Cours Méthodes Numériques et Programmation

Pr. Souad EL BERNOUSSI
Laboratoire de recherche :
*M*athématiques, *I*nformatique et *A*pplications *MIA*
E.mail : s.elbernoussi@fsr.ac.ma

Table des matières

CHAPITRE 1

Méthodes numériques et programmation

1. Introduction

1.1. Analyse numérique ? Etude et la construction d'algorithmes (du nom du mathématicien Al Khawarizmi) de résolution numérique d'un problème donné.

En pratique, l'Analyse Numérique se propose d'étudier les propriétés mathématiques des algorithmes et leur mise en oeuvre (programmation).

1.2. Objectifs ? L'objectif de l'analyse numérique est de :
concevoir et d'étudier
des méthodes de résolution de certains problèmes mathématiques (en général issus de la modélisation de problèmes 'réels'), et dont on cherche à calculer la solution ou son approximation à l'aide d'un ordinateur.

1.3. Enjeux de l'analyse numérique ? Résoudre des problèmes :
- que l'on ne sait pas résoudre autrement
- 'mieux' qu'on ne le faisait avant :
  - plus précisément,
  - moins cher...

1.4. Etique ('Objectifs') de l'analyse numérique.
- Plus vite :
  - complexité des algorithmes
  - complexité des problèmes
- Plus précis :
  - erreur d'arrondi (liées à la machine)
  - erreur d'approximation (liées à l'algorithme)
- Plus fiable :
  - stabilité d'un algorithme
- Facile à programmer :
  - comprendre pour mieux réutiliser

1.5. 1ère Conclusion. La confiance aveugle dans ce que l'on appelle les résultats fournis par l'ordinateur peut être la cause d'erreurs qui peuvent coûter très chères

Alors que faire ?

1.6. Sources d'Erreurs. Il y a en 3 catégories :
- Erreurs liés à la machine,
- Erreurs à la méthode (algorithme),
- Erreurs sur les donnés (résultat d'un calcul approché, d'une mesure physique,...)

2. Arithmétique des calculateurs et Sources d'erreurs

Si sophistiqué qu'il soit , un calculateur ne peut fournir que des réponses approximatives.

Les approximations utilisées dépendent à la fois des contraintes physiques (espace mémoire, vitesse de l'horloge...) et du choix des méthodes retenues par le concepteur du programme .

Le but de ce chapitre est de prendre connaissance de l'impact de ces contraintes et de ces choix méthologiques. Dans certains cas il doit être pris en compte dans l'analyse des résultats dont une utilisation erronée pourrait être coûteuse.

La première contrainte est que le système numérique de l'ordinateur est discret,
c'est à dire qu'il ne comporte qu'un nombre fini de nombres ;
Il en découle que tous les calculs sont entachés d'erreurs.

2.1. Evaluation de l'erreur. Rappelons d'abord quelques notion de base ;
Si X est une quantité à calculer et X^* la valeur calculée, on dit que :

– $X - X^*$ est l'erreur et $\mid E \mid = \mid X - X^* \mid$est l'erreur absolue.

Exemple :

Si $X = 2.224$ *et* $X^* = 2.223$ *alors l'erreur absolue* $\mid E \mid = \mid X - X^* \mid = 2.224 - 2.223 = 0.001$

– $E_r = \left| \frac{X - X^*}{X_r} \right|$ est l'erreur relative,
$X_r \neq 0$. X_r est une valeur de référence pour X. En général ,on prend $X_r = X$.

Exemple :

Si $X = 2.224$ *et* $X^* = 2.223$
alors , si on prend $X_r = X$ *, l'erreur relative* $E_r = \left| \frac{X - X^*}{X_r} \right| = \frac{|X - X^*|}{|X|} = \frac{0.001}{2.224} = 4.496 \times 10^{-4}$

Cependant, si X est la valeur d'une fonction $F(t)$ avec $a \leq t \leq b$, on choisira parfois une valeur de référence globale pour toutes les valeurs de t.

Exemple :

Si $X = \sin(t)$ *avec* $0 \leq t \leq \frac{\pi}{4}$, *on pourra prendre* $X_r = \frac{\sqrt{2}}{2} = \sup_{0 \leq t \leq \frac{\pi}{4}} \sin(t)$.

En général , on ne connait pas le signe de l'erreur de sorte que l'on considère les erreurs absolues et les erreurs relatives absolues.

Les opérations élémentaires propagent des erreurs.
Dans la pratique, on considère que :

1) L'erreur absolue sur une somme est la somme des erreurs absolues.

2) L'erreur relative sur un produit ou un quotient est la somme des ereurs relatives.

On peut estimer l'effet d'une erreur E sur l'argument x d'une fonction $f(x)$ au moyen de la dérivée de $f(x)$. En effet $f(x + E) \simeq f(x) + Ef'(x)$

Exemple :

Calculer la valeur de $(11111111)^2$

La valeur fournie par une petite calculatrice à cinq chiffres est $1,2345x10^{14}$

Mais la réponse exacte est 123456787654321.

La machine a donc tronqué le résultat à 5 chiffres et l'erreur absolue est de $6 * 10^9$.

L'erreur relative est de 0.005% .

Cet exemple montre qu'il faut établir clairement l'objectif visé.

Cet objectif est double ;

1) Nous voulons un bon ordre de grandeur (ici 10^{14}) et avoir le maximum de décimales exactes,

2) Ce maximum ne peut excéder la longueur des mots permis par la machine et dépend donc de la machine

2.2. La mémoire de l'ordinateur : le stockage des nombres. La mémoire d'un ordinateur est formée d'un certain nombre d'unités adessables appelées OCTETS .
Un ordinateur moderne contient des millions voir des milliards d'octets.
Les nombres sont stockés dans un ordinateur comme ENTIERS ou REELS.

2.3. Les nombres entiers : Les nombres entiers sont ceux que l'on utilise d'habitude sauf que le plus grand nombre représentable dépend du nombre d'octets utilisés :

- avec deux (2) octets, on peut représenter les entiers compris entre -32768 et 32767
- avec quatre (4) octets on peut représenterr les entiers compris entre -2147483648 et 2147483647

2.4. Les nombres réels. Dans la mémoire d'un ordinateur, les nombres réels sont représentés en notation flottante.

Cette notation a été introduite pour garder une erreur relative à peu prés constante ; quelque soit l'ordre de gandeur du nombre qu'on manipule.
En notation flottante, un nombre a la forme :

$x = \pm Y \times b^e$

b est la base du système numérique utilisé

Y est la mantisse : une suite de s entier $y_1 y_2 ... y_s$ avec $y_1 \neq 0$ si $x \neq 0$ et $0 \leq y_i \leq (b-1)$

e est l'exposant(un nombre entier relatif)

La norme choisie est celle où la mantisse est comprise entre 0 et 1 et où le premier chiffre après la virgule est différent de zéro.
Calcul de l'erreur

Nous terminons ce chapitre en définissant les notions de troncature et d'arrondie.
Exemple :
En base 10, $x = 1/15 = 0.066666666......$
Dans le cas d'une représentation tronquée nous aurons, pour $s = 5, \;\; fl(x) = 0.66666 * 10^{-1}$.
Remarquez comment nous avons modifié l'exposant afin de respecter la règle qui veut que le premier chiffre de la mantisse ne soit pas nul .

Dans ce cas, l'erreur absolue $X - fl(X)$ est de 6×10^{-7}. L'erreur relative est de l'ordre de 10^{-5}
Dans une représentation tronquée à s chiffres, l'erreur relative maximale est de l'ordre de 10^{-s}

Dans une représentation arrondie, lorsque la première décimale négligée est supérieure à 5, on ajoute 1 à la dernière décimale conservée.
Exemple :
$x = 1/15 = 0.066666666$.
Nous écrirons $\; fl(x) = 0.66667 \times 10^{-1}$
L'erreur absolue serait alors 3.333×10^{-7}*et l'erreur relative serait* 5×10^{-6}
En général, l'erreur relative dans une représentation arrondie à s chiffres est de $5 \times 10^{-(s+1)}$ soit la moitié de celle d'une représentation tronquée.

2.5. Les régles de base du modèle. Pour effectuer une opération sur deux nombres réels, on effectue l'opération sur leurs représentations flottantes et on prend ensuite la représentation flottante du résultat.

2.6. l'addition flottante. $x \oplus y = fl(fl(x) + fl(y))$

2.7. la soustraction flottante. $x \ominus y = fl(fl(x) - fl(y))$

2.8. la multiplication flottante. $x \otimes y = fl(fl(x) \times fl(y))$

2.9. la division flottante. $x \div y = fl(fl(x)/fl(y))$
Chaque opération intermédiaire dans un calcul introduit une nouvelle erreur d'arrrondi ou de troncature.

Dans la pratique, il faudra se souvenir du fait que deux expressions algébriquement équivalentes peuvent fournir des résultats différents et que l'ordre des opérations peut changer les résultats.

Pour l'addition et la soustraction on ne peut effectuer ces 2 opérations que si les exposants sont les mêmes. On transforme le plus petit exposant et donc on ne respecte plus la régle voulant que le premier chiffre de la mantisse ne soit pas nul.

2.10. Quelques remarques sur ce modèle : On constate une déviation importante par rapport aux lois habituelles de l'arithmétique.

$x + (y + z)$ peut être différent de $(x + y) + z$.

Exemple :

Pour 4 chiffres significatifs ($s = 4$) *on a :*

$$(1 + 0.0005) + 0.0005 = 1.000$$

car

$$0.1 \times 10^1 + 0.5. \times 10^{-3} = 0.1. \times 10^1 + 0.00005. \times 10^1 = 0.1 \times 10^1 + 0.0000. \times 10^1 = 0.1 \times 10^1$$

et

$$1 + (0.0005 + 0.0005) = 1.001$$

Ainsi, l'addition flottante n'est pas associative .(TD :Sommation d'une série à termes positifs)

On constate aussi que si y est trés petit par rapport à x, l'addition de x et y donnera seulement x.

Exemple :

L'équation $1 + x = 1$ *a* $x = 0$ *comme unique solution. Mais dans un système à* 10 *chiffres significatifs, elle aura une infinité de solutions* (il *suffit de prendre* $\mid x \mid \prec 5 \times 10^{-11}$)

La distributivité de la multiplication par rapport à l'addition.

Exemple :

Considérons l'opération

$$\begin{gathered} 122 \times (333 + 695) = \\ (122 \times 333) + (122 \times 695) = \\ 125416 \end{gathered}$$

Si nous effectuons ces deux calculs en arithmétique à 3 chiffres ($s = 3$) *et arrondi, nous obtenons :*

$$\begin{gathered} 122 \times (333 + 695) = \\ fl(122) \times fl(1028) = \\ 122 \times 103 \times 10^1 = \\ fl(125660) = 126 \times 10^3 \end{gathered}$$

$$\begin{gathered} (122 \times 333) + (122 \times 695) = \\ fl(40626) + fl(84790) = \\ 406 \times 10^2 + 848 \times 10^2 = \\ fl(406 + 848) \times 10^2 = \\ fl(1254 \times 10^2) = \\ 125 \times 10^3 \end{gathered}$$

Donc la distributivité de la multiplication par rapport à l'addition n'est pas respectée en arithmétique flottante.

3. Propagation des erreurs.

Une étude de la propagation des erreurs d'arrondi permattra d'expliquer ce phénomène.

Soit à calculer e^x à l'aide de son développement en série qui est convergent pour tout x :

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + ...$$

Il est evident que dans la pratique il est impossible d'effectuer la sommation d'une infinité de termes. On arrêtera donc lorsque le terme général $\frac{x^k}{k!}$ devient inférieur à 10^{-t}(on a t digits). Pour x négatif on sait que le reste de la serie est inférieur au premier terme négligé donc à 10^{-t}(puisque la serie est altérnée).

Les calculs suivant sont fait sur ordinateur pour $t = 14$.

x	e^x	S
-10	$4.54.10^{-5}$	$4.54.10^{-5}$
-15	$3.06.10^{-7}$	$3.05.10^{-7}$
-20	$2.06.10^{-9}$	$-1.55.10^{-7}$
-25	$1.39.10^{-11}$	$1.87.10^{-5}$
-30	$9.36.10^{-14}$	$6.25.10^{-4}$

On voit que pour $x \leq -20$ les résultats obtenus sont dépourvus de sens. L'explication de ce phénomène est la suivante : pour $x = -30$ les termes de la serie vont en croissant jusqu'à $\frac{x^{30}}{30!} = 8.10^{11}$ puis ils décroissent et $\frac{x^{107}}{107!} \tilde{} - 9.19.10^{-15}$.

L'erreur absolue sur le terme maximal est de $8.10^{11}.10^{-15} = 8.10^{-4}$. Ainsi le résultat obtenu pour S représente uniquement l'accumulation des erreurs d'arrondi sur les termes de plus grand module de développement en serie.

La propagation des erreurs est un des principaux problèmes en calcul numérique.

Considérons le cas d'une somme :

Dans l'addition, les erreurs absolues s'additionnent. Soit en effet ε_1 et $\varepsilon_2$les erreurs absolues sur x_1 et x_2 .

On peut écrire :

$(x_1 \pm \varepsilon_1) + (x_2 \pm \varepsilon_2) = (x_1 + x_2) \pm (\varepsilon_1 + \varepsilon_2)$

En arithmétique flottante, l'erreur relative δ est à peu près constante et les erreurs absolues peuvent être approximativement explicités par :

$\varepsilon_1 =\mid x_1 \mid \times\delta, \varepsilon_2 =\mid x_2 \mid \times\delta$

Si les nombres en présence ont le même signe, l'erreur relative reste la même que celle qu'on avait x_1 et x_2. En effet

$$\frac{\varepsilon_1 + \varepsilon_2}{|x_1 + x_2|} = \frac{(\mid x_1 \mid + \mid x_2 \mid) \times \delta}{|x_1 + x_2|} = \pm\delta$$

Si par contre les nombres sont de signes différents, l'erreur relative peut être amplifiée de façon spectaculaire.

Dans la multiplication, les erreurs relatives s'additionnent.

