Journal of Optimization Theory and Applications (2025) 206:21
https://doi.org/10.1007/s10957-025-02716-2

Tracking Time Varying Parameters Via Online Simplified Maximum Likelihood

Enrico Bernardi[1] · **Alberto Lanconelli**[1] · **Christopher S. A. Lauria**[1]

Received: 2 September 2024 / Accepted: 3 May 2025

Abstract

Usually, log-likelihood functions fail to satisfy the classical assumptions of strong convexity and Lipschitz-continuity of the gradient (as well as many of their mild counterparts) that are common in general convergence results for stochastic gradient descent algorithms. Therefore, the use of gradient descent schemes to track the maxima of a sequence of objective log-likelihood functions suffers from the lack of theoretical results that guarantee the validity of the method. In this paper, we propose a simplified online scheme to track unknown dynamic parameters that are the optima of a sequence of objective log-likelihood functions. Under a Lipschitz assumption on the time varying optimum we demonstrate that our estimator achieves mean square convergence up to a neighborhood of the optimum, and we establish that the Lipschitz continuity assumption is necessary when a specific desirable property is imposed. The method is inspired by a Taylor expansion of the log-likelihood function around the maximum likelihood estimator, and rigorously justified by the expression for the Riemannian gradient of the log-likelihood of a multivariate Gaussian distribution.

Keywords Stochastic gradient descent · Maxim likelihood · Online optimization · Non-stationary optimization

Mathematics Subject Classification 65K10 · 65K05 · 62F12

Communicated by Josh Taylor.

✉ Alberto Lanconelli
alberto.lanconelli2@unibo.it

Enrico Bernardi
enrico.bernardi@unibo.it

Christopher S. A. Lauria
christopher.lauria2@unibo.it

[1] Università di Bologna, Bologna, Italy

Published online: 21 May 2025

1 Introduction

Optimization problems where the objective function changes through time have been studied in a variety of settings. In a deterministic setting [22], [23], [19], [26] find asymptotic bounds on the estimates of gradient descent schemes. For a general review see [25]. In the signal processing literature algorithms that aim to track a moving target, driven by an unknown stochastic process, over time have been analyzed in [14], see also [18] for a review of some standard techniques. While, in the machine learning domain, time varying online optimization problems have been studied in a multitude of works [32], [6], [30], [7], [10]. These challenges are frequently analyzed within reinforcement learning frameworks [9], [27], particularly in contexts where the environment undergoes continuous variation. Examples include dynamic robotic systems and financial markets, where the optimal policy must be continuously adjusted in response to evolving conditions.

Recently, a time-varying optimization problem that naturally arises in the context of statistical inference, when a time varying parameter is present, was explored in [15]. To effectively follow the trajectory of this time-varying parameter, which represents the optima with respect to a sequence of objective log-likelihood functions, the authors implemented a time-varying stochastic gradient descent scheme demonstrating that the method converges in mean-square error, within a certain neighborhood. A pertinent real-world application of this model is adaptive control in robotics, wherein a robot is required to accurately track a dynamically changing target. The primary objective in such a scenario is to minimize the discrepancy between the robot's estimated position and the actual location of the target.

In the present paper, we introduce a statistically motivated simplification of the stochastic gradient scheme presented in [15]. Our approach does not require the commonly assumed strong convexity of the objective function or the Lipschitz continuity of the gradient. We prove that this algorithm converges to a neighborhood of the time-varying optimum, and the same proof technique applies even when the parameter of interest is a matrix, such as the variance-covariance matrix. Furthermore, we demonstrate that our proposed algorithm eventually outperforms the strategy of simply estimating the unknown parameter at time t using the Maximum likelihood estimator (MLE) for that specific time. Notably, achieving this result necessitates the Lipschitz continuity assumption mentioned above.

2 The Framework

One of the classical problems in Statistical Inference is the estimation of parameters that characterize the distribution of some given random observations. Assuming that these observations, denoted throughout the paper as $\{X_t\}_{t\in\mathbb{N}}$, are independent and identically distributed according to a probability density function $p(\cdot|\lambda)$, $\lambda \in \Lambda$, the aim is to find, for any $n \in \mathbb{N}$, a map

$$(X_1, ..., X_n) \mapsto S_n(X_1, ..., X_n),$$

whose value converges as n tends to infinity, in some suitable probabilistic sense, to the true or pseudo-true unknown parameter $\lambda^\star$. Specifically, if the unknown distribution of the observations is assumed to belong to the parametric family $p(\cdot|\lambda)_{\lambda\in\Lambda}$, we refer to $\lambda^\star$ as the *true* parameter. This implies that $X_t \sim p(\cdot|\lambda^\star)$, and the model is considered to be *correctly specified*. Otherwise, if the law of the observations, say $\tilde{p}$, doesn't coincide with any density from the model $\{p(\cdot|\lambda)\}_{\lambda\in\Lambda}$, then $\lambda^\star$ is interpreted as the value that minimizes the Kullback-Leibler divergence of $\tilde{p}$ from $p(\cdot|\lambda)$, in symbols

$$\lambda^\star := \arg\min_{\lambda\in\Lambda} \mathtt{D}(\tilde{p}(\cdot)||p(\cdot|\lambda)) = \arg\min_{\lambda\in\Lambda} \int_{\mathbb{R}^m} \ln\left(\frac{\tilde{p}(x)}{p(x|\lambda)}\right)\tilde{p}(x)dx.$$

In this case we are in the framework of *mis-specified* models and $\lambda^\star$ is called the *pseudo-true* parameter [29], [1].

The Principle of Maximum Likelihood provides a natural and practical solution to estimate the parameter $\lambda^\star$ using the quantity

$$T_n(X_1, ..., X_n) := \arg\max_{\lambda\in\Lambda}[p(X_1|\lambda)\cdots p(X_n|\lambda)], \quad n \in \mathbb{N}, \tag{2.1}$$

where the function T_n is referred to as the *maximum likelihood estimator*. It is well known (see for instance [29]) that under suitable regularity conditions on the model $\{p(\cdot|\lambda)\}_{\lambda\in\Lambda}$ the maximum likelihood estimator converges an n tends to infinity to the true or pseudo-true parameter $\lambda^\star$, establishing the consistency of the method.

It is important to remark that while the Principle of Maximum Likelihood is a meaningful and theoretically justified criterion for solving the estimation problem, its actual implementation according to the prescription (2.1) is not always straightforward in many cases of interest. Indeed, the analytical solution for the maximum in (2.1) is often not available. A notable example is when $p(x|\lambda)$ represents a mixture of Gaussian distributions, as discussed in [11] and [3]. To overcome this difficulty one might employ a gradient descent algorithm [16] to approximate the maximum. There are at least two different ways to do so:

– if we have a finite number of observations, say $\{X_1, ..., X_N\}$, all available at the outset of our investigation, then the scheme

$$\lambda_{t+1} = \lambda_t + \alpha_t \nabla_\lambda \ln(p(X_1|\lambda_t)\cdots p(X_N|\lambda_t)),$$

produces a sequence $\{\lambda_t\}_{t\in\mathbb{N}}$ converging as t tends to infinity to $T_N(X_1, ..., X_N)$. Here, $\{\alpha_t\}_{t\in\mathbb{N}}$ is the so-called *learning rate*, usually chosen to be proportional to $\frac{1}{t}$, [22], while the objective function is the logarithm of the one in (2.1), as is customary in the Statistical Inference's literature.

– if the observations $\{X_t\}_{t\in\mathbb{N}}$ are available one at a time, then the scheme

$$\lambda_{t+1} = \lambda_t + \alpha_t \nabla_\lambda \ln p(X_t|\lambda_t), \tag{2.2}$$

produces a sequence $\{\lambda_t\}_{t\in\mathbb{N}}$ that converges to $\lambda^\star$ as t tends to infinity. The recursive equation (2.2) is an instance of the class of so-called *stochastic gradient descent*

algorithms, [24], since the gradient $\nabla_\lambda \ln p(X_t|\lambda_t)$ can be interpreted as a *random selection* of the deterministic gradient $\nabla_\lambda \mathtt{D}(\tilde{p}(\cdot)||p(\cdot|\lambda_t))$. In fact, observing that

$$\begin{aligned}\nabla_\lambda \mathtt{D}(\tilde{p}(\cdot)||p(\cdot|\lambda_t)) &= \nabla_\lambda \int_{\mathbb{R}^m} \ln\left(\frac{\tilde{p}(x)}{p(x|\lambda_t)}\right)\tilde{p}(x)dx \\ &= -\nabla_\lambda \int_{\mathbb{R}^m} \ln\left(\frac{p(x|\lambda_t)}{\tilde{p}(x)}\right)\tilde{p}(x)dx \\ &= -\nabla_\lambda \int_{\mathbb{R}^m} \ln\left(p(x|\lambda_t)\right)\tilde{p}(x)dx \\ &= -\nabla_\lambda \mathbb{E}[\ln p(X_t|\lambda_t)] \\ &= -\mathbb{E}[\nabla_\lambda \ln p(X_t|\lambda_t)],\end{aligned}$$

we see that minimizing the function $\lambda \mapsto \mathtt{D}(\tilde{p}(\cdot)||p(\cdot|\lambda))$ through deterministic gradient descent

$$\lambda_{t+1} = \lambda_t - \alpha_t \nabla_\lambda \mathtt{D}(\tilde{p}(\cdot)||p(\cdot|\lambda_t)),$$

can be approximated through the maximization of its random selection $\lambda \mapsto \ln p(X_t|\lambda_t)$.

