2023-2024

APPRENTISSAGE PAR RENFORCEMENT: PROJET

Alex Pierron

Apprentissage Par Renforcement

M2 Mathématiques et Intelligence Artificielle

Université Paris-Saclay / Institut de Mathématiques d'Orsay

22 Janvier 2024

Table des matières

Organisation du travail

Ce travail est organisé en deux documents. Le présent document se concentre sur la Partie I du travail demandé et s'attarde sur les développements et résultats mathématiques.
Un notebook nommé "PIERRON_RL" est disponible et concerne les parties II et III, parties se concentrant sur les simulations numériques et des extensions algorithmiques.

I Etude théorique d'un nouvel algorithmes de points fixes

Soit $d \geqslant 1$, $\mathrm{F} : \mathbb{R}^d \to \mathbb{R}^d$ une application admettant un point fixe $x_* \in \mathbb{R}^d$, $\eta > 0$, et $x_0 \in \mathbb{R}^d$. On définit l'algorithme Ada-FP comme suit :

$$x_{k+1} = x_k + \eta \frac{\mathrm{F}x_k - x_k}{\sqrt{\sum_{\ell=0}^{k} \| \mathrm{F}x_\ell - x_l\|_2^2}}, \quad k \geqslant 0, \tag{Ada-FP}$$

avec la convention $0/0 = 0$. On note pour tout $k \geqslant 0$,

$$u_k = \mathrm{F}x_k - x_k, \tag{1}$$

$$\eta_k = \frac{\eta}{\sqrt{\sum_{\ell=0}^{k} \| \mathrm{F}x_\ell - x_i\|_2^2}}, \tag{2}$$

$$\mathrm{D}_k = \max_{0 \leqslant l \leqslant k} \frac{1}{2} \|x_\ell - x_*\|_2^2 . \tag{3}$$

I.1 Question 1

On veut montrer que

$$\forall k \geqslant 0 \, \|x_{k+1} - x_*\|_2^2 \leqslant \|x_k - x_*\|_2^2 + 2\eta_k u_k^\top (x_k - x_*) + \eta_k^2 \|u_k\|_2^2 .$$

Preuve. Par définition de l'algorithme Ada-FP on a $\forall k \geqslant 0$:

$$\|x_{k+1} - x_*\|_2^2 = \left\| x_k + \eta \frac{\mathrm{F}x_k - x_k}{\sqrt{\sum_{\ell-0}^{k} \| \mathrm{F}x_\ell - x_i\|_2^2}} - x_* \right\|_2^2 = \|x_k + \eta_k u_k - x_*\|_2^2$$

$$\Leftrightarrow \|x_{k+1} - x_*\|_2^2 = \|x_k - x_* + \eta u_k\|_2^2 = \|x_k - x_*\|_2^2 + 2\eta_k u_k^\top (x_k - x_*) + \eta_k^2 \|u_k\|_2^2$$

en développant la norme avec $x = x_k - x_*$ et $y = \eta_k u_k$

D'où on conclut que l'on a bien

$$\forall k \geqslant 0, \|x_{k+1} - x_*\|_2^2 \leqslant \|x_k - x_*\|_2^2 + 2\eta_k u_k^\top (x_k - x_*) + \eta_k^2 \|u_k\|_2^2 . \tag{4}$$

□

I.2 Question 2

On veut déduire de la question 1 que pour tout $k \geqslant 0$,

$$\sum_{l=0}^{k} u_l^\top (x_* - x_l) \leqslant \frac{\mathrm{D}_k}{\eta_k} + \sum_{\ell=0}^{k} \frac{\eta_\ell \|u_\ell\|_2^2}{2} \tag{5}$$

Preuve. Par (4), on a $\forall k \geq 0$:

$$2\eta_k u_k^T(x_* - x_k) \leq \|x_k - x_*\|_2^2 - \|x_{k+1} - x_*\|_2^2 + \eta_k^2 \|u_k\|_2^2$$

$$\Leftrightarrow u_k^T(x_* - x_k) \leq \frac{\|x_k - x_*\|_2^2}{2\eta_k} - \frac{\|x_{k+1} - x_*\|_2^2}{2\eta_k} + \eta_k \frac{\|u_k\|_2^2}{2}$$

On prend ensuite la somme:

$$\sum_{l=0}^{k} u_l^T(x_* - x_l) \leq \sum_{l=0}^{k} \frac{\|x_l - x_*\|_2^2}{2\eta_l} - \frac{\|x_{l+1} - x_*\|_2^2}{2\eta_l} + \eta_l \frac{\|u_l\|_2^2}{2} \tag{6}$$

