

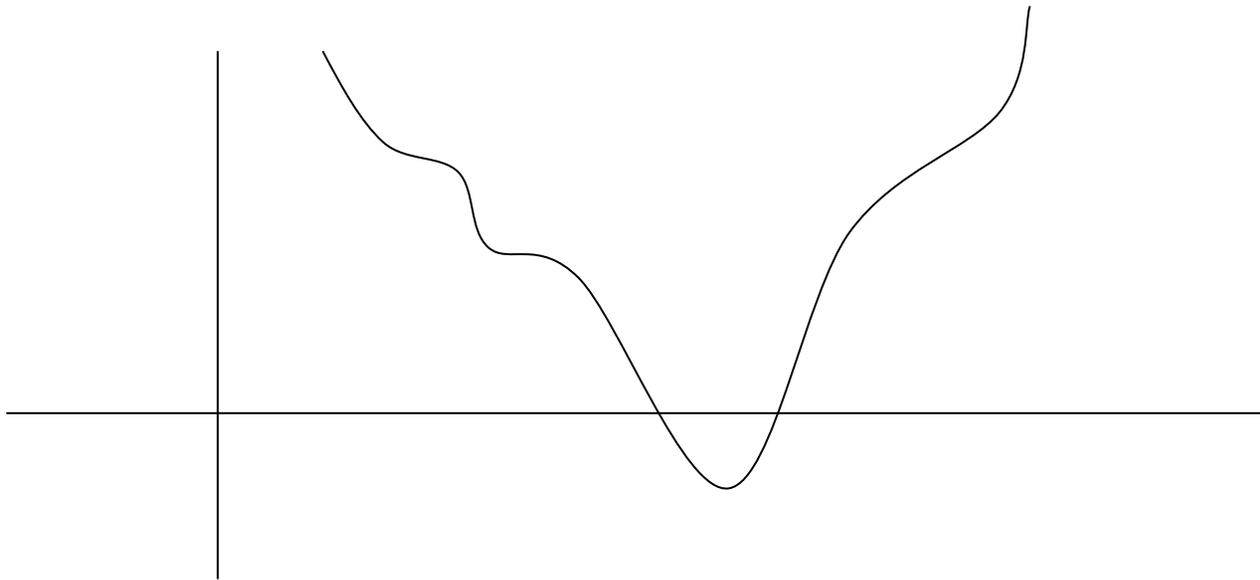
2. Optimisation sans contrainte

Fonctions à une seule variable

2.1. Méthodes n'utilisant que les valeurs des fonctions

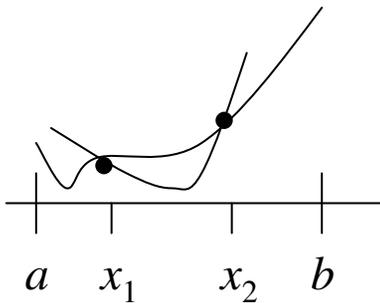
- **Méthode de Fibonacci**

Hypothèse: La fonction f est définie sur l'intervalle $[a, b]$ et est **unimodal**;
i.e., f ne possède qu'un seul minimum local dans $[a, b]$



- L'approche consiste
 - à choisir un certain nombre de points selon une stratégie basée sur les nombres de Fibonacci
 - à évaluer séquentiellement la valeur de la fonction à ces points
 - avec l'objectif de réduire la longueur de l'intervalle contenant le minimum local en se basant sur la propriété d'unimodalité de la fonction

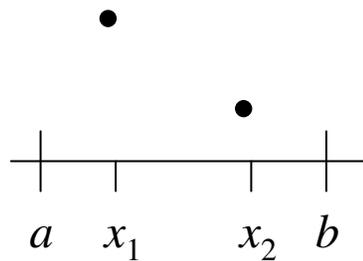
- Étant données les valeurs de la fonction en deux points de l'intervalle, l'unimodalité permet d'identifier une partie de l'intervalle où le minimum ne peut se retrouver