It should be noted that proving convergence for the aforementioned schemes necessitates certain regularity conditions on the log-likelihood function $\lambda \mapsto \ln p(X_t|\lambda)$. These conditions are standard in optimization problems and essentially require the strong convexity and Lipschitz continuity of the gradient (see [22], [21], [5], [31]). However, many important log-likelihood functions do not satisfy these assumptions. Two typical examples illustrate this:

1. When the observations are Bernoulli with parameter q, the function

$$q \mapsto \ln p(x|q) = x\ln(q) + (1-x)\ln(1-q),$$

is neither strongly convex nor does it have a Lipschitz-continuous gradient.

2. When the observations are Gaussian with unknown variance σ^2, the function

$$\sigma^2 \mapsto \ln p(x|\sigma^2) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{x^2}{2\sigma^2},$$

also fails to fulfill these conditions. Of course, in cases where the true parameter is static, employing gradient descent schemes for parameter estimation is unnecessary due to the availability of fully explicit maximum likelihood estimators. However, our focus is on scenarios where the true parameter $\lambda_t^\star$ is time-dependent.

In deterministic analogues to our case, the challenge of tracking the time-varying optimum has been managed by utilizing tracking algorithms such as gradient descent schemes [22], [23], [19]. While in a stochastic setting, a recent paper, [15], applies the stochastic gradient scheme (2.2) to a time-varying generalization of the classical parameter estimation problem. More precisely, the authors assume that for any

$t \in \mathbb{N}$ the random vector X_t possesses a probability density function which depends on the d-dimensional parameter $\lambda_t^\star \in \Lambda \subset \mathbb{R}^d$, in symbols $X_t \sim p(\cdot|\lambda_t^\star)$. Their goal is to shadow the unknown true parameter through time utilizing the sequence of observations $\{X_t\}_{t\in\mathbb{N}}$; notice that for any $t \in \mathbb{N}$ there is only one observation available to estimate $\lambda_t^\star$. From the purview of an optimization problem, this amounts to maximizing an objective function that varies through time; in fact, the maximum of $\lambda \mapsto \nabla_\lambda \mathbb{E}[\ln p(X_t|\lambda)]$ is attained at $\lambda_t^\star$ and such a sequence depends on the time t. We also remark that the time dependent nature of the problem suggests choosing constant learning rates in (2.2), i.e. $\alpha_t = \alpha$ for all $t \in \mathbb{N}$, since vanishing learning rates wouldn't allow to keep up with the dynamics of $\{\lambda_t^\star\}_{t\in\mathbb{N}}$. The authors in [15] prove convergence of the iterative scheme (2.2) to a neighbourhood of the true time varying parameter under the assumptions of strong convexity and Lipschitz continuity of the gradient (the sequence $\{\lambda_t^\star\}_{t\in\mathbb{N}}$ is assumed to be Lipschitz continuous in t as well). It is worth mentioning that this result can be seen as a stochastic version of Theorem 1, Chapter 6.3 in [22] where non stationary optimization problems are investigated. Unfortunately, the assumptions of strong convexity and Lipschitz continuity of the gradient are fairly restrictive in a host of applied settings. These inconveniences force workarounds, such as projecting each iteration on a subset of the parameter space, which increase the complexity of the iterative scheme.

Example 2.1 Consider a robotic system that aims to estimate the position of a moving target over time. At each time step t, the true target position is denoted by $\lambda_t^* \in \mathbb{R}$, which is unknown to the robot. Instead of directly accessing λ_t^*, the robot receives a noisy observation X_t, modeled as:

$$X_t \sim \mathcal{N}(\lambda_t^*, \sigma^2),$$

where λ_t^* is the unknown true position of the target at time t, σ^2 represents the variance in the observations, and X_t is a stochastic observation received at time t. The log-likelihood function for a given estimate λ at time t is

$$\ln p(X_t|\lambda) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(X_t-\lambda)^2}{2\sigma^2}.$$

The expected log-likelihood, ignoring additive constants, attains its maximum at the true parameter value. Specifically, the expectation of the squared deviation is given by

$$\mathbb{E}[(X_t-\lambda)^2] = \sigma^2 + (\lambda_t^*-\lambda)^2.$$

This property allows us to leverage a random selection to iteratively adjust the estimate λ_t toward the true parameter λ_t^* over time, enabling dynamic tracking of the underlying process.

3 Statement of the Main Results

We now introduce some notation and a few standing assumptions that we will use throughout the article. In the sequel we will write $\|\cdot\|$ for the Euclidean norm in $\mathbb{R}^d$ and $\|\cdot\|_{\mathbb{L}^2(\Omega)}$ for $\mathbb{E}[\|\cdot\|^2]^{\frac{1}{2}}$. We consider a sequence $\{X_t\}_{t\in\mathbb{N}}$ of independent random vectors with X_t possessing a probability density function which depends on the d-dimensional parameter $\lambda_t^\star \in \Lambda \subset \mathbb{R}^d$, in symbols $X_t \sim p(\cdot|\lambda_t^\star)$.

Assumption 3.1 *(Unbiased MLE in closed form)* The MLE for the one observation log-likelihood $\ln p(X_t|\lambda)$, i.e.,

$$T(X_t) := \arg\max_{\lambda\in\Lambda} \ln p(X_t|\lambda),$$

is unbiased for $\lambda_t^\star$, i.e. $\mathbb{E}[T(X_t)] = \lambda_t^\star$, and can be written in closed form.

We emphasize that the assumption of having the single observation MLE in closed form is significantly less restrictive than the corresponding requirement for n observations, with $n \in \mathbb{N}$.

Utilizing Assumption 3.1, we are now prepared to propose a scheme that simplifies (2.2) and will be employed in tracking the time-varying parameter $\lambda_t^\star$:

$$\lambda_{t+1} := \lambda_t + \alpha(T(X_t) - \lambda_t), \quad t \in \mathbb{N}, \tag{3.1}$$

where $\lambda_1 \in \mathbb{R}^d$ is an assigned initial condition and $\alpha \in \mathbb{R}$ is positive. Before stating our main results concerning the convergence of the scheme (3.1), we try to motivate our algorithm from three different perspectives:

– **Taylor expansion of the log-likelihood.** An application of the Mean Value Theorem to the function $\lambda \mapsto \nabla_\lambda \ln p(X_t|\lambda)$ from (2.2) gives

$$\begin{aligned}\lambda_{t+1} &= \lambda_t + \alpha\nabla_\lambda \ln p(X_t|\lambda_t)\\ &= \lambda_t + \alpha\nabla_\lambda \ln p(X_t|T(X_t)) + \alpha\mathcal{H}_\lambda[\ln p(X_t|\cdot)](\xi_t)(\lambda_t - T(X_t))\\ &= \lambda_t + \alpha\mathcal{H}_\lambda[-\ln p(X_t|\cdot)](\xi_t)(T(X_t) - \lambda_t),\end{aligned}$$

where we have utilized the definition of $T(X_t)$, the vector ξ_t belongs to the segment $[T(X_t), \lambda_t]$, and $\mathcal{H}_\lambda[\ln p(X_t|\cdot)](\xi_t)$ denotes the Hessian matrix of the function $\lambda \mapsto \ln p(X_t|\lambda)$ evaluated at ξ_t. If we were under the classical conditions used to prove the convergence of stochastic gradient descent algorithms, i.e. strong convexity of $\lambda \to -\ln p(x|\lambda)$ with constant l and Lipschitz-continuity of $\lambda \to -\nabla_\lambda \ln p(x|\lambda)$ with constant L, both independent of x, then we could bound the Hessian matrix as follows $\ell I_d \le \mathcal{H}_\lambda[-\ln p(x|\cdot)](\xi_t) \le L I_d$, where I_d denotes the $d \times d$ identity matrix. Thus replacing the Hessian with βI_d, for some $\ell \le \beta \le L$ to be tuned later, we obtain

$$\lambda_{t+1} = \lambda_t + \alpha\beta(T(X_t) - \lambda_t),$$

that, up to a re-parametrization of the learning rate, is equivalent to equation (3.1). We remark that, in the case of the mean of a Gaussian distribution, equation (2.2) is already in the form (3.1) (with $T(X_t) = X_t$).