On veut maintenant majorer $\sum_{l=0}^{k} \frac{\|x_l - x_*\|_2^2}{2\eta_l} - \frac{\|x_{l+1} - x_*\|_2^2}{2\eta_l}$, on peut réécrire cette somme comme:

$$\sum_{l=0}^{k} \frac{\|x_l - x_*\|_2^2}{2\eta_l} - \frac{\|x_{l+1} - x_*\|_2^2}{2\eta_l} = \frac{1}{2}(\frac{1}{\eta_0} \|x_0 - x_*\|_2^2 - \frac{1}{\eta_k} \|x_{k+1} - x_*\|_2^2 + \sum_{l=1}^{k} \|x_l - x_*\|_2^2 (\frac{1}{\eta_l} - \frac{1}{\eta_{l-1}}))$$

$$\sum_{l=0}^{k} \frac{\|x_l - x_*\|_2^2}{2\eta_l} - \frac{\|x_{l+1} - x_*\|_2^2}{2\eta_l} \leq \frac{1}{\eta_0} \frac{\|x_0 - x_*\|_2^2}{2} + \sum_{l=1}^{k} \frac{\|x_l - x_*\|_2^2}{2} (\frac{1}{\eta_l} - \frac{1}{\eta_{l-1}}) \tag{7}$$

car $- \frac{1}{\eta_k} \|x_{k+1} - x_*\|_2^2$ est négatif.

En introduisant D_k comme défini dans (3), on a $\forall l$ tel que $0 \leq l \leq k$, $\frac{\|x_l - x_*\|_2^2}{2} \leq D_k$. On a donc dans (7):

$$\sum_{l=0}^{k} \frac{\|x_l - x_*\|_2^2}{2\eta_l} - \frac{\|x_{l+1} - x_*\|_2^2}{2\eta_l} \leq \frac{1}{\eta_0} D_k + D_k \sum_{l=1}^{k} (\frac{1}{\eta_l} - \frac{1}{\eta_{l-1}}) \tag{8}$$

Or $\sum_{l=1}^{k} \frac{1}{\eta_l} - \frac{1}{\eta_{l-1}}$ est une somme télescopique, donc on a $\sum_{l=1}^{k} \frac{1}{\eta_l} - \frac{1}{\eta_{l-1}} = \frac{1}{\eta_k} - \frac{1}{\eta_0}$.
En réécrivant (8), on a:

$$\sum_{l=0}^{k} \frac{\|x_l - x_*\|_2^2}{2\eta_l} - \frac{\|x_{l+1} - x_*\|_2^2}{2\eta_l} \leq D_k(\frac{1}{\eta_0} + \frac{1}{\eta_k} - \frac{1}{\eta_0}) = \frac{D_k}{\eta_k} \tag{9}$$

En combinant (6) et (9), on aboutit directement à l'inégalité souhaitée qui est:

$$\forall k \geqslant 0, \sum_{l=0}^{k} u_l^\top (x_* - x_l) \leqslant \frac{\mathrm{D}_k}{\eta_k} + \sum_{\ell=0}^{k} \frac{\eta_\ell \|u_\ell\|_2^2}{2}$$

□

I.3 Questions 3

I.3.1 Question 3.a

Soit $(a_k)_{k\geqslant 0}$ une suite positive. On veut montrer que pour tout $k \geqslant 0$,

$$\sum_{\ell=0}^{k} \frac{a_\ell}{\sqrt{\sum_{m=0}^{\ell} a_m}} \leqslant 2\sqrt{\sum_{\ell=0}^{k} a_\ell},$$

avec la convention $0/0 = 0$.

Preuve. Soit $(a_k)_{k\geqslant 0}$ une suite positive.

Pour tout $n \geq 0$ on introduit $S_n = \sum_{i=0}^{n} a_i$ et on fixe $S_{-1} = 0$. On a alors $\forall l \geq 0$:

$$\frac{a_\ell}{\sqrt{\sum_{m=0}^{\ell} a_m}} = \frac{S_l - S_{l-1}}{\sqrt{S_l}} \text{d'où}$$

$$\sum_{\ell=0}^{k} \frac{a_\ell}{\sqrt{\sum_{m=0}^{\ell} a_m}} = \sum_{\ell=0}^{k} \frac{S_l - S_{l-1}}{\sqrt{S_l}}$$

Or la suite des $(S_k)_{k\geq 0}$ est positive et croissante, on a donc par somme de rectangle à gauche pour la fonction $f : x \to \frac{1}{\sqrt{x}}$ qui est bien définie sur $\mathbb{R}^{+*}$:

$$\sum_{\ell=0}^{k} \frac{S_\ell - S_{\ell-1}}{\sqrt{S_\ell}} \leq \int_0^{S_k} \frac{1}{\sqrt{x}} dx$$

$$\Leftrightarrow \sum_{\ell=0}^{k} \frac{S_\ell - S_{\ell-1}}{\sqrt{S_\ell}} \leq 2\sqrt{S_k}$$

$$\Leftrightarrow \sum_{\ell=0}^{k} \frac{a_\ell}{\sqrt{\sum_{m=0}^{\ell} a_m}} \leqslant 2\sqrt{\sum_{\ell=0}^{k} a_\ell}.$$

□

I.3.2 Question 3.b

On veut déduire de ce qui précède que pour tout $k \geqslant 0$,

$$\sum_{\ell=0}^{k} u_\ell^\top (x_* - x_\ell) \leqslant \left(\eta + \frac{\mathrm{D}_k}{\eta}\right) \sqrt{\sum_{\ell=0}^{k} \|u_\ell\|_2^2}. \tag{10}$$

Preuve. Par la question 2,on a $\forall k \geqslant 0$,

$$\sum_{i=0}^{k} u_e^\top (x_* - x_i) \leqslant \frac{\mathrm{D}_k}{\eta_k} + \sum_{\ell=0}^{k} \frac{\eta_\ell \|u_\ell\|_2^2}{2}$$

Or ici on reconnaît que $a_k = \eta_k \|u_k\|_2^2, \forall k \geq 0$ est une suite positive, on peut donc utiliser le résultat de la question I.3.1 pour cette suite en utilisant que $\eta_k = \dfrac{\eta}{\sqrt{\sum_{\ell=0}^k \| \mathrm{F}x_\ell - x_i\|_2^2}} = \dfrac{\eta}{\sqrt{\sum_{\ell=0}^k \|u_\ell\|_2^2}}$, on a alors:

$$\sum_{i=0}^k u_k^\top (x_* - x_i) \leqslant \frac{\mathrm{D}_k}{\eta_k} + \sum_{\ell=0}^k \frac{\eta_\ell \|u_\ell\|_2^2}{2} \leqslant \frac{\mathrm{D}_k}{\eta_k} + \eta\sqrt{\sum_{\ell=0}^k \|u_\ell\|_2^2}$$

$$\Leftrightarrow \sum_{i=0}^k u_k^\top (x_* - x_i) \leqslant \frac{\mathrm{D}_k}{\eta_k} + \sum_{\ell=0}^k \frac{\eta_\ell \|u_\ell\|_2^2}{2} \leqslant \frac{\mathrm{D}_k}{\eta}\sqrt{\sum_{\ell=0}^k \|u_\ell\|_2^2} + \eta\sqrt{\sum_{\ell=0}^k \|u_\ell\|_2^2}$$

D'où l'on conclut que $\forall k \geq 0, \sum_{\ell=0}^k u_\ell^\top (x_* - x_\ell) \leqslant \left(\eta + \dfrac{\mathrm{D}_k}{\eta}\right)\sqrt{\sum_{\ell=0}^k \|u_\ell\|_2^2}$ □

I.4 Questions 4

On suppose que F est γ_{F}-lipschitzienne pour un certain $0 \leqslant \gamma_{\mathrm{F}} < 1$.

I.4.1 Question 4.a

On veut montrer que que pour tout $k \geqslant 0$,

$$\|\mathrm{F}x_k - x_k\|_2^2 \leqslant 2\left(\mathrm{F}x_k - x_k\right)^\top (x_* - x_k)\,. \tag{11}$$

Preuve. Comme F est γ_{F} lipschitzienne, on a

$$\|\mathrm{F}x_1 - \mathrm{F}x_2\|_2^2 \leqslant \gamma_{\mathrm{F}}^2 \|x_1 - x_2\|_2^2, \forall x_1, x_2 \in \mathbb{R}^p \text{ et on rappelle que } 0 \leqslant \gamma_{\mathrm{F}} < 1$$

.