En effet,soient x_1 et x_2 ,on a :

$(x_1 \pm \varepsilon_1).(x_2 \pm \varepsilon_2) = x_1x_2 \pm x_1\varepsilon_2 \pm x_2\varepsilon_1 \pm \varepsilon_1\varepsilon_2$

Si de plus les erreurs s'écrivent

$\varepsilon_1 = |x_1| \times \delta$
$\varepsilon_2 = |x_2| \times \delta$

on a alors en négligeant certains termes :

$$\frac{|(x_1 \pm \varepsilon_1).(x_2 \pm \varepsilon_2) - x_1x_2|}{|x_1x_2|} = \frac{|\pm x_1\varepsilon_2 \pm x_2\varepsilon_1|}{|x_1x_2|} = \left|\pm\frac{\varepsilon_1}{x_1} \pm \frac{\varepsilon_2}{x_2}\right| \leq 2\delta$$

Des formules équivalentes peuvent donner des résultats différents ; on peut améliorer le résultat en utilisant une formule mathématique équivalente nécessitant des opérations différentes.

3.1. Exemple : *Si on considère les nombres* $\sqrt{7001}$ *et* $\sqrt{7000}$.
En arithmétique flottante à 8 *chiffres , on a :*
$\sqrt{7001}= 0.83671979 \times 10^2$
$\sqrt{7000}= 0.83666003 \times 10^2$
Donc

$$\sqrt{7001}-\sqrt{7000} = fl((0.83671979 - 0.83666003) \times 10^2) = 0.59760000 \times 10^{-2}$$

On peut obtenir un résultat plus précis en utilisant l'identité suivante :

$$\sqrt{x}-\sqrt{y}= (\sqrt{x}-\sqrt{y})\times\frac{\sqrt{x}+\sqrt{y}}{\sqrt{x}+\sqrt{y}}=\frac{x-y}{\sqrt{x}+\sqrt{y}}$$

On obtient alors

$$\frac{1}{\sqrt{7001}+\sqrt{7000}}=\frac{1}{0.16733798 \times 10^3} = 0.59759297 \times 10^{-2}$$

4. Conditionnement et stabilité numérique.

Le fait que certains nombres ne soient pas représentés de façon exacte dans un ordinateur entraine que l'introduction même de donnée d'un problème en machine modifie quelque peu le problème initial ; Il se peut que cette petite variation des données entraine une variation importante des résultats. C'est la notion de conditionnement d'un problème.

On dit qu'un problème est bien (ou mal) conditionné, si une petite variation des données entraine une petite (une grande) variation sur les résultats.

Cette notion de conditionnement est liée au problème mathématique lui même et est indépendante de la méthode utilisée pour le résoudre.

Une autre notion importante en pratique est celle de stabilité numérique. Un problème peut être bien conditionné et la méthode utilisée pour le résoudre peut être sujette à une propagation importante des erreurs numériques.

Ces notions de conditionnement d'un problème et de stabilité numérique d'une méthode de résolution sont fondamentales en analyse numérique. Si un problème est mal conditionné alors la solution exacte du problème tronqué ou arrondi à t digits pourra être très différente de la solution exacte du problème initial. Aucune methode ne pourra rien ; il faudra essayer de donner une autre formulation au problème.

5. Instabilité numérique :

Si les erreurs introduites dans les étapes intermédiaires ont un effet négligeable sur le résultat final,on dira que le calcul ou l'algorithme est numériquement stable. Si des petits changements sur les données entrainent des petits changements sur le résultat. Sinon , on dira que l'algorithme est numériquement instable.

5.1. Exemple. *On veut calculer la valeur de*

$$I_n= \int_0^1 \frac{x^n}{a+x}\,dx$$

où a *est une constante plus grande que* 1, *pour plusieurs valeurs de* n. *pour ce faire, nous allons exprimer* I_n *récursivement, i.e. nous allons exprimer* I_n *en fonction de* n *et* I_{n-1}.

$$
\begin{aligned}
I_n &= \int_0^1 \frac{x^{n-1}(x+a-a)}{a+x}\,dx\\
&= \int_0^1 x^{n-1}\,dx - a\int_0^1 \frac{x^{n-1}}{a+x}\,dx\\
&= \frac{1}{n} - aI_{n-1}\\
&= \sum_{i=0}^{n-1} \frac{(-a)^i}{n-i} + (-a)^n I_0
\end{aligned}
$$

Comme

$I_0 = ln(\frac{1+a}{a})$

On peut calculer I_n pour toutes les valeurs de n.
Mais l'algorithme est numériquement instable car toute erreur dans le calcul de$I_0 = ln(\frac{1+a}{a})$ va se propager.
En effet si on note par I_0^* la valeur approchée de I_0 et si $I_0^* = I_0 + \varepsilon$ alors

$$
\begin{aligned}
I_n^* &= \sum_{i=0}^{n-1} \frac{(-a)^i}{n-i} + (-a)^n\, I_0^*\\
&= \sum_{i=0}^{n-1} \frac{(-a)^i}{n-i} + (-a)^n\,(I_0+\varepsilon)
\end{aligned}
$$

donc $|I_n - I_n^*| \geq a^n\varepsilon$.

5.2. Remarques : *Il y a en fait différentes sources d'erreur. Nous pouvons les classer en 3 catégories :*
- les erreurs liées à l'imprécision des mesures physiques ou au résultat d'un calcul approché
- les erreurs liées à l'algorithme utilisé
- les erreurs de calcul liées à la machine

En général, pour l'objet de notre cours, si le premier chapitre met l'accent sur les erreurs liées à la machine, nous nous interresserons beaucoup plus aux erreurs liées aux méthodes ou encore aux algorithmes utilisés.

6. Série de Travaux Dirigés N I

Exercice 1 : Ecrire le nombre décimale 25,125 en une représentation binaire.

Exercice 2 : En arithmétique flottante avec 3 chiffres significatifs et arrondi, illustrer la non-validité des lois d'associativité et de distributivité.

(On pourra prendre : $x = 854$, $y = 251$ et $z = 852$).

Exercice 3 : En arithmétique flottante, avec $s = 2$, calculer $\sum_{i=1}^{10} \frac{1}{i^2}$.

1) En calculant : $\frac{1}{1} + \frac{1}{4} + \frac{1}{9} + ... \frac{1}{100}$
2) En calculant : $\frac{1}{100} + \frac{1}{81} + ... + \frac{1}{1}$

Quel résultat est le plus précis et pourquoi ?

Exercice 4 : On considère le polynôme $ax^2 + bx + c = 0$ $(a \neq 0)$. On suppose que le discriminant $\Delta \succ 0$. On sait que

(1) $\quad x_1 = \frac{-b+\sqrt{b^2-4ac}}{2a}$ et $x_2 = \frac{-b-\sqrt{b^2-4ac}}{2a}$

1) Vérifier que $x_1 + x_2 = -\frac{b}{a}$ et $x_1 * x_2 = \frac{c}{a}$.
2) Utiliser ce résultat pour montrer que ces racines peuvent aussi s'écrire sous la forme

(2) $\quad x_1 = \frac{-2c}{b+\sqrt{b^2-4ac}}$ et $x_2 = \frac{-2c}{b-\sqrt{b^2-4ac}}$

3) Pour $s = 4$ et trouver les racines de $x^2 + 53.1x + 1 = 0$ en calculant

i) x_1 à partir de (1) et la relation $x_1 * x_2 = \frac{c}{a}$.

ii) x_1 à partir de (2) et la relation $x_1 * x_2 = \frac{c}{a}$.

Quel calcul donne le meilleur résultat et pourquoi ?

iii) Si on clacule d'abord x_2, laquelle des formules (1) et (2) serait-il préférable de choisir et pourquoi ?

Exercice 5 : Résolvez ces deux systèmes linéaires :

$$(1) \qquad \begin{cases} x + y = 2 \\ x + 1.01y = 2.01 \end{cases}$$

$$(2) \qquad \begin{cases} x + y = 2 \\ x + 1.01y = 2.02 \end{cases}$$

Que remarquez-vous ?

Exercice 6 : On cherche les racines de

$p(x) = (x - 1)(x - 2)(x - 10)$

Elles sont évidentes.
Si on développe ce polynome en une valeur approchée pour l'une des racines par exemple 10.1
p(x) devient

$p(x) = (x - 10.1)(x^2 + bx + c)$

Calculer b et c .
En déduire les racines du polynomes du second degré. Que remarquez-vous ?

CHAPITRE 2

Résolution d'une équation f(x)=0

1. Introduction

Soit f une fonction numérique d'une variable réelle.
On cherche les racines simples de l'équation

$$(1) \qquad f(x) = 0$$

– Isoler les racines, c'est à dire trouver un intervalle $[a, b]$ dans lequel α est l'unique racine réelle de (1).
– Trouver cet intervalle : théorème des valeurs intermédiaires :
  – $f(a) * f(b) < 0$ $\quad f$ admet un nombre impair de racines
  – Si $f(a) * f(b) > 0$ $\quad f$ admet un nombre pair de racines
– $f'(x) = 1 + \frac{2e^x}{(e^x-2)^2}$ donc $f'(x) > 0$ pour tout x.
– L'équation a donc 2 racines simples situées de chaque côté de $ln(2)$.
– On vérifie sans problème qu'une première racine appartient à $[-1, 0]$ et la deuxième à $[1, 2]$

On supposera donc désormais avoir trouvé un intervalle $[a, b]$ où f admet une unique racine simple et on supposera que f est définie, continue, et autant de fois continument dérivable que nécessaire.

DÉFINITION 1. *Nous appellerons algoritnme toute méthode de résolution d'un problème donné.*
Pour tout problème, nous avons des données et des résultats.
– *Les données sont appelées paramètres d'entrée(input)*
– *les résultats paramètres de sortie (output)*
Ils constituent l'interface de l'algorithme.

Les algorithmes classiques que nous allons étudier sont les suivants :

(1) Méthode de la bissection
(2) Méthode de Newton-Raphson
(3) Méthode de la sécante
(4) Méthode du point fixe.

2. Méthode de la bissection.

Soit $f(x)$ continue et cherchons p tel que $f(p) = 0$.
– Supposons qu'on a localisé un intervalle $[a, b]$ dans lequel la fonction change de signe.
– on pose $c = \frac{a+b}{2}$,
  – si $f(a) * f(c) \prec 0$ on remplace b par c
  – sinon on remplace a par c,
– on continue cette operation jusqu'à ce qu'on trouve p avec la précision demandée.

BUT :
Trouver une approximation de la solution de $f(x) = 0$ dans $[a, b]$. en construisant une suite d'intervalles $([a_n, b_n])_n$ contenant la racine et tets que a_n ou b_n est le milieu de l'intervalle $[a_{n-1}, b_{n-1}]$.
Entrées : $a, b, \ \epsilon, N_0$
Sortie : la valeur approchée p : $f(p) = 0$

ALGORITHME 2.
(1) *Si $f(a) = 0$ imprimer la solution est a. Si $f(b) = 0$ imprimer la solution est b, aller à 10*
(2) *si $f(b) * f(a) > 0$, imprimer (pas de changement de signe). Aller à 10*

(3) *poser* $N = 1$

(4) *Tant que* $N \leq N_0$,*faire les étapes 5 à 8*

(5) *poser* $p = \frac{a+b}{2}$

(6) *Si* $f(p) = 0$ *ou* $\frac{b-a}{2} \leq \epsilon$, *imprimer* p. *Aller à 10*

(7) *poser* $N = N + 1$

(8) *Si* $f(a) * f(p) > 0$,*alors poser* $a = p$, *sinon poser* $b = p$

(9) *Imprimer aprés* N_0 *itérations l'approximation obtenue est* p *et l'erreur maximale est* $\frac{b-a}{2}$

(10) *Fin*

3. Méthode de Newton-Raphson :

Le principe consiste à construire une suite $(x_n)_n$, telle que x_{n+1} soit l'intersection de la tangente à la courbe de f au point $(x_n, f(x_n))$ avec l'axe horizontal.

On a :

$$\begin{cases} A = (x_0, f(x_0)), B = (x_1, 0) \in axe(Ox) \\ A \text{ et } B \in D : y = ax + b \end{cases}$$

donc

$$\begin{cases} f(x_0) = ax_0 + b \\ 0 = ax_1 + b \end{cases} \Rightarrow \begin{cases} a = f'(x_0) \\ x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \end{cases}$$

BUT

Entrées : une approximation initiale p_0

ε (la précision désirée)

N_0 (le nombre maximum d'itérations)

Sortie : valeur approchée de p ou un message d'échec

ALGORITHME 3. (1) $N = 1$

(2) *Tant que* $N \leq N_0$, *faire les étapes 3 à 6.*

(3) *Poser* $p = p_0 - \frac{f(p_0)}{f'(p_0)}$

(4) *Si* $\mid p - p_0 \mid \leq \varepsilon$ *alors imprimer* p, *aller à l'étape 8.*

(5) *Poser* $N = N + 1$.

(6) *Poser* $p_0 = p$.

(7) *Imprimer la méthode a échoué après* N *itérations.*

(8) *Fin.*

4. Méthode de la sécante

La méthode de Newton-Raphson suppose le calcul de $f'(p)$. On remplaçe dans la méthode de Newton $f'(p_n)$ par

$$\frac{f(p_n) - f(p_{n-1})}{p_n - p_{n-1}}.$$

L'équation de la sécante s'écrit :

$s(x) = f(p_n) + (x - p_n)\frac{f(p_n) - f(p_{n-1})}{p_n - p_{n-1}}$

Si $s(p_{n+1}) = 0$, on en déduit :

$p_{n+1} = p_n - f(p_n)\frac{p_n - p_{n-1}}{f(p_n) - f(p_{n-1})}$

5. Algorithme de la sécante :

BUT et Algorithme de la sécante
Trouver une solution de $f(x) = 0$
Entrées : deux approximations initiales p_0 et p_1
ε (la précision désirée)
N_0 (le nombre maximum d'itérations)
Sortie :la valeur approchée de p ou un message d'échec

(1) poser $N = 1$, $q_0 = f(p_0)$, $q_1 = f(p_1)$

(2) Tant que $N \leq N_0 + 1$,faire les étapes 3 à 6

(3) poser $p = p_1 - q_1 \frac{(p_1 - p_0)}{q_1 - q_0}$

(4) Si $\mid p - p_1 \mid \leq \varepsilon$ alors imprimer p, aller à l'étape 8

(5) Poser $N = N + 1$

(6) Poser $p_0 = p_1$, $q_0 = q_1$, $p_1 = p$,$q_1 = f(p)$

(7) Imprimer la méthode a échoué aprés N_0 itérations

(8) Fin

6. Méthode du point fixe

Nous pouvons observer que la méthode de Newton peut s'interpréter comme $p_{n+1} = g(p_n)$ où

$$g(x) = x - (\frac{f(x)}{f'(x)}).$$

Maintenant , si la fonction $g(x)$ est continue et si l'algorithme converge (c.à.d. $p_n \to p$), on tire de $p_{n+1} = g(p_n)$ que p satisfait l'équation $p = g(p)$; on dit que p est un point fixe de g.
BUT
trouver une solution de $g(x) = x$
Entrées : une approximation initiale p_0
ε(la précision désirée)

N_0 le nombre maximale d'itérations
Sortie : valeur approchée de p ou un message d'échec

6.1. Convergence et ordre de convergence.

DÉFINITION 4. *Soit D une partie de $\mathbb{R}$ et F une application de D dans D. On dit que la fonction F est contractante si*

$$\forall x, y \in D \ , \exists k \in [0, 1[\quad \textit{tel que}$$
$$\mid F(x) - F(y) \mid \leq k \ \mid x - y \mid .$$

k est le coéfficient de contraction ou de Lipschitz de F.