– **Riemannian gradient of a multivariate Gaussian log-likelihood.** In the case where the parameter of interest is the covariance matrix of a Gaussian distribution, an analogue to equation (3.1) for the case of matrices is naturally obtained by utilizing in (2.2) the Riemannian (as opposed to Euclidean) gradient given by the Fisher-Rao metric. Indeed, given a covariance matrix $S \in \mathtt{Sym}^+$ (the class of symmetric non negative definite matrices), the Fisher-Rao metric can be defined as

$$g_S(A, B) = \frac{1}{2}\mathtt{tr}(S^{-1}AS^{-1}B),$$

where A and B belong to $T_S(\mathtt{Sym}+)$ (see [28] for details). Under this metric the induced Riemannian gradient on the one observation m dimensional Gaussian log-likelihood, i.e.

$$\ln p(X|S) = -\frac{m}{2}\ln(2\pi) - \frac{1}{2}\ln S - \frac{1}{2}X^T S^{-1}X, \tag{3.2}$$

satisfies the following relation with the Euclidean gradient (here denoted with ∇_E):

$$\nabla_{g_S} \ln p(X|S) = \frac{1}{2}S[\nabla_E \ln p(X|S) + \nabla_E \ln p(X|S)^T]S. \tag{3.3}$$

Since the Euclidean gradient of (3.2) is

$$\nabla_E \ln p(X|S) = -S^{-1} + S^{-1}XX^T S^{-1},$$

we see using (3.3) that the Riemannian gradient becomes

$$\nabla_{g_S} \ln p(X|S) = S[-S^{-1} + S^{-1}XX^T S^{-1}]S = -S + XX^T.$$

Therefore, rewriting (2.2) for a Gaussian log-likelihood, using the Riemannian gradient, one obtains

$$S_{t+1} = S_t + \alpha\nabla_{g_S} \ln p(X_t|S_t) = S_t + \alpha(X_t X_t^T - S_t), \tag{3.4}$$

which is an instance of Equation (3.1), since $X_t X_t^T$ is the unbiased one observation MLE for the covariance matrix of X_t.

– **Law of Large Numbers.** The well-known Law of Large Numbers states that the arithmetic mean of n independent and identically distributed random variables converges to their common expected value as n approaches infinity. In symbols

$$\lim_{n\to+\infty} \frac{X_1 + \cdots + X_n}{n} = \mathbb{E}[X_1]. \tag{3.5}$$

The sample mean in (3.5) can be rewritten as a stochastic gradient descent scheme of the form (3.1). To this aim we set

$$\lambda_t := \frac{X_1 + \cdots + X_t}{t}, \quad t \in \mathbb{N},$$

and notice that a simple algebraic manipulation gives

$$\begin{aligned}\lambda_{t+1} &= \frac{X_1 + \cdots \dot{+} X_t}{t+1} + \frac{X_{t+1}}{t+1} \\ &= \frac{t}{t+1}\lambda_t + \frac{1}{t+1}X_{t+1} \\ &= \lambda_t + \frac{1}{t+1}(X_{t+1} - \lambda_t).\end{aligned}$$

Therefore, comparing the first and last members above we obtain

$$\lambda_{t+1} = \lambda_t + \frac{1}{t+1}(X_{t+1} - \lambda_t),$$

which corresponds to (3.1) with $\alpha := \frac{1}{t+1}$ and $T(X_{t+1}) := X_{t+1}$. We remark that here the learning rate is vanishing as t increases (we are estimating the static value $\mathbb{E}[X_1]$) and $T(X_{t+1})$ is an unbiased estimator of $\mathbb{E}[X_1]$.

Remark 3.1 Equation (3.1) can also be interpreted as a time-varying stochastic gradient descent algorithm for minimizing a randomly selected quadratic objective function

$$\lambda \mapsto \frac{1}{2}\mathbb{E}_{X_t}[(\lambda - T(X_t))^2].$$

These objective functions will, for each time t, achieve their minimum in $\lambda_t^\star = \mathbb{E}[T(X_t)]$, thus each iteration of equation (3.1) will be a stochastic gradient step towards the respective $\lambda_t^\star$.

Finally Equation (3.1) is similar to the optimization algorithm recently used in [8].

We now introduce a second assumption that will be required to prove the convergence of recursion (3.1).

Assumption 3.2 *(Bounded variance of MLE)* There exists a positive real constant M such that for all $t \in \mathbb{N}$ we have

$$\mathbb{E}[\|T(X_t) - \lambda_t^\star\|^2] = \sum_{j=1}^{d} \mathbb{V}[T_j(X_t)] \leq dM. \tag{3.6}$$

Here, $T_j(X_t)$ stands for the j-th component of the d-dimensional vector $T(X_t)$.

Boundedness of the variance of the one observation MLE is needed to control the noise of the sequence of estimators generated by equation (3.1). The following assumption pertains to the evolution of the time-varying parameter $\{\lambda_t^\star\}_{t\in\mathbb{N}}$ and aligns with the set of assumptions presented in [15].

Assumption 3.3 *(Lipschitz continuity of the true parameter)* There exists a positive constant K such that

$$\|\lambda_{t+1}^\star - \lambda_t^\star\| \leq K \quad \text{for all } t \in \mathbb{N}.$$

Assumption 3.3 has been used in other settings, see for example [25], [7] and [30], since a control on the behaviour of the sequence of true parameters is needed to be able to shadow it. We will demonstrate that this condition is also necessary for the scheme (3.1) to outperform the simple approach of setting $\lambda_t = T(X_t)$, $t \in \mathbb{N}$, as a sequence of estimators for tracking $\{\lambda_t^\star\}_{t\in\mathbb{N}}$. We explicitly note that the choice $\lambda_t = T(X_t)$, $t \in \mathbb{N}$, is not as naive as it might initially appear: given that the observations $\{X_t\}_{t\in\mathbb{N}}$ are independent, so, it may seem improbable to achieve improvement when dealing with a parameter $\{\lambda_t^\star\}_{t\in\mathbb{N}}$ that varies over time.

We can now state our first main theorem.

Theorem 3.4 *Let Assumptions 3.1, 3.2 and 3.3 hold. Then running equation (3.1), with* $0 < \alpha < 1$, *we obtain*

$$\limsup_{t\to\infty} \|\lambda_{t+1} - \lambda_t^\star\|_{\mathbb{L}^2(\Omega)}^2 \leq \frac{\alpha d M}{2-\alpha} + \frac{(1-\alpha)^2 d K^2}{\alpha^2}. \tag{3.7}$$

In addition, the minimum of the right hand side in (3.7) is attained at

$$\alpha = \frac{6\vartheta}{\vartheta^2 + 4\vartheta + 1},$$

where $\vartheta := (1 + 27M/K^2 + 3\sqrt{3}\sqrt{2M/K^2 + 27M^2/K^4})^{1/3}$.

Remark 3.2 Theorem 3.4 is a generalization of Theorem 1, Chapter 6.3 in [22] to stochastic quadratic optimization problems.

Remark 3.3 When $d = 1$, $\mathbb{V}[T(X_t)] = \sigma^2$, and $\lambda_t^\star$ evolves according to a linear function of t, the limit superior becomes a limit and the bound in (3.7) is sharp.

Remark 3.4 As stated in Remark 3.1, equation (3.1) can be interpreted as a time-varying stochastic gradient descent process applied to a sequence of objective functions given by

$$f_t(\lambda) = \frac{1}{2}\mathbb{E}_{X_t}[(\lambda - T(X_t))^2] = (\lambda - \lambda_t^*)^2 + \sigma^2.$$

A fundamental metric for evaluating the effectiveness of a policy in the reinforcement learning setting is the expected regret, [9], [13], [7]. The regret at time t is defined as

$$\mathrm{Reg}(t) = \sum_{i=1}^{t} f_t(\lambda_i) - f_t(\lambda_i^*),$$

where the λ_i are the control variables chosen by the policy. An ideal sequence of control variables $\{\lambda_t\}$ should posses sublinear expected regret, i.e., $\mathbb{E}[Reg(t)] \leq o(t)$. In our framework, for the policy specified by equation (3.1), the expected regret is given by

$$\mathbb{E}[\mathrm{Reg}(t)] = \sum_{i=1}^{t} \mathbb{E}[f_t(\lambda_i)] - f_t(\lambda_i^*) = \sum_{i=1}^{t} \mathbb{E}[(\lambda_i - \lambda_i^*)^2].$$

However, Theorem 3.4 establishes that

$$\sum_{i=1}^{t} \mathbb{E}[(\lambda_i - \lambda_i^*)^2] \in O(t),$$

implying that under its given assumptions, sublinear regret cannot be achieved. Furthermore, modifying the learning rate dynamically over time does not resolve this issue, as it is not possible to simultaneously drive both terms in the bound provided by equation (3.7) to zero. This phenomenon illustrates a fundamental trade-off between observational noise and the evolving nature of the time-varying parameter. In our setting a regret analysis would become feasible if either the observational noise were to vanish or the time-varying parameter were to converge.