$$f(x_1) < f(x_2)$$

et

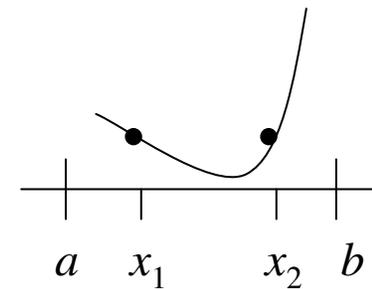
$$x^* \in [a, x_2]$$



$$f(x_1) > f(x_2)$$

et

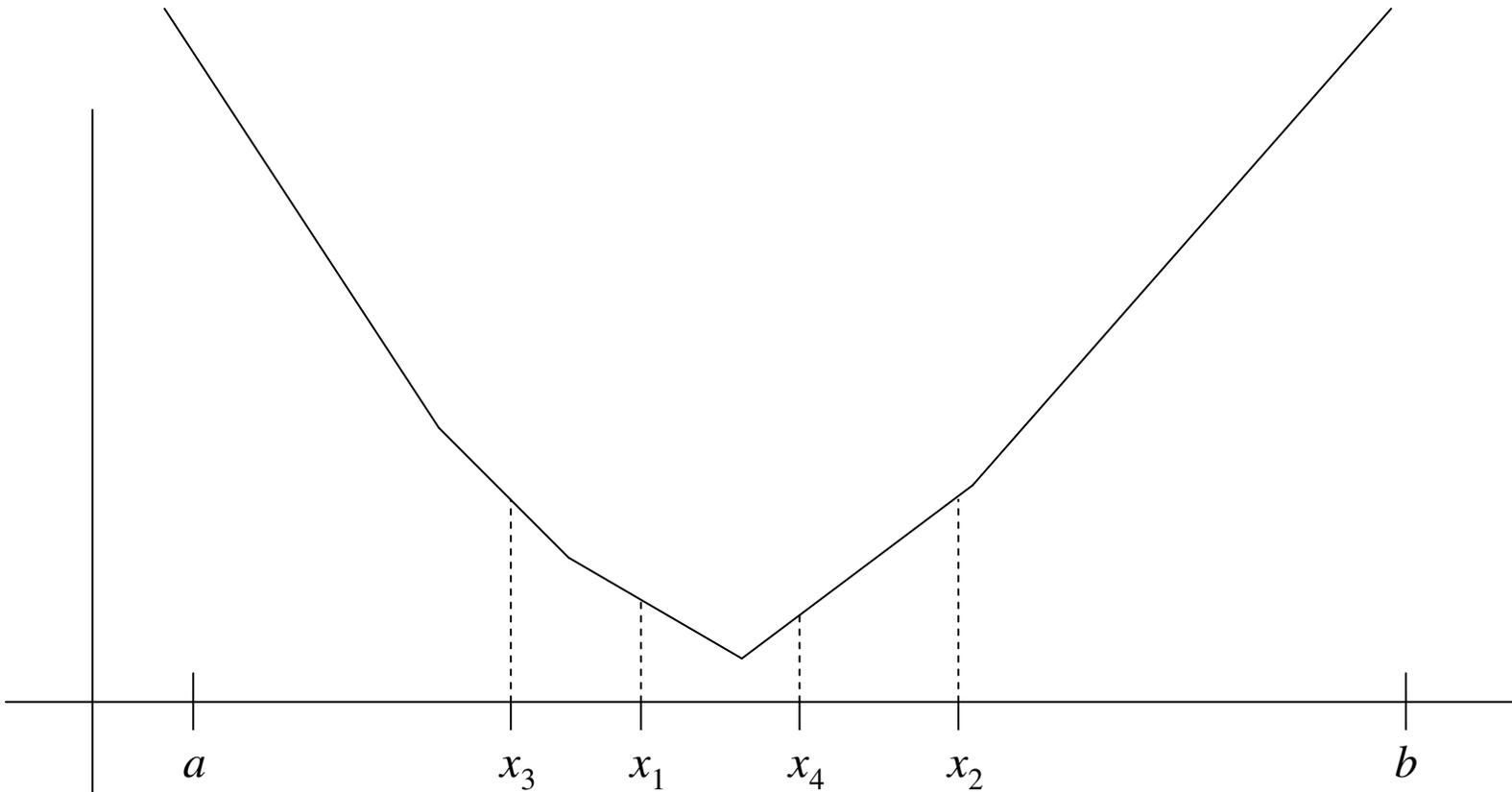
$$x^* \in [x_1, b]$$



$$f(x_1) = f(x_2)$$

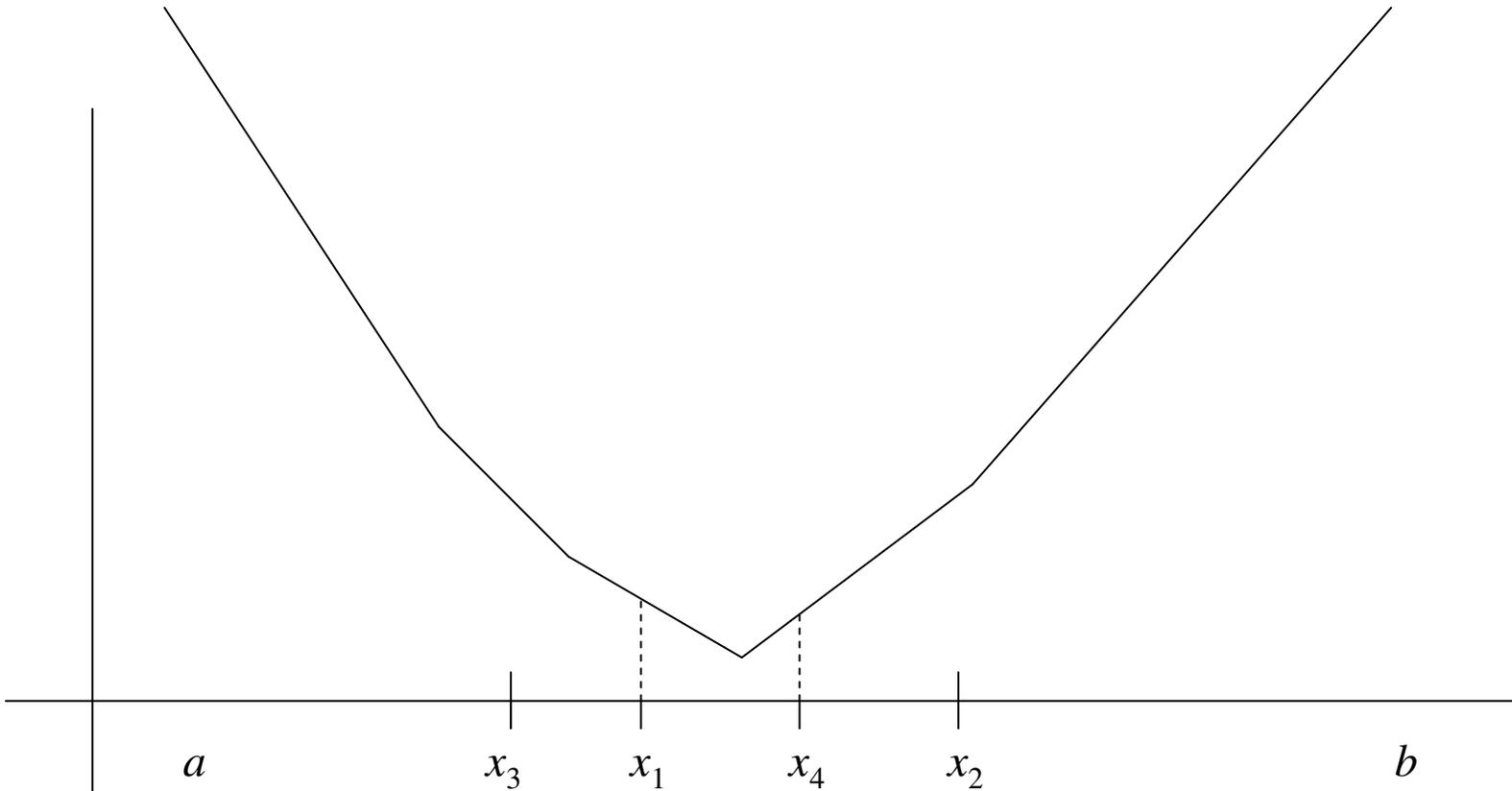
et

$$x^* \in [x_1, x_2]$$



Choisir deux points x_1 et x_2 symétrique et à la même distance de chaque extrémité de l'intervalle $[a, b]$

Choisir le prochain point symétriquement par rapport au point déjà dans l'intervalle résultant.



Choisir le prochain point symétriquement par rapport au point déjà dans l'intervalle résultant.

Stratégie optimale de sélection des points d'évaluation

- Notation:

$d_1 = b - a$, la longueur de l'intervalle initial

d_k = longueur de l'intervalle après avoir utilisé k points d'évaluation

$\{F_k\}$ la suite des nombres de Fibonacci définie comme suit:

$$F_0 = F_1 = 1$$

$$F_n = F_{n-1} + F_{n-2} \quad n = 2, 3, \dots$$

$$\{1, 1, 2, 3, 5, 8, 13, 21, \dots\}$$

Stratégie optimale de sélection des points d'évaluation

Supposons que nous décidons au départ d'utiliser N points d'évaluation.

Procédure se résume comme suit:

- i. Les deux premiers points sont choisis symétriques à une distance $\left(\frac{F_{N-1}}{F_N}\right)d_1$

de chacune des extrémités de l'intervalle $[a, b]$. Une partie de l'intervalle est éliminée en se basant sur l'unimodalité de la fonction.

Il en résulte un intervalle de longueur $d_2 = \left(\frac{F_{N-1}}{F_N}\right)d_1$.

- ii. Le troisième point est choisi symétriquement par rapport au point déjà dans l'intervalle résultant. Ceci engendre un intervalle de longueur

$$d_3 = \left(\frac{F_{N-2}}{F_N}\right)d_1$$

Stratégie optimale de sélection des points d'évaluation

- i. Les deux premiers points sont choisis symétrique à une distance

de chacune des extrémités de l'intervalle $[a, b]$. Une partie de l'intervalle est éliminée en se basant sur l'unimodalité de la fonction.

Il en résulte un intervalle de longueur $d_2 = \left(\frac{F_{N-1}}{F_N}\right)d_1$.

- ii. Le troisième point est choisi symétriquement par rapport au point déjà dans l'intervalle résultant. Ceci engendre un intervalle de longueur

$$d_3 = \left(\frac{F_{N-2}}{F_N}\right)d_1$$

- iii. En général le point suivant est choisi symétriquement par rapport au point déjà dans l'intervalle résultant.

iii. En général le point suivant est choisi symétriquement par rapport au point déjà dans l'intervalle résultant.

Note: Selon iii. , le dernier point N devrait être placé au centre de l'intervalle superposé à celui s'y trouvant déjà. En effet, puisqu'en utilisant cette stratégie de sélection des points d'évaluation, nous avons que

$$d_k = \frac{F_{N-k+1}}{F_N} d_1 \quad k = 2, 3, \dots, N$$

Note: Selon iii. , le dernier point N devrait être placé au centre de l'intervalle superposé à celui s'y trouvant déjà. En effet, puisqu'en utilisant cette stratégie de sélection des points d'évaluation, nous avons que

$$d_k = \frac{F_{N-k+1}}{F_N} d_1 \quad k = 2, 3, \dots, N$$

il s'ensuit que

$$\left. \begin{aligned} d_{N-1} &= \frac{F_{N-(N-1)+1}}{F_N} d_1 = \frac{F_2}{F_N} d_1 = \frac{2}{F_N} d_1 \\ d_N &= \frac{F_{N-N+1}}{F_N} d_1 = \frac{F_1}{F_N} d_1 = \frac{1}{F_N} d_1 \end{aligned} \right\} \Rightarrow d_N = \frac{1}{2} d_{N-1}.$$

Pour remédier à cette situation, le dernier point N est plutôt placé à une distance ε (à gauche ou à droite) de celui s'y trouvant déjà.

