Neural Networks 67 (2015) 140–150

Contents lists available at ScienceDirect

Neural Networks

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

Incremental learning for ν-Support Vector Regression

Bin Gu [a,b,c,d,*], Victor S. Sheng [e], Zhijie Wang [f], Derek Ho [g], Said Osman [h], Shuo Li [f,d]

[a] *Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science & Technology, Nanjing, PR China*
[b] *Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, PR China*
[c] *School of Computer & Software, Nanjing University of Information Science & Technology, Nanjing, PR China*
[d] *Department of Medical Biophysics, University of Western Ontario, London, Ontario, Canada*
[e] *Department of Computer Science, University of Central Arkansas, Conway, AR, USA*
[f] *GE Health Care, London, Ontario, Canada*
[g] *Victoria Hospital, London Health Science Center, London, Ontario, Canada*
[h] *St. Joseph's Health Care, London, Ontario, Canada*

ARTICLE INFO

Article history:
Received 19 May 2014
Received in revised form 28 December 2014
Accepted 20 March 2015
Available online 6 April 2015

Keywords:
Incremental learning
Online learning
ν-Support Vector Regression
Support vector machine

ABSTRACT

The ν-Support Vector Regression (ν-SVR) is an effective regression learning algorithm, which has the advantage of using a parameter ν on controlling the number of support vectors and adjusting the width of the tube automatically. However, compared to ν-Support Vector Classification (ν-SVC) (Schölkopf et al., 2000), ν-SVR introduces an additional linear term into its objective function. Thus, directly applying the accurate on-line ν-SVC algorithm (AONSVM) to ν-SVR will not generate an effective initial solution. It is the main challenge to design an incremental ν-SVR learning algorithm. To overcome this challenge, we propose a special procedure called *initial adjustments* in this paper. This procedure adjusts the weights of ν-SVC based on the Karush–Kuhn–Tucker (KKT) conditions to prepare an initial solution for the incremental learning. Combining the *initial adjustments* with the two steps of AONSVM produces an exact and effective incremental ν-SVR learning algorithm (INSVR). Theoretical analysis has proven the existence of the three key inverse matrices, which are the cornerstones of the three steps of INSVR (including the *initial adjustments*), respectively. The experiments on benchmark datasets demonstrate that INSVR can avoid the infeasible updating paths as far as possible, and successfully converges to the optimal solution. The results also show that INSVR is faster than batch ν-SVR algorithms with both cold and warm starts.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

In real-world regression tasks, such as time-series prediction (e.g. Cao and Tay (2003); Lu, Lee, and Chiu (2009)), training data is usually provided sequentially, in the extreme case, one example at a time, which is an online scenario (Murata, 1998). Batch algorithms seems computationally wasteful as they retrain a learning model from scratch. Incremental learning algorithms are more capable in this case, because the advantage of the incremental learning algorithms is that they incorporate additional training data without re-training the learning model from scratch (Laskov et al., 2006).

ν-Support Vector Regression (ν-SVR) (Schölkopf, Smola, Williamson, & Bartlett, 2000) is an interesting Support Vector Regression (SVR) algorithm, which can automatically adjust the parameter ϵ of the ϵ-insensitive loss function.[1] Given a training sample set $T = \{(x_1, y_1), \ldots, (x_l, y_l)\}$ with $x_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$,

* Corresponding author at: Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science & Technology, Nanjing, PR China, and School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, PR China.
E-mail addresses: jsgubin@nuist.edu.cn (B. Gu), ssheng@uca.edu (V.S. Sheng), zhijie@ualberta.ca (Z. Wang), derek.ho@lhsc.on.ca (D. Ho), sidosman@hotmail.com (S. Osman), Shuo.Li@ge.com (S. Li).

http://dx.doi.org/10.1016/j.neunet.2015.03.013
0893-6080/© 2015 Elsevier Ltd. All rights reserved.

[1] The ϵ-insensitive loss function used in SVR is defined as $|y - f(x)|_\epsilon = \max\{0, |y - f(x)| - \epsilon\}$ for a predicted value $f(x)$ and a true output y, which does not penalize errors below some $\epsilon > 0$, chose a priori. Thus, the region of all (x, y) with $|\{y - f(x)| \le \epsilon\}$ is called ϵ-tube (see Fig. 1).

