
Chapitre 1

Analyse d'erreurs

Sommaire

1.1	Sources d'erreurs	2
1.1.1	Erreur absolue	2
1.1.2	Erreur relative	3
1.1.3	Propagation des erreurs	4
1.2	Représentation des nombres réels	4
1.2.1	Virgule flottante	4
1.2.2	Représentation décimale	5
1.2.3	Arrondi d'un nombre	7
1.2.4	Évaluation des polynômes	8
1.3	Exercices	9

L'analyse numérique est une branche des mathématiques appliquées s'intéressant aux méthodes permettant de résoudre, par des calculs numériques, des problèmes d'analyse mathématique.

L'étude des erreurs forme une partie importante de l'analyse numérique. Dans ce premier chapitre, nous introduisons quelques définitions essentielles sur la notion d'erreurs.

1.1 Sources d'erreurs

Les solutions de problèmes calculées par une méthode numérique sont affectées par des erreurs, qui proviennent de trois sources principales :

- *les erreurs de modélisation*, qui proviennent de l'étape de mathématisation du phénomène physique auquel on s'intéresse. Par exemple en négligeant certains termes physiques.
- *les erreurs d'arrondi*, qui proviennent des erreurs de représentation dues au fait que tout ordinateur travaille en précision finie. Par exemple, la fraction $\frac{1}{3}$ n'a pas de représentation binaire finie, pas plus qu'elle ne possède de représentation décimale finie. On ne pourra donc pas représenter exactement cette fraction, ce qui introduit une erreur.
- *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.

Pour mesurer l'erreur entre la solution fournie par une méthode numérique et la solution du problème que l'on cherche à résoudre, on introduit les notions d'erreur absolue et relative.

1.1.1 Erreur absolue

Soit x un nombre, et x^* une approximation de ce nombre (valeur approchée ou approximative de x).

- Si $x^* > x$, x^* est dite valeur approchée par excès.
- Si $x^* < x$, x^* est dite valeur approchée par défaut.

Dans les deux cas on note

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

Exemple 1

$\frac{1}{3} \simeq 0.33$, $\pi \simeq 3.14$ et $\sqrt{2} \simeq 1.41$ sont des approximations par défaut mais $e \simeq 2.72$ est une approximation par excès. ■

Définition 1 (Erreur absolue)

Soit x un nombre, et x^* une approximation de ce nombre. L'erreur absolue est définie par

$$E_a(x^*) = |x - x^*|. \quad (1.1)$$

Plus l'erreur absolue est petite, plus x est précis.

Dans le cas de quantités mesurées expérimentalement dont on ne connaît que la valeur approximative, on dispose souvent d'une borne supérieure pour l'erreur absolue qui dépend de la précision des instruments de mesure utilisés. Alors qu'en fait on a

$$E_a(x^*) = |x - x^*| \leq \Delta x,$$

ce qui peut également s'écrire

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

et que l'on note parfois

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

Par abus de langage, on confondra la borne Δx avec $E_a(x^*)$.

1.1.2 Erreur relative**Définition 2 (Erreur relative)**

Soit x un nombre non nul, et x^* une approximation de ce nombre. L'erreur relative est définie par

$$E_r(x^*) = \frac{|x - x^*|}{|x|} \simeq \frac{\Delta x}{|x^*|}. \quad (1.2)$$

De plus, en multipliant E_r par 100, on obtient l'erreur relative en pourcentage.

En pratique, il est difficile d'évaluer les erreurs absolue et relative, car on ne connaît généralement pas la valeur exacte de x et l'on n'a que x^* . C'est pourquoi on utilise l'approximation $\frac{\Delta x}{|x^*|}$ pour l'erreur relative et on note $\delta x = \frac{\Delta x}{|x^*|}$.

L'erreur absolue donne une mesure quantitative de l'erreur commise et l'erreur relative en mesure l'importance.

Exemple 2

- $x = \pi = 3.141592 \dots$, $x^* = 3.1419$

$$\text{Erreur absolue} = 3.073464102070211 \times 10^{-4}$$

$$\text{Erreur relative} = 9.783140085199355 \times 10^{-5}.$$

- $y = 10^6$, $y^* = 999996$

$$\text{Erreur absolue} = 4.0 \times 10^0$$

$$\text{Erreur relative} = 4.0 \times 10^6.$$

- $z = 0.000012$, $z^* = 0.000009$

$$\text{Erreur absolue} = 3.0 \times 10^{-6}$$

$$\text{Erreur relative} = 25.0 \times 10^{-2}.$$

■

1.1.3 Propagation des erreurs

Soient x et y deux valeurs positives, x^* et y^* deux valeurs approchées de x et y respectivement.

