

Analyse Numérique 1

Cours, 2^{ième} Année Statistique et Analyse des Données.

Dr. Oulia Bouhoufani

13 décembre 2022

Table des matières

1	Notions de base sur les erreurs	4
1	Introduction	4
2	Sources d'erreurs	4
3	Erreur absolue et erreur relative	5
4	Propagation des erreurs	6
5	Représentation décimale des nombres	7
5.1	Chiffre significatif (c.s)	8
5.2	Chiffre significatif exact (c.s.e)	8
5.3	Troncature et arrondissement d'un nombre	9
6	Exercices	11
2	Résolution numérique d'équations non linéaires	13
1	Résolution numérique d'une équation non linéaire	13
1.1	Introduction	13
1.2	Approximation des racines	13
1.3	Méthode de dichotomie (ou de la bisection)	15
1.4	Méthode des approximations successives	17
1.5	Méthode de Newton-Raphson (ou de la tangente)	19
2	Résolution de système d'équations non linéaires	21
2.1	Méthode du point fixe	21
2.2	Méthode de Newton	22
3	Exercices	23
3	Interpolation polynomiale	25
1	Introduction	25
1.1	Le Problème d'interpolation polynomiale	25
1.2	Existence et unicité de l'interpolant	26
2	Interpolation de Lagrange	27
2.1	Base et polynôme de Lagrange	27
2.2	Exemple et remarques importantes	28
3	Interpolation de Newton	29
3.1	Les différences divisées : "Définitions et propriétés"	29
3.2	Interpolation de Newton en fonction des différences divisées	30
3.3	Cas particulier : "Des pôles équidistants"	33
4	Erreur d'interpolation	34
4.1	Erreur commise (effective)	34

TABLE DES MATIÈRES

4.2	Erreur théorique	34
5	Interpolation optimale "au points de Tchebychev"	36
5.1	Interpolation de Tchebychev dans l'intervalle $[-1,1]$	36
5.2	Interpolation de Tcheby dans un intervalle quelconque de \mathbb{R}	37
6	Exercices	38
4	Approximation au sens des moindres carrés	41
1	Notions préliminaires	41
2	Meilleure approximation dans un espace normé	43
2.1	Formulation du problème	43
2.2	Existence et unicité de la meilleure approximation	43
2.3	Construction de la M.A.S.M.C	43
3	Meilleure approximation d'une fonction au sens des moindres carrés	45
3.1	Approximation polynomiale au sens des M.C "le cas continu"	46
3.2	Approximation polynomiale au sens des M.C "le cas discret"	48
4	Exercices	50
5	Intégration numérique	51
1	Introduction	51
2	Méthodes d'intégration numérique	53
2.1	Méthode de Newton-Côtes	54
2.2	Méthode du trapèze	54
2.3	Méthode des trapèzes (trapèze généralisée)	55
2.4	Méthode de Simpson	56
2.5	Méthode de Simpson généralisée (ou composée)	56
3	Exercices	58
6	Dérivation numérique	59
1	Introduction	59
2	Approximation de la première dérivée	60
2.1	Formules à deux points	60
2.2	Formules à trois points	62
3	Approximation des dérivées d'ordre supérieur	64
4	Exercices	65
	Références	66

Introduction générale

L'Analyse numérique est une discipline à l'interface des mathématiques et de l'informatique. Elle s'intéresse tant aux fondements qu'à la mise en pratique des méthodes permettant de déterminer des solutions approchées à des problèmes d'analyse mathématique.

Plus formellement, l'analyse numérique est l'étude des algorithmes permettant de résoudre numériquement par discrétisation les problèmes de mathématiques continues. Ces algorithmes sont des formulations de certains problèmes issus des sciences physiques, biologiques, économiques, d'ingénierie, ..., etc.

Ce cours est destiné aux étudiants de 2^{ème} année S.A.D. Il a comme objectif de présenter aux étudiants une variété d'outils numériques permettant la résolution effective de certain nombre de problèmes.

Vu le volume horaire consacré à ce module (1 cours et 1 TD) par semaine!, de plus les étudiants ne sont pas de la spécialité mathématiques, tout cela ne nous permet pas de donner tous les détails théoriques concernant les méthodes présentées, alors nous avons volontairement omis la plupart des démonstrations. Le lecteur intéressé devra donc se référer à la littérature pour en savoir plus. En effet, on a consacré nos efforts beaucoup plus sur les techniques de résolution de ces méthodes et aussi à donner des exemples d'applications, pour permettre aux étudiants d'implémenter (sur ordinateur) les méthodes étudiées avec leurs enseignants d'informatique au cours des séances de TP.

Les sujets traités dans ce cours sont :

- Notions de base sur les erreurs
- Résolution numérique d'équations non linéaires
- Interpolation Polynomiale
- Approximation au sens des moindres carrés.
- Intégration Numérique
- Dérivation Numérique

Chapitre 1

Notions de base sur les erreurs

1 Introduction

En mathématiques, l'**analyse des erreurs** est l'étude du type et de quantité d'erreurs, ou d'incertitudes, qui peuvent être présentes dans la solution d'un problème. Cette étude est particulièrement importante dans les domaines appliqués tels que *l'analyse numérique* et les *statistiques*.

Il peut s'agir de l'effet combiné de plusieurs types d'erreurs dans un calcul ; à savoir les erreurs causés par la précision finie des calculs impliquant des valeurs à virgule flottante ou entière, ou l'erreur appelée "erreur de troncature" obtenue soit en jetant (ignorant) les bits les moins significatifs d'une opération arithmétique, soit en utilisant des simplifications qui impliquent, généralement, la troncature d'un développement en série infinie (c.à.d, prendre un nombre fini de ses termes), afin de rendre le calcul possible et pratique.

Ce chapitre devrait donc permettre au lecteur de mieux gérer ces erreurs au sein des processus numériques, afin d'être en mesure de mieux interpréter les résultats obtenus.

2 Sources d'erreurs

Les sources d'erreurs sont innombrables. Elles sont responsables de l'inexactitude de l'approximation obtenue par une méthode numérique. Certaines erreurs sont là avant même de commencer le calcul. Elles peuvent provenir de sources différentes :

- **Erreurs de mesure (ou de données)** : Les quantités mesurées sont toujours approchées. Cela est dû aux erreurs de lecture aux instruments de mesures qui se dérèglent avec le temps, et également aux limitations des instruments (techniques). De plus, les valeurs sont données en précision finie (nombre de chiffres significatifs fini).
- **Erreur au démarrage (due aux calculs précédents)** : Les données initiales d'un calcul peuvent provenir des résultats, d'un calcul précédent, qui eux mêmes sont inexacts. Il y a donc une erreur avant même de démarrer les calculs.
- **Erreurs de troncature (ou d'itération)** : C'est les erreurs associées aux processus infinis en analyse mathématique (par exemple, une série infinie où on doit tronquer cette dernière en prenant un nombre fini de ses termes).

- **Erreurs d'arrondi** : Ces erreurs sont associées au système de numération. Elles sont dues au fait que les nombres ne peuvent être stockés, sur un ordinateur, qu'avec un nombre fini de chiffres. Il en est de même pour les opérations qui introduisent des erreurs supplémentaires. Par exemple, " $\frac{4}{3}$ " sera stocké sur la machine avec la valeur "1.333,333 qui ne comporte que 7 chiffres significatifs correctes, alors que la valeur exacte comporte une décimale infinie.

3 Erreur absolue et erreur relative

Soit x un nombre exact et soit x^* un nombre approché de x (x^* est légèrement différent de x).

- Si $x^* < x$, on dit que x^* est une "valeur approchée par défaut" de x .
- Si $x^* > x$, alors x^* est dit "valeur approchée par excès" de x .

On note, généralement,

$$x^* \simeq x \text{ ou } x^* \approx x.$$

Exemple 3.1 $5/3 \simeq 1,6666$; $\sqrt{2} \simeq 1,42$; $\pi \simeq 3,14$, sont des approximations par défaut, alors que $\exp \simeq 2,7$ est une approximation par excès.

Définition 3.2 (Erreur absolue)

Soit x un nombre exact. On appelle erreur absolue d'un nombre x^* approché de x , tout **majo-**
rant de la valeur absolue de $x - x^*$, et qu'on note Δx :

$$|x - x^*| \leq \Delta x \text{ (3.1)}$$

ou d'une manière équivalente,

$$x^* - \Delta x \leq x \leq x^* + \Delta x$$

et on écrit

$$x = x^* \pm \Delta x$$

qui veut dire que

$$x \in [x^* - \Delta x, x^* + \Delta x].$$

Exemple 3.3 Calculer l'erreur absolue de $\pi \simeq 3,14$, où

$$3,14 \leq \pi \leq 3,15.$$

Corrigé. Ici, on a $|\pi - 3,14| \leq 0,01$. Donc, on peut poser $\Delta x = 0,01$.

Aussi, on a : $3,138 \leq \pi \leq 3,142$, d'où $|\pi - 3,14| \leq 0,002$. Pour cette 2ème approximation, on peut poser $\Delta x = 0,002$.

Evidemment $0,002$ est une meilleure majoration de $|\pi - 3,14|$ que $0,01$.

Remarque 3.4 1. Plus l'erreur absolue est petite, plus x^* est précis.

2. Si le nombre exact 'x' est connu, alors l'erreur absolue Δx est estimée par (3.1).

Définition 3.5 (Erreur Relative)

On appelle erreur relative d'un nombre approché x^* , qu'on note δx , le rapport suivant (lorsque la valeur exacte x est connue) :

$$\delta x = \frac{\Delta x}{|x|}, \quad x \neq 0.$$

- L'erreur relative est souvent exprimée en pourcentage.

Exemple 3.6 Pour la valeur $x = \frac{20}{3}$, la valeur approchée $x_1^* = 6.666667$ est 100 fois plus précise que la valeur approchée $x_2^* = 6.6667$. En effet, on a

$$\Delta_1 x \geq |x - x_1^*| = \left| \frac{20}{3} - 6.666667 \right| = \left| \frac{20}{3} - \frac{6666667}{10^6} \right| = \frac{1}{3} 10^{-6}$$

et

$$\Delta_2 x \geq |x - x_2^*| = \left| \frac{20}{3} - 6.6667 \right| = \left| \frac{20}{3} - \frac{66667}{10^4} \right| = \frac{1}{3} 10^{-4}.$$

Pour l'erreurs relatives correspondantes, on a

$$\delta_1 x = \frac{\Delta_1 x}{|x|} \geq \frac{1}{20} 10^{-6} = 50 \times 10^{-3} \%$$

alors que

$$\delta_2 x = \frac{\Delta_2 x}{|x|} \geq \frac{1}{20} 10^{-4} = 0.5 \times 10^{-3} \%.$$

Donc, la valeur x_2^* est plus précise que x_1^* .

Remarque 3.7 En pratique et puisque $x^* \approx x$, on peut prendre comme majorant de l'erreur relative,

$$\delta x \simeq \frac{\Delta x}{|x^*|}.$$

Par suite,

$$x = x^* \pm |x^*| \delta x.$$

4 Propagation des erreurs

Soient x et y deux quantités exactes et x^*, y^* deux valeurs approchées de x, y , respectivement. On note les erreurs absolues et les erreurs relatives associées à x et y par $\Delta x, \Delta y$ et $\delta x, \delta y$, respectivement. On a les propriétés suivantes.

Erreur de l'Addition :

1. $\Delta(x + y) = \Delta x + \Delta y$.
2. $\delta(x + y) \leq \max\{\delta x, \delta y\}$.

Erreur de la Soustraction :

1. $\Delta(x - y) = \Delta x + \Delta y$.
2. $\delta(x - y) \leq \frac{x^* + y^*}{x^* - y^*} \max\{\delta x, \delta y\}$.

Exemple 4.1 Soient $x^* = 255$ et $y^* = 250$ deux valeurs approchées de x et y , respectivement, et soit $\delta x = \delta y = 10^{-3}$. Calculer $\delta(x - y)$?

Corrigé : Tout d'abord, on a

$$\Delta x = |x^*| \delta x = 255 \times 10^{-3} = 0.255,$$

$$\Delta y = |y^*| \delta y = 250 \times 10^{-3} = 0.250.$$

Ainsi, l'erreur relative est :

$$\delta(x - y) = \frac{\Delta(x - y)}{|x^* - y^*|} = \frac{\Delta x + \Delta y}{5} = 10.1 \times 10^{-2} = 10.1\%.$$

* On remarque que x^* et y^* sont 101 fois plus précis pour x et y , respectivement, que $x^* - y^*$ l'est pour $x - y$.

Erreur de la Multiplication :

1. $\Delta(xy) = |x^*| \Delta y + |y^*| \Delta x.$
2. $\delta(xy) = \delta x + \delta y.$

Erreur de la Division :

1. $\Delta\left(\frac{x}{y}\right) = \frac{|x^*| \Delta y + |y^*| \Delta x}{y^{*2}}$
2. $\delta\left(\frac{x}{y}\right) = \delta x + \delta y.$

Exemple 4.2 *Quelle est la masse volumique d'un objet dont la masse est de $109,47 \text{ g} \pm 0,05 \text{ g}$ et le volume est $12,3 \text{ ml} \pm 0,3 \text{ ml}$?*

Corrigé : Pour la masse volumique, on a $m_v = 109,47 \div 12,3 = 8,90 \text{ g/ml}$.

Pour trouver les incertitudes, il faut utiliser les incertitudes relatives :

$$\Delta m_v = |m_v| \times \delta m_v = 8,9 \times \left[\left(\frac{0,05 \text{ g}}{109,47 \text{ g}} \right) + \left(\frac{0,3 \text{ ml}}{12,3 \text{ ml}} \right) \right] = 0,2.$$

La mesure finale est donc $m_v = (8,9 \pm 0,2) \text{ g/ml}$.

Erreur de la Puissance :

1. $\Delta(x^n) = n |x^{n-1}| \Delta x.$
2. $\delta(x^n) = n \delta x.$

5 Représentation décimale des nombres

Tout nombre réel positif x peut avoir la représentation décimale suivante :

$$x = a_m \times 10^m + a_{m-1} \times 10^{m-1} + \dots + a_{m-n+1} \times 10^{m-n+1} + \dots \quad (5.1)$$

où $a_m \in \mathbb{N}^*$, $a_i \in \{0, 1, 2, \dots, 9\}$ pour $i \neq m$ et le nombre entier ' m ' est le rang supérieur de x (la plus grande puissance de 10).

- Le développement dans (5.1) peut être aussi limité.

Exemple 5.1 *Les deux nombres $\pi = 3.141592\dots$ et 2107.3040 ont les représentations décimales suivantes :*

$$\pi = 3.141592\dots = 3 \times 10^0 + 1 \times 10^{-1} + 4 \times 10^{-2} + 1 \times 10^{-3} + 5 \times 10^{-4} + 9 \times 10^{-5} + 2 \times 10^{-6} + \dots$$

$$2107.3040 = 2 \times 10^3 + 1 \times 10^2 + 0 \times 10^1 + 7 \times 10^0 + 3 \times 10^{-1} + 0 \times 10^{-2} + 4 \times 10^{-3} + 0 \times 10^{-4}.$$

5.1 Chiffre significatif (c.s)

Définition 5.2 On appelle chiffre significatif d'un nombre approché tout chiffre 'a_i' différent du zéro, dans sa représentation décimale.

* Le zéro peut être aussi considéré comme chiffre significatif s'il est "conservé", c.à.d, s'il se trouve entre deux autres chiffres significatifs différents de zéro.

Exemple 5.3 Une approximation à 6 décimales de 0.002070 est comme suit :

$$\underbrace{0.00} \underbrace{2} \underbrace{0} \underbrace{7} \underbrace{0}$$

(a) (b) (c)

(a) : Ne sont pas des chiffres significatifs, car ils ne servent qu'à indiquer les rangs des autres chiffres.

(b) : Etant placé entre les chiffres significatifs 2 et 7, zéro est lui même significatif.

(c) : Ce zéro est significatif du fait que le nombre approché a conservé la décimale du rang 10⁻⁶.

Remarque 5.4 On ne peut pas connaître le nombre de chiffres significatifs du nombre $x = 38100$ écrit sous cette forme. Pour le connaître, il faut avoir, soit sa représentation décimale, ou encore -d'une façon équivalente- son écriture sous la forme : $x, yz... \times 10^s$.

En effet ;

$x = x_1 = 3,81 \times 10^4$ (x_1 à 3 c.s) ou encore $x = x_2 = 3,8100 \times 10^4$ (x_2 à 5 c.s).

5.2 Chiffre significatif exact (c.s.e)

Définition 5.5 On dit que le n^{ième} chiffre significatif après la virgule, d'un nombre approché de x , est exact si

$$\Delta x \leq 0.5 \times 10^{-n}.$$

Exemple 5.6 Soient $x = 2/3$ et $x^* = 0.6671$. Nous avons,

$$\Delta x = \left| \frac{2}{3} - 0.6671 \right| = \left| \frac{20000 - 20013}{3 \times 10^4} \right| = \frac{13}{3} \times 10^{-4} = 0.4333... \times 10^{-3} \leq 0.5 \times 10^{-3}.$$

Donc, les chiffres significatifs 7;6 et 6 de x^* sont exacts.

Propriétés

1. Si un chiffre significatif est exact, alors tous les chiffres à sa gauche sont exacts.
2. Si un chiffre n'est pas exact, alors tous les chiffres à sa droite ne le sont pas.
3. Si $\Delta x \leq 0.5 \times 10^{m-n+1}$, alors le nombre x possède au moins n chiffres exacts, où m est le plus grand rang qui se trouve dans sa représentation décimale.

Proposition 5.7 Si un nombre approché x^* de x possède n chiffres significatifs "c.s" exacts, alors

$$\delta x \leq 5 \times 10^{-n}.$$

5.3 Troncature et arrondissement d'un nombre

Pour approximer $\sqrt{2} = 1.414213562\dots$, on peut considérer la valeur approchée 1.41 ou encore 1.41421, etc... et cela selon le besoin.

- Dans le premier cas, on dit qu'on a tronqué (couper en éliminant une partie) le nombre approché de $\sqrt{2}$, après 2 décimales.
- Dans le second cas, on l'a tronqué après 5 décimales.

Une méthode habituelle pour tronquer un nombre pour ne garder qu'un nombre fini de chiffres significatifs est de l'arrondir.

Règle d'arrondissement

Pour arrondir un nombre jusqu'à n chiffres significatifs, il faut éliminer les chiffres à droite du $n^{\text{ième}}$ chiffre significatif conservé, si on se trouve après la virgule, sinon, on remplace par des zéros. Dans le premier cas, il faut respecter les règles suivantes :

- Si le $(n + 1)^{\text{ième}}$ chiffre significatif est > 5 , on ajoute la valeur "1" au $n^{\text{ième}}$ c.s.
- Si le $(n + 1)^{\text{ième}}$ chiffre significatif est < 5 , le $n^{\text{ième}}$ chiffre significatif (c.s) reste inchangé.
- Si le $(n + 1)^{\text{ième}}$ chiffre significatif = 5, alors deux cas sont possibles :
 1. Tous les chiffres rejetés (situés à droite du $(n + 1)^{\text{ième}}$ c.s) sont des zéros : On applique "la règle du chiffre pair", i.e, le $n^{\text{ième}}$ c.s reste inchangé s'il est pair, et s'il est impair, on lui ajoute 1.
 2. Parmi les chiffres rejetés (situés à droite du $(n + 1)^{\text{ième}}$ c.s), il existe au moins un chiffre qui soit non nul : On ajoute 1 au $n^{\text{ième}}$ chiffre significatif.

Remarque 5.8 1. On n'arrondit que le résultat final (jamais les résultats intermédiaires).
 2. Un nombre correctement arrondi ne possède que des chiffres significatifs exacts.

Exemple 5.9 1. Arrondir $x = 0.7897$ à 3 chiffres significatifs.
 Puisque 7 est plus grand que 5, alors $x \simeq x^* = 0.790$.

2. Arrondir $x = 0.753$ à 2 chiffres significatifs.
 Comme 3 est plus petit que 5, alors $x \simeq x^* = 0.75$.

3. Arrondir $x = 2.5456500$ à 5 chiffres significatifs.
 Puisque tous les chiffres rejetés sont des zéros et le 5^{ième} chiffre étant pair, alors $x \simeq x^* = 2.5456$.

4. Arrondir $x = 2.5453500$ à 5 chiffres significatifs.
 Ici, tous les chiffres rejetés sont des zéros et le 5^{ième} chiffre étant impair, alors $x \simeq x^* = 2.5454$.

5. Arrondir $x = 2.5305023$ à 4 chiffres significatifs.
 Puisque parmi les chiffres rejetés, il existe au moins un qui soit non nul, alors $x \simeq x^* = 2.531$.

