

Cours 9 : Optimisation non linéaire monodimensionnelle sans contraintes

Christophe Gonzales

LIP6 – Université Paris 6, France

- 1 fonctions unimodales
- 2 méthode de la suite de Fibonacci
- 3 méthode dichotomique
- 4 méthode de Newton

Minimum local, global

f : fonction $[a, b] \subset \mathbb{R} \mapsto \mathbb{R}$

minimum global

- f passe par un **minimum (global ou absolu)** en \hat{x} si :

$$x \in [a, b] \implies f(x) \geq f(\hat{x})$$

- minimum **strict** : $x \neq \hat{x} \implies f(x) > f(\hat{x})$

Minimum local, global

f : fonction $[a, b] \subset \mathbb{R} \mapsto \mathbb{R}$

minimum global

- f passe par un **minimum (global ou absolu)** en \hat{x} si :

$$x \in [a, b] \implies f(x) \geq f(\hat{x})$$

- minimum **strict** : $x \neq \hat{x} \implies f(x) > f(\hat{x})$

minimum local

- f passe par un **minimum local** en \hat{x} si :

$$\exists \text{ un voisinage } V \text{ de } \hat{x} \text{ tel que } x \in V \implies f(x) \geq f(\hat{x})$$

- minimum local **strict** : $x \in V \setminus \{\hat{x}\} \implies f(x) > f(\hat{x})$



Dans \mathbb{R} , voisinage V de \hat{x} = ensemble t.q. :
 $V \supseteq$ intervalle ouvert $] \alpha, \beta [\supseteq \{ \hat{x} \}$


Définition d'une fonction unimodale

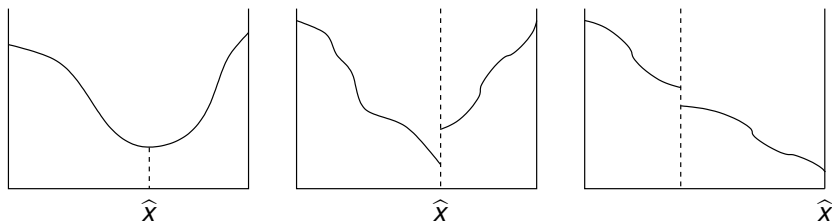
f est **unimodale** lorsqu'il existe $\hat{x} \in [a, b]$ tel que :

$$x^1 < x^2 \leq \hat{x} \implies f(x^1) > f(x^2)$$

$$\hat{x} \leq x^1 < x^2 \implies f(x^1) < f(x^2)$$

\implies une fonction unimodale passe par un minimum strict

 unimodale $\not\Rightarrow$ dérivable



Définition d'une fonction convexe

f est **convexe** si $\forall x, y \in [a, b], \forall \lambda \in [0, 1]$:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Définition d'une fonction strictement convexe

f est **strictement convexe** si $\forall x, y \in [a, b], \forall \lambda \in]0, 1[$:

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

Définition d'une fonction convexe

f est **convexe** si $\forall x, y \in [a, b], \forall \lambda \in [0, 1]$:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Définition d'une fonction strictement convexe

f est **strictement convexe** si $\forall x, y \in [a, b], \forall \lambda \in]0, 1[$:

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

strictement convexe \implies unimodale

Fonctions convexes et strictement convexes

Définition d'une fonction convexe

f est **convexe** si $\forall x, y \in [a, b], \forall \lambda \in [0, 1]$:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Définition d'une fonction strictement convexe

f est **strictement convexe** si $\forall x, y \in [a, b], \forall \lambda \in]0, 1[$:

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

strictement convexe \implies unimodale

convexe $\not\Rightarrow$ unimodale

Fonctions convexes et strictement convexes

Définition d'une fonction convexe

f est **convexe** si $\forall x, y \in [a, b], \forall \lambda \in [0, 1]$:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Définition d'une fonction strictement convexe

f est **strictement convexe** si $\forall x, y \in [a, b], \forall \lambda \in]0, 1[$:

