Chebyshev polynomials here. However, after the study [1], there were few studies or applications using Chebyshev polynomials regression [5–12]. During the same period, most researchers focused on linear regression (e.g. [13,14]) because linear regression methods could be easily applied and transplanted into the case of multi-variable regression. In order to employ the CP-2 for multi-variable regression, Zhang Y. et al. [8] combined Bernstein polynomials with CP-2 to achieve a multi-variable Chebyshev polynomials regression algorithm and defined a cross-validation scheme to prune the redundant polynomial. Although a big contribution was obtained in their study, this hybrid regression method still had two limitations. Firstly, the relationship between the number of regressing coefficients N and the number of regressed attributes t is exponential (i.e. $N = n^t$, where n is the number of polynomials). Therefore, N grows quickly with the increase of the number of regressed variables, which is known as coefficient explosion. It could be ascribed to the large number of different combinations of the serial number of variables and polynomials. Secondly, the accuracy of their algorithm relied mainly on the random grouped result. Compared with the second limitation, coefficient explosion is more vital because the application of polynomials regression is severely restricted by this issue. In this paper, different strategies are proposed to analyze and solve the problem of coefficient explosion in the case of multi-variable regression problem.

To overcome the coefficient explosion, the main purpose of our analysis is to reduce the number of redundant variables. As far as reducing the redundant variables are concerned, two ideas could be addressed. On one hand, each variable could be treated as an individual and contribute to the predicted value separately. Therefore, a variable could be removed from the regression variable set when it is unrelated to the predicted value. Based on this idea, feature selection algorithm seems to be a suitable choice and is designed as an important preprocessing algorithm and treated as a practical way to reduce the complexity of input set. There are some selection algorithms including correlation-based methods [13], uninformative variable elimination [15], successive projections algorithm [16] and mutual-information-based selection [17–20]. Among them, mutual-information-based selection algorithm (MIFS, e.g. [17]) is more attractive than others because not only the knowledge about information entropy but also feature relevance and redundancy are considered. Especially for MIFS-CR proposed by Wang Z. et al. [20], based on the relevance between redundant and original variable set, the relationship between redundant and selected variable set was the focus in their analysis. In our paper, as a preprocessing step, MIFS-CR is combined with CP-2 to build a new regression method. On the other hand, a scheme named cascade regression has become a hot topic. It has been used in some image processing applications, such as face alignment [21–26] and facial object localization [27]. The basic idea of cascade regression is obtaining the accurate prediction by accumulating the outputs that are obtained from several regression steps. In details, through a nonlinear way, each input of the regression step was linked to the residual that is generated from the former regression step. According to the basic idea of cascade regression, a proper nonlinear transformation (e.g. SIFT in [24]) could be treated as a key point. This assumption is valid for two main reasons. Firstly, the nonlinear transformation could be employed to reduce the variable that can contribute to the prediction. Secondly, according to the theory of matrices, the residual error of prediction could not be directly represented by the same matrix. Inspired by partial least squares regression (PLSR) [28,29] that could be treated as a cascade regression model, a novel cascade Chebyshev polynomials regression method named partial Chebyshev polynomials regression method (PCPR) is designed in this paper. In details, during each iteration of PCPR, the main components will be extracted from the combination of residual input and residual output and then used to drive a Chebyshev polynomials regression tool. In this way, the number of components used for one regression tool could be controlled and the redundant variable could be neglected. The main contributions of this paper are concluded as:

A. As a preprocessing step, the feature selection algorithm (i.e. MIFS-CR) proposed by Z. Wang et al. [20] is adopted to reduce the number of variables before the selected variables are employed as the input of Chebyshev polynomials regression method. We call this combined scheme algorithm as MIFS-CR+CP-2. In this case, each variable is treated as an individual and the useful one is picked for the regression.
B. As far as the cascade regression is concerned, a novel model named partial Chebyshev polynomials regression (PCPR) is proposed in this paper. In terms of PCPR, the idea of cascade regression, Chebyshev polynomials and variable pruning are considered and combined as an effective regression model. From the perspective of PCPR, all variables are considered as a unified set and the principle components projection is used as the nonlinear transformation.

In the experiment part, three experiments were conducted. In the first experiment, a family of functions with a different number of variables was adopted to present the superiority of PCPR. In the second experiment, two datasets were employed to show the high accuracy of the feature-selection-based Chebyshev polynomials regression method. In the third experiment, based on the same scheme of two algorithms, different polynomials (e.g. Hermit, Gegenbauer, etc.) were adopted to evaluate the performance of our selected polynomials. As a result, the superiority of two algorithms under different conditions was demonstrated. On one hand, PCPR is the superior tool when each variable makes equal contribution to the prediction problem (e.g. function interpolation). On the other hand, compared with PCPR, feature-selection-based Chebyshev polynomials regression model shows better performance when the number of useful variables is limited in a prediction problem (e.g. real prediction cases).

The paper is organized into following parts: Section 2 introduces the theory of multi-variable Chebyshev polynomials regression algorithm and proposes some details of our method subsequently. Section 3 provides some information about experiments and discusses the results. Section 4 ends up the article with conclusion.