Chp. 4. Minimisation d'une fonction d'une variable

Avertissement! *Dans tout ce chapître, I désigne un intervalle de $I\!R$.*

4.1 Fonctions convexes d'une variable

Définition 9 *Une fonction φ, partout définie sur I, est dite convexe (resp. strictement convexe) sur I si :* $\forall \alpha \in]0,1[, \quad \forall s,t \in I, s \neq t, \quad \varphi(\alpha\, s + (1-\alpha)\, t) \leq \alpha\, \varphi(s) + (1-\alpha)\, \varphi(t) \qquad$ *(resp.$<$)*

Cette définition a une signification géomètrique simple : une fonction φ définie sur un intervalle I de $I\!R$ est convexe (resp. strictement convexe) sur I si la corde joignant deux points quelconques du graphe de φ dont les abscisses sont dans I est toujours située au dessus (resp. strictement au dessus) du graphe de φ.

Exemple 4.1 $\varphi = |\, t \,|$ *est convexe sur $I\!R$*

On dit que φ est *concave* (resp. *strictement concave*) sur I si : $-\varphi$ est convexe (resp. strictement convexe) sur I.

4.2 Propriétés des fonctions convexes définies sur un intervalle I de $\mathbf{R^n}$

- Si φ est convexe sur I, elle y est continue. Si elle convexe sur I et continue sur l'adhèrence J de I, elle est convexe sur J.
- Si φ est dérivable sur I, elle y est $\mathbb{C}$ [1].
- Si φ est $\mathbb{C}$ [1] sur I, elle y est convexe (resp. strictement convexe) si et seulement si φ' est croissante (resp. strictement croissante) sur I.
- La somme de deux fonctions convexes sur I est convexe sur I.
- Si φ est convexe (resp. strictement convexe) sur I et ψ est convexe et *croissante* (resp. *strictement croissante*) sur $J = \varphi(I)$, $\psi \circ \varphi$ est convexe (resp. strictement convexe) sur I.

Attention! *Le produit de deux fonctions convexes sur I n'est pas nécessairement convexe sur I.*

Contre-exemple 4.2 $\varphi = t$ *et :* $\psi = t^2$ *sont convexes sur* $I =]-1, 1\,[$, *mais pas leur produit* $\varphi.\psi = t^3$.

Théorème 4.1 *Si φ est convexe sur I, et admet un point critique $t^\star$ dans l'intérieur de I, $t^\star$ minimise φ sur I.*

Preuve : l'inégalité : $\varphi(\alpha\, s + (1-\alpha)\, t^\star) = \varphi(t^\star + \alpha\,(s - t^\star)) \leq \alpha\, \varphi(s) + (1-\alpha)\, \varphi(t^\star)$ est vérifiée pour tout s dans I, et tout α dans $]0, 1[$, d'où : $\alpha^{-1}\, [\varphi(t^\star + \alpha\,(s - t^\star)) - \varphi(t^\star)] \leq \varphi(s) - \varphi(t^\star)$. En faisant tendre α vers 0, on obtient : $\varphi(t^\star) \leq \varphi(s)$. Donc $t^\star$ minimise φ sur I.

□

4.3 Fonctions unimodales

Définition 10 *Une fonction φ, partout définie sur I, est dite unimodale sur I si :*

- *elle admet un unique minimum $t^\star$ dans l'intérieur de I.*
- *elle est strictement décroissante sur : $I \cap\,]-\infty, t^\star[$ et strictement croissante sur : $I \cap\,]\, t^\star, +\infty[$.*

Théorème 4.2 *Si φ est strictement convexe sur I et atteint son minimum sur I en un point $t^\star$ dans l'intérieur de I, elle est unimodale sur I.*

Preuve : Si : $s < t < t^\star$, il existe α dans l'intervalle $]\, 0,\, 1[$ tel que : $t = \alpha\, s + (1-\alpha)\, t^\star$, d'où :

$$\varphi(t) < \alpha\, \varphi(s) + (1-\alpha)\, \varphi(t^\star) \leq \varphi(s)$$

Donc φ est strictement décroissante sur $I \cap\,]-\infty, t^\star[$. On montre de même que φ est strictement croissante sur $I \cap\,]\, t^\star, +\infty[$.

