DM : Méthode Stochastiques

Convergence based on averaging sequences

16 avril 2021

0.1 Part I

1. Study the convergence of $(\beta_n)_{n\geq N}$.

In this question, we will show that if the sequence $(\beta_n)_{n\geq N}$ converges, it converges to 0. A proof of the convergence of $(\beta_n)_{n\geq N}$ to 0 is provided in the answer to Question 2.

Suppose there exists $\ell \succ 0$ as the limit of the sequence $(\beta_n)_{n\geq N}$.

For sufficiently large n, $\frac{\beta_n}{\beta_{n-1}} = 1 - c_n \longrightarrow 1$. Therefore, $c_n \longrightarrow 0$.

But since $\log(\beta_n) = \sum_{k=N}^{n} \log(1-c_k)$ and $\log(1-c_k) \sim -c_k$, and the series with general term $-c_n$ diverges to $-\infty$.

Hence, by the comparison theorem for series, $\log(\beta_n) \sim -\sum_{k=N}^{n} c_k \longrightarrow -\infty$.

Therefore, $(\beta_n)_{n\geq N}$ converges to 0. **Absurd**.

Thus, *if* $(\beta_n)_{n\geq N}$ *converges, it must converge to* 0.

2. Show that there exists $\ell \in \mathbb{R}$ such that $\beta_n \sim e^{(s_n+\ell)}$.

Since $\sum_{n\geq N} c_n^2 \prec \infty$, we deduce that c_n converges to 0.
Thus, since $\log(1-c_n) \sim -c_n + \frac{c_n^2}{2}$ for sufficiently large n and the series with general term c_n^2 converges, we deduce that :
$\log(\beta_n) + s_n \sim \sum_{k=N}^{n} \frac{c_n^2}{2} \longrightarrow \ell \in \mathbb{R}$, therefore $\beta_n \times e^{s_n} \sim e^{\ell}$.

As a result : $\beta_n \sim e^{(-s_n+\ell)}$.

And in particular, as mentioned in answer 1,

$$\beta_n \longrightarrow 0.$$

3. Give an explicit expression for β_n in the case $c_n = \frac{1}{n}$ and $N \geq 1$.

$\beta_n = \prod_{k=N}^{n}(1-c_k) = \prod_{k=N}^{n}(1-\frac{1}{k}) = \frac{N-1}{n}$.

In particular, if $N = 1$, we have $\beta_n = 0 forall \text{n} \geq 1$.

For any real sequence $(u_n)_{n\geq N}$ *with* $N \geq 1$*, we define its* $(c_n)_{n\geq N}$*-averaging by*

$$\bar{u}_n = c_n u_{n+1} \text{ for } n = N, \text{ and } \bar{u}_n = \beta_{n-1} \sum_{j=N-1}^{n-1} \left(\frac{c_j}{\beta_j}\right) u_{j+1} \text{ for } n \geq N+1.$$

4. Show that for any $n \geq N$, $\bar{u}_{n+1} = (1-c_n)\bar{u}_n + c_n u_{n+1}$. What is $(\bar{u}_n)_{n\geq N}$ in the case where $c_n = \frac{1}{n}$ and $N = 1$.

Remark : If any of the β_j is zero, we make the simplification : $\frac{\beta_{n-1}}{\beta_j} = \prod_{k=j+1}^{n-1}(1-c_k)$ if $n-1 > j$ and $\frac{\beta_{n-1}}{\beta_j} = 1$ if $n-1 = j$. Therefore, the above formula makes sense.

For $n \geq N$, we have :

$$\begin{aligned}
\bar{u}_{n+1} &= \beta_n \sum_{j=N-1}^{n} \left(\frac{c_j}{\beta_j}\right) u_{j+1} \\
&= (1-c_n)\beta_{n-1} \sum_{j=N-1}^{n-1} \left(\frac{c_j}{\beta_j}\right) u_{j+1} + \beta_n \left(c_n \beta_n\right) u_{n+1} \\
&= (1-c_n)\bar{u}_n + c_n u_{n+1}.
\end{aligned}$$