Définition 5.10 On appelle "borne supérieure de l'erreur d'arrondi" d'un nombre approché, la valeur notée Δ_{arr} vérifiant :

$$\Delta_{arr} \leq 0.5 \times 10^{m-n+1},$$

où m est le rang supérieur de x (la plus grande puissance de 10) dans sa représentation décimale, et n est le nombre des chiffres significatifs exacts dans cette représentation.

Après l'arrondissement de la valeur approchée x^* , d'un nombre x , on aura :

$$x = x_{arr}^* \pm (\Delta x + \Delta_{arr}),$$

où x_{arr}^* est le nombre approché arrondi et, en général, $\Delta x \simeq \Delta_{arr}$. Par conséquent,

$$x = x_{arr}^* \pm 2\Delta_{arr}.$$

6 Exercices

"Notion d'erreurs"

Exercice 1.

Soit h la hauteur d'un barrage. Sa valeur mesurée h est $74,260m$. Si l'erreur relative commise sur h est $0,3$, trouver a, b tels que $a \leq H \leq b$.

Exercice 2.

Avec combien de c.s.e faut-il calculer $e^{\sqrt{2}}$, pour que l'erreur relative ne dépasse pas 1% ?

Exercice 3.

On désire approcher $I = \int_0^{\frac{1}{p}} f(t)dt$ où $f(t) = t^3$ et $p \in \mathbb{N}^*$, par la surface S du triangle de sommets $A(0, 0)$, $B(\frac{1}{p}, 0)$ et $C(\frac{1}{p}, f(\frac{1}{p}))$.

Déterminer la valeur minimale de p pour que l'erreur absolue d'approximation ne dépasse pas " $\varepsilon = 10^{-2}$," puis " $\varepsilon = 10^{-4}$ ".

Exercice 4. On considère l'équation :

$$x^2 - 1634x + 2 = 0. \tag{6.1}$$

1. Résoudre l'équation (1.1) en effectuant les calculs avec $N = 10$ chiffres significatifs. (Utiliser le discriminant)
2. Commenter le résultat obtenu et proposer une autre méthode de calcul pour contourner le problème posé.

Exercice 5. Soient les trois nombres réels : $x = 8,22$, $y = 0,00317$ et $z = 0,00432$.

1. Représenter les nombres x, y et z avec virgule flottante.
2. En effectuant les calculs avec 3 c.s, calculer la somme $x + y + z$, en faisant :
 - i) $(x + y) + z$;
 - ii) $x + (y + z)$.
3. Commenter les résultats obtenus. Conclure.

Exercice 6. On veut calculer la surface d'un disque $S = \pi R^2$ où $R = 2,3400$, $\pi = 3,1416$.

En admettant que tous les chiffres de R et π sont exacts,

1. Estimer les erreurs absolue et relative de S .
2. Calculer S en arrondissant au nombre de chiffres significatifs exacts.

Exercice 7.

1. Donner le nombre des chiffres significatifs exacts des nombres 0.0006814369 et 6.17924 , s'ils ont une erreur relative inférieure à 10^{-4} :
2. Si tous les chiffres significatifs de 3724.14 sont exacts. Quelle est son erreur relative ?
3. Si les chiffres significatifs de 476.6 et 3.11918 sont exacts. Quelle est l'erreur absolue (resp. relative) de la somme et combien a-t-elle de chiffres significatifs exacts ?
4. Calculer $\frac{35.4}{0.0042}$, si tous les chiffres significatifs du numérateur et dénominateur sont exacts.
5. Quelle est l'erreur relative (resp. absolue) du rapport ?
6. Arrondir au dernier chiffre significatif exact.

Exercice 8. Soit T la période des petites oscillations du pendule qui est donnée par $T = 2\pi\sqrt{l/g}$, où l est la longueur du pendule et g est la gravité. Supposons que les mesures faites sur T et l ont donné les résultats suivants : $T = \bar{T} \pm \Delta(T) = 1,9360,002s$ et $l = \bar{l} \pm \Delta(l) = 92,950,10cm$.

1. Calculer la gravité g en arrondissant au dernier chiffre significatif exact.
2. Donner l'erreur relative sur g en pourcentage.

Chapitre 2

Résolution numérique d'équations non linéaires

1 Résolution numérique d'une équation non linéaire

1.1 Introduction

Dans ce chapitre, on se propose d'approcher **numériquement** les racines réelles de f (ou les solutions) de l'équation

$$f(x) = 0 \quad \dots \quad (1.1)$$

avec $f : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction non linéaire supposée suffisamment régulière (c.à.d, **définie**, **continue**, et autant de fois **continument dérivable** que nécessaire) sur l'intervalle que nous allons aborder. L'équation (1.1) intervient notamment dans l'étude de plusieurs problèmes motivés par des applications pratiques, pour lesquels des solutions exactes de ce type d'équation ne sont pas explicitement connues. Le but de ce chapitre est de fournir, de manière détaillée, les principes de base des méthodes numériques ; les plus utilisées pour la résolution de l'équation (1.1).

1.2 Approximation des racines

Définition 1.1 *On appelle "racine approchée" de l'équation (1.1) tout point \bar{x} de \mathbb{R} vérifiant $|f(\bar{x})| \leq \varepsilon$, où ε est un réel positif assez petit qui représente la précision ou l'erreur permise pour cette approximation.*

** On peut aussi considérer comme racine approchée de (1.1) tout réel \bar{x} qui vérifie la condition $|\alpha - \bar{x}| \leq \varepsilon$, et cela dans le cas où la racine exacte α est connue.*

Séparation des racines

L'approximation des solutions de l'équation (1.1) nécessite la séparation de ses racines, plus précisément, il s'agit de déterminer des intervalles réels $[a_i, b_i]$ ($i = 1, 2, \dots$) à l'intérieur des quels la fonction f admet **une et une seule** racine.

Pour séparer les racines de l'équation $f(x) = 0$, dans un intervalle donné, il y a deux méthodes :

- Méthode graphique : dans cette méthode, soit on trace le graphe de f et on cherche son intersection avec l'axe (ox), ou bien on décompose f en deux fonctions f_1 et f_2 (simple à étudier), tel que $f \equiv f_1 - f_2$, puis on trace les deux graphes de f_1 et f_2 . Ainsi, les racines de f seront les points d'intersection des graphes de f_1 et f_2 , car pour tout $x \in \mathbb{R}$, on a

$$f(x) = 0 \Leftrightarrow f_1(x) - f_2(x) = 0 \Leftrightarrow f_1(x) = f_2(x).$$

- Méthode analytique : Ici, pour séparer les racines de f , on se base sur le **théorème des valeurs intermédiaires**, plus précisément, on se donne deux réels a et b , si $f(a).f(b) < 0$, alors f admet au moins une racine dans l'intervalle $]a, b[$, et si f est strictement monotone sur $]a, b[$, alors cette racine est unique.

Principe des méthodes numériques

Pour mettre les idées en place, toutes les méthodes que nous présentons ici sont itératives, c'est-à-dire qu'à partir d'une valeur initiale, elles construisent une suite de solutions approchées $(x_k)_k$ avec la propriété que

$$\lim_{k \rightarrow +\infty} x_k = \alpha.$$

Le choix de la méthode appropriée pour la résolution de l'équation (1.1) se base essentiellement sur ses conditions de convergence ainsi que sa vitesse.

Ordre de convergence

Avant de décrire quelques méthodes numériques permettant la résolution de l'équation (1.1), il est utile de définir l'ordre de convergence d'une méthode numérique. Cela permet de comparer son efficacité avec d'autres méthodes.

Définition 1.2 Soit $(x_n)_{n \geq 0}$ une suite de réels convergeant vers une limite α . L'ordre de convergence de la suite $(x_n)_{n \geq 0}$ (ou de la méthode itérative) est le plus grand réel $r \geq 1$ vérifiant

$$\lim_{k \rightarrow +\infty} \frac{|x_{k+1} - \alpha|}{|x_k - \alpha|^r} = l < +\infty.$$

En particulier,

- Si $r = 1$ et $l = 1$, alors la convergence est dite sous-linéaire.
- Si $r = 1$ et $l \in]0; 1[$, alors la convergence est dite linéaire.
- Si $r = 1$ et $l = 0$, la convergence est dite super-linéaire.
- Si $r = 2$, alors la convergence est dite quadratique.

Remarque 1.3 Plus l'ordre de convergence est plus élevé, plus la suite converge rapidement.

* Pour mieux illustrer la notion de vitesse de convergence d'une méthode, on donne l'exemple suivant.

Exemple 1.4 Considérons les suites, définies pour tout $n > 0$, par :

$$1) \quad x_n = 1/n ; 2) \quad y_n = 1/2^n \quad \text{et} \quad z_n = 1/2^{2^n}.$$

Il est clair que

$$\lim_{n \rightarrow +\infty} x_n = \lim_{n \rightarrow +\infty} y_n = \lim_{n \rightarrow +\infty} z_n = 0,$$

en plus, on a

1)

$$\lim_{n \rightarrow +\infty} \frac{|x_{n+1} - 0|}{|x_n - 0|} = \lim_{n \rightarrow +\infty} \frac{n}{n+1} = 1,$$

donc la convergence de $(x_n)_n$ est sous-linéaire.

2)

$$\lim_{k \rightarrow +\infty} \frac{|y_{k+1} - 0|}{|y_k - 0|} = \lim_{k \rightarrow +\infty} \frac{2^k}{2^{k+1}} = \frac{1}{2},$$

d'où la convergence de la suite $\{y_n\}$ est linéaire.

3)

$$\lim_{k \rightarrow +\infty} \frac{|z_{k+1} - 0|}{|z_k - 0|^2} = \lim_{k \rightarrow +\infty} \frac{2^{2^{k+1}}}{2^{2^k}} = 1,$$

donc la convergence de $\{z_k\}$ est quadratique.

Estimation d'erreurs

Chaque méthode itérative a sa propre formule pour calculer l'erreur commise à l'itération $k, k \geq 1$, mais ces formules sont généralement coûteuses.

Les deux formules suivantes :

$$1) \Delta_k = |f(x_k)|, \quad 2) \Delta_k = |x_k - x_{k-1}|$$

sont valables pour toutes les méthodes numériques, mais elles n'expriment pas réellement la vitesse de convergence de ces dernières. Dans tous nos exemples, on utilise la première formule car elle est plus signifiante, en terme de convergence, que la seconde.

Parmi les méthodes itératives les plus connues, on trouve les méthodes suivantes.

1.3 Méthode de dichotomie (ou de la bisection)

Commençons par la méthode la plus simple dite "méthode de dichotomie ou de bisection". Cette méthode est considérée comme étant un procédé systématique de raffinement de la localisation d'une racine. Le mot (dico = deux, tomie = coupe) exprime le principe de la méthode ; qui est la construction d'une suite d'intervalles de plus en plus petits contenant la racine localisée, en utilisant le théorème des valeurs intermédiaires à plusieurs reprises.

Reconsidérons notre équation (1.1), avec f définie et continue sur un intervalle borné $[a, b]$ de \mathbb{R} .

Description de la méthode

1. $i := 0$;
2. Si $f(a) * f(b) < 0$, alors il existe entre a et b au moins une racine de f ;
3. $x_i := \frac{a+b}{2}$; Si $f(x_i) = 0$ alors la racine de f est bien x_i stop FinSi₁ et sortir du programme FinSi₂;
4. Sinon subdiviser $[a, b]$ en deux sous-intervalles $[a, x_i] \cup [x_i, b]$;
5. chercher où se trouve la racine de f en testant les inégalités $f(a) * f(x_i)$ et $f(x_i) * f(b)$, par le théorème des valeurs intermédiaires ;
6. Si la racine est dans $[a, x_i]$, alors $b := x_i$, $i := i + 1$;

7. Si la racine est dans $[x_i, b]$, alors $a := x_i$, $i := i + 1$;
8. On itère le même processus jusqu'à ce qu'on localise la solution.

Test d'arrêt.

Comme l'algorithme est itératif, le test d'arrêt peut être :

- Soit $f(x_i) = 0$
- Soit, en déterminant une valeur approchée près de certain $\varepsilon > 0$, vérifiant l'inégalité

$$|x_k - x_{k-1}| < \varepsilon,$$

à l'itération $k > 0$

- Soit atteindre le nombre N d'itérations calculées à l'avance et donné par l'inégalité

$$N \geq \frac{\ln(b-a) - \ln(\varepsilon)}{\ln 2} - 1$$

- Soit les résultats numériques calculés se stabilisent d'une itération à une autre.

Remarque 1.5 * La convergence de la méthode de dichotomie est garantie dès qu'on détermine l'intervalle contenant la racine exacte.

* La dichotomie est la méthode la plus facile à implémenter sur machine, puisqu'elle nécessite seulement le calcul de f en plusieurs points.

* L'inconvénient majeur de cette méthode est sa convergence lente car elle est linéaire.

Exemple 1.6 Soit l'équation suivante

$$f(x) = x^3 + 2x - 1$$

1. Vérifier que cette équation admet une racine unique α dans $[0, 1]$.
2. Déterminer le nombre minimal d'itérations nécessaires pour calculer une racine approchée de α avec une précision $\varepsilon = 10^{-1}$, en utilisant la méthode de dichotomie, ensuite calculer cette racine approchée.

Corrigé :

1. Application directe du théorème (TVI) sur la fonction f et l'intervalle $[0, 1]$:
 - (a) La fonction f est un polynôme, donc elle est continue sur $[0, 1]$.
 - (b) $f(0)f(1) = -2 < 0$.
 - (c) $f'(x) = 3x^2 + 2 > 0$ sur $]0, 1[$. Ainsi, l'équation admet une racine unique $x^* \in]0, 1[$.
2. Calcul du nombre minimal d'itérations nécessaires pour avoir la précision $\varepsilon = 10^{-1}$. On a, pour tout $k \geq 0$,

$$|x^* - x_k| \leq \frac{b_k - a_k}{2} = \frac{b_{k-1} - a_{k-1}}{2^2} = \dots = \frac{b_0 - a_0}{2^{k+1}} = \frac{b - a}{2^{k+1}}.$$

En posant $\frac{b-a}{2^{k+1}} \leq \varepsilon$, on aura

$$k \geq \frac{\ln(b-a) - \ln(\varepsilon)}{\ln 2} - 1.$$

D'où

$$k \geq \frac{\ln(1-0) - \ln(10^{-1})}{\ln 2} - 1 \simeq 3.$$

3. Calcul de la racine approchée :

k	a_k	b_k	x_k	$f(a_k)f(b_k)$	$\Delta_k = \frac{b_k - a_k}{2}$
0	0	1	0.5	<	0.5
1	0	0.5	0.25	>	0.25
2	0.25	0.5	0.375	>	0.125
3	0.375	0.5	0.4345		0.0625

La racine approchée est $x \simeq 0.4375 \pm 10^{-1}$.

1.4 Méthode des approximations successives

Une des méthodes les plus importantes de résolution numérique des équations non linéaires est "la méthode des approximations successives" dite également "**méthode du point fixe**"; dont le principe est de remplacer l'équation donnée $f(x) = 0$ par une autre équation équivalente de la forme $g(x) = x$. La recherche des zéros (ou racines) de f se remène à la détermination des points fixes de la fonction g .

Définition 1.7 Soit $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction continue sur son domaine de définition qu'on note D_g . On dit que x^* (de D_g) est un point fixe de g ssi $g(x^*) = x^*$.

* Soit $(x_n)_n$ une suite de D_g définie par :

$$x_{n+1} = g(x_n), \quad n \geq 0.$$

Si $(x_n)_n$ converge vers \bar{x} , alors cette limite \bar{x} est un point fixe de g .
En effet, par continuité de g , on a

$$\begin{aligned} \lim_{n \rightarrow \infty} x_{n+1} &= \lim_{n \rightarrow \infty} g(x_n) \\ &= g\left(\lim_{n \rightarrow \infty} x_n\right). \end{aligned} \tag{1.2}$$

Or,

$$\lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} x_n = \bar{x},$$

ainsi (1.2) donne $\bar{x} = g(\bar{x})$. Donc, \bar{x} est un point fixe de g .

Exemple 1.8 L'équation $f(x) = x^2 - x - 3\ln(x) = 0$ peut se réécrire sous la forme $g(x) = x$, avec

1. $g(x) = x^2 - 3\ln(x)$,
2. $g(x) = \sqrt{x + 3\ln(x)}$,
3. $g(x) = \exp x(x - 1)/3$.

* Nous signalons qu'il faut choisir l'écriture qui assure la convergence du processus itératif $(x_{k+1} = g(x_k), k \geq 0)$ vers la solution cherchée, lorsque cette convergence est possible.

Théorème 1.9 ("**Existence et unicité du point fixe**")

i) Soit g une fonction réelle définie sur l'intervalle $[a, b]$ et satisfaisant aux conditions suivantes :

1. $g([a, b]) \subset [a, b]$,

2. Il existe $k \in \mathbb{R}$ ($0 < k < 1$) tel que, pour tout $x, y \in [a, b]$, on a

$$|g(x) - g(y)| \leq k |x - y|.$$

Alors, g admet un point fixe unique $x^* \in [a, b]$. De plus, pour tout $x_0 \in [a, b]$, la suite $(x_n)_n$ définie par : $x_{n+1} = g(x_n), \forall n \geq 0$, converge vers x^* .

ii) La formule suivante donne une majoration de l'erreur commise à l'itération k , lors du calcul de x^* ,

$$|x_k - x^*| \leq \frac{q^k}{1-q} |x_1 - x_0|.$$

iii) Le nombre d'itérations \tilde{k} nécessaires pour trouver un point fixe de g , avec une précision ε , est estimé par :

$$\tilde{k} \geq \frac{\log(\varepsilon) - \log(b-a)}{\log(q)}.$$

Remarque 1.10 Le nombre réel k , dans le théorème précédent, est appelé "rapport de **contraction**", ou bien la "constante de la fonction **contractante** g ".

Proposition 1.11 [1] Soit g une fonction dérivable sur l'intervalle $[a, b]$. Si la dérivée de g vérifie,

$$\max_{x \in [a, b]} |g'(x)| = k < 1,$$

alors g est contractante sur $[a, b]$ (c.à.d, elle vérifie la condition 2 (i) du Théorème 1.5).

Exemple 1.12 Résoudre par la méthode des approximations successives, l'équation :

$$xe^x - 1 = 0,$$

sachant que la solution se situe dans l'intervalle $[0, 1]$.

Corrigé :

1. L'algorithme de la méthode des approximations successives s'écrit :

$$x_{n+1} = g(x_n), n = 0, 1, \dots \text{ et } x_0 \text{ donné}$$

où g est une fonction continue et dérivable sur $[0, 1]$.

Pour que cette méthode "itérative" converge vers la solution, il suffit que $|g'(x)| < 1$, pour x voisin de la solution.

On peut réécrire l'équation donnée sous la forme :

$$x = g(x), \text{ où } g(x) = e^{-x},$$

et comme

$$|g'(x)| = e^{-x} < 1, \forall x \in]0, 1],$$

alors la convergence de la méthode est assurée.

2. En prenant comme point de départ $x_0 = 0$, on obtient la suite des approximations successives :

$x_1 = 1$	$x_8 = 0.560115461$	$x_{15} = 0.567276232$
$x_2 = 0.367879441$	$x_9 = 0.571143115$	$x_{16} = 0.567067898$
$x_3 = 0.692200627$	$x_{10} = 0.564879347$
$x_4 = 0.50004735$	$x_{11} = 0.568428725$	
$x_5 = 0.606243535$	$x_{12} = 0.566414733$	
$x_6 = 0.5453955786$	$x_{13} = 0.567556637$	
$x_7 = 0.579612335$	$x_{14} = 0.566908912$	

La solution est donc $x = 0.567$, si l'on s'arrête avec 3 chiffres après la virgule.

1.5 Méthode de Newton-Raphson (ou de la tangente)

Pour chercher les racines approchées de l'équation (1.1), beaucoup de méthodes de point fixe courantes font le choix de la forme $g(x) = x - u(x)f(x)$, avec u une fonction satisfaisant $0 < |u(x)| < +\infty$, sur le domaine de définition (ou plus généralement sur un intervalle contenant la racine exacte de f). Sous cette hypothèse, on vérifie facilement que toute racine de f est un point fixe de g et vice versa.

