文章编号：2095-2163(2020)12-0028-04 中图分类号：TP181 文献标志码：A

朴素支持张量机的算法研究综述

王 君，杜金星，麻安鹏，杨本娟
（贵州师范大学 数学科学学院，贵阳 550025）

摘 要： 大部分数据的自然表示形式是向量、矩阵或者更高维的数据，支持向量机可以较好地处理向量形式的数据，但是对于高维数据，传统的机器学习算法在将多维数据转化成向量形式时会损失大量的结构信息。因此，研究者提出朴素支持张量机这一类分类器，将多维数据输入进行训练，利用 SMO 算法求解。其中利用 CP 分解、Tucker 分解或者张量核函数等来获取数据的结构信息，这样不但能够获取数据的大部分信息，还可以节省时间成本，减少计算量，又可以求得凸优化函数的全局最优解。本文对这一类分类器的一个算法研究综述，同时指出了算法的优缺点和未来发展的方向。

关键词： 高维数据；朴素支持张量机；CP 分解；Tucker 分解；张量核

A review of algorithms for Naive - supported tensor machines

WANG Jun, DU Jinxing, MA Anpeng, YANG Benjuan
(School of Mathematical Sciences, Guizhou Normal University, Guiyang 550025, China)

【Abstract】 Most of the natural representation forms of data are vectors, matrices or higher-dimensional data. Support vector machines can handle the data in vector form better, but for high-dimensional data, traditional machine learning algorithms will lose a lot of structural information when transforming multidimensional data into vector form. Therefore, the researchers proposed a classifier such as naive support tensor, trained multidimensional data input, and solved it by SMO algorithm, in which CP decomposition, Tucker decomposition or tensor kernel function were used to obtain the structural information of the data. In this way, most of the information of the data can be obtained, the time cost can be saved, the computation can be reduced, and the global optimal solution of convex optimization function can be obtained. This paper is an overview of the algorithm research of this kind of classifier, and points out the advantages and disadvantages of the algorithm and the future development direction.

【Key words】 High-dimensional data; Naive support tensor; CP decomposition; Tucker decomposition; Tensor kernel

0 引 言

支持向量机是一种在统计学习理论基础上发展来的机器学习方法，具有牢固的理论基础、全局最优、解的稀疏性、非线性和泛化等优点。到目前为止支持向量机已经发展得比较完善，现在研究者已经把关注点集中在分类数据的多维性，保持数据的空间结构性等方面。

2005 年，TAO 等研究者就提出可以将基于向量的机器学习方法扩展到张量空间[1]；也有研究者提出了关于矩阵数据的机器学习算法，比如 Cai 等人提出了支持矩阵机，将权重参数定义为 $W \approx u \times v^{\mathrm{T}}$，支持向量机各特征之间是相互独立的，而支持矩阵机认为各个特征之间是线性相关的[2-3]；在 2007 年，T.G. Kolda 和 B.W. Bader 详细介绍了张量分解和应用，包括张量的基本概念、运算法则及张量分解，使得研究者们对张量有了更加深刻的理解[4]；受到支持矩阵机的启发，Tao 等研究者首次提出了分类监督张量框架（STL），其是凸优化和多线性代数运算的结合，并提出了二类支持张量机算法（经典支持张量机算法），通过迭代投影优化算法得出对应的局部最优解，这里的权重张量被定义为秩一张量[5]。但是把张量权重限制在 cp 秩意义下的秩为一的张量，对张量数据有很大的局限性，2012 年 Kotsia 等研究者们提出了高秩支持张量机，对权值张量进行高秩的 cp 分解，首次提出了高秩 STMs、高秩 Σ/ΣwSTMs 和高秩 RMSTMs[6]，这里的权值张量通过 cp 分解，被定义为秩一张量的和，利用迭代的方法对其求解，每次迭代对应一个张量模态投影，而通过 svm 形式的优化问题求解迭代投影，即基于张量的算法是把总问题分解成许多小而简单的问题，

基金项目： 贵州师范大学博士启动项目（085185740001）。

作者简介： 王 君（1995-），女，硕士研究生，主要研究方向：图像处理与机器学习；杜金星（1993-），女，硕士研究生，主要研究方向：图像处理与机器学习；麻安鹏（1995-），女，硕士研究生，主要研究方向：图像处理与机器学习；杨本娟（1982-），女，博士，副教授，主要研究方向：图像处理与模式识别。

