Non-Bayesian Learning in Misspecified Models

Sebastian Bervoets, Mathieu Faure, Ludovic Renou

Aix-Marseille Université & Aix-Marseille Université & Queen Mary University of London

Edinburgh, May 19, 2025

Introduction

- Psychologists, economists and other social scientists have long documented several departures from Bayesian updating:
  - under/over reaction to new information,
  - confirmation bias,
  - base-rate neglect,
  - representativeness heuristic, ...
  - See Benjamin (2019) and Ortoleva (2024) for recent surveys.
- The wording is suggestive of the "sub-optimality" of non-Bayesian updating. (There are also normative arguments in favor of Bayesian updating, e.g., Ghirardato, 2002.)
- This paper offers a more nuanced view (in a non-strategic environment).

Introduction

We revisit the classical learning model of Berk (1966):

- There is a true data-generating process (DGP) p^* an agent is learning about.
- The agent does not know the true DGP, postulates a statistical model, i.e., a set $\mathcal{P}$ of parameterized possible DGPs $p_\theta, \theta \in \Theta$, and follows Bayesian updating.
- At each period n, the agent observes a new piece of information and updates their beliefs about which $p_\theta \in \mathcal{P}$ best explains the data.

Berk (1966)

The agent's predictive DGP converges to the DGP $p_\theta \in \mathcal{P}$, which is the "closest" to the true DGP p^* (Kullback-Leibler divergence).

Introduction

- We study the model of Berk (1966) with the updating rule:

$$(\text{belief at } n+1) = (1-\gamma) \times (\text{belief at } n) + \gamma \times (\text{Bayesian update at } n+1). \quad \text{(ER)}$$

- Edwards (1968) first discussed this rule and coined it "conservative Bayesian updating." Epstein (2006) and Kovach (2023) offer axiomatic characterizations.

- The weight γ may depend on the number of observations n, the current observation and the current belief.

- E.g., to capture the confirmation bias (Rabin and Schrag, 1999), we need the weights to depend on the current belief and observation.

Introduction

An example:

- A risky asset yields either a low return $\underline{x} < 0$ or a high return $\overline{x} > 0$.
- The true probability of the high return is $2/3$ – returns are i.i.d.
- The investor does not know the true DGP and postulates that the probability of $\overline{x}$ is either $9/10$ or $1/10$.
- At each period, the investor observes the realized return and updates their belief about which of the two models best fits the data.

Berk (1966)

The investor's predictive DGP converges to $9/10$, which differs from the true DGP $2/3$.

Introduction

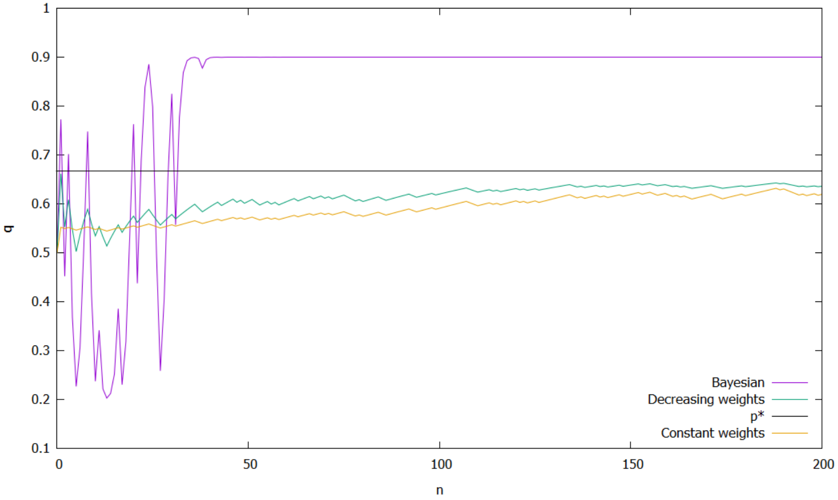


Figure: Example - “Short-run”

