

"COURS DE METHODES NUMERIQUES"

"Sophie DEPEYRE"

"Septembre 2025"

Table des matières

1	Problème de Cauchy pour les équations différentielles	3
1.1	Introduction	3
1.2	Problème de Cauchy	4
1.2.1	Définition 1	4
1.2.2	Première Méthode :	5
1.2.3	Deuxième Méthode :	5
1.3	Existence et unicité des solutions pour un problème de Cauchy	5
1.3.1	Définition 2 - Fonction lipschitzienne :	5
1.3.2	Définition 3 - Domaine de Calcul D convexe :	6
1.3.3	Définition d'une fonction lipschitzienne sur un domaine D convexe	9
1.3.4	Théorème d'existence et d'unicité de la solution d'un problème de Cauchy	10
1.4	Notions de problème bien posé	11
1.4.1	Définition 4 : Problème bien posé	12
1.4.2	Remarques :	12
1.4.3	Application	12
2	Méthodes numériques à un pas	16
2.1	Introduction	16
2.2	Discretisation de l'intervalle $[a, b]$	17
2.3	Méthode d'Euler	17
2.3.1	Construction de la Méthode d'Euler	17
2.3.2	Application	18
2.3.3	Erreur locale	19
2.3.4	Erreur d'approximation globale	21
2.4	Méthode de Taylor	22
2.4.1	Introduction	22
2.4.2	Construction de la Méthode de Taylor	22
2.4.3	Application	24
2.4.4	Précision de la Méthode de Taylor	25
2.5	Méthodes de Runge-Kutta	26
2.5.1	Introduction	26
2.5.2	Construction de la Méthode de Runge-Kutta 2	27
2.5.3	Précision de la Méthode de Runge-Kutta 2	28
2.5.4	Algorithme de la Méthode de Runge-Kutta 2	29
2.5.5	Méthodes de Runge-Kutta 3 et 4	29
2.5.6	Précision des Méthodes de Runge-Kutta	30
3	Propriétés des Méthodes numériques à un pas	31
3.1	Définitions générales	31
3.1.1	Définition 1 - Forme des Méthodes numériques à un pas	31
3.1.2	Remarques	31

3.2	Convergence d'une Méthode numérique à un pas	32
3.2.1	Définition 2 - Convergence	32
3.2.2	Théorème 1 - Convergence d'une Méthode numérique	32
3.3	Consistance d'une Méthode numérique	33
3.3.1	Introduction	33
3.3.2	Définition 3 - Erreur de troncature locale	33
3.3.3	Définition 4 - Consistance d'une Méthode numérique	33
3.3.4	Remarques :	33
3.3.5	Exemple	33
3.3.6	Théorème 2 - Outil pour vérifier la consistance :	35
3.3.7	Exemples	35
3.4	Stabilité théorique d'une Méthode numérique	35
3.4.1	Définition 5 - Stabilité théorique	35
3.4.2	Théorème 3 - Outil pour vérifier la stabilité théorique	36
3.4.3	Exemples	36
3.5	Ordre de précision	38
3.5.1	Définition 6 - Précision d'une Méthode numérique	38
3.5.2	Exemples	38
3.5.3	Erreur d'approximation d'une Méthode numérique	39
4	Interpolation polynômiale	40
4.1	Introduction	40
4.2	Interpolation polynômiale	41
4.2.1	Approximation polynômiale	41
4.2.2	Position du problème d'interpolation	41
4.2.3	Existence et unicité du polynôme d'interpolation	41
4.3	Interpolation de Lagrange	43
4.3.1	Base de Lagrange	43
4.3.2	Propriétés des polynômes de Lagrange	43
4.3.3	Polynôme d'interpolation de Lagrange	44
4.3.4	Exemples d'application	44
4.4	Interpolation de Newton	46
4.4.1	Base de Newton	46
4.4.2	Propriétés des polynômes de Newton	46
4.4.3	Polynôme d'interpolation de Newton	47
4.4.4	Différences divisées	47
4.4.5	Expression finale du polynôme d'interpolation de Newton	49
4.4.6	Réécriture du Polynôme de Newton	50
4.5	Erreur d'interpolation	53
4.5.1	Définition	53
4.5.2	Exemple	53
4.5.3	Phénomène de Runge	54

Chapitre 1

Problème de Cauchy pour les équations différentielles

1.1 Introduction

Dans le cours de Mathématiques de première année, ou dans le cours Compléments de formation Mathématiques, nous avons vu la résolution d'équations différentielles linéaires d'ordre un ou deux. Dans le cas d'équations différentielles linéaires à coefficients constants, ou d'ordre peu élevé, il s'avère possible de déterminer la solution analytique de l'équation différentielle.

Mais dans la plupart des problèmes modélisant des problèmes physiques, il s'avère impossible de déterminer la solution mathématique du problème, que ce soit dans le cas linéaire ou non linéaire.

Par exemple, pour l'équation différentielle linéaire du premier ordre :

$$y'(t) + 2t y(t) = 1,$$

on ne peut pas obtenir la solution analytique de l'équation différentielle.

Celle-ci ne peut pas s'exprimer en termes de fonctions élémentaires : polynômes, exponentielles, logarithmes, puissances, sinus/cosinus, etc...

En effet la solution homogène s'écrit :

$$y_h(t) = \lambda e^{-t^2}, \quad \lambda \in \mathbb{R}$$

Et lorsque l'on cherche la solution particulière y_p vérifiant l'équation différentielle avec second membre, on applique la Méthode de Variation de la Constante, et on obtient :

$$\lambda(t) = \int e^{t^2} dt$$

Or, on ne sait pas calculer une primitive élémentaire de e^{t^2} .

Donc, la solution générale de ce problème linéaire d'ordre un, s'écrit :

$$y(t) = e^{-t^2} \int_0^t e^{s^2} ds,$$

mais ne s'exprime pas analytiquement en termes de fonctions usuelles.

Dans le monde de la physique, la plupart des problèmes rencontrés s'expriment avec des équations différentielles non linéaires.

Par exemple, considérons un problème tout simple : quelle est la vitesse d'un objet qu'on lance en l'air ?