En supposant que la fonction f est de classe C^1 , la méthode de Newton-Raphson fait le choix

$$u(x) = \frac{1}{f'(x)}.$$

Alors, la relation de récurrence définissant cette méthode est :

$$x_{k+1} = g(x_k) = x_k - \frac{f(x_k)}{f'(x_k)}, k \geq 0 \dots (1.3)$$

En effet, en notant par x^* la racine de f (qu'on cherche) et par x_0 une valeur approchée de x^* , en supposant que f est de classe C^2 au voisinage de x^* , alors le développement de Taylor d'ordre 2 de f , au voisinage de x_0 , donnera :

$$f(x) \simeq f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(c)}{2!}(x - x_0)^2, x_0 < c < x^*.$$

Donc, pour $x = x^*$, il vient que

$$f(x^*) = 0 \simeq f(x_0) + \frac{f'(x_0)}{1!}(x^* - x_0) + \frac{f''(c)}{2!}(x^* - x_0)^2.$$

En supposant que $f'(x_0) \neq 0$, on aura

$$x^* - x_0 \simeq -\frac{f(x_0)}{f'(x_0)} - \frac{f''(c)}{2f'(x_0)}(x^* - x_0)^2.$$

Ainsi,

$$x^* \simeq x_0 - \frac{f(x_0)}{f'(x_0)} - \frac{f''(c)}{2f'(x_0)}(x^* - x_0)^2.$$

En négligeant le terme $-\frac{f''(c)}{2f'(x_0)}(x^* - x_0)^2$, on obtient une valeur approchée de x^* , qu'on note par x_1 , où $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$. En itérant le procédé, on trouve la "Formule de récurrence de Newton-Raphson" (1.3).

Interprétation géométriques

Dans la méthode de Newton-Raphson, toute nouvelle approximation de la racine est construite au moyen d'une linéarisation de l'équation $f(x) = 0$, autour de l'approximation précédente. En effet, si l'on remplace $f(x)$ au voisinage du point x_k par l'approximation obtenue par le développement de Taylor du premier ordre de f en x_k , on résout l'équation linéaire résultante :

$$f(x_k) + (x - x_k)f'(x_k) = 0,$$

et en notant sa solution x_{k+1} , on retrouve la relation (1.3). Egalement, il en résulte géométriquement que le point x_{k+1} est l'abscisse du point d'intersection entre la tangente à la courbe de f , au point $(x_k, f(x_k))$, et l'axe (ox) .

Théorème 1.13 *i) Soit f une fonction de classe $C^2([a, b])$ (où $[a, b]$ est un intervalle contenant une racine unique de f). Supposons que les conditions suivantes sont vérifiées :*

1. $f'(x) \neq 0, \forall x \in [a, b]$,
2. f'' garde un signe constant sur $[a, b]$.

Alors, la suite de la méthode de Newton-Raphson engendrée par la relation (1.3) converge vers la solution unique de l'équation $f(x) = 0$, en partant de l'approximation x_0 vérifiant

$$f(x_0).f''(x_0) > 0.$$

ii) L'erreur de cette méthode est estimée par la relation :

$$|x_{k+1} - x^*| \leq \frac{M}{2m} |x_k - x^*|^2,$$

où $m = \min_{a \leq x \leq b} |f'(x)|$, $M = \max_{a \leq x \leq b} |f''(x)|$.

Remarque 1.14 *D'après l'estimation précédente, l'erreur absolue de la $(k+1)^{ième}$ approximation (ou racine) est proportionnelle au carré de l'erreur absolue de la $(k)^{ième}$ approximation. Donc, la méthode de Newton-Raphson est une méthode itérative du deuxième ordre.*

Exemple 1.15 *Considérons l'équation non linéaire*

$$f(x) = x^3 - x - 4 = 0.$$

1. Montrer que cette équation admet une racine unique $x^* \in [1, 2]$.
2. Trouver une valeur approchée de x^* , en appliquant la méthode de Newton-Raphson jusqu'à le troisième itéré, puis calculer l'erreur commise à chaque itération.

Corrigé :

On vérifie facilement que toutes les conditions du Théorème des valeurs intermédiaires sont satisfaites par la fonction donnée " f ", sur l'intervalle $[1, 2]$.

1. Choix du point initial :

Comme $f'(x) = 3x^2 - 1$, $f''(x) = 6x$, alors $f(2).f''(2) = 24 > 0$. Ainsi, on peut prendre $x_0 = 2$ comme un point de départ dans le schéma itératif de la méthode de Newton-Raphson.

2. En utilisant la relation de récurrence (1.3), on obtient

$$x_1 = 1,8182; \quad \Delta_1 = |f(x_1)| = 0,1925 = 1,925 \times 10^{-1},$$

$$x_2 = 1,7966; \quad \Delta_2 = |f(x_2)| = 0,0024 = 2,4 \times 10^{-3},$$

$$x_3 = 1,7963; \quad \Delta_3 = |f(x_3)| = 0,00019 = 1,9 \times 10^{-4}.$$

On remarque que l'erreur commise à l'étape $k = 3$ est la plus petite erreur.

2 Résolution de système d'équations non linéaires

Dans cette section, on propose de résoudre le système de n équations, avec n inconnus x_1, x_2, \dots, x_n , de la forme

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0, \\ f_2(x_1, x_2, \dots, x_n) = 0, \\ \cdot \\ \cdot \\ \cdot \\ f_n(x_1, x_2, \dots, x_n) = 0. \end{cases} \quad (S_1)$$

Le système (S_1) peut toujours se réécrire sous la forme :

$$F(x) = 0; \quad x \in \mathbb{R}^n \text{ et } F : \mathbb{R}^n \longrightarrow \mathbb{R}^n,$$

tels que

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \text{ et } F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \cdot \\ \cdot \\ \cdot \\ f_n(x) \end{pmatrix},$$

avec $F(x) = Ax + b$.

2.1 Méthode du point fixe

Soit $x \in \mathbb{R}^n$ et soit $\omega \in \mathbb{R}^*$. On définit la fonction g_ω par :

$$g_\omega(x) = x + \omega f(x),$$

où $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$.

Remarque : il est bien clair que $f(x) = x \Leftrightarrow g_\omega(x) = 0, \forall x \in \mathbb{R}^n$.

Dans le théorème suivant, on montre que sous certaines hypothèses, on peut déterminer la solution de l'équation $f(x) = 0$, en construisant la suite

$$\begin{cases} x^{(n+1)} = x^{(n)} + \omega f(x^{(n)}), \\ x^{(0)} \in \mathbb{R}^n. \end{cases}$$

Théorème 2.1 Soit $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, vérifiant

1. $\exists \alpha > 0$ tel que, pour tout $x, y \in \mathbb{R}^n$, on a

$$\langle f(x) - f(y), x - y \rangle \leq -\alpha \|x - y\|_2^2,$$

2. $\exists M > 0$, telle que

$$\|f(x) - f(y)\|_2 \leq M \|x - y\|_2, \forall x, y \in \mathbb{R}^n.$$

Alors, la fonction g_ω (définie précédemment) est **contractante**, si et seulement si,

$$0 < \omega < \frac{2\alpha}{M^2}.$$

Vitesse de convergence Soit $(x^n)_{n \geq 0} \subset \mathbb{R}^n$ et $x^* \in \mathbb{R}^n$. Supposons que $(x^n)_n$ converge vers x^* , avec $x_n \neq x^*, \forall x \in \mathbb{R}^n$. On dit que :

- La convergence est linéaire s'il existe $\beta \in]0, 1[$, tel que

$$\lim_{n \rightarrow +\infty} \frac{|x^{(n+1)} - x^*|}{|x^{(n)} - x^*|} = \beta.$$

- La convergence est quadratique s'il existe $l > 0$, tel que

$$\lim_{n \rightarrow +\infty} \frac{|x^{(n+1)} - x^*|}{|x^{(n)} - x^*|^2} = l.$$

2.2 Méthode de Newton

Construire une suite de racines approchées $(x^{(n)})$ qui converge **quadratiquement** vers la racine exacte x^* de l'équation $f(x) = 0$ est bien le principe de la méthode de Newton. L'algorithme de Newton consiste alors à construire une suite de \mathbb{R}^n définie comme suit :

$$(1) \dots \dots \dots \begin{cases} F(x^{(n)}) (x^{(n+1)} - x^{(n)}) = -f(x^{(n)}), n > 0, \\ x^{(0)} \in \mathbb{R}^n, \end{cases}$$

ici $F(x^{(n)})$ représente la matrice différentielle de f au point $x^{(n)}$.

Remarque 2.2 A chaque itération $k \in \mathbb{N}$, de l'algorithme précédent, il faut effectuer les opérations suivantes :

- Calcul de $F(x^{(k)})$,
- Résolution du système d'équations linéaire :

$$F(x^{(k)}) (x^{(k+1)} - x^{(k)}) = -f(x^{(k)})$$

Théorème 2.3 ("Théorème de convergence")

Soit $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ et $x^* \in \mathbb{R}^n$, tel que $f(x^*) = 0$. On munit \mathbb{R}^n d'une norme $\|\cdot\|$ et on suppose que $F(x^{(k)})$ est inversible en tout point x de \mathbb{R}^n . Alors, il existe $b > 0$ et $\beta > 0$, tels que :

1. Si $x^{(0)} \in B(x^*, b) = \{x \in \mathbb{R}^n, \|x - x^*\| < b\}$, alors la suite $(x^{(n)})_n$ de l'algorithme de Newton, notée par (1), est bien définie, en plus, $x^{(n)} \in B(x^*, b), \forall n \geq 0$.
2. Si $x^{(0)} \in B(x^*, b)$ alors la suite (1) (de l'algorithme de Newton) converge vers x^* .
3. Si $x^{(0)} \in B(x^*, b)$ alors $\|x^{(n+1)} - x^*\| \leq \beta \|x^{(n)} - x^*\|, \forall n \geq 0$.

3 Exercices

"Résolution des Equations non linéaires"

Exercice 1.

Séparer les racines de ces deux équations :

$$f(x) = x^4 + 4x + 2 = 0, \quad (3.1)$$

$$g(x) = \cos x - x^2 + 1 = 0. \quad (3.2)$$

Exercice 2. Considérons l'équation,

$$4 \ln x = x$$

1. Montrer que cette équation admet une racine unique appartenant à l'intervalle $]1.2, 2[$.
2. Trouver le nombre d'itérations nécessaires pour calculer une racine approchée de cette équation, en appliquant la méthode de dichotomie, sur l'intervalle $[1.2, 2]$ et avec une précision $\varepsilon = 10^{-1}$.
3. Trouver cette racine approchée avec 5 chiffres significatifs, puis calculer l'erreur commise.

Exercice 3.

Soit l'équation suivante,

$$f(x) = x - \frac{1}{2} \sin x - 1. \quad (E)$$

1. Séparer les racines de l'équation (E) .
2. Déterminer un intervalle de la forme $[m, m + \frac{\pi}{2}]$ (où $m \in \mathbb{Z}$) contenant la racine exacte de (E) .
3. Calculer jusqu' à l'itération $k = 2$ une valeur approchée de la racine, en utilisant la méthode de Newton-Raphson, puis évaluer l'erreur commise à chaque itération.

Exercice 4.

Considérons l'équation

$$h(x) = x^3 + x - 1 = 0. \quad (\tilde{E})$$

1. Montrer que l'équation (\tilde{E}) possède une solution unique dans l'intervalle $[0, 1]$.
2. Utiliser la méthode de Newton-Raphson, pour trouver la racine approchée de (\tilde{E}) avec une précision $\varepsilon = 10^{-3}$.
3. En remarquant que (\tilde{E}) peut se réécrire sous la forme $x = g(x)$, où g est l'une des deux fonctions :

$$g(x) = 1 - x^3, \quad (3.3)$$

$$g(x) = 1 + \frac{1}{1 + x^2}, \quad (3.4)$$

laquelle parmi ces deux fonctions doit-on choisir, pour appliquer la méthode des approximations successives? (justifiez votre choix).

4. Estimer le nombre d'itérations qu'il faut établir, pour trouver la racine approchée de (\tilde{E}) , avec la même précision.

Exercice 5.

On considère le système d'équations non linéaires :

$$\begin{cases} x + y^2 - z^2, \\ y + xz = 19 \\ \frac{x^{(0)}}{2} + xy + y + z = 3.305. \end{cases}$$

1. Montrer que ce système admet une solution proche du vecteur $x^{(0)} = (1, 1, 1)^t$.
 2. Résoudre le système par la méthode du point fixe.
 3. Résoudre le système par la méthode de Newton.
- * On arrête les itérations, dans les deux méthodes, lorsque $\|E(x^{(n)})\| < 10^{-2}$.

Chapitre 3

Interpolation polynomiale

Considérons $\mathbb{P}_n[X]$ l'espace de tous les polynômes d'ordre inférieur ou égal à n ($n \in \mathbb{N}^*$) et a coefficients dans \mathbb{R} ou \mathbb{C} .

1 Introduction

1.1 Le Problème d'interpolation polynomiale

Soit f une fonction connue seulement en $(n + 1)$ points :

$$M_0(x_0, f(x_0)), M_1(x_1, f(x_1)), \dots, M_n(x_n, f(x_n)),$$

dont les abscisses $(x_i)_{i=0}^n$ sont deux à deux distincts.

Le problème d'interpolation polynomiale consiste à trouver un polynôme $P \in \mathbb{P}_n[X]$ dont le graphe passe par les $(n + 1)$ points, c. à. d,

$$\exists ? P \in \mathbb{P}_n[X] / P(x_i) = f(x_i), \forall i = 0, 1, \dots, n.$$

- Si on a que deux points (deux données) : $M_0(x_0, f_0), M_1(x_1, f_1)$, où $f_0, f_1 \in \mathbb{R}$ alors, le polynôme P interpolant ces deux points est de degré ≤ 1 . Donc, on cherche une droite $P(x) = ax + b$, avec $a, b \in \mathbb{R}$.
- Si on a que trois points (ou trois données), le polynôme d'interpolation de ces trois points est de degré ≤ 2 , c.à.d, de la forme $P(x) = ax^2 + bx + c$, $a, b, c \in \mathbb{R}$, pour tout $x \in \mathbb{R}$.
- En général, si on a $(n + 1)$ points, alors le polynôme P interpolant les $(n + 1)$ points est de degré $\leq n$ ($P \in \mathbb{P}_n[X]$). Donc, ils existent $C_0, C_1, \dots, C_n \in \mathbb{R}$, tels que

$$P(x) = C_0 + C_1x + \dots + C_nx^n, \forall x \in \mathbb{R}. \quad (1.1)$$

Remarque 1.1 1. Pour déterminer le polynôme d'interpolation P (dans (1.1)), il faut et il suffit de calculer ses coefficients C_i ($i = 0, 1, \dots, n$).

2. On interpole (ou approximer par un polynôme) une fonction donnée f (par exemple : sin, arctan, ln, exp ..., etc), pour pouvoir calculer la valeur approchée de cette fonction en un point donné $x \in D_f$, tel que x est différent de x_i ($i = 0, n$). On aura besoin du tableau ci-dessous pour calculer les C_i ($i = 0, 1, \dots, n$) et déterminer, ainsi l'interpolant de f .

x_i	x_0	.	.	.	x_n
$f(x_i)$	$f(x_0)$.	.	.	$f(x_n)$

1.2 Existence et unicité de l'interpolant

Théorème 1.2 "Existence et unicité de l'interpolant"

Soit f une fonction connue en $(n + 1)$ points $(x_i, f(x_i))_{i=0}^n$, tels que $(x_i)_{i=0}^n$ sont deux à deux distincts. Alors, il existe un et un seul $P_n \in \mathbb{P}_n[X]$ tel que :

$$P_n(x_i) = f(x_i), \text{ pour tout } i = \overline{0, n}.$$

(Ou bien $P_n(x_i) = y_i$, si la fonction f n'est pas connue).

Preuve : En général, on souhaite trouver un polynôme de degré $\leq n$ qui interpole les noeuds $(x_i, y_i)_{i=0}^n$. Comme $\mathbb{P}_n[X]$ est un espace vectoriel dont la base canonique s'écrit $\{1, X, X^2, \dots, X^n\}$, alors

$$P_n(X) = \sum_{k=0}^n a_k X^k, \text{ où } a_0, a_1, \dots, a_n \in \mathbb{R}.$$

On a,

$$\begin{cases} P_n(x_0) = y_0 \\ P_n(x_1) = y_1 \\ \cdot \\ \cdot \\ P_n(x_n) = y_n \end{cases} \Rightarrow \begin{cases} a_0 + a_1 x_0 + a_2 x_0^2 + \dots + a_n x_0^n = y_0 \\ a_0 + a_1 x_1 + a_2 x_1^2 + \dots + a_n x_1^n = y_1 \\ \cdot \\ \cdot \\ a_0 + a_1 x_n + a_2 x_n^2 + \dots + a_n x_n^n = y_n \end{cases}$$

D'où,

$$\begin{cases} P_n(x_0) = y_0 \\ P_n(x_1) = y_1 \\ \cdot \\ \cdot \\ P_n(x_n) = y_n \end{cases} \Rightarrow \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \cdot \\ \cdot \\ y_n \end{pmatrix}$$

Ainsi, les coefficients $(a_i)_i$ du polynôme d'interpolation s'obtiennent, en résolvant le système linéaire

$$Ba = y,$$

où $(B)_{i,j} = x_i^j, \forall i, j \in \{0, 1, \dots, n\}$, $a = (a_0, a_1, \dots, a_n)^t$ et $y = (y_0, y_1, \dots, y_n)^t \in \mathbb{R}^{n+1}$.

La matrice B est appelée matrice de **Vandermonde**. Le déterminant de B se donne par :

$$\det(B) = \prod_{0 \leq i < j \leq n} (x_j - x_i).$$

Comme $x_i \neq x_j (\forall i \neq j)$, alors $\det(B) \neq 0$. Par conséquent, le système $Ba = y$ admet une et une seule solution donnée par $a^* = B^{-1}y$ ($a^* \in \mathbb{R}^{n+1}$). Ceci prouve l'existence et l'unicité du polynôme de l'interpolation P_n , où $P_n(X) = \sum_{k=0}^n a_k^* X^k, \forall X \in \mathbb{R}$. ■

2 Interpolation de Lagrange

2.1 Base et polynôme de Lagrange

Définition 2.1 "Base de Lagrange"

Pour tout $x \in \mathbb{R}$ et $i = \overline{0, n}$, on définit les polynômes de base de Lagrange associés aux points d'abscisses $(x_i)_{i=0}^n$, et qu'on note par L_i , comme suit :

$$L_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^n \frac{(x-x_k)}{(x_i-x_k)}.$$

Exemple 2.2 Déterminer les polynômes de base de Lagrange associés aux points $x_0 = \frac{1}{2}, x_1 = 1, x_2 = \frac{3}{2}$.

Corrigé : Pour tout $x \in \mathbb{R}$, on a

$$\begin{aligned} L_0(x) &= \prod_{\substack{k=0 \\ k \neq 0}}^2 \frac{(x-x_k)}{(x_0-x_k)} = \frac{(x-x_1)(x-x_2)}{(\frac{1}{2}-1)(\frac{1}{2}-\frac{3}{2})} \\ &= \frac{(x-1)(x-\frac{3}{2})}{(-\frac{1}{2})(-1)} \\ &= 2(x-1)(x-\frac{3}{2}), \end{aligned}$$

$$\begin{aligned} L_1(x) &= \prod_{\substack{k=0 \\ k \neq 1}}^2 \frac{(x-x_k)}{(x_1-x_k)} = \frac{(x-x_0)(x-x_2)}{(1-\frac{1}{2})(1-\frac{3}{2})} \\ &= \frac{(x-\frac{1}{2})(x-\frac{3}{2})}{(\frac{1}{2})(-\frac{1}{2})} \\ &= -4(x-\frac{1}{2})(x-\frac{3}{2}). \end{aligned}$$

$$\begin{aligned} L_2(x) &= \prod_{\substack{k=0 \\ k \neq 2}}^2 \frac{(x-x_k)}{(x_2-x_k)} = \frac{(x-x_0)(x-x_1)}{(\frac{3}{2}-\frac{1}{2})(\frac{3}{2}-1)} \\ &= 2(x-\frac{1}{2})(x-1). \blacksquare \end{aligned}$$

Proposition 2.3 Les polynômes de base de Lagrange, associés aux points d'abscisses $(x_i)_{i=0}^n$, vérifient les propriétés suivantes :

1. Pour tout $i, j \in \{0, 1, \dots, n\}$,

$$L_i(x_j) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (*)$$

2. Pour tout $i \in \{0, 1, \dots, n\}$, L_i est un polynôme de degré n .

3. L'ensemble $\mathcal{B} = \{L_0, L_1, \dots, L_n\}$ est une base pour $\mathbb{P}_n[X]$.

Preuve : Exercice.

(*) δ_{ij} : le coefficient de "Kronecker".

Définition 2.4 "Polynôme d'interpolation de Lagrange"

On définit P_n le polynôme d'interpolation de Lagrange, aux points d'abscisses $(x_i)_{i=1}^n$, par

$$P_n(x) = \sum_{i=0}^n y_i L_i(x), \quad x \in \mathbb{R},$$

où L_i ($i = 1, \dots, n$) sont les polynômes de base de Lagrange, avec $y_i = f(x_i)$, lorsque f est définie explicitement.

2.2 Exemple et remarques importantes

Exemple 2.5 Soit $f(x) = \sqrt{2+x}$, $x \geq -2$.