THEORÈME 5. *Considérons le segment $S = [p_0 - a, p_0 + a] \subset D$; si F est contractante sur S et si $\mid F(p_0) - p_0 \mid \leq (1 - k)\, a$, alors l'itération $p_{n+1} = F(p_n)$ de point initial p_0 , converge vers l'unique point fixe $p \in S$ de F.*

THEORÈME 6. *Si F est différentiable au voisinage d'un point fixe p et si $\mid F'(p) \mid < 1$ alors :*
$\exists V$ voisinage de p tels que $p_0 \in V$ et $p_{n+1} = F(p_n)$ converge vers p.

6.2. Ordre de convergence.

DÉFINITION 7. *Considérons une suite $\{p_n\}$ convergeant vers p et posons $e_n = p_n - p$. On dit dans le cas où $\left\{\left|\frac{e_n}{e_{n-1}}\right|\right\}$ converge, que la suite p_n converge linéairement vers p ou encore que la méthode est du premier ordre.*

Si on a $\left\{\left|\frac{e_n}{(e_{n-1})^k}\right|\right\}$ converge, alors la convergence est dite d'ordre k.

Exemple :

La méthode de Newton est une méthode de type point fixe avec

$$F(x) = x - \frac{f(x)}{f'(x)}.$$

Si x^* est racine simple de $f(x) = 0$, alors $f'(x^*) \neq 0$ et il existe un voisinage V de x^* tel que pour tout $p_0 \in V$, la suite $(p_n)_n$ converge vers x^* et l'ordre de convergence est 2.

Pour déterminer l'ordre de convergence on utilise la formule de Taylors en x^* : $F(x) = F(x^*) + F'(x^*)(x - x^*) + F''(\theta x)\frac{(x-x^*)^2}{2})$.

7. Série de travaux dirigés $N2$

Exercices 1

Résoudre à l'aide de la méthode de bisection $tanx - x = 0$ dans l'intervalle $[4; 4.7]$.

Exercice 2

On considère l'équation

(1) $e^x - 4x = 0$

1) Déterminer le nombre et la position approximative des racines de (1) situées dans .$x \geq 0$

2) Utiliser l'algorithme de bissection pour déterminer la plus petite de ces racines à ε près.(par exemple 10^{-7})

3) Sans faire d'itérations, déterminer combien vous devriez en faire pour calculer la plus grande racine à l'aide de la bissection avec une précision de 10^{-8}, si l'intervalle de départ est $[2; 2, 5]$

Exercice 3

Écrire un algorithme pour calculer par la méthode de Newton la racine k-ième d'un nombre.

Quelle est la valeur de $s = \sqrt{2 + \sqrt{2 + \sqrt{2 + ...}}}$?

Suggestion : écrire .p_{n+1}=G(p_n),p_0=0 Quel est l'ordre de convergence ?

Exercice 4

Écrire 3 méthodes itératives pour la résolution de $x^3 - x - 1$ =0 et vérifier expérimentalement leur convergence avec $x_0 = 1, 5$. Trouver à 10^{-6}près la racine comprise entre 1 et 2. Connaissant la valeur de cette racine, calculer l'ordre de convergence de vos 3 méthodes. Ce résultat coincide-t-il avec l'expérience ?

Exercice 5

Résoudre x^2-1=0 en utilisant la méthode de la sécante avec $x_0 = -3$ et $x_1 = 5/3$. Qu'arrivera-t-il si on choisit et $x_0 = 5/3$ et $x_1 = -3$? Expliquez.

Exercice 6

On considère la fonction $f(x) = x^4 - 2x^3 + 2x - 1$.

1) Calculer l'ordre de convergence de la méthode de Newton pour la racine $x^* = 1$.

2) Etudier la convergence et l'ordre de la méthode de points fixe suivante :

$$x_{n+1} = x_n - 3\frac{f(x_n)}{f'(x_n)}, \text{ avec } x_0 \text{ près de } x^* = 1 \text{ racine de } f.$$

CHAPITRE 3

Résolution de systèmes linéaires

1. Introduction

Un sytème linèaire s'écrit sous la forme :
$(1) \qquad Ax = b$
– A est une matrice $n \times n$ à coefficients réels,
– $b \in R^n$
– $x \in R^n$

Les méthodes de résolution sont de deux types :

(1) Les méthodes directes :
   – obtenir la solution en un nombre fini d'opérations.

(2) Les méthodes itératives :
   – construire une suite $(x_n)_n$ qui converge vers la solution.

Dans ce chapitre nous allons :

(1) Rapeler des notions et notations de base relatives aux systèmes linéaires et aux matrices
(2) Etudier une méthode directe : la méthode de Gauss.
(3) Etudier la décomposition (factorisation) LU.
(4) Etudier des applications : Inverse de matrices,...

2. Rappels sur les systémes linéaires

Si A est inversible alors (1) admet une solution unique

$$x = A^{-1}b$$

Ainsi théoriquement le problème revient à calculer A^{-1}? Mais en pratique ce calcul est difficile.

Méthodes classiques pour résoudre (1) sans calculer A^{-1}.

Exemple

$$(1) \qquad \begin{cases} x + 2y = 5 \\ 2x + y = 4 \end{cases}$$

2.1. La méthode de Cramer : $x = \dfrac{\begin{vmatrix} 5 & 2 \\ 4 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix}} = \frac{-3}{-3} = 1$ et $y = \dfrac{\begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix}} = \frac{-6}{-3} = 2$

2.2. La méthode de substitution (ou d'élimination). $\begin{cases} x + 2y = 5 \\ 2x + y = 4 \end{cases} \Rightarrow \begin{cases} x = -2y + 5 \\ 2x + y = 4 \end{cases} \Rightarrow \begin{cases} x = -2y + 5 \\ 2(-2y + 5) + y = 4 \end{cases}$

$$\Rightarrow \begin{cases} x = -2y + 5 \\ 3y = 6 \end{cases} \Rightarrow \begin{cases} x = -2y + 5 \\ y = 2 \end{cases} \Rightarrow \begin{cases} x = 1 \\ y = 2 \end{cases}$$

Peut-on généraliser ces méthodes pour un systéme de n équations avec $n \in \mathbb{N}$?
Théoriquement OUI mais en pratique cela va nécessiter beaucoup de calculs et de techniques.

3. Méthode Gauss

L'idée de base consiste à transformer (1) en un problème que l'on sait résoudre.

(1) Si la matrice $A = D$ avec D une matrice diagonale, alors on sait résoudre (1).

(2) Si la matrice $A = U$ (ou L) avec U (ou T) triangulaire supérieure (ou inférieure) alors on sait résoudre (1).

Problème : Comment tranformer une matrice en une matrice triangulaire inférieure ou supérieure ?

Cas $n = 3$:
$AX = b$ s'écrit :

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{array}\right.$$

la formeaugmentée

$$(A\ \ b) = \begin{pmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{pmatrix}$$

On suppose que$a_{11} \neq 0$, par élimination, on obtient :

$$(A_1\ \ b_1) = \begin{pmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ 0 & a'_{22} & a'_{23} & b'_2 \\ 0 & a'_{32} & a'_{33} & b'_3 \end{pmatrix}$$

$$(1)\left\{\begin{array}{ll} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 & (l1) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 & (l_2) \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 & (l_3) \end{array}\right.$$

On note par (l_i) la $i^{ème}$ équation du système précedent.
On suppose que $a_{11} \neq 0$,

On pose :

$(l'_2) = a_{11}(l_2) - a_{21}(l'_1)$ et $(l'_3) = a_{11}(l_3) - a_{31}(l'_1)$
Alors (1) s'écrit

$$(2)\left\{\begin{array}{rl} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 & (l1) \\ a'_{22}x_2 + a'_{23}x_3 = b'_2 & (l'_2) \\ a'_{32}x_2 + a'_{33}x_3 = b'_3 & (l'_3) \end{array}\right.$$

On suppose que $a'_{22} \neq 0$. On pose :
$(l''_3) = a'_{22}(l'_3) - a'_{32}(l''_2)$ Alors (2) s'écrit

$$(2)\left\{\begin{array}{rl} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 & (l1) \\ a'_{22}x_2 + a'_{23}x_3 = b'_2 & (l'_2) \\ a''_{33}x_3 = b''_3 & (l''_3) \end{array}\right.$$

Remarque :

(1) Les termes diagonaux sont appelés les pivots

(2) Si un pivot a_{ii} est nul on change de ligne (on permute) de i à n (pivotage partiel)

(3) Cette méthode se généralise assez facilement bien qu'il faut être prudent avec le choix du pivot. En pratique, il faut éviter de prendre des pivots "trop" petits.

3.1. Algorithme : Élimination de Gauss.

3.1.1. *Partie 1 : Réduction à la forme triangulaire.* Entrée A et b Sortie $A = U$ (forme triangulaire), et b.

- Pour $j = 1, ..., (n-1)$
- Pour $i = j+1, ..., n$
    - $l_{ij} \leftarrow \frac{a_{ij}}{a_{jj}}$
    - Pour $k = j+1, ..., n$
        - $a_{ik} \leftarrow a_{ik} - l_{ij}a_{jk}$
        - Fin
    - $b_j \leftarrow b_j - l_{ij}b_j$
- Fin
- Fin

Sortie $A = U$ (forme triangulaire), et b

(1) Poser $j = 1$

(2) Tant que $j \leq n - 1$ faire

(3) Si $a_{jj} = 0$ afficher 'pivot nul' aller à étape 14, sinon

(4) Poser $i = j + 1$

(5) Tant que $i \leq n$ faire

(6) $l_{ij} = \frac{a_{ij}}{a_{jj}}$

(7) Si $l_{ij} = 0$; aller à l'étape 12

(8) Poser $k = j + 1$

(9) Tant que $k \leq n$ faire

(10) $a_{ik} = a_{ik} - l_{ij} a_{jk}$, $k = k + 1$; aller à l'étape 9.

(11) $b_i = b_i - l_{ij} b_j$;

(12) poser $i = i + 1$; Aller à l'étape 5.

(13) $j = j + 1$; Aller à l'étape 2.

(14) Fin

Exemple :

$$\begin{cases} x + y + 3t = 4 \\ 2x + y - z + t = 1 \\ 3x - y - z + 2t = -3 \\ -x + 2y + 3z - t = 4 \end{cases}$$

qui s'écrit encore :

$$\begin{pmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ -3 \\ 4 \end{pmatrix}$$

Nous appliquons l'algorithme à notre exemple en travaillant sur la matrice augmentée.

$$\begin{bmatrix} & & A & & & b \\ 1 & 1 & 0 & 3 & . & 4 \\ 2 & 1 & -1 & 1 & . & 1 \\ 3 & -1 & -1 & 2 & . & -3 \\ -1 & 2 & 3 & -1 & . & 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 & 3 & . & 4 \\ 0 & -1 & -1 & -5 & . & -7 \\ 0 & -4 & -1 & -7 & . & -15 \\ 0 & 3 & 3 & 2 & . & 8 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 & 3 & . & 4 \\ 0 & -1 & -1 & -5 & . & -7 \\ 0 & 0 & 3 & 13 & . & 13 \\ 0 & 0 & 0 & -13 & . & -13 \end{bmatrix}$$

Que l'on peut écrire sous la forme :

$$\begin{cases} x + y + 3t &= 4 \\ -y - z - 5t &= -7 \\ 3z + 13t &= 13 \\ -13t &= -13 \end{cases}$$

Notons que l'étape $j = 3$ nous donnerait $l_{43} = 0$.
Nous avons maintenant un système triangulaire à résoudre.

3.1.2. *Partie 2 : Remontée triangulaire.* Entrée A, b avec A matrice triangulaire supérieure
Sortie x solution du sytème $Ax = b$

(1) $x_n = \frac{b_n}{a_{nn}}$
(2) Pour $i = n-1, n-2, ..., 1$ faire :

$$x_i = \frac{1}{a_{ii}}(b_i - \sum_{j=i+1}^{n} a_{ij}x_j)$$

En appliquant cet algorithme à notre exemple, nous obtenons $x = (-1, 2, 0, 1)$.
Remarque :

(1) Dans la pratique le test (3) de l'algorithme d'élimination de Gauss ne conduit pas à l'arrêt. En fait, si le pivot est nul, on cherche, dans la même colonne, un élément d'indice plus grand non nul, puis on échange les lignes correspondantes. Si ceci est impossible, le système est singulier.
(2) On est parfois amené, pour des raisons de stabilité numérique, à effectuer des échanges de lignes même si le test (3) est négatif (c'est à dire que le pivot est non nul). Ceci conduit à des stratégies dites de pivot que nous n'étudierons pas ici.

