Learning with endogenous information

Subir Bose
University of Leicester

Tassos Magdalinos
University of Southampton

Miltiadis Makris
University of Essex

12 July 2023

Introduction

Economic agents often face sequential decision problems with imperfect information about the impact of their actions. Examples:

- Adopting a new technology: a firm may not know the impact a current innovation will have on the arrival of *profitable* opportunities to innovate in the future.
- Pursuing a novel idea/project: a researcher may have imperfect information about the impact of pursuing novel research on the arrival of future ideas/projects with *high* potential.
- Developing of a medical treatment following a medical breakthrough: there is uncertainty over the treatment's influence on the emergence of future *profitable* breakthroughs.
- Contemplating to consume alcohol/drugs: a consumer may not know whether such consumption can make future *urges* to consume more likely.

Introduction

We study optimal learning when:

(a) data that may carry pertinent information arrives over time *regardless* of past decisions

(b) there is uncertainty about how current choices influence the process that generates future data/information.

- arrival of innovation opportunities
- emergence of research ideas/projects
- arrival of medical or technological breakthroughs
- occurrence of addictive urges to consume.

- Related literature: restless bandit problems; learning by experimentation models.

Contribution of the paper

The learning problem in hand is not standard:

- the data our decision maker (DM) faces sequentially are *dependent* and *heterogeneously distributed* random variables
- Cannot appeal to a law of large numbers (LLN): sample mean does not have to concertos around a number so the usual techniques for establishing asymptotic learning do not apply.
- We overcome this difficulty by establishing a novel *a.s.* positive *asymptotic lower bound* for the sample mean of the data
- Our Bayesian agent learns the truth asymptotically with the posterior probability converging *a.s.* at an exponential rate.

Contribution of the paper

- Our asymptotic bound approach is more generally applicable to studying the asymptotics of a function of a sample mean generated by a nonstationary process (x_t)

$$\psi\left(\bar{x}_n\right)$$

when ψ is either:

  - convex in $\mathbb{R}$ and strictly increasing in $\mathbb{R}_+$
  - concave in $\mathbb{R}$ and strictly decreasing in $\mathbb{R}_+$

- Many Bayesian (learning) problems involve posterior distributions whose asymptotic analysis naturally reduces to the above framework.

Learning

- In each period t, there may exist an opportunity to experiment ($x_t = 1$) or not ($x_t = 0$).
- The arrival of experimentation opportunities over time is governed by the "state of nature" $\theta \in \Theta \equiv \{\theta_0, \theta_1\}$
- The state of nature is a random variable whose realisation is initially unknown to an agent
- Agent observes the history of arrivals of experimentation opportunities and would like to discover the true state of nature
- Question: will the agent will learn asymptotically the true state of nature (θ_0 or θ_1)?
- $x^t = (x_1, ..., x_t)$ be a sample of length t of (observable) experimentation opportunities

Data: experimentation opportunities

- Experimentation *opportunity* at time t: $x_t \in \{0,1\}$
- $x^t = (x_1, ..., x_t)$ observable experimentation *opportunities*
- $\Theta = \{\theta_0, \theta_1\}$ unobserved state of nature: For $p \in (0,1)$
  - i.i.d. under θ_0

$$\mathbb{P}(x_t = 1 | x^{t-1}, \theta_0) = p + \phi_0 f_t(x^{t-1}), \quad \phi_0 \geq 0 \qquad (1)$$

    (typical situation $\phi_0 = 0$)
  - history-dependent under θ_1:

$$\mathbb{P}(x_t = 1 | x^{t-1}, \theta_1) = p + \phi_1 f_t(x^{t-1}), \quad \phi_1 > \phi_0 \qquad (2)$$

  - f_t non-stochastic coordinate-wise increasing satisfying

$$\mathbf{Im}(f_t) = \begin{cases} [-p/\phi_1 + \delta, (1-p)/\phi_1 - \delta], & \text{if } f_1(x^0) > -p/\phi_1 \\ [0, (1-p)/\phi_1 - \delta], & \text{if } f_1(x^0) \geq 0 \end{cases}$$

