

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET
POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique



Centre universitaire Salhi Ahmed -Naâma
Institut des Sciences et technologies
Département de Mathématiques et Informatique

Mémoire de fin d'étude

Pour obtenir le diplôme de
Master En Mathématiques

Spécialité : Analyse fonctionnelle et EDPs

Filière : Mathématiques

Thème

LA RÉOLUTION NUMÉRIQUE DES
PROBLÈMES D'OPTIMISATION PAR LA
MÉTHODE DE NEWTON-RAPHSON

Présenté par :

BELADJANE MEYMOUNA

Soutenu le : 08/07/2019

Devant le jury composé de :

Encadreur : *M^r Kennouza Djamel*

Président : *M^{lle} Moussaoui Fatima*

Examineur : *M^{me} Bellabess Amel*

Pr.C-Univ Salhi Ahmed -Naâma-.

Pr.C-Univ Salhi Ahmed -Naâma-.

Pr.C-Univ Salhi Ahmed -Naâma-.

Année universitaire 2018/2019

Remmercien

Tout d'abord, on doit remercier ALLAH, le tout puissant de nous avoir offert le courage et la volonté, pour réaliser ce modeste travail. On tient, en premier lieu, à exprimer notre profonde gratitude à notre encadreur, Monsieur kennouzza djamel, enseignant au Centre universitaire SALHI Ahmed de Naâma, pour ses directives bénéfiques, sa vision scientifique, son soutien constant et ses conseils. On adresse nos remerciements à Madame Bellabess , enseignante au Centre universitaire SALHI Ahmed de Naâma, d'avoir acceptée de faire partie du jury de ce mémoire et pour le temps et l'attention qu'elle a consacrée à la lecture et à la critique de ce travail. On remercie également Madame Moussaoui , enseignante au Centre universitaire SALHI Ahmed de Naâma, qui a accepté, lui aussi, de faire partie de ce jury et pour le temps de lecture attentive et ses réflexions précieuses. On adresse aussi, nos remerciements à tout le personnel, enseignant et administratif, de département de mathématiques et d'informatique.

DEDICACE

Ce travail est particulièrement dédié à :
Mes parents, tout spécialement ma très chère mère
et mon père.
Mes frères et ma soeur.
Toute ma famille.
Tous mes amis, et à tous ceux qui aiment l'humanité
qui en ce monde on se prive davantage.

**"C'est par la logique que nous prou-
vons, c'est par l'intuition que nous
inventons"**

Sommaire

Table des figures	1
Introduction	2
1 Historique et définitions	3
1.1 Isaac Newton (1643-1727)	3
1.2 Joseph Raphson (1648-1715)	5
1.3 Fonction convexe	6
1.4 Fonction coersive	7
1.5 Probleme d'optimisation	7
1.6 Maximum de vraisemblance	8
1.6.1 Principe	8
1.7 La méthode des moindres carrées	11
1.8 Le développement limité de Taylor	12
1.9 Inverce d'une matrice	12
2 Méthode de Newton Raphson	14
2.1 Introduction	14
2.2 Existence de solution	15
2.3 Unicité de solution	15
2.4 La méthode du Newton-Raphson	16
2.4.1 Algorithme	16
2.4.2 Méthode de N-R pour deux inconnues	18
2.4.3 Convergence de méthode de N-R	20

2.4.4	Vitesse de convergence de la méthode de N-R	22
2.4.5	Test d'arrêt	23
	Application numérique	25
	Conclusion	29
	REFERENCES BIBLIOGRAPHIQUES	30

Table des figures

1.1	Isaac Newton	3
1.2	Joseph Raphson	5
1.3	Droite des moindres carrés	12
2.1	La méthode de N-R	16
2.2	La méthode de N-R	17
2.3	Échec de la méthode de N-R	22
2.4	L'organigramme de résolution	25
2.5	Programme de résolution en MatLab	26

Introduction générale

L'optimisation est un sujet très ancien qui connaît un nouvel essor depuis l'apparition des ordinateurs et dont les méthodes s'appliquent dans de très nombreux domaines : économie, gestion, planification, logistique, automatique, etc.

L'objet de cette mémoire est de présenter et analyser un algorithme permet de calculer, ou plus exactement d'**approcher** la solution des problèmes d'optimisation. Tous les algorithmes étudiés sont effectivement utilisés en pratique pour résoudre sur ordinateur des problèmes concrets d'optimisation.

Ces algorithmes sont aussi tous de nature itérative : à partir d'une donnée initiale x_0 , chaque méthode construit une suite $(x_n)_{n \in \mathbb{N}}$ dont nous montrerons qu'elle converge, sous certaines hypothèses, vers la solution x du problème d'optimisation considéré. Après avoir montré la convergence de ces algorithmes (c'est-à-dire, la convergence de la suite $(x_n)_n$ vers x quel que soit le choix de la donnée initiale x_0), nous dirons aussi un mot de leur vitesse de convergence.

Dans cet mémoire, nous intéressons d'expliquer la méthode de Newton-Raphson qu'est le plus important en pratique (si la fonction est différentiable et que sa dérivée est connue).

Chapitre 1

Historique et définitions

1.1 Isaac Newton (1643-1727)

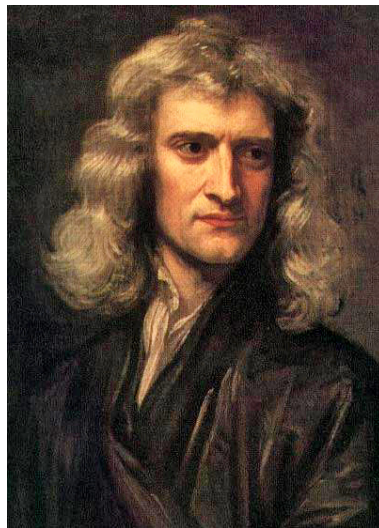


FIGURE 1.1 – Isaac Newton

[3][7] Isaac Newton (né le 4 janvier 1643 et décédé le 31 mars 1727) est un physicien, philosophe, astronome, et mathématicien anglais, considéré comme l'un des plus grands scientifiques de tous les temps. Newton a formulé des lois sur la gravitation universelle et sur les corps en mouvement. Ces lois fondamentales expliquent de quelle façon les objets se déplacent sur terre comme dans les airs. Il a fondé l'op-

tique moderne, étudié le comportement de la lumière, et a construit le premier télescope à miroirs.