- En utilisant cette stratégie de sélection des points d'évaluation,

$$d_k = \frac{F_{N-k+1}}{F_N} d_1 \quad k = 2, 3, \dots, N$$

et il est possible de démontrer que

$$d_N = \frac{F_{N-N+1}}{F_N} d_1 = \frac{F_1}{F_N} d_1 = \frac{d_1}{F_N}$$

est le plus petit intervalle qu'il est possible d'obtenir en utilisant N points d'évaluations

- En utilisant cette stratégie de sélection des points d'évaluation,

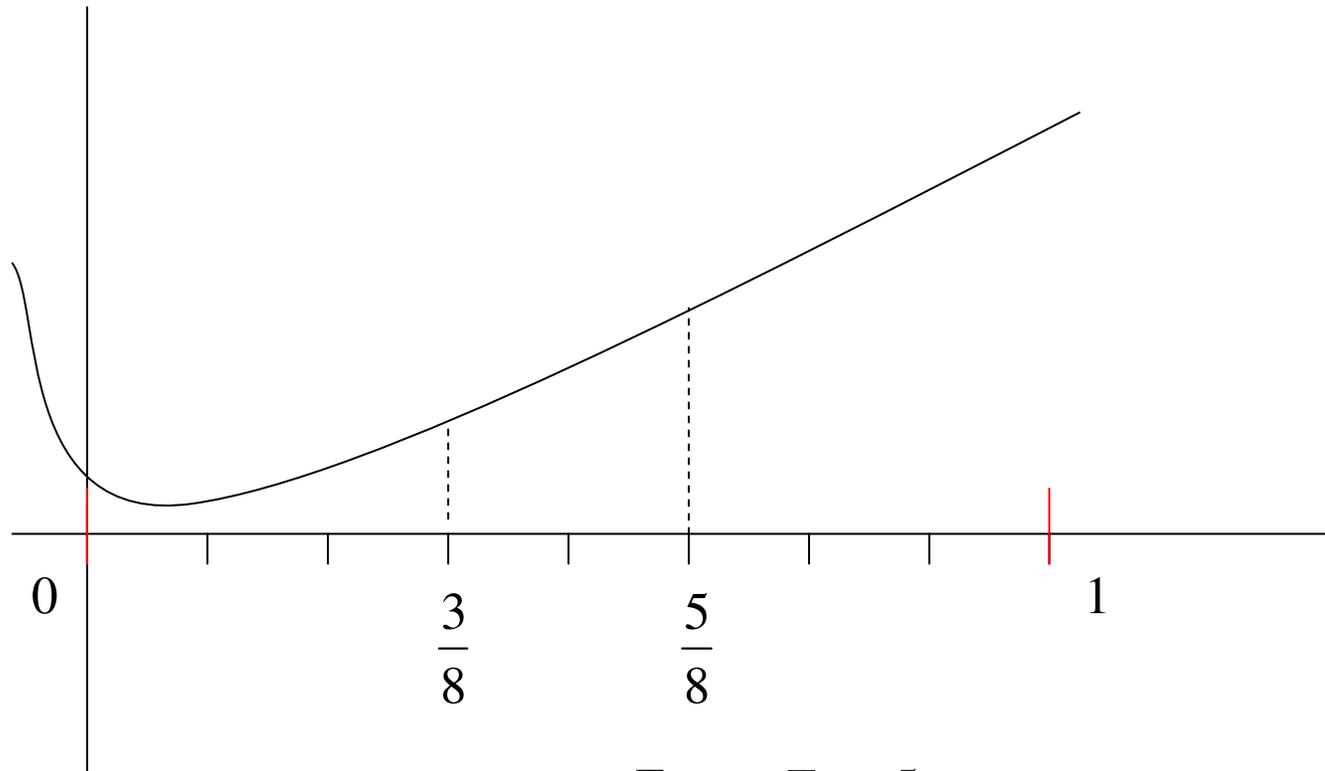
$$d_k = \frac{F_{N-k+1}}{F_N} d_1 \quad k = 2, 3, \dots, N$$

$$d_N = \frac{F_{N-N+1}}{F_N} d_1 = \frac{F_1}{F_N} d_1 = \frac{d_1}{F_N}$$

Ainsi, lorsque le nombre de points d'évaluation N devient très grand pour tendre vers l'infini, la suite des valeurs

$$\{d_k\} \rightarrow 0$$

plus rapidement qu'en utilisant toute autre stratégie.



