Chapter 4

Stochastic Approximation, Endogenous Noise Model

4.1 The Endogenous Noise Model

The first two chapters dealt with analysis of algorithms of the form

$$\theta_{n+1} = \theta_n + \epsilon_n G(\theta_n)$$

designed to "find the zeroes" of a function G, under appropriate assumptions. We will assume here that there is no closed expression for $G(\theta)$, but that:

- we can build *estimators* for $G(\theta)$, and that
- the control variable can be changed at will, while the process is operating (or being simulated).

Chapter 3 presented a methodology to study stochastic approximation of the form

$$\theta_{n+1} = \theta_n + \epsilon_n Y_n \tag{4.1}$$

as $n \to \infty$. The tools of the analysis relied heavily on the decomposition in (3.19). In this chapter we provide the basis for convergence analysis of the same algorithm when the estimates of the function $G(\cdot)$ cannot not be decomposed like in (3.19). Next we provide an illustrating example for this phenomenon.

EXAMPLE 4.1. Suppose that an exercise machine can adjust the resistance $\theta \in [0.5, 1]$ while a person is exercising. Were θ to be kept fixed, the heart rate of the person would attain a steady value $L(\theta)$. The goal is to program an algorithm for the machine's computer such that if the person sets a target heart rate of α, the resistance will adjust to reach the value θ^* such that $L(\theta^*) = \alpha$.

The machine cannot accurately measure the heart rate instantaneously, but instead, the number of heart beats (or pulses) in a time interval of fixed length may be used to estimate for $L(\theta)$. As an example, suppose that for a given person and environment conditions, ξ_n is the number of heart beats in an interval of .3 sec. In order to work this example, we assume the following over-simplified Markov model for the heart beats:

$$\mathbb{P}(\xi_{n+1} = 1 \mid \xi_n = 0, \theta) = p_{0,1}(\theta),$$
$$\mathbb{P}(\xi_{n+1} = 1 \mid \xi_n = 1, \theta) = p_{1,1}(\theta).$$

For each fixed value of θ, the heart rate is defined as the long term average number of heart beats per minute, (200 intervals of .3 sec each):

$$\lim_{N\to\infty}\frac{200}{N}\sum_{n=1}^{N}\xi_n = \lim_{N\to\infty}\frac{200}{N}\sum_{n=1}^{N}\mathbf{1}_{\{\xi_n=1\}} = \lim_{N\to\infty}\frac{200}{N}\sum_{n=1}^{N}\mathbb{E}[\mathbf{1}_{\{\xi_n=1\}}\,|\,\theta] = 200\mu_\theta(1).$$

That is, the heart rate can be conveniently expressed in terms of the stationary distribution $\mu_\theta^\top = (\mu_\theta(0), \mu_\theta(1))$ of the Markov chain, yielding

$$L(\theta) = \mu_\theta(1)\times(60/.3) = 200\mu_\theta(1)$$

as the heart rate (per minute). Using the target tracking formulation we propose the vector field

$$G(\theta) = -(L(\theta)-\alpha) = \alpha - 200\mu_\theta(1).$$

A straightforward choice for Y_n in (4.1) is

$$Y_n = -(200\times\mathbf{1}_{\{\xi_n=1\}} - \alpha).$$

Provided the information up to time $n-1$, the expected outcome of Y_n is

$$\begin{aligned}\mathbb{E}[Y_n\,|\,\xi_{n-1},\theta_n] &= -200\,\mathbb{P}(\xi_n=1\,|\,\xi_{n-1},\theta_n)+\alpha\\ &= -200\,p_{\xi_{n-1},1}(\theta_n)+\alpha\\ &= G(\theta_n)+\beta_n, \end{aligned}\tag{4.2}$$

which imples that

$$\beta_n = 200\left(\mu_{\theta_n}(1) - p_{\xi_{n-1},1}(\theta_n)\right).$$

The exogenous noise approach in Chapter 3 requires that

$$\sum_i \epsilon_i|\beta_i| < \infty,$$

with $\sum_i \epsilon_i = \infty$, which is not satisfied in this example. To see this, notice that $\beta_n/200$ is a binary random variable depending on both θ_n and ξ_n that has the distribution

$$\beta_n/200 = \begin{cases}\mu_{\theta_n}(1) - p_{0,1}(\theta_n) & w.p.\ \mathbb{P}(\xi_{n-1}=0)\\ \mu_{\theta_n}(1) - p_{1,1}(\theta_n) & w.p.\ \mathbb{P}(\xi_{n-1}=1)\end{cases}$$

We observe that the only way that the bias term is null is the case when $p_{0,1}(\theta) = p_{1,1}(\theta) = \mu_\theta(1)$ which corresponds to a "degenerate" Markov chain, since in this case ξ_n is independent of the past. This is the model of independent noise (exogenous) of Chapter 3. Otherwise the values for β_n are non zero and $|\beta_n| \geq 200\,b(\theta_n)$, where the bound $b(\theta_n) = \max(|\mu_{\theta_n}(1) - p_{0,1}(\theta_n)|, |\mu_{\theta_n}(1) - p_{1,1}(\theta_n)|) > 0$.

However, this does by no means imply that the algorithm (4.1) fails to converge to the optimal value. Indeed, as we will show in this chapter, using a more elaborate method of proof we can relax the assumption on $\mathbb{E}[Y_n\,|\,\mathfrak{F}_{n-1}]$ and thereby establish sufficient conditions for convergence.

⁂

Underlying Process: We begin by assuming that for each fixed value of the control variable $\theta \in \Theta$, there is an *underlying* stochastic process $\{\xi_n(\theta)\}$ on $(\Omega, \{\mathfrak{F}_n(\theta)\}, \mathbb{P}_\theta)$ modelling the randomness inherent in the dynamics of the system of interest, where the subscript θ expresses the θ-dependence of the system. This underlying process $\{\xi_n(\theta)\}$ is assumed to be Markovian, that is

$$P_\theta(\xi_n, \cdot) \stackrel{\text{def}}{=} \mathbb{P}_\theta(\xi_{n+1} \in \cdot \mid \xi_n) = \mathbb{P}_\theta(\xi_{n+1} \in \cdot \mid \mathfrak{F}_n(\theta)),$$

and we assume that there is a unique stationary measure $\mu_\theta(\cdot)$. We will use the notation $\mathbb{P}_\theta, \mathbb{E}_\theta$ to indicate probability and expectations with respect to the fixed-θ process $\{\xi_n(\theta)\}$.