Introduction

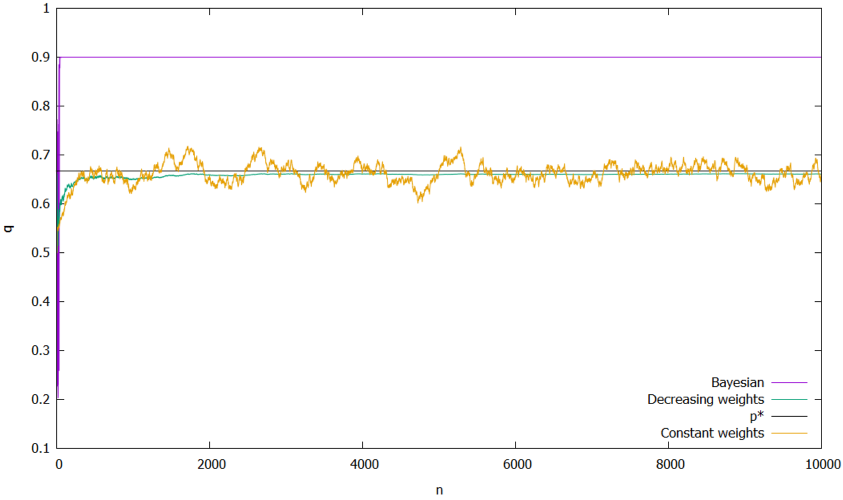


Figure: Example – “Long-run”

Introduction

- We prove that if the agent is sufficiently conservative, then (ER) updating outperforms Bayesian updating, in the sense that the predictive process induced by (ER) comes “closer” to the true data-generating process.

- In the example, the investor’s predictive process converges to the true DGP.

- The intuition is simple:
  - When the model is misspecified (as nearly all models are), a Bayesian agent extracts too much information from their model, which can lead them astray. A more conservative approach limits the severity of this problem.

Related Literature

- Learning in misspecified models: Berk (1966, 1970), Esponda and Pouzo (2016), Heidhues, Strack (2018, 2021), Frick, Iijima, Ischii (2020, 2023, 2024), Bohren and Hauser (2021, 2024), . . .
  - Some of these papers link Bayesian updating in misspecified models with non-Bayesian updating.
  - We can modify the agent's model such that Bayesian updating in that modified model is equivalent to (ER) updating in the original model. The modified model $((1-\gamma)\mathbb{E}_q[p_{\boldsymbol{\theta}}] + \gamma p_\theta)$ depends on belief $q \in \Delta(\Theta)$.
  - However, none of their results are applicable.

- Non-Bayesian learning: Rabin and Schrag (1999), Marinacci (2002), Epstein, Noor and Sandroni (2010), . . .

- Non-Bayesian updating: Edwards (1968), Philipps and Edwards (1966), Kahneman and Tversky (1972, 1973), . . .

Statistical Framework

- At each period $n \in \mathbb{N}$, an agent observes $x_n \in X$.
- The observations $(x_n)_n$ are drawn i.i.d from $p^* \in \Delta(X)$.
- The agent's model is $\mathcal{P} := \{p_\theta \in \Delta(X) : \theta \in \Theta\}$.
- The agent's prior belief is $q_0 \in \operatorname{int}\Delta(\Theta)$.
- All distributions fully supported. All sets finite.

Statistical Framework

- We think of X as consequences (not signals). E.g.,
  - asset returns in financial problems,
  - losses in insurance problems,
  - quantities demanded in pricing problems.
- We can also think of X as signals, with Θ the set of states, $\mathcal{P}^* = \{p^*_\theta : \theta \in \Theta\}$ the true experiment and $p^* := p^*_{\theta^*}$ the true signal distribution. (Interpretations of the results need to be adapted, though.)

Updating

After observing x_{n+1}, the agent updates q_n to:

$$q_{n+1} = (1 - \gamma_{n+1}) q_n + \gamma_{n+1} B(q_n, x_{n+1}), \tag{ER}$$

where

- $B(q_n, x_{n+1})$ is the Bayesian posterior of q_n given the new observation x_{n+1}, and
- γ_{n+1} the (possibly random) positive weight on the Bayesian posterior.

Updating

- Edwards (1968) coined the rule "conservative Bayesianism."
- If $\gamma_{n+1} < 1$ (resp., > 1), the agent "under-reacts" (resp., "over-reacts") to new information.
- Benjamin (2019):

  Stylized Fact 1. Underinference is by far the dominant direction of bias.
- Others departures, such as confirmation bias, are also captured.
- Epstein (2006) and Kovach (2023) offer axiomatizations and extensive discussions.