Déterminer le polynôme de Lagrange interpolant (ou passant par) f aux points d'abscisses : $x_0 = 0, x_1 = 1$ et $x_2 = 2$.

Corrigé : Comme on a trois points d'interpolation, le polynôme de Lagrange associé est d'ordre "2". Soit P_2 ce polynôme. Donc, pour tout $x \geq -2$, on a

$$P_2(x) = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x),$$

où

$$y_0 = f(x_0) = \sqrt{2}, \quad y_1 = f(x_1) = \sqrt{3}, \quad y_2 = f(x_2) = 2,$$

$$\begin{aligned} L_0(x) &= \prod_{k=0, k \neq 0}^2 \frac{(x - x_k)}{(x_0 - x_k)} = \frac{(x - 1)(x - 2)}{(0 - 1)(0 - 2)} \\ &= \frac{1}{2}(x - 1)(x - 2), \end{aligned}$$

$$\begin{aligned} L_1(x) &= \prod_{k=0, k \neq 1}^2 \frac{(x - x_k)}{(x_0 - x_k)} = \frac{(x - 0)(x - 2)}{(1 - 0)(1 - 2)} \\ &= -x(x - 2), \end{aligned}$$

$$\begin{aligned} L_2(x) &= \prod_{k=0, k \neq 2}^2 \frac{(x - x_k)}{(x_0 - x_k)} = \frac{(x - 0)(x - 1)}{(2 - 0)(2 - 1)} \\ &= \frac{1}{2}x(x - 1). \end{aligned}$$

Par suite, il vient que

$$\begin{aligned} P_2(x) &= \sqrt{2}L_0(x) + \sqrt{3}L_1(x) + 2L_2(x) \\ &= \frac{\sqrt{2}}{2}(x - 1)(x - 2) - \sqrt{3}x(x - 2) + x(x - 1) \\ &= (1 - \sqrt{3} - \frac{1}{\sqrt{2}})x^2 + (2\sqrt{3} - \frac{3}{\sqrt{2}} - 1)x + \sqrt{2}, \quad \forall x \geq -2. \blacksquare \end{aligned}$$

Remarque : Ici, on considère comme valeur "exacte" de $\sqrt{2}$ (ou de $\sqrt{3}$) celle obtenue avec la calculatrice, à 3 ou à 4 décimales (c.à.d, $\sqrt{2} = 1,414$ ou $\sqrt{2} = 1,4142$).

Exercice :

1. En utilisant la méthode de Lagrange, déterminer P_3 le polynôme d'interpolation aux points $(x_i, y_i)_i$ donnés par :

x_i	4	6	8	10
y_i	1	3	8	20

2. Calculer la valeur de $P_3(4, 5)$ et $P_3(9, 25)$, à 3 décimales.

Remarque 2.6 1. Si l'une des valeurs y_i est nulle, on a pas besoin de calculer le L_i correspondant à cette valeur (car le produit $y_i L_i(x)$ sera nul, pour tout $x \in \mathbb{R}$).

2. Si on ajoute un autre point d'interpolation x_{n+1} aux points x_0, x_1, \dots, x_n , alors on doit déterminer à nouveau $L_0(x), L_1(x), \dots, L_n(x)$ et on ajoute $L_{n+1}(x)$. "On dit que la méthode de Lagrange s'adapte mal au changement de points".

3 Interpolation de Newton

3.1 Les différences divisées : "Définitions et propriétés"

Définition 3.1 "Les différences divisées (D.D)"

Soit f une fonction définie sur un intervalle $[a, b]$ de \mathbb{R} contenant $(n + 1)$ points distincts : x_0, x_1, \dots, x_n . On définit les différences divisées (D.D) d'ordre 'i'; $i = \overline{0, n}$ de f , aux points x_0, x_1, \dots, x_n , comme suit :

- à l'ordre '0' $\rightarrow f[x_i] = f(x_i); i = \overline{0, n}$.
- à l'ordre '1' $\rightarrow f[x_i, x_j] = \frac{f(x_j) - f(x_i)}{x_j - x_i}; i, j \in \{0, 1, \dots, n\}$ (avec $i \neq j$).
- à l'ordre '2' $\rightarrow f[x_i, x_j, x_l] = \frac{f[x_j, x_l] - f[x_i, x_j]}{x_l - x_i}; i, j, l \in \{0, 1, \dots, n\}$
-
-
-
- à l'ordre $i \geq 3 \rightarrow f[x_0, x_1, \dots, x_i] = \frac{f[x_1, x_2, \dots, x_i] - f[x_0, x_1, \dots, x_{i-1}]}{x_i - x_0}$.
- à l'ordre $n \rightarrow f[x_0, x_1, \dots, x_n] = \frac{f[x_1, x_2, \dots, x_n] - f[x_0, x_1, \dots, x_{n-1}]}{x_n - x_0}$.

Proposition 3.2 "Quelques propriétés des différences divisées"

1. $\forall i, j, k \in \{0, 1, \dots, n\}, f[x_i, x_j] = f[x_j, x_i]$ et $f[x_i, x_j, x_k] = f[x_k, x_j, x_i]$.
2. En général,

$$f[x_i, x_{i+1}, \dots, x_{i+k}] = f[x_{\sigma(i)}, x_{\sigma(i+1)}, \dots, x_{\sigma(i+k)}],$$

où σ est une permutation de l'ensemble $\{i, i + 1, \dots, i + k\}$.

" La valeur d'une différence divisée est indépendante de l'ordre des points $(x_i)_{i=0}^n$ ".

Preuve : Exercice.

Définition 3.3 "Base de Newton"

Les polynômes de base de Newton (notés N_i), associés aux points d'abscisses $(x_i)_{i=1}^n$, sont définis pour tout $x \in \mathbb{R}$ par

- $N_0(x) = 1$,
- $N_1(x) = (x - x_0)$,
- $N_2(x) = (x - x_0)(x - x_1)$,
- \vdots
- $N_i(x) = (x - x_0)(x - x_1) \times \dots \times (x - x_{i-1}) = \prod_{k=0}^{i-1} (x - x_k), \forall i \geq 2$.

Proposition 3.4 Les polynômes de base de Newton, associés aux points d'abscisses $(x_i)_{i=1}^n$, ont les propriétés suivantes :

1. Pour tout $i \in \{0, 1, \dots, n\}$, N_i est un polynôme d'ordre 'i'.
2. $\forall i \geq 1$, N_i admet x_0, x_1, \dots, x_{i-1} comme racines simples.
3. La partie $\mathcal{B}^* = \{N_0, N_1, \dots, N_n\}$ est une base de $\mathbb{P}_n[X]$.

Preuve : Exercice.

3.2 Interpolation de Newton en fonction des différences divisées

Définition 3.5 "Le polynôme d'interpolation de Newton"

On appelle "interpolant de Newton" d'une fonction f aux points d'abscisses $(x_i)_{i=0}^n$, le polynôme P_n donné par :

$$P_n(x) = \sum_{i=0}^n d_i N_i(x),$$

où $x \in \mathbb{R}, d_i = f[x_0, x_1, \dots, x_i]$ (c'est la différence divisée d'ordre 'i', appelée aussi la i^{me} composante de P_n dans la base \mathcal{B}^*), $N_0 = 1$ et $N_i(x) = \prod_{k=0}^{i-1} (x - x_k), i = \overline{1, n}$.

Donc, on a aussi

$$P_n(x) = f(x_0) + \sum_{i=1}^n f[x_0, x_1, \dots, x_i] \prod_{k=0}^{i-1} (x - x_k).$$

Tableau permettant le calcul des différences divisées d'ordre 'i'.

Pour faciliter le calcul des (D.D) d'ordre 'i,' on utilise le tableau suivant :

x_i	$y_i = f(x_i)$	$f[x_i, x_{i+1}]$	$f[x_i, \dots, x_{i+2}]$	\dots	\dots	$f[x_0, x_1, \dots, x_n]$
x_0	$\mathbf{d_0} = f(x_0)$					
		$\mathbf{d_1} = f[x_0, x_1]$				
x_1	$f(x_1)$		$\mathbf{d_2} = f[x_0, x_1, x_2]$			
		$f[x_1, x_2]$		\dots		
x_2	$f(x_2)$		$f[x_1, x_2, x_3]$			$\mathbf{d_n}$
		$f[x_2, x_3]$		\dots		
x_3	$f[x_3]$		$f[x_2, x_3, x_4]$			
\cdot	\cdot	$f[x_3, x_4]$				
\cdot	\cdot	\cdot	\cdot	\cdot		
x_n	$f[x_n]$	\cdot	\cdot	\cdot		

Exemple 3.6 *Considérons la même fonction de l'exercice 1, c.à.d, $f(x) = \sqrt{2+x}, x \geq -2$. Déterminer le polynôme interpolant f aux points d'abssices 0, 1, 2, en utilisant la méthode de Newton.*

Corrigé : Notons par P le polynôme d'interpolation de Newton associé à f , aux points donnés. Alors pour tout $x \geq -2$, on a

$$P(x) = \sum_{i=0}^2 d_i N_i(x),$$

$$= d_0 N_0(x) + d_1 N_1(x) + d_2 N_2(x),$$

où $N_0(x) = 1$, $N_1(x) = (x - x_0) = (x - 0) = x$ et $N_2(x) = (x - x_0)(x - x_1) = x(x - 1)$. Pour calculer les d_i on utilise le tableau suivant :

x_i	$y_i = f(x_i)$	$f[x_i, x_{i+1}]$	$f[x_0, x_1, x_2]$
$x_0 = 0$	$f(x_0) = \sqrt{2} = \mathbf{d_0}$		
		$f[x_0, x_1] = \sqrt{3} - \sqrt{2} = \mathbf{d_1}$	
$x_1 = 1$	$f(x_1) = \sqrt{3}$		$f[x_0, x_1, x_2] = 1 - \sqrt{3} - \frac{1}{\sqrt{2}} = \mathbf{d_2}$
		$f[x_1, x_2] = 2 - \sqrt{3}$	
$x_2 = 2$	$f(x_2) = 2$		

Sachant que

$$d_2 = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = \frac{(2 - \sqrt{3}) - (\sqrt{3} - \sqrt{2})}{2}$$

$$= 1 - \sqrt{3} - \frac{1}{\sqrt{2}}.$$

Ainsi, le polynôme d'interpolation de Newton est :

$$P(x) = \sqrt{2} + (\sqrt{3} - \sqrt{2})x + (1 - \sqrt{3} - \frac{1}{\sqrt{2}})x(x - 1)$$

$$= (1 - \sqrt{3} - \frac{1}{\sqrt{2}})x^2 + (2\sqrt{3} - \frac{3}{\sqrt{2}} - 1)x + \sqrt{2}, \forall x \geq -2. \blacksquare$$

Remarque 3.7 1. Le polynôme d'interpolation obtenu par la Méthode de Newton ou de Lagrange est le même polynôme (C'est unique d'après le Théorème 1.2).

2. Si on rajoute un point d'interpolation supplémentaire " x_{n+1} ", le polynôme de Newton, interpolant f aux points $(x_i)_{i=0}^{n+1}$ se donne, pour tout $x \in \mathbb{R}$, par :

$$P_{n+1}(x) = P_n(x) + f[x_0, x_1, \dots, x_n, x_{n+1}] \prod_{k=0}^n (x - x_k),$$

ici P_n c'est le polynôme de Newton interpolant f aux points $(x_i)_{i=0}^n$.

Exemple 3.8 1. En utilisant la méthode de Newton, déterminer P_3 le polynôme d'interpolation d'une fonction f , passant par les points $(-4, 1245), (-1, 33), (0, 5), (2, 9)$.

2. En déduire P_4 le polynôme d'interpolation passant par les points $(-4, 1245), (-1, 33), (0, 5), (2, 9)$ et $(5, 1335)$.

Corrigé :

1. Le polynôme d'interpolation de Newton passant par les 4 premiers noeuds se donne, pour tout $x \in \mathbb{R}$, par

$$\begin{aligned} P_3(x) &= \sum_{i=0}^3 d_i N_i(x), \\ &= f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ &\quad + f[x_0, x_1, x_2, x_3](x - x_0)(x - x_1)(x - x_2). \end{aligned}$$

Pour calculer les différences divisées (D.D), on forme le tableau suivant :

x_i	$y_i = f(x_i)$	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$	$f[x_i, \dots, x_{i+3}]$	$f[x_0, \dots, x_4]$
$x_0 = -4$	$f(x_0) = 1245$				
		$\frac{33-1245}{-1+4} = -404$			
$x_1 = -1$	$f(x_1) = 33$		$\frac{-28+404}{0+4} = 94$		
		$\frac{5-33}{0+1} = -28$		$\frac{10-94}{2+4} = -14$	
$x_2 = 0$	$f(x_2) = 5$		$\frac{2+28}{2+1} = 10$		$\frac{14+13}{5+4} = 3 = \mathbf{d_4}$
		$\frac{9-5}{2-0} = 2$		$\frac{88-10}{5+1} = 13$	
$x_3 = 2$	$f(x_3) = 9$		$\frac{442-2}{5-0} = 88$		
		$\frac{1335-9}{5-2} = 442$			
$x_4 = 5$	$f(x_5) = 1335$				

Donc, on obtient

$$P_3(x) = 1245 - 404(x + 4) + 94(x + 4)(x + 1) - 14(x + 4)(x + 1)(x - 0).$$

2. D'après la remarque 3.7, on déduit que le polynôme interpolant les 5 noeuds est :

$$\begin{aligned} P_4(x) &= P_3(x) + f[x_0, x_1, x_2, x_3, x_4](x - x_0)(x - x_1)(x - x_2)(x - x_3) \\ &= P_3(x) + 3(x + 4)(x + 1)(x - 0)(x - 2). \end{aligned}$$

Après développement, on obtient

$$P_4(x) = 3x^4 - 5x^3 + 6x^2 - 14x + 5, \forall x \in \mathbb{R}. \blacksquare$$

Remarque 3.9 1. Pour vérifier si le polynôme obtenu est correcte ou non, on teste si $P_4(x_i) = f(x_i), \forall i = \overline{0, 4}$ ou non.

2. La dernière ligne dans le tableau précédent concerne la 2^{ème} question de l'exercice (le calcul de \mathbf{d}_4).

3.3 Cas particulier : "Des pôles équidistants"

Considérons $(n + 1)$ points équidistants $(x_i, y_i)_{i=0}^n$, avec $x_0 \in \mathbb{R}, x_1 = x_0 + h, x_2 = x_0 + 2h, \dots, x_n = x_0 + nh$ où $h > 0$: est le pas du déplacement. On définit aussi les n points équidistants par la relation de récurrence : $x_i = x_{i-1} + h, h > 0, i = \overline{1, n}$.

Définition 3.10 " Les différences finies progressives (D.F.P) "

Soient $(n+1)$ points équidistants $(x_i, y_i)_{i=0}^n$. On définit les différences finies progressives (D.F.P) d'ordre k ($k = \overline{0, n}$) par les expressions suivantes :

- à l'ordre '0' : $\Delta^0 y_i = f(x_i) = y_i, i = \overline{0, n}$.
- à l'ordre '1' : $\Delta^1 y_i = y_{i+1} - y_i, i = \overline{0, n-1}$.
- à l'ordre '2' : $\Delta^2 y_i = \Delta^1 y_{i+1} - \Delta^1 y_i, i = \overline{0, n-2}$.
-
-
-
- à l'ordre $k > 1$: $\Delta^k y_i = \Delta^{k-1} y_{i+1} - \Delta^{k-1} y_i, i = \overline{0, n-k}$.
- à l'ordre n : $\Delta^n y_0 = \Delta^{n-1} y_1 - \Delta^{n-1} y_0$.

Relation entre les différences finies progressives et les différences divisées

Théorème 3.11 Soit f une fonction dont on connaît ses valeurs aux points équidistants $(x_i)_{i=0}^n$. Alors, on a

$$\frac{\Delta^k y_i}{k! h^k} = f[x_i, x_{i+1}, \dots, x_{i+k}], k = \overline{1, n}, i = \overline{0, n-k},$$

où h est le pas de déplacement et $y_i = f(x_i), \forall i \in \{0, \dots, n\}$.

Preuve : Utilisons le raisonnement par récurrence sur k (pour tout $i = \overline{0, n-k}$).

1) La relation est vraie pour $k = 0$. En effet, on a

$$f[x_i] = y_i = \frac{\Delta^0 y_i}{0! h^0}, \forall i = \overline{0, n}.$$

2) Supposons que la relation est vraie jusqu'à l'ordre k et montrons qu'elle est vraie pour $k + 1$. Pour tout $i = \overline{0, n-k}$, on a

$$f[x_i, x_{i+1}, \dots, x_{i+k+1}] = \frac{f[x_{i+1}, x_{i+1}, \dots, x_{i+k+1}] - f[x_i, x_{i+1}, \dots, x_{i+k}]}{x_{i+k+1} - x_i}.$$

Or, par hypothèse on a

$$f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{\Delta^k y_i}{k!h^k},$$

et

$$f[x_{i+1}, x_{i+1}, \dots, x_{i+k+1}] = \frac{\Delta^k y_{i+1}}{k!h^k}.$$

D'autre par

$$x_{i+k+1} - x_i = (x_{i+k+1} - x_{i+k}) + (x_{i+k} - x_{i+k-1}) + (x_{i+k-1} - x_{i+k-2}) + \dots + (x_i - x_{i-1}) = (k+1)h.$$

Ainsi

$$f[x_i, x_{i+1}, \dots, x_{i+k+1}] = \frac{\Delta^k y_{i+1} - \Delta^k y_i}{k!h^k(k+1)h} = \frac{\Delta^{k+1} y_i}{(k+1)!h^{k+1}}. \blacksquare$$

4 Erreur d'interpolation

4.1 Erreur commise (effective)

Définition 4.1 Soit P_n le polynôme interpolant une fonction f aux noeux $(x_i, f(x_i))_{i=0}^n$. On définit l'erreur commise (ou effective), en approximant f par P_n , en un point $x \in [\min x_i, \max x_i]$ par :

$$E_c(x) = |f(x) - P_n(x)|.$$

Remarque 4.2 1. Il est bien clair que si $x = x_i$ (où $i = \overline{0, n}$), alors $E_c(x) = 0$.

2. Comme f n'est pas toujours connue, on introduit la notion de la borne supérieure de l'erreur commise, qu'on note ε et on écrit

$$E_c(x) = |f(x) - P_n(x)| \leq \varepsilon.$$

Par suite, on aura

$$P_n(x) - \varepsilon \leq f(x) \leq P_n(x) + \varepsilon,$$

ou encore : $f(x) \simeq P_n(x) \pm \varepsilon$.

4.2 Erreur théorique

Théorème 4.3 Soit P_n l'interpolant d'une fonction f aux points d'abscisses $(x_i)_{i=0}^n$, tels que :

$$a = x_0 < x_1 < x_2 < \dots < x_n = b \quad (a, b \in \mathbb{R}).$$

Si $f \in C^{n+1}([a, b])$, alors

1. $\forall x \in [a, b], \exists \theta = \theta(x) \in [a, b]$ tel que

$$f(x) - P_n(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!} \times \prod_{i=0}^n (x - x_i). \quad (*)$$

2. En posant $M = \sup_{a \leq x \leq b} |f^{(n+1)}(x)|$, il vient que

$$\begin{aligned} |f(x) - P_n(x)| &\leq \frac{M}{(n+1)!} \times |\Pi_{i=0}^n(x - x_i)| \\ &\leq \frac{M}{(n+1)!} \times (b-a)^n. \end{aligned}$$

Preuve :

1. Soit $x \in [a, b]$.

Cas 1. S'il existe $j \in \{0, 1, \dots, n\}$ tel que $x = x_j$, alors

$$e(x_j) = f(x_j) - P_n(x_j) = 0 \text{ et } \frac{f^{(n+1)}(\theta)}{(n+1)!} \times \Pi_{i=0}^n(x_j - x_i) = 0.$$

D'où le résultat.

