Applied Soft Computing 11 (2011) 5793–5800

Contents lists available at ScienceDirect

Applied Soft Computing

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

Measuring financial risk with generalized asymmetric least squares regression

Yongqiao Wang[a,*], Shouyang Wang[b], K.K. Lai[c]

[a] College of Finance, Zhejiang Gongshang University, Hangzhou 310018, Zhejiang, China
[b] Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080, China
[c] Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong, China

ARTICLE INFO

Article history:
Received 20 May 2009
Received in revised form
25 November 2010
Accepted 20 February 2011
Available online 18 July 2011

Keywords:
Risk measurement
Value-at-risk
Expected shortfall
Kernel trick
Asymmetric least squares regression

ABSTRACT

This paper proposes a generalized asymmetric least squares regression method to estimate Value-at-risk and expected shortfall. By solving an asymmetric least squares regression in a Reproducing Kernel Hilbert Space, the method achieves nonlinear prediction power, while making no assumption on the underlying probability distributions. Two toy datasets are used to demonstrate its nonlinear prediction power. The empirical results on the S&P 500 stock index obviously show that the method is superior to other four benchmark methods.

1. Introduction

Currently Value-at-risk (VaR) and expected shortfall (ES) have become standard risk measures for financial institutions [15]. Value-at-risk (VaR) is defined as the maximum potential loss of a portfolio with a given confidence level over a certain time horizon. ES is simply the average of loss that is larger than VaR. ES is a coherent risk measure whereas VaR is not, due to its non-subadditiveness [3,8]. Though many methods have been suggested, their measurement is still a very challenging statistical problem [5].

Recently Taylor [25] suggested to estimate VaR and ES via linear asymmetric least squares regression (LALSR). LALSR was originally proposed by [2] and was analyzed further by [20] comprehensively. Efron [9] showed that the expectile could be used to estimate VaR and ES because there existed a one-to-one mapping from expectiles to quantiles. Based on the above idea, Taylor [25] applied LALSR to estimate VaR and ES. Its experimental results showed that the LALSR estimation method was competitive with some conventional methods.

Though the LALSR method has many advantages, it fails to handle the nonlinear features of most financial returns. Those nonlinear features include nonnormality, asymmetric cycles, bimodality, nonlinear relationship between lagged variables, variation of prediction performance over the state-space, time irreversibility, sensitivity to initial conditions, and others. They have been well-observed in many financial time series data. See [10,26] for further discussion on this topic.

In this work we propose to use kernel trick [21,22], originally applied in support vector machine, to define a generalization of the asymmetric least squares regression method [25]. Hereafter we call the new method as GALSR. Kernel trick approaches many statistical problems by mapping the data into a high dimensional feature space, where each coordinate corresponds to one feature of the data items, transforming the data into a set of points in a Euclidean space. In that space, a variety of methods can be used to find relations in the data. Since the mapping can be quite general (not necessarily linear, for example), the relations found in this way are accordingly very general. Kernel trick owe its name to the use of kernel functions, which enable them to operate in the feature space without ever computing the coordinates of the data in that space, but rather by simply computing the inner products between the images of all pairs of data in the feature space. This operation is often computationally cheaper than the explicit computation of the coordinates. Kernel trick was first proposed by [28,27]. Since then it has been applied in many traditional statistical problems and enables them with nonlinear power. Support vector machine, as one of the most famous kernel method, has found many successful applications, such as text classification [19], video classification [4], systems identification [18] and banknote recognition [31].

* Corresponding author. Tel.: +86 13216816222.
E-mail addresses: wangyq@zjgsu.edu.cn (Y. Wang), sywang@amss.ac.cn (S. Wang), mskklai@cityu.edu.hk (K.K. Lai).

1568-4946/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2011.02.018

Kernel trick has found many applications in financial risk management, such as foreign exchange prediction [13] and credit risk evaluation [29].

The GALSR estimation method has three advantages. First, it has nonlinear prediction power and can handle nonlinear complexity in financial time series, which is an advantage over the LALSR measurement method [25]. Second, the method is free from any assumptions on underlying probability distributions, which is an advantage over many parametric models. Third, it can estimate VaR and ES at the same time, which is an advantage over many non-parametric models.

This paper is organized as follows. Section 2 introduces how to apply the LALSR method to estimate VaR and ES. Section 3 details the GALSR model, i.e. how to apply kernel trick in asymmetric least squares regression and how to transform it into a quadratic programming problem. Section 4 provides experimental results, both on two artificial data sets and one financial time series. The paper is concluded in Section 5.

2. LALSR method

Let Y be a real-valued random variable with a continuous and strictly increasing distribution function $F_Y(y) = \Pr\{Y \le y\}$. In risk measurement, usually Y is the return on a given asset or portfolio over a certain time horizon. Its quantile function is defined as:

$$Q_Y(\theta) = \inf\{y : F_Y(y) \ge \theta\} \tag{1}$$

where $\theta \in (0, 1)$.

