Expert Systems With Applications 76 (2017) 129–139

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Composite quantile regression neural network with applications

Qifa Xu[a,b], Kai Deng[a], Cuixia Jiang[a,*], Fang Sun[c], Xue Huang[d]

[a] School of Management, Hefei University of Technology, Hefei 230009, Anhui, PR China
[b] Key Laboratory of Process Optimization and Intelligent Decision-making, Ministry of Education, Hefei 230009, Anhui, PR China
[c] Department of Accounting and Information Systems, Queens College, New York, USA
[d] Department of Statistics, Florida State University, Tallahassee 32304, USA

ARTICLE INFO

Article history:
Received 6 December 2015
Revised 29 January 2017
Accepted 30 January 2017
Available online 31 January 2017

Keywords:
Quantile regression
Neural network
Quantile regression neural network
Composite quantile regression
CQRNN

ABSTRACT

In recent years, there has been growing interest in neural network to explore complex patterns. We consider an extension of this framework in composite quantile regression setup and propose a novel composite quantile regression neural network (CQRNN) model. We further construct a differential approximation to the quantile regression loss function, and develop an estimation procedure using standard gradient-based optimization algorithms. The CQRNN model is flexible and efficient to explore potential nonlinear relationships among variables, which we demonstrate both in Monte Carlo simulation studies and three real-world applications. It enhances the nonlinear processing capacity of ANN and enables us to achieve desired results for handling different types of data. In addition, our method also provides an idea to bridge the gap between composite quantile regression and intelligent methods such as ANNs, SVM, etc., which is helpful to improve their robustness, goodness-of-fit and predictive ability.

1. Introduction

Regression analysis plays an important role in exploring the possible relationships among variables. Conventional regression analysis is based on the ordinary least squares (OLS) framework and only focuses on the conditional mean of the response variable given values of covariates. Although OLS is relatively efficient when random error is normally distributed, it fails when the error variance is infinite. As an alternative way, quantile regression (QR) of Koenker and Bassett (1978) has become a popular paradigm to describe the complete conditional distribution information hidden in variables. It considers the entire conditional distribution instead of simply focusing on the center of the distribution. QR has been widely used in various areas such as economics, survival analysis, medicine and environmental sciences, among others. It has been proved that QR avoids the breakdown but has arbitrarily small relative efficiency compared with OLS. To overcome the drawbacks of traditional QR, multiple quantile regression models are simultaneously considered and a composite approach is used in Zou and Yuan (2008) to develop a new quantile regression method called composite quantile regression (CQR). It is worth mentioning that CQR is not only always valid regardless of the error distribution, but also works much better in efficiency comparing with OLS.

Since then, considerable effort has been devoted to developing the extensions of CQR. It is noted by Jiang (2012) that treating the quantile loss at each quantile level equally is not optimal. Hence, a weighted composite quantile regression (WCQR) approach is proposed using a data-driven weighting scheme. Moreover, another two weighting schemes are considered by Zhao and Lian (2015) to further improve the efficiency of CQR. Despite their merits, a drawback with conventional CQR is that they typically can not handle nonlinear quantile regression problems. Given this difficulty, some nonlinear CQR methods including local polynomial CQR of Kai, Li, and Zou (2010) and Li and Li (2016), semiparametric CQR of Kai, Li, and Zou (2011), partially linear additive CQR (A-CQR) of Guo, Tang, Tian, and Zhu (2013), single-index CQR (SI-CQR) of Jiang, Zhou, Qian, and Chen (2013), are extensively investigated. The partially additive linear CQR model can be divided into two components, of which the parametric components are linear in parameters while the nonparametric ones are approximated by polynomial splines. On the other hand, single-index models provide an efficient way to resolve nonlinear estimation problems and avoid the possible curse of dimensionality by assuming that the response is related with a single linear combination of covariates. Two step composite quantile regression is developed by Jiang et al. (2013) to provide an efficient and stable estimation of single-index models where the hidden link function is approximated by a local linear regression.

* Corresponding author.
E-mail addresses: xuqifa1975@gmail.com (Q. Xu), RichardDK@163.com (K. Deng), jiangcuixia@hfut.edu.cn (C. Jiang), fsun@qc.cuny.edu (F. Sun), xhuang@stat.fsu.edu (X. Huang).

http://dx.doi.org/10.1016/j.eswa.2017.01.054
0957-4174/

As an alternative approach, artificial intelligence plays an important role in processing nonlinear problems. It is encouraging that these intelligent methods do not require any prior assumption about underlying function or model. It is a significant advantage over traditional statistical methods such as additive model and single-index model, etc. Artificial neural network (ANN) and support vector machine (SVM) are two well-known techniques to tackle nonlinear problems, refer to Zhao et al. (2015) and Zbikowski (2015). They have been successfully applied to implement nonlinear quantile regression in practice. For instance, a quantile regression neural network (QRNN) method is considered in Taylor (2000) and Cannon (2011) through combining the nice properties of ANN and QR, and a support vector quantile regression (SVQR) approach is proposed in Takeuchi, Le, Sears, and Smola (2006) and Li, Liu, and Zhu (2007) by combining SVM with quantile regression. SVM has already been applied by Shim, Hwang, and Seok (2014) to develop the composite support vector quantile regression (CSVQR) approach. However, there is relatively little literature available on solving CQR in the nonlinear context using ANN.