Target Vector Field: We assume that the target vector field satisifes almost surely

$$G(\theta) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} g(\xi_n(\theta), \theta) = \int g(\xi, \theta)\, \mu_\theta(d\xi), \tag{4.3}$$

where $g(\xi, \theta)$ is is a well defined random function.

EXAMPLE 4.2. Going back to the motivating example at the beginning of this chapter, we can calculate here

$$g(\xi_{n-1}(\theta), \theta) \stackrel{\text{def}}{=} \mathbb{E}[Y_n \mid \mathfrak{F}_{n-1}(\theta)] = -(200\, p_{\xi_{n-1}(\theta),1}(\theta) - \alpha),$$

for the process with fixed value of θ, and by definition of stationary measure,

$$G(\theta) = \int g(x, \theta)\mu(dx) = -\sum_x \mu_\theta(x)(200\, p_{x,1}(\theta) - \alpha) = -200\, \mu_\theta(1) + \alpha = -(L(\theta) - \alpha).$$

⁂

The Model for Y_n: When θ_n varies according to (4.1) we consider the augmented process $\{(\xi_{n-1}, \theta_n)\}$ that satisfies

$$\mathbb{P}(\xi_{n+1} \in B \mid \xi_n, \theta_{n+1}) = P_{\theta_{n+1}}(\xi_n, B),$$

for any measurable set B, and denote the natural filtration of $\{\xi_n\}$ by $\{\mathfrak{F}_n\}$, where we drop the notation of θ for notational convenice. The estimator Y_n is measurable with respect to $\mathfrak{F}_n$, that is, Y_n is constructed as a function of the observed values $(\xi_0, \ldots, \xi_n)$ (the *history* of the process). Because θ_n is a function of θ_{n-1} and of Y_{n-1} it follows that θ_n is measurable with respect to $\mathfrak{F}_{n-1}$.

For our analysis we assume the Markovian property

$$\mathbb{E}[Y_n \mid \mathfrak{F}_{n-1}] = g(\xi_{n-1}, \theta_n) + \beta_n$$

where the error terms β_n are well defined random variables, measurable wrt $\mathfrak{F}_{n-1}$.

EXAMPLE 4.3. Going back to the Example 4.2, we can calculate here

$$g(\xi_{n-1}, \theta_n) \stackrel{\text{def}}{=} \mathbb{E}[Y_n \mid \mathfrak{F}_{n-1}] = -(200\, p_{\xi_{n-1},1}(\theta_n) - \alpha),$$

for $n \geq 0$. Comparing with (4.2) we see that replacing $G(\theta_n)$ by $g(\xi_{n-1}, \theta_n)$ leads to zero bias.

⁂

Comparing with the exogenous model in (3.5), the main difference is that the vector field $G(\theta_n)$ is replaced by $g(\xi_{n-1}, \theta_n)$. The name *endogenous noise* reflects this fact because Y_n cannot be simply determined based on the natural filtration of the Y_n process up to $n-1$ as it was the case in exogenous noise model.

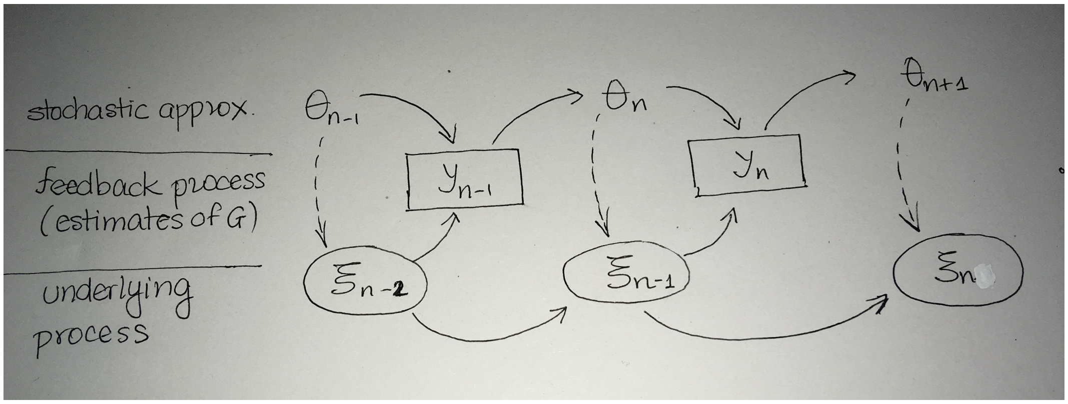


Figure 4.1: Schematic representation of processes in endogenous noise model.

Definition 4.1 *The* endogenous noise model *is determined by a parametrized family of Markov chains* $\{\xi_n(\theta)\}$, $\theta \in \Theta \subset \mathbb{R}^d$ *with a unique stationary measure* $\mu_\theta(\cdot)$ *for each value of* θ. *The target vector field satisfies* $G(\theta) = \int g(\xi, \theta)\, \mu_\theta(d\xi)$ *a.s., for some* instantaneous cost $g(\xi, \theta)$. *Consider the stochastic approximation*

$$\theta_{n+1} = \theta_n + \epsilon_n Y_n,$$

so that the underlying process is a Markov Decision Process with

$$\mathbb{P}(\xi_{n+1} \in B \,|\, \xi_n, \theta_{n+1}) = P_{\theta_{n+1}}(\xi_n, B),$$

and has natural filtration $\{\mathfrak{F}_n\}$. *The* feedback *satisfies*

$$\mathbb{E}[Y_n \,|\, \mathfrak{F}_{n-1}] = g(\xi_{n-1}, \theta_n) + \beta_n$$

where the error terms β_n *are well defined random variables, measurable wrt* $\mathfrak{F}_{n-1}$.

The following Theorem will be stated without proof. The proof is an extension to that of Theorem 3.2, and it uses the methodology of Chapter 6 of [21].

Theorem 4.1 *Consider the endogenous noise model for the underlying process* $\{(\xi_n, \theta_n)\}$, *where* θ_n *is updated according to (3.4). Assume the following conditions hold:*

(a1) The transition probabilities $P_\theta(\xi, \cdot)$ *are continuous in* θ, ξ. *The stationary measure of the fixed-*θ *process* $\{\xi(\theta)\}$, *denoted by* μ_θ, *is unique, and the set* $\{\mu_\theta(\cdot), \theta \in \Theta\}$ *of stationary measures of the fixed-*θ *processes is tight for each compact set* Θ.