Cas 2. Si $x \neq x_j$, pour tout $j \in \{0, 1, \dots, n\}$. Dans ce cas, on considère la fonction suivante :

$$g(t) = f(t) - P_n(t) - e(x) \cdot \Pi_{i=0}^n \left(\frac{t - x_i}{x - x_i} \right), \forall t \in [a, b].$$

Il est clair que $g \in C^{n+1}([a, b])$, $g(x_i) = 0$ pour tout $i \in \{0, 1, \dots, n\}$ et que $g(x) = 0$. Donc, g admet $(n+2)$ racines distinctes dans $[a, b]$. D'après le théorème de Rolle, la fonction g' admet $(n+1)$ racines distinctes dans $[a, b]$. En appliquant le théorème de Rolle sur g' , on déduit que g'' admet n racines distinctes dans $[a, b]$. Ainsi, on pourra vérifier que $g^{(n+1)}$ possède "une" racine dans $[a, b]$, qu'on note par θ . On a

$$g^{(n+1)}(\theta) = 0 \Leftrightarrow f^{(n+1)}(\theta) - e(x) \cdot \Pi_{i=0}^n \left(\frac{1}{x - x_i} \right) (n+1)! = 0.$$

D'où

$$e(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!} \times \Pi_{i=0}^n(x - x_i).$$

2. Posant $M = \sup_{a \leq x \leq b} |f^{(n+1)}(x)|$. D'après (1), $\exists \theta \in [a, b]$ tel que

$$\begin{aligned} |f(x) - P_n(x)| &= \frac{|f^{(n+1)}(\theta)|}{(n+1)!} \times |\Pi_{i=0}^n(x - x_i)| \\ &\leq \sup_{a \leq \theta \leq b} |f^{(n+1)}(\theta)| \times \frac{|\Pi_{i=0}^n(x - x_i)|}{(n+1)!} \\ &\leq \frac{M}{(n+1)!} \times (b-a)^n. \quad \blacksquare \end{aligned}$$

Exemple 4.4 Avec quelle précision peut-on calculer $\sqrt{115}$ à l'aide de la formule de Lagrange pour interpoler la fonction $f(x) = \sqrt{x}$ aux points : $x_0 = 100, x_1 = 121, x_2 = 144$.

Corrigé :

Soit P_2 le polynôme de Lagrange passant par f en $(x_i)_{i=0}^2$. Alors, pour tout $x \in [100, 144]$, on a

$$|f(x) - P_2(x)| \leq \frac{M}{3!} \times |\Pi_{i=0}^2(x - x_i)|, \text{ où } M = \sup_{x_0 \leq x \leq x_2} |f^{(3)}(x)|.$$

Pour $x = 115$, il vient que

$$|\Pi_{i=0}^2(115 - x_i)| = |(115 - 100)(115 - 121)(115 - 144)| = 2610.$$

D'autre part, les dérivées de $f : x \rightarrow \sqrt{x}$, jusqu'à l'ordre '3' sont :

$$f'(x) = \frac{1}{2\sqrt{x}}, f''(x) = -\frac{1}{4}x^{-\frac{3}{2}}, f^{(3)}(x) = \frac{3}{8}x^{-\frac{3}{2}}, \forall x > 0.$$

Comme la fonction $f^{(3)}$ est décroissante sur $[100, 144]$, alors $M = |f^{(3)}(100)| = \frac{3}{8}10^{-5}$.
Donc,

$$|f(x) - P_2(x)| \leq \frac{3}{8 \times 3!}10^{-5} \times 2610 = 0,16 \times 10^{-4}.$$

C'est la précision voulue, ou encore l'estimation de l'erreur théorique au point $x = 115$. ■

Relation entre les différences divisées et les dérivées d'ordre "n"

Lemme 4.5 Soit $f \in C^n([a, b])$ et $(x_i)_{i=0}^n \subset [a, b]$ tels que $(x_i)_i$ sont deux à deux distincts. Alors, $\exists \theta \in [\min_{0 \leq i \leq n} x_i, \max_{0 \leq i \leq n} x_i]$ vérifiant

$$f[x_0, x_1, \dots, x_n] = \frac{f^{(n)}(\theta)}{n!}.$$

Idée de la preuve : Il suffit d'appliquer le théorème de Rolle n fois sur la fonction

$$d : x \rightarrow d(x) = (f - P_n)(x),$$

où P_n est le polynôme de Newton interpolant f , aux point $x_i, i = \overline{0, n}$.

Remarque 4.6 1. L'erreur d'interpolation (effective ou théorique) se donne par les mêmes relations, quelque soit la méthode utilisée (de Lagrange, de Newton ou autres méthodes).

2. D'après la relation (*), donnée au Théorème 4.3, rien n'assure que l'erreur d'interpolation tend vers '0' lorsque n tend vers ∞ .
3. Cependant, l'erreur d'interpolation dépend de deux termes :

(a) $M = \sup_{a \leq x \leq b} |f^{(n+1)}(x)|$

(b) $\max_{a \leq x \leq b} \prod_{i=0}^n |x - x_i|.$

Pour le terme (a), on ne peut pas le minimiser tant que f est fixée dès le début. Par contre, le terme (b), qui dépend de l'emplacement du point x avec les points d'interpolations $(x_i)_i$, on peut le changer de sorte que l'erreur d'interpolation soit minimale (ou optimale). Cela est exactement l'objectif du paragraphe suivant.

5 Interpolation optimale "au points de Tchebychev"

5.1 Interpolation de Tchebychev dans l'intervalle [-1,1]

Définition 5.1 " Polynôme de Tchebychev "

Soit $n \in \mathbb{N}$. On définit le polynôme de Tchebychev d'ordre n , qu'on note T_n , comme suit,

$$T_n : [-1, 1] \rightarrow [-1, 1]$$

$$x \mapsto T_n(x) = \cos(n \arccos(x)).$$

* Si on pose $x = \cos \theta$, où $\theta \in [0, 2\pi]$, on aura pour tout $n \in \mathbb{N}$

$$T_n(x) = T_n(\cos \theta) = \cos(n\theta).$$

* En développant la fonction $\cos(n\theta)$, on obtient, aussi, pour tout $n \in \mathbb{N}$ et $x \in [-1, 1]$,

$$T_n(x) = 2^{n-1}x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_1x + a_0.$$

Proposition 5.2 Les polynômes de Tchebychev $(T_n)_n$ ont les propriétés suivantes :

1. $\forall n \in \mathbb{N}^*, \forall x \in [-1, 1]$,

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

2. Les racines de chaque polynôme T_n (ou les points de Tcheby) sont :

$$\hat{x}_i = \cos\left(\frac{2i+1}{2n} \cdot \pi\right), i = \overline{0, n-1}.$$

Par suite, $\forall x \in [-1, 1]$, on a

$$T_n(x) = 2^{n-1}(x - \hat{x}_0)(x - \hat{x}_1) \times \dots \times (x - \hat{x}_{n-1}).$$

3. Les extrémums de chaque polynôme T_n , notés par $(z_i)_{i=\overline{0, n}}$, sont :

$$z_i = \cos\left(\frac{i\pi}{n}\right), i = \overline{0, n},$$

avec $T_n(z_i) = \pm 1$.

Preuve : Exercice. **Indication pour 1 :** Utiliser la relation :

$$\cos((n+1)\theta) + \cos((n-1)\theta) = 2 \cos \theta \cdot \cos n\theta$$

.

Théorème 5.3 Pour tout $n \in \mathbb{N}^*$, les points $(\hat{x}_i)_{i=\overline{0, n-1}}$; racines de T_n sont les meilleurs points d'interpolation polynomiale, dans l'intervalle $[-1, 1]$. Autrement dit,

$$\max_{-1 \leq x \leq 1} |x - \hat{x}_i| \leq \max_{-1 \leq x \leq 1} |x - x_i|,$$

où $(x_i)_{i=\overline{0, n-1}}$ sont des points d'interpolation quelconques de $[-1, 1]$.

5.2 Interpolation de Tcheby dans un intervalle quelconque de \mathbb{R}

Remarque 5.4 1. Les n points d'interpolation de Tchebychev, dans un intervalle quelconque $[a, b]$ de \mathbb{R} , sont donnés par la relation suivante :

$$\hat{x}_i = \frac{b+a}{2} + \frac{b-a}{2} \cos\left(\frac{2i+1}{2n} \cdot \pi\right), i = \overline{0, n-1}.$$

2. Soit $E_t(x)$ l'erreur théorique en interpolant une fonction f en un point x , aux " n " points de Tchebychev. Alors, on a :

$$|E_t(x)| = |f(x) - P_n(x)| \leq \frac{M}{(n+1)!2^{n-1}},$$

où $M = \sup_{\hat{x}_0 \leq x \leq \hat{x}_n} |f^{(n+1)}(x)|$.

6 Exercices

"Sur la méthode de Cramer"

Exercice 1. Donnons la forme de la matrice A dans les cas suivants :

1. $a_{ij} = 0, \forall i, j = \overline{1, n}, i \neq j \Leftrightarrow A$ est diagonale.
2. $a_{ij} = 0, \forall i, j = \overline{1, n}, i < j \Leftrightarrow A$ est triangulaire inférieure.
3. $a_{ij} = 0, \forall i, j = \overline{1, n}, i > j \Leftrightarrow A$ est triangulaire supérieure.

$$4. A = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{11} & a_{22} & a_{23} & \dots & 0 & 0 & 0 & 0 \\ 0 & a_{31} & a_{32} & a_{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & a_{k(k-1)} & a_{kk} & a_{k(k+1)} & 0 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{n(n-1)} & a_{nn} \end{pmatrix} \quad (A \text{ est dite tri-diagonale}).$$

Exercice 2. Résoudre les 3 systèmes linéaires par la "Méthode de Cramer" :

1.

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = b; \text{ où } A = \begin{pmatrix} 2 & 1 & 5 \\ 6 & 3 & 2 \\ 0 & 4 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Ce système admet une solution unique puisque $\det(A) = 104 \neq 0$. En appliquant la méthode de Cramer, sa solution $(x, y, z)^t$ se donne par :

$$x = \frac{\begin{vmatrix} 1 & 1 & 5 \\ 0 & 3 & 2 \\ 1 & 4 & 1 \end{vmatrix}}{\det(A)} = -\frac{9}{52}, y = \frac{\begin{vmatrix} 2 & 1 & 5 \\ 6 & 0 & 2 \\ 0 & 1 & 1 \end{vmatrix}}{\det(A)} = \frac{5}{26}, z = \frac{\begin{vmatrix} 2 & 1 & 1 \\ 6 & 3 & 0 \\ 0 & 4 & 1 \end{vmatrix}}{\det(A)} = \frac{3}{13}.$$

2.

$$A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = b; \text{ où } A = \begin{pmatrix} 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \\ 1 & 2 & 4 & 8 \\ 1 & 4 & 16 & 64 \end{pmatrix}, b = \begin{pmatrix} -14 \\ -\frac{11}{4} \\ -8 \\ -29 \end{pmatrix}$$

On a $\det(A) = 720 \neq 0$, alors ce système admet une solution unique. En suivant les mêmes étapes que dans 1, on obtient :

$$x_1 = \frac{\begin{vmatrix} -14 & -2 & 4 & -8 \\ -11/4 & -1 & 1 & -1 \\ -8 & 2 & 4 & 8 \\ -29 & 4 & 16 & 64 \end{vmatrix}}{720} = 1, x_2 = \frac{\begin{vmatrix} 1 & -14 & 4 & -8 \\ 1 & -11/4 & 1 & -1 \\ 1 & -8 & 4 & 8 \\ 1 & -29 & 16 & 64 \end{vmatrix}}{720} = 1/2, x_3 = -3, x_4 = 1/4.$$

Donc, la solution du 2ème système linéaire est $(x_1, x_2, x_3, x_4) = (1, 1/2, -3, 1/4)$.

3. La résolution d'un système linéaire de 4 équations de type " $P_3(x_j) = y_j, j = \overline{1, 3}; P_3(x) = \sum_{i=0}^3 a_i x^i$ ", par la méthode de Cramer donne : $P_3(x) = \frac{53}{30}x^3 - 7x^2 + \frac{253}{30}x - 1$.

"Interpolation Polynomiale"

Dans tout ce qui suit $\mathbb{P}_n[X]$ désigne l'espace des Polynômes de degré inférieur ou égal à n .

Exercice 1

- (1) Déterminer $P \in \mathbb{P}_3[X]$ le polynôme d'interpolation de Lagrange satisfaisant au tableau ci-dessous.
- (2) Donner une valeur approchée de $f(4.2)$ par cette interpolation.

x	0	2	3	5
$f(x)$	-1	2	9	87

Exercice 2 Soit $f(x) = \ln x$, $x \in \mathbb{R}_+^*$.

- (1) Par la méthode de Lagrange, estimer la valeur de $\ln(0.60)$ avec :

x	0.40	0.50	0.70	0.80
$f(x)$	-0.916291	-0.699314	-0.356675	-0.223144

- (2) Calculer $E_e(0.60)$ l'erreur effectivement commise en approximant f par le polynôme de Lagrange au point $x = 0.60$. (On considère comme valeur "exacte" celle obtenue grâce à votre calculatrice).
- (3) Estimer $E_t(0.60)$: l'erreur théorique au point $x = 0.60$. Commenter.

Exercice 3

- (a) Réécrire la formule d'interpolation de Lagrange dans le cas où les points d'appui $(x_i)_{0 \leq i \leq N}$ sont équidistants.
- (b) Utiliser le même tableau qu'à l'exercice précédent ; complété par l'"exacte" valeur de $\ln(0.6)$, pour estimer la valeur de $\ln(0.54)$.

Exercice 4 Soit $n \in \mathbb{N}^*$, $x \in \mathbb{R}$.

- (1) Calculer l'erreur théorique en interpolant la fonction $g : x \mapsto x^n$, aux points $x_i = \frac{i}{n}$; $i = 0, 1, \dots, n$. Expliquer le résultat.
- (2) Donner une expression simple de $f(x)$;

$$f(x) = L_1(x) + 2^k L_2(x) + \dots + n^k L_n(x),$$

où $1 \leq k \leq n$ et L_0, L_1, \dots, L_n sont les polynômes de base de Lagrange associés aux points $x_i = \frac{i}{n}$; $i = 0, 1, \dots, n$.

Exercice 5

Soit la fonction $f(x) = e^{-\frac{x}{10}}$ définie sur l'intervalle $[1, 4]$ par la table :

x_i	1	2	3	4
$f(x_i)$	0.905	0.819	0.741	0.670

- (1) Calculer l'interpolant de f par la méthode de Newton.
- (2) Estimer la valeur d'erreur au point $x = 1,5$ si $|f^{(4)}(x)| \leq 10^{-2}$.

Exercice 6

- (1) Compléter la table de différences divisées suivante :

x_i	$f(x_i)$	$f[x_i, x_{i+1}]$	$f[x_i, \dots, x_{i+2}]$	$f[x_i, \dots, x_{i+3}]$
1,9	0,94630			
		-0,127975		
1,5	0,99749		?	
		-0,314725		?
2,3	0,74571		?	
		-0,795824		
2,7	0,42738			

- (2) Calculer une approximation de $f(1,8)$, en utilisant le polynôme de Newton passant par les 3 premiers points.
- (3) Donner une estimation de l'erreur de l'interpolation en $x = 1,8$.
- (4) Sachant que $f(x) = \sin x$, estimer l'erreur théorique en $x = 1,8$.
- (5) Quel polynôme est le plus précis, celui trouvé en 2, ou le polynôme de Lagrange passant par f en $x = 1,5; 1,9$ et $2,3$? Justifier.

Exercice 7 Considérons les polynômes de Tchebychev définis par :

$$T_n(x) = \cos(n \cdot \arccos x); \quad x \in [-1, 1] \text{ et } n \geq 0.$$

- 1) Vérifier que : $T_0(x) = 1$, $T_1(x) = x$,
et que $\forall n \geq 1, T_{n+1}(x) = 2x.T_n(x) - T_{n-1}(x)$.

Exercice 8

On considère la fonction $f(x) = \frac{1}{1+x^2}$ définie sur $[-1, 1]$.

- a) Déterminer l'interpolant de f aux points optimaux ($n = 3$).
- b) Au point $x = 2/3$, comparer les erreurs obtenues par l'interpolation de f aux points équidistants et celle aux points de Tchebychev.

Chapitre 4

Approximation au sens des moindres carrés

Dans ce chapitre, on désigne par "M.A.S.M.C" : la meilleure approximation au sens des moindres carrés.

Introduction

Soit f une fonction définie sur un intervalle $[a, b]$ de \mathbb{R} , explicitement ou connue seulement en un nombre fini de points $(x_i)_{(i=\overline{0, n})}$ de $[a, b]$.

Le problème d'approximation de f au sens des **moindres carrés (A.S.M.C)** consiste à déterminer une fonction φ^* d'écriture simple (un polynôme , une exponentielle,...,etc) de telle sorte que " la distance entre f est φ^* soit **minimale** ".

1 Notions préliminaires

Définition 1.1 Soit E un espace vectoriel sur un corps \mathbb{K} (où $\mathbb{K} = \mathbb{R}$ ou \mathbb{C}).

L'application $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{R}$ est dite **produit scalaire** sur E , si et seulement si elle vérifie ce qui suit :

1. $\langle x, x \rangle \geq 0, \forall x \in E$ et $(\langle x, x \rangle = 0 \Leftrightarrow x = 0)$.
2. $\langle x, y \rangle = \langle y, x \rangle, \forall x, y \in E$.
3. $\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle, \forall x, y, z \in E, \forall \alpha, \beta \in \mathbb{K}$.

Remarque 1.2 Un espace vectoriel muni d'un produit scalaire est appelé "un espace **préhilbertien**".

Exemples :

1. $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ est un espace préhilbertien avec le produit scalaire : $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$.
2. $(C[a, b], \langle \cdot, \cdot \rangle)$ (où $C([a, b])$ est l'espace vectoriel des fonctions continues sur l'intervalle réel $[a, b]$) est un espace préhilbertien avec le produit scalaire : $\langle f, g \rangle = \int_a^b f(x)g(x).w(x)dx$, où w est une fonction appelée " fonction poids".

Définition 1.3 Soit E un espace vectoriel sur \mathbb{K} (où $\mathbb{K} = \mathbb{R}$ ou \mathbb{C}). Une application $N : E \rightarrow \mathbb{R}$ est appelée **norme** si, et seulement si les trois propriétés suivantes sont vérifiées :

1. $N(x) = 0 \Leftrightarrow x = 0, \forall x \in E$.
2. $N(\lambda x) = |\lambda| N(x), \forall x \in E, \forall \lambda \in \mathbb{K}$.
3. $N(x + y) \leq N(x) + N(y), \forall x, y \in E$.

Vocabulaire et remarques :

Une norme se note par $N(\cdot)$ ou $\|\cdot\|$.

Si $\|\cdot\|$ est une norme sur l'espace vectoriel E , alors le couple $(E, \|\cdot\|)$ est appelé un **espace vectoriel normé**.

" La norme la plus connue " :

Un produit scalaire $\langle \cdot, \cdot \rangle$ sur $E \times E$ permet de définir une norme sur E par la relation :

$$\forall x \in E, \|x\| = \sqrt{\langle x, x \rangle},$$

ce qui est équivalent aussi à :

$$\forall x \in E, \|x\|^2 = \langle x, x \rangle.$$

Définition 1.4 " Vecteurs Orthogonaux "

Soit $(E, \langle \cdot, \cdot \rangle)$ est un espace préhilbertien. On dit que x, y de E sont **" orthogonaux "** par rapport au produit scalaire $\langle \cdot, \cdot \rangle$ ssi (si et seulement si) : $\langle x, y \rangle = 0$ et on note $x \perp y$.

Définition 1.5 Dans un espace préhilbertien $(E, \langle \cdot, \cdot \rangle)$, un système (ou ensemble) de vecteurs $\{V_1, V_2, \dots, V_m\}$ de E est dit orthogonal, par rapport au produit scalaire $\langle \cdot, \cdot \rangle$, ssi $\forall i, j = \overline{0, m}$, on a

$$\langle V_i, V_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ \|V_i\|^2 & \text{si } i = j, \end{cases}$$

en plus, si on a $\|V_i\|^2 = 1, \forall i, j = \overline{0, m}$ alors, on dit que le système $\{V_1, V_2, \dots, V_m\}$ est **" orthonormal "**.

Exemple 1.6 1. La base canonique $(e_i)_{(i=\overline{1, n})}$ de \mathbb{R}^n ($e_1 = (1, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), e_3 = (0, 0, 1, \dots, 0), \dots$) munie du produit scalaire euclidien (c.à.d : $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$) est une base orthonormée de l'espace \mathbb{R}^n , car $\forall i, j = \overline{0, n}$, on a

$$\langle e_i, e_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

2. Les polynômes $(T_n)_{n \geq 0}$ de Tchebychev (étudiés au chapitre 1) forment un système orthogonal sur $[a, b]$, par rapport à la fonction poids $w(x) = \frac{1}{\sqrt{1-x^2}}$ et au produit scalaire défini par : $\forall m, n \in \mathbb{N}$,

$$\begin{aligned} \langle T_m, T_n \rangle &= \int_{-1}^1 T_m(x) T_n(x) \cdot \frac{1}{\sqrt{1-x^2}} dx \\ &= \begin{cases} 0 & \text{si } n \neq m \\ \pi & \text{si } n = m = 0 \\ \frac{\pi}{2} & \text{sinon (c.à.d : } n = m \text{ et } m \neq 0.) \end{cases} \end{aligned}$$

2 Meilleure approximation dans un espace normé

2.1 Formulation du problème

Définition 2.1 Soit $f \in E$ (où $(E, \|\cdot\|)$: un espace vectoriel normé) et soit G un sous-espace vectoriel de E . On dit que φ^* de G est une meilleure approximation de f dans G , ssi

$$\|f - \varphi^*\| = \min_{\varphi \in G} \|f - \varphi\| \dots\dots(P)$$

Remarque 2.2 1. Si $f \in G$, alors $\varphi^* = f$.