The same proof technique utilized to prove Theorem 3.4 can be used to show the convergence, up to a neighbourhood, of equation (3.4); in this case the Euclidean norm must be replaced by the Frobenius norm.

It is also straightforward to extend the proof of Theorem 3.4 to the case in which we have $N \in \mathbb{N}$ observations from the same parametric distribution at each time t. In such a case the MLE at time t would be $T_N(X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(N)})$ where $X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(N)}$ are independent and identically distributed with parameter $\lambda_t^\star$. Substituting this estimator into the proof in place of $T(X_t)$ does not introduce modifications to the argument. In the presence of multiple observations at each time t the variance of the MLE $T_N(X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(N)})$ will decrease with N. Consequently, this reduction in variance will lead to a smaller value of M in Assumption 3.2, thereby reducing the asymptotic tracking error bounds given in equation (3.7).

The next corollary follows immediately from Theorem 3.4.

Corollary 3.1 *Let Assumptions 3.1 and 3.3 hold. If there exists a positive real constant* σ^2 *such that* $\mathbb{E}[\|T(X_t) - \lambda_t^\star\|^2] = d\sigma^2$ *for all* $t \in \mathbb{N}$ *(this is stronger than Assumption 3.2), then we can always find an* $0 < \alpha < 1$ *such that*

$$\limsup_{t\to+\infty} \|\lambda_{t+1} - \lambda_t^\star\|^2_{\mathbb{L}^2(\Omega)} < d\sigma^2. \tag{3.8}$$

Corollary 3.1 is of practical significance as it indicates that running equation (3.1) will eventually yield a superior estimator, in terms of lower mean square error, compared to using $T(X_t)$ as an estimator for $\lambda_t^\star$ at each time t. Specifically, the right-hand side of (3.8) corresponds exactly to the mean square error of $T(X_t)$. Thus, Corollary 3.1 states that for any K, it is always possible to find an α that improves upon the single-observation estimator in terms of mean squared error. As K increases, the value of α will approach 1; however, it is only in the limit as $K \to \infty$ that equation (3.1) reduces to $\lambda_t = T(X_t)$. In the case where multiple observations $X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(N)}$ are available at each time t, the comparison still holds true with the maximum likelihood estimator $T_N(X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(N)})$.

We now state and prove the converse of Corollary 3.1.

Theorem 3.5 *Let Assumption 3.1 hold and assume that there exists an α such that running equation (3.1) with it we obtain*

$$\limsup_{t\to+\infty} \|\lambda_{t+1} - \lambda_t^\star\|^2_{\mathbb{L}^2(\Omega)} < d\sigma^2.$$

Then, there exists a positive constant K such that for all $t \in \mathbb{N}$ we have

$$\|\lambda_{t+1}^\star - \lambda_t^\star\| \leq K.$$

Theorem 3.5 shows that Assumption 3.3 is necessary to be able to, eventually, perform better than the one observation MLE. This result resemble a stochastic counterpart to ones obtained in the deterministic setting, as for example in [19].

In conclusion, we extend Theorem 3.4 to a more general setting in which the MLE is approximated rather than computed exactly. This approximation introduces a potential bias in the estimation process. Specifically, we make the following assumption:

Assumption 3.6 Suppose that instead of the exact MLE for the single-observation log-likelihood function $\ln p(X_t|\lambda)$, we only have access to an approximation. That is, at each time step t, the available estimator takes the form $T(X_t) + \epsilon_t$, where the noise term ϵ_t has mean m and second moment m_1. Furthermore, we assume that, for all t, ϵ_t is independent of all observations X_t.

Under this assumption, it remains possible to establish an upper bound on the asymptotic tracking error using a proof technique analogous to the one employed in the exact MLE case. For ease of notation, we present the corresponding theorem in the one dimensional setting.

Theorem 3.7 *Let Assumptions 3.6, 3.2 and 3.3 hold. Then running the recursive equation*

$$\lambda_{t+1} := \lambda_t + \alpha(T(X_t) + \epsilon_t - \lambda_t), \quad t \in \mathbb{N}, \tag{3.9}$$

with $0 < \alpha < 1$, *we obtain*

$$\limsup_{t\to\infty} \mathbb{E}[(\lambda_{t+1} - \lambda_t^\star)^2] \leq \alpha(\sigma^2 + m_1) + (1-\alpha)m^2 + \frac{1-\alpha}{\alpha^2}\left(K + \frac{\alpha}{1-\alpha}|m|\right)^2.$$

The tracking error bound established in Theorem 3.7 can be again compared to the mean squared error of the single-observation estimator, that now is $\mathbb{E}\left[(T(X_t) + \epsilon_t - \lambda_t^*)^2\right] = \sigma^2 + m_1$. Notably, in this setting, there is no guarantee that an appropriate choice of α will always lead to an improvement when utilizing the approximation to the single-observation MLE. The extent of improvement is highly dependent on the bias term m; if the approximation is excessively inaccurate, the algorithm will no longer be able to compensate for the error.

In the next section, we provide numerical illustrations of our findings, with an emphasis on a connection to the econometric literature. Specifically, we observe that when the parameter of interest is the variance of a Gaussian distribution, the iterative scheme we propose belongs to the family of GARCH(1,1) models [4]. The final section will be dedicated to the proofs of our main results.

4 Examples and Numerical Illustrations

In this section we present a set of examples alongside relevant simulations.

4.1 The Mean of a Gaussian

Consider the setting of Example 2.1, where the observations are Gaussian with a time-varying mean and a common variance σ^2, i.e.,

$$X_t \sim \mathcal{N}(\lambda_t^*, \sigma^2).$$

In Figure 1, we present a comparison of the algorithm described by (3.1) when the mean λ_t^* evolves linearly, against the single-observation MLE, which in this case is simply X_t. Even from a visual inspection, it appears that after a sufficient number of iterations, λ_t outperforms the single-observation MLE on average.

Using the same data-generating process as in Figure 1, we also conduct a Monte Carlo analysis to examine the distance between the estimator defined by equation (3.1) and the true parameter value. The results are presented in Figure 2.

4.2 The Bernoulli Distribution

The log-likelihood of a Bernoulli distribution, i.e.,

$$p \mapsto \ln(p^x(1-p)^{(1-x)}) = x\ln(x) + (1-x)\ln(1-p),$$

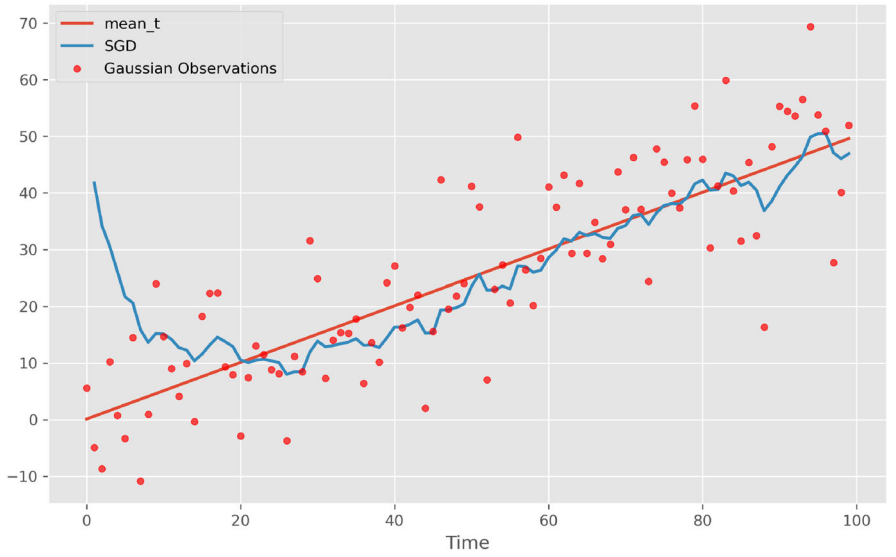


Fig. 1 The simulation was conducted with the mean of the data-generating process following the equation $\lambda_t^* = 0.1 + 0.5t$, while the standard deviation of the Gaussian noise was set to $\sigma = 10$. The initial value for equation (3.1) was chosen as $\lambda_0 = 50$, and the learning rate was fixed at $\alpha = 0.15$

is neither strongly convex nor does it have a Lipschitz continuous gradient. So the main theorem of [15] can't be straightforwardly applied to show that a time varying stochastic gradient descent scheme would track a time varying parameter $\{p_t^\star\}_{t\in\mathbb{N}}$ through time. However, the MLE of the parameter p is known in closed form and is unbiased; for a single observation log-likelihood, we have $T(X_t) = X_t$. Consequently, the iterative scheme given by (3.1) can be implemented, and Theorem 3.4 provides a theoretical guarantee of its convergence to a neighborhood of the true parameter. In Figure 3, we simulate the algorithm given by (3.1) when p evolves linearly over time.