VaR is defined as the maximum potential loss in value of a portfolio of financial instruments with a given confidence level $1-\theta$ over a certain time horizon. The VaR of the random return Y with confidence level $1-\theta$ is

$$\text{VaR}_Y(1-\theta) = -Q_Y(\theta). \tag{2}$$

ES is defined as the conditional expectation of the loss given that it exceeds the VaR. So the ES with confidence level $1-\theta$ is

$$\text{ES}_Y(1-\theta) = -\mathbb{E}(Y|Y \le Q_Y(\theta)) \tag{3}$$

2.1. Expectiles

Asymmetric least squares regression solves the following minimization problem

$$\min_m \mathbb{E}(|\tau - \mathbf{1}(Y < m)|(Y-m)^2), \tag{4}$$

where the expectation $\mathbb{E}$ is taken with respect to the random variable Y, $\tau \in (0,1)$ is the asymmetric parameter that tunes the weights, and $\mathbf{1}(A)$ denotes the indicator function for the event A. In the paper, we call the optimal solution of (4) as τ-expectile $\mu_Y(\tau)$. The central case, $\tau = 1/2$, when (4) becomes the ordinary least squares regression, gives $\mathbb{E}Y$.

Instead of estimating VaR and ES from τ-expectile, Kuan et al. [17] suggested that τ-expectile $\mu_Y(\tau)$ itself could be a financial risk measure. In this case, τ can be regarded as level of prudentiality.

2.2. From expectiles to VaR and ES

It must be emphasized that θ-expectile is not necessarily equal to θ-quantile. So we cannot obtain θ-quantile directly by solving (4) with parameter setting $\tau = \theta$. But Efron [9] shows that the expectile can be used to estimate VaR and ES because there exists a one-to-one mapping from expectiles to quantiles. The existence of a one-to-one mapping from expectiles to quantiles is supported by theoretical work of [1,14]. Empirical support for Efron's proposal is provided by [12,23], using macroeconomic data and absolute financial returns, respectively.

Efron [9] suggested that the asymmetric parameter τ should be set such that the proportion below the τ-expectile is θ, i.e.

$$F(\mu_Y(\tau)) = \theta. \tag{5}$$

Because the distribution function F is unknown, τ should be tuned such that the proportion of observations that are below $\mu_Y(\tau)$ is θ.

The estimation of ES from expectiles can be followed from the following theorem. If $\mu_Y(\tau)$ solves (4) and $F(\mu_Y(\tau)) = \theta$, we have

$$\mathbb{E}(Y|Y \le Q_Y(\theta)) = \left(1 + \frac{1}{\theta}\frac{\tau}{1-2\tau}\right)\mu_Y(\tau) - \frac{1}{\theta}\frac{\tau}{1-2\tau}\mathbb{E}Y. \tag{6}$$

Due to the implicit dependence between τ and θ, it is not necessarily that the conditional expectation $\mathbb{E}(Y|Y \le Q_Y(\theta))$ linearly depends on expectile $\mu_Y(\tau)$.

2.3. Conditional expectiles and LALSR

In usual applications the task is to predict Y's quantiles and expectations conditional on some prior information $\mathbf{X}$, which means that we are tackling the conditional variable $Y|\mathbf{X}$. In financial risk measurement the predictor $\mathbf{X}$ is all information available to the modeler at the time of prediction. The above theoretical analysis can be easily extended to the case of conditional random variable, just by replacing the independent random variable Y with the conditional random variable $Y|\mathbf{X}$.

Though one can predict VaR and ES by deducing the distribution of the conditional random variable $Y|\mathbf{X}$ from previous observations $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_N, y_N)$, $\mathbf{x}_i \in \Re^d$, $y \in \Re$, such endeavor often surrounds great computational complexity, due to the fact that high-dimension distributions are very difficult to model. Instead of inferring the distribution of $Y|\mathbf{X}$, the LALSR method directly build the functional relations between the predictors $\mathbf{X}$ and the expectiles $\mu_{Y|\mathbf{x}}(\tau)$ by the following minimization problem

$$\min_{\mu_{Y|\mathbf{x}}(\tau) \in \mathcal{H}} \sum_{i=1}^{N} \rho_\tau(y_i - \mu_{Y|\mathbf{x}}(\tau)), \tag{7}$$

where $\mathcal{H}$ is the function space from $\Re^d \to \Re$ and $\rho_\tau(\cdot)$ is the asymmetric least error loss function

$$\rho_\tau(r) = \begin{cases} \tau r^2 & \text{if } r > 0 \\ (1-\tau)r^2 & \text{otherwise.} \end{cases} \tag{8}$$

To make the above minimization problem tractable, Newey and Powell [20] suggested a linear function form of $\mu_{Y|\mathbf{x}}(\tau)$

$$\mu_{Y|\mathbf{x}}(\tau) = b + \mathbf{w}^T\mathbf{x}, \tag{9}$$

where $b \in \mathbb{R}$ and $\mathbf{w} \in \mathbb{R}^d$ were the parameters to be determined according to the following optimization problem

$$\min_{b,\mathbf{w}} \sum_{i=1}^{N} \rho_\tau(y_i - b - \mathbf{w}^T\mathbf{x}_i). \tag{10}$$

Based on the above idea, Taylor [25] detailed algorithm based on LALSR. But the linear functional form fails to handle many nonlinear features in financial time series. The motivation of this paper is to define a nonlinear generalization of Taylor [25].

3. GALSR method

In this section we will apply kernel trick to define a nonlinear generalization of LALSR [25]. We will detail how to transform the estimation problem to a quadratic program. The complete algorithm is detailed in Section 3.6.