Learning and Earning in Finance – a Stochastic Control Approach

Ola Andersson (Department of Economics)
Erik Ekström (Department of Mathematics)

The purpose of the current project is to use stochastic control methods and applied and experimental game theory to study various problems in economics and operations research. The unifying theme of the considered problems is **incomplete information where the learning rate of the unknown information is controlled**, thereby creating a trade-off between learning and earning (or exploration vs exploitation).

**Background on the Project**

In classical **stochastic control** theory one deals with the existence of uncertainty by assuming that the evolution of a system is governed by noise. For example, a basic controlled system in continuous time can be described by an equation

\[ dX_t = u_t \, dt + dW_t, \]  

(1)

where the noise is modeled using a Brownian motion \( W \) and the control variable \( u_t \) is chosen in an adaptive way by the controller. Stochastic control aims to design the time path of the controlled system that performs a certain control task with minimal average cost (alternatively, with maximal average reward). For example, a classical instance is the linear-quadratic control problem

\[
\inf_u \mathbb{E} \left[ \int_0^\infty e^{-rt} (X_t^2 + cu_t^2) \, dt \right]
\]

of choosing a control \( u \) to keep the process \( X \) in (1) close to 0. Other stochastic control problems are **stopping problems**, where the control variable does not affect the dynamics of the system, but instead determines when to stop the experiment. Natural examples of stopping problems include statistical experiments, where stopping is designed so as to balance the obtained precision in the experiment against the cost of collecting further data, and financial examples of finding an optimal time to enter a certain business opportunity.

In many applications of stochastic control theory, **incomplete information** is a natural additional ingredient. A basic model for incomplete information is Brownian motion

\[ X_t = \mu t + W_t \]

with unknown drift \( \mu \), where \( \mu \) is a random variable which cannot be observed directly. Instead, observations of the process \( X \) yield noisy observations of the drift. In a stopping problem with incomplete information, there is a trade-off between continuing (to collect more information about the underlying model) and stopping (to collect the pay-off).

The current project addresses **control problems with incomplete information where the learning rate is controlled**. A few instances of such problems are described below.
(i) In dynamic allocation problems, a decision maker allocates resources between two (or several) projects with unknown profitability. A particular case is A/B-testing, where two (or more) variants (A and B) of a single variable are tested. A classical example is a company directing online customers to two different versions of their web page, with the aim of learning which version is more efficient (gives a larger probability that the customer buys the company’s product). In doing so, however, there is an inherent trade-off between learning and earning (or exploration and exploitation).

(ii) Stochastic game theory with asymmetric information: in many real-world situations, several decision makers work under competition, thus forcing agents to act strategically. Moreover, different agents have different information, and the actions chosen will in general reveal information to the other parties. One example is in hiring processes, and more generally employer-employee relations, where the employee has more knowledge about their true capacity.

(iii) In the dynamic pricing problem, a seller sequentially sets prices based on prior beliefs of an unknown demand function as well as on observations of accepted and rejected prices in the past. In this way, when setting prices the seller faces a trade-off between optimal learning and earning.

(iv) In sorting algorithms for lists where each comparison is subject to randomness (for example, a tennis tournament, where a match is won by the stronger player only with some probability \( p > 0.5 \)), there is an inherent trade-off between precision (achieved by playing many matches, as well as selecting pairs to match that give the most information) and cost of sampling.

(v) Some investment situations exhibit a feature of learning-by-doing, in the sense that the more you are invested into a project, the faster you learn about the profitability of the project (the ability to distinguish idiosyncratic noise from systemic noise increases). Thus an uninvested agent may be more eager to increase their level of investment than an invested one, and a trade-off between learning (by investing in a non-myopic way) and myopic earning arises.

Notably, while these formulations share the feature of control of the learning rate, they also exhibit distinct differences. For example, in (ii), (iii) and (iv), the magnitude of the learning rate is controlled, whereas (i) and (iv) amounts to choosing the direction of learning.

Problems of this type are notoriously difficult and rarely studied in the applied mathematics literature. In fact, to the best of our knowledge, existing literature is limited to less than a handful sources. On the other hand, the economics literature is richer, but often limited to application specific problem formulations in discrete time. The goal of the project is twofold:

(i) To set up and study appropriate mathematical models for various problems of exploration and exploitation using methods of stochastic control. To begin with, application specific problems will be studied, but with the long-term goal of developing a systematic approach to stochastic problems with incomplete information and controllable learning rate, thereby filling gaps in existing literature.

(ii) To study robustness of models, for example by investigating the effect of changes in preferences. Also, while the Bayesian approach seems viable in general, the robustness of results with respect to other optimization criteria is also of interest. Additionally, to design experiments that verify (or contradict) predictions made from analytic considerations.
The successful candidate will be based at the Department of Mathematics, with Professor Erik Ekström as main advisor, and with the second advisor Associate Professor Ola Andersson at the Department of Economics. The project will be performed as a team effort, combining a mathematical background in stochastic control (EE) with expertise in experimental and behavioral economics (OA). In particular, instrumental for the project is a successful combination of two strands of literature (mathematics and economics), as well as relevant training of the successful candidate in both disciplines.

Description of the candidate

The proposed project is challenging and requires the ability to acquire a strong competence in several diverse fields, such as stochastic analysis, Bayesian statistical theory, experimental design, Economic theory, game theory and modelling, and the successful candidate is expected to have a background in at least one of these areas. The collection of methods that will be employed to study the suggested optimization problems consists of both analytic considerations within stochastic control theory and experimental methods in economics and game theory.

Funding

The successful candidate will be funded jointly by CIM and the Department of Mathematics.