通讯作者： 杨本娟 Email：bj.yang@ hotmail.com

收稿日期： 2020-10-19

每一个小问题都被典型的定义在某一个张量模态。2011 年 Kotsia 提出了支持 tucker 机。

支持张量机需要通过迭代投影优化的算法求解，计算量很大，而且求得的解是局部最优解，不是全局最优解。针对这样的问题，有研究者提出了朴素支持张量机，即张量输入到 svm，利用 smo 算法求解。2013 年，Hao 等研究者提出了一个线性高阶支持张量机算法，将张量数据 cp 分解求得内积，再根据 smo 算法求得决策函数[7]。由于 tucker 分解是 cp 分解的一般化形式，且 CP 分解需要提前给出张量的秩才能分解，因此在 2019 年李云雷提出基于 tucker 分解的支持张量机[8]；吴振宇等研究者又相继提出基于 tucker 分解的半监督支持张量机，对张量数据进行 tucker 分解求取内积，然后利用 smo 算法求解[9]。

监督张量模型需要大量的带标签的训练样本，但收集这些样本需要高劳动力和时间成本，在真实世界中提供足够的标记训练数据是不现实的。Miller DJ 等研究者从数据分布估计分析得出结论，分类器的泛化能力可以随着未标记数据的显著增加而提高[10]。目前半监督支持向量机已经发展得比较完善，Shifei Ding 等研究者又发表了关于半监督支持向量机的研究综述，但是向张量形式的扩展还比较少[11]。

之前的研究都集中在张量数据的多线性关系，但真实样本数据的潜在结构是非线性的，后面的研究者开始运用核函数或提出张量核来挖掘数据的非线性结构。在 2014 年 He L 等研究者提出 DUSK 算法，在 linear SHTM 基础上的非线性扩展，可以解决复杂的非线性数据[12]；也有研究者提出张量核，挖掘数据的非线性关系[13]。

1　经典支持张量机模型

支持张量机是支持向量机的多维扩展，SVM 是向量输入，STM 是张量输入，把张量转化成向量的形式忽视了数据的原始结构信息，而输入张量数据却保持了数据的原始结构信息，支持向量机是要最大化支持向量到超平面的几何距离，支持张量机同样如此，只是向多维扩展。经典支持张量机模型如式(1)：

$$\begin{cases} \min\limits_{\vec{w}_k \mid \prod_{k=1}^{M}, b, \vec{\xi} \geqslant \frac{1}{2} \mid} \| \bigotimes_{k=1}^{M} w_k \|_{Fro}^2 + C \sum\limits_{i=1}^{N} \xi_i \\ \text{s.t. } y_i [X_i \prod\limits_{k=1}^{M} \times_k \vec{w}_k + b] \geqslant 1 - \xi_i, \xi_i \geqslant 0, 1 \leqslant i \leqslant N. \end{cases} \tag{1}$$

通过迭代投影的方法得到权值张量 W_k 和偏置项 b。

2　朴素支持张量机

2.1　朴素支持张量机的基本模型

朴素支持张量机是在经典 SVM 的理论基础上，把输入数据用张量表示，然后直接处理张量数据，利用 SMO 算法求出最终的最优分类器。

假设有 N 个训练 $X_i \in R^{I_1 \times I_2 \times \cdots \times I_M} (1 \leqslant i \leqslant N)$ 和对应的类标签 $y_i \in \{-1, 1\}$，朴素支持张量机要找到最优的投影向量 $W \in R^{I_1 \times I_2 \times \cdots \times I_M}$ 和偏项 $b \in R$，可以通过解决下面的凸优化问题，式(2)：

$$\begin{cases} \min\limits_{\vec{w}, b, \vec{\xi} \geqslant 0} \dfrac{1}{2} \| W \|_{Fro}^2 + C \sum\limits_{i=1}^{N} \xi_i \\ \text{s.t. } y_i [\vec{W}^T \vec{x}_i + b] \geqslant 1 - \xi_i, \xi_i \geqslant 0, 1 \leqslant i \leqslant N. \end{cases} \tag{2}$$

其中，$\vec{\xi} = [\xi_1, \xi_2, \cdots \xi_N] \in R^N$，是在解决线性不可分问题时所有松弛变量组成的向量，ξ_i 是第 i 个样本的边界误差，当解决的是线性可分的样本时，则 $\vec{\xi} = 0$。

