Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: https://www.tandfonline.com/loi/tstf20

Personalised medicine with multiple treatments: a PhD thesis abstract

Zhilan Lou

To cite this article: Zhilan Lou (2017) Personalised medicine with multiple treatments: a PhD thesis abstract, Statistical Theory and Related Fields, 1:2, 182-184, DOI: 10.1080/24754269.2017.1396426

To link to this article: https://doi.org/10.1080/24754269.2017.1396426

Published online: 08 Nov 2017.

STATISTICAL THEORY AND RELATED FIELDS, 2017
VOL. 1, NO. 2, 182–184
https://doi.org/10.1080/24754269.2017.1396426

Taylor & Francis
Taylor & Francis Group

Personalised medicine with multiple treatments: a PhD thesis abstract

Zhilan Lou

School of Statistics, East China Normal University, Shanghai, China

ABSTRACT

When there is substantial heterogeneity of treatment effectiveness for comparative treatment selection, it is crucial to identify individualised treatment rules for patients who have heterogeneous responses to treatment. Existing approaches include directly modelling clinical outcome by defining the optimal treatment rule according to the interactions between treatment and covariates and outcome weighted approach that uses clinical outcome as weights to maximise a target function whose value directly reflects correct treatment assignment. All existing articles of estimating individualised treatment rules are all assuming just two treatment assignments. Here we propose an outcome weighted learning approach that uses a vector hinge loss to extend estimating individualised treatment rules in multi-category treatments case. The consistency of the resulting estimator is shown. We also demonstrate the performance of our approach in simulation studies and a real data analysis.

ARTICLE HISTORY
Received 7 October 2017
Accepted 21 October 2017

KEYWORDS
Heterogeneity of treatment effectiveness; individualised treatment rule; risk bound; RKHS; weighted multi-category support vector machine

Personalised medicine provides the right treatment to the right individual patient according to patient characteristics such as demographics, genomic information, treatment and outcome history, and so on; patients with different characteristics have significant heterogeneity in their responses to treatments. Thus, it becomes an increasingly important research topic among clinical and intervention scientists in establishing an evidence-based personalised treatment assignment rule as a function of patient characteristics to optimise patient responses. The traditional approach is finding an optimal treatment assignment rule through the estimation of the response expectation conditioned on patient characteristics treated as covariates. However, the estimation of the main treatment effect (which does not affect the optimality of treatments) interferes with the estimation of covariate-treatment interaction (which affects the optimality), and hence the latter cannot be estimated accurately. Efforts have been made to separate the main effect from the covariate-treatment interaction effect, either through multiple testing strategies (Su, Tsai, Wang, Nickerson, & Li, 2009) or through prediction (Foster, Taylor, & Ruberg, 2011).

A more recent alternative approach is to by-pass the estimation of the covariate-treatment interaction and directly search an optimal treatment assignment rule by maximising the expected clinical outcome related with different treatments. This approach is named as outcome weighted learning in Zhang, Tsiatis, Davidian, Zhang, and Laber (2012) and Zhao, Zeng, Rush, and Kosorok (2012). However, almost all existing outcome weighted learning methods are for the case of two treatments. In the case of multiple treatments, although some outcome weighted learning methods can still be applied, they are not optimal and their theoretical properties such as the risk-consistency (Zhao et al., 2012) are unknown. The purpose of this PhD thesis is to develop an outcome weighted learning method for three or more treatments, and to study its theoretical and empirical properties.

The expected clinical outcome involves the 0-1 loss, which is difficult to maximise due to its discontinuity and nonconvexity. In the case of two treatments, the 0-1 loss is replaced by a convex surrogate loss, the hinge loss. We utilise a vector hinge loss used by Lee, Lin, and Wahba (2004) in multicategory support vector machine. We prove that maximising the expected clinical outcome is equivalent to minimising the risk under the convex vector hinge loss weighted by clinical outcomes. This is called Fisher consistency and justifies the validity of using the vector hinge loss.

All existing approaches so far are limited to equal losses, i.e. the misclassification costs are equal. The case of unequal losses may be encountered in real world problems, especially in medical applications. For example, if treatment A is more expensive, toxic, or laborious than treatment B, then we may only prefer treatment A when the benefit under treatment A is larger than that under treatment B to a certain factor. We develop a framework to extend the outcome weighted learning to unequal loss case, and establish the Fisher consistency when a vector hinge unequal-loss function is used.

Although the vector hinge loss is continuous and convex, the minimisation of the corresponding risk

CONTACT Zhilan Lou 52153001012@ecnu.cn

function is still quite hard in application because the loss is not smooth enough and we do not make any parametric assumption. We then adopt the idea in Lee et al. (2004), i.e. we add a penalty term in the objective risk function and restrict our solution to the Reproducing Kernel Hilbert Space (RKHS). Using a representer theorem, the original minimisation problem turns into a quadratic programming problem with some equality and inequality constraints. In practice, the quadratic programming problem can be easily carried out via available software packages.