Exemple :

$$\left\{\begin{array}{rcr} 2x+6y+10z &=& 0 \\ x+3y+3z &=& 2 \\ 3x+14y+28z &=& -8 \end{array}\right. \Leftrightarrow$$

$$\begin{pmatrix} 2 & 6 & 10 \\ 1 & 3 & 3 \\ 3 & 14 & 28 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ -8 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 6 & 10 & 0 \\ 1 & 3 & 3 & 2 \\ 3 & 14 & 28 & -8 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & 6 & 10 & 0 \\ 0 & 0 & -4 & 4 \\ 0 & 5 & 13 & -8 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & 6 & 10 & 0 \\ 0 & 5 & 13 & -8 \\ 0 & 0 & -4 & 4 \end{pmatrix}$$

En utilisant la remonté on trouve :

$$\left\{\begin{array}{c} z = \frac{4}{-4} = -1 \\ y = \frac{1}{5}(-8 - 13 \times (-1)) = 1 \\ x = \frac{1}{2}(-6 \times 1 - 10 \times (-1)) = 2 \end{array}\right. \Rightarrow x^* = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$$

Méthode de Gauss avec normalisation :Elle consiste à normaliser le pivot : On a :

$$(1)\left\{\begin{array}{ll} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 & (l_1) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 & (l_2) \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 & (l_3) \end{array}\right.$$

On note par (l_i) la $i^{ème}$ équation du système précedent.
On suppose que $a_{11} \neq 0$,
(l_1) s'écrit : $x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3 = \frac{b_1}{a_{11}}$ (l'_1)
Si on pose :
$(l'_2) = (l_2) - a_{21}(l'_1)$
et
$(l'_3) = (l_3) - a_{31}(l'_1)$
Alors (1) s'écrit

$$(2)\left\{\begin{array}{rl} x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3 = \frac{b_1}{a_{11}} & (l'_1) \\ a'_{22}x_2 + a'_{23}x_3 = b'_2 & (l'_2) \\ a'_{32}x_2 + a'_{33}x_3 = b'_3 & (l'_3) \end{array}\right.$$

On suppose que $a'_{22} \neq 0$,
(l'_2) s'écrit $x_2 + a''_{23}x_3 = b''_2$ (l''_2)
Si on pose :
$(l''_3) = (l'_3) - a'_{32}(l''_2)$
si $a''_{33} \neq 0$ on pose (l'''_3) $x_3 = \frac{b''_3}{a''_{33}}$
Alors (2) s'écrit

$$(2)\left\{\begin{array}{r} x_1+\frac{a_{12}}{a_{11}}x_2+\frac{a_{13}}{a_{11}}x_3=\frac{b_1}{a_{11}} \quad (l_1') \\ x_2+a_{23}''x_3=b_2'' \quad (l_2'') \\ x_3=\frac{b_3''}{a_{33}''} \quad (l_3''') \end{array}\right.$$

4. Factorisation LU

Supposons que dans l'élimination de Gauss on n'utilise aucune stratégie de pivotage et que tous les pivots $a_{k,k}^{(k)} \neq 0$.

Dans ce cas le passage de $A^{(k)} \to A^{(k+1)}$ $(1 \leq k \leq N-1)$ revient à multiplier à gauche la matrice $A^{(k)}$ par la matrice $N \times N$:

$$E^{(k)}=\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & \cdots & \\ \vdots & \ddots & \ddots & & & & \vdots \\ \vdots & & 0 & 1 & \cdots & \cdots & 0 \\ \vdots & & \vdots & -l_{k+1,k} & \ddots & & \vdots \\ \vdots & & \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & -l_{N,k} & 0 & \cdots & 1 \end{pmatrix} \text{Avec } l_{i,k}=\frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}}, \text{ pour } k+1 \leq i \leq N.$$

La matrice $A^{(N)}$, qui est triangulaire supérieure, est alors égale à

$A^{(N)}=MA$

avec $M=E^{(N-1)}E^{(N-2)}...E^{(1)}$

M est le produit de matrices triangulaires inférieures, donc M est aussi triangulaire inférieure, on a $\det(M)=\prod_{i=1}^{N-1}\det(E^{(i)})=1$ et l'inverse de M est aussi triangulaire inférieure.

En posant $U=A^{(N)}$ et $L=M^{-1}$, on a $A=LU$.

$$(E^{(k)})^{-1}=\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & \cdots & \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & 0 & 1 & \ddots & \cdots & 0 \\ \vdots & & \vdots & l_{k+1,k} & \ddots & \ddots & \vdots \\ \vdots & & \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & l_{N,k} & 0 & ..0 & 1 \end{pmatrix}$$

$L=M^{-1}=(E^{(1)})^{-1}(E^{(2)})^{-1}.....(E^{(N-1)})^{-1}$

$$L=\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ l_{2,1} & 1 & 0 & \cdots & \cdots & \cdots & \\ \vdots & \ddots & \ddots & & & & \vdots \\ \vdots & & \ddots & 1 & \cdots & \cdots & 0 \\ \vdots & & \vdots & l_{k+1,k} & \ddots & & \vdots \\ \vdots & & \vdots & \vdots & & \ddots & \vdots \\ l_{N,1} & \cdots & \cdots & l_{N,k} & \cdots & \cdots l_{N,N-1} & 1 \end{pmatrix}$$

Exemple :

$$A=\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix}=A^{(1)}$$

$$E^{(1)}=\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \frac{4}{5} & 0 & 1 \end{pmatrix};A^{(2)}=\begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & \frac{18}{5} & \frac{9}{5} \end{pmatrix}$$

$$E^{(2)}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{9}{20} & 1 \end{pmatrix};A^{(3)}=\begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & \frac{9}{4} \end{pmatrix}=U$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{-4}{5} & \frac{-9}{20} & 1 \end{pmatrix} \; et \; A = LU$$

THEORÈME 8. *Soit $A = (a_{i,j})_{1 \leq i,j \leq N}$ une matrice carrée d'ordre N telle que les N sous-matrices de A :*

$$\begin{pmatrix} a_{11} & \cdots & \cdots & a_{1k} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{k1} & \cdots & \cdots & a_{kk} \end{pmatrix}, 1 \leq k \leq N$$

soient inversibles,
alors il existe une matrice triangulaire inférieure $L = (l_{ij})_{1 \leq i \leq N; 1 \leq j \leq N}$, avec
$l_{ii} = 1$ $(1 \leq i \leq N)$, et une matrice triangulaire supérieure U telles que $A = LU$.
De plus cette factorisation est unique.

Si $A = LU$ on peut résoudre : $Ax = b$ en résolvant

(1) $Lz = b$

(2) $Ux = z$

$$(1)\ Ax = b \Leftrightarrow \begin{cases} (2)\ Lz = b \\ (3)\ Ux = z \end{cases}$$

Dans ce cas $Ax = LUx = L(Ux) = Lz = b$.

Les systèmes (2) et (3) étant triangulaires, la résolution ne nécessite que l'exécution d'une remontée et d'une descente triangulaire.

Exemple :

$$\begin{pmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{pmatrix}.$$

Exemple :

$$A = LU = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{pmatrix}, \text{ on résoud le système } Ax = \begin{pmatrix} 4 \\ 1 \\ -3 \\ 4 \end{pmatrix}.$$

$$Lz = b \Rightarrow \begin{cases} z_1 & = & 4 \\ 2z_1 + z_2 & = & 1 \\ 3z_1 + 4z_2 + z_3 & = & -3 \\ -z_1 - 3z_2 + z_4 & = & 4 \end{cases} \Rightarrow z = \begin{pmatrix} 4 \\ -7 \\ 13 \\ -13 \end{pmatrix}.$$

$$Ux = z \Rightarrow \begin{cases} -13x_4 & = & -13 \\ 3x_3 + 13x_4 & = & 13 \\ -x_2 - x_3 - 5x_4 & = & -7 \\ x_1 + x_2 + 3x_4 & = & 4 \end{cases} \Rightarrow x = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}.$$

5. Mesure des erreurs

L'utilisation d'un calculateur pour implanter les algorithmes étudiés conduira inévitablement à des erreurs. Pour mesurer celles-ci, nous devons mesurer la distance entre le vecteur représentant la solution exacte $x = (x_1, ..., x_n)$ et le vecteur $\hat{x} = (\hat{x}_1, ..., \hat{x}_n)$ représentant la solution approchée. Nous pouvons, pour ce faire, utiliser la "longueur" usuelle de R^n i.e. :

$$\|x\|_2 = \{\sum_1^n x_i^2\}^{\frac{1}{2}}$$

pourtant, dans la pratique on lui préfère souvent la longueur

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

Par exemple si $x = (1, -7, 2, 4)$ alors $\|x\|_\infty = 7$.

Exemple :

Si $x = (1, 1, 1, 1)$ alors $\|x\|_\infty = 1$ si $\hat{x} = (1.01, 1.1, 1, 1)$, on a

$$\|x - \hat{x}\|_\infty = 0.1$$

Considérons alors le système

$$\begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 8 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}$$

dont la solution exacte est $x = (1, 1, 1, 1)$.

Si dans le membre de droite nous remplaçons b par :

$$\hat{b} = (32.06; 22.87; 33.07; 30.89)$$

nons obtenons

$$\hat{x} = (9.19; -12.59, 4.49, -1.09)$$

C'est-à-dire qu'une erreur relative de l'ordre de :

$$\frac{\|b - \hat{b}\|_\infty}{\|b\|_\infty} = 3 * 10^{-1}$$

sur b a entraîné une erreur relative de l'ordre de

$$\frac{\|x - \hat{x}\|_\infty}{\|x\|_\infty} = 13.52$$

sur la solution.

Nous devons donc soupçonner que l'application de l'arithmétique finie à la résolution d'un tel système serait désastreuse. L'étude de cette question dépasse le cadre de ce programme.

6. Méthodes itératives

On va voir un type de méthodes itératives de résolution du système linéaire $Ax = b$ sous la forme :

(3) $$\begin{cases} x^{(0)} \text{ vecteur arbitraire,} \\ x^{(k+1)} = Bx^{(k)} + c, k \geq 0 \end{cases}$$

lorsque

$$Ax = b \Longleftrightarrow x = Bx + c,$$

la matrice B et le vecteur c sont en fonction de A et b.

Définition

La méthode itérative (3) est convergente si

$$\lim_{k \to +\infty} x^{(k)} = x, \ \forall x^{(0)}.$$

Remarque :

Si on Pose $e^{(k)} = x^{(k)} - x$, pour $k = 0, 1, ...$

Comme $x = Bx + c$ et $x^{(k+1)} = Bx^{(k)} + c$,on a

$e^{(k)} = Bx^{(k-1)} - Bx = Be^{(k-1)} = ... = B^k e^{(0)}$

$\lim\limits_{k \to +\infty} x^{(k)} = x \Longleftrightarrow \lim\limits_{k \to +\infty} e^{(k)} = 0 \Longleftrightarrow \lim\limits_{k \to +\infty} B^k e^{(0)} = 0$

Donc

La méthode itérative (3) est convergente si

$$\lim_{k \to +\infty} B^k v = 0, \ \forall v$$

ce qui équivaut à

$$\lim_{k \to +\infty} ||B^k v|| = 0, \ \forall v$$

pour toute norme vectorielle $||.||$.

6.1. Convergence des méthodes itératives :

THEORÈME 9. *Les propositions suivantes sont équivalentes :*

(1)la méthode itérative (3) *est convergente ;*

(2)$\rho(B) < 1$;

(3)$||B|| < 1$ *pour au moins une norme matricielle subordonnée* $||.||$.

6.2. Méthode de Jacobi, de Gauss-Seidel. Ces méthodes sont des cas particuliers de la méthode suivante : $A = M - N$ avec M $inversible$ et $assez\ simple$. On aurait alors :

$$\begin{aligned} Ax = b &\Longleftrightarrow Mx = Nx + b \\ &\Longleftrightarrow x = M^{-1}Nx + M^{-1}b \\ &\Longleftrightarrow x = Bx + c, \end{aligned}$$

avec $B = M^{-1}N$ et $c = M^{-1}b$.

6.2.1. *Méthode de Jacobi :* En posant $A = D - (E + F)$,
$M = D$ est la diagonale de A et $N = E + F$.

$$Ax = b \Longleftrightarrow Dx = (E + F)x + b.$$

On suppose que D est inversible, c'est à dire $a_{ii} \neq 0, 1 \leq i \leq N$.
La matrice

$$J = D^{-1}(E + F) = I_N - D^{-1}A$$

est appelée la matrice de Jacobi.

$$\left\{ \begin{array}{c} x^{(0)} \text{ donné}, \\ Dx^{(k+1)} = (E + F)x^{(k)} + b, k \geq 0 \end{array} \right.$$

A chaque étape, on calcule les N composantes $x_1^{(k+1)}, .., x_N^{(k+1)}$ du vecteur $x^{(k+1)}$:

$$\left\{ \begin{array}{rcl} a_{1,1}x_1^{(k+1)} &=& -a_{1,2}x_2^{(k)} - ... - a_{1,N}x_N^{(k)} + b_1 \\ a_{2,2}x_2^{(k+1)} &=& -a_{21}x_1^{(k)} - a_{2,3}x_3^{(k)} ... - a_{2,N}x_N^{(k)} + b_2 \\ && \vdots \\ a_{N,N}x_N^{(k+1)} &=& -a_{N,1}^{(k)}x_1^{(k)} - ... - a_{N,N-1}^{(k)}x_{N-1}^{(k)} + b_N \end{array} \right.$$

6.2.2. *Méthode de Gauss-Seidel.* M est la partie triangulaire inférieure de A : $M = D - E$ et $N = F$.

On pourrait améliorer la méthode précédente en utilisant les quantités déjà calculées, on calcule successivent les N composantes
$x_1^{(k+1)}, .., x_N^{(k+1)}$ du vecteur $x^{(k+1)}$:

$$\left\{ \begin{array}{rcl} a_{1,1}x_1^{(k+1)} &=& -a_{1,2}x_2^{(k)} - ... - a_{1,N}x_N^{(k)} + b_1 \\ a_{2,2}x_2^{(k+1)} &=& -a_{21}x_1^{(k+1)} - a_{2,3}x_3^{(k)} ... - a_{2,N}x_N^{(k)} + b_2 \\ && \vdots \\ a_{N,N}x_N^{(k+1)} &=& -a_{N,1}x_1^{(k+1)} - . \quad .. - a_{N,N-1}x_{N-1}^{(k+1)} + b_N \end{array} \right.$$

qui revient à écrire :

$$Dx^{(k+1)} = Ex^{(k+1)} + Fx^{(k)} + b$$

ou encore

$$(D - E)x^{(k+1)} = Fx^{(k)} + b$$

$$\Updownarrow$$

$$x^{(k+1)} = (D - E)^{-1}Fx^{(k)} + (D - E)^{-1}b$$

$\mathcal{L}_1 = (D - E)^{-1}F$ est la matrice de Gauss-Seidel.
Elle est inversible si $a_{ii} \neq 0, 1 \leq i \leq N$.

