CHAPTER 1

Analyse d'erreurs

1.1. Introduction

Les théories mathématiques permettent de résoudre de façon analytique un certain nombre de problèmes (intégration, équations algébriques ou différentielles, géometrie, ...). La solution est exacte. L'analyse numérique se distingue des méthodes pures classiques. Pour un problème donné, il est possible d'utiliser plusieurs techniques numériques de résolution qui donnent lieu à différents algorithmes, mais le resultat est mathématiquement approximatif. Ces routines numériques dépendent de certains paramètres influençant le résultat final. Ceci nous amène à une partie incontournable de l'analyse numérique qui est contenir les impacts des erreurs introduites, qui proviennent de trois origines principales:

(1) les erreurs de modélisation,
(2) les erreurs de représentation sur ordinateur,
(3) les erreurs de troncature.

Les erreurs de modélisation proviennent de l'étape de mathématisation du phénomène physique. Cette étape consiste à mettre sous forme d'équations (différentielles, stochastiques, ...) le phénomène en négligeant ses composantes les moins importantes ou qui rendent la résolution numérique trop difficile.

Les erreurs de représentation sur ordinateur proviennent de la représentation binaire des nombres sur machine. Par exemple, la fraction $1/3$ n'a pas de représentation binaire exacte, car elle ne possède pas de représentation décimale finie. Ces erreurs, même petites au départ, s'accumulement et se propagent dans les calculs dégradant ainsi le résultat final.

Enfin, les erreurs de troncature proviennent principalement de l'utilisation du développement de Taylor, qui permet par exemple de remplacer une équation différentielle par une équation algébrique. Le développement de Taylor est le principal outil mathématique du numéricien. Il est donc primordial d'en maîtriser l'énoncé et ses conséquences.

1.2. Terminologie

Avant d'aller plus loin, un peu de terminologie est inevitable pour décrire les concepts.

Definition 1. Soit $x \in \mathbb{R}$ et x^*son approximation.

(1) L'erreur absolue est definie par

$$\Delta x = |x - x^*|$$

(2) L'erreur relative est definie par

$$E_r(x) = \frac{\Delta x}{|x|}$$

(3) En multipliant par 100 le nombre $E_r(x)$ on obtient l'erreur en pourcentage.

Cela nous amène à parler des chiffres significatifs.

Definition 2. Si l'erreur absolue vérifie

$$\Delta x \leq 0.5\, 10^m$$

alors le chiffre m est dit significatif ainsi que tous ceux à sa gauche.

EXAMPLE. Soit $x = \pi$ et une approximation $x^* = \frac{22}{7} = 3,142857.....$ Le clacul donne $\Delta x = 0,00126 \leq 0.5\,10^{-2}$. On conclut le millième est significatif donc on a 3 chiffres significatifs $3,14$.

Si on retient $x^* = 3,1416$ comme approximation de π. on a: $\Delta x = 0,73\,10^{-5} \leq 0.5\,10^{-4}$. Le chiffre 6 et tous ceux à sa gauche. Donc $3,1416$ contient 5 chiffres significatifs.

EXERCISE. Le poids d'une personne est mesuré 90,567 kg par une machine précise. Tous les chiffres sont significatifs. Trouver l'incertitude de la machine (Borne sup pour Δx).

1.3. Les erreurs de troncature

Les erreurs de troncature constituent la principale catégorie d'erreurs en analyse numérique. Elles proviennent essentiellement du développement de Taylor. D'abord, l'ordre d'une méthode dépend du nombre de termes utilisés dans les développements de Taylor appropriés. Le développement de Taylor peut être vu comme un problème d'approximation d'une fonction $f(x)$ au voisinage d'un point x_0 par un polynôme $P_n(x)$ de degré n.

DEFINITION 3. Soit $f \in \mathcal{C}^n(\mathbb{R})$. Le polynôme de Taylor de f de degré n au voisinage de x_0 est donné par

$$P_n(x) = f(x_0) + f^{'}(x_0)(x - x_0) + \frac{f^{(2)}(x_0)}{2!}(x - x_0)^2 + ... + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n,$$

où $f^{(n)}(x_0)$ est la dérivée d'ordre n de $f(x)$ au point x_0.

On choisit généralement le point x_0 où on développe le polynôme de Taylor de façon à ce que l'on puisse facilement évaluer la fonction $f(x)$ ainsi que ses dérivées. Ce polynôme donne une approximation de la fonction $f(x)$ au voisinage de x_0, il y a donc une erreur d'approximation.