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

strictement convexe \implies unimodale

convexe $\not\Rightarrow$ unimodale
unimodale $\not\Rightarrow$ convexe

Lemme 1

- fonction unimodale
- minimum local \implies minimum global

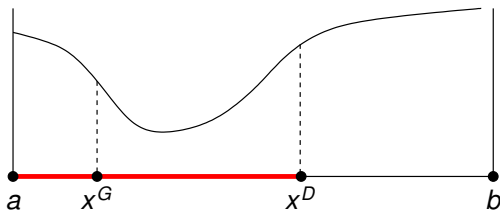
Unimodalité et recherche de minima

Lemme 1

- fonction unimodale
- minimum local \implies minimum global

Lemme 2

- $f : [a, b] \rightarrow \mathbb{R}$ unimodale
- x^G, x^D tels que $a < x^G < x^D < b$ alors :
 - $f(x^G) < f(x^D) \implies \hat{x} \in [a, x^D]$



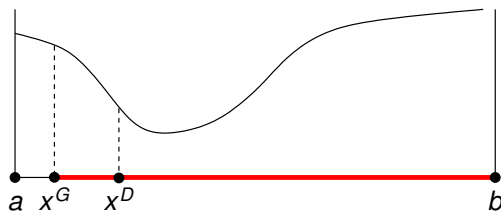
Unimodalité et recherche de minima

Lemme 1

- fonction unimodale
- minimum local \implies minimum global

Lemme 2

- $f : [a, b] \rightarrow \mathbb{R}$ unimodale
- x^G, x^D tels que $a < x^G < x^D < b$ alors :
 - $f(x^G) < f(x^D) \implies \hat{x} \in [a, x^D]$
 - $f(x^G) > f(x^D) \implies \hat{x} \in [x^G, b]$



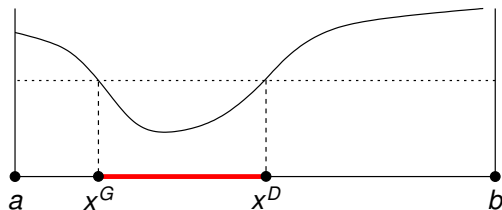
Unimodalité et recherche de minima

Lemme 1

- fonction unimodale
- minimum local \implies minimum global

Lemme 2

- $f : [a, b] \rightarrow \mathbb{R}$ unimodale
- x^G, x^D tels que $a < x^G < x^D < b$ alors :
 - $f(x^G) < f(x^D) \implies \hat{x} \in [a, x^D]$
 - $f(x^G) > f(x^D) \implies \hat{x} \in [x^G, b]$
 - $f(x^G) = f(x^D) \implies \hat{x} \in [x^G, x^D]$.



Rappel : suite de Fibonacci

- $F_0 = F_1 = 1$

- $F_{k+2} = F_{k+1} + F_k$

- $\frac{F_{k+1}}{F_k} \rightarrow \frac{1 + \sqrt{5}}{2}$

Rappel : suite de Fibonacci

- $F_0 = F_1 = 1$

- $F_{k+2} = F_{k+1} + F_k$

- $\frac{F_{k+1}}{F_k} \rightarrow \frac{1 + \sqrt{5}}{2}$

méthode de la suite de Fibonacci : utiliser Fibonacci pour trouver **rapidement** le minimum d'une fonction unimodale sur un intervalle $[a, b]$

méthode de la suite de Fibonacci

- on se donne N = nombre total de fois où l'on évaluera f en un point
- initialisation : $a_1 = a$, $b_1 = b$
- itérations :

$$x_k^G = a_k + \frac{F_{N-k}}{F_{N+2-k}}(b_k - a_k) \quad x_k^D = a_k + \frac{F_{N+1-k}}{F_{N+2-k}}(b_k - a_k)$$

- $f(x_k^G) < f(x_k^D) \implies a_{k+1} = a_k$ et $b_{k+1} = x_k^D$
- $f(x_k^G) > f(x_k^D) \implies a_{k+1} = x_k^G$ et $b_{k+1} = b_k$
- $f(x_k^G) = f(x_k^D) \implies a_{k+2} = x_k^G$ et $b_{k+2} = x_k^D$
- arrêt : longueur de l'intervalle après N évaluations de f :

$$\leq \prod_{k=1}^{N-1} \frac{F_{N+1-k}}{F_{N+2-k}}(b - a) = \frac{b - a}{F_{N+1}}$$

calcul de x_{k+1}^D et x_{k+1}^G : cas $f(x_k^G) < f(x_k^D)$

- $a_{k+1} = a_k$ et $b_{k+1} = x_k^D$
- $$\begin{aligned}x_{k+1}^D &= a_{k+1} + \frac{F_{N+1-(k+1)}}{F_{N+2-(k+1)}}(b_{k+1} - a_{k+1}) \\ &= a_k + \frac{F_{N-k}}{F_{N+1-k}}(x_k^D - a_k) \\ &= a_k + \frac{F_{N-k}}{F_{N+1-k}} \times \frac{F_{N+1-k}}{F_{N+2-k}}(b_k - a_k) \\ &= a_k + \frac{F_{N-k}}{F_{N+2-k}}(b_k - a_k) \\ &= x_k^G\end{aligned}$$