Le projectile vérifie l'équation différentielle suivante, où l'inconnue est sa vitesse en fonction du temps :

$$\begin{cases} mv'(t) = mg - k|v|v \\ v(t=0) = v_0 \end{cases}$$

où m désigne la masse du projectile, g la pesanteur et k le coefficient de frottement de l'air.

Il s'agit d'une équation différentielle d'ordre un, non-linéaire à cause du terme $|v|v$. Et on peut pas déterminer de solution analytique de ce problème, pourtant très simple, à cause de la non-linéarité $|v|v$.

On doit alors faire appel à des méthodes de résolution numérique, pour calculer une solution approchée au moyen d'ordinateurs.

Dans ce cours, nous allons présenter plusieurs méthodes numériques approchant ce type de problème physique. On calculera alors des solutions approchées, appelées solutions numériques.

La question qui vient alors, est de montrer que la solution numérique du problème approché, est bien proche de la solution exacte de l'équation différentielle, sous certaines conditions que l'on va expliciter dans ce cours.

Pour l'ensemble des chapitres, on considérera un problème à valeur initiale, c'est-à-dire un problème d'évolution d'une certaine variable $y(t)$ au cours du temps.

Le problème sera modélisé par une équation différentielle d'ordre 1, linéaire ou non linéaire. $y(t)$ pourra représenter par exemple : la position ou la vitesse d'un objet, ou alors la température d'un système. La variable t désignera le temps.

On adjoint à l'équation différentielle une condition initiale, précisant l'état du système à l'instant initial : par exemple, la position ou la vitesse d'un objet à l'instant initial.

1.2 Problème de Cauchy

1.2.1 Définition 1

On appelle **problème de Cauchy** le problème suivant :

Soit $\alpha \in \mathbb{R}$ donné.

Déterminer la fonction $y : [a, b] \longrightarrow \mathbb{R}$ tel que :

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(a) = \alpha \end{cases} \quad (1.1)$$

La variable t représente le temps, et l'inconnue est la fonction $y(t)$ qui dépend du temps.

Les données sont l'instant initial a , la condition initiale α , et la fonction : $f : [a, b] \times \mathbb{R} \longrightarrow \mathbb{R}$.

On connaît l'état du système à l'instant initial, grâce à la condition initiale.

Le problème de Cauchy modélise un problème d'évolution au cours d'un temps fini : t est borné car $t \in [a, b]$.

Comme on l'a vu dans l'Introduction, il existe deux méthodes pour résoudre ce problème.

1.2.2 Première Méthode :

La première méthode consiste à déterminer la solution analytique du problème de Cauchy, lorsqu'on le peut.

Exemple :

$$\begin{cases} y'(t) = t.y(t), & t \in [0, 1] \\ y(0) = 1 \end{cases}$$

Ici, on a : $a = 0$, $b = 1$, $\alpha = 1$ et $f(t, y(t)) = t.y(t)$. On peut résoudre ce problème explicitement facilement, et montrer que le système admet une unique solution :

$$y(t) = e^{t^2/2}.$$

On remarque que si on retire la condition initiale, la solution n'est plus unique.

Dans ce cas, il existe une infinités de solutions :

$$y(t) = \lambda e^{t^2/2}, \quad \lambda \in \mathbb{R}.$$

On constate dans cet exemple que la solution d'un problème de Cauchy, si elle existe, est unique.

1.2.3 Deuxième Méthode :

Comme on l'a vu, la plupart des équations différentielles ordinaires ne peuvent pas être résolues explicitement en termes de fonctions simples.

On va donc construire une méthode numérique qui approche le problème de Cauchy, et s'assurer que la solution approchée numérique est proche de la solution exacte.

1.3 Existence et unicité des solutions pour un problème de Cauchy

1.3.1 Définition 2 - Fonction lipschitzienne :

Soit $f(t, y)$ une fonction de deux variables t et y définies sur un ensemble $D \subset \mathbb{R}^2$.

S'il existe une constante $L > 0$ telle que :

$$\forall (t, y_1) \in D, \forall (t, y_2) \in D, \quad |f(t, y_1) - f(t, y_2)| \leq L |y_1 - y_2|,$$

alors on dira que la fonction $f(t, y)$ est lipschitzienne par-rapport à y , de constante L sur D .

La constante L est appelée Constante de Lipschitz pour la fonction f .

Exemple :

Soit $D \subset \mathbb{R}^2$ tel que :

$$D = \{ (t, y) \text{ tels que : } 1 \leq t \leq 2; \quad -3 \leq y \leq 4 \}$$

Soit la fonction $f(t, y) = t |y|$.

Montrer que la fonction $f(t, y)$ est lipschitzienne par-rapport à y sur D .

Démonstration :

Soient $(t, y_1) \in D$ et $(t, y_2) \in D$. Alors :

$$|f(t, y_1) - f(t, y_2)| = |t|y_1 - t|y_2| = |t| \cdot ||y_1| - |y_2|| \leq |t| \cdot |y_1 - y_2| \leq 2|y_1 - y_2|$$

car $1 \leq t \leq 2$.

La fonction $f(t, y)$ est lipschitzienne par-rapport à y de constante $L = 2$ sur D .