Ses travaux en mathématiques l'ont conduit à l'invention d'une branche des mathématiques nommée calculus (calcul différentiel et intégral, également développé par le mathématicien allemand Gottfried Wilhelm Leibniz).

Newton a exposé ses idées dans plusieurs publications et deux d'entre elles, *Philosophiae Naturalis Principia Mathematica* (Principes mathématiques de la philosophie naturelle, en 1687) et *Opticks* (traité d'optique, en 1704) sont considérées parmi les plus grands travaux scientifiques jamais réalisés.

Les contributions révolutionnaires d'Isaac Newton ont permis d'expliquer une grande partie du monde qui nous entoure en termes mathématiques, et ont permis de se rendre compte que la science était également en mesure d'expliquer un grand nombre d'autres phénomènes.

Newton était capable de formuler des théories mathématiques pour expliquer des faits connus. Il utilisait ses théories mathématiques pour prédire le comportement d'objets dans diverses situations, puis comparait ses prédictions avec ce qu'il avait observé dans les expériences qu'il avait menées. Newton utilisait ensuite ces résultats afin de vérifier, et modifier si besoin, ses théories. Il a réussi à faire de l'explication de propriétés physiques un moyen de prédiction !

1.2 Joseph Raphson (1648-1715)



FIGURE 1.2 – Joseph Raphson

Né à Middlesex, en Angleterre, en 1648, Joseph Raphson se rendit au Jesus College Cambridge où il obtint son diplôme de maîtrise en arts en 1692. En tant qu'étudiant un peu plus âgé (43 ans), Raphson fut étonnamment nommé membre de la Royal Society en 1691, un an avant sa graduation. Cet honneur découle directement de son livre publié en 1690, intitulé *Analysis aequationum universalis*. Ce livre s'est principalement concentré sur la méthode de Newton pour approximer les racines d'une équation, d'où la méthode de Newton-Raphson. La méthode des flux de Newton décrit la même méthode et des exemples pour approximer les racines des équations. Cependant, elle a été écrite en 1671 et n'a pas été publiée avant 1736. Joseph Raphson a donc publié ce matériau et sa méthode 50 ans avant Newton !

Bien que la relation de Raphson avec Newton ne soit pas bien comprise, il est connu que ce dernier était autorisé par lui-même à visualiser et à étudier périodiquement

les travaux mathématiques de Newton. Une telle pratique n'était pas courante. En fait, Raphson et Edmund Halley ont collaboré avec Newton à la publication des travaux de Newton du début des années 1670 sur les courbes en quadrature, les fluxions et une sorte de mathématique connue maintenant sous le nom de calcul, mais ils n'ont pas pu le faire avant 1704. La publication résultante de Newton donne une explication détaillée de la méthode de fluxion de Newton. Raphson a également été autorisé à consulter les documents mathématiques de Newton en 1711, via Roger Cotes et William Jones. En conséquence, Raphson a écrit une Histoire de Fluxions, qui a été imprimée mystérieusement un an après la mort de Raphson.

Ce travail appuyait fortement le calcul de Newton sur la seule revendication de Leibniz d'avoir été l'inventeur du calcul. Les vues de Newton sur le contenu du livre de Raphson étaient presque certainement favorables, puisque Newton avait publié des lettres échangées entre lui et Leibnitz sur le débat sur l'identité exacte de l'inventeur du calcul dans l'annexe de la réimpression d'une histoire de fluxions.

1.3 Fonction convexe

Dans cette section, E désigne un espace vectoriel réel, de dimension finie ou non, C est une partie convexe non vide de E , et l'on se donne une fonction $f : C \subset E \rightarrow \mathbb{R}$

Définition 1.3.1. [6] La fonction f est dite convexe si, pour tous $x, y \in C$ et $\lambda \in [0, 1]$, on a :

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y) \quad (1.1)$$

Si cette inégalité est stricte dès que $x \neq y$ et $\lambda \neq 0, 1$, f est dite strictement convexe.

Théorème 1. [6] Si I est ouvert de \mathbb{R} , une fonction dérivable $f : I \rightarrow \mathbb{R}$ est convexe (resp. strictement convexe) si, et seulement si, f' est croissante (resp. strictement croissante).

Théorème 2. [6] Une fonction $f : I \rightarrow \mathbb{R}$ deux fois dérivable est convexe si, et seulement si, $f'' \geq 0$ sur I . L'intervalle I est ici quelconque.

1.4 Fonction coersive

Définition 1.4.1. [6] Soit $\Omega \subset \mathbb{R}^n$ un ensemble non borné et $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$. On dit que f est coercive sur Ω si on a :

$$\lim_{\substack{\|x\| \rightarrow +\infty \\ x \in \Omega}} f(x) = +\infty$$

ou de manière plus précise :

$$\forall \nu \in \mathbb{R}, \quad \exists \rho \geq 0; \quad (x \in \Omega \text{ et } \|x\| \geq \rho) \quad \Rightarrow \quad f(x) \geq \nu.$$

1.5 Probleme d'optimisation

L'optimisation est une branche des mathématiques cherchant à modéliser, à analyser et à résoudre analytiquement ou numériquement les problèmes qui consistent à minimiser ou maximiser une fonction sur un ensemble.

L'optimisation joue un rôle important en recherche opérationnelle (domaine à la frontière entre l'informatique, les mathématiques et l'économie), dans les mathématiques appliquées (fondamentales pour l'industrie et l'ingénierie), en analyse et en analyse numérique, en statistique pour l'estimation du maximum de vraisemblance d'une distribution, pour la recherche de stratégies dans le cadre de la théorie des jeux, ou encore en théorie du contrôle et de la commande.