\implies seul $f(x_{k+1}^G)$ doit être évalué à l'itération $k + 1$

taille de l'intervalle $[a_{k+1}, b_{k+1}]$: cas $f(x_k^G) < f(x_k^D)$

- $a_{k+1} = a_k$ et $b_{k+1} = x_k^D$

- $b_{k+1} - a_{k+1} = x_k^D - a_k = \frac{F_{N+1-k}}{F_{N+2-k}}(b_k - a_k)$

$$\implies \frac{b_{k+1} - a_{k+1}}{b_k - a_k} = \frac{F_{N+1-k}}{F_{N+2-k}}$$

taille de l'intervalle $[a_{k+1}, b_{k+1}]$: cas $f(x_k^G) < f(x_k^D)$

- $a_{k+1} = a_k$ et $b_{k+1} = x_k^D$
- $b_{k+1} - a_{k+1} = x_k^D - a_k = \frac{F_{N+1-k}}{F_{N+2-k}}(b_k - a_k)$

$$\Rightarrow \frac{b_{k+1} - a_{k+1}}{b_k - a_k} = \frac{F_{N+1-k}}{F_{N+2-k}}$$

Par symétrie :

cas $f(x_k^G) > f(x_k^D)$

- seul $f(x_{k+1}^D)$ doit être évalué à l'itération $k + 1$

- $\frac{b_{k+1} - a_{k+1}}{b_k - a_k} = \frac{F_{N+1-k}}{F_{N+2-k}}$

taille de l'intervalle $[a_{k+2}, b_{k+2}]$: cas $f(x_k^G) = f(x_k^D)$


- $a_{k+2} = x_k^G$ et $b_{k+2} = x_k^D$

- $$\begin{aligned} \frac{b_{k+2} - a_{k+2}}{b_k - a_k} &= \frac{F_{N+1-k} - F_{N-k}}{F_{N+2-k}} = \frac{F_{N-1-k}}{F_{N+2-k}} \\ &= \frac{F_{N-1-k}}{F_{N-k}} \left[\frac{F_{N-k}}{F_{N+1-k}} \times \frac{F_{N+1-k}}{F_{N+2-k}} \right] \\ &\leq \frac{F_{N-k}}{F_{N+1-k}} \times \frac{F_{N+1-k}}{F_{N+2-k}} \end{aligned}$$

taille de l'intervalle $[a_{k+2}, b_{k+2}]$: cas $f(x_k^G) = f(x_k^D)$

● $a_{k+2} = x_k^G$ et $b_{k+2} = x_k^D$

●
$$\begin{aligned} \frac{b_{k+2} - a_{k+2}}{b_k - a_k} &= \frac{F_{N+1-k} - F_{N-k}}{F_{N+2-k}} = \frac{F_{N-1-k}}{F_{N+2-k}} \\ &= \frac{F_{N-1-k}}{F_{N-k}} \left[\frac{F_{N-k}}{F_{N+1-k}} \times \frac{F_{N+1-k}}{F_{N+2-k}} \right] \\ &\leq \frac{F_{N-k}}{F_{N+1-k}} \times \frac{F_{N+1-k}}{F_{N+2-k}} \end{aligned}$$

 évaluation de f en x_{k+2}^G et x_{k+2}^D mais intervalle plus réduit que pour 2 itérations des cas $f(x_k^G) \neq f(x_k^D)$

Résumé :