Remarque :

La constante de Lipschitz L est la plus petite valeur strictement positive telle que :

$$|f(t, y_1) - f(t, y_2)| \leq L |y_1 - y_2|, \quad \text{sur } D.$$

Rappel sur les inégalités triangulaires :

On a :

$$\forall (x_1, x_2) \in \mathbb{R}^2, \left| |x_1| - |x_2| \right| \leq |x_1 - x_2| \leq |x_1| + |x_2|$$

1.3.2 Définition 3 - Domaine de Calcul D convexe :

Un ensemble $D \subset \mathbb{R}^2$ est dit convexe si et seulement si :

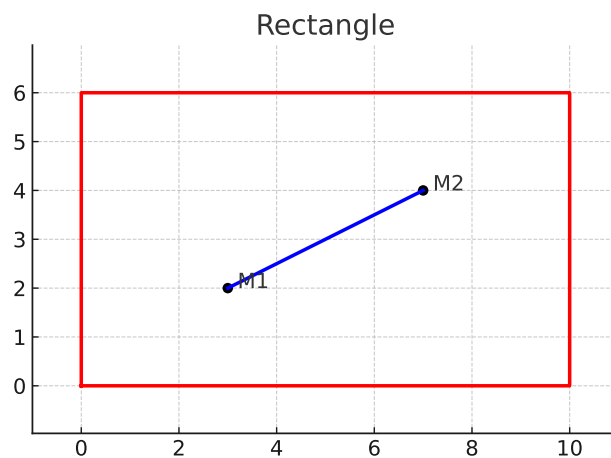
$$\forall (t_1, y_1) \in D, \forall (t_2, y_2) \in D, \forall \lambda \in [0, 1],$$

le point de coordonnées $((1 - \lambda)t_1 + \lambda t_2, (1 - \lambda)y_1 + \lambda y_2)$ appartient aussi à D .

Définition géométrique :

Le domaine D est convexe si et seulement si :

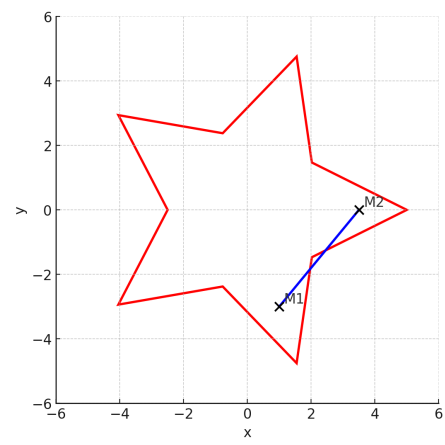
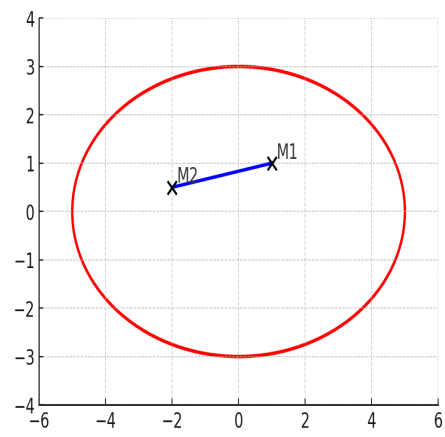
pour tout point $M_1 \in D$, et pour tout point $M_2 \in D$, le segment $[M_1, M_2]$ appartient à D .

Exemples de domaines convexes ou non convexes :

Un rectangle, un triangle, une ellipse, un cercle sont des domaines convexes.

Un domaine D en forme d'étoile n'est pas un domaine convexe.

Dès que le domaine D présente au moins un angle rentré, alors il n'est pas convexe.

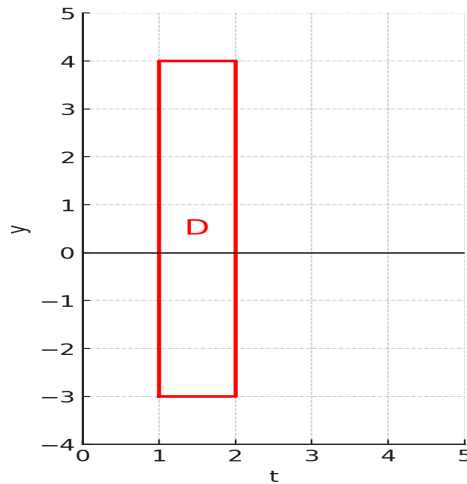


Exemple :

On considère le domaine suivant :

$$D = \{ (t, y) \text{ tels que : } 1 \leq t \leq 2 ; \quad -3 \leq y \leq 4 \}$$

On peut tracer le domaine D dans le repère (t, y) : on obtient un domaine de calcul rectangulaire.



Donc D est bien convexe.

Démontrons-le à l'aide de la définition mathématique :

Soient deux points M_1 et M_2 de D.

On pose $M_1 = (t_1, y_1) \in D$ et $M_2 = (t_2, y_2) \in D$.

Soit un paramètre $\lambda \in [0, 1]$.

Par définition du domaine de calcul D, on a : $1 \leq t_1 \leq 2$ et $-3 \leq y_1 \leq 4$.

Comme $\lambda \geq 0$, on a :

$$\begin{aligned} 1 - \lambda &\leq (1 - \lambda)t_1 \leq 2(1 - \lambda) \\ \lambda &\leq \lambda t_2 \leq 2\lambda. \end{aligned}$$

En additionnant les 2 inégalités terme à terme, on obtient le résultat attendu :

$$1 \leq (1 - \lambda)t_1 + \lambda t_2 \leq 2$$

De même, par définition du domaine D, on a : $-3 \leq y_1 \leq 4$ et $-3 \leq y_2 \leq 4$.

Comme $\lambda \geq 0$, on a :

$$\begin{aligned} -3(1 - \lambda) &\leq (1 - \lambda)y_1 \leq 4(1 - \lambda) \\ -3\lambda &\leq \lambda y_2 \leq 4\lambda. \end{aligned}$$

En additionnant les 2 inégalités terme à terme, on obtient ainsi :

$$-3 \leq (1 - \lambda)y_1 + \lambda y_2 \leq 4$$

Nous avons donc montré que le point de coordonnées : $((1 - \lambda)t_1 + \lambda t_2, (1 - \lambda)y_1 + \lambda y_2)$ appartient bien à D.

Donc D est bien convexe.

1.3.3 Définition d'une fonction lipschitzienne sur un domaine D convexe

Théorème 1 :

Soit $f(t, y)$ une fonction définie sur un ensemble convexe $D \subset \mathbb{R}^2$.

Supposons que la fonction $f(t, y)$ possède une dérivée partielle par-rapport à y , qui est continue sur D .

S'il existe une constante $L > 0$ telle que :

$$\left| \frac{\partial f}{\partial y}(t, y) \right| \leq L, \quad \forall (t, y) \in D,$$

alors la fonction $f(t, y)$ est lipschitzienne par-rapport à y , de constante L sur D .