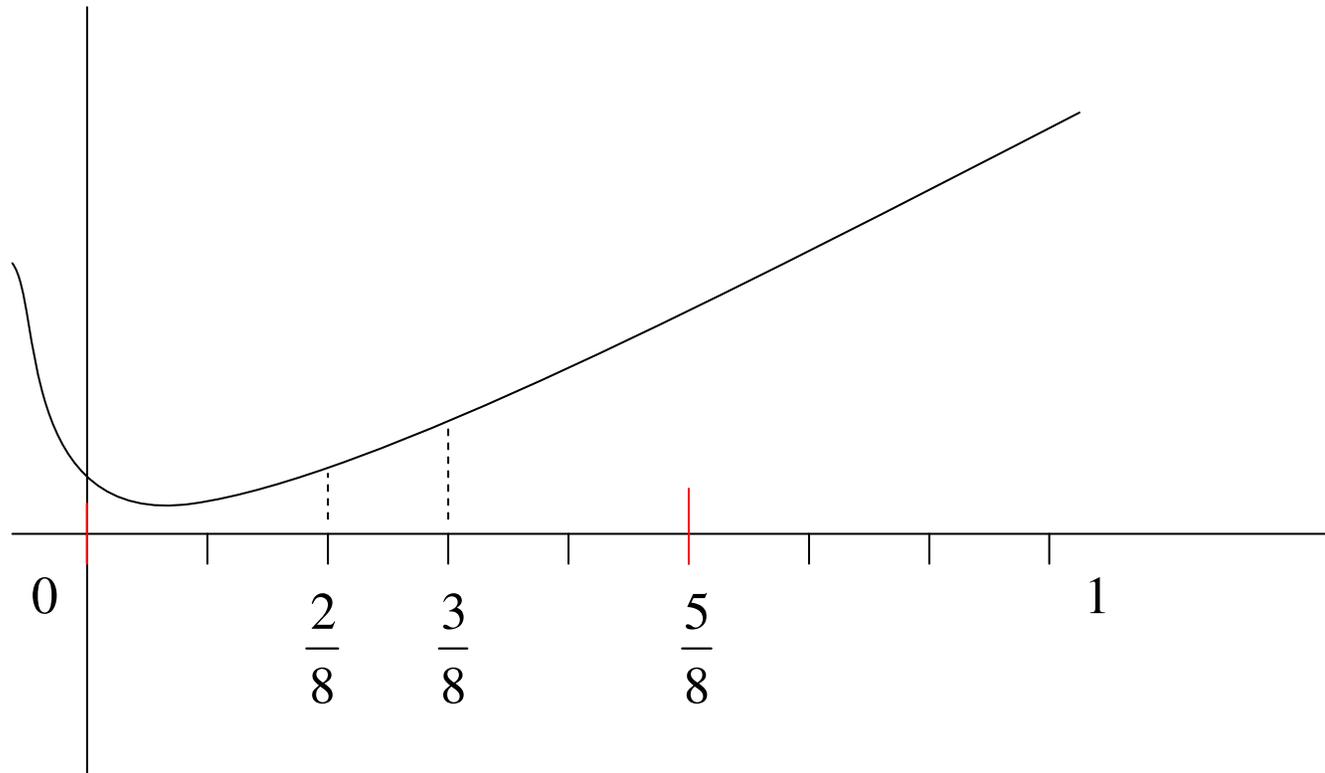
$$N = 5$$

$$\{1, 1, 2, 3, 5, 8\}$$

$$\frac{F_{N-1}}{F_N} = \frac{F_4}{F_5} = \frac{5}{8}$$

$$d_2 = \frac{F_{5-1}}{F_5} d_1 = \frac{5}{8}$$

$$d_k = \frac{F_{N-k+1}}{F_N} d_1 \quad k = 2, 3, \dots, N$$

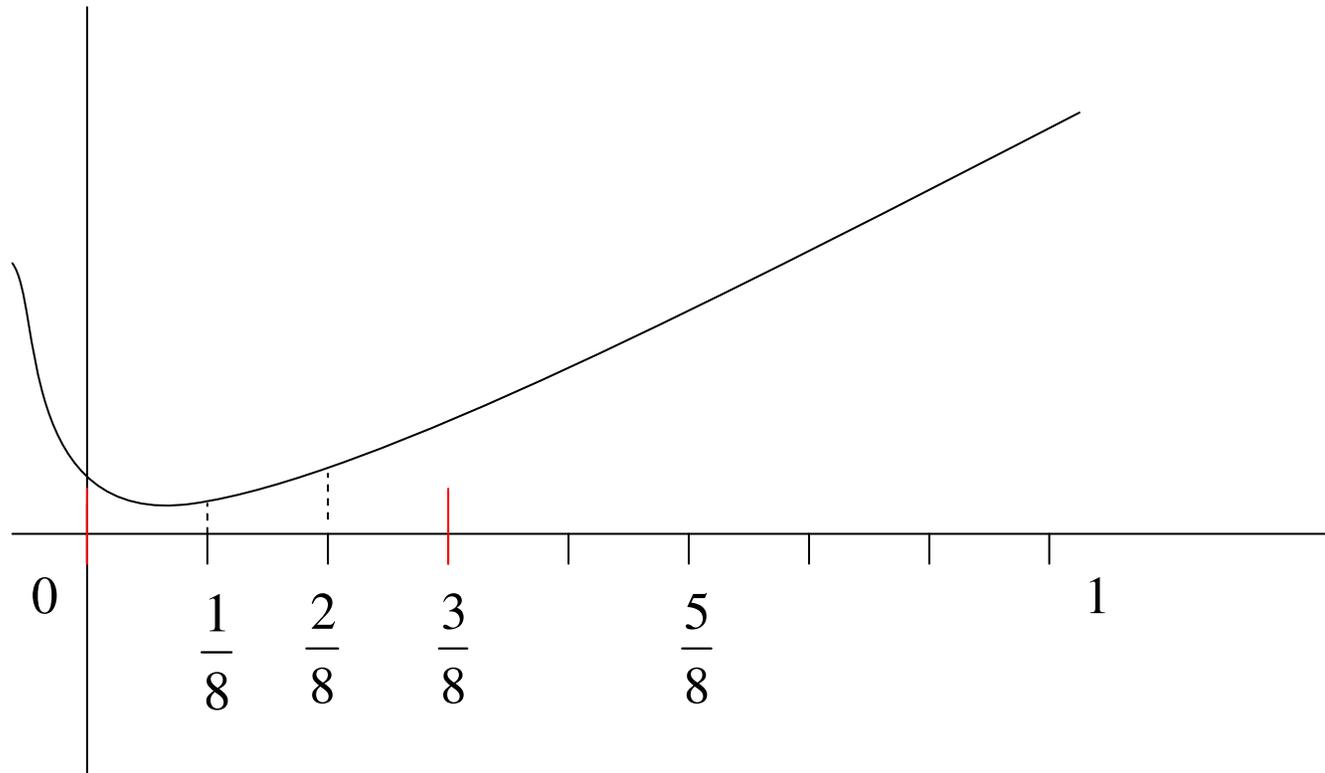


$N = 5$

$\{1, 1, 2, 3, 5, 8\}$

$$d_3 = \frac{F_{5-2}}{F_5} d_1 = \frac{3}{8}$$

$$d_k = \frac{F_{N-k+1}}{F_N} d_1 \quad k = 2, 3, \dots, N$$

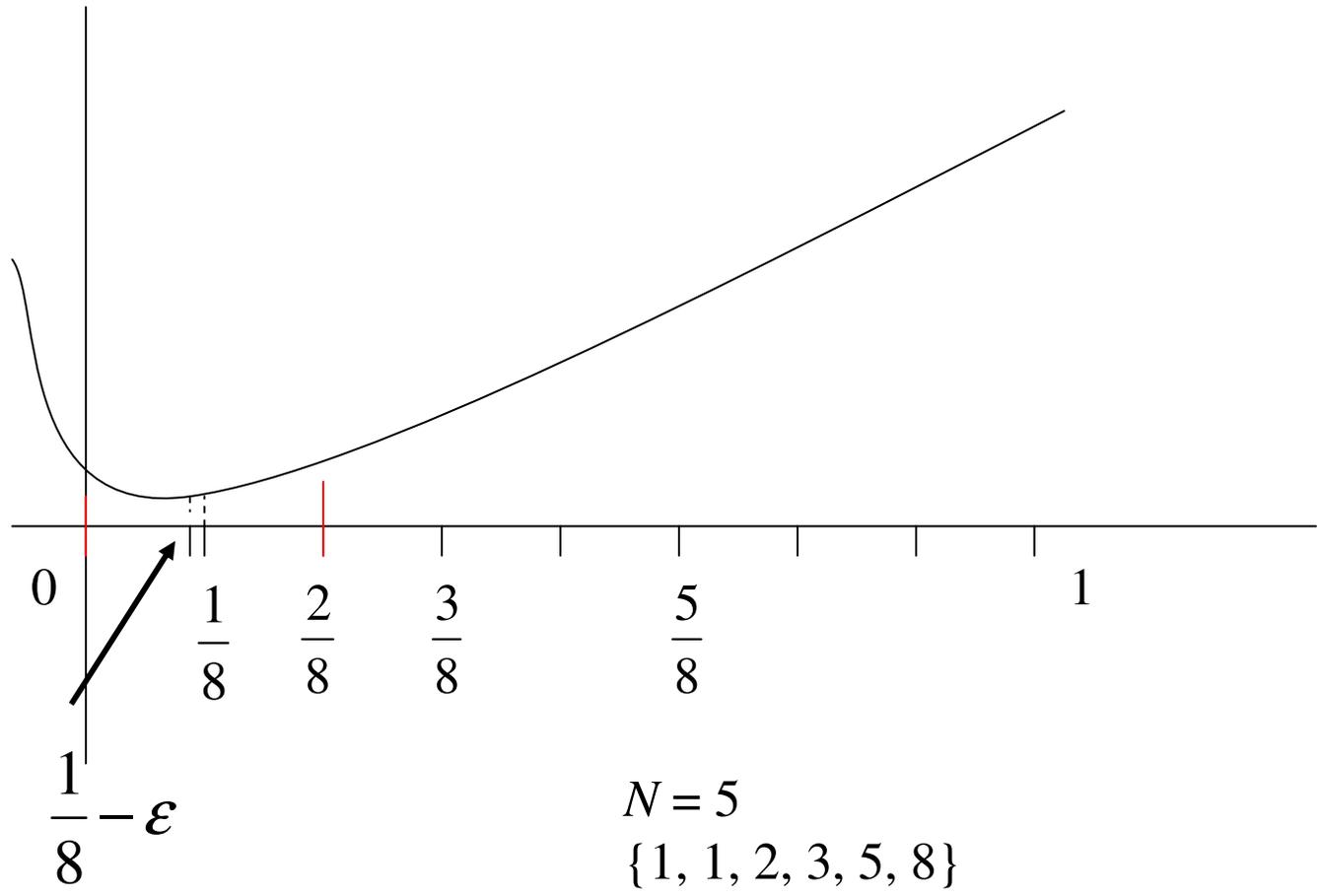


$N = 5$

$\{1, 1, 2, 3, 5, 8\}$

$$d_4 = \frac{F_{5-3}}{F_5} d_1 = \frac{2}{8}$$

$$d_k = \frac{F_{N-k+1}}{F_N} d_1 \quad k = 2, 3, \dots, N$$



$$d_5 = \frac{F_{5-4}}{F_5} d_1 = \frac{1}{8}$$

$$d_k = \frac{F_{N-k+1}}{F_N} d_1 \quad k = 2, 3, \dots, N$$

- **Méthode de la section dorée**

(nombre d'or $\tau = 1.618\dots$)

La méthode de la section dorée utilise la même stratégie que la méthode de Fibonacci pour sélectionner les points d'évaluation, **mais** le nombre de points d'évaluation n'est pas spécifié au départ.

Pour spécifier les deux premiers points, nous procédons comme dans la

méthode de Fibonacci en les prenant symétriques à une distance $\frac{F_{N-1}}{F_N} d_1$ de chaque extrémité en considérant que $N \rightarrow \infty$.

- La méthode de la section dorée utilise la même stratégie que la méthode de Fibonacci pour sélectionner les points d'évaluation, **mais** le nombre de points d'évaluation n'est pas spécifié au départ.

Pour spécifier les deux premiers points, nous procédons comme dans la

méthode de Fibonacci en les prenant symétrique à une distance $\frac{F_{N-1}}{F_N} d_1$ de chaque extrémité en considérant que $N \rightarrow \infty$.