Beaucoup de systèmes susceptibles d'être décrits par un modèle mathématique sont optimisés. La qualité des résultats et des prédictions dépend de la pertinence du modèle, de l'efficacité de l'algorithme et des moyens pour le traitement numérique. tous les problèmes d'optimisation que nous considérerons peuvent être exprimés de la façon suivante :

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{sous la contrainte} \quad : x \in X \quad (1.2)$$

où X est un sous-ensemble de \mathbb{R}^n . On pourra écrire aussi

$$\min_{x \in X} f(x)$$

Les variables $x = (x_1; \dots; x_n)$ sont appelées "variables d'optimisation" ou variables de décision. La fonction $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est appelée fonction objectif et l'ensemble X ensemble des contraintes. Le problème (1.2) est dit réalisable si $X \neq \emptyset$.

1.6 Maximum de vraisemblance

l'estimateur du maximum de vraisemblance est un estimateur (statistique) utilisé pour inférer les paramètres de la loi de probabilité d'un échantillon donné. Cette méthode a été développée par le statisticien Ronald Aylmer Fisher en 1922.

1.6.1 Principe

Soit une famille paramétrique de distributions de probabilités D_θ dont les éléments sont associés soit à une densité de probabilité connue (distribution continue), soit à une fonction de masse connue (distribution discrète), notée $f(x|\theta)$. On tire un échantillon de n valeurs x_1, x_2, \dots, x_n

de la distribution, et l'on calcule la densité de probabilité associée aux données observées :

$$f_{\theta}(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i | \theta)$$

Ceci étant une fonction de θ avec x_1, x_2, \dots, x_n fixés, c'est une vraisemblance.

$$L(\theta) = f_{\theta}(x_1, \dots, x_n; \theta)$$

Lorsque θ n'est pas observable, la méthode du maximum de vraisemblance utilise les valeurs de θ qui maximisent $L(\theta)$ estimateur de θ : c'est l'estimateur du maximum de vraisemblance de θ noté $\hat{\theta}$. Par exemple dans le cas du produit discret, on effectue un tirage de n valeurs, il faut donc trouver le paramètre qui maximise la probabilité d'avoir tiré ce tirage.

Cette méthode se distingue de la recherche d'un estimateur non biaisé de θ , ce qui ne donne pas nécessairement la valeur la plus probable pour θ .

L'estimateur du maximum de vraisemblance peut exister et être unique, ne pas être unique, ou ne pas exister.

Définition 1.6.1. Soit X une variable aléatoire réelle, de loi discrète ou continue, dont on veut estimer un paramètre θ . On note D_{θ} cette famille de lois paramétriques. Alors on définit une fonction f telle que :

$$f(x; \theta) = \begin{cases} f_{\theta}(x) & \text{Si } X \text{ est une v.a. continue,} \\ P_{\theta}(X = x) & \text{Si } X \text{ est une v.a. discrète.} \end{cases}$$

$f_{\theta}(x)$ représente la densité de X (où θ apparaît) et $P_{\theta}(X = x)$ représente une probabilité discrète (où θ apparaît).

On appelle vraisemblance de θ au vu des observations $(x_1, \dots, x_i, \dots, x_n)$ d'un n -échantillon indépendamment et

identiquement distribué selon la loi D_θ , le nombre :

$$L(x_1, \dots, x_i, \dots, x_n; \theta) = f(x_1; \theta) \times f(x_2; \theta) \times \dots \times f(x_n; \theta) = \prod_{i=1}^n f(x_i | \theta)$$

On cherche à trouver le maximum de cette vraisemblance pour que les probabilités des réalisations observées soient aussi maximum. Ceci est un problème d'optimisation. On utilise généralement le fait que si L est dérivable (ce qui n'est pas toujours le cas) et si L admet un maximum global en une valeur $\theta = \hat{\theta}$, alors la dérivée première s'annule en $\theta = \hat{\theta}$ et que la dérivée seconde est négative. Réciproquement, si la dérivée première s'annule en $\theta = \hat{\theta}$ et que la dérivée seconde est strictement négative en $\theta = \hat{\theta}$, alors $\theta = \hat{\theta}$ est un maximum local de $L(x_1, \dots, x_i, \dots, x_n; \theta)$. Il est alors nécessaire de vérifier qu'il s'agit bien d'un maximum global. La vraisemblance étant positive et le logarithme népérien une fonction croissante, il est équivalent et souvent plus simple de maximiser le logarithme népérien de la vraisemblance (le produit se transforme en somme, ce qui est plus simple à dériver). On peut facilement construire la statistique $Y_n = \Theta$ qui est l'estimateur voulu.

Ainsi en pratique :

- . La condition nécessaire

$$\frac{\partial L(x_1, \dots, x_i, \dots, x_n; \theta)}{\partial \theta} = 0$$

ou

$$\frac{\partial \ln L(x_1, \dots, x_i, \dots, x_n; \theta)}{\partial \theta} = 0$$

permet de trouver la valeur $\theta = \hat{\theta}$.

- . $\theta = \hat{\theta}$ est un maximum local si la condition suffisante est remplie au point critique $\theta = \hat{\theta}$:

$$\frac{\partial^2 L(x_1, \dots, x_i, \dots, x_n; \theta)}{\partial \theta^2} < 0$$

ou

$$\frac{\partial^2 \ln L(x_1, \dots, x_i, \dots, x_n; \theta)}{\partial \theta^2} < 0$$

Pour simplifier, dans les cas de lois continues, où parfois la densité de probabilité est nulle sur un certain intervalle, on peut omettre d'écrire la vraisemblance pour cet intervalle uniquement.

