In PASCAL 2008 Workshop on *Approximate Inference in Stochastic Processes and Dynamical Systems*, Cumberland Lodge, UK.

Information Evolution of Optimal Learning

Roman V. Belavkin

School of Computing Science, Middlesex University, London NW4 4BT, UK

It is widely accepted that learning is closely related to theories of optimisation and information. Indeed, there is no need to learn if there is nothing to optimise; if one possesses full information, then there is simply nothing new to learn. The paper considers learning as an optimisation problem with dynamical information constraints. Unlike the standard approach in the optimal control theory, where the solutions are given by the Hamilton–Jacobi–Bellman equation for Markov time evolution, the optimal solution is presented as the system of canonical Euler equations defining the optimal information–utility trajectory in the conjugate space. The optimal trajectory is parameterised by the information–utility constraints, which are illustrated on examples for finite and infinite–dimensional cases.

Without uncertainty, optimisation corresponds to a simple choice problem, where a preference relation (total preorder) is defined over the underlying set Ω. If one can represent the preference relation by some real *utility* function $u : \Omega \to \overline{\mathbb{R}}$, then the optimal solution corresponds to the extremum (e.g. the maximum) of this function. The utility function may take the form of a Lagrangian incorporating many objectives, and such problems have been well–studied in variational analysis and theory of optimal control [1, 2].

When the states cannot be observed directly, the problem is usually formulated as optimisation under uncertainty, and the methods of Bayesian estimation, dynamic programming and stochastic control are applied [3, 4]. A significant development in this field was the theory of optimal non–linear filtering [5] and conditional Markov processes [6], where general solutions were first described by stochastic differential equations. Particular cases of these equations, such as the discrete and linear equations [7], have become important in applications.

Despite the successes of this approach, its limitations are becoming apparent in some learning problems. In particular, the standard approach is based on the classical theory of statistical estimation, which is an asymptotic theory. This means that its predictions are true only if the probability distributions are known with sufficient precision. Furthermore, many problems of learning and choice under uncertainty violate some of the basic assumptions of the weak low of large numbers, such as the independence of trials, stationary distributions and so on. It is suggested that these limitations of asymptotic learning theories can be overcome by considering dynamic information constraints.

The problems considering constraints on entropy or Shannon information first appeared in information theory [8–10], where they were used to determine maximum channel capacity and information value. Here, I first present a generalised form of these problems by finding conditional extremum in partially ordered Banach space. Next, this general form will be used to prove a result about the interior and exterior extrema. This result will be then applied to a dynamical setting to define the optimal learning trajectory. The results will be illustrated on examples for dynamical constraints on entropy for finite and continuous sets, and determine the optimal trajectories in the corresponding finite and infinite–dimensional Banach spaces.

The relation of the optimal learning trajectory to other methods will be discussed. In particular, it will be shown that the solution converges to the standard asymptotic techniques as information increases. It will be shown also that some existing stochastic techniques, such as simulated annealing [11], correspond to a particular case of the solution, and the optimal trajectory can be used to define the optimal annealing schedule. The connection between the information dynamics and the theory of conditional Markov processes will be considered.

The implementation of the theory in several applications will be reviewed. In particular, it will be shown how the theory was applied by the author to cognitive models of animal learning and human decision–making [12], and how it helped to model several psychological phenomena. Another application is in the study of optimal action selection strategies and adaptation of agents in stochastic environments [13]. The third application is stochastic meta–control of symbolic rule learning by neural cell–assemblies.

References

1. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., Mishchenko, E.F.: The mathematical theory of optimal processes. John Wiley and Sons (1962) Translated from Russian.
2. Gelfand, I.M., Fomin, S.V.: Calculus of variations. Prentice–Hall (1963) Translated from Russian.
3. Wald, A.: Statistical Decision Functions. John Wiley & Sons, New York (1950)
4. Bellman, R.E.: Dynamic Programming. Princeton University Press, Princeton, NJ (1957)
5. Stratonovich, R.L.: Optimum nonlinear systems which bring about a separation of a signal with constant parameters from noise. Radiofizika **2**(6) (1959) 892–901
6. Stratonovich, R.L.: Conditional Markov processes. Theory of Probability and its Applications (USSR) **5**(2) (1960) 156–178
7. Kalman, R.E., Bucy, R.S.: New results in linear filtering and prediction theory. Transactions ASME, Series D–J. Basic Engineering. **83** (December 1961) 94–107
8. Shannon, C.E.: A mathematical theory of communication. Bell System Technical Journal **27** (July and October 1948) 379–423 and 623–656
9. Kolmogorov, A.N.: The theory of information transmission. In: Meeting of the USSR Academy of Sciences on Scientific Problems of Production Automatisation, 1956, USSR Academy of Sciences (1957) 66–99 In Russian.
10. Stratonovich, R.L.: O tsennosti informatsii. Izvestiya of USSR Academy of Sciences, Technical Cybernetics **5** (1965) 3–12 In Russian.
11. Kirkpatrick, S., Gelatt, C.D., Vecchi, J.M.P.: Optimization by simulated annealing. Science **220**(4598) (May 1983) 671–680
12. Belavkin, R.V., Ritter, F.E.: The use of entropy for analysis and control of cognitive models. In Detje, F., Dörner, D., Schaub, H., eds.: Proceedings of the Fifth International Conference on Cognitive Modeling, Bamberg, Germany, Universitäts–Verlag Bamberg (April 2003) 21–26 ISBN 3-933463-15-7.
13. Belavkin, R.V.: Acting irrationally to improve performance in stochastic worlds. In Bramer, M., Coenen, F., Allen, T., eds.: Proceedings of AI–2005, the 25th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence. Volume XXII of Research and Development in Intelligent Systems., Cambridge, BCS, Springer (December 2005) 305–316 ISBN 1-84628-225-X.