Il est possible de démontrer que $\lim_{N \rightarrow \infty} \frac{F_{N-1}}{F_N} = \frac{1}{\tau}$.

Par conséquent les deux premiers points sont choisis symétriquement à une distance $\frac{1}{\tau} d_1$ de chaque extrémité.

2.2 Méthode utilisant les dérivées

- **Méthode de Newton**

Rappel: Formule de Taylor d'ordre n

Il existe un point z entre x et x^k tel que

$$f(x) = f(x^k) + f'(x^k)(x - x^k) + \frac{f''(x^k)}{2!}(x - x^k)^2 + \dots \\ + \frac{f^{(n-1)}(x^k)}{(n-1)!}(x - x^k)^{n-1} + \frac{f^{(n)}(z)}{n!}(x - x^k)^n$$

- **Méthode de Newton**

Hypothèses: Aux points d'évaluation x^k , il est possible d'évaluer $f(x^k)$, $f'(x^k)$, $f''(x^k)$, et de plus $f''(x^k) \neq 0$.

Étant donné un point d'évaluation x^k , considérons l'approximation quadratique suivante de f à x^k :

$$q_k(x) = f(x^k) + f'(x^k)(x - x^k) + \frac{f''(x^k)}{2!}(x - x^k)^2$$

$$f(x) = f(x^k) + f'(x^k)(x - x^k) + \frac{f''(z)}{2!}(x - x^k)^2$$

z entre x et x^k

$$q_k(x) = f(x^k) + f'(x^k)(x - x^k) + \frac{f''(x^k)}{2!}(x - x^k)^2$$

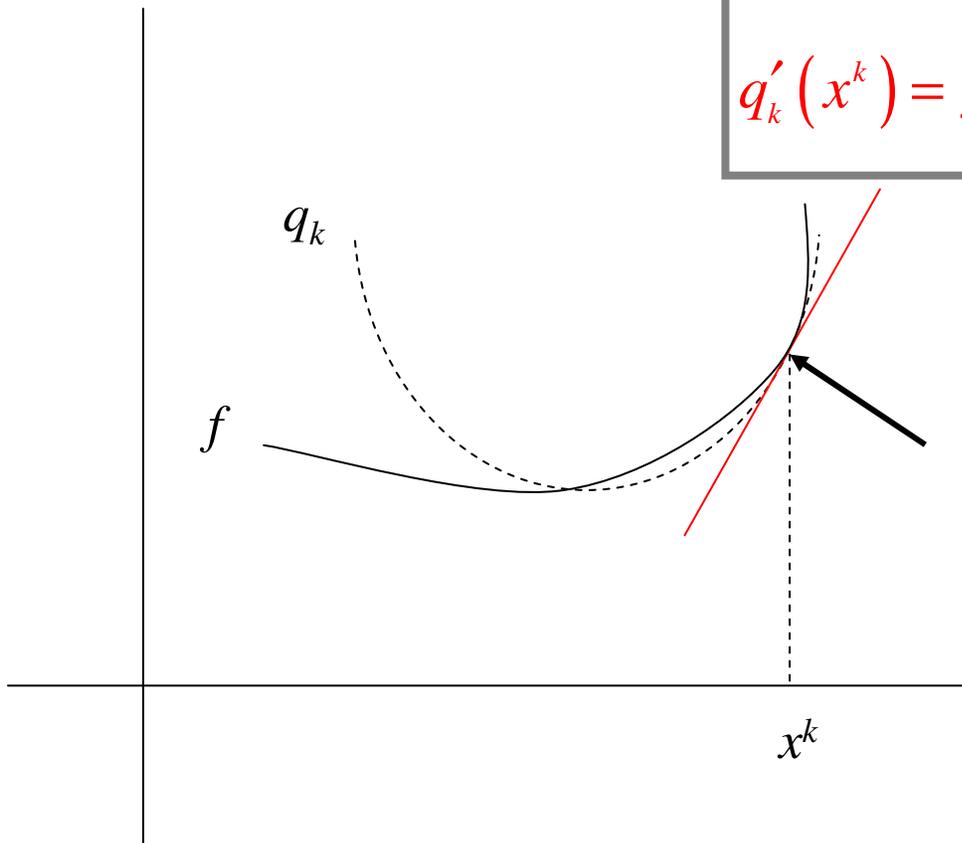
$$q_k(x^k) = f(x^k) + f'(x^k)(x^k - x^k) + \frac{f''(x^k)}{2!}(x^k - x^k)^2 = f(x^k)$$

$$q'_k(x) = f'(x^k) + \cancel{2} \frac{f''(x^k)}{\cancel{2}!}(x - x^k)$$

$$q'_k(x^k) = f'(x^k) + \cancel{2} \frac{f''(x^k)}{\cancel{2}!}(x^k - x^k) = f'(x^k)$$

$$q''_k(x) = f''(x^k)$$

$$q''_k(x^k) = f''(x^k)$$

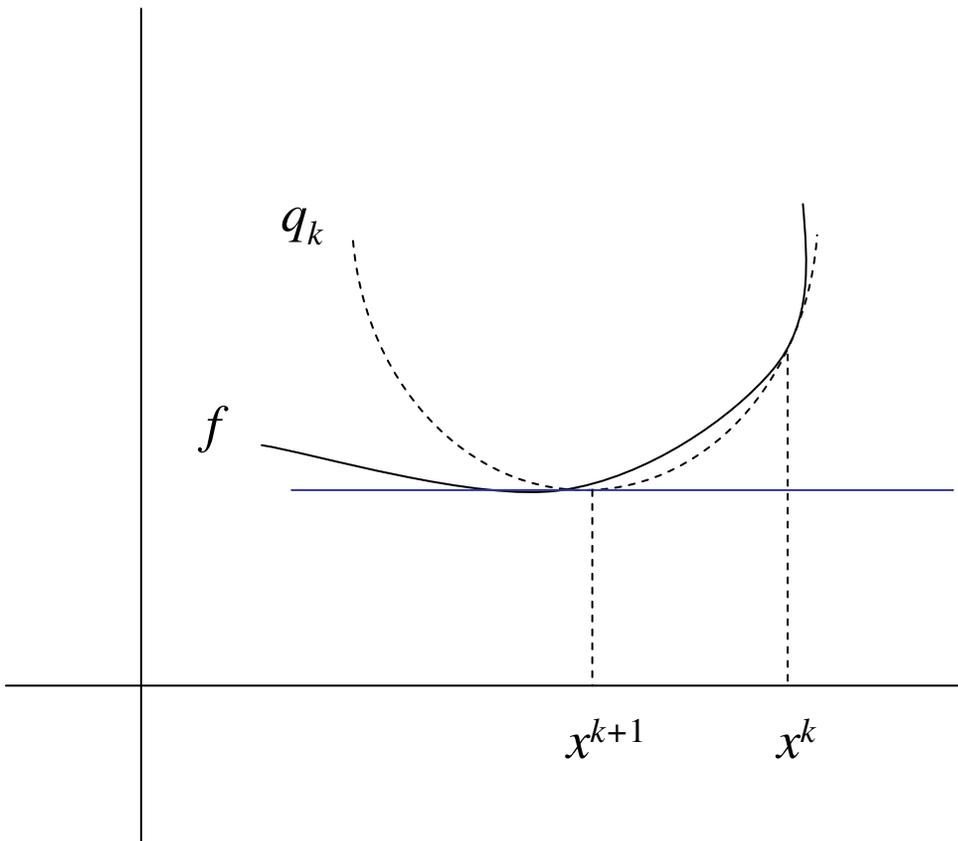


La méthode itérative de Newton détermine le prochain point d'évaluation x^{k+1} en remplaçant f par q_k et en annulant $q'_k(x)$:

$$0 = q'_k(x^{k+1}) = f'(x^k) + f''(x^k)(x^{k+1} - x^k) \Leftrightarrow$$

$$x^{k+1} - x^k = -\frac{f'(x^k)}{f''(x^k)} \quad \Leftrightarrow$$

$$x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}$$



Convergence de la méthode de Newton

Théorème 2.1: Soit f une fonction possédant des dérivées continues d'ordre 3. Supposons que x^* satisfait les conditions $f'(x^*) = 0$ et $f''(x^*) \neq 0$. Si x^0 (le point initial) est choisi suffisamment près de x^* , alors la suite des points $\{x^k\}$ générés par la méthode de Newton converge vers x^* avec un ordre de convergence d'au moins 2.