□

Corollaire 4.3 *Supposons : φ strictement convexe sur $[0, +\infty[$, φ dérivable en 0, $\varphi'(0) < 0$, et : $\varphi(t) \to +\infty$ lorsque $t \to +\infty$. Alors φ est unimodale sur $[\,0, +\infty[$.*

On dit qu'une fonction d'une variable φ partout définie sur $I\!R$ est *coercive* sur $I\!R$ si : $\varphi(t) \to +\infty$ dès que : $|\, t\,| \to +\infty$.

Corollaire 4.4 *Si φ est strictement convexe et coercive sur $I\!R$, elle y est unimodale.*

4.4 Utilisation de la section dorée

On utilise l'inverse : $\alpha = \dfrac{\sqrt{5}-1}{2}$ ($\simeq 0.618$) du *nombre d'or* : $\dfrac{\sqrt{5}+1}{2}$:

```
GoldenSearch(φ, a0,b0, tolerance)
a ← a0
b ← b0
c ← α a + (1 − α) b
d ← a + b − c
Tant que : b − a > 2 ∗ tolerance
   Si : φ(c) < φ(d)     b = d ;   d = c ;   c = a + b − d ;   (elimination de b)
   Sinon :              a = c ;   c = d ;   d = a + b − c ;   (elimination de a)

Retourner : (a + b)/2
```

Théorème 4.5 *Si φ est unimodale sur l'intervalle : $[\,a_0, b_0\,]$, l'algorithme* `GoldenSearch` converge vers le minimum de φ sur cet intervalle.

Preuve : L'intervalle d'incertitude est réduit, à chaque étape, d'environ 40%.

4.5 Interpolation quadratique

On approche itérativement le graphe de la fonction à minimiser par une parabole. L'abscisse du sommet de la parabole fournit, à chaque étape, une nouvelle estimation du minimum cherché :

QuadSearch(φ, a$_0$,b$_0$, tolerance)
a $\leftarrow$ a$_0$
b $\leftarrow$ b$_0$
c $\leftarrow$ $(a+b)/2$
Repeter :

$$\text{d} \leftarrow \frac{1}{2}(a+b) + \frac{1}{2}\,\frac{(\varphi(a)-\varphi(b))\,(b-c)(c-a)}{(b-c)\,\varphi(a)+(c-a)\,\varphi(b)+(a-b)\,\varphi(c)}$$

Si : $\varphi(c) < \varphi(d)$
si : c $< d$ b $= d$ sinon : a $= d$
sinon :
si : c $< d$ a $= c$ c $= d$
sinon : b $= c$ c $= d$
Tant que : d $- c >$ tolerance

Retourner : d

Attention! *L'algorithme* `QuadSearch` n'est pas nécessairement convergent, même si φ est unimodale sur l'intervalle $[\,a_0, b_0\,]$.

Contre-exemple 4.3 $\varphi = t^3 - t$*,* $a_0 = 0$*, et :* $b_0 = 1$*.*

Mais il donne rapidement une approximation du minimum d'autant meilleure que le graphe de φ est proche d'une parabole.

4.6 Méthode de Newton unidimensionnelle

La méthode de Newton est un algorithme de descente unidimensionnel utilisant la direction de descente de Newton :

```
NewtonSearch(φ, t₀, tolerance)
t ← t₀
Tant que : φ''(t) φ'(t)² >  tolerance :
   t ← t − φ'(t)/ φ''(t)
Retourner : t
```

Lorsque le calcul de φ'' est possible et pas trop coûteux, la méthode de Newton est remarquablement efficace. On a, par exemple :

Théorème 4.6 *Supposons que φ est de classe $\mathcal{C}^2$ sur l'intérieur de son domaine de définition, et admet un minimum local $t^\star$. Dans chacun des cas suivants :*