THEOREM 4. *Soit* $f \in \mathcal{C}^{n+1}(\mathbb{R})$. *On a* $f(x) = P_n(x) + R_n(x)$ *où* $P_n(x)$ *est le polynôme de Taylor de degré* n *et* $R_n(x)$ *est l'erreur d'approximation commise*

$$R_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}(x - x_0)^{n+1}$$

avec ξ_x *un nombre compris entre* x *et* x_0.

Remarquons que le terme d'erreur $R_n(x)$ devient de plus en plus grand lorsque x s'éloigne de x_0 sur la base de $(x - x_0)^{n+1}$. Aussi, le point ξ_x existe mais n'est pas connu exactement, donc le terme d'erreur ne peut pas etre évalué exactement.

Notons qu'on commet une erreur de troncature chaque fois que l'on utilise le développement de Taylor et que l'on néglige le terme d'erreur.

Une forme courante du développement de Taylor est obtenue en écrivant $x = x_0 + h$ et on a

$$P_n(h) = P_n(x_0 + h) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} h^k \qquad et \qquad R_n(x_0 + h) = \frac{f^{(n+1)}(\xi_h)}{(n+1)!} h^{n+1}$$

EXAMPLE 5. On examine $f(x) = e^x$ au voisinage de $x_0 = 0$. Le polynôme de Taylor de degré n est

$$e^{x_0+h} = e^h \simeq P_n(h) = \sum_{k=0}^{n} \frac{e^0}{k!} h^k = 1 + h + \frac{h^2}{2!} + ... + \frac{h^n}{n!}$$

Le terme erreur est

$$R_n(h) = \frac{f^{(n+1)}(\xi_h)}{(n+1)!} h^{n+1} = \frac{e^{\xi_h}}{(n+1)!} h^{n+1} \quad \xi_h \in (0, h).$$

On peut trouver une borne sup $M_n(h)$ de l'erreur

$$R_n(h) \leq M_n(h) = \frac{e^h}{(n+1)!} h^{n+1}$$

Un tableau de valeurs est etabli pour $h = 0,1$. La valeur exacte est $e^{0.1}$

n	$P_n(h)$	Err absolue	$M_n(h)$	nbre de chiffres significatifs
0	1	$0.105\,10^{0}$	$0.111\,10^{0}$	position 0→1 chiffre
1	1.1	$0.517\,10^{-2}$	$0.552\,10^{-2}$	position -3→2 chiffres
2	1.105	$0.171\,10^{-3}$	$0.184\,10^{-3}$	position -3→4 chiffres
3	1.1051667	$0.420\,10^{-5}$	$0.460\,10^{-5}$	position -5→6 chiffres

La précision de l'approximation s'améliore avec le degré du polynôme au prix d'une augmentation du nombre de calculs. Prenons maintenant $h \leftarrow h/2$, c.a.d, $h = 0,05$. On retrouve

$$P_3(0.05) = 1.05112708 \quad et \quad M_3(0.05) = 0.2\,10^{-6} \ \rightarrow 7\, chiffres\ significatifs$$

La nouvelle erreur absolue $\Delta(0.05) = 0.263\,10^{-6}$ et le rapport

$$\frac{\Delta(0.1)}{\Delta(0.05)} = 16.4$$

L'erreur absolue $\Delta(0.05)$ est environ 16 fois plus petite que $\Delta(0.1)$ en passant de h à $h/2$. ce rapport n'est pas le fruit du hasard. Le reste de la théorie l'explique.

Definition 6. On dit que $E(h)$ est un grand ordre de h^n au voisinage de 0 (notation $V(0)$) et on note $E(h) = \mathcal{O}(h^n)$ s'il existe une constante $C > 0$ au $V(0)$ tel que

$$|E(h)| \leq C\, h^n.$$

Remark. Quand $E(h) = \mathcal{O}(h^n)$, on dit que

- $E(h)$ est dominé par h^n,
- $E(h)$ est un grand O de h^n.