(a2) There is a jointly continuous measurable function $g(\xi, \theta)$ *(it may depend on* n*, but we will introduce the simplest model here), such that*

$$\mathbb{E}[Y_n \,|\, \mathfrak{F}_{n-1}] = g(\xi_{n-1}, \theta_n) + \beta_n,$$

and there exists a continuous and bounded function $G \colon \mathbb{R}^d \to \mathbb{R}^d$ *such that, for every fixed* θ,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} g(\xi_n(\theta), \theta) = G(\theta), \qquad w.p.1$$

and $\sup_{\theta\in\Theta} G(\theta) < \infty$.

(a3) $V_n = \mathbb{E}[(Y_n - \mathbb{E}(Y_n \,|\, \mathfrak{F}_{n-1}))^2]$ *satisfies:* $\sum_{i=0}^{\infty} \epsilon_i^2 V_i < \infty$.

(a4) $\sum_{i=0}^{\infty} \epsilon_i = +\infty$.

(a5) The error terms are asymptotically negligible, that is: $\sum_{i=0}^{\infty} \epsilon_i \|\beta_i\| < \infty \; w.p.1$

(a6) The ODE

$$\frac{d\vartheta(t)}{dt} = G(\vartheta(t)) \tag{4.4}$$

has a unique limit for each initial condition. Let S denote the set of stable points of this ODE and assume that $S \neq \emptyset$.

Then there exists a null set $N \subset \Omega$ such that, for every $\omega \neq N$, $\{\theta_n(\omega)\}$ converges, and $\theta_n(\omega) \to S$. In particular, if θ^* is the only asymptotically stable point of the ODE, then $\theta_n \to \theta^*$ a.s.

EXAMPLE 4.4. Consider again the problem of tracking the heart rate while a person is exercising, with parameters given by the following simple model:

$$\mathbb{P}(\xi_{n+1} = 1 \,|\, \xi_n = 0, \theta) = \theta,$$
$$\mathbb{P}(\xi_{n+1} = 1 \,|\, \xi_n = 1, \theta) = \theta^2.$$

For any $\theta \in [0.5, 1]$, the fixed-θ process is a recurrent Markov chain with state space $\{0, 1\}$. We calculate the stationary measure:

$$\mu_\theta(1) = \mu_\theta(1)\theta^2 + (1 - \mu_\theta(1))\theta \Longrightarrow \mu_\theta(1) = \frac{\theta}{1 + \theta(1 - \theta)}.$$

Thus the function $L(\theta) = 200\,\mu_\theta(1)$ is increasing in θ. To see this, it suffices to verify that

$$\frac{d\mu_\theta(1)}{d\theta} = \frac{1 + \theta - \theta^2 - (\theta - \theta^2)}{(1 + \theta - \theta^2)^2} = \frac{1 + \theta^2}{(1 + \theta - \theta^2)^2} > 0.$$

Thus, for any $\alpha \in [80, 200]$ there is a unique value θ^* such that $L(\theta^*) = \alpha$.

Assumptions (a1) and (a2) are verified for this problem as can be seen in Example 4.2. In this case the measurements Y_n have no bias terms β_n and they are absolutely bounded by 200, so they have uniformly bounded variance. Any stepsize sequence satisfying the standard conditions, see (3.2), will ensure (a3), (a4) and (a5).

Because $L(\theta)$ is continuously differentiable and bounded, then (a6) is satisfied. For each $\alpha \in [80, 200]$, there is a unique stable point of (4.4). Therefore our learning algorithm (3.4) will converge to the optimal setting - provided that the person exercises long enough, of course!

⁂

4.2 Constant Step Size, Weak Convergence

In this section we study the stochastic version of (2.8) with constant stepsize:

$$\theta_{n+1}^{\epsilon} = \theta_n^{\epsilon} + \epsilon\, Y_n^{\epsilon}, \tag{4.5}$$

where $\epsilon > 0$ is kept constant for all $n \in \mathbb{N}$. Given our analysis in Section 3.3, it is apparent that the Martingale M_n and error noise processes B_n cannot possibly converge to zero a.s., as $n \to \infty$, with non decreasing stepsizes.

In order to find the zeroes of $G(\cdot)$, under the appropriate conditions, Theorem 2.3 establishes that the iterative method (2.8) with constant stepsize approaches an ODE whose limits are the desired stationary points. As in Theorem 2.3, the sense in which the algorithm (4.5) converges is in terms of the interpolated processes, as $\epsilon \to 0$ (not as $n \to \infty$). In addition, the stochastic processes converge *in distribution* to the solution of the ODE. In this section we present an analysis of the constant stepsize algorithm that "averages out" the noise to recover the result of the deterministic case Theorem 2.3.

EXAMPLE 4.5. To illustrate the concept of convergence, we revisit here the problem of finding an economic equilibrium of Exercise 3.6. In that example, using the data given in the exercise for $N = 100$, the ODE can be approximated by a deterministic recursion $\theta_{n+1} = \theta_n + \epsilon G(\theta_n)$. Figure 4.2 demonstrating this result was obtained at $\epsilon = 0.5$ using Mathematica.