- *φ'' est strictement positive et croissante sur $]\,t^\star, t_0\,[$.*
- *φ'' est strictement positive et décroissante sur $]\,t_0, t^\star[$.*
- *φ'' est strictement positive (sauf peut être en $t^\star$) et croissante sur $]\,t_0, +\infty\,[$, avec : $t_0 < t^\star$.*
- *φ'' est strictement positive (sauf peut être en $t^\star$) et décroissante sur $]\,-\infty, t_0\,[$, avec $t^\star < t_0$.*

l'algorithme `NewtonSearch` *converge. Si en outre φ est $\mathcal{C}^3$ et : $\varphi''(t^\star) \neq 0$, sa vitesse de convergence est quadratique.*

Preuve : Par construction :

$$(1) \qquad t_{k+1} - t^\star \;=\; t_k - t^\star - \varphi'(t_k)/\varphi''(t_k) \;=\; \frac{(t_k - t^\star)\,\varphi''(t_k) - \varphi'(t_k) + \varphi'(t^\star)}{\varphi''(t_k)}$$

En appliquant Taylor-Lagrange à la fonction φ sur l'intervalle : $[\,t_k, t^\star]$, on déduit :

$$(2) \qquad t_{k+1} - t^\star \;=\; (t_k - t^\star)\;\; \frac{\varphi''(t_k) - \varphi''(\theta_k)}{\varphi''(t_k)}$$

avec : $\theta_k \in]\,t_k, t^\star[$, que l'on récrit :

$$(3) \qquad t_{k+1} - t^\star \;=\; (t_k - t^\star)^2 \;\; \frac{t_k - \theta_k}{t_k - t^\star} \;\; \frac{1}{\varphi''(t_k)} \;\; \frac{\varphi''(t_k) - \varphi''(\theta_k)}{t_k - \theta_k}$$

En utilisant (3), on vérifie que, dans chacun des cas, la suite t_k construite par l'algorithme de Newton est bien définie et reste, pour $k \geq 1$, toute entière d'un même côté de $t^\star$. En utilisant (2), on montre alors qu'elle est monotone et bornée, donc convergente. Finalement (1) implique que sa limite ne peut être que $t^\star$.

Pour estimer la vitesse de convergence, on combine (1) et :

$$(4) \qquad \begin{aligned} |\,(t_k - t^\star)\,\varphi''(t_k) - \varphi'(t_k) + \varphi'(t^\star)\,| \;&=\; \left| \int_{t^\star}^{t_k} (\varphi''(t_k) - \varphi''(u)]\; du \right| \\ &\leq\; \int_{t^\star}^{t_k} L(t_k - u)\, du = \frac{L}{2}\,(t_k - t^\star)^2 \end{aligned}$$

où : L est une constante de Lipschitz de φ'' sur un intervalle borné quelconque contenant la suite t_k. On en déduit, pour k suffisamment grand :

$$(5) \qquad |\,t_{k+1} - t^\star\,| \;\leq \frac{L}{2\,c}\;|\,t_k - t^\star\,|^{\,2}$$

où c est une constante positive arbitraire minorant strictement $\varphi''(t_k)$.

□

Du théorème 4.6, on déduit directement :

Corollaire 4.7 *Supposons φ de classe $\mathcal{C}^2$ sur l'intervalle ouvert $]\,a,b\,[$ $(-\infty \leq a < b \leq +\infty)$, admettant un minimum local $t^\star$ dans $]\,a,b\,[$. Si φ" est strictement positive sur $]\,a,b\,[$, sauf peut être en $t^\star$, décroissante sur $]\,a,t^\star[$ et croissante sur $]t^\star\!,b\,[$, l'algorithme* `NewtonSearch` converge pour toute initialisation dans $]\,a,b\,[$. Si en outre φ est de classe $\mathcal{C}^3$ sur $]\,a,b\,[$ et : $\varphi(t^\star) \neq 0$, la convergence est quadratique.