Inspired by the attractive properties of CQR and ANN, this paper proposes a new regression method, called composite quantile regression neural network (CQRNN), which adds neural network structure to CQR. It is an artificial neural network extension of linear composite quantile regression. We first apply ANN technique to approximate the potential nonlinear relationship in CQRNN model. With this approximation, the nonlinear component is represented by a combination of predictors with appropriate weights and biases. Consequently, training CSVQR model becomes a problem of estimating the coefficients in the combinations. We then use the CQR approach to estimate the coefficients due to its appealing properties and performance. The CQRNN model can be estimated via optimizing an approximate loss function and standard gradient based optimization algorithms. Our CQRNN model can also be viewed as a composite version of the QRNN model of Cannon (2011). It synthetically utilizes multiple regression quantiles modeled by the QRNN model and consequently produces a more efficient estimator from a theoretical point of view. To demonstrate the efficacy of the proposed CQRNN model we conduct three Monte Carlo simulation studies and three real-world applications. The numerical results show that the model is flexible to explore potential nonlinear relationships among variables and outperforms many conventional methods in terms of goodness-of-fit and predictive ability.

This paper contributes to the literature on intelligent modeling in several ways. First, the CQRNN model enhances the nonlinear processing capacity of ANN and enables us to achieve desired results for handling different types of data. Second, our method also provides an idea to bridge the gap between composite quantile regression and intelligent methods such as ANNs, SVM, etc., which is helpful to improve their robustness, goodness-of-fit and predictive ability. In addition, as an intelligent technique, the complexity of CQRNN model is also determined by the number of hidden layer nodes. The problem of overfitting the model is solved by adding a penalty term in the empirical approximate loss function. We further provide a model selection scheme to select the optimal number of nodes and the regularization parameter. In this paper, the CQRNN model is designed under a multilayer perceptron framework. Similar to most intelligent methods, the CQRNN is a time-consuming model. From the expert systems perspective, there are some hybrid systems that have implemented neural networks as part of a larger whole construction of a combination system. A potential way is to refine the CQRNN model by reducing features or variables via some feature extraction methods. Serving the extracted significant variables as the input nodes of the CQRNN could improve the performance of hybrid system in terms of accuracy as well as computational time.

The rest of this paper is organized as follows. In Section 2 we present the CQRNN model in detail. In Section 3 we investigate the finite sample performance of the proposed CQRNN method through Monte Carlo simulation studies. Three real world applications with benchmark data sets in Section 4 highlight the efficacy of our proposed method. We make brief conclusions in Section 5.

2. Composite quantile regression neural network

In this section we propose the composite quantile regression neural network (CQRNN) model through adding ANN structure to composite quantile regression method. We first discuss the model setup, and then provide a simple conditional quantile estimator using a standard gradient based algorithm. Moreover, we construct an extended BIC criterion for model selection.

2.1. Model setup

In practice, we often confront the problem of exploring the true relationship between a response y and predictors $\boldsymbol{x} \equiv (x_1, x_2, \cdots, x_p)'$. From a statistical perspective, this problem can be stated as a stochastic model

$$y = f(x_1, x_2, \cdots, x_p) + \epsilon \equiv f(\boldsymbol{x}; \boldsymbol{\theta}) + \epsilon, \tag{1}$$

where $\boldsymbol{\theta}$ is a vector with parameters to be estimated, ϵ is a random error. This model is typically solved using regression techniques. To this end, nonlinear OLS is a classical one to explore the potential relationship f and has been widely used in numerous applications. As a supplementary way, quantile regression is able to present a complete description of the entire conditional distribution of y given $\boldsymbol{x}$ instead of only considering the conditional mean. Many nonlinear quantile regression methods, including locally polynomial QR, partially linear additive QR, and B-spline QR, are developed to estimate the nonlinear relationship f. In addition, artificial neural network (ANN) has also been employed by Taylor (2000) and Cannon (2011) to develop the quantile regression neural network (QRNN), which is flexible to implement a nonlinear quantile regression without prior specification of the form of the relationships. It turns out that QRNN is more efficient than aforementioned conventional nonlinear quantile regression methods in terms of estimation and prediction accuracy.

In particular, the QRNN objective aims for minimizing the empirical loss function

$$EQRNNL(\tau) = \frac{1}{N}\sum_{i=1}^{N} \rho_\tau(y_i - Q_{y_i}(\tau)), \tag{2}$$

where $\rho_\tau(u) \equiv \tau u I(u \geq 0) + (\tau - 1) u I(u < 0)$ is the check function defined in Koenker and Bassett (1978) for $0 \leq \tau \leq 1$, $Q_{y_i}(\tau)$ is the conditional quantile of y at τ and can be estimated in the following two steps. First, an output from the jth hidden layer node $g_{j,l}(\tau)$ is given by applying a sigmoid transfer function $f^{(h)}$ to the inner product between the predictors and the hidden layer weights $w_{jl}^{(h)}$ plus the hidden layer bias $b_j^{(h)}$

$$g_{j,l}(\tau) = f^{(h)}\left(\sum_{j=1}^{p} x_{ij} w_{jl}^{(h)} + b_l^{(h)}\right), \tag{3}$$

where $\boldsymbol{w}^{(h)} \equiv (w_{11}^{(h)}, w_{12}^{(h)}, \cdots, w_{p,L}^{(h)})'$ is a weight vector of hidden layer, $\boldsymbol{b}^{(h)} \equiv (b_1^{(h)}, b_2^{(h)}, \cdots, b_L^{(h)})'$ is a bias vector of the hidden layer, and $f^{(h)}$ denotes a sigmoid transfer function often using the hyperbolic tangent. Second, an estimate of the conditional τ th conditional quantile of response is consequently given by

$$Q_{y_i}(\tau) = f^{(o)}\left(\sum_{l=1}^{L} w_l^{(o)} g_{j,l}(\tau) + b^{(o)}\right), \tag{4}$$