通过把上面的凸优化问题添加对偶算子，转化成对偶函数如式(3)：

$$\begin{cases} \min \dfrac{1}{2} \sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} \alpha_{i\alpha j} y_i y_j < x_i, x_j > - \sum\limits_{i=1}^{N} \alpha_i \\ \text{s.t. } \sum\limits_{i=1}^{N} \alpha_i y_i = 0, 0 \leqslant \alpha_i \leqslant C, i = 1, 2, \cdots, N \end{cases} \tag{3}$$

最终得到分类函数，式(4)：

$$f(x) = \langle \omega, X \rangle + b = \langle \sum_{i=1}^{N} \alpha_i y_i X_i, X \rangle + b. \tag{4}$$

但是在这种情况下，输入数据用张量表示求出的结果和经典的 SVM 算法是相同的，因此没有任何意义。后面的研究者提出了利用 CP 分解和 tucker 分解以及张量核的方式来获取数据的原始结构信息和多线性关系，使得朴素支持张量机有了意义。

2.2　获取支持张量机的结构性信息模型

CP 分解和 Tucker 分解应用的范围很广，Tucker 分解可以用于提取数据特征(降维)、去除噪声、图像恢复、数据压缩和数据存储等，CP 分解可以用于数据去燥等。最经典的朴素支持张量机模型是利用 CP 分解和 Tucker 分解来获取数据的结构信息。下面介绍它的基本模型：

(1)以 CP 分解为核心，获取数据结构性的支持张量机，式(5)。

$$\langle X_i, X_j \rangle = \langle \sum_{p=1}^{R} \bigotimes_{i=1}^{M} x_{ip}, \sum_{q=1}^{R} \bigotimes_{j=1}^{M} x_{jq} \rangle = \sum_{p=1}^{R} \sum_{q=1}^{R} \prod_{n=1}^{N} \langle x_{ip}, x_{jq} \rangle. \tag{5}$$

(2)以 Tucker 分解为核心，获取数据结构性的

支持张量机,式(6):

$$\langle \mathcal{X}_i, \mathcal{X}_j \rangle = \langle G_i \bigotimes_{p=1}^{M} U_{ip}, G_j \bigotimes_{q=1}^{M} U_{jq} \rangle. \quad (6)$$

公式(5)是对输入张量数据进行 CP 分解求取数据的内积。2013 年,Hao 等研究者提出了一个线性高阶支持张量机算法,就是运用公式(5)求取内积,再利用 SMO 算法优化求解,这是线性的算法,只是获取了结构信息。

公式(6)是对输入张量数据进行 Tucker 分解,求取数据的内积。2019 年吴振宇等研究者提出基于 tucker 分解的半监督支持张量机,对张量数据进行 tucker 分解求取内积,再利用 SMO 算法优化求解。

2.3 获取支持张量机的非线性关系模型

之前的工作大多是利用线性的决策边界分离数据,但实际上真实数据的潜在结构是非线性的,因此利用 cp 分解等多线性分解方法来近似张量数据复杂的非线性结构是很困难的,而且也不能完全的分离数据。因此,后续的研究者提出了利用张量核,或添加核函数来获取数据的非线性关系。

2014 年,He L 等研究者提出 DUSK 算法,是在 linear SHTM 基础上的非线性扩展,先利用非线性的投影把输入的数据投影到更高维的特征空间,再利用公式(5)对数据进行 CP 分解,这样就保存了数据的原始结构信息,又捕捉到了数据的非线性关系。

加入 RBF 核,式(7):

$$K\langle \mathcal{X}_i, \mathcal{X}_j \rangle = K\langle \sum_{p=1}^{R} \bigotimes_{i=1}^{M} x_{ip}, \sum_{q=1}^{R} \bigotimes_{j=1}^{M} x_{jq} \rangle = \sum_{p=1}^{R} \sum_{q=1}^{R} \exp\left(-\sigma \sum_{n=1}^{M} \| x_{ip} - x_{jq} \|^2\right). \quad (7)$$

2015 年,BogusławCyganek 等研究者提出了利用弦距离核的支持向量机对多维数据进行分类,即在 SVM 上利用弦距离核对图像分类,弦距离核如式(8):

$$K\langle A, B \rangle = \prod_{j=1}^{M} K_j \langle A, B \rangle = \prod_{j=1}^{M} \exp\left(-\frac{1}{2\sigma^2} \| D_{A,1}^{j} D_{A,1}^{j\,\mathrm{T}} - D_{B,1}^{j} D_{B,1}^{j\,\mathrm{T}} \|_F^2\right). \quad (8)$$