1.7 La méthode des moindres carrés

La méthode des moindres carrés, indépendamment élaborée par Legendre et Gauss au début du XIXe siècle, permet de comparer des données expérimentales, généralement entachées d'erreurs de mesure, à un modèle mathématique censé décrire ces données. On considère un nuage de points $M_i(x_i, y_i)$ que l'on désire ajuster au mieux par une courbe mathématique (c) de type $f(x) : x \rightarrow y$ dont on devra choisir le type de façon pertinente eu égard au phénomène étudié. On recherche les paramètres de f , fonction affine, polynôme, exponentielle, etc., minimisant la somme des carrés des distances entre y_i et $f(x_i)$, autrement dit :

$$T = \sum_{i=1}^n [y_i - f(x_i)]^2$$

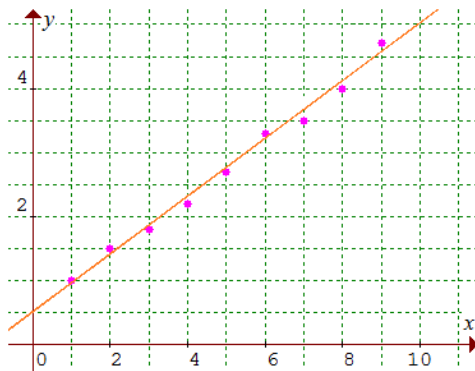


FIGURE 1.3 – Droite des moindres carrés

1.8 Le développement limité de Taylor

La série de Taylor est basée sur le résultat mathématique suivant : toute fonction $u(x)$ peut être exprimée à l'aide d'un polynôme ayant un nombre infini de termes. C'est le développement en série de cette fonction autour d'un point connu x_0 voisin de x :

$$u(x) = u(x_0) + (x - x_0)u'(x_0) + \frac{(x - x_0)^2}{2!}u''(x_0) + \dots + \frac{(x - x_0)^n}{n!}u^{(n)}(x_0) + \theta((x - x_0)^n)$$

1.9 Inverse d'une matrice

Définition 1.9.1. Transposée d'une matrice :

Soit $A = [a_{ij}]_{i,j=1}^{n,m} \in M_{n,m}(K)$. On appelle transposée de A , la matrice A^T dont les lignes sont les colonnes de A . Autrement dit :

$$A^T = [b_{ji}]_{i,j=1}^{m,n} \in M_{m,n}(K) \quad \text{avec} \quad b_{ji} = a_{ij}$$

Théorème 3. *Le déterminant d'une matrice carrée et celui de sa transposée sont égaux.*

Définition 1.9.2. Soit $A = [a_{ij}]_{i,j=1}^n \in M_n(K)$ une matrice carrée. On appelle :

1. mineur d'ordre (i,j) le scalaire : $|A_{ij}|$, où A_{ij} est la matrice obtenue en éliminant la ligne i et la colonne j .
2. cofacteur d'ordre $(i; j)$ le scalaire : $(-1)^{i+j}|A_{ij}|$.
3. la matrice des cofacteurs de A est appelée comatrice de A . On la note $\text{com}(A)$.

Théorème 4. [5] Soit $A = [a_{ij}]_{i,j=1}^n \in M_n(K)$ une matrice carrée. Alors,

$$A \text{ inversible} \Leftrightarrow |A| \neq 0$$

Dans ce cas, la matrice inverse est donnée par la formule :

$$A^{-1} = \frac{1}{|A|}(\text{com}(A))^T$$

Chapitre 2

Méthode de Newton Raphson

2.1 Introduction

En analyse numérique, la méthode de Newton ou méthode de Newton-Raphson est une méthode de résolution de l'équation $f(x) = 0$ et, dans son application la plus simple, un algorithme efficace pour trouver numériquement une approximation précise d'un zéro (ou racine) d'une fonction réelle d'une variable réelle. Cette méthode doit son nom aux mathématiciens anglais Isaac Newton (1643-1727) et Joseph Raphson (peut-être 1648-1715), qui furent les premiers à la décrire pour la recherche des solutions d'une équation polynomiale. Thomas Simpson (1710-1761) élargit considérablement le domaine d'application de l'algorithme en montrant, grâce à la notion de dérivée, comment on pouvait l'utiliser pour calculer une solution d'une équation non linéaire, pouvant ne pas être un polynôme, et d'un système formé de telles équations.

2.2 Existence de solution

Théorème 5. *Soit $f : \Omega \subset \mathbb{R}^n$ un ensemble non-vidé et fermé et $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue. On suppose :*

- *Soit Ω est borné.*
- *Soit Ω est non borné et f est une fonction coercive.*

Alors il existe au moins un point minimum de f sur Ω (c'est à dire, $\exists x^ \in \Omega$ tel que $f(x^*) \leq f(x); \forall x \in \Omega$).*

Remarque 2.2.1. *Si Ω est un ensemble convexe fermé non vide de \mathbb{R}^n et si f est une fonction continue, coercive, strictement convexe sur Ω , alors le problème (1.2) a une solution et une seule.*

2.3 Unicité de solution

Si la fonction numérique $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est strictement convexe sur un ensemble convexe Ω et si f admet un point minimisant, ie :

$$f(a) = \inf_{x \in \Omega} f(x)$$

alors a est unique.

2.4 La méthode du Newton-Raphson

2.4.1 Algorithme

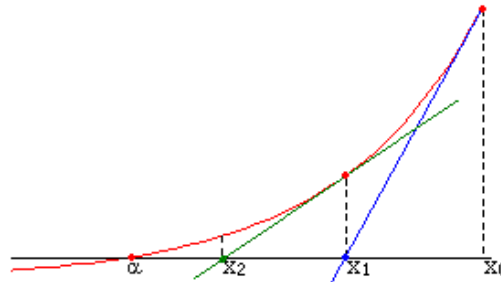


FIGURE 2.1 – La méthode de N-R

La méthode de Newton- Raphson [4][5][6] est une méthode forte qu'utilise pour résoudre des équations de la forme $f(x) = 0$. On notons que x une racine exacte cherche et que x_0 une valeur approchée de x . On suppose que $f(x)$ être développer en série de Taylor autour de x_0 , on a :

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2 f''(s)}{2!}, \quad \forall s \in]x, x_0[$$

Si $x < x_0$ et comme :

$$f(x) = 0$$

$$f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2 f''(s)}{2!} = 0$$

D'ou l'on calcule si $f'(x_0) \neq 0$

$$(x - x_0)f'(x_0) = -f(x_0) - \frac{(x - x_0)^2 f''(s)}{2!}$$

$$x - x_0 = -\frac{f(x_0)}{f'(x_0)} - \frac{f''(s)}{f'(x_0)} \frac{(x - x_0)^2}{2!}$$

Donc

$$x = x_0 - \frac{f(x_0)}{f'(x_0)} - \frac{f''(s)}{f'(x_0)} \frac{(x - x_0)^2}{2!}$$

En négligent le reste $-\frac{f''(s)}{2!} \frac{(x-x_0)^2}{2!}$ qui ne représente plus la racine exacte x , mais une racine approché x_1 par rapport à x_0 . En répétant le procédé on trouve la formule de de récurrence :

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad , n \in \mathbb{N}$$

que l'on appelle la formule de récurrence de Newton-Raphson (noté N-R).