Naturally, an important theoretical question is, as the sample size of the training data increases to infinity, whether the solution in RKHS converges to the optimal solution of the original problem that is hard to compute, or whether the risk evaluated at the solution in RKHS converges to the risk of the unknown optimal rule, which is called risk-consistency. We first establish a relationship between the excess risk under the vector hinge loss and the excess risk under the expected clinical outcome. Using this result, we prove the risk-consistency of our proposed solution for both equal and unequal losses, under some minor conditions. Our proof is more rigorous than some existing similar proofs in the following aspects. The first one is the technical requirement that the range of covariate values is compact. We show that this requirement can be relaxed by applying a one-to-one bounded transformation and our solution is invariant with this transformation. The second one is that the RKHS we use needs to be dense in the space of continuous functions, which can be achieved by using the so-called universal kernel in RKHS.

With technology advances, the number of measured covariates nowadays is often very large, even comparable with the sample size. However, the number of covariate actually related with the response is usually small. This suggests the necessity of covariate selection or screening, or dimension reduction in the process of constructing the optimal rule. Note that the set of covariates having effects on the response may be different under different treatments. Hence, it is better to perform covariate screening or dimension reduction separately in different treatments. Our method can actually handle the situation where the covariate sets in different treatments are different, and it remains to be risk-consistent if the covariate screening or the dimension reduction is consistent. Since our approach does not make any assumption on the expected response conditioned on covariates and treatment, we apply some model free feature screening procedure for both categorical and continuous covariates.

Besides the theoretical derivation, we carry out many simulation studies to evaluate the fixed sample performance of the proposed method in both low dimensional covariate and high dimensional covariate cases. Our proposed method is compared to two other methods. One method compares two treatments at a time by applying the method of Zhao et al. (2012) and then finds the winner in all paired comparisons. The other method is the weighted tree method in Zhang et al. (2012), which is proposed for two treatments but can be applied to multiple treatments. We consider two comparison criteria. One is the misclassification error rate and the other is the magnitude of the excess risk, both approximated by an independent validation data set in the simulation. We consider many scenarios in low dimensional case, including linear or nonlinear boundary, complex main effect structure, different kind of covariates, Bayes error, three or four treatments and so on. The performance of our proposed method is better and the improvement can be tremendous in some cases. Our results also show that the weighted tree method does not obtain the optimal solution. In high dimensional case, we study the performance of the proposed method after covariate screening and compare it with the oracle method assuming that we know exactly which covariates should be included or excluded. The performance of the proposed method with screening is close to the oracle method when the sample size is large. Our results also show the necessity of covariate screening. In addition, our results show that covariate screening separately under different treatments is more beneficial in our framework.

For the choice of the kernel function, we suggest a convex linear combination of a Gaussian kernel and a spline kernel. The goal is to adaptively choose a better kernel function for better empirical performance.

Finally, we apply the proposed method to a real data set from a cancer behavioural study with four treatment arms. We define a comparative treatment effect to measure the increase of average outcome when assigning patients according to the proposed rule. The m-out-of-n bootstrap method is used to construct a confidence interval for the quantity. The results show that the estimated treatment assignment rule by using our proposed method seems to be able to enhance the treatment effect.

Acknowledgements

The author would like to thank Jun Shao and Menggang Yu for their help with preparing the manuscript. This work was supported by the Chinese 111 Project [grant number B14019] (for Lou and Shao).

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

Zhilan Lou is a PhD candidate, School of Statistics, East China Normal University, Shanghai, China.

References

Foster, J. C., Taylor, J. M., & Ruberg, S. J. (2011). Subgroup identification from randomized clinical trial data. *Statistics in Medicine, 30*(24), 2867–2880.

Lee, Y., Lin, Y., & Wahba, G. (2004). Multicategory support vector machines: Theory and application to the classification of microarray data and satellite radiance data. *Journal of the American Statistical Association, 99*, 67–81.

Su, X., Tsai, C.-L., Wang, H., Nickerson, D. M., & Li, B. (2009). Subgroup analysis via recursive partitioning. *The Journal of Machine Learning Research, 10*, 141–158.

Zhao, Y., Zeng, D., Rush, A. J., & Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Stat, 107*(499), 1106–1118.

Zhang, B, Tsiatis, A., Davidian, M., Zhang, M., & Laber, E. (2012). Estimating optimal treatment regimes from a classification perspective. *Stat, 1*, 103–114.