Learning

- The agent's predictive process is $(\mathbb{E}_{q_n}[p_{\boldsymbol{\theta}}])_n \in \Delta(X)^{\mathbb{N}}$.
- Two types of learning:
  - Learning to <u>predict</u>, i.e., whether $(\mathbb{E}_{q_n}[p_{\boldsymbol{\theta}}])_n \to p^*$.
  - Learning the "<u>state</u>," i.e., whether $(\mu_n)_n \in \Delta\Delta(X)^{\mathbb{N}} \to \delta_{\{p^*\}}$, where μ_n is the distribution over $\Delta(X)$ induced by q_n:

    $$\mu_n(p) = q_n(\{\theta : p_\theta = p\}).$$

    When we view Θ as states and X as signals, this is essentially equivalent to $(q_n)_n \to \delta_{\{\theta^*\}}$ (up to identification). (Recall $p^* := p^*_{\theta *}$.)
- We mostly focus on the former, but our results are also applicable to the latter.

The Main Result

Benjamin (2019) documents that:

- The more observations an agent has, the less the agent reacts to new information (compared to the Bayesian reaction) – Stylized Fact 2.
- We assume $(\gamma_n)_n$ is a decreasing sequence:

$$\sum_n \gamma_n = +\infty, \sum_n \gamma_n^2 < +\infty.$$

- The agent keeps learning, but at an increasingly slower rate. E.g., $\gamma_n = 1/n$.
- We will discuss other assumptions later on. E.g., $\gamma_n = 1/\sqrt{n}$ also works.

The Main Result

For all $p \in \Delta(X)$, let $V(p) := \mathbb{E}_{p^*}[\log p] = \sum_x p^*(x) \log p(x)$.

$V(p)$ is the (negative of) cross-entropy of p with respect to p^*. (KL divergence is entropy of p^* minus $V(p)$.)

Let $V^* = \max_{p \in \mathrm{co}\mathcal{P}} V(p)$ and $C^* = \{q \in \Delta(\Theta) : V(\mathbb{E}_q[p_\theta]) = V^*\}$.

Main Theorem

For all interior priors, the belief process (as a limit set) converges almost surely to C^*.

The Main Result

- In words, the belief process converges to the distributions, which minimize the KL divergence between p^* and the convex hull of $\mathcal{P}$.
- Bayesian updating converges to the distributions, which minimize the KL divergence between p^* and $\mathcal{P}$.
- It is *as if* we "convexify" the set of models and apply Bayesian updating to that larger model.
- In particular, if $p^* \in \mathrm{co}\mathcal{P} \setminus \mathcal{P}$, the non-Bayesian predictive process converges to the true data generating process, while the Bayesian's one does not. (Recall the introductory example.)
- In the introductory example, the belief converges to $(17/24, 7/24)$ since $2/3 = (9/10)(17/24) + (1/10)(7/24)$.

A Remark: Noisy Updates

Suppose that the agent uses $b(q_n, x_{n+1})$ to update, where $b(q_n, x_{n+1})$ is a noisy version of $B(q_n, x_{n+1})$, centered around $B(q_n, x_{n+1})$.

E.g., the agent uses $B(q_n, x_{n+1})$ + additive noise, with zero mean.

The theorem continues to hold with noisy updates.

“Misspecification Test”

- Assume that the model is identified, i.e., $(p_\theta \neq p_{\theta'})$ for all $(\theta \neq \theta')$.
- If the process $(q_n)_n$ does not converge to a mass point on a parameter θ, then the model is misspecified.
- To test the null hypothesis that the model is correctly specified, an idea is thus to compare the distance of q_n to $\delta_{\{\theta\}}$ and reject the null when it is above a threshold t for all θ.
- The optimal choice of n and t trades off type I and II errors.
- It would be interesting to develop such a test.

A Brief Detour: Stochastic Approximation

Suppose that we want to find a root of continuous f: $f(a^*) = 0$.

- If there exists a neighborhood $\mathcal{O}$ of a^* such that $f(a)(a - a^*) > 0$ for all $a \in \mathcal{O} \setminus \{a^*\}$, then the scheme

$$a_{n+1} = a_n - \gamma f(a_n)$$

  converges to a^*, for all $a_0 \in \mathcal{O}$. $\approx$ Newton's method/gradient descent. (Choose $\gamma < \sup_{a \in \mathcal{O}} f(a)^2$.)

- If we only observe $f(a_n) + \varepsilon_n$, where ε is a white noise, then Robbins and Monro (1951) proved that the scheme

$$a_{n+1} = a_n - \gamma_n \Big(\underbrace{f(a_n)}_{\text{deterministic}} + \underbrace{\varepsilon_n}_{\text{noise}} \Big),$$

  also converges to a^*.