- *If $c_n = \frac{1}{n+1}$ and $N = 1$, and $n \geq N+1$:*

We have $\beta_n = \frac{N}{n}$ and $\frac{\beta_{n-1}}{\beta_j} = \frac{j}{n-1}$, so :

$$\begin{aligned}
\bar{u}_n &= \beta_{n-1} \sum_{j=N-1}^{n-1} \left(\frac{c_j}{\beta_j}\right) u_{j+1} \\
&= \sum_{j=0}^{n-1} \frac{1}{j} \frac{j}{n} u_{j+1} \\
&= \frac{1}{n} \sum_{j=0}^{n-1} u_{j+1} \\
&= \frac{1}{n} \sum_{j=1}^{n} u_j.
\end{aligned}$$

0.2 Part II

We now consider a sequence $(\theta_n)_{n\geq N}$ defined on an open convex set $O \subset \mathbb{R}^d$ satisfying the recursion for any $n \in \mathbb{N}$:

$$\theta_{n+1} = \theta_n + \gamma_{n+1} h(\theta_n) + c_n u_{n+1} + \tilde{\eta}_{n+1}.$$

$$\tilde{\theta}_{n+1} = \tilde{\theta}_n + \gamma_{n+1} h(\theta_n) + c_n \bar{u}_n + \tilde{\eta}_{n+1}.$$

1. For $n \geq 0$, we have $\tilde{\theta}_{n+1} - \theta_{n+1} = \tilde{\theta}_n - \theta_n + c_n(\bar{u}_n - u_{n+1})$.

Since $c_n(\bar{u}_n - u_{n+1}) = \bar{u}_n - \bar{u}_{n+1}$, we can write :

$$\tilde{\theta}_{n+1} - \theta_{n+1} = \tilde{\theta}_n - \theta_n + \bar{u}_n - \bar{u}_{n+1}.$$

Hence,

$$\tilde{\theta}_n - \theta_n = \tilde{\theta}_0 - \theta_0 + \bar{u}_0 - \bar{u}_n.$$

If we set $\tilde{\theta}_0 - \theta_0 + \bar{u}_0 = 0$,

since $\bar{u}_n \longrightarrow 0$, we deduce that $\tilde{\theta}_n - \theta_n \longrightarrow 0$.

2. Show that there exists a sequence of random variables $(\nu_n)_{n\in\mathbb{N}}$ such that almost surely for any $n \in \mathbb{N}$

$$\tilde{\theta}_{n+1} = \tilde{\theta}_n + \gamma_{n+1} h(\tilde{\theta}_n) + \eta_{n+1}, \quad \text{and} \quad \sum_n \frac{\|\eta_{n+1}\|^2}{\gamma_{n+1}} < +\infty.$$

We set $\eta_0 = \bar{\eta}_0$, $\eta_{n+1} = \gamma_{n+1}(h(\theta_n) - h(\tilde{\theta}_n)) + c_n\bar{u}_n + \tilde{\eta}_{n+1}$.

Since h is Lipschitz, there exists a constant $L_h > 0$ such that :

$$\|h(\theta_n) - h(\tilde{\theta}_n)\| \leq L_h \|\theta_n - \tilde{\theta}_n\|.$$

Given that $\theta_n - \tilde{\theta}_n = -\bar{u}_n$, we deduce that $\|h(\theta_n) - h(\tilde{\theta}_n)\| \leq L_h\|\bar{u}_n\|$. Hence :

$$\frac{\|\eta_{n+1}\|^2}{\gamma_{n+1}} \leq 6\left[\frac{\|\bar{\eta}_{n+1}\|^2}{\gamma_{n+1}} + \max(L_h, 1)\left(\gamma_{n+1} + \frac{c_n^2}{\gamma_{n+1}}\right)\|\bar{u}_n\|^2\right].$$

By condition (28), we deduce that :

$$\sum_n \frac{\|\eta_{n+1}\|^2}{\gamma_{n+1}} < +\infty.$$

3. Assume here that for any θ,

$$\|\nabla V(\theta)\|^2 \leq C(1+V(\theta)), \quad \text{and} \quad \langle h(\theta), \nabla V(\theta)\rangle \leq -\lambda\|\nabla V(\theta)\|^2.$$

By condition (30) and the assumptions made about the functions h and V, as discussed in the course, we have :

$(V(\tilde{\theta}_n))_n$ converges and $\sum_n \gamma_n \|\nabla V(\tilde{\theta}_n)\|^2 < +\infty$.