Du point de vue géométrique, la méthode consiste en un extrapolation lineaire partant d'un point x_n proche de x en extrapole le long de tangente en $(x_n, f(x_n))$ jusqu'à son point d'intersection avec l'axe (ox) et ce point est appelé (x_{n+1})

Ayant : $tg(\alpha) = \frac{f(x_n)}{x_n - x_{n+1}} = f'(x_n)$

On trouve de nouveau :

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad , n \in \mathbb{N}$$

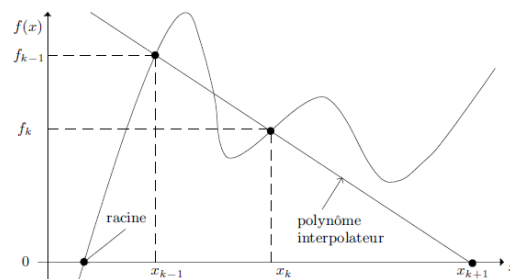


FIGURE 2.2 – La méthode de N-R

2.4.2 Méthode de N-R pour deux inconnues

On veut résoudre le système d'équation non lineaire :

$$\begin{cases} F(x, y) = 0 \\ G(x, y) = 0. \end{cases} \quad (2.1)$$

ou F et G sont des fonctions données des variables independantes x,y.

dans ce cas la méthode de N-R est analoguie celle qui est présent précédent (2.4.1). Aprer avoir choisi une solution approchée (x_0, y_0) c'est à dire une approximation initiale, on pose :

$$\begin{cases} x = x_0 + \varepsilon_0 \\ y = y_0 + \tau_0 \end{cases}$$

Et l'on développe en série de Taylor les deux fonctions F et G au point (x_0, y_0) puisque $(x_0 + \varepsilon_0, y_0 + \tau_0)$ est une solution de système d'équation (2.1). On résulte :

$$0 = F(x_0 + \varepsilon_0, y_0 + \tau_0) = F(x_0, y_0) + \varepsilon_0 \left(\frac{\partial F}{\partial x} \right) + \tau_0 \left(\frac{\partial F}{\partial y} \right) + \begin{matrix} \text{terme} \\ \text{d'ordre} \\ \text{superieur} \end{matrix}$$

$$0 = G(x_0 + \varepsilon_0, y_0 + \tau_0) = G(x_0, y_0) + \varepsilon_0 \left(\frac{\partial G}{\partial x} \right) + \tau_0 \left(\frac{\partial G}{\partial y} \right) + \begin{matrix} \text{terme} \\ \text{d'ordre} \\ \text{superieur} \end{matrix}$$

ou sous forme matricielle :

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix}_{(x_0, y_0)} + \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix}_{(x_0, y_0)} \begin{pmatrix} \varepsilon_0 \\ \tau_0 \end{pmatrix} + \begin{matrix} \text{terme} \\ \text{d'ordre} \\ \text{superieur} \end{matrix}$$

En supposant l'existence de l'inverse de la matrice (c'est à

dire : $\begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix}^{-1}$ existe). Donc

$$-\begin{pmatrix} F \\ G \end{pmatrix}_{(x_0, y_0)} = \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix}_{(x_0, y_0)} \begin{pmatrix} \varepsilon_0 \\ \tau_0 \end{pmatrix} + \begin{matrix} \text{terme} \\ \text{d'ordre} \\ \text{superieur} \end{matrix}$$

$$\begin{pmatrix} \varepsilon_0 \\ \tau_0 \end{pmatrix} = - \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix}_{(x_0, y_0)}^{-1} \begin{pmatrix} F \\ G \end{pmatrix}_{(x_0, y_0)} + \begin{matrix} \text{terme} \\ \text{d'ordre} \\ \text{superieur} \end{matrix}$$

On a :

$$\begin{pmatrix} \varepsilon_0 \\ \tau_0 \end{pmatrix} = \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = - \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix}_{(x_0, y_0)}^{-1} \begin{pmatrix} F \\ G \end{pmatrix}_{(x_0, y_0)} + \begin{matrix} \text{terme} \\ \text{d'ordre} \\ \text{superieur} \end{matrix}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix}_{(x_0, y_0)}^{-1} \begin{pmatrix} F \\ G \end{pmatrix}_{(x_0, y_0)} + \begin{matrix} \text{terme} \\ \text{d'ordre} \\ \text{superieur} \end{matrix}$$

L'orsqu'on néglige les termes d'ordre superieure, la solution(x,y) devient une approximation qui sera prise comme nouveau point de départ (x_1, y_1) de la prochaine approximation :

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix}_{(x_0, y_0)}^{-1} \begin{pmatrix} F \\ G \end{pmatrix}_{(x_0, y_0)}$$

Ainsi on trouve la formule de N-R pour deux inconnues :

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix}_{(x_n, y_n)}^{-1} \begin{pmatrix} F \\ G \end{pmatrix}_{(x_n, y_n)}$$

Remarque 2.4.1. Pour que la méthode de N-R soit applicable. Il faut que la fonction f vérifie :

N1 f une fonction continue de $[a, b]$.

N2 f soit monotone sur $[a, b]$.

N3 f est dérivable et $f'(x) \neq 0$.

N4 $f(a)f(b) < 0$.

N5 f est deux fois dérivable et f'' admet un min et un max sur $[a, b]$.