Exemple :

Soit le domaine :

$$D = \{ (t, y) \text{ tels que : } 1 \leq t \leq 2 ; \quad -3 \leq y \leq 4 \}$$

Soit la fonction :

$$f(t, y) = t^2 \cos(t)y.$$

Montrons que la fonction $f(t, y)$ est lipschitzienne par-rapport à y sur D , à l'aide du Théorème 1.

D est bien convexe : nous l'avons démontré dans la section précédente.

La fonction $f(t, y)$ est de classe \mathcal{C}^∞ sur $\mathbb{R} \times \mathbb{R}$, donc sur D .

Elle possède donc une dérivée partielle par-rapport à y , $\frac{\partial f}{\partial y}$, qui est continue sur D .

On a :

$$\frac{\partial f}{\partial y}(t, y) = t^2 \cos(t)$$

et

$$\left| \frac{\partial f}{\partial y}(t, y) \right| = \left| t^2 \cos(t) \right| = t^2 |\cos(t)| \leq t^2 \leq 4,$$

car $1 \leq t \leq 2$.

Donc $f(t, y)$ est bien lipschitzienne par-rapport à y , de constante $L=4$, sur D .

Remarque :

Il est souvent très important de déterminer si la fonction $f(t, y)$ intervenant dans un problème de Cauchy est lipschitzienne par-rapport à y , et la condition donnée par le Théorème 1 est plus facile à appliquer que celle de la Définition 2.

Cependant, le Théorème 1 ne donne qu'une condition suffisante.

En effet, il faut que $f(t, y)$ possède une dérivée partielle par-rapport à y $\frac{\partial f}{\partial y}$, $\forall (t, y) \in D$, et que celle-ci soit continue sur D .

Mais il existe de nombreuses fonctions lipschitziennes qui ne sont pas de classe \mathcal{C}^1 , ou alors qui ne sont pas partout dérivables.

Ainsi, on peut montrer que la fonction $f(t, y) = t|y|$ satisfait la condition de Lipschitz avec la Définition 2 sur D :

$$D = \{ (t, y) \text{ tels que : } 1 \leq t \leq 2 ; \quad -3 \leq y \leq 4 \}.$$

Mais, on ne peut pas appliquer le Théorème 1, car $f(t, y)$ n'admet pas de dérivée partielle par-rapport à y , pour $y=0$ (car $|y|$ n'est pas dérivable en $y=0$).

1.3.4 Théorème d'existence et d'unicité de la solution d'un problème de Cauchy

Théorème 2 : Théorème de Cauchy - Lipschitz

Soit une fonction $f(t, y)$ définie sur un ensemble D convexe :

$$D = \{ (t, y) \text{ tels que : } a \leq t \leq b ; \quad -\infty < y < +\infty \}.$$

On suppose que :

1. $f(t, y)$ est continue sur D .
2. $f(t, y)$ est lipschitzienne par-rapport à y sur D .

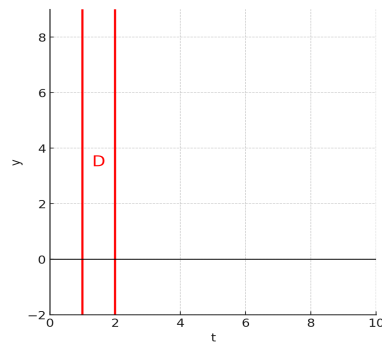
Alors le problème de Cauchy :

$$\begin{cases} y'(t) = f(t, y(t)); & a \leq t \leq b \\ y(a) = \alpha \end{cases} \quad (1.2)$$

admet une solution unique $y(t)$ de classe \mathcal{C}^1 sur $[a, b]$.

Remarque :

L'ensemble D est bien convexe : il s'agit d'un tube infini dans le repère (t, y) .



Pour tout point $M_1 \in D$, et pour tout point $M_2 \in D$, le segment $[M_1, M_2]$ appartient bien à D .

On fera la démonstration en TD.

Exemple :

Soit le problème de Cauchy :

$$\begin{cases} y'(t) = 1 + t \sin(ty); & 0 \leq t \leq 2 \\ y(0) = 0 \end{cases} \quad (1.3)$$

Soit l'ensemble :

$$D = \{ (t, y) \text{ tels que : } 0 \leq t \leq 2; \quad -\infty < y < +\infty \}.$$

Montrer que le problème de Cauchy (1.3) admet une solution unique $y(t)$.

Démonstration :

1. L'ensemble D est bien convexe : il s'agit d'un tube infini dans le repère (t, y) .
2. Soit $f(t, y) = 1 + t \sin(ty)$.
 $\sin(ty)$ est de classe \mathcal{C}^∞ sur $\mathbb{R} \times \mathbb{R}$, donc $f(t, y)$ est de classe \mathcal{C}^∞ sur $\mathbb{R} \times \mathbb{R}$, donc sur D .
Donc $f(t, y)$ est bien continue sur D .
3. Montrons que $f(t, y)$ est lipschitzienne par-rapport à y , sur D .
Comme f est de classe \mathcal{C}^∞ sur D , elle possède une dérivée partielle par-rapport à y , qui est continue sur D . On peut donc appliquer le Théorème 1 pour montrer que $f(t, y)$ est lipschitzienne par-rapport à y , sur D .

On a :

$$\frac{\partial f}{\partial y}(t, y) = t^2 \cos(ty)$$

et

$$\left| \frac{\partial f}{\partial y}(t, y) \right| = \left| t^2 \cos(ty) \right| = t^2 |\cos(ty)| \leq t^2 \leq 4,$$

car $0 \leq t \leq 2$.

Donc $f(t, y)$ est lipschitzienne par-rapport à y , de constante $L = 4$, sur D .

4. **Conclusion :** La fonction $f(t, y)$ vérifie toutes les hypothèses du Théorème 2, et l'ensemble D est bien convexe.

Donc le problème de Cauchy (1.3) admet une solution unique $y(t)$ de classe \mathcal{C}^1 sur $[0, 2]$

1.4 Notions de problème bien posé

Maintenant que nous avons vu quelles sont les conditions requises, pour que le problème de Cauchy admette une solution unique, nous nous posons une deuxième question.