3 朴素支持张量机的优点和缺点

就朴素支持张量机本身来说是没有意义的,但是只要利用 CP 分解、Tucker 分解、或者利用张量核等算法来获取数据的结构信息和非线性关系,这就使得其具有了意义。

3.1 优点

能够获取数据的大部分信息,而且这样的方法既可以节省时间成本、减少计算量,又可以求得凸优化函数的全局最优解。经典支持张量机算法需要通过迭代投影方法求得局部最优解,但是运用朴素支持张量机的算法根据 CP 分解、Tucker 分解、张量核等算法求得了张量内积后,就可以用 SMO 算法调用 libsvm 工具包求得全局最优解,既张量形式表示了数据,又节省了时间成本,减少了计算量。

3.2 缺点

运用 SMO 算法求解时,涉及到求核矩阵。每次模型训练计算训练集和测试集核矩阵时,都需要所有数据计算,这就使得在数据量过大时,如果要更新数据(加入新的数据或减小数据),每次训练模型都要计算所有的数据,会造成很大的计算量。

4 结束语

本文对朴素支持张量机算法发展过程做了研究与总结,了解到朴素支持张量机本身来说是没有意义的,但是只要利用 CP 分解、Tucker 分解、或者利用张量核等算法来获取数据的结构信息和非线性关系,这就使其具有了意义。未来的工作可以向这 3 个方向发展:

(1)朴素支持张量机可以向在线支持张量机扩展,主要是针对于数据更新的情况;

(2)数据在不断地更新,要对图像分类就必须用到大量带标签的数据,给数据标标签需要耗费大量的时间与金钱,而且还存在标注误差。获取支持张量机的非线性关系模型还没向半监督扩展,之前的研究大多数都是监督学习方式,为了符合社会发展需要,会向半监督方向扩展;

(3)线性支持张量机不能完全的分离数据,会向非线性方向扩展。

参考文献

[1] TAO D, LI X, WU X, et al. Supervised tensor learning[J]. Knowledge and Information Systems, 2005, 13(1):450-457.

[2] LUO L, XIE Y, ZHANG Z, et al. Support matrix machines[C]//International conference on machine learning, 2015: 938-947.

[3] KOLDA T G, BADER B W. Tensor Decompositions and Applications[J]. Siam Review, 2009, 51(3):455-500.

[4] Cai, Deng, He, Xiaofei, Wen, Jirong, et al. Support Tensor Machines for Text Categorization[J]. International Journal of Academic Research in Business & Social Sciences, 2006, 2(12): 2222-6990.

[5] D Tao, X Li, W Hu, S Maybank, and X Wu. Supervised Tensor Learning[C]//Knowledge and Information Systems, 2007:1-42.

[6] KOTSIA I, GUO W, PATRAS I. Higher rank Support Tensor Machines for visual recognition[J]. Pattern Recognition, 2012, 45(12):4192-4203.

[7] HAO Z, HE L, CHEN B, et al. A Linear Support Higher-Order

Tensor Machine for Classification[J]. IEEE Transactions on Image Processing,2013,22(7):2911-2920.
[8] 李云雷. 基于 HOG 与 STM 的行人检测系统[D]. 大连:大连理工大学.2019.
[9]吴振宇,李云雷,吴凡. 基于 Tucker 分解的半监督支持张量机[J]. 计算机科学,2019,46(9):195-200.
[10]MILLER D J,UYAR H S. A mixture of experts classifier with learning based on both labelled and unlabelled data[C]//Advances in neural information processing systems. 1997: 571-577.
[11]DING S, ZHU Z, ZHANG X. An overview on semi-supervised support vector machine[J]. Neural Computing & Applications, 2015, 28(5):1-10.
[12] HE L, KONG X, YU P S, et al. Dusk: a dual structure-preserving kernel for supervised tensor learning with applications to neuroimages[C]//Proceedings of the 2014 SIAM international conference on data mining,2014:127-135.
[13] Bogusław Cyganek, Bartosz Krawczyk, Michał Wozniak. Multidimensional data classification with chordal distance based kernel and Support Vector Machines[J]. Engineering Application of Artificial Intelligence.2015,46:10-22.