Data: experimentation

- Experiment at time t: $d_t = 1$; do not experiment $d_t = 0$
- Experimentation in period t takes place ($d_t = 1$) when:
  - an opportunity for experimentation arises ($x_t = 1$)
  - and the extent of past experimentation is sufficiently low: $f_t(x^{t-1})$ below some threshold

$$d_t = x_t \mathbf{1}\{f_t(x^{t-1}) < \overline{f}_t(x^{t-1})\} \tag{3}$$

$$h_* := \inf_{t \in \mathbb{N}} \overline{f}_t(x^{t-1}) > 0 \quad a.s. \tag{4}$$

- The experimentation rule (3) could be:
  - exogenously given (as an experimentation heuristic, exogenously set criteria for "go ahead")
  - the result of optimal decision making by a Bayesian agent faced with a learning-by-experimentation problem where the occurrence of an opportunity to experiment generates a short-run net benefit from experimentation but, depending on the state of nature, could also give rise to costs: e.g. harmful consumption under stochastic habit formation

Assumption

$$f_t(x^{t-1}) = g(f_{t-1}(x^{t-2}), d_{t-1}) \quad t \geq 2$$

where g is a differentiable function satisfying:

(a) when $\mathbf{Im}(f_t) = [0, (1-p)/\phi_1 - \delta]$,

$$g(0,0) = 0, \quad g(0,1) > 0, \quad \inf_{z \geq 0} \frac{\partial g(z,y)}{\partial z} \geq 0, \quad \inf_{z \geq 0} \frac{\partial g(z,0)}{\partial z} > 0 \tag{5}$$

(b) when $\mathbf{Im}(f_t) = [-p/\phi_1 + \delta, (1-p)/\phi_1 - \delta]$,

$$\gamma := \sup_{z<0} \sup_{y \geq 0} \frac{\partial g(z,y)}{\partial z} \in [0,1) \tag{6}$$

in addition to (5).

- (5) implies $g(\cdot, y)$ that is increasing and that both $g(\cdot, 0)$ and $g(0, \cdot)$ are strictly increasing in $\mathbb{R}_+$
- (6) asymptotically eliminates negative initial condition effects.

Posterior distribution

- Bayesian agent observes experimentation opportunities x^t.
- Posterior belief of the agent that $\theta = \theta_1$, given sample x^t:

$$\mu(x^t) = \frac{\mu(x^{t-1})\mathbb{P}(x^t|x^{t-1},\theta_1)}{\mu(x^{t-1})\mathbb{P}(x^t|x^{t-1},\theta_1) + (1-\mu(x^{t-1}))\mathbb{P}(x^t|x^{t-1},\theta_0)}$$

- Backward recursion:

$$\begin{aligned}\mu(x^{t-1})\mathbb{P}(x^t|x^{t-1},\theta_1) &= \mu_0 \prod_{j=1}^{t} \mathbb{P}\left(x^j|x^{j-1},\theta_1\right) \\ &= \prod_{j=1}^{t} \left(p+\phi_1 h_j\right)^{x_j} \left(1-p-\phi_1 h_j\right)^{1-x_j}\end{aligned}$$

μ_0 prior belief that $\theta = \theta_1$ given the null history x_0 and

$$h_j = f_j(x^{j-1}) \text{ and } \overline{h}_j = \overline{f}_j(x^{j-1}).$$

Sufficiency of asymptotic bounds

- Posterior belief of the agent that $\theta = \theta_1$, given sample x^t:

$$\mu(x^t) = \left(1 + \frac{1-\mu_0}{\mu_0} \exp\left\{-t\tilde{\pi}_t\right\}\right)^{-1}, \quad \tilde{\pi}_t = \frac{1}{t}\sum_{j=1}^{t} q\left(x_j, h_j\right)$$

  where

$$q\left(x, y\right) = x \log\left(\frac{p+\phi_1 y}{p+\phi_0 y}\right) + (1-x)\log\left(\frac{1-p-\phi_1 y}{1-p-\phi_0 y}\right).$$

