Optimisation à finalité statistique

Salim Lardjane

Université de Bretagne-Sud

Cours 4 - Algorithmes de recherche directe : cas unidimensionnel

Algorithmes de recherche directe

Les algorithmes de recherche directe sont des méthodes d'optimisation ayant la caractéristique de ne requérir l'évaluation *explicite* d'aucune dérivée partielle de la fonction à optimiser.

Elles se basent uniquement sur les valeurs de la fonction calculées au cours des itérations.

Dans certains cas, ces valeurs sont utilisées pour obtenir des *approximations numériques* des dérivées de la fonction à optimiser.

Algorithmes de recherche directe

Dans d'autres cas, elles sont utilisées pour ajuster des polynômes de degré deux ou trois à la fonction au voisinage du minimum ou du maximum.

On abordera dans la suite d'abord le cas d'un seul paramètre (cas univarié ou unidimensionnel) et ensuite le cas de plusieurs paramètre (cas multivarié ou multidimensionnel).

Cas unidimensionnel

Les algorithmes de recherche directe pour la minimisation d'une fonction d'un seul paramètre peuvent être classés en deux catégories :

(i) ceux qui déterminent un *intervalle* contenant le minimum

(ii) ceux qui spécifient la position du minimum à l'aide d'une approximation ponctuelle de celui-ci.

Cas unidimensionnel

Afin d'appliquer les algorithmes de type (i), on suppose donné un intervalle initial *contenant le minimum* et que la fonction est *unimodale* (elle admet un seul extremum) sur celui-ci.

On cherche alors le minimum sur un intervalle $a < \theta < b$ en évaluant la fonction en certains points bien choisis de l'intervalle.

Les méthodes de type (ii), quant à elles, procèdent par évaluation de quelques valeurs de la fonction en des points particuliers afin d'approximer celle-ci localement par un polynôme simple.

La position du minimum est alors approximée par la position du minimum du polynôme, cette dernière étant relativement simple à calculer.

Algorithme de Fibonacci

On suppose que le minimum recherché est dans intervalle $[\theta^1, \theta^2]$ *connu* et que l'on doit choisir deux points dans cet intervalle de façon à avoir

$$\theta^1 < \theta^3 < \theta^4 < \theta^2$$

Comme la fonction est supposée unimodale :

- Si $f(\theta^3) > f(\theta^4)$, alors le minimum se trouve dans $[\theta^3, \theta^2]$
- Si $f(\theta_3) < f(\theta_4)$, alors le minimum se trouve dans $[\theta^1, \theta^4]$

Algorithme de Fibonacci

On peut itérer cette procédure et réduire la largeur de l'intervalle contenant le minimum en évaluant la fonction en d'autres points de l'intervalle dont on dispose à chaque étape.

On se pose alors la question suivante : Comment choisir les valeurs θ pour lesquelles on évalue la fonction ?

L'idée est d'utiliser, pour effectuer ce choix, les valeurs de la fonction déjà calculées.

Algorithme de Fibonacci

Si on spécifie que l'on ne s'autorise *qu'un nombre fixé* n *d'évaluations de la fonction*, alors on peut démontrer que la procédure la plus efficace est celle de Fibonacci.

L'efficacité signifie dans ce contexte *qu'aucune autre méthode ne garantit une diminution de la largeur de l'intervalle aussi importante pour* n *évaluations*.

Algorithme de Fibonacci

De façon plus précise, l'algorithme de Fibonacci utilise une suite d'entiers positifs, dite *de Fibonacci*, définie par les équations

$$\begin{aligned} F_0 &= F_1 = 1 \\ F_i &= F_{i-1} + F_{i-2} \end{aligned}$$

On obtient ainsi les nombres 1, 1, 2, 3, 5, 8, 13, ...

Si on suppose que $[\theta^1, \theta^2]$ est l'intervalle initial, alors la première étape de l'algorithme de Fibonacci consiste à calculer

$$I_1 = \frac{F_{n-1}}{F_n} \cdot I_0$$

où $I_0 = \theta^2 - \theta^1$.