Schölkopf et al. (2000) considered the following primal problem:

$$\begin{aligned} \min_{w,\epsilon,b,\xi_i^{(*)}} \quad & \frac{1}{2}\langle w, w\rangle + C \cdot \left(\nu\epsilon + \frac{1}{l}\sum_{i=1}^{l}\left(\xi_i + \xi_i^*\right)\right) \\ s.t. \quad & (\langle w, \phi(x_i)\rangle + b) - y_i \leq \epsilon + \xi_i, \\ & y_i - (\langle w, \phi(x_i)\rangle + b) \leq \epsilon + \xi_i^*, \\ & \xi_i^{(*)} \geq 0, \qquad \epsilon \geq 0,\ i = 1, \ldots, l. \end{aligned} \tag{1}$$

The corresponding dual is:

$$\begin{aligned} \min_{\alpha,\alpha^*} \quad & \frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i, x_j) - \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)y_i \\ s.t. \quad & \sum_{i=1}^{l}\left(\alpha_i^* - \alpha_i\right) = 0, \qquad \sum_{i=1}^{l}\left(\alpha_i^* + \alpha_i\right) \leq C\nu, \\ & 0 \leq \alpha_i^{(*)} \leq \frac{C}{l},\ i = 1, \ldots, l \end{aligned} \tag{2}$$

where, following Schölkopf et al. (2000), training samples x_i are mapped into a high dimensional reproducing kernel Hilbert space (RKHS) (Schölkopf & Smola, 2001) by the transformation function ϕ. $K(x_i, x_j) = \langle\phi(x_i), \phi(x_j)\rangle$, $\langle\cdot,\cdot\rangle$ denotes inner product in RKHS. $^{(*)}$ is a shorthand implying both the variables with and without asterisks. C is the regularization constant, and ν is the introduced proportion parameter with $0 \leq \nu \leq 1$, which lets one control the number of support vectors and errors. To be more precise, they proved that ν is an upper bound on the fraction of margin errors, and a lower bound on the fraction of support vectors. In addition, with probability 1, asymptotically, ν equals both fractions.

Compared with ϵ-Support Vector Regression (ϵ-SVR) (Smola & Schölkopf, 2003), ν-SVR introduces two complications: the first one is that the box constraints are related to the size of the training sample set, and the second one is that one more inequality constraint is introduced in the formulation. Compared with ν-Support Vector Classification (ν-SVC) (Schölkopf et al., 2000), ν-SVR introduces an additional linear term into the objective function of (2). To sum up, the formulation of ν-SVR is more complicated than the formulations of ϵ-SVR and ν-SVC.

Early studies about SVR mostly focus on solving large-scale problems. For example, Chang and Lin (2001, 2002) gave SMO algorithm and implementation for training ϵ-SVR. Tsang, Kwok, and Zurada (2006) proposed core vector regression for training very large regression problems. Shalev-Shwartz, Singer, Srebro, and Cotter (2011) proposed stochastic sub-gradient descent algorithm with explicit feature mapping for training ϵ-SVR. Ho and Lin (2012) and Wang and Lin (2014) proposed coordinate descent algorithm for linear L1 and L2 SVR. Due to the complications in the formulation of ν-SVR as mentioned above, there are still no effective methods proposed for solving incremental ν-SVR learning.