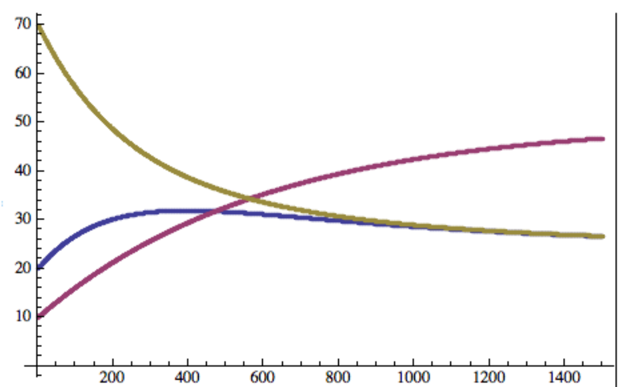


Figure 4.2: Trajectory of the ODE

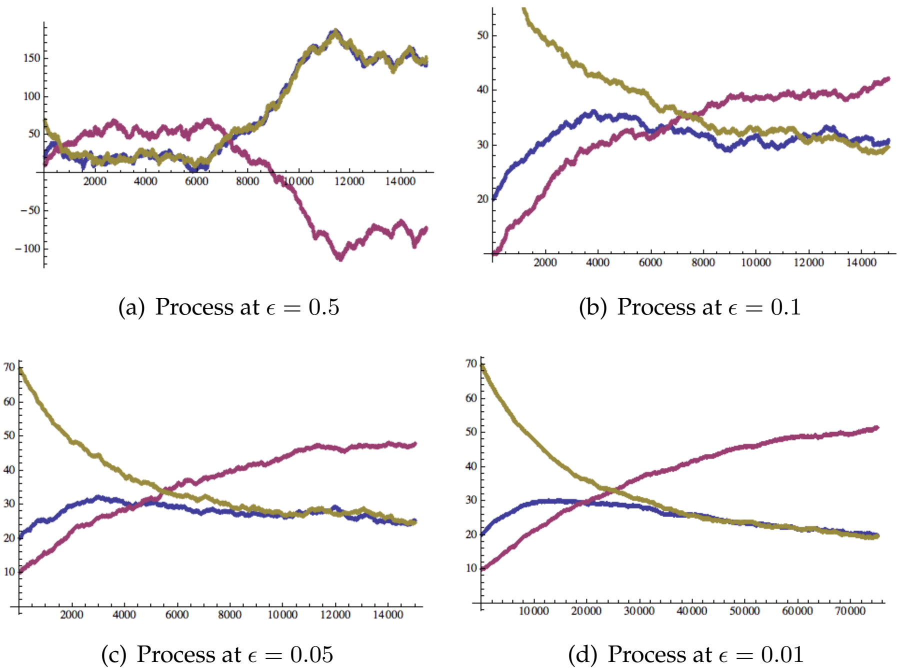
Suppose that we do not know the exact formulas for the individual travel times given the allocation θ, only that they are affine functions of θ. Estimates can be obtained by simulation, which adds a random non-negative noise with Gamma distribution. That is,

$$\widehat{T}_i(\theta) = T_i(\theta) + \gamma_i,$$

where $\gamma_i, i = 1, 2, 3$ are independent with Gamma distribution with parameters $(2, 0.5)$, so that $\mathbb{E}[\gamma_i] = 1$. Notice that although this creates biased estimates for T, the noise in the estimation of G has zero mean, using:

$$Y_n^\epsilon(i) = -\left(\widehat{T}_i(\theta_n^\epsilon) - \frac{1}{3}\sum_k \widehat{T}_k(\theta_n^\epsilon)\right),$$

The following graphs were obtained using decreasing $\epsilon = 0.5, 0.1, 0.05, 0.01$. The behaviour is very erratic for $\epsilon = 0.5$ compared to the ODE, even for this relatively small error. Because we are not truncating the iterates to ensure that the traffic flow be non negative, the sequence θ_n^ϵ had negative components in numerous experiments, and when this happens, it does not return to positive values. In contrast, when the deterministic algorithm is started at negative values, the sequence always comes back to the stable point.



(a) Process at $\epsilon = 0.5$ (b) Process at $\epsilon = 0.1$

(c) Process at $\epsilon = 0.05$ (d) Process at $\epsilon = 0.01$

Figure 4.3: Results of constant step size processes, as the step size is made smaller

This example also illustrates the relationship between the time scale of the limit ODE and the value of ϵ. Because the time is defined in terms of $t = n\epsilon$, the number of iterations needed to attain the same time grows as $1/\epsilon$.

***

We will work with the (more general) endogenous noise model of Section 4.1. Because the exogenous noise model is a particular case of this more general model, our results extend to that case as well.

In this section we will follow the common notation adding the superindex ϵ to the various variables to stress the fact that ϵ is the parameter with respect to which we will study the limiting behaviour. The process $\{(\xi_n^\epsilon, \theta_n^\epsilon)\}$ defines the structure of the model as a Markov process, where $\mathfrak{F}_n^\epsilon = \sigma((\theta_{i+1}^\epsilon, \xi_i^\epsilon); i = 0, \ldots, n)$ is the filtration, the transition probabilities are given by:

$$\mathbb{P}(\xi_{n+1}^\epsilon \in B \,|\, \mathfrak{F}_n^\epsilon) = P_{\theta_{n+1}^\epsilon}(\xi_n^\epsilon, B),$$

for any measurable set $B \subset \Omega$ and θ_{n+1}^ϵ updates according to (4.5), where we assume that Y_n^ϵ is measurable w.r.t. $\mathfrak{F}_n^\epsilon$. We will use the notation:

$$\mathbb{E}\left[Y_n^\epsilon \,|\, \mathfrak{F}_{n-1}^\epsilon\right] = g(\xi_{n-1}^\epsilon, \theta_n^\epsilon), \tag{4.6}$$

for a well defined function g (it may depend on n, but we will introduce the simplest model here).

Definition 4.2 *The* interpolation processes $\vartheta^\epsilon(\cdot)$ *are defined for each* $\epsilon > 0$ *by:*

$$\vartheta^\epsilon(t) = \theta_n^\epsilon, \quad t \in [t_n, t_{n+1}), \quad t_n = n\,\epsilon \tag{4.7}$$

and the iteration counting process *of (3.13) becomes* $m(t) = \lfloor t/\epsilon \rfloor$.

The following result is the extension of Theorem 2.3 and the reader is encouraged to revise the proof in Chapter 2 that prepares much of the ground work to be done in this section where the noise structure is endogenous. This creates much complexity in the treatment of the algorithm, and it is somewhat surprising that one can recover the same result as for the deterministic case with relatively general conditions on the stochastic model.

As mentioned in Chapter 2, the ODE limit provides more insight about the iterative method than results that characterise only the limit points $\lim_{n\to\infty} \theta_n$. The limit occurs as $\epsilon \to 0$, which of course is never attained (not unlike $n \to \infty$, which is never attained either), and so we must interpret the limit of θ_n in terms of the algorithms being "close" to a deterministic trajectory when "ϵ is small enough", as illustrated in Example 4.5. When the conditions of Theorem 4.2 below hold, we know that our iterative method is close to the solution of

$$\frac{dx(t)}{dt} = G(x(t)),$$

where the vector field G in (4.3) is a stationary average mean drift of the algorithm. This is helpful in practice in many ways. First, we may wish to define a target or desired ODE that will have as stable points the optimal values of our problem. Second, we may be able to find formulations such as changes in variables or in time scale that will make it more convenient for the estimation of G, or that may achieve faster convergence to the stable points. For example, in the target tracking problem, instead of using the gradient of the distance we simply use $L(\theta) - \alpha$, which is easier to estimate and under monotonicity conditions has the same stable points as the gradient. The ODE trajectories not only determine the limit points, but provide a global understanding on how the algorithm gets there.

Theorem 4.2 *Let $G\colon \mathbb{R}^d \to \mathbb{R}^d$ be a given target vector field. Assume the following conditions hold:*

(a1) The transition probabilities of the fixed-θ processes $P_\theta(\xi, \cdot)$ are weakly continuous (uniformly) wr.t. (θ, ξ). That is, for each bounded and continuous real-valued function f, and any $\rho > 0$, there is a $\delta > 0$ such that:

$$\left| \int f(x) P_\theta(\xi, dx) - \int f(x) P_{\theta'}(\xi', dx) \right| \leq \rho$$

whenever $\|(\xi, \theta) - (\xi', \theta')\| < \delta$.