Algorithme de Fibonacci

Les deux valeurs définissant l'intervalle suivant sont alors choisies comme

$$\begin{aligned} \theta^3 &= \theta_1 + I_1 \\ \theta^4 &= \theta_2 - I_1 \end{aligned}$$

Une partie de cet intervalle est alors éliminée en utilisant l'hypothèse d'unimodalité et le processus est répété dans le nouvel intervalle.

En examinant le ratio I_n/I_0, on peut *déterminer* la nombre d'étapes requises pour obtenir une précision donnée.

Algorithme de Fibonacci

Une alternative à l'algorithme de Fibonacci ne nécessitant pas la spécification initiale du nombre d'évaluations est l'algorithme dit de *la section dorée*. Hormis celui-ci, il existe encore d'autres méthodes.

Interpolation quadratique

Si on connaît la valeur d'une fonction $f(\theta)$ en trois points distincts θ_1, θ_2, θ_3, on peut approximer $f(\theta)$ par la fonction

$$h(\theta) = A\theta^2 + B\theta + C$$

où A, B, C sont déterminés par les équations

$$\left\{ \begin{array}{rcl} A\theta_1^2 + B\theta_1 + C & = & f(\theta_1) \\ A\theta_2^2 + B\theta_2 + C & = & f(\theta_2) \\ A\theta_3^2 + B\theta_3 + C & = & f(\theta_3) \end{array} \right.$$

Interpolation quadratique

La solution de cette équation peut être obtenue explicitement.

Par exemple,

$$A = -\frac{(\theta_2 - \theta_3)f(\theta_1) + (\theta_3 - \theta_1)f(\theta_2) + (\theta_1 - \theta_2)f(\theta_3)}{(\theta_1 - \theta_2)(\theta_2 - \theta_3)(\theta_3 - \theta_4)}$$

On peut obtenir de même des formules pour B et C.

En différentiant h par rapport à θ, on obtient

$$\frac{dh}{d\theta}(\theta) = 2A\theta + B$$

ce qui nous donne la position du minimum (en supposant $A > 0$)

$$\widetilde{\theta} = -\frac{B}{2A}$$

d'où une expression *explicite* de $\widetilde{\theta}$.

Interpolation quadratique

En pratique, pour utiliser cette approche, on se donne une *approximation initiale* θ_0 de la position du minimum et un *pas* δ fixé.

On part alors des points $\theta_1 = \theta_0$ et $\theta_2 = \theta_0 + \delta$ et on évalue $f(\theta_1)$ et $f(\theta_2)$.

Si $f(\theta_1) \leq f(\theta_2)$, on prend un troisième point $\theta_3 = \theta_0 - \delta$.

Si $f(\theta_1) > f(\theta_2)$, on pose $\theta_3 = \theta_0 + 2\delta$. On évalue alors $f(\theta_3)$ et on utilise les résultats précédents pour obtenir une nouvelle approximation du minimum.

On itère alors la procédure en prenant cette nouvelle approximation comme point de départ, en réduisant éventuellement le pas δ.

Interpolation quadratique

Le plus grand intérêt des méthodes d'approximation quadratique est leur utilité dans le cas multivarié.

En effet, il est souvent nécessaire dans ce cadre de déterminer le minimum de $f(\theta)$ le long d'une *droite* $\theta_0 + \lambda d$ où θ_0 est un point initial et où d spécifie une direction.

Les valeurs de $f(\theta_0 + \lambda d)$ le long de cette droite sont fonction d'*un seul paramètre* λ.

On peut donc utiliser l'algorithme d'interpolation quadratique.

Notons qu'une procédure analogue à l'interpolation quadratique mais qui est plus efficace consiste à approximer la fonction par un polynôme de degré trois. Cette méthode est fréquemment utilisée dans le contexte multivarié.