Est-ce que de faibles changements dans le problème de Cauchy (1.2) n'entraînent qu'une variation dans la solution exacte $y(t)$, qui reste faible ?

En effet, notre but est de construire plusieurs méthodes numériques qui vont approcher le problème de Cauchy (1.2).

Et lorsque l'on approche un problème de Cauchy par une méthode numérique, on commet des

erreurs par-rapport au problème de Cauchy initial.

On souhaite alors s'assurer que la solution numérique obtenue avec la méthode numérique, soit proche de la solution exacte du problème de Cauchy.

On dira que le problème de Cauchy initial (1.2) est bien posé si de petites perturbations du problème de Cauchy (1.2) n'entraînent qu'une faible variation sur la solution exacte $y(t)$.

1.4.1 Définition 4 : Problème bien posé

Le problème de Cauchy :

$$\begin{cases} y'(t) = f(t, y(t)); & a \leq t \leq b \\ y(a) = \alpha \end{cases} \quad (1.4)$$

est dit bien posé si et seulement si :

1. Il admet une solution unique $y(t)$.
2. La solution $y(t)$ dépend continûment des données initiales, c'est-à-dire que $y(t)$ dépend continûment de la fonction $f(t, y)$ et de la condition α .

Comment traduit-on cette définition en mathématiques ?

Le problème sera dit bien posé si et seulement si :

$\exists K > 0, \forall \epsilon > 0, \forall \epsilon_0 > 0$, avec $|\epsilon_0| < \epsilon$, et $\forall \delta(t)$, une fonction continue qui satisfait $|\delta(t)| < \epsilon$,

le problème perturbé défini comme suit :

$$\begin{cases} z'(t) = f(t, z(t)) + \delta(t); & a \leq t \leq b \\ z(a) = \alpha + \epsilon_0 \end{cases} \quad (P)$$

admet une solution unique $z(t)$, et cette solution satisfait :

$$|z(t) - y(t)| \leq K \cdot \epsilon \quad \forall t \in [a, b].$$

1.4.2 Remarques :

1. La constante K doit être indépendante de ϵ et de t .
2. $\epsilon > 0$ est un réel infiniment petit, qui a pour vocation de tendre vers 0.
3. Pour obtenir le problème perturbé (P), on perturbe la fonction $f(t, y)$ en ajoutant une erreur $\delta(t)$, et on perturbe la condition initiale α en ajoutant une erreur ϵ_0 .
On impose : $|\delta(t)| < \epsilon$ et $|\epsilon_0| < \epsilon$ pour simuler de faibles perturbations numériques sur la fonction $f(t, y)$ et sur la condition initiale α .
4. Que signifie $|z(t) - y(t)| \leq K \cdot \epsilon$?
La solution numérique $z(t)$ du problème perturbé (P) doit converger vers la solution exacte du problème de Cauchy (1.4), lorsque ϵ tend vers 0.

1.4.3 Application

On considère le problème de Cauchy :

$$\begin{cases} y'(t) = y(t) - t^2 + 1; & 0 \leq t \leq 2 \\ y(0) = \frac{1}{2} \end{cases} \quad (1.5)$$

Montrer que le problème (1.5) admet une solution unique $y(t)$.

Démonstration :

Soit l'ensemble

$$D = \{ (t, y) \text{ tels que } : 0 \leq t \leq 2 ; \quad -\infty < y < +\infty \}.$$

1. D est bien convexe, car il s'agit d'un tube infini dans le repère (t, y) .
2. Soit $f(t, y) = y - t^2 + 1$.
 $f(t, y)$ est de classe \mathcal{C}^∞ sur $\mathbb{R} \times \mathbb{R}$, donc sur D. Donc $f(t, y)$ est bien continue sur D.
3. Comme $f(t, y)$ est de classe \mathcal{C}^∞ sur D, elle possède une dérivée partielle par-rapport à y, qui est continue sur D. On a :

$$\frac{\partial f}{\partial y}(t, y) = 1$$

et

$$\left| \frac{\partial f}{\partial y}(t, y) \right| = 1 \leq 1.$$

Donc $f(t, y)$ est lipschitzienne par-rapport à y, de constante L=1, sur D.

4. Conclusion : Le problème de Cauchy (1.5) admet une solution unique $y(t)$.

On considère le problème perturbé :

$$\begin{cases} z'(t) = z(t) - t^2 + 1 + \delta; & 0 \leq t \leq 2, & |\delta| < \epsilon \\ z(a) = \alpha + \epsilon_0, & & |\epsilon_0| < \epsilon \end{cases} \quad (P)$$

1. Calculer la solution exacte du problème de Cauchy (1.5).
2. Calculer la solution du problème perturbé (P).
3. Montrer qu'il existe $K > 0$, $\forall \epsilon > 0$, $|z(t) - y(t)| \leq K.\epsilon \quad \forall t \in [0, 2]$.

Démonstration :

1. Solution exacte du Problème de Cauchy :

On résout :

$$y' - y = -t^2 + 1$$

Equation homogène :

$$y' - y = 0$$

La solution homogène s'écrit : $y_h(t) = \lambda e^t$, $\lambda \in \mathbb{R}$

Equation avec second membre :

$$y' - y = -t^2 + 1$$

On cherche une solution particulière y_p .

Comme le second membre est un polynôme de degré 2, la solution particulière y_p est aussi un polynôme de degré 2.

On pose : $y_p = at^2 + bt + c$.

On a : $y'_p = 2at + b$.

On remplace y_p et y'_p dans l'équation avec second membre :

$$2at + b - at^2 - bt - c = -t^2 + 1 \iff -at^2 + (2a - b)t + (b - c) = -t^2 + 1$$

On identifie et on obtient un système linéaire :

$$\begin{cases} -a = -1 \\ 2a - b = 0 \\ b - c = 1 \end{cases}$$

Ce qui nous donne :

$$\begin{cases} a = 1 \\ b = 2 \\ c = 1 \end{cases}$$

D'où y_p s'écrit :

$$y_p(t) = t^2 + 2t + 1 = (t + 1)^2$$

Solution générale :

$$y(t) = y_h(t) + y_p(t) = \lambda e^t + (t + 1)^2, \quad \lambda \in \mathbb{R}$$

On obtient une infinité de solutions $y(t)$ car λ varie dans \mathbb{R} .