Exemples : Etude de la convergence des méthodes de Jacobi et de Gauss-Seidel dans le cas où la matrice du système linéaire est :

1- $A_1 = \begin{pmatrix} -4 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & -4 \end{pmatrix}$;

Les deux méthodes convergent.

2- $A_2 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$;

la méthode de Jacobi diverge et La méthode de Gauss-Seidel converge.

3- $A = \begin{pmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}$.

la méthode de Jacobi converge et la méthode de Gauss-Seidel diverge.

subsection**convergence des méthodes de Jacobi, de Gauss-Seidel**

THEORÈME 10. *Soit A une matrice symétrique définie positve.*
On suppose que

$$A = M - N, \textit{ avec } M \textit{ inversible.}$$

Si la matrice symétrique

$$M^T + N$$

est définie positive,
alors

$$\rho(M^{-1}N) < 1.$$

Exemple :

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

1- Pour tout $u \in \mathbb{R}^N$,

$$u^T A u = u_1^2 + u_N^2 + \sum_{i=2}^{N} (u_i - u_{i-1})^2$$

$\Longrightarrow$ la matrice symétrique A est définie positive.
2- Pour la méthode de Jacobi :

$$M = D \text{ et } N = E + F :$$

$A = M - N$ est symétrique définie positive, car

$$u^T(M^T + N)u = u_1^2 + u_N^2 + \sum_{i=2}^{N} (u_i + u_{i-1})^2$$
$$> 0, \text{ si } u \neq 0.$$

$M^T + N$ est définie positive $\Longrightarrow \rho(M^{-1}N) < 1$.

CHAPITRE 4

Interpolation

1. Introduction

Nous abordons dans ce chapitre un nouveau type de problème, faisant intervenir la notion d'approximation d'une fonction.
Exemples :
1) D'aprés la Formule de Taylor à l'ordre 5 de la fonction $\sin(x)$, on a :

$$\forall x \in Vois(0), \sin(x) \simeq x - \frac{x^3}{3!} + \frac{x^5}{5!} + \sin^{(6)}(\xi)\frac{x^6}{6!}$$

L'erreur commise serait de l'ordre de $\sin^{(6)}(\xi)\frac{x^6}{6!}$ Ainsi :
– Si $N = 3, \sin(0.1) = (0.1) - \frac{(0.1)^3}{3!} = 9.9833 \times 10^{-2}$
– Si $N = 5, \sin(0.1) = 0.1 - \frac{(0.1)^3}{3!} + \frac{(0.1)^5}{5!} = 9.9833 \times 10^{-2}$
Avec le logiciel Maple on a : $\sin(0.1) = 9.9833 \times 10^{-2}$
2) Avec les cours d'analyse I et II, on ne connait pas d'expression explicite de $I = \int_0^1 e^{-x^2}dx$
Cependant d'aprés :
– La formule du trapèze $I = \int_0^1 e^{-x^2}dx \simeq \frac{f(0)+f(1)}{2} = \frac{1+e^{-1}}{2} = 0.68394$
– La formule de Simpson : $I = \int_0^1 e^{-x^2}dx \simeq \frac{1}{6}\left[f(0) + 4f(\frac{1}{2}) + f(1)\right] = \frac{1}{6}(1 + 4e^{-\frac{1}{4}} + e^{-1}) = 0.74718$
Avec le Logiciel Maple, on a : $\int_0^1 e^{-x^2}dx = 0.74682$
On ne connait pas à ce niveau du cours l'expression explicite de l'erreur.

La notion d'approximation d'une fonction consiste à remplacer un problème donné par un problème voisin.

La question fondamentale serais de savoir la qualité de cette approximation.

Remarque : En pratique la fonction f est connue explicitement, ou seulement par ses valeurs en quelques points.

La notion d'interpolation polynomiale est la façon la plus simple d'obtenir une telle approximation.

Théorème : Soit f une fonction continue dans $[a, b] \subset IR$, alors pour tout $\epsilon > 0$ donné, il existe un polynôme P_n de degré n tel que

$$\underset{x\in[a,b]}{Max} |f(x) - P_n(x)| < \epsilon$$

(1) L'interpolation polynômiale est un outil pour la construction des méthodes d'intégration numérique ou des méthodes d'approximation des équations différentielles.

(2) L'interpolation par les fonctions splines est largement utilisée dans tous les programmes de dessin assisté par ordinateur, conception assistée par ordinateur ou plus généralement de graphisme.

(3) Les séries de Fourier et leur analogue discret, la transformation de Fourier discrète : Elles sont un moyen très utile pour l'approximation des fonctions périodiques.

Remarque :
– Pour les équations aux dérivées partielles, la méthode des éléments finis, un des outils de base de l'ingénierie moderne, utilise de façon essentielle l'interpolation multi-dimensionnelle
– Une façon naturelle d'approcher les fonctions périodiques est d'utiliser les polynômes trigonométrique.

Nous allons nous limiter à l'introduction de l'interpolation Polynômiale. Elle consiste à déterminer un polynôme $P_n(x)$ de degré n qui puisse remplacer lors des applications la fonction $f(x)$.
De plus, c'est un outil efficace pour :
– Calculer, pour x donné, une approximation de $f(x)$ en calculant $P_n(x)$
– Construire :

(1) des méthodes d'intégration numérique

(2) des méthodes de différentation

(3) des méthodes d'approximation des équations différentielles

(4) ...

Le principe est simple, le procédé est le suivant :
– On choisit (ou on se donne) $(n+1)$ points $x_0, x_1, ..., x_n$.
– On calcule $y_0 = f(x_0), ..., y_n = f(x_n)$
ou on se donne (x_i, y_i), $i = 0, ..., n$.
– On cherche un polynôme de degré n tel que $P_n(x_i) = y_i$, $i = 0, ..., n$.

Remarque :

(1) Les points $(x_i, y_i)_{i=0,n}$ sont appelés points d'interpolation.

(2) Si la fonction f est connue seulement par ses valeurs en quelques points, les $(n+1)$ points $x_0, x_1, ..., x_n$ sont fixés..

(3) Si on veut que $P_m(x_i) = f(x_i)$ et $P'_m(x_i) = f'(x_i)$, $i = 0, ..., n$, on obtient l'interpolation dite d'Hermite

Il existe plusieurs techniques pour calculer $P_n(x)$. Les plus connues sont celles de Lagrange et de Newton-Côtes. Nous allons en fait le faire de deux façons :

(1) Une méthode directe basée sur la résolution d'un systéme linéaire

(2) Une méthode itérative due à Lagrange.

(3) Une méthode itérative d'Hermite.

Nous terminerons ce chapitre par :

(1) Une brève discution sur l'erreur d'interpolation polynomiale

(2) Une brève description du principe de la méthode itérée de Newton-Côtes

2. Une méthode directe basée sur la résolution d'un systéme linéaire :

– On se donne $(n+1)$ points $x_0, x_1, ..., x_n$.
– On calcule $y_0 = f(x_0), ..., y_n = f(x_n)$.
– On cherche un polynôme de degré n tel que $P_n(x_i) = y_i$, $i = 0, ..., n$.

Écrivons explicitement $P_n(x_i) = y_i$.

$$a_n x_i^n + a_{n-1} x_i^{n-1} + ... + a_1 x_i + a_0 = y_i, \qquad i = 0, ..., n$$

Sous forme matricielle :

$$\begin{pmatrix} x_0^n & x_0^{n-1} & \cdots & x_0 & 1 \\ x_1^n & x_1^{n-1} & \cdots & x_1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n^n & x_n^{n-1} & \cdots & x_n & 1 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_0 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Matrice de type Vandermonde. Son déterminant est

$$\det = \prod_{i<j} (x_i - x_j)$$

On a $\det \neq 0$ si tous les x_i sont distincts. On peut donc trouver un unique vecteur de coefficients $(a_n, ..., a_0)$ résolvant le problème.

3. Une méthode itérative : Méthode de Lagrange

3.1. Interpolation Linéaire : On considère deux points $(x_0, y_0), (x_1, y_1)$ avec :

$$\begin{cases} x_0 \neq x_1 \\ y_0 = f(x_0) \text{ et } y_1 = f(x_1). \end{cases}$$

Pour déterminer le polynôme $P_1(x) = ax + b$) qui passe par deux points distincts $(x_0, y_0), (x_1, y_1)$ $(x_0 \neq x_1)$. On peut :
1) Résoudre le système d'équations :

$$\begin{cases} ax_0 + b &= y_0 \\ ax_1 + b &= y_1 \end{cases}$$

d'où

$$\left\{\begin{array}{lll} a & = & \frac{(y_1-y_0)}{(x_1-x_0)} \\ b & = & y_0 - ax_0 = \frac{x_1y_0-x_0y_1}{x_1-x_0} \end{array}\right.$$

On a :

$$P_1(x) = \frac{(y_1-y_0)}{(x_1-x_0)}x + (\frac{x_1y_0-x_0y_1}{x_1-x_0})$$

et

$$P_1(x_0) = y_0 \text{ et } P_1(x_1) = y_1$$

2) Poser

$$L_0(x) = \frac{x-x_1}{x_0-x_1}$$
$$L_1(x) = \frac{x-x_0}{x_1-x_0}$$

On a :

$$L_k(x_i) = \left\{\begin{array}{lll} 0 & \text{si} & i \neq k \\ 1 & \text{si} & i = k \end{array}\right.$$

Ainsi,

$$\begin{aligned} P_1(x) &= y_0L_0(x) + y_1L_1(x) \\ &= y_0\frac{(x-x_1)}{(x_0-x_1)} + y_1(\frac{x-x_0}{x_1-x_0}) \\ &= \frac{(y_1-y_0)}{(x_1-x_0)}x + (\frac{x_1y_0-x_0y_1}{x_1-x_0}) \end{aligned}$$

On a :

$$P_1(x_0) = y_0 \text{ et } P_1(x_1) = y_1$$

car

$$L_k(x_i) = \left\{\begin{array}{lll} 0 & \text{si} & i \neq k \\ 1 & \text{si} & i = k \end{array}\right.$$

Ces deux procédés déterminent évidemment le même polynôme de dégré 1 (la même droite).

Si maintenant, on veut déterminer le polynome de degré 2 qui passe par trois (3) points distincts alors :

i) la première expression de $P_1(x)$ est inadéquate (il faut refaire les calculs)

ii) la dexième expression se prête assez facilement à une généralisation par récurrence.

Exemple :

Déterminer le polynôme d'interpolation $P_1(x)$ de degré 1 tel que
$P_1(x_i) = f(x_i), i = 0, 1$
avec $y_i = f(x_i)$ $i = 0, 1$, $(x_0, y_0) = (0, 1), (x_1, y_1) = (2, 5)$
D'aprés la méthode de Lagrange,

$$\begin{aligned} P_1(x) &= y_0L_0(x) + y_1L_1(x) \\ &= y_0\frac{(x-x_1)}{(x_0-x_1)} + y_1(\frac{x-x_0}{x_1-x_0}) \\ &= 1\frac{(x-2)}{(0-2)} + 5\frac{(x-0)}{(2-0)} \\ &= 1\frac{(x-2)}{(-2)} + 5\frac{(x)}{(2)} = 2x + 1 \end{aligned}$$

3.2. Interpolation parabolique. On considére trois points $(x_0, y_0), (x_1, y_1)$ et (x_2, y_2) avec :

$$\begin{cases} x_0 \neq x_1 \text{ ,et } x_0 \neq x_2 \;\; et \;\; x_1 \neq x_2 \\ y_0 = f(x_0), y_1 = f(x_1) \text{ et } y_2 = f(x_2). \end{cases}$$

Pour déterminer le polynôme $P_2(x)$ de dégré 2, d'équation $y = ax^2+bx+c$ qui passe par trois points distincts (x_0, y_0), (x_1, y_1) et (x_2, y_2), il suffit de poser :

$$L_0(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}$$
$$L_1(x) = \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}$$
$$L_2(x) = \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}$$

On a :

$$L_k(x_i) = \begin{cases} 0 & \text{si} \quad i \neq k \\ 1 & \text{si} \quad i = k \end{cases}$$

Ainsi

$$\begin{aligned} P_2(x) &= y_0L_0(x) + y_1L_1(x) + y_2L_2(x) \\ &= y_0\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + \\ & y_1\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + \\ & y_2\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} \end{aligned}$$

est le polynôme d'interpolation polynômiale associé.

Exemple :

Déterminer le polynôme d'interpolation $P_2(x)$ de degré 2 tel que
$P_2(x_i) = f(x_i), i = 0, 1$ et 2
avec $y_i = f(x_i)$ $i = 0, 1$ et 2 , $(x_0, y_0) = (0, 1), (x_1, y_1) = (1, 2)$ et $(x_2, y_2) = (2, 5)$

D'aprés la méthode de Lagrange,

$$\begin{aligned} P_2(x) &= y_0L_0(x) + y_1L_1(x) + y_2L_2(x) \\ &= y_0\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} \\ &+ y_1\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + y_2\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} \\ &= 1\frac{(x-1)(x-2)}{(-1)(-2)} + 2\frac{(x)(x-2)}{(1)(-1)} + 5\frac{(x)(x-1)}{(2)(1)} \\ &= 1\frac{(x-1)(x-2)}{2} + 2\frac{(x)(x-2)}{-1} + 5\frac{(x)(x-1)}{(2)(1)} \\ &= x^2+1 \end{aligned}$$

Remarque :

(1) Pour calculer $P_2(x)$,on n'a pas utilisé le polynome $P_1(x)$ calculé dans l'exemple précédent et pourtant on avait deux points communs.