(上接第 27 页)

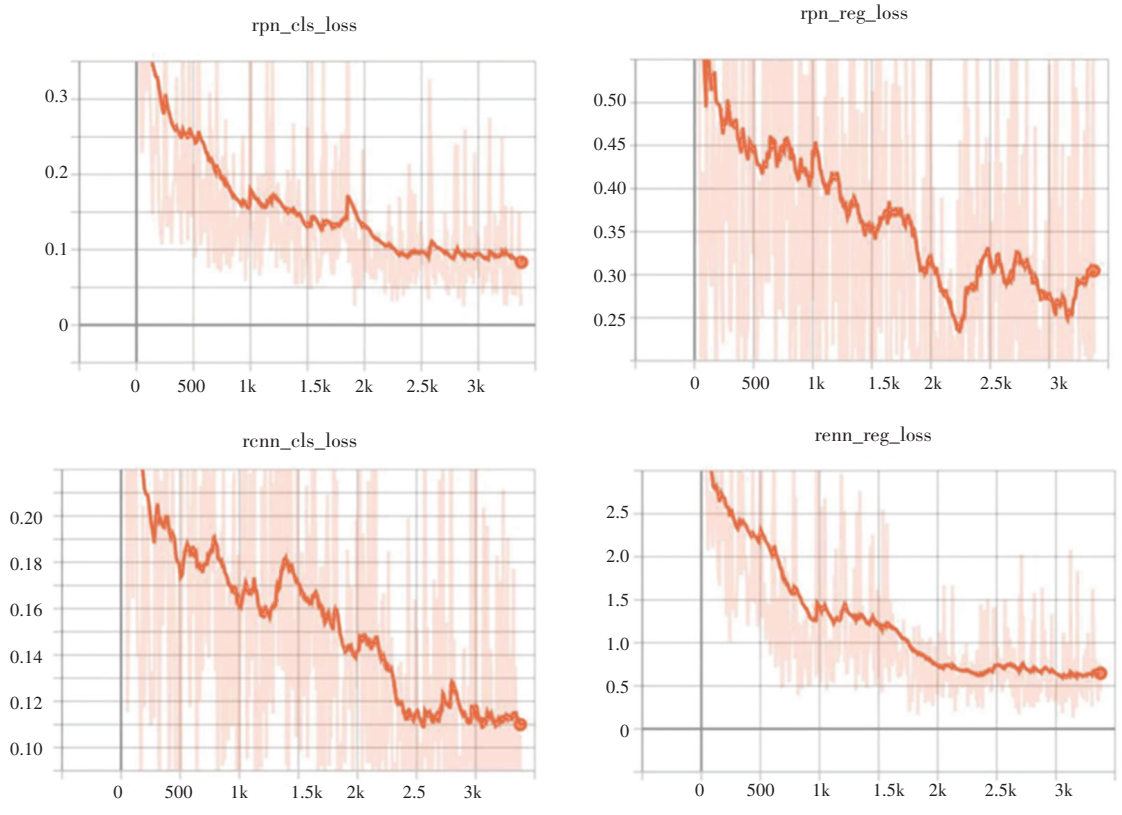

图 5　训练 Loss 损失图

Fig. 5　Training loss figure

参考文献

[1] 张远. 印刷体文档表格识别技术研究[D]. 长沙. 湖南大学,2018.
[2] 王文华. 浅谈 OCR 技术的发展和应用[J]. 福建电脑,2012,(6):56-56.
[3] 司明. 表格识别的研究[D]. 西安.西安科技大学,2009.
[4] 潘军. 复杂表格文档预处理与文本提取算法研究[D]. 北京.北京交通大学,2017.
[5] 卞静潇. 复杂版面文档图像表格与图的提取及分析[D]. 西安.西安电子科技大学,2015.
[6] 常亮,邓小明,周明全,等. 图像理解中的卷积神经网络[J]. 自动化学报, 2016, 42(9):1300-1312
[7] 李青宇. 快速高效的深度神经网络目标检测方法研究[D]. 北京.北京交通大学,2019.
[8] 郭佳. 基于图像的表格识别算法与自动录入系统[D]. 北京.北京邮电大学,2018.
[9] 张昊玥. 非结构化文档的版面分析及表格提取[D]. 北京.北京交通大学,2019.
[10]朱维松. 基于距离变换的纤维骨架提取算法研究[D]. 上海.东华大学,2008.
[11]Ren Shaoqing,He Kaiming, et al.Faster R-CNN:Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(6):1137-1149.
[12]黄继鹏,史颖欢,高阳. 面向小目标的多尺度 Faster-RCNN 检测算法[J]. 计算机研究与发展, 2019, 56(2):319-327.