(a2) The stationary measure μ_θ is unique, and the set $\{\mu_\theta(\cdot), \theta \in \Theta\}$ of stationary measures of the fixed-θ processes is tight for each compact set Θ.

(a3) There exists a continuous function $g(\cdot, \theta)$ such that for any θ it holds with probability one that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} g(\xi_n(\theta), \theta) = \int g(x, \theta) \mu_\theta(dx) = G(\theta),$$

and $\sup_{\theta\in\Theta} G(\theta) < \infty$ for each compact set Θ.

(a4) The sequence $\{Y_n^\epsilon, \epsilon > 0\}$ is uniformly integrable, that is:

$$\sup_{n,\epsilon} \lim_{K\to\infty} \mathbb{E}\Big[\|Y_n^\epsilon\| \mathbf{1}_{\{\|Y_n^\epsilon\| \geq K\}} \Big] = 0.$$

(a5) The set $\{(\xi_n^\epsilon, \theta_n^\epsilon); \epsilon > 0\}$ is tight.

Then the interpolated processes $\vartheta^\epsilon(\cdot)$ of (4.7) converge in distribution, as $\epsilon \to 0$, to a limit process which is continuous w.p.1 and satisfies the ODE (2.10):

$$\frac{dx(t)}{dt} = G(x(t)).$$

REMARK. Assumptions (a1), (a2) and (a3) in Theorem 4.2 refer specifically to the process with fixed θ and are usually easier to verify than assumptions (a4) and (a5). The uniform integrability assumption is often replaced by the weaker condition that $\{Y_n^\epsilon\}$ has a uniformly bounded variance, which implies (a4) and is often easier to verify with most models. For the latter, it is common to use a stochastically dominating process to find uniform bounds. Example 4.6 later on will illustrate this method. The uniform integrability assumption is the stochastic counterpart to the conditions put forward in Theorem 2.4 for establishing the ODE result for all $t \geq 0$.

Proof: The method of proof follows three stages and will be using the same ideas presented in Chapter 2, as well as adequate treatment of the noise.

Telescopic sum and integral representation: From the SA algorithm (4.5),

$$\vartheta^\epsilon(t+s) - \vartheta^\epsilon(t) = \epsilon \sum_{n=m(t)}^{m(t+s)} Y_n^\epsilon.$$

From (a4), the processes $\{\vartheta^\epsilon(\cdot), \epsilon > 0\}$ are tight in the space $D[0,\infty)$ of cadlag processes under the Skorokhod topology. Moreover, using Theorem B.13 any limit process $\vartheta(\cdot)$ as $\epsilon \to 0$ is Lipschitz continuous w.p.1. Theorem B.13 is the stochastic analogue of Ascoli Arzelà Theorem. We now apply the compactness argument of proof. From now until further notice, pick a (weakly) convergent subsequence, labelled by $(\epsilon_k, k \to \infty)$, so that

$$\vartheta^{\epsilon_k}(\cdot) \stackrel{\mathcal{L}}{\Longrightarrow} \vartheta(\cdot)$$

where $\vartheta(\cdot)$ is Lipschitz continuous w.p.1. Because of our already cumbersome indexing notation, we will omit the subindex k for the chosen ϵ-subsequence.

Rewrite now the telescopic sum regrouping terms into larger subintervals, each containing n_ϵ intervals of size ϵ each (see Figure 4.4), that is,

$$\delta_\epsilon = \epsilon\, n_\epsilon,$$

yielding

$$\vartheta^\epsilon(t+s) - \vartheta^\epsilon(t) = \sum_{l=\lfloor t/\delta_\epsilon \rfloor}^{\lfloor (t+s)/\delta_\epsilon \rfloor} \epsilon \sum_{n=ln_\epsilon}^{(l+1)n_\epsilon - 1} Y_n^\epsilon + \rho_1(\epsilon) = \sum_{l=\lfloor t/\delta_\epsilon \rfloor}^{\lfloor (t+s)/\delta_\epsilon \rfloor} \delta_\epsilon \left(\frac{1}{n_\epsilon} \sum_{n=ln_\epsilon}^{(l+1)n_\epsilon - 1} Y_n^\epsilon \right) + \rho_1(\epsilon),$$

where ρ_1 is the end point error from the larger intervals of size δ_ϵ, and it satisfies:

$$\|\rho_1(\epsilon)\| \leq \|\vartheta^\epsilon(t+s) - \vartheta^\epsilon(\delta_\epsilon \lfloor (t+s)/\delta_\epsilon \rfloor)\| + \|\vartheta^\epsilon(t) - \vartheta^\epsilon(\delta_\epsilon \lfloor t/\delta_\epsilon \rfloor)\|.$$

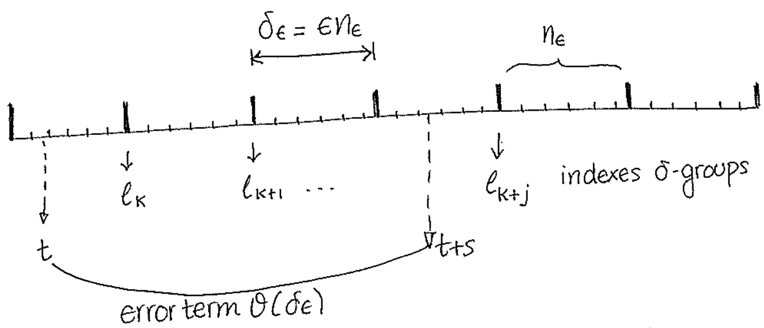
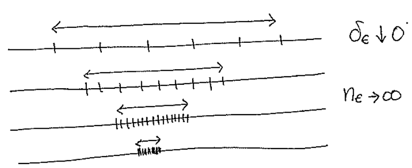


Figure 4.4: Grouping n_ϵ terms into batches for averages.



Choose now the regrouping such that:

$$\lim_{\epsilon\to 0} \delta_\epsilon = 0, \quad \lim_{\epsilon\to 0} n_\epsilon = +\infty,$$

as illustrated in the left, and notice that an obvious choice is $\delta_\epsilon = \sqrt{\epsilon}, n_\epsilon = 1/\sqrt{\epsilon}$.