- The following asymptotic bounds are sufficient for asymptotic learning:

$$\liminf_{t\to\infty} \tilde{\pi}_t > 0 \qquad \mathbb{P}_{\theta_1}\text{-a.s.}$$

$$\limsup_{t\to\infty} \tilde{\pi}_t < 0 \qquad \mathbb{P}_{\theta_0}\text{-a.s.}$$

- No need for law of large numbers for $\tilde{\pi}_t$

Derivation of asymptotic bounds

- Approximation by sample mean of C.E's: for $i \in \{0,1\}$

$$\tilde{\pi}_t = \frac{1}{t}\sum_{j=1}^{t} \mathbb{E}_{\theta_i}\left(q\left(x_j, h_j\right)\middle| x^{j-1}\right) + o_{a.s.}\left(1\right)$$

- Using expression for $\mathbb{P}\left(x_j = 1 | x^{j-1}, \theta_i\right)$:

$$\mathbb{E}_{\theta_i}\left(q\left(x_j, h_j\right)\middle| x^{j-1}\right) = \psi_1\left(h_j\right)\mathbf{1}\left\{\theta = \theta_1\right\} + \psi_0\left(h_j\right)\mathbf{1}\left\{\theta = \theta_0\right\}$$

$$\begin{aligned}
\psi_1\left(y\right) &= \left(p+\phi_1 y\right)\log\left(\frac{p+\phi_1 y}{p+\phi_0 y}\right) + \left(1-p-\phi_1 y\right)\log\left(\frac{1-p-\phi_1 y}{1-p-\phi_0 y}\right) \\
\psi_0\left(y\right) &= \left(p+\phi_0 y\right)\log\left(\frac{p+\phi_1 y}{p+\phi_0 y}\right) + \left(1-p-\phi_0 y\right)\log\left(\frac{1-p-\phi_1 y}{1-p-\phi_0 y}\right)
\end{aligned}$$

- ψ_1 convex, strictly increasing on $\mathbb{R}_+$, $\psi_1\left(0\right) = 0$
- ψ_0 concave, strictly decreasing on $\mathbb{R}_+$, $\psi_0\left(0\right) = 0$

Reduction by convexity

- Law of motion

$$\mathbb{P}(x_t = 1 | x^{t-1}, \theta_i) = p + \phi_i h_t$$

  implies an asymptotic lower bound for $\mathbb{E}\left(h_t \middle| x^{t-2}\right)$
- We need to pass to $n^{-1}\sum_{t=1}^{n}\psi_i\left(h_t\right)$

Lemma

Let $\mathcal{G}_j$ be a filtration, Y_j be $\mathcal{G}_j$-adapted, ξ_j be $\mathcal{G}_{j-1}$-adapted,

$$\mathbb{E}\left(Y_j \middle| \mathcal{G}_{j-1}\right) \geq \kappa + \xi_j \ \text{ a.s.}, \quad \text{some } \kappa > 0,\ \xi_n \to_{a.s.} 0$$

If ψ is convex on $\mathbb{R}$, strictly increasing on $\mathbb{R}_+$ with $\psi(0) = 0$, continuous at κ and $\sup_{j\in\mathbb{N}}\left\|\psi\left(Y_j\right)\right\|_{L_2} < \infty$, then

$$\liminf_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\psi\left(Y_j\right) \geq \psi(\kappa) > 0 \ \text{ a.s.}$$

Main result

- Lower bound:

$$\kappa = \min\left\{\beta p, g\left(h_*, 0\right)\right\} > 0$$

$h_* = \inf_{t\in\mathbb{N}} \overline{f}_t(x^{t-1}) > 0$, $\beta \in (0,1)$ a proportion.

- Apply the Lemma with $Y_j = h_j$ and $\psi \in \{\psi_1, -\psi_0\}$

Theorem

Under Bayesian updating, the posterior belief that $\theta = \theta_i$ converges almost surely to unity $\mathbb{P}_{\theta_i}$-a.s. with exponential rate. In particular, there exists $\rho > 0$ such that, as $t \to \infty$:

$$1 - \mu(x^t) = O\left(e^{-\rho t}\right) \quad a.s. \quad when\ \theta = \theta_1$$

$$\mu(x^t) = O\left(e^{-\rho t}\right) \quad a.s. \quad when\ \theta = \theta_0$$