(2) $L_i(x), i = 0, 1, 2$ sont des polynômes de degré 2 :
- $L_0(x) = \frac{(x-1)(x-2)}{(-1)(-2)} = \frac{1}{2}(x-1)(x-2) = \frac{1}{2}x^2 - \frac{3}{2}x + 1$
- $L_1(x) = \frac{(x)(x-2)}{(1)(-1)} = -x(x-2) = -x^2 + 2x$
- $L_2(x) = \frac{(x)(x-1)}{(2)(1)} = \frac{1}{2}x(x-1) = \frac{1}{2}x^2 - \frac{1}{2}x$

3.3. Interpolation de Lagrange.

- On choisit $n+1$ points $x_0, x_1, ..., x_n$.
- On calcule $y_0 = f(x_0), ..., y_n = f(x_n)$.
- On cherche un polynôme de degré n tel que $P_n(x_i) = y_i$, $i = 0, ..., n$.

On introduit les coefficients d'interpolation de Lagrange.

$$L_k(x) = \frac{(x-x_0)...(x-x_{k-1})(x-x_{k+1})...(x-x_n)}{(x_k-x_0)...(x_k-x_{k-1})(x_k-x_{k+1}...(x_k-x_n)}$$

$$L_k(x) = \prod_{j=0\ j\neq k}^{j=n} \frac{(x-x_j)}{(x_k-x_j)}$$

$L_k(x)$ est un polynôme de degré n,

$$L_k(x_i) = \begin{cases} 0 & \text{si} \quad i \neq k \\ 1 & \text{si} \quad i = k \end{cases}$$

Donc

$$P(x) = y_0 L_0(x) + y_1 L_1(x) + ... + y_n L_n(x) = \sum_{k=0}^{n} y_k L_k(x)$$

est un polynôme de degré n qui vérifie bien $P(x_i) = y_i$

Propriété : Le Polynôme d'interpolation polynômiale est unique.

En effet si $P(x)$ et $Q(x)$ sont deux polynômes d'interpolation alors :

$P(x) - Q(x)$ est un polynôme de degré n pour lequel

$$P(x_i) - Q(x_i) = 0, \qquad i = 0, ..., n.$$

Ce polynôme de degré $\leq n$ ayant $n+1$ racines, il est identiquement nul.

Exemple :

On suppose que $f(x) = \sqrt[3]{x}$ et que $(x_0, y_0) = (0,0), (x_1, y_1) = (1,1)$ et $(x_2, y_2) = (8,2)$

1) Déterminer le polynôme $P_2(x)$ d'interpolation polynômiale qui passent par les points $(x_i, y_i)_{i=0,2}$

D'aprés la méthode de Lagrange,

$$\begin{aligned} P_2(x) &= y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x) \\ &= y_0 \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} \\ &+ y_1 \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + y_2 \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} \\ &= 0 \frac{(x-1)(x-2)}{(0-1)(0-2)} + 1 \frac{(x-0)(x-8)}{(1-0)(1-8)} + 2 \frac{(x-0)(x-1)}{(8-0)(8-1)} \\ P_2(x) &= 1 \frac{x(x-8)}{-7} + 2 \frac{x(x-1)}{56} = -\frac{3}{28}x^2 + \frac{31}{28}x \end{aligned}$$

On a bien $P_2(0) = 0, P_2(1) = 1$ et $P_2(8) = -\frac{3}{28}(8)^2 + \frac{31}{28}8 = 2$

2) Calculer $P_2(x)$ et $f(x) = \sqrt[3]{x}$ pour $x = 0.5, 0.95, 1, 1.5$ et 3. On a :

x	$f(x)$	$P_2(x) = -\frac{3}{28}x^2 + \frac{31}{28}x$
0.5	0.7937	0.52679
0.95	0.98305	0.95509
1	1	1
1.5	1.1447	1.4196
3	1.4422	$\frac{33}{14} = 2.3571$

Remarque :

1) En pratique, on utilise l'interpolation polynômiale avec des polynômes de dégré n assez grand ou l'interpolation polynômiale par morceaux. Ainsi dans l'exemple précedent, il faut augmenter le nombre de points d'interpolations.

2) Si les valeurs y_k sont des valeurs expérimentales. L'interpolation polynomiale est une technique peu appropriée pour de telles situations. Les polynômes de degré élevé sont sensibles à la perturbation des données.

3) La méthode de Lagrange s'adapte mal au changement du nombre de points $(x_i, y_i)_i$. On ne peut utiliser les coefficients de Lagrange si on passe de n à $(n+1)$ points.

4) **Phénomène de RUNGE** (fonction de Runge) : L'interpolation polynômiale ne fournit pas une bonne approximation de la fonction $f(x) = \frac{1}{1+25x^2}$. Si on augmente le nombre de points d'interpolation le resultat devient plus mauvais.

4. Interpolation Itérée de Newton-Côtes

- On choisit $n+1$ points $x_0, x_1, ..., x_n$.
- On calcule $y_0 = f(x_0), ..., y_n = f(x_n)$.
- On cherche un polynôme de degré n tel que $P_n(x_i) = y_i$, $i = 0, ..., n$.

L'Interpolation Itérée de Newton-Côtes est un procédé itératif qui permet de calculer le polynôme d'interpolation $P_n(x)$ de dégré n basé sur $(n+1)$ points $(x_i, y_i)_{i=0,n}$ à partir du polynôme d'interpolation $P_{(n-1)}(x)$ de dégré $(n-1)$ basé sur n points $(x_i, y_i)_{i=0,(n-1)}$, en posant :

$$P_n(x) = P_{(n-1)}(x) + C(x),\ n \geq 1$$

avec

$$C(x) = a_n(x-x_0)(x-x_1)...(x-x_{(n-1)})$$
$$a_n = \sum_{k=0}^{n} \frac{f(x_k)}{(x_k-x_0)...(x_k-x_{(k-1)})(x_k-x_{(k+1)})...(x_k-x_n)}$$

Les coéfficients a_n sont appelés différences divisées d'ordre n de la fonction f, on note :

$$a_n = f\left[x_0, x_1, ..., x_n\right]$$

- On appelle "différence divisée d'ordre 0 de f en un point x" :

$$f\left[x\right] = f(x)$$

- Différence "divisée d'ordre 1 de f en x et y" :

$$\begin{aligned} f\left[x,y\right] &= \frac{f\left[x\right]-f\left[y\right]}{x-y} \\ &\mathit{on}\text{ a} \\ f\left[x,y\right] &= \frac{f(x)}{x-y}+\frac{f(y)}{y-x} \end{aligned}$$

- Différence "divisée d'ordre 2 de f en x, y et z" :

$$\begin{aligned} f\left[x,y,z\right] &= \frac{f\left[x,y\right]-f\left[y,z\right]}{x-z} \\ &= \frac{f(x)}{(x-y)(x-z)}+\frac{f(y)}{(y-x)(y-z)} \end{aligned}$$
$$+\frac{f(z)}{(z-x)(z-y)}$$

et plus généralement :

$$f\left[x_1, x_2, ..., x_n\right] = \sum_{i=1}^{n} \frac{f(x_i)}{\prod\limits_{\substack{k=1\\k\neq i}}^{n}(x_i-x_k)}$$

Remarque :
Les différences divisées sont indépendants de l'ordre des points.

Quel est le lien entre $f(x)$ et les différences divisées ?
Soit x un point autre que les $n+1$ points x_i, $i = 1, ..., n$. On a

$$\begin{aligned} f\left[x,x_0\right] &= \frac{f(x)-f\left[x_0\right]}{x-x_0} \\ &d'o\grave{u} \\ f(x) &= f\left[x_0\right]+(x-x_0)\, f\left[x,x_0\right] \end{aligned}$$

mais comme

$$f\left[x, x_0, x_1\right]=\frac{f\left[x, x_0\right]-f\left[x_0, x_1\right]}{x-x_1}$$

alors

$$f(x)=f\left[x_0\right]+\left(x-x_0\right) f\left[x_0, x_1\right]-\left(x-x_0\right)\left(x-x_1\right) f\left[x, x_0, x_1\right]$$

en continuant ainsi de proche en proche on obtient :

$$\begin{aligned}
f(x) &= f\left[x_0\right]+\left(x-x_0\right) f\left[x_0, x_1\right] \\
&+ \ldots+\left(x-x_0\right) \ldots\left(x-x_{n-1}\right) f\left[x_0, \ldots, x_n\right]+ \\
& \left(x-x_0\right) \ldots\left(x-x_n\right) f\left[x, x_0, \ldots, x_n\right]
\end{aligned}$$

on vérifie que

$$f(x)=P_n(x)+L(x) f\left[x, x_0, \ldots, x_n\right]$$

où $P_n(x)$ est un polynôme de degré n tel que $P_n(x_i)=f(x_i)$, pour $i=0, \ldots, n$. C'est donc le polynôme d'interpolation de Lagrange, on l'appelle le polynôme de Newton.

5. Erreur d'Interpolation polynomiale :

L'erreur commise lors d'une interpolation est une question fondamentale en analyse numérique :

- elle renseigne à priori sur la nature de cette erreur
- elle fournit des informations sur les termes qui y participent
- elle permet d'avoir un ordre de grandeur de l'erreur commise.

Théorème :
Soient f une fonction de classe C^{n+1} dans I et , $(x_i)_{i=0,n}$ $(n+1)$ points distincts dans I avec $x_0<x_1<\ldots<x_n$
Alors pour tout $x \in\left[x_0, x_n\right]$, il existe $\zeta=\zeta(x)$ tel que

$$f(x)-P_n(x)=\frac{f^{(n+1)}(\zeta)}{(n+1)!}\left(x-x_0\right)\left(x-x_1\right) \ldots\left(x-x_n\right)=\frac{f^{(n+1)}(\zeta)}{(n+1)!} L(x)$$

$$P_n(x)=y_0 L_0(x)+y_1 L_1(x)+\ldots+y_n L_n(x)=\sum_{k=0}^n y_k L_k(x)$$

$$L_k(x)=\prod_{j=0\ j \neq k}^n \frac{\left(x-x_j\right)}{\left(x_k-x_j\right)}$$

$$L(x)=\left(x-x_0\right)\left(x-x_1\right) \ldots\left(x-x_n\right)$$

Remarque :
1) Cette formule montre que :
   i) l'erreur est nulle pour $x=x_i$ i.e. x est un point d'interpolation.
   ii) l'erreur dépend de la fonction considérée (de $f^{(n+1)}$) et des points d'interpolations $(x_i)_i$.
2) Cette formule d'erreur permet de trouver des formules d'erreur pour l'intégration numérique et la differentiabilité numérique.

Dans le cas de l'erreur d'interpolation à partir de la forme de Newton, on a :

$$f(x)-P_n(x)=L(x) . f[x, x_0, \ldots, x_n] .$$

Comme on a la même fonction f selon les mêmes points x_i pour $i=0, \ldots, n$, il s'agit de deux formes du même polynôme, et l'erreur d'interpolation est la même, d'où

$$f(x)-P_n(x)=\frac{f^{(n+1)}(\zeta)}{(n+1)!} L(x)=L(x) . f[x, x_0, \ldots, x_n] .$$

Exemple :
Déterminer l'erreur d'interpolation polynomiale dans le cas de l'interpolation parabolique
On approche la fonction $f(x)$ par la parabole passant par les points $(x_0, y_0)=(0,1),(x_1, y_1)=(1,2)$ et $(x_2, y_2)=(2,5)$.
Le polynôme d'interpolation $P_2(x)$ de degré 2 tel que $P_2(x_i)=f(x_i), i=0,1$ et 2
avec $y_i=f(x_i)$ $i=0,1$ et 2 , $(x_0, y_0)=(0,1),(x_1, y_1)=(1,2)$ et $(x_2, y_2)=(2,5)$
D'aprés la méthode de Lagrange,

$$
\begin{aligned}
P_2(x) &= y_0L_0(x) + y_1L_1(x) + y_2L_2(x) \\
&= 1\frac{(x-1)(x-2)}{(-1)(-2)} + 2\frac{(x)(x-2)}{(1)(-1)} + 5\frac{(x)(x-1)}{(2)(1)} \\
&= x^2 + 1
\end{aligned}
$$

D'après le théorème précédent,

$$
\begin{aligned}
f(x) - P_2(x) &= \frac{f^{(3)}(\zeta)}{3!}(x-x_0)(x-x_1)(x-x_2) \\
&= \frac{f^{(3)}(\zeta)}{3!}x(x-1)(x-2)
\end{aligned}
$$

Si $\left|f'^{(3)}(x)\right| \leq M$ alors

$$
\begin{aligned}
&\forall x \in [0,2]\ ,\ \ |f(x) - P_2(x)| \\
|f(x) - P_2(x)| &\leq \frac{M}{6}\,|x(x-1)(x-2)| \\
&\leq \frac{M}{6}x(x-1)(x-2) \\
&\leq 6.4 * 10^{-2} * M.
\end{aligned}
$$

(le maximum de $u(x) = x(x-1)(x-2)$ est atteint en $x^* = \frac{3-\sqrt{3}}{3}$; d'où $\frac{1}{6}u(x^*) = \frac{1}{6}\frac{3-\sqrt{3}}{3}\left(\frac{3-\sqrt{3}}{3} - 1\right)\left(\frac{3-\sqrt{3}}{3} - 2\right) = 0.0\,641\,5 \sim 6.4 * 10^{-2}$).

6. Interpolation d'Hermite

L'interpolation de Lagrange, qui fournit facilement un polynôme prenant des valeurs données, présente l'inconvénient dans certains cas de donner une qualité médiocre : entre les points d'interpolation la différence entre une fonction et son polynôme d'interpolation peut être grande, même si le nombre de points tend vers l'infini (phénomène dit de Runge).
Pour remédier à cela on peut essayer d'utiliser non seulement les valeurs d'une fonction mais aussi celles de ses dérivées : c'est l'interpolation d'Hermite.