2.4.3 Convergence de méthode de N-R

Reprenons la méthode de N-R

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \in \mathbb{N}$$

Posons :

$$x_{n+1} = g(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$$

c'est à dire : $g(x) = x - \frac{f(x)}{f'(x)}$

Si $|g'(x)| < 1 \forall x \in v(x^*)$ avec x^* la racine de $f(x) = 0$ les conditions suivantes sont satisfaites (les conditions de théorème du point fixe $\varphi(x) = x$ telle que : φ définie sur $[a, b]$ dans \mathbb{R}) :

- φ est continue et dérivable sur $[a, b]$.
- φ prend ses valeurs dans $[a, b] (\forall x \in [a, b], \varphi(x) \in [a, b])$.
- $\exists \alpha \in]0, 1[; \forall x \in [a, b], |\varphi'(x)| \leq \alpha$, on dira que φ est une contraction stricte.

Ainsi on a convergence de la suite $(x_n)_n$ vers une racine x^* de l'équation $f(x) = 0$.

En cherchant x^* dans $[a, b]$ on a :

$$|g'(x)| = \left| \left(x - \frac{f(x)}{f'(x)} \right)' \right| = \left| \frac{f(x)f''(x)}{(f'(x))^2} \right| < 1$$

Supposons que :

$$\forall x \in [a, b], |f'(x)| \geq m > 0 \text{ et } |f''(x)| \leq M$$

On développant $f(x)$ autour de la solution x approchée x_n , on à :

$$f(x) = f(x_n) + (x - x_n)f'(x_n) + \frac{(x - x_n)^2 f''(s)}{2}$$

où $\varphi_n \in [x, x_n]$ et en divisant cette équation par $f'(x_n)$ ($f'(x_n) \neq 0$).

$$\frac{-f(x_n)}{f'(x_n)} = (x - x_n) + \frac{(x - x_n)^2 f''(s)}{2f'(x_n)}$$

Selons de la formule de N-R, on a : $\frac{-f(x_n)}{f'(x_n)} = x - x_n$, ainsi on arrive à :

$$x_{n+1} - x^* = \frac{(x_n - x^*)^2 f''(s)}{2f'(x_n)}$$

Et comme on à : $|f'(x)| \geq m > 0$ et $|f''(x)| \leq M$ on trouve l' inégalité :

$$|x_{n+1} - x^*| \leq \frac{M}{2m} |x_n - x^*|^2$$

On constate que l'erreur la $(n + 1)^{ieme}$ approximation. On dit que la méthode de N-R est itérative du deuxième ordre.

Si on pose : $\frac{M}{2m} = \alpha$, on a :

$$|x_{n+1} - x^*| \leq \alpha |x_n - x^*|^2$$

$$\alpha|x_n - x^*| \leq (\alpha|x_n - x^*|)^2 \quad (2.2)$$

$$\leq (\alpha|x_{n-1} - x^*|)^{2^2} \quad (2.3)$$

$$\vdots \quad (2.4)$$

$$\leq (\alpha|x_0 - x^*|)^{2^{n-1}} \quad (2.5)$$

Comme $|x_0 - x^*| \leq (b - a)$, on arrive à l'inégalité :

$$|x_n - x^*| < \frac{(\alpha(b - a))^{2^n}}{\alpha} \quad (2.6)$$

$$< \alpha^{2^n - n} (b - a)^{2^n} \quad (2.7)$$

Il y a donc convergence vers x^* , si $\frac{M}{2m}(b - a) < 1$ d'où il faut toujours choisir $[a, b]$ suffisamment petit c-à-d :

$$(b - a) < \frac{2m}{M}.$$

Remarque 2.4.2. *La méthode de Newton-Raphson peut ne pas converger vers une solution(2.3)*

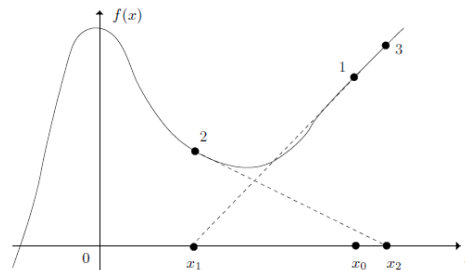


FIGURE 2.3 – Échec de la méthode de N-R

2.4.4 Vitese de convergence de la méthode de N-R

On a $\varphi(x) = x - \frac{f(x)}{f'(x)} = c$, d'où

$$\varphi'(x) = \frac{f(x)f''(x)}{(f'(x))^2}$$

Et donc au point c donne $\varphi'(c) = 0$

Par la formule de Taylor à l'ordre 2 : $\exists \mu \in [a, b]$

$$\begin{aligned}\varphi(x) &= \varphi(c) + (x - c)\varphi'(c) + \frac{1}{2}(x - c)^2\varphi''(\mu) \\ &= \varphi(c) + \frac{1}{2}(x - c)^2\varphi''(\mu) \\ &= c + \frac{1}{2}(x - c)^2\varphi''(\mu)\end{aligned}$$

$$\begin{aligned}|\varphi(x) - c| &= \frac{1}{2}(x - c)^2\varphi''(\mu) \\ &\leq \frac{1}{2}M(x - c)^2\end{aligned}$$

avec ($M = \max_{x \in [a, b]} |\varphi''(\mu)|$)

On applique cette majoration aux itérées x_n

$$\begin{aligned}\frac{1}{2}M|x_n - c| &\leq \left(\left| \frac{1}{2}M(x_{n-1} - c) \right| \right)^2 \\ &\leq \left(\left| \frac{1}{2}M(x_{n-2} - c) \right| \right)^{2^2} \\ &\leq \vdots \\ &\leq \left(\left| \frac{1}{2}M(x_0 - c) \right| \right)^{2^{n-1}} \\ &\leq \left(\left| \frac{1}{2}M(b - a) \right| \right)^{2^n}\end{aligned}$$