Alors on a les propriétés suivantes.

$$\Delta(x + y) = \Delta x + \Delta y, \quad \delta(x + y) \leq \max(\delta x, \delta y),$$

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

$$\Delta(xy) = x^* \Delta y + y^* \Delta x, \quad \delta(xy) = \delta x + \delta y,$$

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

1.2 Représentation des nombres réels**1.2.1 Virgule flottante**

La virgule flottante est une méthode d'écriture de nombres réels fréquemment utilisée dans les ordinateurs. Elle consiste à représenter un nombre réel par :

- un signe (égal à -1 ou 1),
- une mantisse (aussi appelée significande),
- et un exposant (entier relatif, généralement borné).

Un tel triplet représente le nombre réel :

$$\text{signe} \times \text{mantisse} \times b^{\text{exposant}},$$

où b est la base de représentation, généralement 2 sur ordinateur, mais aussi 8 ou 16 sur certaines anciennes machines, 10 sur de nombreuses calculatrices, ou éventuellement toute autre valeur.

En faisant varier l'exposant, on fait flotter la virgule.

Exemple 3

Le nombre 1.3254 peut être représenté en virgule flottante comme suit

$$1.3254 = \underbrace{13254}_{\text{mantisse}} \times 10^{\overbrace{-4}^{\text{exposant}}}.$$

■

1.2.2 Représentation décimale

Tout nombre réel positif x peut être représenté sous la forme d'une représentation décimale (en base 10) de développement fini ou infini

$$x = a_m 10^m + a_{m-1} 10^{m-1} + \cdots + a_{m-n+1} 10^{m-n+1} + \cdots,$$

avec $a_i \in \{0, 1, \dots, 9\}$ pour $i \neq m$ et $a_m \neq 0$ où m est le rang supérieur de x (la plus grande puissance de 10).

Exemple 4

Cas d'un développement fini :

$$1404.1978 = 1 \times 10^3 + 4 \times 10^2 + 0 \times 10^1 + 4 \times 10^0 + 1 \times 10^{-1} + 9 \times 10^{-2} + 7 \times 10^{-3} + 8 \times 10^{-4}.$$

Cas d'un développement infini :

$$\frac{1}{7} = 0.14286 \cdots = 1 \times 10^{-1} + 4 \times 10^{-2} + 2 \times 10^{-3} + 8 \times 10^{-4} + 6 \times 10^{-5} + \cdots.$$

■

Dans la pratique on n'utilise que des approximations par des nombres qui ont des représentations décimales de développement fini

$$x \simeq x^* = a_m 10^m + a_{m-1} 10^{m-1} + \cdots + a_{m-n} 10^{m-n}, \quad a_m \neq 0.$$

Pour mesurer la précision d'une approximation, on introduit la notion de chiffres significatifs (c.s), qui sont les chiffres conservés a_i .

Chiffres significatifs

Définition 3 (Chiffre significatif)

On appelle chiffre significatif d'un nombre approché, tout chiffre dans sa représentation décimale différent de zéro, et un zéro s'il se trouve entre deux chiffres significatifs, ou s'il constitue un chiffre conservé.

Exemple 5

Une approximation à 5 décimales de 0.02010 est

$$\underline{0.02010}$$

- Les zéros soulignés ne sont pas significatifs car ils ne servent qu'à indiquer les rangs des autres chiffres.
- Le zéro double souligné étant placé entre les chiffres significatifs 2 et 1, zéro est lui même un chiffre significatif.
- Le zéro triple souligné traduit le fait que le nombre approché a conservé la décimale 10^{-5} est un chiffre significatif. ■

Remarque 4

L'écriture en virgule flottante détermine le nombre des chiffres significatifs, par exemple 3.52×10^4 a trois chiffres significatifs mais 3.5200×10^4 a cinq chiffres significatifs.

Définition 5 (Chiffre significatif exact)

Un chiffre significatif d'un nombre approché x^* est dit exact (c.s.e) si l'erreur absolue de x^* vérifie

$$\Delta x \leq 0.5 \times 10^k$$

avec k est le rang de ce chiffre significatif.