Let us pay our attention to the exact incremental and decremental SVM algorithm (Cauwenberghs & Poggio, 2001) (hereinafter referred to as the C&P algorithm). Since the C&P algorithm was proposed by Cauwenberghs and Poggio in 2001, further studies mainly focus on two aspects. One is focusing on the C&P algorithm itself. For example, Gu, Wang, and Chen (2008) and Laskov et al. (2006) provided more detailed theoretical analysis for it. Gâlmeanu and Andonie (2008) addressed some implementation issues. Karasuyama and Takeuchi (2010) proposed an extension version which can update multiple samples simultaneously. The other applies the C&P algorithm to solve other problems. For example, Gretton and Desobry (2003) and Laskov et al. (2006) applied it to implementing an incremental one-class SVM algorithm. Martin (2002) and Ma, Theiler, and Perkins (2003) introduced it to ϵ-SVR (Vapnik, 1998) and developed an accurate on-line support vector regression (AOSVR). Recently, Gu et al. (2012) introduced the C&P algorithm to ν-SVC and proposed an effective accurate on-line ν-SVC algorithm (AONSVM), which includes the *relaxed adiabatic incremental adjustments* and the *strict restoration adjustments*. Further, Gu and Sheng (2013) proved the feasibility and finite convergence of AONSVM. Because great resemblance exists in ν-SVR and ν-SVC, in this paper, we wish to design an exact and effective incremental ν-SVR algorithm based on AONSVM.

As ν-SVR has an additional linear term in the objective function compared with ν-SVC, directly applying AONSVM to ν-SVR will not generate an effective initial solution for the incremental ν-SVR learning. To address this issue, we propose a new incremental ν-SVR algorithm (collectively called INSVR) based on AONSVM. In addition to the basic steps of AONSVM (i.e., the relaxed adiabatic incremental adjustments and the strict restoration adjustments), INSVR has an especial adjusting process (i.e. *initial adjustments*), which is used to address the complications of the ν-SVR formulation and to prepare the initial solution before the incremental learning. Through theoretical analysis, we can show the existence of the three key inverse matrices, which are the cornerstone of the *initial adjustments*, the *relaxed adiabatic incremental adjustments*, and the *strict restoration adjustments*, respectively. The experiments on benchmark datasets demonstrate that INSVR can avoid the infeasible updating path as far as possible, and successfully converge to the optimal solution. The results also show that INSVR is faster than batch ν-SVR algorithms with both cold and warm starts.

The rest of this paper is organized as follows. In Section 2, we modify the formulation of ν-SVR and give its KKT conditions. The INSVR algorithm is presented in Section 3. The experimental setup, results and discussions are presented in Section 4. The last section gives some concluding remarks.

Notation: To make the notations easier to follow, we give a summary of the notations in the following list.

α_i, g_i	The ith element of the vector α and g.
α_c, y_c, z_c	The weight, output, and label of the candidate extended sample (x_c, y_c, z_c).
Δ	The amount of the change of each variable.
$\boxed{\epsilon'}, \boxed{\Delta\epsilon'}$	If $\lvert\sum_{i\in S_S} z_i\rvert = \lvert S_S\rvert$, $\boxed{\epsilon'}$ and $\boxed{\Delta\epsilon'}$ stands for ϵ' and $\Delta\epsilon'$, respectively. Otherwise, they will be ignored.
$Q_{S_S S_S}$	The submatrix of Q with the rows and columns indexed by S_S.
$\widetilde{Q}_{\setminus\widetilde{M}^2}$	The submatrix of $\widetilde{Q}$ with deleting the rows and columns indexed by $\widetilde{M}$.
$\check{R}_{t*}, \check{R}_{*t}$	The row and the column of a matrix $\check{R}$ corresponding to the sample (x_t, y_t, z_t), respectively.
$\mathbf{0}, \mathbf{1}$	The vectors having all the elements equal to 0 and 1, respectively, with proper dimension.
$\mathbf{z}_{S_S}, \mathbf{u}_{S_S}$	A $\lvert S_S\rvert$-dimensional column vector with all equal to z_i, and $z_i y_i$ respectively.
$\det(\cdot)$	The determinant of a square matrix.
$\mathrm{cols}(\cdot)$	The number of columns of a matrix.
$\mathrm{rank}(\cdot)$	The rank of a matrix.

2. Modified formulation of ν-SVR

Obviously, the correlation between the box constraints and the size of the training sample set makes it difficult to design an incremental ν-SVR learning algorithm. To obtain an equivalent formulation, whose box constraints are independent to the size of the training sample set, we multiply the objective function of (1) by the size of the training sample set. Thus, we consider the following primal problem:

متن کامل مقاله ← دریافت فوری

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات