Minimisation par recuit simulée.

Énoncé du TP

Soit la fonction f :

$$\begin{aligned} f &: \mathbb{R} \to \mathbb{R} \\ x &\mapsto (x-1.0)(x-0.5)(x-0.3)(x+0.5) \end{aligned}$$

On souhaite minimiser f en utilisant l'algorithme du recuit simulé en fixant le nombre d'itération à $N = 10000$ et en utilisant le noyau de communication qui propose un nouvelle état de la façon suivante :

$$X \leftarrow X + \delta X \text{ avec } \delta X \sim U([-0.5, 0.5])$$

La variable aléatoire δX est définie par une loi uniforme continue sur le segment $[-0.5, 0.5]$, donc le cadre théorique impliquant un ensemble discret ne s'applique pas directement, cependant,la génération de nombre aléatoire sur un ordinateur est bien un processus discret, donc on peut appliquer la théorie vue en cours. Caractérisez l'action de la température sur la valeur de sur la valeur du minimum global.

Nom1 Prénom1
Nom2 Prénom2

Minimisation de fonction non convexe par recuit simulé : influence de la température.

1 Introduction/Sujet

(poser le contexte) Nous cherchons a minimiser la fonction suivante :

$$f : \mathbb{R} \to \mathbb{R}$$
$$x \mapsto (x - 1.0)(x - 0.5)(x - 0.3)(x + 0.5)$$

pour cela, nous utilisons un algorithme d'optimisation non convexe : le recuit simulé. Le principe de cette méthode consiste à proposer des solutions aléatoirement et d'accepter toutes les bonnes solutions (celles qui minimisent f) mais aussi une fraction contrôlée des mauvaises solutions pour ne pas être bloqué par des minima locaux. De façon plus formelle, voici comme se décompose la méthode :

Choisir un état initial $x^{(0)} \in \mathbb{R}$ et une température initiale
for *n=1 à 10000* **do**
  $\rho \leftarrow x^{(n-1)} + \delta x$ avec $\delta x \sim U([-0.5, 0.5])$
  if $f(\rho) \leq f(x^{(n-1)})$ **then**
    $x^{(n)} \leftarrow \rho$
  else
    $x^{(n)} \leftarrow \rho$ avec la probabilité $e^{-\frac{\Delta U}{T}}$
  end
  mettre à jour T
end

Algorithm 1: Recuit simulé pour trouver l'argument du minimum de f

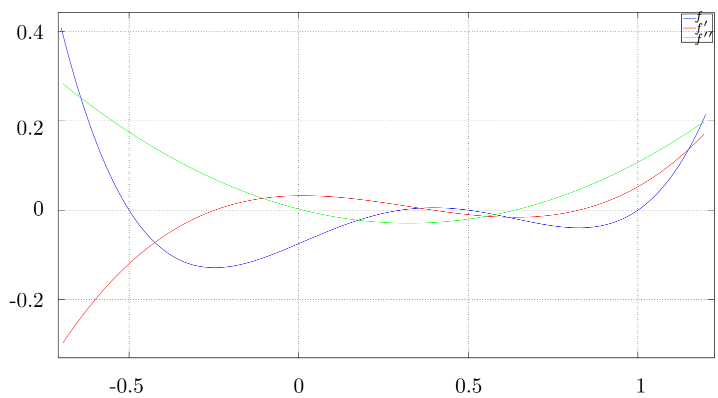

Figure 1: Représentation de la fonction (bleue) à minimiser et de ses dérivée (rouge) et dérivée seconde (vert). La dérivée est multipliée par 10 et la dérivée seconde par 1000.

2 Expériences/Questions

(si les questions sont clairement formulées, suivez le plan des questions avec la numérotation, sinon, il faut analyser le problème)

Dans un premier temps, nous observons que f a deux minima locaux (dont un est global) et un maximum car f' s'annule trois fois (polynôme de degrés 3) et f'' est positive pour deux de ces points, comme illustré sur la figure 1.

2.1 Effet de la température

Pour mettre en évidence les effets de la température sur la recherche (de l'argument) du minimum, nous testons cette algorithme avec plusieurs couples de températures initiales et finales qui correspondent aux niveaux d'acceptations qui sont listés dans le Tableau 1 On notera, avant

Table 1: Récapitulatif des arguments des expériences. Chacune est répétée 30 fois et les résultats sont les moyennes des résultats avec leurs écarts-types

$T_{\text{init}}/T_{\text{fin}}$	$8.0/10^{-5}$	$10^{21}/10^{-5}$	$8.10^{-3}/10^{-5}$	$8.0/0.1$	$8.0/10^{-21}$
tau d'acceptation initial	0.92 ± 0.02	1.0 ± 0.0	0.23 ± 0.03	0.915 ± 0.03	0.88 ± 0.03
tau d'acceptation final	0.008 ± 0.004	0.011 ± 0.005	0.006 ± 0.004	0.73 ± 0.03	0.0 ± 0.0
argument	-0.248 ± 0.002	-0.069 ± 0.41	0.29 ± 0.54	0.17 ± 0.4	-0.17 ± 0.27
minimum	-0.128907 ± 0.00001	-0.11 ± 0.03	-0.08 ± -0.04	-0.06 ± 0.05	-0.123 ± 0.022

toute interprétation, que les tau d'acceptation sont estimés sur 300 itérations de l'algorithme, ce qui a deux conséquences : la précision des taux est limitée à $\frac{1}{300}$ et la température a le temps d'évoluée ce qui fait aussi evoluer le tau d'acceptation ($e^{\frac{\Delta f}{T}}$).

Dans un premier temps, nous considéreront les températures "bien choisies", ce qui veut dire que les tau d'acceptation des mauvaises configurations est au départ dans l'intervalle $[0.6, 0.95]$ et à l'arrivée dans $[10^{-5}, 10^{-3}]$. Nous avons choisie le couple $(8.0/10^{-5})$, qui n'est pas complètement dans le critère énoncé précédemment, cependant l'algorithme converge et il tend vers les bonnes estimations. On observe que la valeur de l'estimation de l'argument minimisant f est précise comparativement aux autres configurations. On peut faire le même commentaire pour la valeur du minimum estimé.

Pour toutes les autres configurations, la valeur de l'argument du minimum est mal estimé, il suffit de comparer ces valeur avec un estimation graphique faire à partir de la figure 1 pour s'en convaincre. (dans ce cas, on peut utiliser le graphique comme argument car il fait office de contre-exemple). En effet, l'algorithme peut se faire piéger dans des bassin d'attraction locaux. Pour les couples $(8.10^{-3}/10^{-5})$, $(10^{21}/10^{-5})$ et $(8.0/10^{-21})$, la valeur de l'argument du minimum est précisément estimée pour chaque expérience car peu de mauvaise configurations sont acceptées à la fin des itérations. Cependant, la vitesse de décroissance de T n'est pas bien adaptée pour visiter touts les bassins d'attractions, donc, le recuit sort aléatoirement l'un des deux minima. Le couple $(8.0/0.1)$, quant à lui, ne permet pas une bonne estimation car le tau d'acceptation final est trop élevé, donc l'algorithme ne converge pas.

3 Conclusion/ouverture

On peut conclure qu'une bonne estimation des températures est une conditions nécessaire pour faire converger le recuit vers un minimum global. Cette conclusion n'est, cependant, valable que pour cet exemple qui ne présente aucun caractère général. Il serait intéressant d'étudier aussi l'influence du nombre d'itération et de la définition du noyau de communication.