2.4.5 Test d'arrêt

Une fois construite la suite (x_n) converge vers le réel α vérifiant $g(\alpha) = \alpha$, et une fois fixée la tolérance ϵ , nous cherchons le premier entier n_0 vérifiant :

$$|x_{n_0+1} - x_{n_0}| < \epsilon$$

Si on note $e_n = x_n - \alpha$ l'erreur à l'itération n , on a :

$$e_{n+1} = x_{n+1} - \alpha = g(x_n) - g(\alpha) = g'(c_n)e_n$$

avec c_n un réel entre x_n et α donné par le théorème des accroissements finis. Par conséquent,

$$x_{n+1} - x_n = (x_{n+1} - \alpha) - (x_n - \alpha) = e_{n+1} - e_n = (g'(c_n) - 1)e_n$$

Or si n est suffisamment grand,

$$g'(c_n) \approx g'(\alpha) = 0$$

et donc :

$$e_n \approx x_{n+1} - x_n$$

L'erreur qu'on commet lorsque l'on adopte ce critère est donc plus petite que la tolérance ϵ fixée

$$|x_{n+1} - x_n| = |g'(c_n) - 1||x_n - \alpha|$$

Application numérique

Voici l'organigramme correspondant à notre résolution

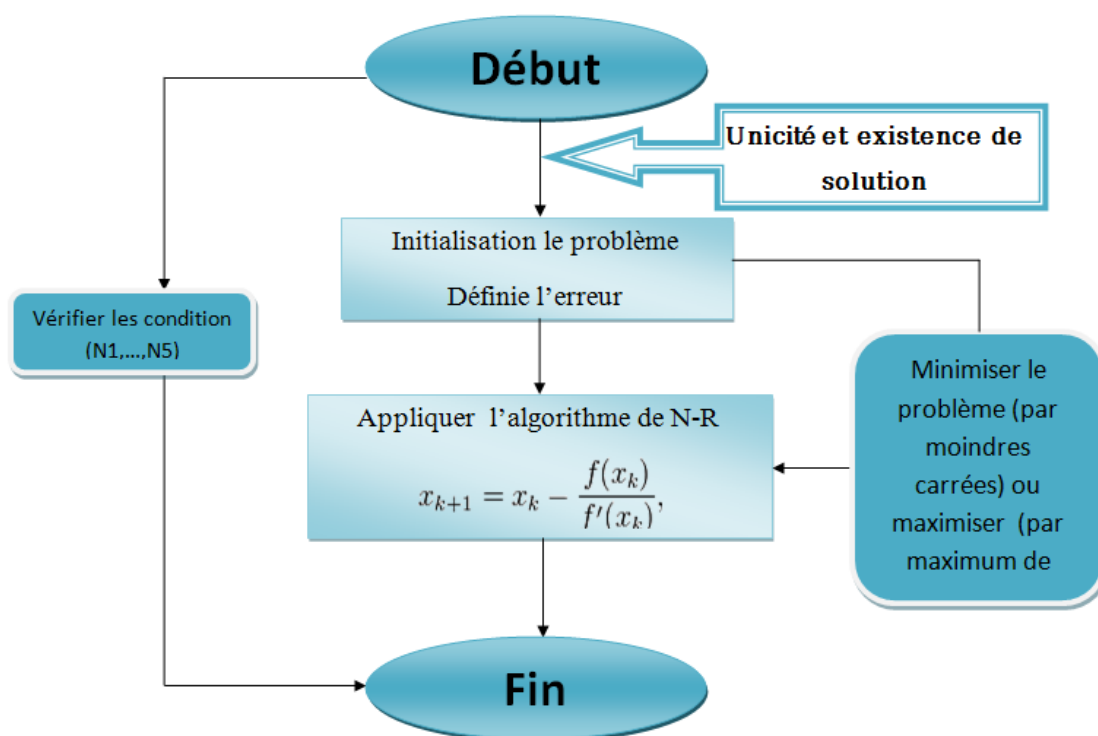


FIGURE 2.4 – L'organigramme de résolution

Donc on va programmer cette méthode itérative sous Mat-Lab

```

% Optimisation : Newton-Raphson
close all;
clear all;
clc;
syms x;
fun=input('Entrer f(x): \n'); %Demander f(x).
f=inline(fun);
m=diff(f(x),1); %Dérivée de la fonction f(x).
fl=inline(m);
t=diff(f(x),2); %Dérivée deuxième de la fonction f(x).
f2=inline(t);
x=input('Entrer la valeur initiale: \n'); % Demander x0.
xx(1)=x;
yy(1)=x;
x0=x;
a=input('Entrer la valeur de borne infieur de l intervale: \n');
b=input('Entrer la valeur de borne superieur de l intervale : \n');
err=input('Préciser l erreur: \n'); % Demander l'erreur.
for i=1:100 %* cette boucle est pour trouver un mirnum.
x=x-(fl(x)/f2(x));
j=i+1;
xx(j)=x; % stocker la valeur de x pour eviter l'écrasage de la nouvelle
% valeur quand la boucle roule et donc utiliser
% cette valeur pour mettre a jour le test d'arret.
ii=i;
Err=abs(xx(j)-xx(j-1)); %
if Err<err % test d'arret.
break;
end
end
FF=f(x);
disp(['valeur optimal de: ' num2str(x) ',est' num2str(FF) ', le nombre d iteration est:' num2str(ii) ])
for i=1:100 %* cette boucle est pour trouver la racine de fonction.
y=x0-(f(x0)/fl(x0));
j=i+1;
yy(j)=y; % stocker la valeur de y pour eviter l'écrasage de la nouvelle
% valeur quand la boucle roule et donc utiliser
% cette valeur pour mettre a jour le test d'arret.
ii=i;
Err=abs(yy(j)-yy(j-1));
if Err<err % test d'arret.
break;
end
end
KK=f(y);
disp(['x= ' num2str(y) ', f(x)= ' num2str(KK) ', le nombre d iteration est:' num2str(ii) ])
x=a:0.1:b;
F=f(x);
plot(x,F)
title('Optimisation:Newton-Raphson')
hold on
% plot( 1.4276 , 1.7757 , 'o ') % pour sélectionner le minimum de f(x).
% plot( 2.7721 , -0.046068 , '*') % pour sélectionner le racine de f(x).
grid on
xlabel('x')
ylabel('y')
hold off