Under this choice, $\|\rho_1(\epsilon)\| \overset{\mathcal{L}}{\Longrightarrow} 0$, uniformly in t (exercise). If the updates were iid, we could use now the Law of Large Numbers for the averages of the Y_n^ϵ and (a3) to obtain the desired result as an integral. Because our model is more realistic, it allows for feedback dependencies between the control variables θ_n^ϵ and the updating variables Y_n^ϵ, and we proceed with a more careful analysis. By construction, the process $\{\vartheta^\epsilon(\cdot)\}$ is adapted to the filtration $\{\mathfrak{F}^\epsilon_{\lfloor t/\epsilon\rfloor}\}$, and therefore for any indices k, l with $l \geq \lfloor t/\delta_\epsilon\rfloor$ and $k \leq \lfloor t/\epsilon\rfloor$,

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{n_\epsilon}\sum_{n=ln_\epsilon}^{(l+1)n_\epsilon-1} Y_n^\epsilon \,\middle|\, \mathfrak{F}_k^\epsilon\right] &= \mathbb{E}\left[\mathbb{E}\left[\frac{1}{n_\epsilon}\sum_{n=ln_\epsilon}^{(l+1)n_\epsilon-1} Y_n^\epsilon \,\middle|\, \mathfrak{F}^\epsilon_{ln_\epsilon-1}\right] \,\middle|\, \mathfrak{F}_k^\epsilon\right]\\
&= \mathbb{E}\left[\mathbb{E}\left[\frac{1}{n_\epsilon}\sum_{n=ln_\epsilon}^{(l+1)n_\epsilon-1} \mathbb{E}(Y_n^\epsilon \,|\, \mathfrak{F}_{n-1}) \,\middle|\, \mathfrak{F}^\epsilon_{ln_\epsilon-1}\right] \,\middle|\, \mathfrak{F}_k^\epsilon\right],\\
&= \mathbb{E}\left[\mathbb{E}\left[\frac{1}{n_\epsilon}\sum_{n=ln_\epsilon}^{(l+1)n_\epsilon-1} g(\xi^\epsilon_{n-1}, \theta_n^\epsilon) \,\middle|\, \mathfrak{F}^\epsilon_{ln_\epsilon-1}\right] \,\middle|\, \mathfrak{F}_k^\epsilon\right]
\end{aligned}$$

where we have used (4.6). To express our telescopic sum as an integral, define now the piecewise constant function:

$$\mathcal{G}^\epsilon(u) = \frac{1}{n_\epsilon}\sum_{n=ln_\epsilon-1}^{(l+1)n_\epsilon} \mathbb{E}\left[g(\xi^\epsilon_{n-1}, \theta_n^\epsilon \,|\, \mathfrak{F}^\epsilon_{ln_\epsilon-1}\right], \quad \text{for } u \in [l\delta_\epsilon, (l+1)\delta_\epsilon)$$

and notice that $\mathcal{G}^\epsilon(l\delta_\epsilon)$ is a random variable depending on $(\theta^\epsilon_{ln_\epsilon}, \xi^\epsilon_{ln_\epsilon})$. Then we have:

$$\mathbb{E}\left[\vartheta^\epsilon(t+s) - \vartheta^\epsilon(t) \,|\, \mathfrak{F}_{\lfloor t/\epsilon \rfloor})\right] = \int_t^{t+s} \mathcal{G}^\epsilon(u)\, du + \rho_1(\epsilon) \tag{4.8}$$

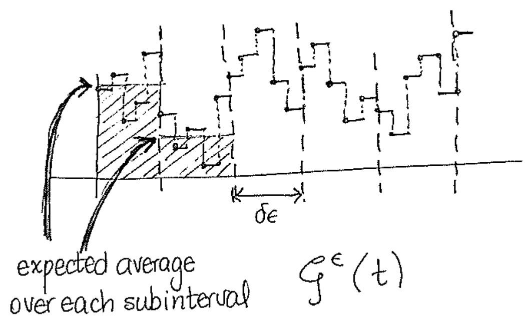


Figure 4.5: Visualization of integral representation.

The idea is to characterize the limit function $\vartheta(\cdot)$ along the chosen convergent subsequence.

The Averaging Result: In this stage of the proof, it is shown that assumptions (a1), (a2), (a3) and (a5) imply that for each u,

$$\mathcal{G}^\epsilon(u) \overset{\mathcal{L}}{\Longrightarrow} G(\vartheta(u)), \tag{4.9}$$

as $\epsilon \to 0$ along the chosen weakly convergent subsequence. We give now the idea of the proof. Fix a value $u \in \mathbb{R}^+$. Then for each $\epsilon > 0$ there is a unique index ℓ_ϵ such that $\ell_\epsilon n_\epsilon \le u < (\ell_\epsilon + 1)n_\epsilon$. Use Skorohod representation Theorem B.12 (Appendix B) to define a sequence $\tilde{\vartheta}^\epsilon(\cdot)$ such that it converges to the chosen limit w.p.1. Because the limit is a.s. Lipschitz continuous and $\delta_\epsilon \to 0$, then

$$\sup\Big(\|\tilde{\vartheta}^\epsilon(s) - \vartheta(u)\| \,:\, s \in [\ell_\epsilon\delta_\epsilon, (\ell_\epsilon+1)\delta_\epsilon)\Big) \to 0,$$

Use now Assumption (a1) and continuity of g to replace the process $\{(\xi^\epsilon_{n-1}, \theta^\epsilon_n)\}$ in the definition of $\mathcal{G}^\epsilon$ by the process at fixed value $\vartheta(u)$, that is,

$$\mathcal{G}^\epsilon(u) \approx \frac{1}{n_\epsilon} \sum_{n=\ell_\epsilon n_\epsilon - 1}^{(\ell_\epsilon+1)n_\epsilon} \mathbb{E}\left[g(\xi_{n-1}(\vartheta(u)), \vartheta(u)) \,|\, \mathfrak{F}^\epsilon_{\ell_\epsilon n_\epsilon - 1}\right].$$

Assumptions (a2) and (a5) are needed in order to ascertain that the initial value $(\xi^\epsilon_{\ell_\epsilon n_\epsilon - 1}, \theta^\epsilon_{\ell_\epsilon n_\epsilon})$ has a limit distribution (the mass "does not go away", see Appendix B, Theorem B.11) and that this limit distribution is actually the stationary measure $\mu_{\vartheta(u)}$. Finally, use (a3) to establish the claim (4.9).

The Martingale Representation: The proof is now completed by showing that (4.8) and (4.9) imply that the limit of the chosen weakly convergent subsequence satisfies:

$$\vartheta(t+s) - \vartheta(t) = \int_t^{t+s} G(\vartheta(u))\, du, \tag{4.10}$$

where ϑ is the limit process of the chosen subsequence. By (a5) G is bounded and continuous, thus this limit is the (unique) solution to the ODE (2.10), and because this limit is the same for any such convergent subsequence, the claim follows.