- $f(x_k^G) < f(x_k^D) \implies$ seul $f(x_{k+1}^G)$ doit être évalué

$$\frac{b_{k+1} - a_{k+1}}{b_k - a_k} = \frac{F_{N+1-k}}{F_{N+2-k}}$$

- $f(x_k^G) > f(x_k^D) \implies$ seul $f(x_{k+1}^D)$ doit être évalué

$$\frac{b_{k+1} - a_{k+1}}{b_k - a_k} = \frac{F_{N+1-k}}{F_{N+2-k}}$$

- $f(x_k^G) = f(x_k^D) \implies f(x_{k+2}^G)$ et $f(x_{k+2}^D)$ doivent être évalués

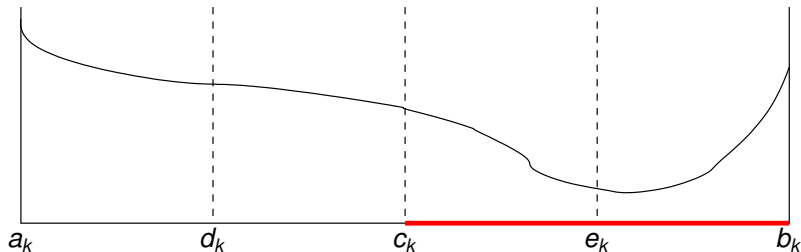
$$\frac{b_{k+2} - a_{k+2}}{b_k - a_k} \leq \frac{F_{N-k}}{F_{N+1-k}} \times \frac{F_{N+1-k}}{F_{N+2-k}}$$

$$\text{à l'arrêt : } b_N - a_N \leq \prod_{j=1}^{N-1} \frac{F_{N+1-j}}{F_{N+2-j}} (b - a) = \frac{b - a}{F_{N+1}}$$

La méthode dichotomique

L'algorithme par dichotomie

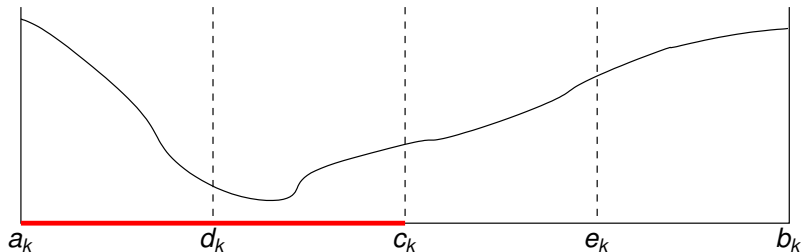
- à l'itération k : intervalle $[a_k, b_k]$
- $d_k = \frac{3a_k + b_k}{4}$ $c_k = \frac{a_k + b_k}{2}$ $e_k = \frac{a_k + 3b_k}{4}$
- $f(c_k) > f(e_k) \implies a_{k+1} = c_k$ et $b_{k+1} = b_k$



La méthode dichotomique

L'algorithme par dichotomie

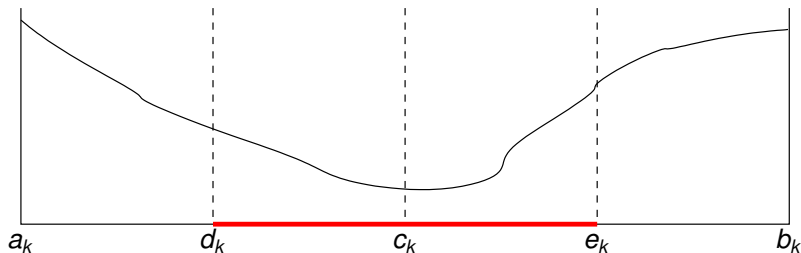
- à l'itération k : intervalle $[a_k, b_k]$
- $d_k = \frac{3a_k + b_k}{4}$ $c_k = \frac{a_k + b_k}{2}$ $e_k = \frac{a_k + 3b_k}{4}$
- $f(c_k) > f(e_k) \implies a_{k+1} = c_k$ et $b_{k+1} = b_k$
 $f(d_k) < f(c_k) \implies a_{k+1} = a_k$ et $b_{k+1} = c_k$