Exemple 4.4 *Si $\varphi = t^4 + t^2$, l'algorithme converge, quelle que soit l'initialisation t_0 dans $\mathbb{R}$, vers 0, et la vitesse de convergence est quadratique. Si $\varphi = t^4$, la convergence reste assurée pour toute initialisation, mais la vitesse de convergence est seulement linéaire (explication : $\varphi''(0) = 0$).*

Corollaire 4.8 *Supposons φ de classe $\mathcal{C}^2$ sur l'intervalle ouvert $]\,a,+\infty[$ (resp. $]-\infty,a[$), et admettant un minimum local $t^\star$ dans $]\,a,+\infty[$ (resp. $]-\infty,a[$). Si φ" est strictement positive sur $]\,a,+\infty[$ (resp. $]-\infty,a[$), sauf peut être en $t^\star$, et croissante sur $]\,a,+\infty[$ (resp. décroissante sur $]-\infty,a[$), l'algorithme* `Newtonsearch` *converge, pour toute initialisation t_0 dans $]\,a,+\infty[$ (resp. $]-\infty,a[$). Si en outre φ est de classe $\mathcal{C}^3$ sur $]\,a,b\,[$ et : $\varphi(t^\star) \neq 0$, la convergence est quadratique.*

Exemple 4.5 *On peut calculer l'unique racine $t^\star$ de l'équation : $t - e^{-t} = 0$ en appliquant* `NewtonSearch` *à : $\varphi = (1/2)\,t^2 + e^{-t}$. L'algorithme converge, quelle que soit l'initialisation t_0 dans $\mathbb{R}$, et sa vitesse de convergence est quadratique.*

Attention! *Le résultat est faux si on inverse le sens de monotonie.*

Contre-exemple 4.6 *$\varphi = \dfrac{1}{t} + t$ est $\mathcal{C}^2$ sur $]0,+\infty[$, φ'' est strictement positive et décroissante sur $]0,+\infty[$. L'algorithme* `NewtonSearch` *converge vers l'unique minimum $t^\star = 1$ de φ sur $]0,+\infty[$ dès que : $0 < t_0 < t^\star$ (thm (4.6)), mais : $t_1 = -1$ si $t_0 = 2$!*

On peut cependant énoncer :

Théorème 4.9 *Si φ est $\mathcal{C}^3$, strictement convexe et unimodale sur $[a,b]$, et φ'' est uniformément minorée par une constante strictement positive c sur $[a,b]$,* `NewtonSearch` *converge pour toute initialisation t_0 dans $[a,b]$ vérifiant : $|\,t_0 - t^\star| < 2\,c/L$, où : L est une constante de Lipschitz de φ'' sur $[a,b]$.*

Preuve : On reprend les arguments de la démonstration précédente, en déduisant (5) de (1) et (4) pour établir directement que la suite t_k construite par l'algorithme est alors bien définie et converge vers $t^\star$.

□

En reprenant le contre-exemple (4.6) précédent, on voit que φ est $\mathcal{C}^3$ sur $]0,+\infty[$, et : $\varphi''(1) = 2$. L'algorithme `NewtonSearch` convergera donc encore pour toute initialisation $t_0 > 1$ *suffisamment proche* de 1. Mais $t_0 = 2$ est déja « trop grand ».

Sans sortir de l'intervalle sur lequel la fonction à minimiser est définie, comme dans (4.6), l'algorithme peut aussi se tromper de point critique lorsque l'initialisation est trop « loin » du minimum local :

Contre-exemple 4.7 *$\varphi = -\dfrac{1}{t} + \cos t$ est $\mathcal{C}^3$, strictement convexe et unimodale sur $]2,\pi[$. Elle atteint son minimum en $t^\star \simeq 3.033$, et sa dérivée seconde φ'' est minorée par : $c = 0.16$ et croissante* *sur* $]2, \pi[$. *Mais l'algorithme* `NewtonSearch`, *initialisé avec* t_0 *voisin de* 2, *converge vers un maximum local de* φ *voisin de* 6.3 *: le point* 2 *est ici « trop loin » de* $t^\star$.