Interpretation: Lorsque h est assez petit, si h décroit $E(h)$ décroit proportionellemnt selon h^n. Plus n est grand plus la decroissance est rapide. Donc $\mathcal{O}(h^3)$ decroit plus vite que $\mathcal{O}(h^2)$. Pour se donner une idée sur une fonction $E(h) = \mathcal{O}(h^n)$, c.a.d,

$|E(h)| \leq C\, h^n$, si on divise h par 2, alors $E(h)$ decroit d'un facteur 2^n car

$$\left|E(\frac{h}{2})\right| \leq C\, \left(\frac{h}{2}\right)^n = C\, \frac{h^n}{2^n}.$$

Revenons au développement de Taylor au $V(0)$ pour $f \in \mathcal{C}^{n+1}(\mathbb{R})$, l'erreur absolue est

$$\Delta_n h = |f(h) - P_n(h)| = |R_n(h)| = \left|\frac{f^{(n+1)}(\xi_h)}{(n+1)!}\right| h^{n+1} \leq M\, h^{n+1}$$

Dans l'exemple precedent, on a $n = 3$, et on a passe de h à $h/2$, l'erreur decroit approximativement d'un facteur 2^{n+1}, c.a.d, $2^4 = 16$ ce qui explique le facteur obtenu précédemment.

DEFINITION 7. Une approximation dont l'erreur absolue Δh est $\mathcal{O}(h^n)$ est dite d'ordre n

Le polynome de Taylor est une approximation au moins $\mathcal{O}(h^{n+1})$

REMARK 8. il y a parfois confusion entre le degré et l'ordre du polynôme de Taylor. Il faut s'assurer de bien distinguer ces deux notions.

EXAMPLE 9. Calculons le developpement de Taylor d'ordre 5 au $V(0)$ de $f(x) = sin(x)$.

On sait que le polynome de Taylor est d'ordre $(n+1)$. On prend $n = 4$ et on calcule de f a $f^{(4)}$. On obtient

$$sin(h) = h - \frac{h^3}{6} + \frac{cos(\xi_h)}{5!} h^5$$

on déduit que le polynôme de Taylor $P_3(h)$ est de degré 3 mais d'ordre 5. Si $h' = \frac{h}{2}$ alors $sin(h)$ décroit de $2^5 = 32$.

1.4. Propagation des erreurs.

Que peut-on dire, par exemple, de la précision des résultats obtenus lorsqu'on additionne où qu'on multiplie des valeurs connues avec une précision limitée? Plus généralement, si

$$x = x^* \pm \Delta x \quad \text{et} \quad y = y^* \pm \Delta y$$

quelle sera la précision d'une fonction d'une seule variable $f(x^*)$ ou de deux variables $g(x^*, y^*)$? Ici le développement de Taylor apporte la solution comme suit. Commençons par le cas d'une variable.

$$f(x) = f(x^* \pm \Delta x) = f(x^*) \pm f^{'}(x^*)\Delta x + \mathcal{O}\left((\Delta x)^2\right)$$

Négligeons les termes de degré plus grand ou égal à 2, on obtient:

$$|f(x) - f(x^*)| = \Delta f \simeq \left| f^{'}(x^*) \right| \Delta x$$

Example 10. On a mesuré la longueur d'un côté d'une boîte cubique et on a obtenu $l^* = 10,2\,cm$ avec une précision de l'ordre du millimètre ($\Delta l = 0.1\,cm$). On cherche le volume v de cette boîte. Dans ce cas, $v = f(l) = l^3$ et l'erreur liée au volume est:

$$\Delta v = \left| f^{'}(l^*) \right| \Delta l = 3\,(l^*)^2\,\Delta l = 3\,(10.2)^2 \times 0.1 = 31.212 \leq 0.5\,10^2$$

La valeur approximative du volume est $v = (10.2)^3 = 1061.2\,cm^3$, dont seuls les deux premiers chiffres sont significatifs 1 et 0 (c.a.d à partir de la position 2).

On traite les fonctions de plusieurs variables en faisant appel au développement de Taylor en plusieurs variables. Nous donnons le résultat en dimension 3 seulement, car le cas général ne pose aucune difficulté supplémentaire.