Pour avoir une solution unique, on applique la condition initiale du Problème de Cauchy.

$$y(0) = \frac{1}{2} \iff \lambda + 1 = \frac{1}{2} \iff \lambda = -\frac{1}{2}$$

Conclusion : La solution exacte du problème de Cauchy (1.5) s'écrit :

$$y(t) = -\frac{1}{2}e^t + (t + 1)^2$$

2. Solution exacte du Problème perturbé (P) :

On résout :

$$z' - z = -t^2 + 1 + \delta, \quad |\delta| < \epsilon$$

Equation homogène :

$$z' - z = 0$$

La solution homogène s'écrit : $z_h(t) = \lambda e^t, \quad \lambda \in \mathbb{R}$

Equation avec second membre :

$$z' - z = -t^2 + 1 + \delta$$

On cherche une solution particulière z_p .

Comme le second membre est un polynôme de degré 2 et $\delta \in \mathbb{R}$, la solution particulière z_p est aussi un polynôme de degré 2.

On pose : $z_p = at^2 + bt + c$.

On a : $z'_p = 2at + b$.

On remplace z_p et z'_p dans l'équation avec second membre :

$$2at + b - at^2 - bt - c = -t^2 + 1 + \delta \iff -at^2 + (2a - b)t + (b - c) = -t^2 + 1 + \delta$$

On identifie et on obtient un système linéaire :

$$\begin{cases} -a = -1 \\ 2a - b = 0 \\ b - c = 1 + \delta \end{cases}$$

Ce qui nous donne :

$$\begin{cases} a = 1 \\ b = 2 \\ c = 1 - \delta \end{cases}$$

D'où z_p s'écrit :

$$z_p(t) = t^2 + 2t + 1 - \delta = (t+1)^2 - \delta$$

Solution générale :

$$z(t) = z_h(t) + z_p(t) = \lambda e^t + (t+1)^2 - \delta, \quad \lambda \in \mathbb{R}$$

On obtient une infinité de solutions $z(t)$ car λ varie dans \mathbb{R} .

Pour avoir une solution unique, on applique la condition initiale du Problème perturbé (P).

$$z(0) = \frac{1}{2} + \epsilon_0 \iff \lambda + 1 - \delta = \frac{1}{2} + \epsilon_0 \iff \lambda = -\frac{1}{2} + \epsilon_0 + \delta$$

Conclusion : La solution exacte du problème perturbé (P) s'écrit :

$$z(t) = \left(\delta + \epsilon_0 - \frac{1}{2} \right) e^t + (t+1)^2 - \delta$$

3. Montrons que le problème de Cauchy (1.5) est bien posé :

Il faut montrer :

$$\exists K > 0, \forall \epsilon > 0, |z(t) - y(t)| \leq K \cdot \epsilon \quad \forall t \in [0, 2].$$

Calculons $|z(t) - y(t)|$:

$$|z(t) - y(t)| = \left| \left(\delta + \epsilon_0 - \frac{1}{2} \right) e^t - \delta + \frac{1}{2} e^t \right| = \left| (\delta + \epsilon_0) e^t - \delta \right|$$

Or :

$$\left| (\delta + \epsilon_0) e^t - \delta \right| \leq \left| (\delta + \epsilon_0) e^t \right| + |\delta| \leq \left| \delta + \epsilon_0 \right| e^t + |\delta|$$

Or $0 \leq t \leq 2$ et $e^0 = 1 \leq e^t \leq e^2$

Donc on a :

$$|z(t) - y(t)| \leq (|\delta| + |\epsilon_0|) e^2 + |\delta|$$

Comme $|\delta| < \epsilon$ et $|\epsilon_0| < \epsilon$, on obtient :

$$|z(t) - y(t)| \leq (2\epsilon e^2 + \epsilon)$$

et donc :

$$|z(t) - y(t)| \leq (2e^2 + 1) \epsilon$$

Donc : $K = 2e^2 + 1 > 0$.

On a bien trouvé une constante K indépendante de ϵ et de t , telle que :

$$\forall \epsilon > 0, |z(t) - y(t)| \leq K \cdot \epsilon$$

Conclusion : La solution $z(t)$ du problème perturbé (P) converge vers la solution exacte du problème de Cauchy $y(t)$, quand ϵ tend vers 0.

Le problème de Cauchy est bien posé, car de petites perturbations sur la fonction $f(t, y)$ et sur la condition initiale α , n'entraînent que de faibles variations sur la solution du problème de Cauchy $y(t)$.

Chapitre 2

Méthodes numériques à un pas

2.1 Introduction

On considère le problème de Cauchy suivant :

$$\begin{cases} y'(t) = f(t, y(t)); & a \leq t \leq b \\ y(a) = \alpha \end{cases} \quad (2.1)$$

Nous avons vu que dans la plupart des problèmes, on ne sait pas calculer la solution exacte du problème de Cauchy (2.1).

Nous allons donc construire une méthode numérique, appelée encore schéma numérique, qui permet d'obtenir une solution approchée de la solution du problème de Cauchy $y(t)$.

On suppose que la solution $y(t)$ est une fonction suffisamment régulière.

L'idée est simple : afin d'approcher la solution $y(t)$ pour t variant dans l'intervalle $[a, b]$, il suffit de connaître certaines valeurs de la solution en certains points.

On suppose donc que l'on connaît la solution $y(t_i)$, en certains points $t_i \in [a, b]$.