2. Le problème (P) est équivalent au problème suivant :

$$\exists \varphi^* \in G \text{ tel que } \|f - \varphi^*\|^2 = \min_{\varphi \in G} \|f - \varphi\|^2$$

ce qui est équivalent à

$$\exists \varphi^* \in G \text{ tel que } \langle f - \varphi^*, f - \varphi^* \rangle = \min_{\varphi \in G} \langle f - \varphi, f - \varphi \rangle .$$

2.2 Existence et unicité de la meilleure approximation

Théorème 2.3 Soient E un espace vectoriel normé et G un sous-espace vectoriel de E . Si $\dim G < +\infty$, alors tout élément $f \in E$ admet au moins une approximation dans G , au sens de la Définition 2.1.

Théorème 2.4 "Propriété caractéristique de la M.A.S.M.C"

Supposons que les hypothèses du théorème précédent sont satisfaites. On dit que φ^* ($\varphi^* \in G$) est une meilleure approximation (au sens des moindres carrés) d'un élément f ($f \in E$) dans G , si et seulement si :

$$\langle f - \varphi^*, \varphi \rangle = 0, \forall \varphi \in G$$

c.à.d. que le vecteur $f - \varphi^*$ est orthogonal à chaque vecteur φ de G (donc orthogonal à G).

Remarque 2.5 On peut voir le problème de la M.A.S.M.C comme un problème de **projection orthogonale** d'un élément $f \in E$, sur le sous-espace vectoriel (s.e.v) G . Dans ce cas, $f - \varphi^*$ est un vecteur orthogonal à G .

Théorème 2.6 "Unicité de la M.A.S.M.C"

Si la M.A.S.M.C d'un vecteur f de E dans un s.e.v G de E existe, alors elle est **unique**.

2.3 Construction de la M.A.S.M.C

Soient E un espace vectoriel normé et G un sous-espace vectoriel de E avec $\dim G < +\infty$. Soit $B = \{\varphi_1, \varphi_2, \dots, \varphi_m\}$ une base de G (c.à.d. $B \subset G$; B libre et génératrice). Considérons φ^* la M.A.S.M.C de f ($f \in E$) dans G . Puisque $\varphi^* \in G$ et B est une partie génératrice de G , alors

$$\exists a_1, a_2, \dots, a_m \in \mathbb{R}, \text{ tels que } \varphi^* = \sum_{i=1}^m a_i \varphi_i.$$

Du Théorème 2.4, on a $\langle f - \varphi^*, \varphi \rangle = 0, \forall \varphi \in G$. En substituant la valeur de φ^* , cela donne : $\langle f - \sum_{i=1}^m a_i \varphi_i, \varphi_j \rangle = 0, \forall j = \overline{1, m}$ (car $(\varphi_j)_{j=\overline{1, m}}$ est une base de G). Il résulte que

$\langle f, \varphi_j \rangle - \langle f, \sum_{i=1}^m a_i \langle \varphi_i, \varphi_j \rangle = 0, \forall j = \overline{1, m}$. Ainsi, on obtient

$$\langle f, \varphi_j \rangle = \sum_{i=1}^m a_i \langle \varphi_i, \varphi_j \rangle, \forall j = \overline{1, m}.$$

Cette dernière relation donne le système (de "m" équations linéaires) suivant :

$$\begin{cases} \sum_{i=1}^m a_i \langle \varphi_i, \varphi_1 \rangle = \langle f, \varphi_1 \rangle \\ \sum_{i=1}^m a_i \langle \varphi_i, \varphi_2 \rangle = \langle f, \varphi_2 \rangle \\ \cdot \\ \cdot \\ \sum_{i=1}^m a_i \langle \varphi_i, \varphi_m \rangle = \langle f, \varphi_m \rangle \end{cases} \dots\dots(S)$$

$$(S) \iff \begin{pmatrix} \langle \varphi_1, \varphi_1 \rangle & \langle \varphi_2, \varphi_1 \rangle & \langle \varphi_3, \varphi_1 \rangle & \dots & \langle \varphi_m, \varphi_1 \rangle \\ \langle \varphi_1, \varphi_2 \rangle & \langle \varphi_2, \varphi_2 \rangle & \langle \varphi_3, \varphi_2 \rangle & \dots & \langle \varphi_m, \varphi_2 \rangle \\ \cdot \\ \cdot \\ \langle \varphi_1, \varphi_m \rangle & \langle \varphi_2, \varphi_m \rangle & \langle \varphi_3, \varphi_m \rangle & \dots & \langle \varphi_m, \varphi_m \rangle \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_m \end{pmatrix} = \begin{pmatrix} \langle f, \varphi_1 \rangle \\ \langle f, \varphi_2 \rangle \\ \cdot \\ \cdot \\ \langle f, \varphi_m \rangle \end{pmatrix}$$

Donc, pour calculer les $a_i, i = 1, 2, \dots, m$ (les composantes de notre approximation φ^* , dans la base B), il faut et il suffit de résoudre le système linéaire (S) qui est de la forme $M \times a = b$, où

$$(M)_{i,j} = \langle \varphi_i, \varphi_j \rangle, \forall i, j \in \{1, 2, \dots, m\}, a = (a_1, a_2, \dots, a_m)^t \text{ et } b_i = \langle f, \varphi_i \rangle, \forall i = \overline{1, m}.$$

Comme les $\varphi_i \neq 0, \forall i = 1, 2, \dots, m$, on peut vérifier facilement que le $\det M \neq 0$. Ainsi, le système (S) admet une et une seule solution $a^* = M^{-1} \times b$. L'approximation qu'on cherche à calculer est alors,

$$\varphi^* = \sum_{i=1}^m a_i^* \varphi_i.$$

Remarque 2.7 Si la famille (ou la partie) $B = \{\varphi_1, \varphi_2, \dots, \varphi_m\}$ est orthogonale, alors le système (S) s'écrit comme suit :

$$\begin{pmatrix} \|\varphi_1\|^2 & 0 & 0 & \dots & 0 \\ 0 & \|\varphi_2\|^2 & 0 & \dots & 0 \\ \cdot \\ \cdot \\ 0 & 0 & 0 & \dots & \|\varphi_m\|^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_m \end{pmatrix} = \begin{pmatrix} \langle f, \varphi_1 \rangle \\ \langle f, \varphi_2 \rangle \\ \cdot \\ \cdot \\ \langle f, \varphi_m \rangle \end{pmatrix}$$

ce qui donne : $a_i = \frac{\langle f, \varphi_i \rangle}{\|\varphi_i\|^2}, \forall i = \overline{1, m}$. Par suite, on aura

$$\varphi^* = \sum_{i=1}^m \left(\frac{\langle f, \varphi_i \rangle}{\|\varphi_i\|^2} \right) \times \varphi_i \dots (**)$$

Remarque 2.8 Si la base $B = \{\varphi_1, \varphi_2, \dots, \varphi_m\}$ n'est pas orthogonale, alors la matrice M du système (S) devient pleine (pas beaucoup de zéros), donc elle sera difficile à inverser!. Ainsi, la résolution du système (S) devient très difficile!.

L'orthogonalité de Gram-Schmidt

Si la famille $B = \{\varphi_1, \varphi_2, \dots, \varphi_m\}$ d'un espace préhilbertien $(E, \langle \cdot, \cdot \rangle)$ n'est pas orthogonale, on peut construire une autre famille $\{e_1, e_2, \dots, e_m\} \subset E$, à partir de $\{\varphi_1, \varphi_2, \dots, \varphi_m\}$ tel que $\{e_1, e_2, \dots, e_m\}$ soit orthogonale par rapport au même produit scalaire $\langle \cdot, \cdot \rangle$, comme suit :

$$\begin{cases} e_1 = \varphi_1 \\ \forall i = 2, \dots, m \\ e_i = \varphi_i - \sum_{j=1}^{i-1} \left(\frac{\langle \varphi_i, e_j \rangle}{\|e_j\|^2} \right) \times e_j \end{cases}$$

Exemple 2.9 Soit $E = C([0, 1])$ l'espace muni du produit scalaire :

$$\langle f, g \rangle = \int_0^1 f(x)g(x)dx, \forall f, g \in E$$

et soit $B = \{1, x\}$ la base canonique de $\mathbb{P}_1[X]$. Construire à partir de B une autre base "orthogonale" par rapport à ce produit scalaire ?

Corrigé :

On peut vérifier par un simple calcul que B n'est pas orthogonale par rapport au produit scalaire donné. D'après l'orthogonalité de Gram Schmidt, on peut construire $B' = \{e_1, e_2\}$ une autre base de $\mathbb{P}_1[X]$; tel que B' soit **orthogonale**. En effet, on a

$$\begin{cases} e_1 = 1 \text{ (où } 1 : x \mapsto 1(x) = 1, \forall x \in \mathbb{R}) \\ e_2 = \frac{\langle x, 1 \rangle}{\|1\|^2} = x - \frac{1}{2}. \end{cases}$$

Donc, la partie $\{1, x - \frac{1}{2}\}$ est bien orthogonale, tant que

$$\langle e_1, e_2 \rangle = \int_0^1 1 \times (x - \frac{1}{2})dx = \frac{1}{2} [x^2 - x]_0^1 = 0. \blacksquare$$

3 Meilleure approximation d'une fonction au sens des moindres carrés

a) Le cas continu :

Soit $E = C([a, b])$ et soit le produit scalaire défini par :

$$\forall f, g \in E, \langle f, g \rangle = \int f(x)g(x)w(x)dx$$

où : w est une fonction strictement positive sur $[a, b]$ et intégrable sur $]a, b[$, appelée fonction "poids". On définit sur E la norme suivante :

$$\|f\| = \sqrt{\langle f, f \rangle} = \left(\int_a^b [f(x)]^2 w(x)dx \right)^{1/2} \dots\dots(*)$$

Remarque 3.1 Si la fonction poids n'est pas donnée, alors on prend $w(x) = 1, \forall x \in [a, b]$

Définition 3.2 Soient $f \in E = C([a, b])$ et G un sous espace vectoriel de dimension finie de E . On dit que $\varphi^* \in G$ est une meilleure approximation au sens des moindres carrés (M.A.S.M.C) "le cas continu" de la fonction f , dans G ssi :

$$\|f - \varphi^*\|^2 = \min_{\varphi \in G} \|f - \varphi\|^2,$$

où $\|\cdot\|$ est la norme définie par (*).

3.1 Approximation polynomiale au sens des M.C "le cas continu"

Soit $f \in E = C([a, b])$ et considérons $G = \mathbb{P}_m[X]$. Ici, on cherche à approximer, au sens des moindres carrés, la fonction f par P^* de G . Or l'espace $G = \mathbb{P}_m[X]$ est engendré par la partie $A = \{1, x, x^2, x^3, \dots, x^m\}$ (Puisque A est une base de G , donc elle est génératrice). Par conséquent,

$$\exists a_0, a_1, a_2, \dots, a_m \in \mathbb{R}, \text{ tels que } P^*(x) = \sum_{i=0}^m a_i x^i, \forall x \in [a, b].$$

D'après le Théorème 2.4 (propriété d'orthogonalité), on a $\langle f - P^*, P \rangle = 0, \forall P \in G$. En substituant la valeur de P^* , en tout $x \in [a, b]$, on aura $\langle f - \sum_{i=0}^m a_i x^i, x^j \rangle = 0, \forall j = \overline{0, m}$ (car $(x^j)_{j=\overline{0, m}}$ est une base de G). Ainsi, on obtient $\langle f, x^j \rangle = \sum_{i=0}^m a_i \langle x^i, x^j \rangle, \forall j = \overline{0, m}$. Ce qui donne le système linéaire, de $m + 1$ équations, suivant :

$$\begin{cases} \sum_{i=0}^m a_i \langle x^i, 1 \rangle = \langle f, 1 \rangle \\ \sum_{i=0}^m a_i \langle x^i, x \rangle = \langle f, x \rangle \\ \cdot \\ \cdot \\ \sum_{i=0}^m a_i \langle x^i, x^m \rangle = \langle f, x^m \rangle \end{cases} \dots\dots\dots(\tilde{S})$$

Par définition du produit scalaire, on a : $\langle x^i, x^j \rangle = \int_a^b x^i x^j w(x) dx = \int_a^b x^{i+j} w(x) dx$ et $\langle f, x^j \rangle = \int_a^b x^j f(x) w(x) dx, \forall i, j = 0, 1, \dots, m$. Par conséquent

$$(\tilde{S}) \iff \begin{cases} \sum_{i=0}^m a_i \int_a^b x^i w(x) dx = \int_a^b f(x) w(x) dx \\ \sum_{i=0}^m a_i \int_a^b x^{i+1} w(x) dx = \int_a^b x f(x) w(x) dx \\ \cdot \\ \cdot \\ \sum_{i=0}^m a_i \int_a^b x^{i+m} w(x) dx = \int_a^b x^m f(x) w(x) dx \end{cases}$$

* Si on prend $w \equiv 1$, sur $[a, b]$, alors on trouve

$$(\tilde{S}) \iff \begin{pmatrix} \int_a^b 1 dx & \int_a^b x dx & \int_a^b x^2 dx & \dots & \int_a^b x^m dx \\ \int_a^b x dx & \int_a^b x^2 dx & \int_a^b x^3 dx & \dots & \int_a^b x^{m+1} dx \\ \cdot \\ \cdot \\ \int_a^b x^m dx & \int_a^b x^{m+1} dx & \int_a^b x^{m+2} dx & \dots & \int_a^b x^{2m} dx \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_m \end{pmatrix} = \begin{pmatrix} \int_a^b f(x) dx \\ \int_a^b x f(x) dx \\ \cdot \\ \cdot \\ \int_a^b x^m f(x) dx \end{pmatrix}$$

On a $(\tilde{S}) \iff \tilde{M} \times a = \tilde{b}$, avec $(\tilde{M})_{i,j} = \langle x^i, x^j \rangle, \forall i, j \in \{0, 1, \dots, m\}, a = (a_0, a_1, \dots, a_m)^t$ et $(\tilde{b})_i = \langle f, x^i \rangle, \forall i = \overline{0, m}$.

Comme la matrice \tilde{M} est symétrique, alors elle est semblable à une matrice diagonale dont les coefficients $\langle x^i, x^i \rangle \neq 0, \forall i = 0, 1, \dots, m$ (car les $x^i \neq 0$). Donc, \tilde{M} est inversible et par suite le système (\tilde{S}) admet une solution unique donnée par $a^* = \tilde{M}^{-1} \times (\tilde{b})$. L'approximation polynomiale de f est alors définie par

$$P^*(x) = \sum_{i=0}^m a_i^* x^i, \text{ pour tout } x \in [a, b].$$

Remarque 3.3 1. Le calcul des composantes de la matrice \tilde{M} et du vecteur \tilde{b} nécessite la plupart du temps des intégrations numériques.

2. La base $A = \{1, x, x^2, x^3, \dots, x^m\}$ dite la base canonique de l'espace $\mathbb{P}_m[X]$ conduit à une matrice (\tilde{M}) pleine (du système $\tilde{M} \times a = \tilde{b}$), donc elle sera difficile à inverser !. Ainsi, il est préférable de considérer une autre base pour $\mathbb{P}_m[X]$, de sorte qu'elle soit orthogonale par rapport au produit scalaire considéré.

Exemple 3.4 Soit la fonction $f(x) = x^3$, définie sur $[0, 1]$. Déterminer $P_1^* \in \mathbb{P}_1[X]$; l'approximation de f au sens des moindres carrés.

Corrigé :

f est définie continuellement sur $[0, 1]$ (et non pas dans un nombre fini de points), donc on est dans le cas continu. Comme $P_1^* \in \mathbb{P}_1[X]$, alors $\exists a_0, a_1 \in \mathbb{R}, P_1^*(x) = a_0 + a_1x, \forall x \in [0, 1]$. (ici la base de $G = \mathbb{P}_1[X]$ est $\{1, x\}$). Donc, a_0, a_1 sont solutions du système linéaire suivant :

$$\begin{aligned} (\langle f - P_1^*, X^j \rangle = 0, j = \overline{0, 1}) &\iff \begin{cases} \langle P_1^*, 1 \rangle = \langle f, 1 \rangle \\ \langle P_1^*, X \rangle = \langle f, X \rangle \end{cases} \\ &\iff \begin{cases} \int_0^1 (a_0 + a_1x) dx = \int_0^1 f(x) dx \\ \int_0^1 x(a_0 + a_1x) dx = \int_0^1 xf(x) dx \end{cases} \\ &\iff \begin{pmatrix} \int_0^1 1 dx & \int_0^1 x dx \\ \int_0^1 x dx & \int_0^1 x^2 dx \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \int_0^1 x^3 dx \\ \int_0^1 x^4 dx \end{pmatrix} \\ &\iff \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 1/4 \\ 1/5 \end{pmatrix} \\ &\iff \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} -1/5 \\ 9/10 \end{pmatrix} \end{aligned}$$

Ainsi, l'approximation de f au sens des moindres carrés (le cas "continu"), dans l'espace $\mathbb{P}_1[X]$ est : $P_1^*(x) = \frac{9}{10}x - \frac{1}{5}$. ■

Remarque 3.5 1. La fonction f à approximer est en général une fonction quelconque (pas nécessairement un polynôme), mais pour simplifier les calculs, on a pris $f \in \mathbb{P}_3[X]$ dans l'exemple précédent.

2. La base $\{1, x\}$ de $\mathbb{P}_1[X]$ n'est pas orthogonale, donc la matrice du système précédent n'était pas diagonale. Par la méthode d'orthogonalité de Gram Schmidt, on peut obtenir $\{1, x - \frac{1}{2}\}$ comme base orthogonale de $\mathbb{P}_1[X]$ (voir l'exercice 1). Ainsi, le calcul de P_1^* peut se faire aussi comme suit :

On pose $P_1^*(x) = b_0 + b_1(x - \frac{1}{2}), \forall x \in [0, 1]$. D'après la relation (**) (page 19), on trouve $b_0 = \frac{\langle f, e_0 \rangle}{\|e_0\|^2} = \frac{1}{4}$ et $b_1 = \frac{\langle f, e_1 \rangle}{\|e_1\|^2} = \frac{9}{5}$. Ainsi,

$$P_1^*(x) = b_0 + b_1(x - \frac{1}{2}) = \frac{1}{4} + \frac{9}{5}(x - \frac{1}{2}) = \frac{9}{10}x - \frac{1}{5}, \forall x \in [0, 1].$$

b) Le cas discret :

Définition 3.6 Soit f une fonction dont on connaît ses valeurs $f(x_i)_{i=\overline{0,n}}$. On dit que φ^* est une approximation au sens des moindres carrés " le cas discret " de la fonction f dans un espace vectoriel G , ssi :

$$\|f - \varphi^*\|^2 = \min_{\varphi \in G} \|f - \varphi\|^2,$$

où $\|\cdot\|$ est la norme définie par :

$$\|f\| = \sqrt{\langle f, f \rangle} = \left(\sum_{i=0}^n [f(x_i)]^2 w(x_i) \right)^{\frac{1}{2}}$$

* Cette norme est définie à partir du produit scalaire suivant :

$$\langle f, g \rangle = \sum_{i=0}^n f(x_i)g(x_i)w(x_i),$$

où : f, g sont deux fonctions définies discrètement (c.à.d; définies en un nombre fini de points $(x_i)_{i=\overline{0,n}}$) et w est la fonction " Poids " (strictement positive).

Remarque 3.7 Comme dans le cas continu, si la fonction poids w n'est pas donnée, alors on la considère $w \equiv 1$.

3.2 Approximation polynomiale au sens des M.C "le cas discret"

Dans ce cas particulier, nous prenons $G = \mathbb{P}_m[X]$.

L'approximation au sens des moindres carrés d'une fonction f , définie discrètement en $(x_i)_{i=\overline{0,n}}$ (à condition que $m \leq n$), par un polynôme $\hat{P} \in G$ se fait comme suit :

* Pour tout $x \in [x_0, x_n]$, on pose

$$\hat{P}(x) = \sum_{i=0}^m a_i x^i, \quad a_0, a_1, a_2, \dots, a_m \in \mathbb{R},$$

tant que $A = \{1, x, x^2, x^3, \dots, x^m\}$ est une base de $G = \mathbb{P}_m[X]$.

