Dynamic benchmark targeting ☆

Karl H. Schlag [a], Andriy Zapechelnyuk [b],*

[a] *Department of Economics, University of Vienna, Hohenstaufengasse 9, 1010 Vienna, Austria*
[b] *Adam Smith Business School, University of Glasgow, University Avenue, Glasgow G12 8QQ, UK*

Abstract

We study decision making in complex discrete-time dynamic environments where Bayesian optimization is intractable. A decision maker is equipped with a finite set of benchmark strategies. She aims to perform similarly to or better than each of these benchmarks. Furthermore, she cannot commit to any decision rule, hence she must satisfy this goal at all times and after every history. We find such a rule for a sufficiently patient decision maker and show that it necessitates not to rely too much on observations from distant past. In this sense we find that it can be optimal to forget.
© 2017 Published by Elsevier Inc.

1. Introduction

We are concerned with decision making in discrete-time dynamic environments that are hard to predict and to model explicitly, due to complexity or lack of information.

How would a firm optimally choose its inventories if the demand for its product is stochastic and subject to unpredictable structural breaks?

How would a police department decide about the number of police cars and their patrol routes if crimes do not follow any stationary pattern?

How should patients in an emergency ward be assigned to doctors if there is no discernible system in arrival of patients with different urgency of medical attention?

In economics, the standard approach to dynamic decision making involves modeling the environment as a specific stochastic process and then optimizing within this model. Unknown parameters of the process are estimated by statistical methods. However, this approach typically comes with several problems. Different assumptions on the underlying stochastic process lead to different solutions, and the true environment is never known. Explicit and tractable solutions only exist for simplest scenarios. Complex models that include more realistic features, such as structural breaks at unknown time, easily make the problem intractable. Tractable models often cannot approximate the real environment, resulting in serious errors in decision making.

An alternative approach that is popular in machine learning considers decision making with expert advice and the well known no-regret problem.[1] It can deal with environments of arbitrary complexity—in fact, the modeler does not even need to know anything about the environment. In this approach, the decision maker is equipped with a finite set of benchmark strategies or experts that she uses as targets. Her objective is to perform similarly to or better than each of them, without making any specific assumptions about the environment. These benchmark strategies could be simple heuristic decision making rules, standard practices in the given situation, solutions to the problem under specific assumptions about the environment, or strategies of experts who know more about the environment than the decision maker does. However, this approach has two caveats from an economist's point of view. First, a decision maker is infinitely patient, there is no discounting of payoffs in practically all papers. Second, the decision maker has the power to commit to a decision rule, as the performance is only measured at the outset.

This paper addresses these two caveats by inserting a new pair of elements into decision making with expert advice: *discounting of future payoffs* and *dynamic consistency*. We refer to our methodology as *dynamic benchmark targeting*. We design a decision-making rule that dynamically combines benchmark strategies and achieves a similar or superior present-value performance to each of them in all environments, at each point of time, provided the decision maker is sufficiently patient.

Freedom of choice and the absence of legal institutions that hardwire the behavior make dynamic consistency a necessity. Dynamic consistency, while being a standard assumption for economists (see Strotz, 1956; Rubinstein, 1998), is a novel feature that our paper introduces to the literature on decision making with expert advice. A decision rule is dynamically consistent if it performs well at any point in time, not only ex-ante. The decision maker does not commit to any course of actions from the start. She asks herself in every period, after every history, whether the previously chosen strategy will continue to perform well enough relative to the set targets and whether she should continue using it. However, all the literature on decision making with expert advice assumes commitment to a particular strategy from the start, and thus ignores dynamic consistency.[2] All but two decision rules used in this literature are not dynamically consistent. The two exceptions are discussed at the end of this section.

Discounting of future payoffs is the paradigm in economic decision making in which one is forward-looking and considers tradeoffs over time. The literature on decision making with expert

---

[1] For a survey of this literature see Cesa-Bianchi and Lugosi (2006).

[2] Some papers consider infinitely patient decision makers who care about long-run average streams of payoffs, so the dynamic consistency issue does not arise (e.g., Hart and Mas-Colell, 2000, 2001).