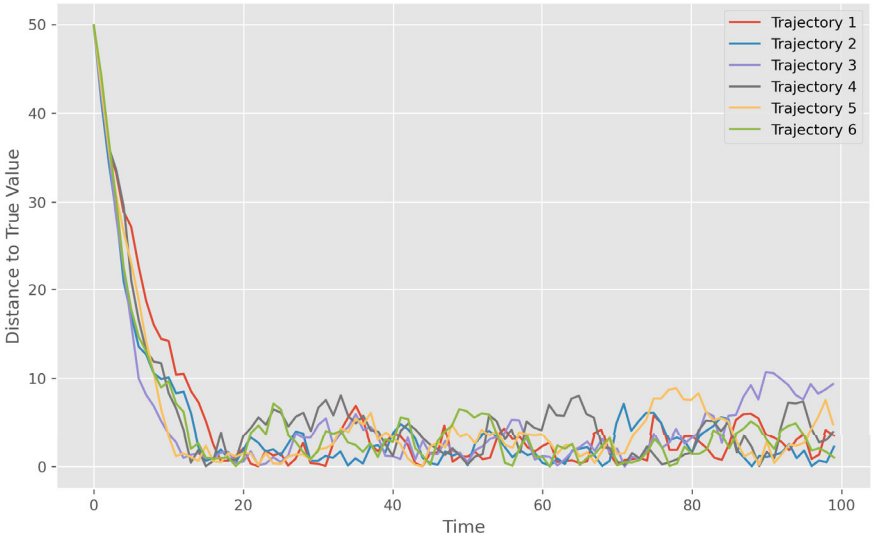
4.3 The Covariance Matrix of a Multivariate Gaussian Distribution

In this section, we analyze the tracking of an evolving covariance matrix in the context of a multivariate Gaussian distribution. Specifically, we consider a sequence of independent Gaussian observations X_t with a time-varying covariance matrix and apply the iterative update given by equation (3.4)

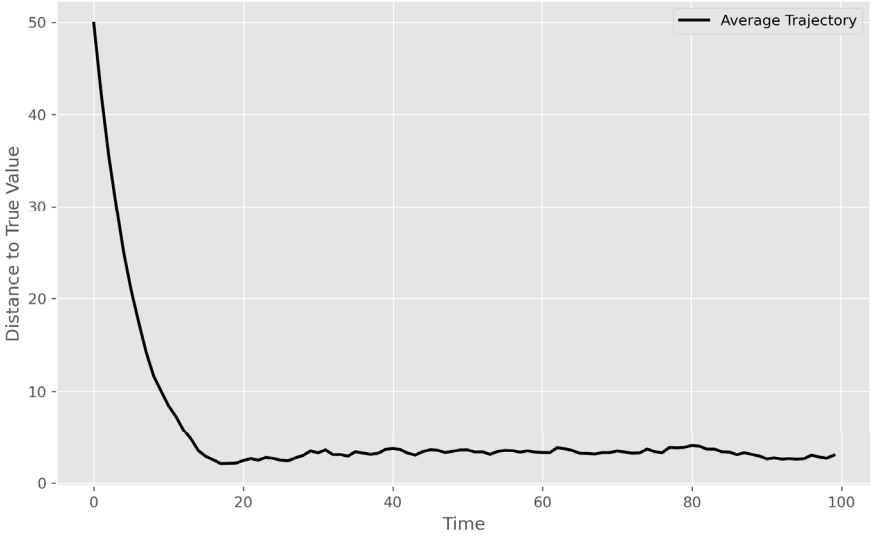
$$S_{t+1} = S_t + \alpha(X_t X_t^T - S_t).$$

Now S_t represents the estimated covariance matrix at time t.

To evaluate the performance of this approach, we conduct a numerical simulation and visualize the results in Figure 4. The figure illustrates the behavior of the Frobenius norm of S_t in comparison to the Frobenius norm of the true covariance matrix of X_t. The simulation results indicate that, over time, the estimated covariance matrix S_t



(a) Plotting the distances to the true mean for 6 different trajectories with starting point $\lambda_0 = 50$ and learning rate $\alpha = 0.15$.



(b) The average of the distance of 100 different trajectories to the true mean.

Fig. 2 A Monte Carlo analysis of the distance to the true parameter through time

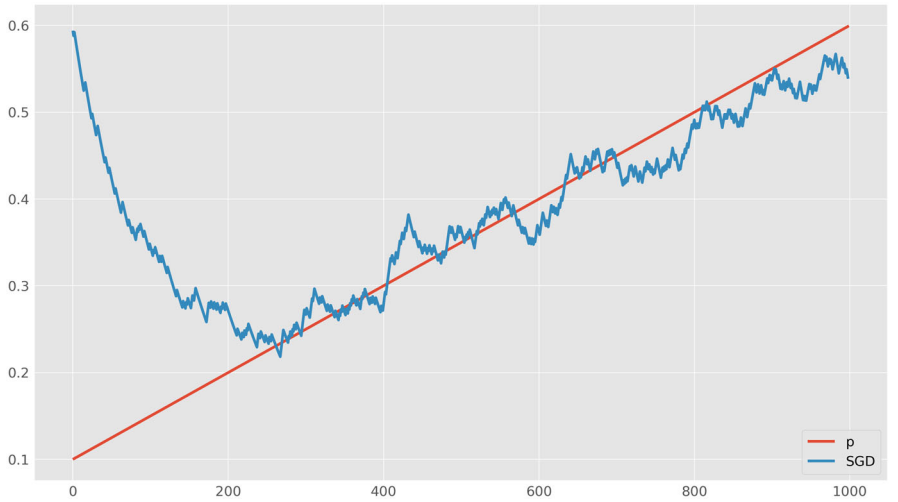


Fig. 3 The simulation was conducted with the mean of the data-generating process following the equation $\lambda_t^* = 0.1 + 0.0005t$. The initial value for equation (3.1) was chosen as $\lambda_0 = 0.6$, and the learning rate was fixed at $\alpha = 0.01$

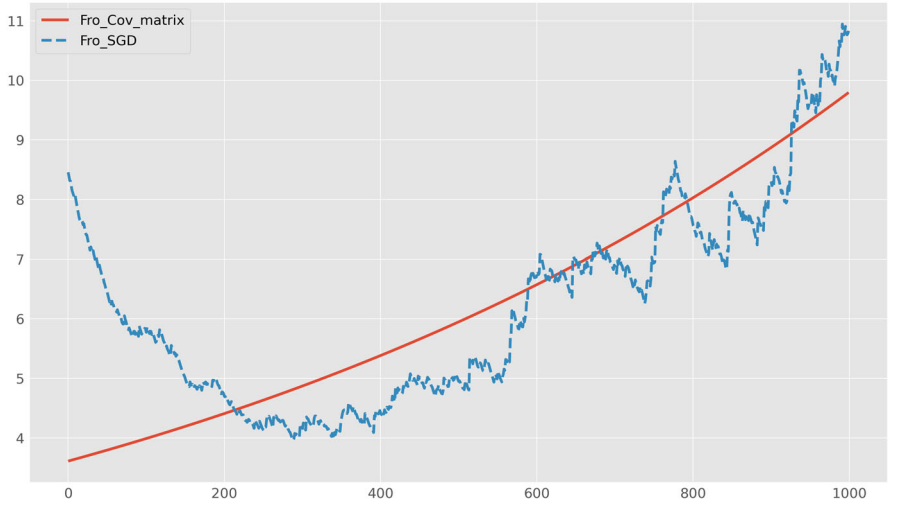


Fig. 4 The simulation was conducted with the covariance matrix S_t^* of the data-generating process following the equation $S_t^* = 1.001 S_t^*$. The initial value for equation (3.1) was chosen as $S_t = \left(\begin{smallmatrix} 6 & 0 \\ 0 & 6 \end{smallmatrix}\right)$, and the learning rate was fixed at $\alpha = 0.01$

effectively captures the underlying dynamics of the true covariance matrix, demonstrating the efficacy of the proposed update rule in the multi-dimensional setting.