D'après la propriété d'orthogonalité (Théorème 2.4), on a $\langle f - \hat{P}, P \rangle = 0, \forall P \in G$. Ce qui donne : $\langle f - \sum_{i=0}^m a_i x^i, x^j \rangle = 0, \forall j = \overline{0, m}$ (car $(x^j)_{j=\overline{0, m}}$ est une base de G). Ainsi, on obtient

$$\langle f, x^j \rangle = \sum_{i=0}^m a_i \langle x^i, x^j \rangle, \forall j = \overline{0, m}.$$

i.e.

$$\sum_{i=0}^n f(x_i)(x_i)^j w(x_i) = \sum_{i=0}^m a_i \sum_{j=0}^n x_j^{i+j}, \forall j = \overline{0, m} \dots (R)$$

En considérons $w \equiv 1$, ceci donne le système linéaire, de $m + 1$ équations, suivant :

$$(R) \iff \begin{pmatrix} \sum_{j=0}^n 1 & \sum_{j=0}^n x_j & \sum_{j=0}^n x_j^2 & \dots & \sum_{j=0}^n x_j^m \\ \sum_{j=0}^n x_j & \sum_{j=0}^n x_j^2 & \sum_{j=0}^n x_j^3 & \dots & \sum_{j=0}^n x_j^{m+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{j=0}^n x_j^m & \sum_{j=0}^n x_j^{m+1} & \sum_{j=0}^n x_j^{m+2} & \dots & \sum_{j=0}^n x_j^{2m} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^n f(x_i) \\ \sum_{i=0}^n x_i f(x_i) \\ \vdots \\ \sum_{i=0}^n x_i^m f(x_i) \end{pmatrix}$$

* Le système d'équations linéaires (R) est appelé **Système d'équations normales**. Il admet une solution unique $a^* = (a_0^*, a_1^*, \dots, a_m^*)^t \in \mathbb{R}^{m+1}$ (puisque c'est un cas particulier du système (S)). Par suite, l'approximation de f sera $\hat{P}(x) = \sum_{i=0}^m a_i^* x^i, \forall x \in [x_0, x_n]$. Ainsi, on a approximé une fonction définie discrètement (en $n+1$ points) par un polynôme de degré $\leq m$ (avec $m \leq n$).

Exemple 3.8 Soit f connue en $x_0 = 0, x_1 = 1$ et $x_2 = 2$, avec

$$y_0 = f(x_0) = 1, y_1 = f(x_1) = 0, y_2 = f(x_2) = 2.$$

Déterminer la meilleure approximation de f au sens des moindres carrés, dans l'espace $\mathbb{P}_1[X]$, relativement à la base canonique de cet espace ?

Corrigé :

Ici la fonction f à approximer n'est pas définie continument, mais seulement discrètement, aux points x_0, x_1, x_2 . Donc, on est dans le cas discret ("le système (R) ").

Comme notre approximation est dans $\mathbb{P}_1[X]$ avec la base canonique $\{1, x\}$, alors on la note par P_1 avec $P_1(x) = a_0 + a_1x, \forall x \in [x_0, x_2] = [0, 2]$.

Dans cet exemple, on a $m = 1 \leq n = 2$, d'après le Théorème 2.4 (la propriété d'orthogonalité), les coefficients a_0, a_1 du polynôme P_1 sont solutions d'un système linéaire (de $m + 1 = 2$ équations) suivant :

$$\begin{pmatrix} \sum_{j=0}^2 1 & \sum_{j=0}^2 x_j \\ \sum_{j=0}^2 x_j & \sum_{j=0}^2 x_j^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^2 f(x_i) \\ \sum_{i=0}^2 x_i f(x_i) \end{pmatrix} \iff \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

La résolution de ce système donne : $a = (\frac{1}{2}, \frac{1}{2})^t$. D'où $P_1(x) = \frac{1}{2}(x + 1)$. ■

4 Exercices

"Approximation au sens des moindres carrés"

Exercice 1.

A partir de la base canonique $\{1, x, x^2\}$ construire une base orthogonale pour l'espace vectoriel $\mathbb{P}_2[X]$, en utilisant l'algorithme de Gram Schmidt.

Exercice 2. Trouver le polynôme P^* qui réalise le minimum suivant :

$$\min_{P \in \mathbb{P}_2[X]} \left(\int_{-1}^1 [P(x) - |x|]^2 dx \right).$$

Indication : utiliser la base orthogonale obtenue dans l'exercice 1.

Exercice 3.

Soit $P^* \in \mathbb{P}_m[X]$ le meilleur approximant, au sens des moindres carrés, de la fonction f dont on connaît les valeurs $\{f(x_i)\}_{i=0}^n$.

Montrer que si $m = n$ alors, P^* n'est que l'interpolant de f aux points $(x_i)_{i=0}^n$.

Exercice 4. À partir des données suivantes :

x_i	1	1.1	1.8	2.2	2.5	3.5	3.7	4
$g(x_i)$	6.008	5.257	9.549	11.098	15.722	27.130	28.828	33.772

1. Calculer la droite (respectivement la parabole) qui réalise la M.A.S.M.C de g .
2. Calculer les erreurs obtenues dans les deux cas. Que remarque t-on ?
3. Justifier géométriquement les résultats obtenus.

Exercice 5.

Soit f une fonction passant par les points suivants :

$(-2, 17); (-1, 4); (0, 3); (1, 8)$ et $(2, 61)$. Relativement à ces points faire :

1. Déterminer le polynôme $P^* \in \mathbb{P}_2[x]$ de la M.A.S.M.C discrète de f .
2. Donner la table des différences divisées de f et en déduire le polynôme interpolant f .
3. Calculer $f(0.5)$ par ces deux approximations. Conclure.

Exercice 6.

Soit la fonction $f(x) = x^3 - 1$ définie sur $[-1, 1]$.

1. Déterminer P l'interpolant de f aux points $\{-1, -0.5, 0.5, 1\}$.
2. Déterminer $P'_3 \in \mathbb{P}_3[X]$: la M.A.S.M.C discrète de P relativement aux points $\{-1, -0.5, 0.5, 1\}$.
3. Déterminer $P^*_3 \in \mathbb{P}_3[X]$: la M.A.S.M.C continue de P'_3 . Conclure.

Chapitre 5

Intégration numérique

1 Introduction

Soit f une fonction intégrable sur $[a, b]$; $a, b \in \mathbb{R}$. On pose

$$I(f) = \int_a^b f(x) dx.$$

Le problème d'intégration numérique d'une fonction (donnée explicitement ou connue seulement en un nombre fini de points) consiste à **approcher** son intégrale $I(f)$ par la somme suivante :

$$J(f) = \sum_{i=0}^n C_i f(x_i), \dots (5.1)$$

où les coefficients $C_i \in \mathbb{R}, i = \overline{0, n}$ sont appelés les poids, les points $x_i \in [a, b], i = \overline{0, n}$ sont dites les noeuds.

On écrit

$$I(f) = J(f) + R(f),$$

avec $R(f)$ est l'erreur de l'intégration numérique.

Les méthodes d'intégration numériques se répartissent en deux grandes catégories :

1. Méthodes interpolatoires.

Dans ces méthodes, la fonction f (à intégrer) est remplacée par son polynôme d'interpolation. Donc, les noeuds $(x_i)_i$ sont connus (fixés), et on cherche la valeur de chaque poids $C_i, i = 0, 1, \dots, n$.

2. Méthodes de Gauss.

Dans ces méthodes, on utilise des polynômes orthogonaux pour lesquels les points de subdivision de l'intervalle $[a, b]$ (c.à.d, les noeuds) sont imposés.

Définition 1.1 "Formule de Quadrature"

Le calcul numérique d'une intégrale simple s'appelle quadrature mécanique, et toute formule correspondante à ce calcul s'appelle **formule de quadrature**.

Remarque 1.2 La formule définie dans (5.1) est une formule de quadrature.

Remarque 1.3 Pour obtenir une formule de quadrature, il suffit de remplacer la fonction f par son approximation 'h' qui doit être facile à intégrer, ce qui est le cas si la fonction 'h' est un polynôme.

Si $h \in G$, où G est un espace vectoriel engendré par la famille $\{\varphi_1, \varphi_2, \dots, \varphi_m\}$, alors on aura

$$I(h) = \sum_{i=0}^n a_i \int_a^b \varphi(x), \quad (1.1)$$

où les $(a_i)_{i=0}^n$ sont les coefficients du polynôme h et $I(f)$ est dite "formule de quadrature", réalisant une approximation de $I(f)$. On écrit :

$$I(f) \approx I(h) \text{ ou bien } I(f) = I(h) + R(f),$$

où $R(f)$ est l'erreur de la formule (1.1).

Définition 1.4 "Formule de Quadrature Composée"

Si on décompose (ou subdivise) l'intervalle de l'intégration en un nombre fini de sous-intervalles, puis on utilise sur chaque sous-intervalle une formule de quadrature, alors la formule induite est dite : formule de quadrature composée ou composite.

Définition 1.5 "Degré de précision"

Soit $m \in \mathbb{N}$. On dit que m est le degré de précision (ou d'exactitude) d'une formule de quadrature si : $R(p_k) = 0, \forall p_k \in \mathbb{P}_m[X]$, et pour au moins un $p \in \mathbb{P}_{m+1}[X]$, on a $R(p) \neq 0$, où R est l'erreur de cette formule de quadrature.

Remarque 1.6 Si le degré de précision d'une formule de quadrature est plus élevé, on dit que cette formule est plus précise (ou exacte).

Théorème 1.7 Le degré de précision d'une formule est égale à m , ssi on a

$$R(x^k) = 0, \forall k = \overline{0, m} \text{ et } R(x^{m+1}) \neq 0,$$

avec R : l'erreur de la formule de quadrature considérée.

Preuve : Exercice.

Définition 1.8 "Ordre d'une formule de quadrature"

Soit $m \in \mathbb{N}$. Une formule de quadrature est dite de degré m si son degré de précision est égale à ce m .

Définition 1.9 "Convergence d'une formule de quadrature"

On dit qu'une formule de quadrature $J(f) = \sum_{i=0}^n C_i f(x_i)$ est convergente vers la valeur exacte de l'intégrale ssi,

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^n C_i f(x_i) = \int_a^b f(x) dx.$$

2 Méthodes d'intégration numérique

Méthodes interpolatoire : Dans ces méthodes, on remplace f la fonction à intégrer par ' P_n ' son polynôme d'interpolation, aux points donnés $(x_i)_{i=\overline{0,n}} \subset [a, b]$, pour obtenir une formule de quadrature de la forme

$$J(f) = \sum_{i=0}^n C_i A_i. \quad (2.1)$$

Exemples :

- Si on interpole f avec la méthode de Newton, on aura, pour tout $x \in [a, b]$,

$$P_n(x) = \sum_{i=0}^n N_i(x) f[x_0, x_1, \dots, x_i].$$

Par suite,

$$J(f) = \int_a^b P_n(x) dx = \sum_{i=0}^n C_i f[x_0, x_1, \dots, x_i], \quad (2.2)$$

avec $C_i = \int_a^b N_i(x) dx$.

* L'intérêt ici est que les coefficients $C_i (i = \overline{0, n})$ ne dépendent pas du nombre de points d'interpolation, c.à.d, on peut calculer de nouveaux termes de la somme (2.2), en ajoutant d'autres points "sans modifier les coefficients précédents". Ceci n'est pas le cas, si on interpole f avec la méthode de Lagrange.

- Si l'interpolation de f est faite par la méthode de Lagrange, alors on a pour tout $x \in [a, b]$, $P_n(x) = \sum_{i=0}^n f(x_i) \times L_i(x)$, ainsi, on obtient

$$J(f) = \int_a^b P_n(x) dx = \sum_{i=0}^n C_i f(x_i), \quad (2.3)$$

où $C_i = \int_a^b L_i(x) dx$,

où les $(L_i)_i$ sont les polynômes de base de Lagrange relatifs aux points $(x_i)_i$.

* L'intérêt ici (ou l'avantage) est que la formule (2.3) est sous la formule généralisée (2.2), où les seuls inconnus sont les coefficients $C_i (i = \overline{0, n})$. Par contre dans (2.2), les inconnus sont les $C_i (i = \overline{0, n})$ et les $f[x_0, x_1, \dots, x_i]$.

Théorème 2.1 *Toute formule de quadrature interpolatoire, utilisant $(n + 1)$ noeuds distincts, admet un degré de précision aux moins égal à 'n' (c.à.d, le degré $\geq n$).*

Preuve : Exercice.

Exemples de formules de quadrature interpolatoires

2.1 Méthode de Newton-Côtes

Ce sont les méthodes interpolatoires par la méthode de Lagrange, où les points d'interpolation $(x_i)_{i=0,n}$ sont équidistants. c.à.d, où $x_i = x_0 + ih$, pour tout $i = 1, \dots, n$, avec $h = \frac{x_n - x_0}{n}$: c'est le pas du déplacement, $n \in \mathbb{N}^*$: c'est le nombre des sous-intervalles de $[x_0, x_n]$.

Si on pose $P_n(x) = \sum_{i=0}^n f(x_i)L_i(x)$: l'interpolant de Lagrange de f aux points équidistants $(x_i = x_0 + ih, \forall i = 0, \dots, n)$, on aura pour tout $i = 0, \dots, n$ et $x \in \mathbb{R}$,

$$L_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^n \frac{(x - x_k)}{(x_i - x_k)} \\ = (-1)^{n-i} \frac{\prod_{k=0}^n (t - k)}{(t - i)i!(n - i)!}.$$

Ce qui donne, pour tout $i = 0, \dots, n$,

$$C_i = \int_{x_0}^{x_n} L_i(x)dx = \frac{(-1)^{n-i}h}{i!(n-i)!} \int_0^n \frac{\prod_{j=0}^n (t-j)}{(t-i)} dt.$$

Remarque

Pour vérifier la formule précédente, il suffit de poser $x = x_0 + th$ (et que $dx = h \times dt$).

Ainsi, il en résulte (de la relation (2.3)), la formule de **Newton-Côtes** :

$$J(f) = (b - a) \sum_{i=0}^n H_i f(x_i) \dots \dots (**)$$

où $b = x_n, a = x_0$ et les " $H_i = \frac{C_i}{b-a}$ ", $i = \overline{0, n}$ (H_i sont appelés les coefficients de Côtes).

2.2 Méthode du trapèze

Cette méthode est un cas particulier de celle de Newton-Côtes, où $n = 1$. Donc, d'après la formule de Newton-Côtes, on obtient :

$$J_t(f) = (b - a) (H_0 f(x_0) + H_1 f(x_1)),$$

avec $H_0 = - \int_0^1 \frac{t(t-1)}{t} dt = \frac{1}{2}$ et $H_1 = \int_0^1 \frac{t(t-1)}{t-1} dt = \frac{1}{2}$. D'où,

$$J_t(f) = \frac{(b-a)}{2} [f(a) + f(b)].$$

* Si $f \in C^2([a, b])$, l'erreur (ou le reste) de cette méthode se donne par

$$R_t(f) = -\frac{h^3}{12} f^{(2)}(\xi), \quad \xi \in [a, b],$$

avec $h = b - a$.

Ainsi, l'ordre de la méthode du trapèze (simple) est bien $m = 1$.

2.3 Méthode des trapèzes (trapèze généralisée)

Pour minimiser l'erreur de la formule du trapèze, on peut subdiviser l'intervalle de l'intégration $[a, b]$ en ' n ' parties égales $[x_0, x_1], \dots, [x_{n-1}, x_n]$, puis appliquer sur chacune d'elles la formule du trapèze simple. Posons $h = \frac{b-a}{n}$, alors il vient que :

$$\begin{aligned} \int_a^b f(x)dx &= \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \dots + \int_{x_{n-1}}^{x_n} f(x)dx \\ &\simeq \frac{h}{2} [(f(x_0) + f(x_1)) + (f(x_1) + f(x_2)) + \dots + (f(x_{n-1}) + f(x_n))]. \end{aligned}$$

Donc, la formule du trapèze composée se donne par :

$$J_T(f) = \frac{h}{2} [f(x_0) + f(x_n) + 2 \sum_{i=1}^{n-1} f(x_i)],$$

où $h = (b - a)/2$.

* Pour $f \in \mathbb{C}^2([a, b])$, le reste (ou l'erreur) de cette formule est le suivant :

$$R_T(f) = \frac{-nh^3}{12} f^{(2)}(\xi) = \frac{-(b-a)h^2}{12} f^{(2)}(\xi) = \frac{-(b-a)^3}{12n^2} f^{(2)}(\xi),$$

pour $\xi \in [a, b]$.

Exemple 2.2 Déterminez une valeur approximative de $I_1 = \int_0^1 e^{-x^2} dx$, en utilisant les méthodes du trapèze (simple) et trapèze composée ($n = 2$), puis estimer l'erreur d'intégration de chaque méthode.

Corrigé :

a) Trapèze simple.

$$I_1 = \int_0^1 e^{-x^2} dx \simeq \frac{(b-a)}{2} [f(a) + f(b)] = \frac{1}{2}(e^{-1} + e^{-0}) = \frac{e^{-1} + 1}{2} \simeq \mathbf{0,6839}.$$

*Estimons l'erreur théorique :

$$R_1(f) = -\frac{h^3}{12} f^{(2)}(\xi), \quad \xi \in [a, b] \text{ et } h = b - a = 1. \text{ D'où}$$

$$|R_1(f)| \leq \frac{1}{12} \sup_{0 \leq \xi \leq 1} |f^{(2)}(\xi)|.$$

Or $f(x) = e^{-x^2}$ et $f''(x) = 2e^{-x^2}(2x^2 - 1)$, alors $\sup_{0 \leq \xi \leq 1} |f^{(2)}(\xi)| = |f''(1)| = 0,7358$. Ainsi $|R_1(f)| \leq \frac{0,7358}{12} \simeq \mathbf{0,0613}$.

b) Trapèze composée ($n = 2$).

$$I_1 \simeq \frac{h}{2} \left[f(a) + f(b) + 2 \sum_{i=1}^{n-1} f(x_i) \right], \quad h = (b - a)/n.$$

Par suite, $I_1 \simeq \frac{1}{4} [f(0) + f(1) + 2f(\frac{1}{2})]$, puisque $h = (b - a)/2 = \frac{1}{2}$, $x_0 = a = 0$, $x_1 = x_0 + h = \frac{1}{2}$ et $x_2 = b = 1$. Ainsi

$$I_1 \simeq \frac{1}{2} \left[e^{-0} + e^{-1} + 2e^{-\frac{1}{4}} \right] \simeq \mathbf{0,7319}.$$

* Estimons l'erreur théorique :

$$|R_2(f)| = \left| -\frac{nh^3}{12} f^{(2)}(\xi) \right| \leq \left| -\frac{2(1/2)^3}{12} 0,7358 \right| \simeq 0,0153.$$

Remarquons que la méthode des trapèzes ($n = 2$) est plus précise que celle du trapèze simple, puisque $R_2(f) = 0,0153 \leq R_1(f) = 0,0613$. ■

2.4 Méthode de Simpson

Cette formule est aussi un cas particulier de celle de Newton-Cotes, où $n = 2$. Donc, en utilisant la relation (**), on obtient

$$J_s(f) = \frac{b-a}{6} (f(a) + 4f(\frac{a+b}{2}) + f(b)) \simeq \int_a^b f(x)dx,$$

ou bien

$$J_s(f) = \frac{h}{3} (f(a) + 4f(\frac{a+b}{2}) + f(b)),$$

avec $h = \frac{b-a}{2}$.

2.5 Méthode de Simpson généralisée (ou composée)

D'une manière analogue à la méthode du trapèze, on peut généraliser la formule de Simpson, en subdivisant l'intervalle $[a, b]$ à $n = 2k$ intervalles partiels égaux ($k \in \mathbb{N}^*$) et 'n' doit être paire, puis on applique la formule de Simpson (simple), sur chacun de ces intervalles :

$$[x_0, \underbrace{x_0+2h}_{x_2}], [\underbrace{x_0+2h}_{x_2}, \underbrace{x_0+4h}_{x_4}], \dots, [\underbrace{x_0+(n-2)h}_{x_{n-2}}, x_n],$$

avec $x_0 = a, x_n = b$ et $h = \frac{b-a}{n}$.

$$\begin{aligned} \int_a^b f(x)dx &= \int_{x_0}^{x_2} f(x)dx + \int_{x_2}^{x_4} f(x)dx + \dots + \int_{x_{n-2}}^{x_n} f(x)dx \\ &\simeq \frac{h}{3} [f(x_0) + 4f(x_1) + f(x_2)] + \dots + \frac{h}{3} [f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)]. \end{aligned}$$

Ce qui donne,

$$J_S(f) \simeq \frac{h}{3} \left[f(x_0) + f(x_n) + 4 \times \sum_{i=1}^{n-1} (i:\text{impair}) f(x_i) + 2 \times \sum_{i=2}^{n-2} (i:\text{pair}) f(x_i) \right].$$

L'erreur d'intégration. Si $f \in \mathbb{C}^4([a, b])$, alors

1. L'erreur de Simpson simple se donne par :

$$R_s(f) = \frac{-h^5}{90} f^{(4)}(\xi),$$

où $h = (b - a)/2$.