La méthode dichotomique

L'algorithme par dichotomie

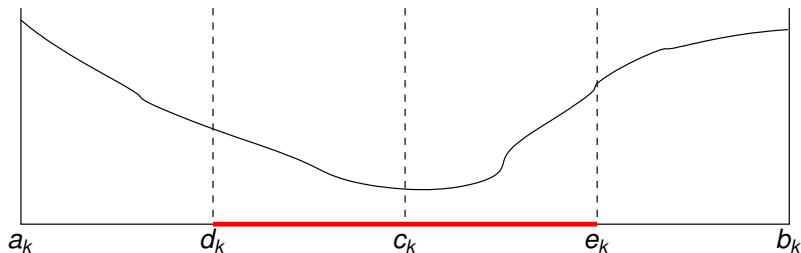
- à l'itération k : intervalle $[a_k, b_k]$
- $d_k = \frac{3a_k + b_k}{4}$ $c_k = \frac{a_k + b_k}{2}$ $e_k = \frac{a_k + 3b_k}{4}$
- $f(c_k) > f(e_k) \implies a_{k+1} = c_k$ et $b_{k+1} = b_k$
 $f(d_k) < f(c_k) \implies a_{k+1} = a_k$ et $b_{k+1} = c_k$
sinon $a_{k+1} = d_k$ et $b_{k+1} = e_k$



La méthode dichotomique

L'algorithme par dichotomie

- à l'itération k : intervalle $[a_k, b_k]$
- $d_k = \frac{3a_k + b_k}{4}$ $c_k = \frac{a_k + b_k}{2}$ $e_k = \frac{a_k + 3b_k}{4}$
- $f(c_k) > f(e_k) \implies a_{k+1} = c_k$ et $b_{k+1} = b_k$
 $f(d_k) < f(c_k) \implies a_{k+1} = a_k$ et $b_{k+1} = c_k$
sinon $a_{k+1} = d_k$ et $b_{k+1} = e_k$
- arrêt : quand $b_k - a_k \leq \epsilon$



Fonctions dérivables

$f : [a, b] \subset \mathbb{R} \mapsto \mathbb{R}$ dérivable. f' = dérivée de f

Point stationnaire

$\hat{x} \in]a, b[$: point stationnaire si $f'(\hat{x}) = 0$

$f : [a, b] \subset \mathbb{R} \mapsto \mathbb{R}$ dérivable. f' = dérivée de f

Point stationnaire

$\hat{x} \in]a, b[$: point stationnaire si $f'(\hat{x}) = 0$

Lemme 3

- f dérivable
- $\hat{x} \in]a, b[$
- \hat{x} minimum local de $f \implies \hat{x}$ = point stationnaire

$f : [a, b] \subset \mathbb{R} \mapsto \mathbb{R}$ dérivable. f' = dérivée de f

Point stationnaire

$\hat{x} \in]a, b[$: point stationnaire si $f'(\hat{x}) = 0$

Lemme 3

- f dérivable
- $\hat{x} \in]a, b[$
- \hat{x} minimum local de $f \implies \hat{x}$ = point stationnaire

Lemme 4

- f dérivable et convexe
- $\hat{x} \in]a, b[$
- \hat{x} minimum local de $f \iff \hat{x}$ = point stationnaire

Méthode dichotomique

- f unimodale, dérivable
- $a_1 = a, b_1 = b$
- $f'(a_k) < 0$ et $f'(b_k) > 0 \implies \hat{x} \in [a_k, b_k]$
- calcul de $f' \left(\frac{a_k + b_k}{2} \right)$
- $f' \left(\frac{a_k + b_k}{2} \right) > 0 \implies \hat{x} \in \left[a_k, \frac{a_k + b_k}{2} \right]$
- $f' \left(\frac{a_k + b_k}{2} \right) < 0 \implies \hat{x} \in \left[\frac{a_k + b_k}{2}, b_k \right]$
- $f' \left(\frac{a_k + b_k}{2} \right) = 0 \implies \hat{x} = \frac{a_k + b_k}{2}$

Méthode de Newton (1/3)

fonction de classe C^2

- $f : [a, b] \subset \mathbb{R} \mapsto \mathbb{R}$
- f : 2 fois dérivable
- f'' continue

Méthode de Newton (1/3)