On se donne $(x_i, f^{(k)}(x_i))$, pour $i = 0, ..., n, k = 0, ..., m_i$ où $m_i \in \mathbb{N}$.
En posant $N = \sum_{i=0}^{n}(m_i + 1)$, on peut montrer que si les noeuds x_i sont distincts, il existe un unique polynôme $H_{N-1} \in \mathbb{P}_{N-1}$, appelé polynôme dinterpolation d'Hermite, tel que

$H_{N-1}^{(k)}(x_i) = y_i^{(k)}, i = 0, ..., n, k = 0, ..., m_i.$

Ce polynôme s'écrit
$H_{N-1}(x) = \sum_{i=0}^{n}\sum_{k=0}^{m_i} y_i^{(k)}L_{ik}(x)$ où $y_i^{(k)} = f^{(k)}(x_i), i = 0, ..., n, k = 0, ..., m_i$. Les fonctions $L_{ik}(x)$sont appelées polynômes caractéristiques d'Hermite et sont définies par les relations

$\frac{d^p}{x^p}(L_{ik})(x_j) = \begin{cases} 1\ si\ i = j, k = p \\ \quad 0 \text{ sinon} \end{cases}$

6.1. Interpolation d'Hermite, cas où $k = 1$.

En définissant les polynômes

$l_{i0}(x) = \prod_{\substack{s=0 \\ s\neq i}}^{n}(\frac{x-x_s}{x_i-x_s})^2, i = 0, ..., n$

$l_{i1}(x) = (x - x_i)\prod_{\substack{s=0 \\ s\neq i}}^{n}(\frac{x-x_s}{x_i-x_s})^2, i = 0, ..., n$

et en posant
$L_{i1}(x) = l_{i1}(x)$, pour $i = 0, ..., n$,
$L_{i0}(x) = l_{i0}(x) - l_{i0}^{'}(xi)L_{i1}(x).$

Soit f une fonction de classe C^{2n+2} sur un intervalle I contenant les points deux à deux distincts $x_i; i = 0; ...; 2n+1$, rangés dans l'ordre croissant. Alors pour tout $x \in [x_0; x_{2n+1}]$ il existe (au moins) un réel ξ_x dans ce même intervalle tel que :
Concernant l'erreur d'interpolation, on a l'estimation $f(x) - H_{2n+1}(x) = \frac{f^{(2n+1)}(\xi_x)}{2n+1!}\Omega_N(x)$, où $\xi_x \in I(x; x_0, ..., x_n)$ et Ω_N est le polynôme de degré (2n+1) défini par $\Omega_{2n+1}(x) = (x-x_0)^2(x-x_1)^2...(x-x_n)^2$.

6.1.1. *Interpolation d'Hermite : Exemple.* Soit f une fonction de classe C^2 sur l'intervalle I = [0 ; 1]. Soit H(x) un polynôme de degré 3 tel que :
$H(0) = f(0)$
$H(1) = f(1)$
$H'(0) = f'(0)$
$H'(1) = f'(1)$
Déterminons H. En appliquant les formules , on trouve
$H(x) = f(0)(1-x)^2(1+2x) + f(1)x^2(3-2x) + f'(0)x(1-x)^2 + f'(1)x^2(x-1)$

$z_i \;\; x_i$	$w_i \;\; y_i$	$d_i = \frac{w_{i+1}-w_i}{z_{i+1}-z_i}$	$dd_i = \frac{d_{i+1}-d_i}{z_{i+2}-z_i}$	$ddd_i = \frac{dd_{i+1}-dd_i}{z_{i+3}-z_i}$
$z_1 = x_1$	$w_1 = y_1$			
		$d_1 = y'_1$		
$z_2 = x_1$	$w_2 = y_1$		$dd_1 = \frac{d_2-d_1}{z_3-z_1}$	
		$d_2 = \frac{w_3-w_2}{z_3-z_2}$		$ddd_1 = \frac{dd_2-dd_1}{z_4-z_1}$
$z_3 = x_2$	$w_3 = y_2$		$dd_2 = \frac{d_3-d_2}{z_4-z_2}$	
		$d_3 = y'_2$		
$z_4 = x_2$	$w_4 = y_2$			

CHAPITRE 5

Dérivation et Integration

1. Introduction :

Si f est une fonction dérivable sur $[a,b]$, la dérivée en $c \in]a,b[$ est définie par :

$$f'(c) = \lim_{h\to 0} \frac{\Delta f(c)}{h}$$
$$\text{où } \Delta f(c) \ = f(c+h) - f(c)$$

Si f est une fonction continue sur $[a,b]$, l'integrale de f sur $[a,b]$ est définie par

$$\int_a^b f(x)dx = \lim_{h\to 0} R(h)$$
$$\text{où } \ R(h) = \sum_{k=1}^{n} f(a+kh).h$$

$R(h)$ est la somme de Riemann avec $h = \frac{b-a}{n}$.

– Il existe des fonctions simples comme $\frac{\sin x}{x}$ ou $\sqrt{\cos^2 x + 3\sin^2 x}$ qui n'ont pas de primitive connue.

– f peut-être connue seulement en quelques points et sa formule est inconnue (exp : résultats experimentaux,...),

Comment peut-on integrer de telles fonctions entre a et b?

Si $P(x)$ est une approximation de f dans l'intervalle $[a,b]$, nous nous proposons d'étudier les approximations :

$$f'(y) \approx P'(y) \qquad y \in [a,b]$$
$$et$$
$$\int_a^b f(x)dx \approx \int_a^b P(x)dx.$$

2. Dérivation.

2.1. Dérivée première. La dérivation numérique nous permet de trouver une estimation de la dérivée ou de la pente d'une fonction, en utilisant seulement un ensemble discret de points. Soit f une fonction connue seulement par sa valeur en $(n+1)$ points donnés $x_i \ \ i = 0, 1, ..., n$ distincts.

On suppose connue la valeur de la fonction en x_{i-1}, x_i et x_{i+1}; on pose $f(x_{i-1}) = y_{i-1}, f(x_i) = y_i$ et $f(x_{i+1}) = y_{i+1}$.

Si on suppose que l'espace entre deux points successifs est constant, donc on pose $h = x_i - x_{i-1} = x_{i+1} - x_i$.

Alors les formules standarts en deux points sont :

Formule de difference progressive :

$$f'(x_i) \approx \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}.$$

Formule de difference régressive :

$$f'(x_i) \approx \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} = \frac{y_i - y_{i-1}}{x_i - x_{i-1}}.$$

Formule de difference centrale :

$$f'(x_i) \approx \frac{f(x_{i+1}) - f(x_{i-1})}{x_{i+1} - x_{i-1}} = \frac{y_{i+1} - y_{i-1}}{x_{i+1} - x_{i-1}}.$$

Exemple :

Pour illustrer les trois formules, considérons les données suivantes :

$(x_0, y_0) = (1, 2); (x_1, y_1) = (2, 4); (x_2, y_2) = (3, 8); (x_3, y_3) = (4, 16)$ et $(x_4, y_4) = (5, 32)$.

Nous voulons estimer la valeur de $f'(x_2)$.

(1) Progressive : $f'(x) \approx \frac{f(x_3)-f(x_2)}{x_3-x_2} = \frac{16-8}{4-3} = 8.$

(2) Regressive : $f'(x) \approx \frac{f(x_2)-f(x_1)}{x_2-x1} = \frac{8-4}{3-2} = 4.$

(3) Centrale : $f'(x) \approx \frac{f(x_3)-f(x_1)}{x_3-x1} = \frac{16-4}{4-2} = 6.$

Les données ont été calculé pour la fonction $f(x) = 2^x$. $f'(x) = 2^x \ln(2)$ et pour $x = 3$ $f'(3) = 2^3 \ln(2) = 5.544$.

Remarque :

En utilisant la formule de Taylor :

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(\eta).$$
$$x \leq \eta \leq x+h$$

Formule progressive :

$$h = x_{i+1} - x_i$$
$$f'(x_i) = \frac{f(x_{i+1}) - f(x_i)}{h} - \frac{h}{2} f''(\eta)$$
$$x_i \leq \eta \leq x_{i+1}$$

l'erreur est $\frac{h}{2} f''(\eta)$ donc en $O(h)$.

Cette formule peut être trouvée aussi en utilisant le polynôme d'interpolation de Lagrange pour les points $(x_i, f(x_i))$ et $(x_{i+1}, f(x_{i+1}))$. Formule regressive :

$$h = x_i - x_{i-1}$$
$$f'(x_i) = \frac{f(x_i) - f(x_{i-1})}{h} + \frac{h}{2} f''(\eta)$$
$$x_{i-1} \leq \eta \leq x_i$$

La formule de différence centrale de la dérivée en x_i peut être trouvée en utilisant la formule de Taylor d'ordre 3 avec $h = x_{i+1} - x_i = x_i - x_{i-1}$

$$f(x_{i+1}) = f(x_i) + hf'(x_i) + \frac{h^2}{2} f''(x_i) + \frac{h^3}{3!} f'''(\eta_1)$$
$$f(x_{i-1}) = f(x_i) - hf'(x_i) + \frac{h^2}{2} f''(x_i) - \frac{h^3}{3!} f'''(\eta_2)$$
$$x_i \leq \eta_1 \leq x_{i+1}, \; x_{i-1} \leq \eta_2 \leq x_i$$

si on suppose que f''' est continue sur $[x_{i-1}, x_{i+1}]$ on peut ecrire la formule suivante :

$$f'(x_i) = \frac{f(x_{i+1}) - f(x_{i-1})}{2h} + \frac{h^2}{6} f'''(\eta)$$
$$x_{i-1} \leq \eta \leq x_{i+1}$$

l'erreur est $\frac{h^2}{6} f'''(\eta)$ donc en $O(h^2)$. La formule de différence centrale peut aussi être trouvée à partir du polynôme d'intérpolation de Lagrange en 3 points.

On peut interpoler les données par un polynome au lieu d'utiliser la droite, nous obtenons alors les formules de différence qui utilisent plus de deux points. On suppose que le pas h est constant.

Formule de différence progressive utilisant trois points :

$$f'(x_i) \approx \frac{-f(x_{i+2}) + 4f(x_{i+1}) - 3f(x_i)}{x_{i+2} - x_i}$$

Formule de différence régressive utilisant trois points :

$$f'(x_i) \approx \frac{3f(x_i) - 4f(x_{i-1}) + f(x_{i-2})}{x_i - x_{i-2}}$$

Exemple : Formules de différence en trois points :

En utilisant les données de l'exemple précédent, on trouve :

$f'(x_i) \approx \frac{-32+4(16)-3(8)}{2} = 4$ progressive.

$f'(x_i) \approx \frac{3(8)-4(4)+2}{2} = 5$ regressive.

2.2. Formule générale en trois points. La formule d'approximation en 3 points de la dérivée première, basée sur le polynôme d'interpolation de Lagrange, n'utilise pas des points équidistants.

Etant donné trois points (x_1, y_1); (x_2, y_2) et (x_3, y_3) avec $x_1 < x_2 < x_3$, la formule suivante permet d'approcher la dérivée en un point $x \in [x_1, x_3]$. Les dérivées aux points x_i sont les suivantes :

$$f'(x_1) = \frac{2x_1 - x_2 - x_3}{(x_1 - x_2)(x_1 - x_3)} y_1 +$$
$$\frac{x_1 - x_3}{(x_2 - x_1)(x_2 - x_3)} y_2 + \frac{x_1 - x_2}{(x_3 - x_1)(x_3 - x_2)} y_3;$$
$$f'(x_2) = \frac{x_2 - x_3}{(x_1 - x_2)(x_1 - x_3)} y_1 +$$
$$\frac{2x_2 - x_1 - x_3}{(x_2 - x_1)(x_2 - x_3)} y_2 + \frac{x_2 - x_1}{(x_3 - x_1)(x_3 - x_2)} y_3;$$
$$f'(x_3) = \frac{x_3 - x_2}{(x_1 - x_2)(x_1 - x_3)} y_1 +$$
$$\frac{x_3 - x_1}{(x_2 - x_1)(x_2 - x_3)} y_2 + \frac{2x_3 - x_2 - x_1}{(x_3 - x_1)(x_3 - x_2)} y_3;$$

Le polynôme de Lagrange est donnée par

$$P(x) = L_1(x)y_1 + L_2(x)y_2 + L_3(x)y_3$$

où

$$L_1(x) = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)}$$
$$L_2(x) = \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)}$$
$$L_3(x) = \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}$$

L'approximation de la dérivée première est donnée par $f'(x) \approx P'(x)$, qui peut s'ecrire

$$P'(x) = L_1'(x)y_1 + L_2'(x)y_2 + L_3'(x)y_3$$

où

$$L_1'(x) = \frac{2x - x_2 - x_3}{(x_1 - x_2)(x_1 - x_3)}$$
$$L_2'(x) = \frac{2x - x_1 - x_3}{(x_2 - x_1)(x_2 - x_3)}$$
$$L_3'(x) = \frac{2x - x_1 - x_2}{(x_3 - x_1)(x_3 - x_2)}$$

donc

$$f'(x) = \frac{2x - x_2 - x_3}{(x_1 - x_2)(x_1 - x_3)} y_1 + \frac{2x - x_1 - x_3}{(x_2 - x_1)(x_2 - x_3)} y_2 + \frac{2x - x_1 - x_2}{(x_3 - x_1)(x_3 - x_2)} y_3.$$

2.3. Dérivées d'ordre supérieur. Les formules de dérivées d'ordre supérieur, peuvent être trouvées à partir des dérivées du polynôme de Lagrange ou en utilisant les formules de Taylor.

Par exemple, étant donné 3 points x_{i-1}, x_i, x_{i+1} équidistants, la formule de la dérivée seconde est donnée par :

$$f''(x_i) = \frac{1}{h^2}[f(x_{i+1}) - 2f(x_i) + f(x_{i-1})]$$

l'erreur est en $O(h^2)$.

Dérivée seconde à partir du polynôme de Taylor.

$f(x + h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{3!} f'''(x) + \frac{h^4}{4!} f^{(4)}(\eta_1)$
$f(x - h) = f(x) - hf'(x) + \frac{h^2}{2} f''(x) - \frac{h^3}{3!} f'''(x) + \frac{h^4}{4!} f^{(4)}(\eta_2)$
$x \leq \eta_1 \leq x + h$ et $x - h \leq \eta_2 \leq x$.

$$f''(x) \simeq \frac{f(x + h) + f(x - h) - 2f(x)}{h^2}$$

l'erreur est en $O(h^2)$.

Pour obtenir les formules de la troisième et la quatrième dérivée, on prend une combinaison linéaire des développement de Taylor, pour $f(x+2h), f(x+h), f(x-h)$ et $f(x-2h)$.