Preuve. Soient $k_1 > 0$ et $k_2 > 0$ tels que $|f'''(x^*)| < k_1$ et $|f''(x^*)| > k_2$. Par continuité de f''' et de f'' , il existe un scalaire $\varepsilon_1 > 0$ tel que pour tout $\xi \in B_{\varepsilon_1}(x^*) = \{\xi : |\xi - x^*| < \varepsilon_1\}$

$$|f'''(\xi)| < k_1 \text{ et } |f''(\xi)| > k_2.$$

Preuve. Soient $k_1 > 0$ et $k_2 > 0$ tels que $|f'''(x^*)| < k_1$ et $|f''(x^*)| > k_2$.

Par continuité de f''' et de f'' , il existe un scalaire $\varepsilon_1 > 0$ tel que pour

tout $\xi \in B_{\varepsilon_1}(x^*) = \{\xi : |\xi - x^*| < \varepsilon_1\}$

$$|f'''(\xi)| < k_1 \text{ et } |f''(\xi)| > k_2.$$

Puisque $f'(x^*) = 0$, alors la relation définissant x^{k+1}

$$x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}$$

devient

$$x^{k+1} = x^k - \frac{f'(x^k) - f'(x^*)}{f''(x^k)}$$

$$x^{k+1} - x^* = x^k - x^* - \frac{f'(x^k) - f'(x^*)}{f''(x^k)}$$

$$x^{k+1} - x^* = \frac{f''(x^k)(x^k - x^*) - f'(x^k) + f'(x^*)}{f''(x^k)}. \quad (2.1)$$

$$x^{k+1} - x^* = \frac{f''(x^k)(x^k - x^*) - f'(x^k) + f'(x^*)}{f''(x^k)}. \quad (2.1)$$

Se référant à la formule de Taylor d'ordre 2 appliquée à f' ,
il existe un γ entre x^* et x^k tel que

$$f'(x^*) = f'(x^k) + f''(x^k)(x^* - x^k) + \frac{f'''(\gamma)}{2!}(x^* - x^k)^2$$

ce qui s'écrit également

$$-f''(x^k)(x^* - x^k) - f'(x^k) + f'(x^*) = \frac{f'''(\gamma)}{2!}(x^* - x^k)^2$$

ou encore

$$f''(x^k)(x^k - x^*) - f'(x^k) + f'(x^*) = \frac{f'''(\gamma)}{2!}(x^* - x^k)^2.$$

Substituant dans (2.1),

$$x^{k+1} - x^* = \frac{1}{2} \frac{f'''(\gamma)}{f''(x^k)} (x^* - x^k)^2 \Rightarrow |x^{k+1} - x^*| = \frac{1}{2} \frac{|f'''(\gamma)|}{|f''(x^k)|} (x^* - x^k)^2. \quad (2.2)$$

$$x^{k+1} - x^* = \frac{1}{2} \frac{f'''(\gamma)}{f''(x^k)} (x^* - x^k)^2 \Rightarrow |x^{k+1} - x^*| = \frac{1}{2} \frac{|f'''(\gamma)|}{|f''(x^k)|} (x^* - x^k)^2. \quad (2.2)$$

Posons maintenant

$$\varepsilon = \min \left\{ \varepsilon_1, \frac{2k_2}{k_1} \right\}.$$

Démontrons maintenant que partant de n'importe lequel point $x^0 \in B_\varepsilon(x^*)$, nous convergeons vers x^* et que la convergence est d'ordre au moins 2.

En effet si $|x^k - x^*| < \varepsilon$, alors

i) le point γ dans la formule de développement de Taylor a la propriété que $\gamma \in B_{\varepsilon_1}(x^*)$, et ainsi (par définition de k_1 et k_2)

Se référant à la formule de Taylor d'ordre 2 appliquée à f' , il existe un γ entre x^* et x^k tel que

$$f'(x^*) = f'(x^k) + f''(x^k)(x^* - x^k) + \frac{f'''(\gamma)}{2!} (x^* - x^k)^2$$

Convergence de la méthode de Newton

Théorème 2.1: Soit f une fonction possédant des dérivées continues d'ordre 3. Supposons que x^* satisfait les conditions $f'(x^*) = 0$ et $f''(x^*) \neq 0$. Si x^0 (le point initial) est choisi suffisamment près de x^* , alors la suite des points $\{x^k\}$ générés par la méthode de Newton converge vers x^* avec un ordre de convergence d'au moins 2.

Preuve. Soient $k_1 > 0$ et $k_2 > 0$ tels que $|f'''(x^*)| < k_1$ et $|f''(x^*)| > k_2$. Par continuité de f''' et de f'' , il existe un scalaire $\varepsilon_1 > 0$ tel que pour tout $\xi \in B_{\varepsilon_1}(x^*) = \{\xi : |\xi - x^*| < \varepsilon_1\}$

$$|f'''(\xi)| < k_1 \text{ et } |f''(\xi)| > k_2.$$

$$x^{k+1} - x^* = \frac{1}{2} \frac{f'''(\gamma)}{f''(x^k)} (x^* - x^k)^2 \Rightarrow |x^{k+1} - x^*| = \frac{1}{2} \frac{|f'''(\gamma)|}{|f''(x^k)|} (x^* - x^k)^2. \quad (2.2)$$

Posons maintenant

$$\varepsilon = \min \left\{ \varepsilon_1, \frac{2k_2}{k_1} \right\}.$$

Démontrons maintenant que partant de n'importe lequel point $x^0 \in B_\varepsilon(x^*)$, nous convergeons vers x^* et que la convergence est d'ordre au moins 2.

En effet si $|x^k - x^*| < \varepsilon$, alors

i) le point γ dans la formule de développement de Taylor a la propriété que $\gamma \in B_{\varepsilon_1}(x^*)$, et ainsi (par définition de k_1 et k_2)

$$|f'''(\gamma)| < k_1 \text{ et } |f''(\gamma)| > k_2.$$

Puisque $|x^k - x^*| < \varepsilon$, alors également $x^k \in B_{\varepsilon_1}(x^*)$ et

$$|f'''(x^k)| < k_1 \text{ et } |f''(x^k)| > k_2.$$

Substituant dans (2.2),

$$|x^{k+1} - x^*| < \frac{1}{2} \frac{k_1}{k_2} (x^* - x^k)^2;$$

$$\varepsilon = \min \left\{ \varepsilon_1, \frac{2k_2}{k_1} \right\}$$

En effet si $|x^k - x^*| < \varepsilon$, alors

i) le point γ dans la formule de développement de Taylor a la propriété que $\gamma \in B_{\varepsilon_1}(x^*)$, et ainsi (par définition de k_1 et k_2)

$$|f'''(\gamma)| < k_1 \text{ et } |f''(\gamma)| > k_2,$$

Puisque $|x^k - x^*| < \varepsilon$, alors également $x^k \in B_{\varepsilon_1}(x^*)$ et

$$|f'''(x^k)| < k_1 \text{ et } |f''(x^k)| > k_2.$$

Substituant dans (2.2)

$$|x^{k+1} - x^*| < \frac{1}{2} \frac{k_1}{k_2} (x^* - x^k)^2;$$

ii) également

$$|x^k - x^*| < \varepsilon \leq \frac{2k_2}{k_1}$$

$$|x^{k+1} - x^*| < \frac{1}{2} \frac{k_1}{k_2} (x^* - x^k)^2;$$

ii) également

$$|x^k - x^*| < \varepsilon \leq \frac{2k_2}{k_1}$$

et par conséquent

$$|x^k - x^*| \frac{1}{2} \frac{k_1}{k_2} < 1.$$

Il s'ensuit que

$$|x^{k+1} - x^*| < \frac{1}{2} \frac{k_1}{k_2} (x^* - x^k)^2 = |x^k - x^*| \boxed{|x^k - x^*| \frac{1}{2} \frac{k_1}{k_2}}$$

$$|x^{k+1} - x^*| < |x^k - x^*|,$$

et la méthode converge.