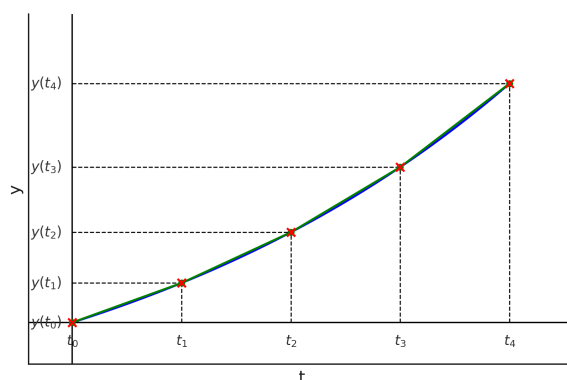


FIGURE 2.1 – Approximation numérique d'une fonction

En reliant le point $y(t_i)$ au point $y(t_{i+1})$, on obtient une approximation de la courbe représentative de la fonction $y(t)$.

2.2 Discrétisation de l'intervalle [a,b]

Une Méthode numérique repose donc sur un principe de discrétisation de l'intervalle $[a, b]$.

On subdivise l'intervalle $[a, b]$ en n sous-intervalles $]t_i, t_{i+1}]$ pour $i = 0, 1, 2, \dots, n - 1$.

On choisit une subdivision uniforme de l'intervalle $[a, b]$.

On pose $h = t_{i+1} - t_i$, $\forall i = 0, 1, 2, \dots, n - 1$. Le pas h est constant, et on l'appelle pas de discrétisation de l'intervalle $[a, b]$, ou encore pas de temps.



FIGURE 2.2 – Discrétisation de l'intervalle [a,b]

On voit sur la Figure 2 qu'il y a $n + 1$ points de discrétisation t_i .

On définit la subdivision de l'intervalle $[a, b]$ par :

$$\begin{cases} t_0 = a \\ t_{i+1} = t_i + h, \quad \forall i = 0, 1, \dots, n - 1 \\ t_n = b \end{cases}$$

On a donc :

$$t_i = t_0 + ih = a + ih, \quad \forall i = 0, 1, \dots, n$$

Nous remarquons sur la Figure 2.1, que plus le pas de discrétisation h est petit, meilleure sera l'approximation de la solution $y(t)$.

On choisira toujours un pas de discrétisation h très petit devant 1 : $h \ll 1$.

2.3 Méthode d'Euler

La première méthode numérique que nous allons voir s'appelle la Méthode d'Euler. Cette méthode est très largement utilisée dans les codes de calculs, car elle est facile à programmer.

2.3.1 Construction de la Méthode d'Euler

La Méthode d'Euler repose sur un développement limité d'ordre 1 de la solution $y(t)$ au point t_i .

On a donc :

$$y(t) = y(t_i) + y'(t_i)(t - t_i) + o(t - t_i)$$

On va réécrire le développement limité pour $t = t_{i+1}$:

$$y(t_{i+1}) = y(t_i) + y'(t_i)(t_{i+1} - t_i) + o(t_{i+1} - t_i)$$

Or, par définition du pas de discrétisation h , on a :

$$h = t_{i+1} - t_i$$

On peut donc écrire :

$$y(t_{i+1}) = y(t_i) + hy'(t_i) + o(h) \quad (2.2)$$

Or $y(t)$ est la solution du problème de Cauchy qui vérifie :

$$y'(t) = f(t, y(t)), \quad a \leq t \leq b$$

Donc pour $t = t_i$, on a :

$$y'(t_i) = f(t_i, y(t_i)) \quad (2.3)$$

En remplaçant l'équation (2.3) dans l'équation (2.2), il vient :

$$y(t_{i+1}) = y(t_i) + hf(t_i, y(t_i)) + o(h)$$

Définition 1 - Méthode d'Euler

La Méthode d'Euler construit la suite d'approximations numériques w_i avec $0 \leq i \leq n$, qui approche les valeurs de la solution exacte $y(t_i)$ telle que :

$$\begin{cases} w_{i+1} = w_i + hf(t_i, w_i) & \forall i = 0, 1, 2, \dots, n-1 \\ w_0 = y(t_0) = \alpha \end{cases} \quad (2.4)$$

2.3.2 Application

On considère le problème de Cauchy suivant :

$$\begin{cases} y'(t) = y(t) - t^2 + 1; & 0 \leq t \leq 2 \\ y(0) = \frac{1}{2} \end{cases} \quad (2.5)$$

Ce problème admet une solution $y(t)$ que l'on peut calculer :

$$y(t) = (t+1)^2 - \frac{1}{2}e^t$$

Nous allons approcher cette solution $y(t)$ par la Méthode d'Euler.

On a :

$$f(t, y) = y - t^2 + 1$$

On choisit comme pas de discrétisation $h = 0.2$.

Comment s'écrit la discrétisation de l'intervalle $[0, 2]$ avec un pas $h = 0.2$?

On a donc :

$$\begin{cases} t_0 = 0 \\ t_i = t_0 + ih = 0 + 0.2i = 0.2i, \quad \forall i = 0, 1, \dots, 10 \end{cases}$$

Appliquons l'Algorithme de la Méthode d'Euler (2.4) en faisant varier i de 0 à 10.

$$w_0 = y(t_0) = y(0) = \frac{1}{2}$$

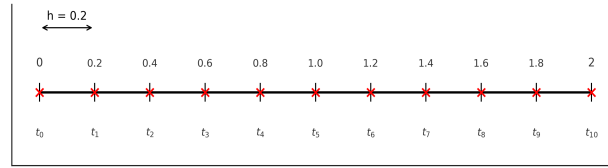


FIGURE 2.3 – Discrétisation de l'intervalle $[0,2]$

Pour initialiser la Méthode d'Euler, on applique à w_0 la valeur de la condition initiale du problème de Cauchy $\alpha = \frac{1}{2}$.

Calculons w_1 :

$$w_1 = w_0 + hf(t_0, w_0) = w_0 + h.(w_0 - t_0^2 + 1) = \frac{1}{2} + 0.2 \left(\frac{1}{2} - 0 + 1 \right) = 0.8$$

De la même façon, on calcule w_2 :

$$w_2 = w_1 + hf(t_1, w_1) = w_1 + h.(w_1 - t_1^2 + 1) = 0.8 + 0.2 (0.8 - 0.2^2 + 1) = 1.152$$

Calculons w_3 et w_4 :

$$w_3 = w_2 + hf(t_2, w_2) = w_2 + h.(w_2 - t_2^2 + 1) = 1.152 + 0.2 (1.152 - 0.4^2 + 1) = 1.5504$$

$$w_4 = w_3 + hf(t_3, w_3) = w_3 + h.(w_3 - t_3^2 + 1) = 1.5504 + 0.2 (1.5504 - 0.6^2 + 1) = 1.9884$$

On peut ainsi, de proche en proche, calculer les valeurs w_5, w_6, w_7, w_8, w_9 , et w_{10} .