4.4 Link to the GARCH(1,1) Model

We notice that, in the univariate Gaussian case, equation (3.4) becomes

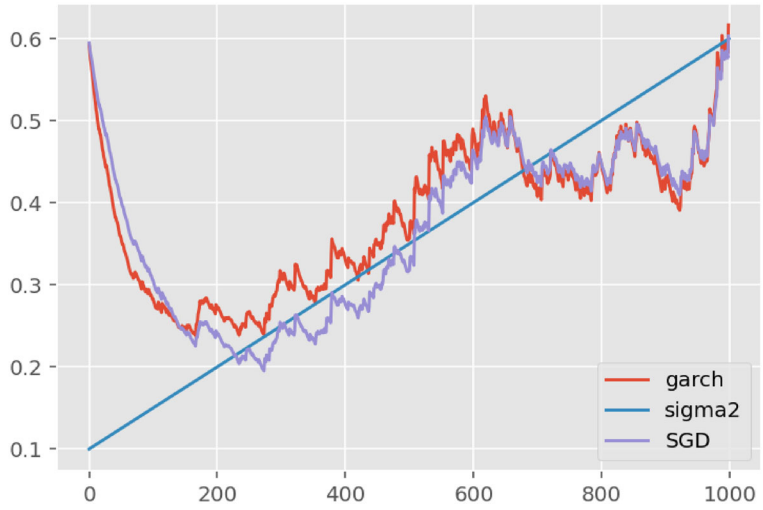


Fig. 5 The data-generating process consists of independent Gaussian observations with variance given by $\sigma_t^{*2} = 0.1 + 0.0005t$. The parameters of the GARCH(1,1) process are estimated using maximum likelihood after observing the entire dataset. The initial value is set to 0.6 for both the GARCH process and the gradient descent procedure

$$\sigma_{t+1}^2 = \sigma_t^2 + \alpha(X_t^2 - \sigma_t^2) = (1-\alpha)\sigma_t^2 + \alpha X_t^2,$$

which has the form of a GARCH(1,1) model, [4],

$$\sigma_{t+1}^2 = w + \gamma\sigma_t^2 + \beta X_t^2,$$

with $w = 0$, $\gamma = (1-\alpha)$, $\beta = \alpha$. The GARCH(1,1) was introduced to model the time varying volatility of financial time series and has since gained traction as an industry benchmark, [12]. In the econometric literature it is a recognized empirical fact that when estimating the parameters of a GARCH(1,1) one often finds $w \approx 0$, $\gamma + \beta \approx 1$, see for instance, [20] [2], [17]. Thus, in applied settings, the GARCH(1,1) is of the form of equation (3.1). In this context, Theorem 3.4 indicates that, under model misspecification, the GARCH(1,1) model, as estimated in practice (with $w \approx 0$ and $\gamma + \beta \approx 1$), converges to a neighborhood of the time-varying unconditional variance. This observation could provide additional insight into why empirical estimates often result in $w \approx 0$ and $\gamma + \beta \approx 1$. We highlight this finding for further research. In Figure 5, after having estimated its parameters, we run a GARCH(1,1) and compare its evolution to the iterations given by equation 3.1.

5 Conclusions

In this paper, we have introduced a simplified online scheme for tracking dynamic parameters that evolve as the optima of a sequence of log-likelihood functions.

Given the challenges associated with applying stochastic gradient descent in this setting-namely, the failure of log-likelihood functions to satisfy strong convexity and Lipschitz-continuity of the gradient-our approach provides a theoretically grounded alternative that ensures mean square convergence up to a neighborhood of the true parameter trajectory. These findings contribute to the broader understanding of online estimation in non-convex settings and suggest potential avenues for further research.

6 Proofs of the Main Results

6.1 Proof of Theorem 3.4

We subtract $\lambda_t^\star$ from both sides of our recursive equation (3.1), i.e.

$$\lambda_{t+1} - \lambda_t^\star = (1-\alpha)(\lambda_t - \lambda_t^\star) + \alpha(T(X_t) - \lambda_t^\star), \tag{6.1}$$

and we take the squared Euclidean norm to get

$$\begin{aligned}\|\lambda_{t+1} - \lambda_t^\star\|^2 = (1-\alpha)^2\|\lambda_t - \lambda_t^\star\|^2 + 2\alpha(1-\alpha)\langle\lambda_t - \lambda_t^\star, T(X_t) - \lambda_t^\star\rangle \\ + \alpha^2\|T(X_t) - \lambda_t^\star\|^2.\end{aligned}$$

Taking conditional expectations we obtain

$$\begin{aligned}\mathbb{E}[\|\lambda_{t+1} - \lambda_t^\star\|^2|\mathcal{F}_{t-1}] = (1-\alpha)^2\mathbb{E}[\|\lambda_t - \lambda_t^\star\|^2|\mathcal{F}_{t-1}] \\ + \alpha^2\mathbb{E}[\|T(X_t) - \lambda_t^\star\|^2|\mathcal{F}_{t-1}]\end{aligned} \tag{6.2}$$

here we exploited independence of the observations and Assumption 3.1. Computing expectations of both sides of equation (6.2) yields

$$S_{t+1} = \beta^2\mathbb{E}[\|\lambda_t - \lambda_t^\star\|^2] + \alpha^2\sigma_t^2, \tag{6.3}$$

where for notational convenience we set

$$\begin{aligned}&S_{t+1} := \mathbb{E}[\|\lambda_{t+1} - \lambda_t^\star\|^2], \quad \sigma_t^2 := \mathbb{V}[T(X_t)] = \mathbb{E}[\|T(X_t) - \lambda_t^\star\|^2], \\ &\text{and} \quad \beta := (1-\alpha);\end{aligned}$$

moreover, elaborating on the term $\mathbb{E}[\|\lambda_t - \lambda_t^\star\|^2]$, we can write

$$\begin{aligned}\mathbb{E}[\|\lambda_t - \lambda_t^\star\|^2] &= \mathbb{E}[\|\lambda_t - \lambda_{t-1}^\star + \lambda_{t-1}^\star - \lambda_t^\star\|^2] \\ &= S_t + \|K_t\|^2 - 2\langle K_t, v_{t-1}\rangle,\end{aligned}$$

where we introduced the additional notation

$$v_t := \mathbb{E}[\lambda_{t+1} - \lambda_t^\star] \quad \text{and} \quad K_t := \lambda_t^\star - \lambda_{t-1}^\star.$$

Hence, equation (6.3) now reads

$$S_{t+1} = \beta^2[S_t + \|K_t\|^2 - 2\langle K_t, v_{t-1}\rangle] + \alpha^2\sigma_t^2. \tag{6.4}$$

We now study the sequence $\{v_t\}_{t\in\mathbb{N}}$; taking the expectation in (6.1) gives

$$v_t = \beta v_{t-1} - \beta K_t, \tag{6.5}$$

and assuming for simplicity that $v_0 = 0$, we can solve equation (6.5) to get

$$v_t = -\beta \sum_{h=1}^{t} \beta^{t-h} K_h.$$

A substitution of the last expression in (6.4) yields

$$\begin{aligned}
S_{t+1} &= \beta^2 \left[S_t + \|K_t\|^2 + 2\langle K_t, \sum_{h=1}^{t-1} \beta^{t-h} K_h \rangle \right] + \alpha^2\sigma_t^2 \\
&= \beta^2 S_t + \beta^2 \left[\left\| K_t + \sum_{h=1}^{t-1} \beta^{t-h} K_h \right\|^2 - \left\| \sum_{h=1}^{t-1} \beta^{t-h} K_h \right\|^2 \right] + \alpha^2\sigma_t^2 \\
&= \beta^2 S_t + \beta^{2+2t} \left[\left\| \beta^{-t} K_t + \sum_{h=1}^{t-1} \beta^{-h} K_h \right\|^2 - \left\| \sum_{h=1}^{t-1} \beta^{-h} K_h \right\|^2 \right] + \alpha^2\sigma_t^2 \\
&= \beta^2 S_t + \beta^{2+2t} \left[\|c_t\|^2 - \|c_{t-1}\|^2 \right] + \alpha^2\sigma_t^2,
\end{aligned}$$

where

$$c_t := \beta^{-t} K_t + \sum_{h=1}^{t-1} \beta^{-h} K_h = \sum_{h=1}^{t} \beta^{-h} K_h.$$

This implies that S_t solves

$$\begin{aligned}
&S_{t+1} = \beta^2 S_t + \beta^{2+2t} \left[\|c_t\|^2 - \|c_{t-1}\|^2 \right] + \alpha^2\sigma_t^2; \\
&S_1 = 0,
\end{aligned}$$

whose solution is given by

$$S_t = \alpha^2 \sum_{i=0}^{t-1} \beta^{2i} \sigma_{t-1-i}^2 + \beta^{2t} \|c_{t-1}\|^2.$$

Now, by virtue of Assumption 3.2 we conclude that

$$S_t \leq \frac{\alpha^2 dM(1-\beta^{2(t-1)})}{1-\beta^2} + \beta^{2t}\|c_{t-1}\|^2. \tag{6.6}$$