To show (4.10), define the process:

$$M^\epsilon(t) = \vartheta^\epsilon(t) - \vartheta^\epsilon(0) - \int_0^t \mathcal{G}(\vartheta^\epsilon(u))\, du.$$

We will now use the martingale characterization Theorem B.17 (Appendix B). Notice that for any $p \in \mathbb{N}, s_i \leq t; i \leq p$, and any bounded and continuous function h,

$$\mathbb{E}\left[h(\vartheta^\epsilon(s_i), i \leq p)\left[\vartheta^\epsilon(t+s) - \vartheta^\epsilon(t) - \int_t^{t+s} \mathcal{G}^\epsilon(u)\, du\right]\right] =$$
$$\mathbb{E}\left[h(\vartheta^\epsilon(s_i), i \leq p)\mathbb{E}\left[\vartheta^\epsilon(t+s) - \vartheta^\epsilon(t) - \int_t^{t+s} \mathcal{G}^\epsilon(u)\, du \,\middle|\, \mathfrak{F}^\epsilon_{\lfloor t/\epsilon \rfloor}\right]\right] = \mathbb{E}[\rho_1(\epsilon)]$$

for all ϵ such that $s_p \leq \epsilon\lfloor t/\epsilon \rfloor < t$. Along the chosen subsequence, $\mathbb{E}\|\rho_1(\epsilon)\| \to 0$, and $\vartheta^\epsilon \stackrel{\mathcal{L}}{\Longrightarrow} \vartheta$, so that

$$\mathbb{E}\left[h(\vartheta(s_i), i \leq p)\left[\vartheta(t+s) - \vartheta(t) - \int_t^{t+s} G(\vartheta(u))\, du\right]\right] = 0,$$

thus by Theorem B.17 (Appendix B), $M^\epsilon \stackrel{\mathcal{L}}{\Longrightarrow} M$ where:

$$M(t) = \vartheta(t+s) - \vartheta(t) - \int_t^{t+s} G(\vartheta(u))\, du$$

is a Martingale. Finally, by assumption (a3), $G(\cdot)$ is bounded and continuous, so that Lipschitz continuity of the limit process $\vartheta(\cdot)$ implies Lipschitz continuity of $M(\cdot)$. Using now Theorem B.16 (Appendix B), because $M(\cdot)$ is a Martingale which is Lipschitz continuous a.s. and which satisfies $M(0) \equiv 0$, then $M(t) \equiv 0$ a.s., which establishes that the limit process $\vartheta(\cdot)$ satisfies the ODE:

$$\frac{dx(t)}{dt} = G(x(t)).$$

QED

EXAMPLE 4.6. Consider a $GI/G/1$ queueing system: inter-arrival times are iid $\{A_n\}$, and there is a single server that uses an amount $S_n(\theta)$ of time to serve the n-th client. Clients queue up until they find the server free (first come, first serve discipline). Assume that $\{S_n(\theta)\}$ is an iid sequence with $\theta = \mathbb{E}[S_n(\theta)]$. Assume that $\mathbb{E}[A_1] < \infty, \mathsf{Var}(A_1) < \infty$, and that the service time distribution has a Lebesgue density $f_\theta(s)$ which is continuously differentiable in θ. It is well known that the queue process is *stable* only when

$$\theta \in \mathcal{S} \stackrel{\text{def}}{=} \{\theta \in \mathbb{R}^+ : \theta < \mathbb{E}[A_1]\}, \tag{4.11}$$

and we will also assume that $\mathsf{Var}(S_1(\theta)) < \infty$ for $\theta \in \mathcal{S}$. By stable, we mean that the queue process is ergodic and has a unique stationary distribution.

The consecutive waiting times of customers constitute a Markov process given by the so-called Lindley recursion. Let ξ_n be the waiting time of customer n, then:

$$\xi_n = (\xi_{n-1} + S_{n-1}(\theta) - A_n)_+, \tag{4.12}$$

where $(x)_+ = \max(0, x)$ is the positive part of the number x. It is well known that for stable parameters $\theta \in \mathcal{S}$ the Markov chain is Harris recurrent and ergodic, with a unique stationary measure. We will discuss this example in detail in part II and will provide a proof of (4.12) in Example 6.10.

The problem is to adjust the server's speed so that the stationary average wait per client is no greater than a given amount α, while minimising the cost of operation, which increases as θ decreases. Because it is costly to speed up the service, the solution to the problem is given when the constraint is active. For $\theta \in \mathcal{S}$, let

$$L(\theta) = \lim_{n\to\infty} \mathbb{E}_\theta[\xi_n] = \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} \mathbb{E}[\xi_n]$$

be the stationary waiting time, which is strictly increasing. Then we seek θ^* such that $L(\theta^*) = \alpha$. For this model, there is a unique value $\theta^* \in \mathcal{S}$ that solves the problem. Therefore we use a target tracking scheme to optimize the server. The target vector field is $G(\theta) = L(\theta) = \alpha$.

The difficulty is that while the queue is operating and adjusting θ, but the arrival distribution is not known, only the consecutive arrival times can be observed. Thus the server must adjust θ to solve the inverse problem $L(\theta^*) = \alpha$. Following our tracking examples, we propose to use

$$\theta_{n+1} = \theta_n - \epsilon(\xi_n - \alpha),$$

so that $Y_n = -(\xi_n - \alpha)$. Call $\mathfrak{F}_n$ the σ-algebra generated by $(\theta_0; \xi_1, \ldots, \xi_n)$. There are several important differences between this problem and the Robbins-Monro model:

- The feedback is biased, that is,

  $$\mathbb{E}[Y_n \,|\, \mathfrak{F}_{n-1}] = -\mathbb{E}[(\xi_{n-1} + S_n(\theta_n) - A_n)_+ \,|\, \xi_{n-1}, \theta_n] - \alpha \neq L(\theta_n) - \alpha = G(\theta_n),$$

  but the expectation when ξ_{n-1} has the stationary distribution is unbiased.

- Y_n does not have a bounded support.

- Consecutive observations are correlated: the process ξ_n and the control values θ_n are coupled in a Markov process.

We therefore use the endogenous noise model to analyse convergence of the algorithm. In order to apply Theorem 4.2, we need to verify the assumptions. Assumption (a1) is left as an exercise, and (a2) follows for $\Theta \subset \mathcal{S}$. Assumption (a3) follows from the ergodicity of $\{\xi(\theta)\}$ and the definition of $G(\theta)$, so we need now verification of (a4) and (a5). Tightness and uniform integrability are usually the most difficult conditions to verify in practice.

