PTML 10: 16/06/2022

TABLE DES MATIÈRES

1 Aitken's Δ^2 process 1
1.1 Presentation . . . 1
1.2 Equations . . . 2
1.3 Simulation . . . 2
2 Richardson extrapolation 3
2.1 Presentation . . . 3
2.2 Simulation in $\mathbb{R}^2$ . . . 3
2.3 Application to logistic regression . . . 3
2.3.1 Simulation . . . 3

INTRODUCTION

In this session we study some methods that accelerate the convergence of some iterative algorithms to their limit.

These methods combine iterates of the algorithms to produce a another sequence in parallel, that sometimes converges faster to the limit.

1 AITKEN'S Δ^2 PROCESS

1.1 Presentation

Aitken's process is one of the simplest of these methods. We consider a sequence $(x_k)_{k\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. The idea is to locally model the sequence as a first-order auto-regressive sequence, which means finding, for each $k \in \mathbb{R}$, two numbers a_k and b_k, such that :

$$\begin{cases} x_{k+1} = a_k x_k + b_k \\ x_{k+2} = a_k x_{k+1} + b_k \end{cases}$$

and then compute the limit of the sequence $(y_i^k)_{i\in\mathbb{N}}$ defined by

$$y_{i+1}^k = a_k y_i^k + b_k \tag{1}$$

we note l_k the limit of $(y_i^k)_{i\in\mathbb{N}}$. This defines another sequence $(l_k)_{k\in\mathbb{N}}$, that might converge faster to the limit of $(x_k)_{k\in\mathbb{N}}$, if this limit exists.

Note that in order to observe an acceleration, this model does not need to be exact, it can also be true only asymptotically.

1.2 Equations

Exercice 1 : Assuming that if iterate is different $x_k \neq x_{k+1}$, solve the linear system defining the locally auto-regressive process for a given k (which means find a_k and b_k)

Exercice 2 : For a given k, what would be a sufficient condition that ensures that the sequence $(y_i^k)_{i\in}$ has a limit ?

1.3 Simulation

We will apply Aitken's method to the Leibniz formula, a method to approximate π as the limit of the following sequence :

$$x_k = 4\sum_{j=0}^{k} \frac{(-1)^j}{2j+1} \tag{2}$$

this is one of the most famous applications of the method.

Exercice 3 : Is the condition from question 2 verified ?

Exercice 4 : Run a simulation that applies the method to the sequence $(x_k)_{k\in\mathbb{N}}$. You should observe something like figure 1, i.e. an acceleration of the convergence to the limit.

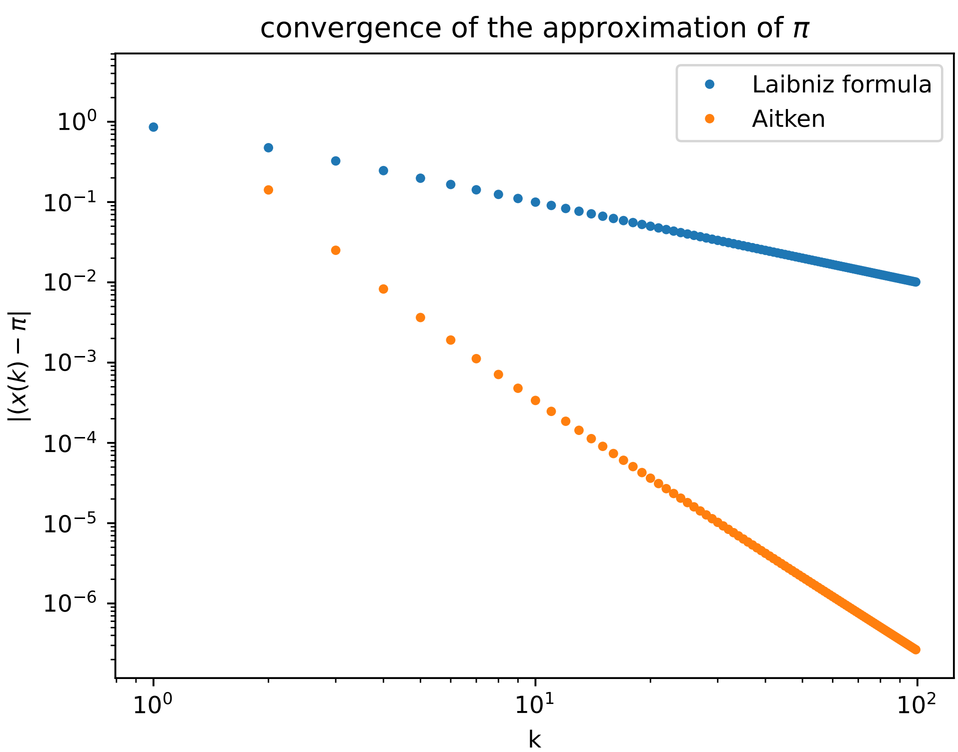

FIGURE 1 – Aitken's Δ^2 process accelerates the convergence to the π.

2 RICHARDSON EXTRAPOLATION

2.1 Presentation

We consider $(x_k)_{k\in\mathbb{N}}$ of points of $\mathbb{R}^d$ and we assume that

$$x_k = x_* + \frac{1}{k}\Delta + \mathcal{O}(\frac{1}{k^2}) \tag{3}$$

With $\Delta \in \mathbb{R}^d$. Hence, $(x_k)_{k\in\mathbb{N}} \to x_*$.

Exercice 5 : Show that for k even :

$$2x_k - x_{k/2} = x_* + \mathcal{O}(\frac{1}{k^2}) \tag{4}$$

Hence, the sequence defined by $y_k = 2x_k - x_{k/2}$ might converge faster to the limit x_*. This method is called Richardson extrapolation This method is called Richardson extrapolation

Exercice 6 : Can Richardson extrapolation have a strong negative impact ? what is the worst-case loss of performance ?

2.2 Simulation in $\mathbb{R}^2$

Exercice 7 : To get familiar with the method, apply it to your own sequence of points in $\mathbb{R}^2$.

2.3 Application to logistic regression

When performing gradient descent, for instance on logistic regression, it is possible to average the obtained iterates $x_j \in \mathbb{R}^d$.

$$z_k = \frac{1}{k}\sum_{j=0}^{k-1} x_j \tag{5}$$

This provides robustness to noise, however the initial conditions are not forgotten very fast. A workaround is **tail-averaging**, which means only taking the last half of the iterates. If k is even, we then define a new sequence t_k.

$$t_k = \frac{1}{k/2}\sum_{j=k/1}^{k-1} x_j \tag{6}$$

Exercice 8 : What is the link with Richardson extrapolation ?

2.3.1 *Simulation*

Exercice 9 : In some contexts, like logistic regression, as explained in [Bach, 2021] , it is possible to show that the sequence z verifies 3. Run a simulation that applies the method and look for a configuration where Richardson extrapolation accelerates convergence.

You have some useful functions for logistic regression in TP4.

RÉFÉRENCES

[Bach, 2021] Bach, F. (2021). On the Effectiveness of Richardson Extrapolation in Data Science. SIAM Journal on Mathematics of Data Science, 3(4) :1251–1277.