De plus, la convergence est d'ordre au moins 2 puisque

$$\frac{|x^{k+1} - x^*|}{|x^* - x^k|^2} < \frac{1}{2} \frac{k_1}{k_2}.$$

□

- **Note:** hypothèse $f''(x^*) \neq 0$ assure que x^* est un maximum ou un minimum de f , et qu'il est possible d'utiliser un $k_2 \neq 0$.
- Méthode également utilisée pour déterminer un point où la fonction s'annule. Il suffit de considérer la fonction $g(x) = f'(x)$

$$x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}$$

$$x^{k+1} = x^k - \frac{g(x^k)}{g'(x^k)} \Leftrightarrow \frac{g(x^k)}{g'(x^k)} = x^k - x^{k+1}$$

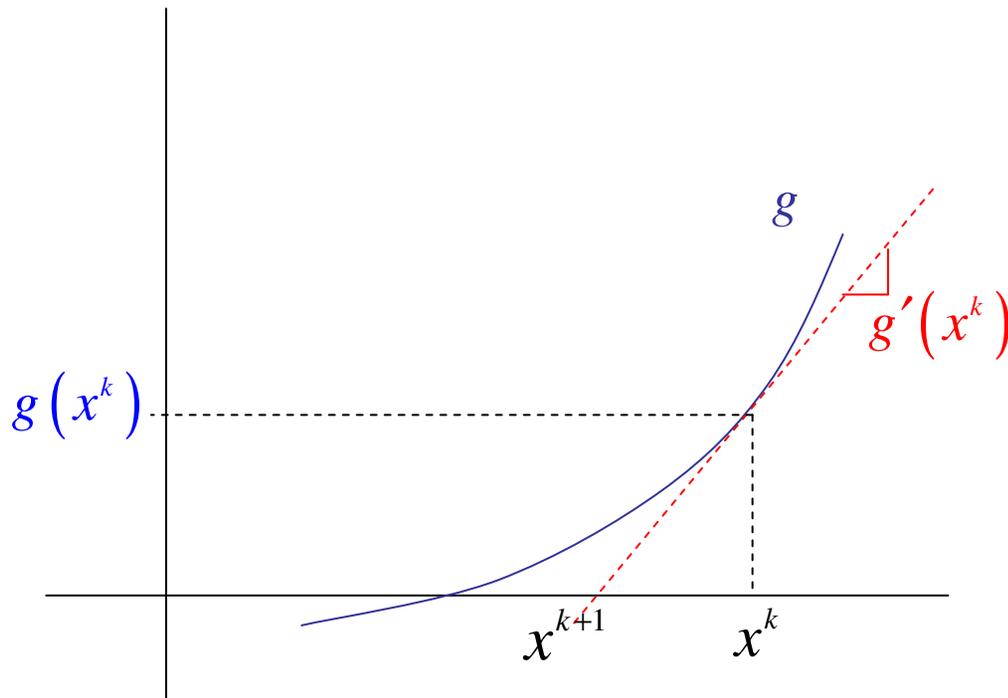
ou encore

$$g'(x^k) = \frac{g(x^k) - 0}{x^k - x^{k+1}}$$

$$x^{k+1} = x^k - \frac{g(x^k)}{g'(x^k)}$$

ou encore

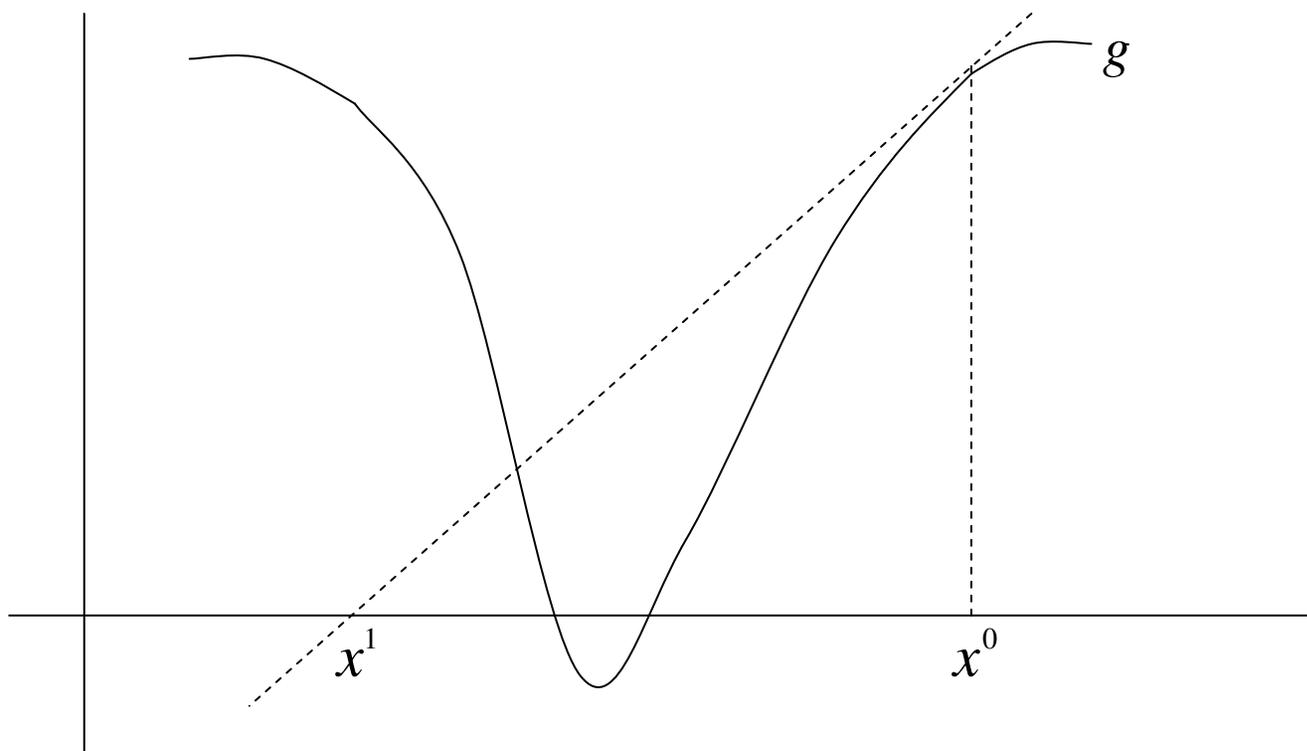
$$g'(x^k) = \frac{g(x^k) - 0}{x^k - x^{k+1}}$$



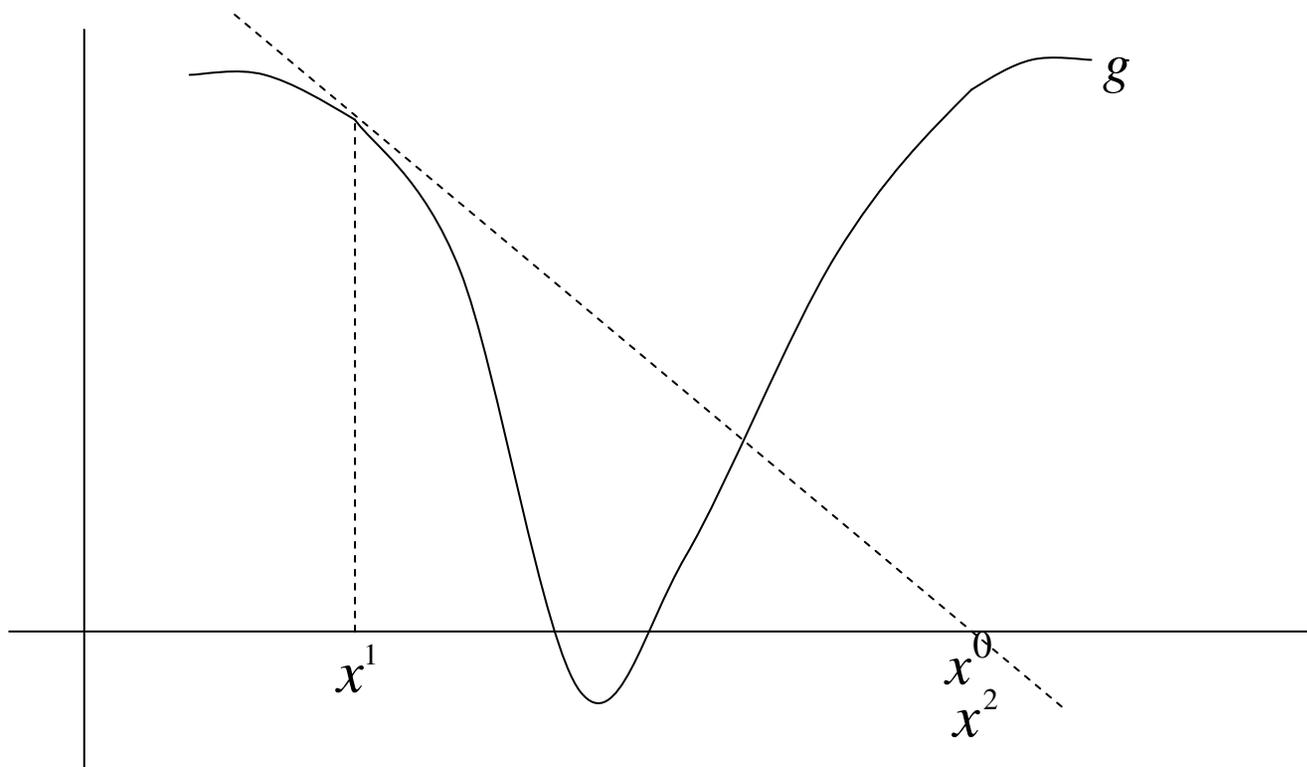
Interprétation géométrique:
 x^{k+1} choisi de telle sorte que
 de la droite passant par les
 points $(x^k, g(x^k))$ et $(x^{k+1}, 0)$
 a pour pente $g'(x^k)$, celle de
 g au point x^k .