```

FIGURE 2.5 – Programme de résolution en MatLab

Exemple :

On va appliquer cette méthode pour la recherche des racines de la fonction non linéaire suivante :

$$f(x) = 2\sin(x) - (0,1)x^2$$

dans $x \in [0, 10]$

le résultat de l'exécution avec l'erreur=0.01 et la valeur initiale $x_0 = 2.5$

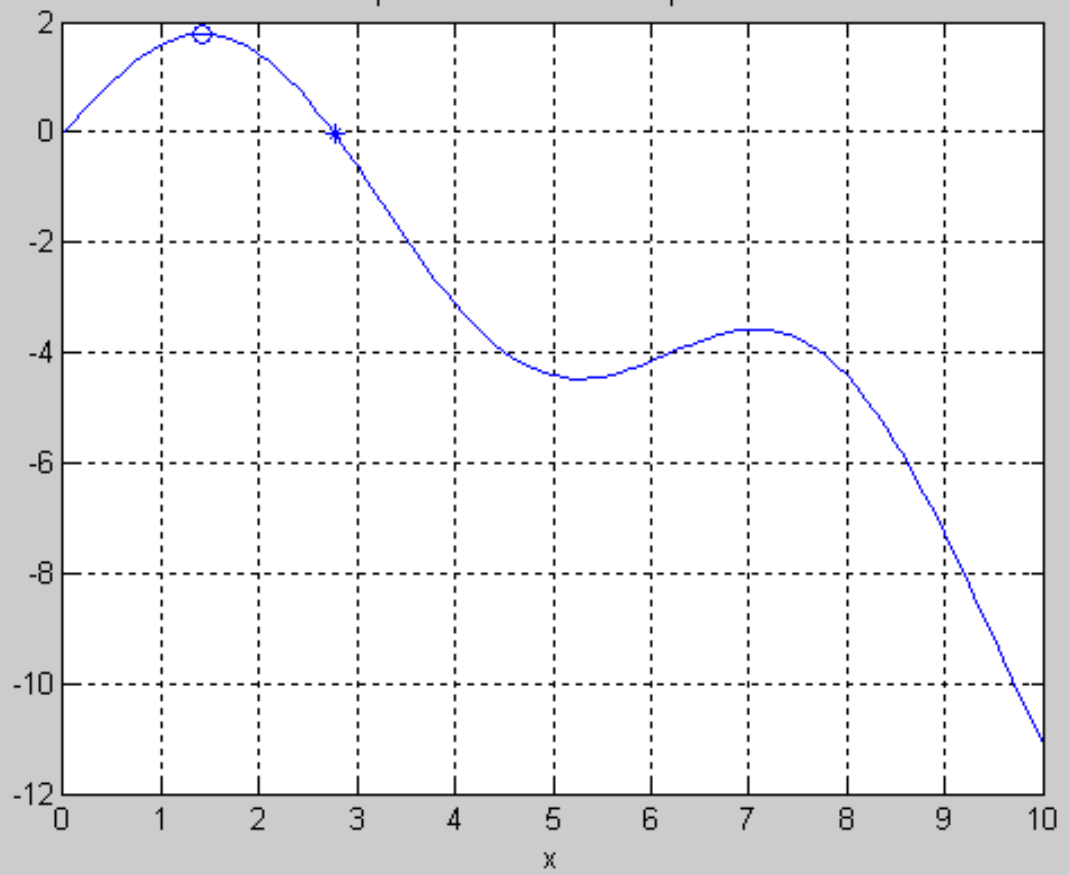
```
Entrer f(x):
2.*sin(x)-0.1.*x.^2
Entrer la valeur initiale:
2.5

x0 =

    2.5000

Entrer la valeur de borne infieur de l intervale:
0
Entrer la valeur de borne superieur de l intervale :
10
Préciser l erreur:
0.01
valeur optimal de: 1.4276,est: 1.7757 , le nombre d iteration est: 4
x= 2.7721, f(x)= -0.046068 , le nombre d iteration est: 2
>>
```

Optimisation: Newton-Raphson



Conclusion

Se trouve dans cette mémoire la méthode numérique de Newton -Raphson utilisée pour la résolution des problèmes d'optimisation. Le but de cette mémoire est de décrire l'algorithme utilisé pour résoudre des équations non linéaires du type :

$$f(x) = 0$$

La méthode de Newton -Raphson est la plus ancienne et peut être la base des méthodes numériques d'approximation des équations non linéaires, y'a beaucoup de groupes de recherche à travers le monde qui travaillent dans cette méthode riche. Alors l'apprentissage de cette méthode va permettre l'étudiant de comprendre la base mathématique des méthodes numériques.

L'avantage de cette technique est converger vite à condition de faire un bon choix du point de départ et ça demande des efforts de recherche supplémentaire, au contraire déjà comme tu dis faut que la fonction soit dérivable et en plus si on choisit un point de départ trop éloigné du zéro qu'on cherche bah si la fonction se comporte bizarrement autour du point choisi au départ la méthode va converger super lentement, voire pas converger du tout.



REFERENCES BIBLIOGRAPHIQUES

- [1] Allaire, G. : Analyse numérique et optimisation : Une introduction à la modélisation mathématique et à la simulation numérique. École polytechnique 2004 ou Éditions Ellipses, Paris, 2005.
- [2] Chokri, Bekkey, Zouhaire, Helali : Résolution numérique de l'équation $f(x)=0$,
- [3] D. Brewster. : Memoirs of the Life, Writings, and Discoveries of Sir Isaac Newton : Volume 1, Dunod 17 avril 2001.
- [4] Eric Walter. :Méthodes numériques et optimisation, un guide du consommateur, Fluctuant Nec Merguntur, Paris, 2015.
- [5] Jean-Pierre Demailly : Analyse numérique et équations différentielles, éditions des Presses Universitaires de Grenoble, 1996.
- [6] J.P. Ramis, A. Warusfel. : Mathématiques, Tout-en-un pour la licence, Dunod Paris.
- [7] J.Glieck. : Isaac Newton , Un destin fabuleux, Dunod 19 septembre 2005.