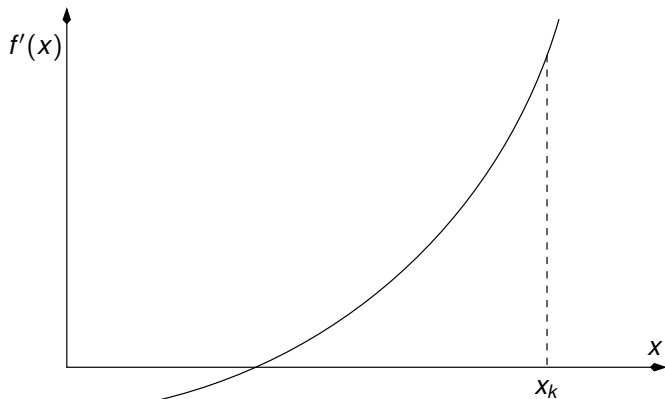
fonction de classe C^2

- $f : [a, b] \subset \mathbb{R} \mapsto \mathbb{R}$
- f : 2 fois dérivable
- f'' continue

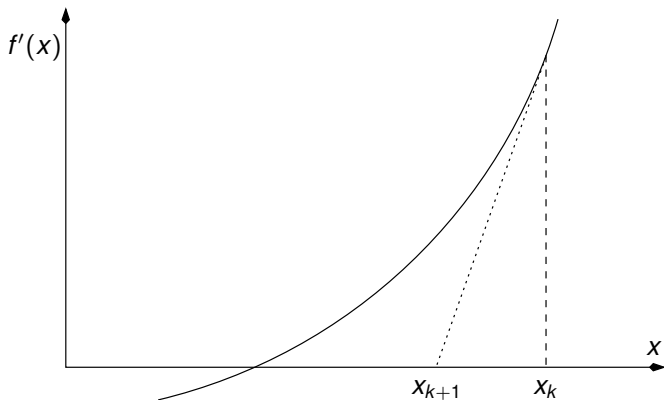
Méthode de Newton

- *principe* : engendrer une suite de points (x^k) tendant vers un point stationnaire
- point stationnaire : $f'(\hat{x}) = 0$
- itération k : f' est remplacée par sa linéarisée en x^k :
$$l(x) = f'(x^k) + [x - x^k]f''(x^k)$$
- x^{k+1} déterminé par $l(x^{k+1}) = 0$:
$$\implies x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}$$

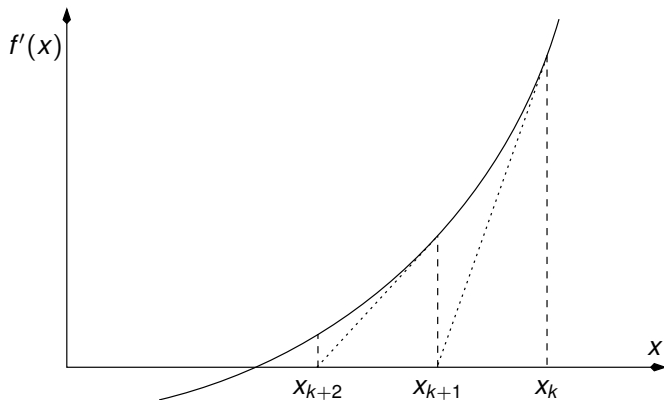
Méthode de Newton (2/3)



Méthode de Newton (2/3)



Méthode de Newton (2/3)



Conditions suffisantes de convergence de la méthode

Conditions suffisantes de convergence à partir d'un point de départ x^0 quelconque :

- f de classe C^3 et $f'(a) \times f'(b) < 0$
- $f''(x) > 0 \forall x$ (\implies stricte convexité)
- $0 \leq 1 - \frac{d}{dx} \left[\frac{f'(x)}{f''(x)} \right] \leq q < 1 \forall x$

\implies taux de convergence quadratique :

$$\exists \beta > 0 \text{ tel que } \|x^{k+1} - \hat{x}\| \leq \beta \|x^k - \hat{x}\|^2$$


Conditions suffisantes de convergence de la méthode

Conditions suffisantes de convergence à partir d'un point de départ x^0 quelconque :

- f de classe C^3 et $f'(a) \times f'(b) < 0$
- $f''(x) > 0 \forall x$ (\implies stricte convexité)
- $0 \leq 1 - \frac{d}{dx} \left[\frac{f'(x)}{f''(x)} \right] \leq q < 1 \forall x$

\implies taux de convergence quadratique :

$$\exists \beta > 0 \text{ tel que } \|x^{k+1} - \hat{x}\| \leq \beta \|x^k - \hat{x}\|^2$$

 conditions très restrictives

en pratique : on applique Newton même si ces conditions ne sont pas satisfaites