On a en particulier $\forall k \geq 0$:

$$\|\mathrm{F}x_k - \mathrm{F}x_*\|_2^2 = \|\mathrm{F}x_k - x_k - \mathrm{F}x_* + x_k\|_2^2 = \|\mathrm{F}x_k - x_k\|_2^2 + 2(\mathrm{F}x_k - x_k)^T(x_k - \mathrm{F}x_*) + \|x_k - \mathrm{F}x_*\|_2^2$$

or $\|\mathrm{F}x_k - \mathrm{F}x_*\|_2^2 \leqslant \gamma_{\mathrm{F}}^2 \|x_k - x_*\|_2^2$, donc on a:

$$\|\mathrm{F}x_k - x_k\|_2^2 + 2(\mathrm{F}x_k - x_k)^T(x_k - \mathrm{F}x_*) + \|x_k - \mathrm{F}x_*\|_2^2 \leqslant \gamma_{\mathrm{F}}^2 \|x_k - x_*\|_2^2$$

Comme x_* est un point fixe, on a $\mathrm{F}x_* = x_*$, d'où:

$$\|\mathrm{F}x_k - x_k\|_2^2 + 2(\mathrm{F}x_k - x_k)^T(x_k - x_*) + \|x_k - x_*\|_2^2 \leqslant \gamma_{\mathrm{F}}^2 \|x_k - x_*\|_2^2$$

$$\Leftrightarrow \|\mathrm{F}x_k - x_k\|_2^2 + 2(\mathrm{F}x_k - x_k)^T(x_k - x_*) \leqslant (\gamma_{\mathrm{F}}^2 - 1)\|x_k - x_*\|_2^2 \leqslant 0, \text{ car } 0 \leqslant \gamma_{\mathrm{F}} < 1$$

$$\Leftrightarrow \|\mathrm{F}x_k - x_k\|_2^2 \leqslant 2(\mathrm{F}x_k - x_k)^T(x_* - x_k)$$

□

I.4.2 Question 4.b

On veut déduire de ce qui précède que pour tout $k \geqslant 0$,

$$\min_{0 \leq l \leq k} \| \mathrm{F}x_\ell - x_\ell\|_2 \leqslant \frac{2}{\sqrt{k}}\left(\eta + \frac{\mathrm{D}_k}{\eta}\right) \tag{12}$$

Preuve. Soit $k \geq 0$, on pose l^* l'indice tel que $\min_{0 \leq l \leq k} \| \mathrm{F}x_\ell - x_\ell\|_2 = \min_{0 \leq l \leq k} \|u_l\|_2 = \|u_{l^*}\|_2$.Par construction, $l^* \in \{1, ..., k\}$.

Tout d'abord, on remarque que:

$$k \|u_{l^*}\|_2^2 \leq \sum_{l=0}^{k} \|u_l\|_2^2$$

$$\|u_{l^*}\|_2^2 \leq \sum_{l=0}^{k} \frac{\|u_l\|_2^2}{k} \tag{13}$$

par définition de u_{l^*}. On a alors par (10) et (11) combinés:

$$\sum_{l=0}^{k} \frac{\|u_l\|_2^2}{k} \leq \frac{2}{k} \sum_{l=0}^{k} u_l^T (x_k - x_*) \leq \frac{2}{k}(\eta + \frac{D_k}{\eta})\sqrt{\sum_{l=0}^{k} \|u_l\|_2^2}$$

$$\Leftrightarrow \frac{\sqrt{\sum_{l=0}^{k} \|u_l\|_2^2}}{k} \leq \frac{2}{k}(\eta + \frac{D_k}{\eta}) \tag{14}$$

En utilisant la racine carré pour (13) et en multipliant (14) par $\sqrt{k}$, on a:

$$\|u_{l^*}\|_2 \leq \frac{\sqrt{\sum_{l=0}^{k} \|u_l\|_2^2}}{\sqrt{k}} \leq \frac{2}{\sqrt{k}}(\eta + \frac{D_k}{\eta})$$

$$\text{soit } \|u_{l^*}\|_2 \leq \frac{2}{\sqrt{k}}(\eta + \frac{D_k}{\eta})$$

Ce qui est bien l'inégalité que l'on souhaite démontrer. □

I.4.3 Question 4.c

On souhaite exhiber une ou des propriétés sur $\min_{0 \leqslant \ell \leqslant k} \|x_\ell - x_*\|_2$. On a: $\forall k \geq 0$

$$\|x_k - x_*\|_2 = \|x_k - \mathrm{F}x_k + \mathrm{F}x_k - \mathrm{F}x_* + \mathrm{F}x_* - x_*\|_2 \leq \|x_k - \mathrm{F}x_k + \mathrm{F}x_k - \mathrm{F}x_*\|_2 + \|\mathrm{F}x_* - x_*\|_2$$

or comme x_* est le point fixe de F, on a que $\|\mathrm{F}x_* - x_*\|_2 = 0$, on peut alors écrire:

$$\|x_k - x_*\|_2 \leq \|x_k - \mathrm{F}x_k + \mathrm{F}x_k - \mathrm{F}x_*\|_2 \leq \|x_k - \mathrm{F}x_k\|_2 + \|\mathrm{F}x_k - \mathrm{F}x_*\|_2 \tag{15}$$