Furthermore, an application of Hölder's inequality with ℓ^∞ and ℓ^1-norms allows for the upper bound

$$\begin{aligned}\beta^{2t}\|c_{t-1}\|^2 &= \beta^2 \left\| K_{t-1} + \sum_{h=1}^{t-1} \beta^{t-1-h} K_h \right\|^2 = \beta^2 \sum_{i=1}^{d} \langle \mathbf{K}, \boldsymbol{\beta} \rangle^2 \\ &\leq \sum_{i=1}^{d} \|\mathbf{K}\|_\infty^2 \|\boldsymbol{\beta}\|_1^2 = \beta^2 dK^2 \left(\frac{1-\beta^{t-1}}{1-\beta} \right)^2,\end{aligned}$$

where $\mathbf{K} := (K_{t-1}, K_{t-2}, \cdots, K_1)$ and $\boldsymbol{\beta} = (1, \beta, \cdots, \beta^{t-2})$. By means of this estimate we deduce from inequality (6.6) that

$$S_t \leq \frac{\alpha^2 dM(1-\beta^{2(t-1)})}{1-\beta^2} + \beta^2 dK^2 \left(\frac{1-\beta^{t-1}}{1-\beta} \right)^2,$$

and thus

$$\limsup_{t\to\infty} \|\lambda_{t+1} - \lambda_t^\star\|_{\mathbb{L}^2(\Omega)}^2 \leq \frac{\alpha dM}{2-\alpha} + \frac{(1-\alpha)^2 dK^2}{\alpha^2}.$$

To find the optimal α it is enough to minimize the function

$$\alpha \mapsto \psi(\alpha) := \frac{\alpha}{2-\alpha} + \theta^2 \frac{(1-\alpha)^2}{\alpha^2},$$

where $\theta := K/\sqrt{M}$; in fact,

$$\frac{\alpha dM}{2-\alpha} + \frac{(1-\alpha)^2 dK^2}{\alpha^2} = dM \left[\frac{\alpha}{2-\alpha} + \theta^2 \frac{(1-\alpha)^2}{\alpha^2} \right] = dM\psi(\alpha).$$

We now introduce a new variable y through the equation $\alpha = 1/(1+y) =: g(y)$; notice that $\alpha \in (0,1)$ implies $y \in (0,\infty)$. Moreover, since $g'(y) < 0$ for all $y \in (0,\infty)$, we can affirm that

$$\frac{d}{dy}\psi(g(y)) = 0 \quad \text{if and only if} \quad \frac{d}{d\alpha}\psi(\alpha) = 0.$$

From

$$\psi(g(y)) = \theta^2 y^2 + \frac{1}{1+2y},$$

we obtain

$$\frac{d}{dy}\psi(g(y)) = 2\theta^2 y - \frac{2}{(1+2y)^2},$$

and the critical points are found at $(1+2y)^2 y = 1/\theta^2$. Notice also that in such a point we have that

$$\psi(g(y)) = \theta^2 y^2 + \frac{1}{1+2y} = \frac{y}{(1+2y)^2} = \frac{3y+1}{1+4y+4y^2} < 1.$$

so at the optimal α we will have

$$\frac{\alpha d M}{2-\alpha} + \frac{(1-\alpha)^2 d K^2}{\alpha^2} = dM\psi(\alpha) < dM.$$

An application of Cardano's formula entails that the unique real positive solution to $(1+2y)^2 y = 1/\theta^2$ is given by

$$y = \frac{1}{3}\left[-1 + \frac{1}{2\vartheta} + \frac{\vartheta}{2}\right],$$

where $\vartheta := (1 + 27/\theta^2 + 3\sqrt{3}\sqrt{2/\theta^2 + 27/\theta^4})^{1/3}$. Thus, the explicit value of the optimal α is

$$\alpha = \frac{1}{1 + -1/3 + 1/6\vartheta + \vartheta/6} = \frac{6\vartheta}{\vartheta^2 + 4\vartheta + 1}$$

and at this value $0 < \alpha < 1$, since $\vartheta > 0$ for any θ, and

$$6\vartheta < \vartheta^2 + 4\vartheta + 1 \iff (\vartheta - 1)^2 > 0.$$

6.2 Proof of Corollary 3.1

From Theorem 3.4, substituting dM with $d\sigma^2$, we get

$$\limsup_{t\to\infty} \|\lambda_{t+1} - \lambda_t^\star\|^2_{\mathbb{L}^2(\Omega)} \leq \frac{\alpha d\sigma^2}{2-\alpha} + \frac{(1-\alpha)^2 d K^2}{\alpha^2};$$

therefore, the statement will follow if we prove that

$$\frac{\alpha d\sigma^2}{2-\alpha} + \frac{(1-\alpha)^2 d K^2}{\alpha^2} \leq d\sigma^2. \tag{6.7}$$

But inequality (6.7) is equivalent to

$$\frac{K^2}{\sigma^2} \leq \frac{2\alpha^2}{(1-\alpha)(2-\alpha)},$$

and, since the right hand side above diverges to $+\infty$ as α approaches 1, we can certainly find, for any K and σ^2, a value in the interval $(0, 1)$ for the learning rate α that fulfills inequality (6.7).

6.3 Proof of Theorem 3.5

Let $\sigma_t^2 = d\sigma^2$ for all $t \in \mathbb{N}$; then, repeating the same calculations of Theorem 3.4, we deduce that S_t solves

$$\begin{aligned} &S_{t+1} = \beta^2 S_t + \beta^{2+2t}\left[\|c_t\|^2 - \|c_{t-1}\|^2\right] + \alpha^2 d\sigma^2; \\ &S_1 = 0, \end{aligned}$$

which has solution

$$S_t = \alpha^2 \sum_{i=0}^{t-1} \beta^{2i} d\sigma^2 + \beta^{2t}\|c_{t-1}\|^2 = \frac{\alpha^2 d\sigma^2(1-\beta^{2(t-1)})}{1-\beta^2} + \beta^{2t}\|c_{t-1}\|^2.$$

Now, by assumption we have that $\limsup_{t\to\infty} S_t < d\sigma^2$; this means that there exists an $N \in \mathbb{N}$ such that $t > N$ implies $S_t < d\sigma^2$, i.e,

$$\beta^{2t}\|c_{t-1}\|^2 < d\sigma^2\left[\frac{2\alpha\beta}{1-\beta^2}\right] + \frac{\alpha^2 d\sigma^2 \beta^{2(t-1)}}{1-\beta^2},$$

where we have utilized the relation $\beta = 1-\alpha$. Moreover, if $t-1 > N$, we also have

$$\beta^{2t+2}\|c_t\|^2 < d\sigma^2\left[\frac{2\alpha\beta}{1-\beta^2}\right] + \frac{\alpha^2 d\sigma^2 \beta^{2t}}{1-\beta^2}. \tag{6.8}$$

From the definition of c_t we get

$$\beta^{2t+2}\|c_t\|^2 = \beta^2 \left\| K_t + \sum_{h=1}^{t-1} \beta^{t-h} K_h \right\|^2, \quad \beta^{2t}\|c_{t-1}\|^2 = \left\| \sum_{h=1}^{t-1} \beta^{t-h} K_h \right\|^2,$$

thus, from (6.8), we obtain

$$\left\| K_t + \sum_{h=1}^{t-1} \beta^{t-h} K_h \right\|^2 < d\sigma^2\left[\frac{2\alpha}{\beta(1-\beta^2)}\right] + \frac{\alpha^2 d\sigma^2 \beta^{2t-2}}{1-\beta^2}.$$

Therefore,

$$\begin{aligned}\|K_t\| &= \left\| K_t + \sum_{h=1}^{t-1} \beta^{t-h} K_h - \sum_{h=1}^{t-1} \beta^{t-h} K_h \right\| \\ &\leq \left(d\sigma^2 \left[\frac{2\alpha}{\beta(1-\beta^2)} \right] + \frac{\alpha^2 d\sigma^2 \beta^{2t-2}}{1-\beta^2} \right)^{1/2} \\ &\quad + \left(d\sigma^2 \left[\frac{2\alpha\beta}{1-\beta^2} \right] + \frac{\alpha^2 d\sigma^2 \beta^{2t-2}}{1-\beta^2} \right)^{1/2},\end{aligned}$$

6.4 Proof of Theorem 3.7

We subtract $\lambda_t^\star$ from both sides of equation (3.9)

$$\lambda_{t+1} - \lambda_t^\star = (1-\alpha)(\lambda_t - \lambda_t^\star) + \alpha(T(X_t) + \epsilon_t - \lambda_t^\star),$$

and we take the square to get

$$\begin{aligned}(\lambda_{t+1} - \lambda_t^\star)^2 &= (1-\alpha)^2(\lambda_t - \lambda_t^\star)^2 + 2\alpha(1-\alpha)(\lambda_t - \lambda_t^\star)(T(X_t) - \lambda_t^\star + \epsilon_t) \\ &\quad + \alpha^2(T(X_t) + \epsilon_t - \lambda_t^\star)^2.\end{aligned}$$

Taking conditional expectations (with respect to the sigma algebra generated by both the X_t and ϵ_t up to time $t-1$) we obtain

$$\begin{aligned}\mathbb{E}[(\lambda_{t+1} - \lambda_t^\star)^2 | \mathcal{F}_{t-1}] &= (1-\alpha)^2 \mathbb{E}[(\lambda_t - \lambda_t^\star)^2 | \mathcal{F}_{t-1}] \\ &\quad + 2\alpha(1-\alpha)(\lambda_t - \lambda_t^\star)m + \alpha^2(\sigma^2 + m_1)\end{aligned} \tag{6.9}$$

here we exploited the independence of the observations, the independence of the noise terms from the observations, and Assumption 3.1.