We illustrate with this example a useful technique that parallels Lyapunov stability ideas and uses convergence of super (sub)-martingales to ensure recurrence of the process $\{\xi_n\}$ when θ_n follows (4.5). Propose a stochastic Lyapunov function, given $\mathfrak{F}_n$:

$$V(\theta_n) = \frac{1}{2}(\xi_n - \alpha)^2,$$

which given $\mathfrak{F}_n$, is a function of θ_n and of the exogenous random variable A_{n+1}.

Consider the evolution of the process $\{\xi_n\}$ within a single busy cycle. Let $\tau = \min(n > 0 \colon \xi_n = 0)$ (that is, for all $n < \tau, S_n(\theta_n) > A_{n+1} - \xi_{n-1}$) and consider the stopped process $V(\theta_{\tau\wedge n})$. Use a Taylor approximation with remainder for $V(\cdot)$ to obtain, for $n \leq \tau$;

$$\begin{aligned}\mathbb{E}\left[V(\theta_{n+1}) - V(\theta_n) \,|\, \mathfrak{F}_n\right] &= \mathbb{E}\left[(\xi_n - \alpha)(\theta_{n+1} - \theta_n)S'(\theta_n) \,|\, \mathfrak{F}_n\right] + \mathcal{O}(\epsilon^2) \\ &= -\epsilon\,\mathbb{E}\left[(\xi_n - \alpha)^2 S'(\theta_n) \,|\, \mathfrak{F}_{n-1}\right] + \mathcal{O}(\epsilon^2)\end{aligned}$$

which for small enough ϵ is non-positive, proving that $V(\cdot)$ is a non negative local super-martingale within cycles. On the other hand, $V(\theta_\tau) = 0.5\alpha^2$, a constant. Thus, the state $\{\xi_n = 0\}$ must be a recurrent state. This shows that $\{\xi_n\}$ are tight, thus uniformly integrable, and also that $\theta_n \in \mathcal{S}$ infinitely often. This condition implies (a5), and using the assumption of finite variance for stable mean service times, it also implies (a4).

⁂

4.3 Exercises

EXERCISE 4.1. Consider the following queueing problem. Consecutive inter-arrival times are continuous random variables, iid, with finite moments and unit mean. The mean service time is $\theta > 0$. Let $f_\theta(\cdot)$ be the well defined density of the corresponding service time $S_n(\theta)$ and assume that $\mathsf{Var}(S_n(\theta)) < \infty$ for all finite values of θ. We wish to minimise the cost of operation $C(\theta) = 1/\theta^2$ while satisfying $L(\theta) \stackrel{\text{def}}{=} \mathbb{P}(W(\theta) > w) \leq \alpha$, where $W(\theta)$ is a random variable with the stationary waiting time distribution and $\alpha \in (0,1)$.

(a) Argue by using the results of Chapter 1, that at the optimal value θ^* the constraint must be active, that is, $L(\theta^*) = \alpha$.

(b) Let $\{\xi_n\}$ be the sequence of consecutive waiting times. Lindley's equation gives the dynamics of the waiting time process:
$$\xi_n = (\xi_{n-1} + S_n(\theta_n) - A_n)_+, \tag{4.13}$$
where A_n is the inter-arrival time between customers n and $n+1$, and $S_n(\theta)$ is the service time of customer n, given θ_n. Discuss the validity of the stochastic approximation procedure:
$$\theta_{n+1} = \theta_n - \epsilon_n Y_n,$$
for Y_n an estimator of $L(\theta) - \alpha$ obtained observing the process $\{\xi_n\}$. Specify your model (what will you use for Y_n) and verify the assumptions of Theorem 4.1, assuming that $\theta_n < 1$ infinitely often for this procedure.

(c) Instead of using one observation of the process ξ_n to produce Y_n, consider using an estimation interval with T observations. That is, use $\xi_{nT}, \ldots, \xi_{(n+1)T-1}$ to produce the estimate Y_n of the n-th interval. Write a program to simulate the queue under the service time control and experiment with various values of T. Plot the results and discuss.

EXERCISE 4.2. Show that $\|\rho_1(\epsilon)\| \stackrel{\mathcal{L}}{\Longrightarrow} 0$ in the proof of Theorem 4.2: first establish that $0 < t - \delta_\epsilon \lfloor t/\delta_\epsilon \rfloor < \delta_\epsilon$, then argue that the distribution of the random variable $X^\epsilon = \|\vartheta^\epsilon(t) - \vartheta^\epsilon(\delta_\epsilon \lfloor t/\delta_\epsilon \rfloor)\|$ can be made arbitrarily close to that of $X = \|\vartheta(t) - \vartheta(\delta_\epsilon \lfloor t/\delta_\epsilon \rfloor)\|$ and use the fact that the limit function $\vartheta(\cdot)$ is a.s. Lipschitz continuous. Justify why the result holds uniformly in t.

EXERCISE 4.3. Consider the model of the "automatic learning" exercise machine that adjusts the resistance to each user so as to enable them to reach their desired target heart rate (refer to Example 6, 7 and 8.

(a) Suppose that your algorithm considers N intervals of .3 sec in order to get a better estimate of the person's heart rate, so that only every N measurements , θ_n changes:

$$\theta^{\epsilon}_{n+1} = \theta^{\epsilon}_n + \epsilon \left(200 \sum_{i=n-N+1}^{n} \xi_i - N\,\alpha \right) \mathbf{1}_{\{(n/N)\in\mathbb{N}\}},$$

for $\alpha = 120$. Define the interpolation process $\vartheta^{\epsilon}(\cdot)$ as usual, that is:

$$\vartheta^{\epsilon}(t) = \theta_{m(t)}, \quad m(t) = \left\lfloor \frac{t}{\epsilon} \right\rfloor$$

Using Theorem 4.2, find the limiting ODE for this process. What is the dependency of the behaviour of the limiting ODE on N?

(b) Program the procedure for $N = 1, 10, 20$ and show the plots with $\epsilon = 0.001$. Discuss your results.

(c) Program the procedure $\theta_{n+1} = \theta_n + \epsilon_n Y_n$ with decreasing step sizes $\epsilon_n = 1/n$, plot and discuss the results. If you were to patent the algorithm for the fitness industry, what scheme would you choose and why?

EXERCISE 4.4. Show that for the problem in Example 4.6 assumption (a1) in Theorem 4.2 holds for any compact set in the stability region.