Et comme F est η_F lipschitzienne, on a:

$$\|x_k - x_*\|_2 \leq \|x_k - \mathrm{F}x_k\|_2 + \|\mathrm{F}x_k - \mathrm{F}x_*\|_2 \leq \|x_k - \mathrm{F}x_k\|_2 + \eta_\mathrm{F} \|x_k - x_*\|_2$$

$$\Leftrightarrow \|x_k - x_*\|_2 \leq \frac{\|x_k - \mathrm{F}x_k\|_2}{1-\eta_\mathrm{F}}$$

Or par (12), on a que

$$\forall k \geqslant 0, \min_{0 \leq l \leq k} \| \mathrm{F}x_\ell - x_\ell\|_2 \leqslant \frac{2}{\sqrt{k}}\left(\eta + \frac{\mathrm{D}_k}{\eta}\right)$$

d'où l'on obtient

$$\min_{0 \leqslant \ell \leqslant k} \|x_\ell - x_*\|_2 \leqslant \frac{2}{\sqrt{k}}\left(\eta + \frac{\mathrm{D}_k}{\eta}\right)\frac{1}{1-\eta_\mathrm{F}} \tag{16}$$

On a donc finalement

Theorem 1. *Soit* F *une fonction* η_F*-lipschitzienne. Alors pour tout* $k \geq 0$*,on a:*

$$\min_{0 \leqslant \ell \leqslant k} \|x_\ell - x_*\|_2 \leq \frac{2}{\sqrt{k}}\left(\eta + \frac{\mathrm{D}_k}{\eta}\right)\frac{1}{1-\eta_\mathrm{F}} \tag{17}$$

où D_k *est défini par* $\mathrm{D}_k = \max_{0 \leqslant l \leqslant k} \frac{1}{2}\|x_\ell - x_*\|_2^2$

I.5 Question 5

On rappelle que dans un contexte de programmation dynamique (i.e. où la dynamique de transition p du MDP est disponible sous forme explicite), les itérations valeur (synchrones) pour l'évaluation d'une politique $\pi \in \Pi_0$ (resp. pour le contrôle) sont données par

$$v_{k+1} = \mathrm{B}_\pi v_k, \quad k \geqslant 1, \tag{$\mathrm{VI}_\pi^{(\mathrm{V})}$}$$

et

$$v_{k+1} = \mathrm{B}_* v_k, \quad k \geqslant 1, \tag{$\mathrm{VI}_*^{(\mathrm{V})}$}$$

En utilisant (Ada-FP), on veut définir des algorithmes aux itérations valeurs synchrones classiques ($\mathrm{VI}_\pi^{(\mathrm{V})}$) et ($\mathrm{VI}_*^{(\mathrm{V})}$).

Sur la base de (Ada-FP), on propose les deux algorithmes (Ada-$\mathrm{VI}_\pi^{(\mathrm{V})}$) et (Ada-$\mathrm{VI}_*^{(\mathrm{V})}$) définis par:

$$v_{k+1} = v_k + \eta\frac{\mathrm{B}_\pi v_k - v_k}{\sqrt{\sum_{\ell=0}^k \| \mathrm{B}_\pi x_\ell - x_l\|_2^2}}, \quad k \geqslant 1 \tag{Ada-$\mathrm{VI}_\pi^{(\mathrm{V})}$}$$

et

$$v_{k+1} = v_k + \eta\frac{\mathrm{B}_* v_k - v_k}{\sqrt{\sum_{\ell=0}^k \| \mathrm{B}_* x_\ell - x_l\|_2^2}}, \quad k \geqslant 1 \tag{Ada-$\mathrm{VI}_*^{(\mathrm{V})}$}$$

Avec η un hyperparamètre strictement positif. Ce choix est motivé par le fait que B_π et B_* sont des fonctions admettant un point fixe comme vue en cours.
l'algorithme (Ada-FP) peut donc s'appliquer à ces fonctions.