2. L'erreur de Simpson généralisée est :

$$R_S(f) = \frac{-kh^5}{90} f^{(4)}(\xi)$$

ou encore,

$$R_S(f) = \frac{-(b-a)h^4}{180} f^{(4)}(\xi) = \frac{-(b-a)^5}{180n^4} f^{(4)}(\xi).$$

3 Exercices

"Intégration numérique"

Exercice 1. Soient :

$$I_1 = \int_0^1 e^{-x^2} dx, \quad I_2 = \int_0^\pi \sin x dx.$$

1. Déterminer une valeur approximative de I_1 , en utilisant la méthode du trapèze, puis par celle du trapèze généralisée ($n = 2$). Estimer l'erreur théorique à chaque fois.
2. Calculer l'intégrale I_2 par la méthode de Simpson, puis par celle de Simpson composite, où le nombre de subdivisions de l'intervalle $[0, \pi]$ est $n = 4$. Estimer l'erreur théorique de chaque méthode.
3. Comparer les résultats obtenus dans (2) avec la valeur exacte de l'intégrale I_2 .

Exercice 2.

Calculer 'n' le nombre de subdivisions nécessaires de l'intervalle d'intégration $[-\pi, \pi]$, pour évaluer à 0.5×10^{-3} près, grâce à la méthode de Simpson, l'intégrale : $I = \int_{-\pi}^{\pi} \cos x dx$.

Exercice 3.

Déterminer 'n' le nombre de subdivisions nécessaires des intervalles d'intégration pour évaluer à 0.5×10^{-6} près, par la méthode des trapèzes, les intégrales :

$$J_1 = \int_1^0 \frac{dx}{1 + e^x}, \quad J_2 = \int_1^0 \frac{dx}{1 + \sin x}.$$

Exercice 4.

Soit $f : [-1, 1] \rightarrow \mathbb{R}$ une fonction continue.

1. Calculer les coefficients des formules de quadrature suivantes :

$$\int_{-1}^1 f(x) dx \simeq \lambda_0 f\left(-\frac{1}{3}\right) + \lambda_1 f\left(\frac{1}{3}\right), \text{ exacte pour } f \in \mathbb{P}_1[X], \quad (1)$$

$$\int_{-1}^1 f(x) dx \simeq \lambda_0 f\left(-\frac{3}{5}\right) + \lambda_1 f\left(-\frac{1}{5}\right) + \lambda_2 f\left(\frac{1}{5}\right) + \lambda_3 f\left(\frac{3}{5}\right), \text{ exacte pour } f \in \mathbb{P}_3[X].$$

2. Donner la formule (1) pour un intervalle d'intégration $[a, b]$ quelconque de \mathbb{R} , en utilisant le changement de variable suivant :
 $\varphi : [0, 1] \rightarrow [a, b]; \varphi(x) = (b - a)x + a.$
3. Construire la formule composée correspondante à la première question avec des noeuds équidistants.

Chapitre 6

Dérivation numérique

1 Introduction

Comme pour l'intégrale, on voudrait être en mesure d'évaluer numériquement la dérivée d'une fonction difficile à manipuler, ou qui n'est donnée que par un ensemble discret de valeurs (données expérimentales par exemple), ou bien intervenant dans des équations différentielles. Ce problème de dérivation numérique est très commun en ingénierie et en analyse numérique (c'est la base des méthodes de différences finies).

On considère une fonction $f : [a, b] \rightarrow \mathbb{R}$ de classe suffisamment élevée, $x \in]a, b[$ fixé. On veut approcher (au mieux!) les dérivées de la fonction f au point x .

Or, la dérivée première de f en x est définie par

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

Donc, une idée simple pour calculer numériquement la 1^{ère} dérivée de f , au point x , consiste à se donner une valeur $|h| \ll 1$, et à calculer l'une des quantités $\nabla_h f(x)$; $\bar{\nabla}_h f(x)$ ou $\delta_{2h} f(x)$, avec

$$\nabla_h f(x) = f(x+h) - f(x),$$

$$\bar{\nabla}_h f(x) = f(x) - f(x-h),$$

$$\delta_{2h} f(x) = f(x+h) - f(x-h).$$

Une des méthodes les plus anciennes utilisées pour obtenir des formules de dérivation "numérique" consiste à construire des quotients différentiels à l'aide des développements de Taylor.

Définition 1.1 *i) On définit l'erreur commise lorsqu'on remplace la dérivée de f en x , par l'une des formules d'approximation ci-dessous, par*

$$E_h = f'_{\text{exacte}}(x) - f'_{\text{approximative}}(x).$$

ii) On dit que la formule de dérivation est d'ordre n , s'il existe une constante $C > 0$, dépendant de f telle que,

$$|E_h| < Ch^n,$$

et on écrira fréquemment de façon abrégée $|E_h| \sim O(h^n)$.

iii) La constante C est dite "facteur de convergence de la méthode".

2 Approximation de la première dérivée

2.1 Formules à deux points

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction suffisamment dérivable sur $[a, b]$. Fixons $x \in]a, b[$ et $h > 0$, tels que $x - h, x + h \in]a, b[$.

- Le développement de Taylor d'ordre 1 de f , à droite de x , donne

$$f(x + h) = f(x) + f'(x)h + \frac{1}{2}f''(\xi)h^2,$$

où $\xi \in]x, x + h[$. Ainsi,

$$f'(x) = \frac{f(x + h) - f(x)}{h} + \frac{h}{2}f''(\xi),$$

ce qui donne la **formule de différence progressive** :

$$f'(x) \simeq f'_{hd}(x) = \frac{f(x+h)-f(x)}{h} = \frac{\nabla_h f(x)}{h} \dots\dots \textbf{(FDP)}$$

avec une erreur d'approximation,

$$E_1 = -\frac{h}{2}f''(\xi).$$

- Le développement de Taylor d'ordre 1 de f , à gauche de x , donne

$$f(x - h) = f(x) - f'(x)h + \frac{1}{2}f''(\xi)h^2,$$

avec $\xi \in]x - h, x[$. Ce qui donne la **formule de différence régressive** :

$$f'(x) \simeq f'_{hg}(x) = \frac{f(x)-f(x-h)}{h} = \frac{\bar{\nabla}_h f(x)}{h} \dots\dots \textbf{(FDR)}$$

dont l'erreur d'approximation est :

$$E_2 = \frac{h}{2}f''(\xi).$$

Remarque 2.1 Les formules obtenues **(FDP)** et **(FDR)** sont deux approximations **d'ordre 1** de la dérivée première ; dont l'erreur commise, à chaque fois, E_1 et E_2 tendent vers 0 quand h tend vers 0.

- Soient, maintenant, $h_1, h_2 > 0$ tels que $x + h_1, x - h_2 \in]a, b[$. Augmentons l'ordre du développement de Taylor de f autour de x , pour avoir

$$f(x + h_1) = f(x) + f'(x)h_1 + \frac{1}{2}f''(\xi)h_1^2 + \frac{h_1^3}{6}f'''(\xi_1), \quad \xi_1 \in]x, x + h_1[,$$

$$f(x - h_2) = f(x) - f'(x)h_2 + \frac{1}{2}f''(\xi)h_2^2 - \frac{h_2^3}{6}f'''(\xi_2), \quad \xi_2 \in]x - h_2, x[.$$

Par soustraction, on trouve

$$f(x + h_1) - f(x - h_2) = f'(x)(h_1 + h_2) + \frac{1}{6} [h_1^3 f'''(\xi_1) + h_2^3 f'''(\xi_2)],$$

ce qui donne la **formule de différence centrée** :

$$f'(x) \simeq \frac{h_2}{h_1(h_1+h_2)} f(x + h_1) - \frac{h_1}{h_2(h_1+h_2)} f(x - h_2) + \left(\frac{1}{h_2} - \frac{1}{h_1} \right) f(x).$$

Le cas de points équidistants :

Quand on est dans le cas où $h_1 = h_2 = h$, la formule de différence centrée se simplifie notablement, pour donner la **formule centrée-points équidistants** :

$$f'(x) \simeq f'_{hc}(x) = \frac{f(x+h) - f(x-h)}{2h} = \frac{\delta_{2h} f(x)}{2h} \dots\dots \textbf{(FDC)}$$

avec une erreur commise,

$$E = -\frac{h^2}{12} [f'''(\xi_1) + f'''(\xi_2)] = -\frac{h^2}{6} f'''(\xi), \quad \xi \in]x - h, x + h[.$$

Remarque 2.2 1) Les deux formules ci-dessus de différence centrée sont des formules d'approximation **d'ordre 2** et par suite, plus précises que les deux premières formules décentrées (**FDP**) et (**FDR**).

2) La formule centrée (**FDC**) n'est autre que la moyenne des deux formules décentrées, dans le cas de points équidistants.

3) La dérivation numérique est une opération très instable, c'est à dire très sensible aux erreurs d'arrondi (soustraction entre des termes voisins).

Exemple 2.3 Pour illustrer les formules ci-dessus, considérons les données suivantes :

$$(x_0, f(x_0)) = (1, 7/2), (x_1, f(x_1)) = (2, 3) \text{ et } (x_2, f(x_2)) = (3, 3/2).$$

Estimer la valeur de $f'(2)$, en utilisant les formules ci-dessus, puis évaluer l'erreur commise à chaque fois.

Corrigé :

1) La formules de différence progressive (**FDP**) :

$$f'(2) \simeq \frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{\frac{3}{2} - 3}{3 - 2} = -\frac{3}{2},$$

2) La formules de différence régressive (**FDR**) :

$$f'(2) \simeq \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{3 - \frac{7}{2}}{2 - 1} = -\frac{1}{2},$$

3) La formules de différence centrée (**FDC**) :

$$f'(2) \simeq \frac{f(x_2) - f(x_0)}{x_2 - x_0} = \frac{\frac{3}{2} - \frac{7}{2}}{3 - 1} = -1.$$

*Les données ont été calculées pour la fonction $f : x \mapsto -\frac{1}{12}x^3 + \frac{1}{12}x + \frac{7}{2}$, donc, on a $f'(x) = -\frac{1}{4}x^2 + \frac{1}{12}$ et $f'(2) = -\frac{11}{12}$.

* Evaluons les erreurs d'approximation commise, avec chaque formule.

$$|E_1| = |f'(2) - f'_{hd}(2)| = \left| -\frac{11}{12} + \frac{3}{2} \right| = \frac{7}{12},$$

$$|E_2| = |f'(2) - f'_{hg}(2)| = \left| -\frac{11}{12} + \frac{1}{2} \right| = \frac{5}{12},$$

$$|E| = |f'(2) - f'_{hc}(2)| = \left| -\frac{11}{12} + 1 \right| = \frac{1}{12}.$$

Ainsi, la meilleure formule pour approcher $f'(2)$ est celle de **(FDC)**. ■

Remarque 2.4 Les formules à deux points peuvent également être obtenues à partir du polynôme d'interpolation de Lagrange.

Etant donné deux points $(x_0, f(x_0))$ et $(x_1, f(x_1))$. Le polynôme de degré 1 passant par ces deux points est :

$$P_1(x) = \frac{x - x_1}{x_0 - x_1} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1).$$

Une approximation de la dérivée f' , en tout point x , peut alors être obtenue en dérivant directement l'interpolant.

$$f'(x) \simeq P'_1(x) = \frac{1}{x_0 - x_1} f(x_0) + \frac{1}{x_1 - x_0} f(x_1).$$

En choisissant $(x_0 = x, x_1 = x + h)$; $(x_1 = x, x_0 = x - h)$ et $(x_1 = x + h, x_0 = x - h)$, on obtient les approximations **(FDP)**, **(FDR)** et **(FDC)**, respectivement.

2.2 Formules à trois points

Pour approximer la dérivée de f en un point $x \in]a, b[$, il est possible de développer d'autres formules, en effectuant un développement de Taylor de f , autour de x , avec un pas $2h$.

*Par exemple,

$$f(x + h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{6} f'''(\xi_1), \quad \xi_1 \in]x, x + h[, \quad (2.1)$$

$$f(x + 2h) = f(x) + 2hf'(x) + 2h^2 f''(x) + \frac{4}{3} h^3 f'''(\xi_2), \quad \xi_2 \in]x, x + 2h[. \quad (2.2)$$

En combinant les relations (2.1) et (2.2), on obtient

$$f'(x) = \frac{4f(x + h) - f(x + 2h) - 3f(x)}{2h} + h^2 \left(\frac{2}{3} f'''(\xi_1) - \frac{1}{3} f'''(\xi_2) \right). \quad (2.3)$$

Si f''' (la dérivée du troisième ordre) est continue sur l'intervalle $[x, x + 2h]$ alors, d'après le théorème des valeurs intermédiaires, il existe un $\xi \in]x, x + 2h[$, vérifiant

$$f'''(\xi) = 2f'''(\xi_1) - f'''(\xi_2).$$

Par conséquent (2.3) devient de la forme

$$f'(x) = \frac{4f(x+h) - f(x+2h) - 3f(x)}{2h} + \frac{f'''(\xi)}{3}h^2, \quad \xi \in]x, x+2h[. \quad (2.4)$$

Ce qui donne la **formule de différence progressive à trois points** :

$$f'(x) \simeq \frac{4f(x+h) - f(x+2h) - 3f(x)}{2h}.$$

* De même, en écrivant le développement de Taylor en $x-h$ et $x-2h$, on obtient la **formule de différence régressive à trois points** :

$$f'(x) \simeq \frac{3f(x) - 4f(x-h) + f(x-2h)}{2h}.$$

L'erreur d'approximation commise, par les deux formules précédente se donne par :

$$|E| = \frac{h^2}{3} f'''(\xi), \quad \xi \in]x-2h, x+2h[.$$

* Enfin, en effectuant le développement de Taylor en $x-h$ et $x+h$, on aura la **formule de différence centrée à trois points** :

$$f'(x) \simeq \frac{f(x+h) - f(x-h)}{2h}.$$

Exemple 2.5 *Considérons la fonction $f : x \mapsto -\frac{1}{12}x^3 + \frac{1}{12}x + \frac{7}{2}$ de l'exemple précédent pour les données :*

$$(x_0, f(x_0)) = (0, 7/2); (x_1, f(x_1)) = (1, 7/2); (x_2, f(x_2)) = (2, 3);$$

$$(x_3, f(x_3)) = (3, 3/2); (x_4, f(x_4)) = (4, -3/2).$$

* Formule progressive à trois points (équidistants) :

$$f'(2) \simeq \frac{-f(4) + 4f(3) - 3f(2)}{2} = \frac{\frac{3}{2} + 6 - 9}{2} = -\frac{3}{4},$$

* Formule régressive à trois points (équidistants) :

$$f'(2) \simeq \frac{3f(2) - 4f(1) + f(0)}{2} = \frac{9 - 14 + \frac{7}{2}}{2} = -\frac{3}{4},$$

Remarque 2.6 *Les trois formules à trois points (quelconques ou équidistants) peuvent également être obtenues à partir du polynôme de Lagrange interpolant ces trois points.*

Interprétation géométrique :

Pour les différences à "deux points", on estime la dérivée "approchée" par la pente du segment de droite joignant les deux points. Dans le cas des différences à trois points, on détermine le polynôme d'interpolation de degré 2 (passants par ces trois points), dont la pente en x_0 , en x_1 et en x_2 donne, respectivement, les différences progressive, centrée et régressive.

3 Approximation des dérivées d'ordre supérieur

Pour obtenir une approximation de la dérivée seconde (ou d'ordre supérieur) nous procédons de la même façon, c.à.d, en utilisant le développement limité de Taylor d'ordre 4 (ou d'ordre supérieur), ou bien à partir des dérivées du polynôme de Lagrange passant par des points donnés $(x_i, f(x_i))$. Par exemple, pour la dérivée seconde, on reprend le polynôme de degré 2, passant par trois points (déjà utilisé pour calculer la dérivée première) :

$$P_1(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x - x_2)}f(x_0) + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x - x_2)}f(x_1) + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x - x_1)}f(x_2).$$

L'approximation de la dérivée seconde, en tout point $x \in]a, b[$, se donne par :

$$\begin{aligned} f''(x) &\simeq P_2''(x) \\ &= \frac{2}{(x_0 - x_1)(x - x_2)}f(x_0) + \frac{2}{(x_1 - x_0)(x - x_2)}f(x_1) + \frac{2}{(x_1 - x_0)(x - x_2)}f(x_2). \end{aligned}$$

Ainsi, en prenant $(x_0 = x, x_1 = x + h, x_2 = x + 2h); (x_0 = x - 2h, x_1 = x - h, x_2 = x)$ et $(x_0 = x - h, x_1 = x, x_2 = x + h)$, respectivement, on obtient les approximations suivantes :

- **formule de différence progressive :**

$$f''(x) \simeq \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2},$$

- **formule de différence régressive :**

$$f''(x) \simeq \frac{f(x-2h) - 2f(x-h) + f(x)}{h^2},$$

- **formule de différence centrée :**

$$f''(x) \simeq \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}.$$

Exemple 3.1 *Considérons la fonction $f : x \mapsto e^{2x}$. Estimons la valeur de $f''(0)$, en choisissant $h = 0,0312$.*

Corrigé :

1. *Formule progressive :*

$$f''(0) \simeq \frac{f(0 + 2h) - 2f(0 + h) + f(0)}{h^2} = 4.2589,$$

2. *Formule régressive :*

$$f''(0) \simeq \frac{f(x - 2h) - 2f(x - h) + f(x)}{h^2} = 3.7592,$$

3. *Formule centrée :*

$$f''(0) \simeq \frac{f(0 - 2h) - 2f(0 - h) + f(0)}{h^2} = 4.0013.$$

En comparant ces résultats avec la valeur exacte de $f''(0)$ qui est : $4e^{2 \times 0} = 4$, on conclut que la formule centrée est la plus précise.

4 Exercices

"Dérivation numérique"

Exercice 1.

Considérer la fonction $x \mapsto f(x) = 2^x, x \in [1, 5]$ passant par les points :

$$(x_0, y_0) = (1, 2); (x_1, y_1) = (2, 4); (x_2, y_2) = (3, 8); (x_3, y_3) = (4, 16) \text{ et } (x_4, y_4) = (5, 32).$$

1. Approcher la valeur $f'(x_2)$, en utilisant les trois formules vues en cours (formules à deux points).
2. Evaluer les erreurs d'approximation, sachant que $f'(x) = (\ln 2)2^x, x \in [1, 5]$.
3. Quelle est la meilleure formule pour approcher $f'(3)$?

Exercice 2. Soit la fonction $f(x) = x^2 + x + \sin x$ et considérons le point $x = 0$.

Tracer un tableau permettant de voir l'évolution des dérivées premières numériques, en $x = 0$, en fonction de h et de comparer les résultats des 3 formules de différences (à deux points équidistants).

* Prenner $h = 0.5, 0.25, 0.0125$ puis $h = 0.0123$.

Exercice 3.

Refaire la même question de l'exercice 2, en utilisant les deux formules progressive et régressive à trois points équidistants.

Exercice 4. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction de classe $\mathbb{C}^4[a, b]$ (ou plus). On se fixe un nombre $h > 0$ "assez petit", et un point (quelconque) $x \in]a, b[$.

1. Montrer que le rapport A approche une dérivée de f que l'on déterminera, où

$$A = \frac{-2f(x+2h) + 32f(x+h) - 60f(x) + 32f(x-h) - 2f(x-2h)}{24h^2}$$

2. Donner l'ordre de précision de cette approximation.
3. Commenter les résultats obtenus.

Exercice 5. Considérons $g : [a, b] \rightarrow \mathbb{R}$ une fonction de classe $\mathbb{C}^5[a, b]$. On se fixe un nombre $h > 0$ "assez petit", ainsi qu'un point (quelconque) $x \in]a, b[$. Soit le rapport

$$B = \frac{f(x+3h) - 3f(x+h) + 3f(x-h) - f(x-3h)}{8h^3}.$$

1. Montrer que le rapport B approche une dérivée de g que l'on déterminera.
2. Donner l'ordre de précision de cette approximation.
3. Vérifier que A coïncide avec une formule de dérivation numérique que l'on donnera.

Références

- [1] C. Boulonne, Analyse numérique et approximation, Notes de cours. Université de Lille (2008-2009).
- [2] P.G. Ciarlet, Introduction à l'analyse numérique. *Comput. Math. App* **56,4** (2008), 874–882.
- [3] J. P. Demally, Analyse Numérique et équations différentielles. *Differential Integral Equations* **19** (2006), 1235–1270.
- [4] G. Faccanoni, M33, Analyse numérique : Recueil d'exercices corrigés et aide-mémoire (2013-2014).
- [5] E. Walter, *Méthodes numériques et optimisation : Un guide de consommateur*. Springer International Publishing AG 2014.