Variable selection for dynamic treatment regimes

Raphael Fonteneau
University of Liège
raphael.fonteneau@ulg.ac.be

Louis Wehenkel
University of Liège
L.Wehenkel@ulg.ac.be

Damien Ernst
University of Liège
dernst@ulg.ac.be

Nowadays, many illnesses as for example HIV/AIDS, cancer or psychological diseases are seen by the medical community as being chronic-like diseases. For treating such diseases, physicians often adopt explicit, operationalized series of decision rules specifying how treatment level and type should vary overtime. These rules are referred to in the medical community as dynamic treatment regime or DTR in short. Designing DTR for such diseases is a challenging issue. Among the difficulties encountered, we can mention the poor compliance to treatments due to the side effects associated to some drugs (e.g., chemotherapies can decrease significantly the quality of life of some patients), the decrease of treatment efficiency with time (e.g., apparition of drug-resistant HIV viruses after several years of treatment) and the enormous cost of administrating drugs to patients over periods ranging sometimes to tens of years. To a large extent DTR are nowadays based on clinical judgment and medical instinct rather than on a formal and systematic data-driven process that could reveal itself to be more efficient. These latter ten years, one has seen the emergence among the biostatistics community of a research field addressing specifically problems of inference of DTR from clinical data. While this research field is still young, encouraging results have already been published. We mention for example [1] where the authors propose such an approach for designing treatments for psychotic patients.

One common approach in biostatistics to infer DTR from the data collected through some (randomized multi-stage) clinical trials is to formalize this inference problem as an optimal control problem for which most of the information available on the 'system dynamics' (the system is here the patient and the input of the system is the treatment) is contained in the clinical data. This problem of inference of (close-to) optimal policies from real-life data has been vastly studied in Reinforcement Learning (RL), a subfield of machine learning (see e.g., [2]). The common approach in RL is to process these data to output closed-loop policies which usually take their values on the clinical indicator space which can be large-dimensional (for example, the clinical data analysed in [1] contain around 60 indicators). Using policies outputted by these RL algorithms as such can thus be non-practical for the physicians who prefer to have DTR based on a few indicators rather than on the large set of variables monitored through the clinical trial. In this research, we address the problem of inference of good policies defined on a small subset of indicators, a problem that we have chosen to refer to as variable selection for policy representation in RL.

The problem of variable selection has already been considered by many authors in the machine learning community but mostly for Supervised Learning (SL). For example, in [3] the authors propose a Bayesian approach for selecting the most informative variables, while decision or regression trees are used in [4]. Only a handful of papers address this problem in the RL context (see e.g., [5]) and mostly for reducing the input space for RL algorithms in order to leverage their generalization capabilities rather than for finding policies defined on a small number of indicators.

Our approach for variable selection is based on a new class of RL algorithms named fitted Q iteration which reformulates the problem of inference of good policies from data as a sequence of supervised learning problems [2]. While not entering into the details, this reformulation of the problem is then exploited by our approach to use state-of-the-art techniques in variable selection in SL for the RL problem. The approach is validated on several examples and the results found are encouraging.

References

[1] S. Murphy, "An experimental design for the development of adaptative treatment strategies," *Statistics in Medicine*, vol. 24, pp. 1455–1481, 2005.

[2] D. Ernst, P. Geurts, and L. Wehenkel, "Tree-based batch mode reinforcement learning," *Journal of Machine Learning Research*, vol. 6, pp. 503–556, 2005.

[3] E. George and R. McCulloch, "Approaches for Bayesian variable selection." *Statistica Sinica*, vol. 7, 2, pp. 339–373, 1997.

[4] C. Strobl, "Variable selection in classification trees based on imprecise probabilities," in *4th International Symposium on Imprecise Probabilities and Their Applications, Pittsburgh, Pennsylvania*, 2005. [Online]. Available: http://www.sipta.org/isipta05/proceedings/022.html

[5] L. Gunter, J. Zhu, and S. Murphy, *Artificial Intelligence in Medicine.* Springer Berlin / Heidelberg, 2007, vol. 4594/2007, ch. Variable Selection for Optimal Decision Making, pp. 149–154.