A Brief Detour: Stochastic Approximation

Intuition:

- For n large, the impact of the noise becomes negligible and the iterates track the trajectories of the dynamical system:

$$\dot{a}(t) = f(a(t)).$$

- Indeed, if we extrapolate the iterates at "times" $\gamma_1, \gamma_1 + \gamma_2, \ldots, \sum_{i=1}^{n} \gamma_i, \ldots$, the time step becomes negligible as n increases (thus, approaching continuous time), and the impact of the noise vanishes (because $\sum_{i=1}^{n} \gamma_i^2$ becomes small).

- We can then study the stability of the critical points of f.

Sketch of Proof

We can rewrite (ER) as the Robbins-Monro scheme:

$$q_{n+1} - q_n = \gamma_{n+1}\Big(\underbrace{\left(\mathbb{E}_{p^*}[B(q_n, \mathbf{x}_{n+1})] - q_n\right)}_{:=H(q_n),\text{deterministic}} + \underbrace{\left(B(q_n, x_{n+1}) - \mathbb{E}_{p^*}[B(q_n, \mathbf{x}_{n+1})]\right)}_{\text{martingale difference, noise}}\Big).$$

From standard arguments in stochastic approximation theory, the associated ODE is:

$$\dot{q}(t) = H(q(t)).$$

We then study the critical/rest points of H and their stability.

Sketch of Proof

The θ-th coordinate of $H(q)$ is:

$$H(q)_\theta = q_\theta \left(\sum_x p^*(x) \frac{p_\theta(x)}{\sum_{\theta'} q_{\theta'} p_{\theta'}(x)} - 1 \right).$$

The solutions to $H(q) = 0$ thus correspond to the Kuhn-Tucker conditions of the constrained problems:

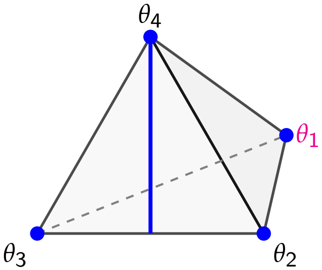
$$\max_{q \in \Delta(\Theta)} V\left(\sum_{\theta'} q_{\theta'} p_{\theta'}\right) = \sum_x p^*(x) \log\left(\sum_{\theta'} q_{\theta'} p_{\theta'}(x)\right),$$

subject to $q_\theta > 0$ iff $\theta \in \widehat{\Theta} \subseteq \Theta$. (Each $\widehat{\Theta}$ represents a "face" of $\Delta(\Theta)$.)

Sketch of Proof

We prove that the critical points can be partitioned into finitely many non-empty compact and convex components, with the additional property that V is constant on each component.

EXAMPLE. $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}, X = \{x_1, x_2, x_3\}$, and
$p_{\theta_1} = \frac{1}{4}(2,1,1),\ p_{\theta_2} = \frac{1}{4}(1,2,1),\ p_{\theta_3} = \frac{1}{4}(1,1,2),\ p_{\theta_4} = \frac{1}{8}(2,3,3), p^* = \frac{1}{8}(6,1,1)$.



- Four components, with a continuum on the boundary.

Sketch of Proof

None of the classical methods in stochastic approximation theory, e.g., Pemantle (1990), Brandiere and Duflo (1996) and Brandiere (1998), are applicable.

Pemantle (1990, p.700-701) states that

> *"if there are uncountably many unstable points, then $(q_n)_n$ can still converge to one of them even though each point has probability zero being hit. In this case, all one might conclude is that the law of the limit has no point masses at unstable critical points."*

Intuitively, the risk is to hop from one unstable point to another, indefinitely.

We show that this cannot happen: Points in component $C \neq C^*$ are repelled in a "uniform" direction, towards higher cross-entropy. Details

Updating Weights

Benaim (1999) offers other conditions on the updating weights for our main result to hold.

E.g., γ_n is $\mathcal{F}_{n-1}$-measurable with $\mathbb{E}[\sum_n \gamma_n] = +\infty$ and $\mathbb{E}[\sum_n \gamma_n^2] < +\infty$.

We discuss other, less trivial, conditions:

- Constant weights,
- Observation-dependent weights.

Constant Weights

Assume $\gamma_n = \gamma$ for all n.

As shown in introduction, the belief process does not converge in general.

The process defines a Markov chain on the space of beliefs $\Delta(\Theta)$.