x^{k+1} à l'intersection de l'axe des x et la tangente de g au point x^k

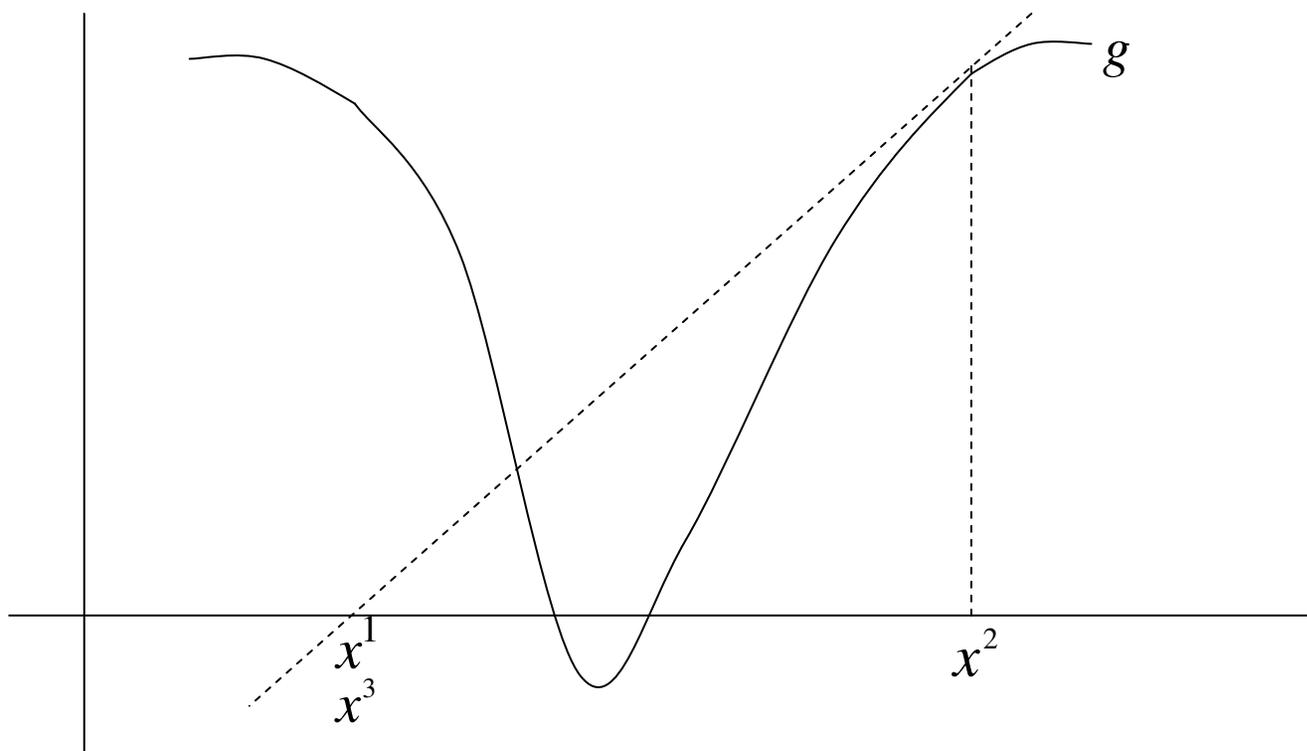
- Importance de l'hypothèse que x^0 soit suffisamment près de x^*



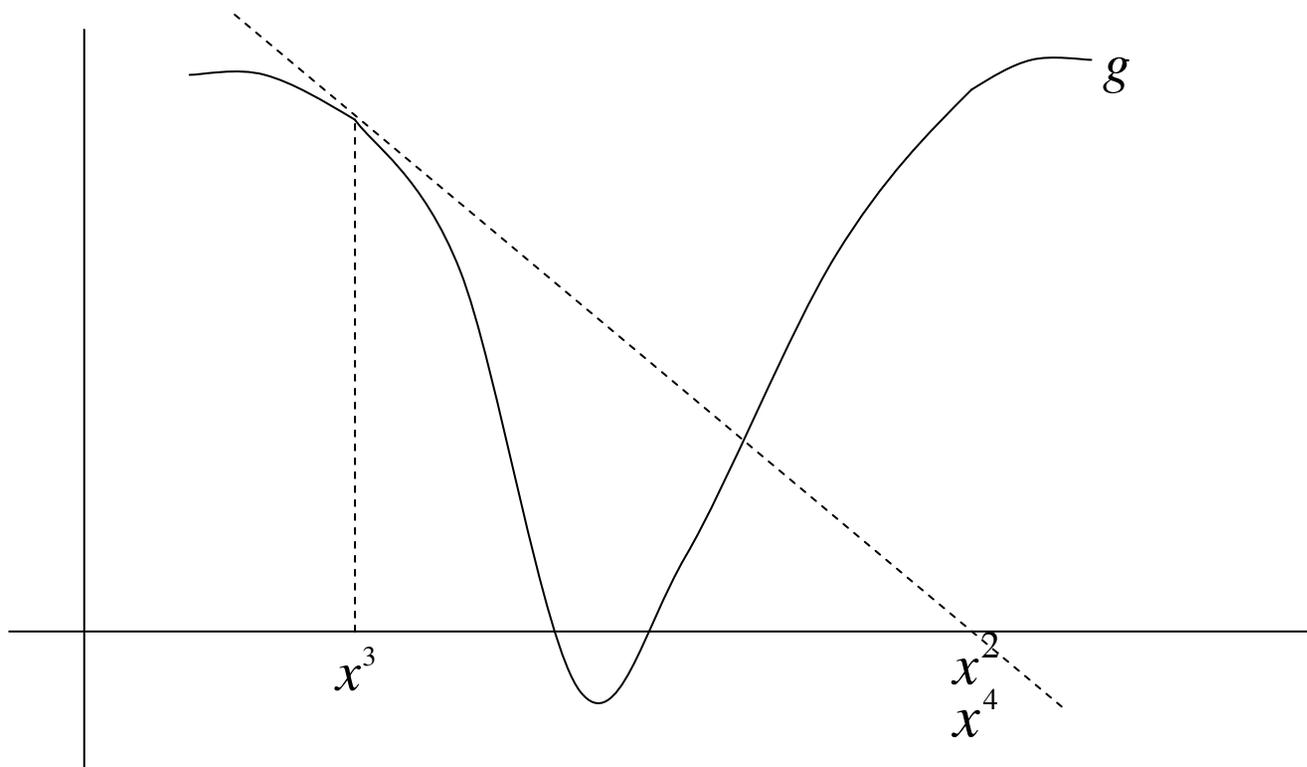
- Importance de l'hypothèse que x^0 soit suffisamment près de x^*



- Importance de l'hypothèse que x^0 soit suffisamment près de x^*



- Importance de l'hypothèse que x^0 soit suffisamment près de x^*



- Importance de l'hypothèse que x^0 soit suffisamment près de x^*

