Support Tensor Machines

Tiago Filipe Sousa Gonçalves

9^{th} December 2018

1 Assignment

Support Vector Machines (SVM) is one of the most representative and successful classification algorithms [11], achieving a huge success in pattern classification. The problem, however, relies on the fact that a regular SVM model, since it is based on vector inputs, can not directly handle matrices or higher dimensional data structures, i.e. tensors, which happen to be very common in real world problems [1]. A tensor is a multi-dimensional array that generalizes matrix representation and its dimension is named as mode or way [4, 3]. Usually, when dealing with high dimensional data inputs, SVM need each sample to be reshaped into a vector, but, if the training data size is relatively small (when compared with the feature vector dimension), it may show poor classification performance due to overfitting [12, 10, 7]. Bearing this in mind, and taking into account the potentiality that tensor data structures have in machine learning applications [2, 8], it was just a matter of time until the extension of standard SVM to tensor formulations was achieved, leading, thus, to significant performance improvements [9, 6]: it was the beginning of Support Tensor Machines (STM). For this to be accomplished, there was a need to adapt this supervised tensor learning methodology, by changing vector inputs into tensor inputs and decomposing the correspondent weight vector into a rank-1 tensor, which is trained by the alternating projection method [9]; based on this approach, it was possible to extend the linear SVM to a general tensor form, i.e. STM [9]. These preliminary achievements lead researchers to think that problems such as the *curse of dimensionality* and the overfitting problem in traditional SVM was solved; despite solving the overfitting problem, STM were showing poor classification accuracy and further studies revealed that *curse of dimensionality*, related to the tunning of several parameters was still a issue [6, 5]. It can be concluded, however, that, when one is dealing with data where the number of attributes surpasses the number of samples, using STM would be a valid solution that could lead to good results, and surpass some of the obstacles related with SVM. Actually, more recently, in order to overcome STM problems, a new concept, called Suport Tensor Train Machine (STTM) has been introduced by Chen *et al* [1].

References

[1] Cong Chen, Kim Batselier, Ching-Yun Ko, and Ngai Wong. A support tensor train machine. *arXiv preprint arXiv:1804.06114*, 2018.

[2] Zhongming Chen, Kim Batselier, Johan AK Suykens, and Ngai Wong. Parallelized tensor train learning of polynomial classifiers. *IEEE Transactions on Neural Networks and Learning Systems*, 2017.

[3] Lifang He, Chun-Ta Lu, Guixiang Ma, Shen Wang, Linlin Shen, S Yu Philip, and Ann B Ragin. Kernelized support tensor machines. In *International Conference on Machine Learning*, pages 1442–1451, 2017.

[4] Tamara G Kolda and Brett W Bader. Tensor decompositions and applications. *SIAM review*, 51(3):455–500, 2009.

[5] Irene Kotsia, Weiwei Guo, and Ioannis Patras. Higher rank support tensor machines for visual recognition. *Pattern Recognition*, 45(12):4192–4203, 2012.

[6] Irene Kotsia and Ioannis Patras. Support tucker machines. In *CVPR*, pages 633–640. Citeseer, 2011.

[7] Jing Li, Nigel Allinson, Dacheng Tao, and Xuelong Li. Multitraining support vector machine for image retrieval. *IEEE Transactions on Image Processing*, 15(11):3597–3601, 2006.

[8] Alexander Novikov, Dmitrii Podoprikhin, Anton Osokin, and Dmitry P Vetrov. Tensorizing neural networks. In *Advances in Neural Information Processing Systems*, pages 442–450, 2015.

[9] Dacheng Tao, Xuelong Li, Weiming Hu, Stephen Maybank, and Xindong Wu. Supervised tensor learning. In *Data Mining, Fifth IEEE International Conference on*, pages 8–pp. IEEE, 2005.

[10] Dacheng Tao, Xiaoou Tang, Xuelong Li, and Xindong Wu. Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval. *IEEE transactions on pattern analysis and machine intelligence*, 28(7):1088–1099, 2006.

[11] Vladimir Vapnik. *The nature of statistical learning theory*. Springer science & business media, 2013.

[12] Shuicheng Yan, Dong Xu, Qiang Yang, Lei Zhang, Xiaoou Tang, and Hong-Jiang Zhang. Multilinear discriminant analysis for face recognition. *IEEE Transactions on Image Processing*, 16(1):212–220, 2007.