La table suivante donne différentes formules centrales toutes en $O(h^2)$:

$$
\begin{aligned}
f'(x_i) &\simeq \frac{1}{2h}\left[f(x_{i+1}) - f(x_{i-1})\right] \\
f''(x_i) &\simeq \frac{1}{h^2}\left[f(x_{i+1}) - 2f(x_i) + f(x_{i-1})\right] \\
f'''(x_i) &\simeq \frac{1}{2h^3}\left[f(x_{i+2}) - 2f(x_{i+1}) + 2f(x_{i-1}) - f(x_{i-2})\right] \\
f^{(4)}(x_i) &\simeq \frac{1}{h^4}\left[f(x_{i+2}) - 4f(x_{i+1}) + 6f(x_i) - 4f(x_{i-1}) + f(x_{i-2})\right].
\end{aligned}
$$

En utilisant les polynômes d'interpolation de Lagrange les dérivées d'ordre p sont calculées par :

$$f^{(p)}(\alpha) \sim \sum_{i=0}^{n} A_i(\alpha) f(x_i)$$

où

$$
\begin{aligned}
A_i(\alpha) &= L_i^{(p)}(\alpha) \quad p \leq n \\
\sum_{i=0}^{n} A_i(\alpha) x_i^k &= 0 \quad 0 \leq k \leq p-1 \\
\sum_{i=0}^{n} A_i(\alpha) x_i^k &= k(k-1)...(k-p+1)\alpha^{k-p} \quad p \leq k \leq n.
\end{aligned}
$$

Remarque :

- **:** La fomule est exacte pour les polynômes de degrés $\leq n$.
- **:** Le système linéaire donnant les $A_i(\alpha)$ a un déterminant de type Vandermonde différent de zéro si les x_i sont distincts.
- **:** Les $A_i(\alpha)$ sont indépendants de f et peuvent être calculés une fois pour toutes.

2.4. Etude de l'erreur commise. D'aprés le chapitre précédent, si f est connue en $(n+1)$ points $x_i, i = 0, ..., n$ alors $f(x) = P_n(x) + e(x)$, où $e(x)$ est l'erreur d'interpolation. En dérivant on obtient

$$
\begin{aligned}
f'(x) &= P_n'(x) + e'(x) \\
&= \sum_{i=0}^{i=n} A_i(x).f(x_i) + e'(x) \\
\text{et} \quad e'(x) &= \frac{d}{dx}\left(\frac{1}{(n+1)!}L(x).f^{(n+1)}(\xi_x)\right) = \frac{d}{dx}\left(L(x).f[x_0, ..., x_n, x]\right) \\
&= \frac{1}{(n+1)!}L'(x).f^{(n+1)}(\xi_x) + \frac{1}{(n+1)!}L(x).\frac{d}{dx}\left(f^{(n+1)}(\xi_x)\right)
\end{aligned}
$$

Remarque

L'erreur de dérivation est nulle si f est un polynôme de degré inférieur ou égale à n.

Si on prend pour x un point x_i, le second terme de la dérnière somme s'annule, sinon il faut connaître $\frac{d}{dx}\left(f^{(n+1)}(\xi_x)\right)$, ce qui est difficile car la fonction $x \to \xi_x$ étant inconnue.

On peut donner une forme si f est $n+2$ fois dérivable en utilisant la notion de différence. En effet

$$
\begin{aligned}
\frac{d}{dx}\left(f^{(n+1)}(\xi_x)\right) &= \frac{d}{dx}\left(f[x_0, ..., x_n, x]\right) \\
&= \lim_{h\to 0} \frac{f[x_0, ..., x_n, x+h] - f[x_0, ..., x_n, x]}{h} \\
&= \lim_{h\to 0} f[x_0, ..., x_n, x, x+h] \\
&= \lim_{h\to 0} \frac{1}{(n+2)!} f^{(n+2)}(\theta_{x,h}).
\end{aligned}
$$

On constate qu'on devra se contenter d'une estimation

$$\mid e(x) \mid \leq \frac{1}{(n+1)!} \mid L'(x) \mid M_{n+1} + \frac{1}{(n+2)!} \mid L(x) \mid M_{n+2}.$$

3. Méthodes numériques d'intégration.

Soit $f : [a,b] \to \mathbb{R}$, une fonction continue donnée. On désire approcher numériquement la quantité $\int_a^b f(x)dx$.

3.1. Formules fermées. On appelle ainsi les formules quand la fonction f est continue sur lintervalle $[a,b]$.

Les points d'interpolation x_i verifient $a = x_0 < x_1 < ... < x_{n-1} < x_n = b$.

La formule des rectangles est une formule dite à un point $x_0 = a$. Le polynôme d'interpolation associé est $P_0(x) = f(a)$ et $L(x) = x - a$ pour tout x appartenant à $[a,b]$. D'où

$$I(f) \simeq I(P_0) = f(a)(b-a).$$

L'interprétation graphique consiste donc à remplacer $\int_a^b f(x)dx$ par l'aire du rectangle de base $[a,b]$ et de hauteur $f(a)$.

La formule des trapèzes est une formule à 2 points : $x_0 = a$ et $x_1 = b$. Le polynôme de Lagrange associé à ces deux points est $P_1(x) = f(a)\left(\frac{x-b}{a-b}\right) + f(b)\left(\frac{x-a}{b-a}\right)$.D'où

$$I(f) \simeq I(P_1) = \int_a^b P_1(x)dx = \frac{f(a)+f(b)}{2}(b-a).$$

La formule de Simpson est une formule à trois points $x_0 = a$, $x_1 = \frac{a+b}{2}$ et $x_2 = b$: . Le polynôme associé à ces trois points est $P_2(x) = f(a)L_0(x) + f(\frac{a+b}{2})L_1(x) + f(b)L_3(x)$. Notons que

$$L_0(x) = \frac{(x-x_1)(x-b)}{(a-x_1)(a-b)} \Rightarrow \int_a^b L_0(x)dx = \frac{(b-a)}{6},$$

$$L_1(x) = \frac{(x-a)(x-b)}{(x_1-a)(x_1-b)} \Rightarrow \int_a^b L_1(x)dx = \frac{4(b-a)}{6},$$

$$L_2(x) = \frac{(x-x_1)(x-a)}{(b-x_1)(b-a)} \Rightarrow \int_a^b L_2(x)dx = \frac{(b-a)}{6},$$

On tire donc la formule suivante :

$$I(f) \simeq I(P_2) = \frac{(b-a)}{6}\left[f(a) + 4f(\frac{a+b}{2}) + f(b)\right].$$

On appelle ainsi les formules quand la fonction f est continue sur l'intervalle $]a,b[$. Les points d'interpolation x_i verifient $a < x_0 < x_1 < ... < x_{n-1} < x_n < b$.

Il en existe une infinité.

– Une à 1 points avec $x_0 = \frac{a+b}{2}$:

$$I(f) \simeq (b-a)f(\frac{a+b}{2})$$

Cette formule est exacte pour tout polynôme de degré 1.

– Une à 2 points avec $x_0 = \frac{2a+b}{3}$ et $x_1 = \frac{a+2b}{3}$:

$$I(f) \simeq \frac{b-a}{2}\left(f\left(\frac{2a+b}{3}\right) + f\left(\frac{2b+a}{3}\right)\right).$$

Cette formule est exacte pour tout polynôme de degré 1.

– Une à 3 points avec $x_0 = \frac{3a+b}{4}$ et $x_1 = \frac{a+b}{2}$ et $x_2 = \frac{3b+a}{4}$:

$$I(f) \simeq \frac{b-a}{6}\left(4f\left(\frac{3a+b}{4}\right) + 2f\left(\frac{a+b}{2}\right) - 2f\left(\frac{a+3b}{4}\right)\right).$$

Cette formule est exacte pour tous les polynômes de degré 2.

3.2. Etude générale de l'erreur commise. Estimer l'erreur $E(f) = I(f) - I(P_n)$ avec précision ? ?.

Si f est suffisamment dérivable, on a

$$E(f) = I(f - P_n) = \int_a^b \left[\frac{1}{(n+1)!}f^{(n+1)}(\xi_x)L(x)\right]dx.$$

Théorème : Supposons que $E(f) = 0$ pour les polynômes de degré au plus n et que la fonction $f \in C^{n+1}([a,b])$. On dit alors que la méthode est d'ordre $n+1$. Si on pose $M_{n+1} = \max\limits_{x\in[a,b]} \mid f^{(n+1)}(x) \mid$, Une première estimation de l'erreur est

$$\mid E(f) \mid \leq \frac{1}{(n+1)!}M_{n+1}\int_a^b \mid L(x) \mid dx.$$

Théorème : En plus des hypothèses du Th précédent, on suppose que le polynôme $L(x)$ ne change pas de signe sur $[a, b]$, alors en utilisant le Th de la moyènne pour $E(f)$, on obtient

$$E(f) = \frac{1}{(n+1)!} f^{(n+1)}(\eta) \int_a^b L(x)dx.$$
$$\eta \in [a, b]$$

En utilisant ce dernier Théorème on peut estimer les erreurs des méthodes vues ci-dessus.

– **Pour la formule du rectangle on a :**

$$E(f) = f'(\eta) \int_a^b (x-a)dx = f'(\eta)\frac{(b-a)^2}{2} \qquad \eta \in [a, b]$$

cette méthode est d'ordre 1.

– **Pour la formule du trapèze on a :**

$$E(f) = \frac{1}{2} f''(\eta) \int_a^b (x-a)(x-b)dx = -\frac{f''(\eta)}{12}(b-a)^3$$

la méthode de Trapèze est dordre 2.

– **Pour la formule de Simpson on a :**

$$E(f) = -\frac{f^{(4)}(\eta)}{90} \left[\frac{b-a}{2}\right]^5,$$

la méthode de Simpson est dordre 4.

Exemple :

$I = \int_0^1 e^{-x^2} dx,\ a = 0, \frac{a+b}{2} = \frac{1}{2}, b = 1, f(0) = 1, f(\frac{1}{2}) = .7788, f(1) = .36788.$

(1) Rectangle : $I \simeq f(0) = 1.$

(2) Trapèze : $I \simeq \left[\frac{f(0)+f(1)}{2}\right] = .68393.$

(3) Simpson : $I \simeq \frac{1}{6}\left[f(0) + 4f(\frac{1}{2}) + f(1)\right] = .74718.$

(4) La valeur de I à 5 décimales est .74718.

3.3. Formules composées. Plutôt que d'augmenter le degré du polynôme d'interpolation, on peut obtenir une formule d'integration en découpant l'intervalle d'intégration en sous-intervalles et en appliquant des formules simples sur chacun des sous-intervalles.

Si n est entier, posons

$$h = \frac{b-a}{n}, x_k = a + kh,\ k = 0, ..., n.$$

alors

$$I(f) = \int_a^b f(x)dx = \sum_{k=0}^{n-1} \left(\int_{x_k}^{x_{k+1}} f(x)dx\right)$$
$$= \sum_{k=0}^{n-1} \left[\left(\frac{f(x_k) + f(x_{k+1})}{2}\right) h - \frac{h^3}{12} f''(\eta_k)\right],$$

où $\eta_k \in [x_k, x_{k+1}], k = 0, ..., n-1$

Développant et regroupant les termes qui apparaissent 2 fois, on obtient

$$I(f) = \frac{h}{2}\left[f(a) + 2\sum_{k=1}^{n-1} f(a+kh) + f(b)\right] - \frac{h^3}{12}\sum_{k=0}^{n-1} f''(\eta_k)$$

En appliquant le Th des valeurs intermédiaires, on peut réécrire l'erreur sous la forme

$$E(f) = -\frac{nh^3}{12} f''(\eta) = -\frac{(b-a)}{12} f''(\eta)h^2.$$

Ceci nous donne la formule du trapèze composée pour laquelle l'approximation est donnée par :

$$T_n(f) = \frac{h}{2}\left[f(a) + 2\sum_{k=1}^{n-1} f(a+kh) + f(b)\right]$$

et l'erreur par

$$ET(f) = -\frac{(b-a)}{12} f''(\eta) h^2.$$

Supposons que n soit pair, groupant les intervalles 2 à 2 et appliquant la formule de Simpson sur $[x_i, x_{i+2}]$, on obtient

$$I(f) = \frac{h}{3}\left[f(a) + 4\sum_{k\ impair} f(a+kh) + 2\sum_{k\ pair} f(a+kh) + f(b)\right] - \frac{n}{2}\frac{f^{(4)}(\eta)}{90} h^5.$$

Ceci nous conduit à la formule de Simpson composée pour laquelle l'approximation est donnée par

$$S_n(f) = \frac{h}{3}\left[f(a) + 4\sum_{k\ impair} f(a+kh) + 2\sum_{k\ pair} f(a+kh) + f(b)\right]$$

et l'erreur par

$$ES(f) = -f^{(4)}(\eta)\frac{(b-a)}{180} h^4.$$

Exemple : Déterminer $\int_0^1 e^{-x^2} dx$.

Si n désigne le nombre des intervalles utilisés.

n	$T_n(f)$	$ET(f)$
2	.73137	.015
4	.74298	3.84×10^{-3}
8	.74658	9.58×10^{-4}
16	.74676	1.39×10^{-4}
32	.74680	5.98×10^{-5}

Si nous désirons obtenir 6 décimales exactes, il nous faut déterminer h tel que

(1) $$\max_{0\leq\eta\leq 1} \mid f''(\eta) \mid \frac{h^2}{12} \leq 5 \times 10^{-7},$$

Pour une partition régulière $x_k = kh, h = \frac{1}{n}$; donc nous cherchons n tel que

$$n^2 \geq \frac{1}{12}\max_{0\leq\eta\leq 1} \mid f''(\eta) \mid \frac{1}{5 \times 10^{-7}}.$$

or $f''(x) = e^{-x^2}(4x^2 - 2)$ et $f'''(x) = e^{-x^2} 4x(3 - 2x^2)$. Puisque $f'''(x)$ ne change pas de signe sur $[0; 1]$,

$$\max_{0\leq\eta\leq 1} \mid f''(\eta) \mid = \max\left\{\mid f''(0) \mid, \mid f''(1) \mid\right\} = 2.$$

On voit que (1) sera satisfaite si

$$n^2 \geq \frac{10^6}{3}, \quad n > 578.$$

Remarque Dans le choix de la précision demandée, il faut tenir compte des erreurs d'arrondi et de l'accumutation des erreurs