The process admits multiple invariant distributions. E.g., $\delta_{\{\delta_{\{\theta\}}\}}$ is an invariant distribution for each θ.

We study the convergence of its occupation measure as γ varies.

Constant Weights

We show that, for each γ, the occupation measure converges to an invariant distribution $\pi^\gamma \in \Delta(\Delta(\Theta))$ of the chain.

Theorem

As $\gamma \to 0$, the support of the invariant distribution π^γ concentrates on C^*

Intuitively, at an invariant measure, the expected growth rate of the belief must be zero.

But beliefs grow if the invariant measure does not concentrate on C^* – they grow towards higher cross-entropy.

Observation-dependent Weights

Confirmation biases and similar biases require the weights to depend on realized observations and current beliefs.

The canonical model (Rabin and Schrag, 1999) : Two states A and B, two observations a and b, $p_A(a) > 1/2$, and $p_B(b) > 1/2$.

The observation a (resp., b) confirms A (resp., B).

If the agent believes that A is more likely, the agent reacts more to observing a than b.

We can incorporate this behavioral assumption into the weights $\gamma_n(x, q)$.

Observation-dependent Weights

Without any additional restriction, little can be said.

We assume $\gamma_n(x, q) = \gamma_n \times \alpha(x, q)$.

If $\alpha(x, q)$ is independent of q, the analysis is the same with

$$p^*_\alpha(x) := \frac{p^*(x)\alpha(x)}{\sum_{x'} p^*(x')\alpha(x')}.$$

If $\alpha(x, q)$ is piece-wise constant in q, the analysis is *locally* the same as above.

Generalized Bayes Rule

The updated belief $q_{n+1}(\theta)$ is proportional to

$$q_n(\theta) p_\theta(x_{n+1})^{\gamma_{n+1}},$$

when the agent observes x_{n+1}.

This rule is popular in statistics and machine learning (e.g., Grünwald, 2017, 2020), also captures some form of under-reaction, and corresponds to Bayes rule when $\gamma_{n+1} = 1$.

We can derive this rule as the solution to:

$$\min_{q \in \Delta(\Theta)} KL(q \| q_n) - \gamma_{n+1} \mathbb{E}_q[\log p_{\boldsymbol{\theta}}(x)].$$

Generalized Bayes Rule

It has the same long-run properties as Bayes rule.

For simplicity, assume that $V(p_{\theta^*}) > V(p_\theta)$ for all θ.

Proposition

The beliefs induced by the generalized Bayes rule converges to $\delta_{\{\theta^*\}}$ almost surely.

Intuitively, the rule under-reacts, but not sufficiently so (log vs linear scale).


THANK
YOU

Sketch of Proof

Recall that each component of $\{q : H(q) = 0\}$ is convex (and not merely connected, as typically the case).

It follows that each component has a maximally supported $\widehat{q}$, that is, if $\widehat{q}_\theta = 0$, then $q_\theta = 0$ for all q in the component.

Fix a component C other than C^*, the component that maximizes V. Choose maximally supported $\widehat{q} \in C$.

We show that there exists $q^* \in C^*$ such that $q^*_\theta > 0$ and $\widehat{q}_\theta = 0$ for some θ.

By construction, the cross-entropy is strictly increasing along the segment $[\widehat{q}, q^*]$.

This implies that $\dot{q}_\theta > 0$ for all q in the neighborhood of $\widehat{q}$.

This essentially gives us a uniform direction for all q in the component.

Back

Generalized Bayes rule: Details

Assume $\gamma_n = \gamma$ for all n

Fix a sequence of realized observations $(x_1, \ldots, x_n)$. We have:

$$\log\left(\frac{q_n(\theta^*)}{q_n(\theta)}\right) = \log\left(\frac{q_0(\theta^*)}{q_0(\theta)}\right) + \gamma \sum_{i=1}^{n} \log\left(\frac{p_{\theta^*}(x_i)}{p_\theta(x_i)}\right)$$

$$= \log\left(\frac{q_0(\theta^*)}{q_0(\theta)}\right) + n\gamma \underbrace{\left[\frac{1}{n}\sum_{i=1}^{n} \log p_{\theta^*}(x_i) - \frac{1}{n}\sum_{i=1}^{n} \log p_\theta(x_i)\right]}_{\to V(p_{\theta^*}) - V(p_\theta)}$$

Thus, $q_n(\theta^*) \to 1$, almost surely.