We now bound the double product term of equation (6.9) as follows

$$2\alpha(1-\alpha)(\lambda_t - \lambda_t^\star)m \leq \alpha(1-\alpha)((\lambda_t - \lambda_t^\star)^2 + m^2),$$

resulting in

$$\mathbb{E}[(\lambda_{t+1} - \lambda_t^\star)^2 | \mathcal{F}_{t-1}] = (1-\alpha)\mathbb{E}[(\lambda_t - \lambda_t^\star)^2 | \mathcal{F}_{t-1}] + \alpha(1-\alpha)m^2 + \alpha^2(\sigma^2 + m_1).$$

At this point, taking total expectations, calling $\tilde{\sigma}^2 := \alpha(1-\alpha)m^2 + \alpha^2(\sigma^2 + m_1)$, and utilizing the same notation that was introduced in the proof of Theorem 3.4 we have

$$S_{t+1} = \beta \mathbb{E}[(\lambda_t - \lambda_t^\star)^2] + \tilde{\sigma}^2.$$

So we can follow the same derivation as in the proof of Theorem 3.4 to obtain

$$\limsup_{t\to\infty} \mathbb{E}[(\lambda_{t+1} - \lambda_t^\star)^2] \leq \alpha(\sigma^2 + m_1) + (1-\alpha)m^2 + \frac{1-\alpha}{\alpha^2}\left(K + \frac{\alpha}{1-\alpha}|m|\right)^2.$$

Funding Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement.

References

1. Akaike, H.: Information theory and an extension of the maximum likelihood principle. In: E. Parzen, K. Tanabe, G. Kitagawa (eds.) Selected Papers of Hirotugu Akaike, pp. 199–213. Springer New York, New York, NY (1998). https://doi.org/10.1007/978-1-4612-1694-0_15
2. Anton, S.: Evaluating the forecasting performance of garch models. evidence from romania. Procedia - Social and Behavioral Sciences **62**, 1006–1010 (2012).https://doi.org/10.1016/j.sbspro.2012.09.171
3. Bishop, C.M.: Pattern Recognition and Machine Learning, vol. 4. Springer (2006)
4. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics **31**(3), 307–327 (1986). https://doi.org/10.1016/0304-4076(86)90063-1
5. Bottou, L., Curtis, F., Nocedal, J.: Optimization methods for large-scale machine learning. SIAM Review **60**(2), 223–311 (2018). https://doi.org/10.1137/16M1080173
6. Bubeck, S., Stoltz, G., Szepesvári, C., Munos, R.: Online optimization in x-armed bandits. In: D. Koller, D. Schuurmans, Y. Bengio, L. Bottou (eds.) Advances in Neural Information Processing Systems, vol. 21. Curran Associates, Inc. (2008). https://proceedings.neurips.cc/paper_files/paper/2008/file/f387624df552cea2f369918c5e1e12bc-Paper.pdf
7. Cao, X., Zhang, J., Poor, H.V.: On the time-varying distributions of online stochastic optimization. In: 2019 American Control Conference (ACC), pp. 1494–1500 (2019). https://doi.org/10.23919/ACC.2019.8814889
8. Cavraro, G., Dall'Anese, E., Comden, J., Bernstein, A.: Online state estimation for time-varying systems. IEEE Transactions on Automatic Control **67**(10), 5424–5431 (2021)
9. Cesa-Bianchi, N., Lugosi, G.: Prediction, Learning, and Games. Cambridge University Press (2006)
10. Cutler, J., Drusvyatskiy, D., Harchaoui, Z.: Stochastic optimization under time drift: iterate averaging, step-decay schedules, and high probability guarantees. In: M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, J.W. Vaughan (eds.) Advances in Neural Information Processing Systems, vol. 34, pp. 11859–11869. Curran Associates, Inc. (2021). https://proceedings.neurips.cc/paper/2021/file/62e7f2e090fe150ef8deb4466fdc81b3-Paper.pdf
11. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the em algorithm. Journal of the Royal Statistical Society: Series B (Methodological) **39**(1), 1–22 (1977)
12. Hansen, P.R., Lunde, A.: A forecast comparison of volatility models: does anything beat a garch(1,1)? Journal of Applied Econometrics **20**(7), 873–889 (2005) https://doi.org/10.1002/jae.800. https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.800
13. Jadbabaie, A., Rakhlin, A., Shahrampour, S., Sridharan, K.: Online Optimization : Competing with Dynamic Comparators. In: G. Lebanon, S.V.N. Vishwanathan (eds.) Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics, *Proceedings of Machine Learning*

Research, vol. 38, pp. 398–406. PMLR, San Diego, California, USA (2015). https://proceedings.mlr.press/v38/jadbabaie15.html
14. Kushner, H., Yin, G.: Stochastic Approximation and Recursive Algorithms and Applications. Stochastic Modelling and Applied Probability. Springer New York (2003). https://books.google.it/books?id=EC2w1SaPb7YC
15. Lanconelli, A., Lauria, C.S.A.: Maximum likelihood with a time varying parameter. Statistical Papers **65**(4), 2555–2566 (2024). https://doi.org/10.1007/s00362-023-01497-y
16. Lemaréchal, C.: Cauchy and the gradient method. Doc Math Extra **251**(254), 10 (2012)
17. Lima, F., Pimenta Junior, T., Gaio, L.: Volatility behaviour of bric capital markets in the 2008 international financial crisis. African Journal of Business Management **8**, 373 (2014). https://doi.org/10.5897/AJBM2013.7162
18. Ljung, L., Gunnarsson, S.: Adaptation and tracking in system identification-a survey. Automatica **26**(1), 7–21 (1990) https://doi.org/10.1016/0005-1098(90)90154-A. https://www.sciencedirect.com/science/article/pii/000510989090154A
19. Madden, L., Becker, S., Dall'Anese, E.: Bounds for the tracking error of first-order online optimization methods. Journal of Optimization Theory and Applications **189**, 437–457 (2021)
20. Mikosch, T., Stric, C.: Limit theory for the sample autocorrelations and extremes of a garch(1,1) process. The Annals of Statistics **28**(5), 1427–1451 (2000) https://doi.org/10.1214/aos/1015957401. https://projecteuclid.org/euclid.aos/1015957401
21. Nesterov, Y.: Introductory Lectures on Convex Optimization: A Basic Course, 1st edn. Springer Publishing Company, Incorporated (2014)
22. Polyak, B.T.: Introduction to Optimization. Translations Series in Mathematics and Engineering. Optimization Software (1987). https://www.amazon.com/dp/0911575146
23. Popkov, A.Y.: Gradient methods for nonstationary unconstrained optimization problems. Automation and Remote Control **66**(6), 883–891 (2005). https://doi.org/10.1007/s10513-005-0132-z
24. Robbins, H., Monro, S.: A stochastic approximation method. The Annals of Mathematical Statistics **22**(3), 400–407 (1951). http://www.jstor.org/stable/2236626
25. Simonetto, A., Dall'Anese, E., Paternain, S., Leus, G., Giannakis, G.B.: Time-varying convex optimization: Time-structured algorithms and applications. Proceedings of the IEEE **108**(11), 2032–2048 (2020)
26. Simonetto, A., Dall'Anese, E.: Prediction-correction algorithms for time-varying constrained optimization. IEEE Transactions on Signal Processing **65**(20), 5481–5494 (2017). https://doi.org/10.1109/TSP.2017.2728498
27. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction, second edn. The MIT Press (2018). http://incompleteideas.net/book/the-book-2nd.html
28. Thanwerdas, Y., Pennec, X.: O (n)-invariant riemannian metrics on spd matrices. Linear Algebra and its Applications **661**, 163–201 (2023)
29. White, H.: Maximum Likelihood Estimation of Misspecified Models. Econometrica **50**(1), 1–25 (1982). https://ideas.repec.org/a/ecm/emetrp/v50y1982i1p1-25.html
30. Wilson, C., Veeravalli, V.V., Nedić, A.: Adaptive sequential stochastic optimization. IEEE Transactions on Automatic Control **64**(2), 496–509 (2019). https://doi.org/10.1109/TAC.2018.2816168
31. Wojtowytsch, S.: Stochastic gradient descent with noise of machine learning type part i: Discrete time analysis. Journal of Nonlinear Science **33**(3), 45 (2023)
32. Zinkevich, M.: Online convex programming and generalized infinitesimal gradient ascent. In: T. Fawcett, N. Mishra (eds.) Proceedings of the 20th International Conference on Machine Learning (ICML-2003), pp. 928–936. AAAI Press (2003)