Ainsi : Le $n^{\text{ième}}$ chiffre significatif après la virgule est exact si $\Delta x \leq 0.5 \times 10^{-n}$. Le $n^{\text{ième}}$ chiffre significatif avant la virgule est exact si $\Delta x \leq 0.5 \times 10^{n-1}$.

Exemple 6

Soit $x = 35.97$ et $x^* = 36.00$. Ici $m = 1$ car $x = 3 \times 10^1 + 6 \times 10^0 + 0 \times 10^{-1} + 0 \times 10^{-2}$. On a

$$\Delta x = |x - x^*| = |35.97 - 36.00| = 0.03 \leq 0.5 \times 10^{-1}.$$

Alors $m - n + 1 = k = -1$ qui donne $n = m + 2 = 3$ et donc x^* est une approximation de x avec trois chiffres significatifs exacts. ■

Remarque 6

- Si un chiffre significatif est exact, tous les chiffres significatifs à sa gauche sont exacts.
- Si un chiffre significatif n'est pas exact, tous ceux à sa droite ne le sont pas.

1.2.3 Arrondi d'un nombre

Un arrondi d'un nombre est une valeur approchée de ce nombre obtenue, à partir de son développement décimal, en réduisant le nombre de chiffres significatifs.

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, puis on procède de la manière suivante :

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

Exemple 7

1. L'arrondi de $x = 0.254$ à 2 chiffres significatifs est $x^* = 0.25$ car $4 < 5$.
2. L'arrondi de $x = 0.4368$ à 3 chiffres significatifs est $x^* = 0.437$ car $8 > 5$.
3. L'arrondi de $x = 1.534500$ à 4 chiffres significatifs est $x^* = 1.534$ car tous les chiffres rejetés sont des zéros et le $4^{\text{ième}}$ chiffre étant pair.
4. L'arrondi de $x = 1.5347500$ à 5 chiffres significatifs est $x^* = 1.5348$ car tous les chiffres rejetés sont des zéros et le $5^{\text{ième}}$ chiffre étant impair.

5. L'arrondi de $x = 23.6050420$ à 4 chiffres significatifs est $x^* = 23.61$ car parmi les chiffres rejeté il existe au moins un qui soit non nul. ■

1.2.4 Évaluation des polynômes

Il est très fréquent d'avoir à évaluer des polynômes de degré élevé en analyse numérique. Il est donc important de pouvoir les évaluer rapidement et de la façon la plus stable possible du point de vue de l'arithmétique flottante. C'est ce que permet l'algorithme de *Horner* appelé aussi algorithme de multiplication imbriquée. Pour évaluer un polynôme de la forme

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n$$

en un point x quelconque, il suffit de regrouper les termes de la façon suivante

$$p(x) = a_0 + x(a_1 + x(a_2 + x(a_3 + \cdots + x(a_{n-1} + a_nx) \cdots)))$$

Exemple 8

Soit le polynôme $p(x) = 2 + 4x + 5x^2 + 3x^3$ qui nécessite 6 multiplications et 3 additions. En suivant le mode de regroupement suivant l'algorithme de Horner, on obtient $p(x) = 2 + x(4 + x(5 + 3x))$ qui nécessite seulement 3 multiplications et 3 additions. On réduit donc le nombre d'opérations nécessaires. ■

Algorithme 1 : Schéma de Horner pour évaluer un polynôme p .

- Données :** 1) Les coefficients a_i de p dans l'ordre décroissant des exposants des x^i .
2) Un point x en lequel à évaluer le polynôme p .

1 **début**

2 $n :=$ degré du polynôme p

3 $Q := a_n$

4 **pour** $i \leftarrow 1$ à n **faire**

5 $Q := Q * x + a_{n-i}$

Résultat : Q qui est égal à $p(x)$ sous la forme d'un polynôme de Horner.

Code Matlab 1 : Évaluation d'un polynôme par la méthode de Horner

```
1 function f = horner(a,x)
2 %Entrée:
```

```
3 % a=[a_n a_{n-1}... a_0] coefficients d'un polynôme p de degré n dans ...  
   l'ordre des puissances décroissantes de x.  
4 % x : un point en lequel à évaluer le polynôme p.  
5 %Sortie:  
6 % f : la valeur du polynôme p en x par la méthode de Horner.  
7 n = length(a)-1;  
8 m = n+1;  
9 g = a(1);  
10 for j = 2:m  
11     g = g*x + a(j);  
12 end  
13 f = g;
```

1.3 Exercices

En rédaction!