La méthode de trichotomie

Tristan Vaccon

24 septembre 2012

La méthode de trichotomie permet de calculer le minimum d'une fonction convexe définie sur $[a, b] \subset \mathbb{R}$, avec une convergence géométrique, en $\frac{\sqrt{5}-1}{2}$.

Conseil important sur le déroulement de ce développement : pour chaque cas considéré, faire un dessin de la situation générique de la courbe et des points dans ce cas aide énormément à rendre compréhensible et digeste le développement. Ces dessins s'improvisent assez mal.

Démonstration. Tout d'abord, une fonction convexe sur $[a, b]$ est continue, et $[a, b]$ est un compact, donc une fonction convexe sur $[a, b]$ admet un minimum.

Le lemme suivant donne l'étape de "trichotomie" permettant l'algorithme qui va permettre de déterminer un minimum de f.

Lemme 1. *Soit $I = [a, b] \subset \mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction convexe. Soit c, d deux éléments de $]a, b[$ avec $c < d$. On va montrer que si l'on connaît $f(a,), f(b), f(c)$ et $f(d)$, on peut déterminer si le minimum de f sur $[a, b]$ est atteint sur $[a, d]$ ou sur $[c, b]$ (éventuellement sur les deux...).*

Démonstration. Soit c et d deux points de $[a, b]$, avec $a < c < d < b$. On pose $(x_0, \ldots, x_3) = (a, c, d, b)$, et μ_i la pente de la corde entre les points $(x_{i-1}, f(x_{i-1}))$ et $(x_i, f(x_i))$, i dans $\{1, 2, 3\}$. f étant convexe, si $x \in]a, b[$, la fonction $]x - a, b - x[\to \mathbb{R}$, $y \mapsto \frac{f(y)-f(x)}{y-x}$, est croissante, et on en déduit que $\mu_1 \leq \mu_2 \leq \mu_3$.

Cela nous donne de plus que si $\mu_{i+1} \leqslant 0$, alors $f(x_{i+1}) \leqslant f(x_i)$ et f est décroissante sur $[x_0, x_i]$. En effet, s'il existe $u_0 < v_0 < x_i$ tel que $f(u_0) < f(v_0)$, alors $\frac{f(u_0)-f(v_0)}{u_0-v_0} > 0$ et on en déduit par croissance la contradiction :

$$0 \geqslant \mu_{i+1} = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \geqslant \frac{f(x_{i+1}) - f(u_0)}{x_{i+1} - u_0} \geqslant \frac{f(v_0) - f(u_0)}{v_0 - u_0} > 0$$

De la même manière, si $\mu_i \geqslant 0$, alors f est croissante sur $[x_{i+1}, x_3]$.

On a alors quatre cas à traiter :

- Si tous les μ_i sont négatifs, f est décroissante sur $[x_0, x_2]$, et donc atteint son minimum dans $[x_2, x_3] \subset [x_1, x_3]$.

– Sinon, si $\mu_2 \leqslant 0$, alors f est décroissante sur $[x_0, x_1]$, et donc atteint son minimum dans $[x_1, x_3]$.
– Sinon, si $\mu_1 \leqslant 0$ mais $\mu_2 > 0$, alors f est croissante sur $[x_2, x_3]$, et donc atteint son minimum dans $[x_0, x_2]$.
– Sinon, $\mu_2 > 0$, alors f est croissante sur $[x_1, x_3]$, et donc atteint son minimum dans $[x_0, x_1] \subset [x_0, x_2]$.

Dans tout les cas, f atteint son minimum sur $[a, d] = [x_0, x_2]$ ou sur $[c, b] = [x_1, x_3]$, et on peut déterminer un intervalle parmi ces deux où il est atteint, connaissant $f(a), f(b), f(c)$ et $f(d)$. □

On peut maintenant considérer l'algorithme de trichotomie. On se donne $t \in \left[0, \frac{1}{2}\right]$, et on débute avec $(a, b) = (a_0, b_0)$. Par récurrence, si on sait que f atteint son minimum sur $[a_n, b_n]$, alors en posant $c_n = \frac{a_n+b_n}{2} - t\frac{b_n-a_n}{2}$ et $d_n = \frac{a_n+b_n}{2} + t\frac{b_n-a_n}{2}$, on peut déterminer avec le lemme un intervalle parmi $[a_n, d_n]$ et $[c_n, b_n]$ tel que f atteint son minimum sur cet intervalle, qu sera $[a_{n+1}, b_{n+1}]$.

La longueur des intervalles est majorée par $(t + \frac{1}{2})(b_n - a_n)$ et ainsi par récurrence,

$$|b_n - a_n| \leqslant (t + \frac{1}{2})^n |b_0 - a_0|.$$

Ainsi, a priori, au bout de $2n + 2$ évaluations de f, on connaît le minimum de f à $(t + \frac{1}{2})^n |b - a|$ près.

On peut chercher t tel que l'on puisse réutiliser les calculs de c_n et d_n lorsqu'on considère $[a_{n+1}, b_{n+1}]$, ce qui épargnerai une évaluation à chaque étape. Ceci peut s'écrire $c_n = d_{n+1}$ si le minimum de f est bien atteint sur $[a_n, d_n]$ et $d_n = c_{n+1}$ s'il est atteint sur $[c_n, b_n]$.

On peut se ramener à $[a, b] = [0, 1]$, $n = 1$ et alors $x_1 = \frac{1}{2} - t$ et $x_2 = \frac{1}{2} + t$. On peut considérer le premier cas où $[a_1, b_1] = \left[0, t + \frac{1}{2}\right]$.

Alors $c_2 = (t + \frac{1}{2})(-t + \frac{1}{2}) = \frac{1}{4} - t^2$ et $d_2 = (t + \frac{1}{2})^2$. On souhaite avoir $x_1 = c_1 = d_2$, soit

$$(t + \frac{1}{2})^2 = \frac{1}{2} - t.$$

Soit encore :

$$t^2 + 2t - \frac{1}{4}.$$

Avec la condition $t \in \left]0, \frac{1}{2}\right[$, ceci donne $t = -1 + \frac{\sqrt{5}}{2}$.

Par symétrie (par rapport à $\frac{1}{2}$), on vérifie que ce t convient bien dans le second cas, pour $d_1 = c_2$.

Au final, par cette méthode, on connaît le minimum de f à $(t + \frac{1}{2})^n |b - a|$ près en n itérations, et en ayant fait $n + 3$ évaluations de f. Ceci donne en particulier une convergence en $(t + \frac{1}{2})^n$, soit une convergence géométrique, en $\frac{\sqrt{5}-1}{2}$. □

Références

[1] Chambert-Loir, Antoine & Fermigier, Stéfane Analyse 2, exercices. Agrégation de mathématiques