Since V and ∇V are Lipschitz, and $\theta_n - \tilde{\theta}_n \longrightarrow 0$, we deduce that :

- $(V(\theta_n))_n$ converges to the same limit as $(V(\tilde{\theta}_n))_n$. Indeed,

$$\|\nabla V(\theta_n) - \nabla V(\tilde{\theta}_n)\|^2 \leq L_V^2 \|\theta_n - \tilde{\theta}_n\|^2 \longrightarrow 0.$$

And
$$\gamma_n \|\nabla V(\theta_n)\|^2 \leq 2\left[\gamma_n \|\nabla V(\tilde{\theta}_n)\|^2 + L_{\nabla V}^2 \gamma_n \|\theta_n - \tilde{\theta}_n\|^2\right].$$

Since $\theta_n - \tilde{\theta}_n = -\bar{u}_n$ and $\sum_n \gamma_n \|\bar{u}_n\|^2 \leq +\infty$ by (28), we deduce that :

$$\sum_n \gamma_n \|\nabla V(\theta_n)\|^2 < +\infty.$$

0.3 Part III

$$\theta_{n+1} = \theta_n + \frac{\gamma}{n}\{h(\theta_n) + \epsilon_{n+1}\}.$$

1. We define $M_0 = 0$, $M_{n+1} = \sum_{i=0}^{n} \frac{\epsilon_{i+1}}{(i+1)^{1/2+\tau}}$.

We have that $(M_n)_n$ is a martingale with respect to the filtration $(\mathcal{F}_n)_n$.

Since $\mathbb{E}[\|M_{n+1}\|^2] = \mathbb{E}[\mathbb{E}[\|M_{n+1}\|^2 | \mathcal{F}_n]] \leq 2\sigma_e^2$, by the bounded convergence theorem for martingales in L^2, $(M_n)_n$ converges almost surely.

To show that $\lim_{n\to+\infty}(n+1)^{\frac{1}{2}+\tau} \sum_{i=0}^{n} \epsilon_{i+1} = 0$, it suffices to apply Kronecker's lemma.

Indeed, we have :

- $\frac{\epsilon_{n+1}}{(n+1)^{1/2+\tau}}$ is the general term of a convergent series (M_n).
- $((n+1)^{1/2+\tau})_n$ is an increasing sequence of positive real numbers diverging to infinity.

Therefore, by Kronecker's lemma :

$$\lim_{n\to+\infty}(n+1)^{\frac{1}{2}+\tau}\sum_{i=0}^{n}\epsilon_{i+1}=0.$$

2. The sequence $(\gamma_n)_n$ satisfies the condition (γ) : $\sum_n \gamma_n = +\infty$, and $\sum_n \gamma_n^2 < +\infty$.

Since $(\theta_n)_n$ satisfies the hypothesis $(u_n, \tilde{\eta}_n)$, we deduce from **Part II** that $(V(\theta_n))_n$ converges almost surely and $\sum_n \gamma_n \|\nabla V(\theta_n)\|^2 < +\infty$.

0.4 Part IV - Applications.

1. The Robbins-Monro algorithm converges in the case of dosage. According to the course, the dosage case is written as :

- 0 is strongly attractive for the mean field h.

R-M : $\theta_{n+1} = \theta_n + \gamma_n h(\theta_n) + \gamma_n Z_{n+1}$

Where $(Z_n)_n$ is an i.i.d. sequence with zero mean.
Since h is assumed to be Lipschitz and $(\gamma_n)_n$ satisfies the hypothesis (γ), if we take :

$V(\theta) = \|\theta - \theta^*\|^2$ as the Lyapunov function, we have, due to the strong attraction of 0 for h, that h satisfies (31) as well. Therefore, by **III-3**, we have :

$(V(\theta_n))_n$ converges almost surely, and $\sum_n \gamma_n \|\nabla V(\theta_n)\|^2 < +\infty$.

Let us show that the limit of $(V(\theta_n))_n$ is 0.

Assume by contradiction that $V(\theta_n) = \|\theta_n - \theta^*\|^2 \sim \ell \in \mathbb{R}_+^*$.

Since $\|\nabla V(\theta_n)\|^2 = 4V(\theta_n)$, we would then have $\gamma_n \|\nabla V(\theta_n)\|^2 \sim 4\ell\gamma_n$, which is a term of a divergent series, contradicting $\sum_n \gamma_n \|\nabla V(\theta_n)\|^2 < +\infty$.

Therefore, $V(\theta_n) \longrightarrow 0$, and consequently :

$$\theta_n \longrightarrow \theta^*.$$