THEOREM 11. *Soit $f(x, y, z)$ une fonction de 3 variables x, y, z dont les valeurs approchées sont x^*, y^*, z^* avec une precision $\Delta x, \Delta y, \Delta z$. L'erreur absolue est*

$$\Delta f = \left|\frac{\partial f(x^*, y^*, z^*)}{\partial x}\right| \Delta x + \left|\frac{\partial f(x^*, y^*, z^*)}{\partial y}\right| \Delta y + \left|\frac{\partial f(x^*, y^*, z^*)}{\partial z}\right| \Delta z.$$

EXAMPLE 12. Un signal électrique est donné par

$$V = A\, sin(\omega t - \phi)$$

où V est la tension, A est l'amplitude du signal ($A^* = 100V$), ω est la fréquence ($\omega^* = 3rad/s$), ϕ est le déphasage ($\phi^* = 0,55\, rad$) et t est le temps ($t^* = 0,001s$). En supposant que A et ω sont connus exactement ($A = A^*, \omega = \omega^*$) et que ϕ^* et t^* possèdent respectivement 2 et 1 chiffres significatifs, il s'agit d'évaluer l'erreur absolue liée à V ainsi que le nombre de chiffres significatifs.

Puisque A et ω sont connus exactement, on sait immédiatement que $\Delta A = \Delta\omega = 0$ et qu'ils n'ont aucune contribution à l'erreur liée à V. Par ailleurs:

$$\Delta\phi = 0.5\, 10^{-2} \text{ et } \Delta t = 0.5\, 10^{-3}$$

L'incertitude sur V est

$$\Delta V = \left|\frac{\partial V(t^*, \phi^*)}{\partial t}\right| \Delta t + \left|\frac{\partial V(t^*, \phi^*)}{\partial \phi}\right| \Delta\phi$$

autrement

$$\Delta V = |A^*\omega^*\, cos(\omega^* t^* - \phi^*)|\, \Delta t + |-A^*\, cos(\omega^* t^* - \phi^*)|\, \Delta\phi$$

finalement

$$\Delta V = 256,22666235 \times (0,5\, 10^{-3}) + |-85,4088745| \times (0,5\, 10^{-2}) = 0,555 = 0,0555\, 10^1$$

La tension approximative est

$$V^* = A^* \, sin(\omega^* t^* - \phi^*) = -52,0127$$

Le resultat a un seul chiffre significatif qui est le 5!

1.5. Arithmétique flottante

En système décimal, la notation flottante à m chiffres d'un nombre $x \in \mathbb{R}$ est definie par

$$fl_m(x) = 0, d_1 d_2 ... d_m \, \times \, 10^l \quad l \in \mathbb{Z}.$$

Le partie $0, d_1 d_2 ... d_m$ s'appelle la mantisse. Arrondir à $n < m$ chiffres significatifs consiste à ajouter 5 au chiffre de position $(n+1)$ de la mantisse ensuite tronquer au chiffre de position n.

EXAMPLE 13. Arrondir à $n = 4$ chiffres significatifs.

$\frac{1}{3} = 0,33333333... \implies fl(\frac{1}{3}) = 0,3333,$
$fl(12,4551) = fl(0,124551 \, 10^2) = 0,1246 \, 10^2 = 12,46$

Les opérations élémentaires en arithmétique flotannte sont l'addition, la soustraction, la multiplication et la division. Soit x et y, deux nombres réels, on a

$$x + y \to fl\Big[fl(x) + fl(y)\Big]$$

$$x - y \to fl\Big[fl(x) - fl(y)\Big]$$

$$x \, \times \, y \to fl\Big[fl(x) \, \times \, fl(y)\Big]$$

$$x/y \to fl\Big[fl(x)/fl(y)\Big]$$

Example 14. 1- Calculer $1/3 \times 3$ en prenant $n = 4$. On a $1/3 = 0,3333\,3$ donc $fl(1/3) = 0,3333$

Aussi $3 = 0,3000\,0\ .10^1$ donc $fl(3) = 0,3000\,.10^1$.

La définition donne

$$\frac{1}{3} \times 3 = fl\Big[0,3333\,10^0 \times 0,3000\,10^1\Big] = fl\Big[0,09999000\,10^1\Big] = fl\Big[0,9999000\,10^0\Big] = 0,9999$$

On remarque une petite perte de précision par rapport à la valeur exacte de cette opération qui est 1.

2- Arrondir à 4 chiffres $0,4035\,10^6 + 0,1978\,10^4$ cette addition flottante. On a

$$0,4035\,10^6 + 0,1978\,10^4 = fl(0,4035\,10^6 + 0,1978\,10^4) = fl(0,4035\,10^6 + 0,001978\,10^6) = fl(0$$

donc

$$fl_4(0,4035\,10^6 + 0,1978\,10^4) = 0,4055\,10^6$$