La question que l'on se pose alors, est la suivante :

Est-ce que la valeur obtenue par la Méthode d'Euler w_i est proche de la solution exacte du problème de Cauchy au point t_i : $y(t_i)$?

Pour cela, on définit l'erreur locale au point t_i .

2.3.3 Erreur locale

On appelle erreur locale au point t_i :

$$e_i = |y(t_i) - w_i|, \quad \forall i = 0, 1, 2, \dots, n$$

Dans l'exemple précédent, on résume les résultats obtenus dans un tableau pour $i = 0, \dots, 5$:

t_i	Solution exacte $y(t_i)$	Solution numérique w_i	Erreurs locales $ y(t_i) - w_i $
0	$\frac{1}{2}$	$\frac{1}{2}$	0
0.2	0.829298	0.8	0.029298
0.4	1.214087	1.152	0.062087
0.6	1.648940	1.5504	0.098540
0.8	2.127229	1.98848	0.138749
1	2.640859	2.458176	0.182683

On remarque dans le tableau, que les erreurs locales augmentent quand t_i augmente, et qu'elles augmentent d'une manière linéaire.

L'erreur locale e_i dépend de deux choses :

1. l'erreur de troncature du développement limité dans la Méthode d'Euler :
 Pour construire la Méthode d'Euler, on a effectué seulement un développement limité d'ordre 1.
 On verra par la suite que la Méthode d'Euler est d'ordre 1 en temps.
 Cela signifie que l'erreur locale e_i est de l'ordre du pas de discrétisation h .
 Cela signifie aussi que si on divise le pas h par 2, les erreurs locales e_i seront grosso modo divisées par 2.
2. le cumul des erreurs :
 Comme l'algorithme de la Méthode d'Euler est un algorithme itératif, les erreurs numériques s'accumulent au fur et à mesure des itérations. Ce que l'on souhaite, c'est que ce cumul des erreurs reste limité et ne s'amplifie pas outre-mesure au cours du calcul.
 Pour garantir cela, nous avons un Théorème.

Théorème 1 - Borne sur les erreurs locales

Soit $f(t, y)$ une fonction continue et lipschitzienne sur D de rapport $L > 0$:

$$D = \{ (t, y) \text{ tels que : } a \leq t \leq b ; \quad -\infty < y < +\infty \}$$

Supposons de plus, que la dérivée seconde de la solution du problème de Cauchy $y(t)$ est bornée sur $[a, b]$:

$$\exists M > 0, \quad \forall t \in [a, b], \quad |y''(t)| \leq M.$$

Soient $(w_i)_{0 \leq i \leq n}$ les solutions approchées de $(y(t_i))_{0 \leq i \leq n}$ par la Méthode d'Euler.

Alors, on a l'estimation d'erreur suivante :

$$\forall 0 \leq i \leq n, \quad e_i = |y(t_i) - w_i| \leq \frac{hM}{2L} \left(e^{L(t_i-a)} - 1 \right)$$

Remarque :

Par ce Théorème, on obtient une borne sur les erreurs locales e_i produites par la Méthode d'Euler et on voit que cette borne dépend linéairement du pas h . On obtient :

$$\forall 0 \leq i \leq n, \quad e_i = |y(t_i) - w_i| \leq C.h, \quad \text{avec } C > 0$$

Exemple :

Reprenons l'Exemple précédent :

$$\begin{cases} y'(t) = y(t) - t^2 + 1; & 0 \leq t \leq 2 \\ y(0) = \frac{1}{2} \end{cases}$$

L'ensemble D est convexe car :

$$D = \{ (t, y) \text{ tels que : } 0 \leq t \leq 2 ; \quad -\infty < y < +\infty \}$$

On a $f(t, y) = y^2 - t^2 + 1$.

La fonction $f(t, y)$ est de classe C^∞ sur $\mathbb{R} \times \mathbb{R}$, donc sur D .

Comme $f(t, y)$ est de classe C^∞ sur D , elle admet une dérivée partielle par-rapport à y , qui est continue sur D .

$$\frac{\partial f}{\partial y}(t, y) = 1 \implies \left| \frac{\partial f}{\partial y}(t, y) \right| = 1 \leq 1, \quad \forall (t, y) \in D,$$

Donc $f(t, y)$ est lipschitzienne par-rapport à y , de constante $L=1$.

Est-ce que la dérivée seconde de la solution $y(t)$ est bornée sur $[0, 2]$?

La solution exacte $y(t)$ s'écrit :

$$y(t) = (t+1)^2 - \frac{1}{2}e^t$$

$y(t)$ est de classe C^∞ sur \mathbb{R} , donc sur $[0, 2]$. Donc sa dérivée seconde existe, et on a :

$$y'(t) = 2(t+1) - \frac{1}{2}e^t$$

$$y''(t) = 2 - \frac{1}{2}e^t$$

Or la dérivée seconde $y''(t)$ est bornée, et on a :

$$|y''(t)| = |2 - \frac{1}{2}e^t| \leq \frac{1}{2}e^2 - 2, \quad \forall t \in [0, 2]$$

Donc il existe une constante $M = \frac{1}{2}e^2 - 2 > 0$, $\forall t \in [0, 2]$, $|y''(t)| \leq M$.

Appliquons le Théorème 1 au Problème de Cauchy.

On a : $h = 0.2$; $L = 1$ et $M = \frac{1}{2}e^2 - 2$ Donc $\frac{hM}{2L} = 0.1 \left(\frac{1}{2}e^2 - 2 \right)$.

Dans ce cas, on obtient l'estimation d'erreur suivante :

$$\forall 0 \leq i \leq 10, e_i = |y(t_i) - w_i| \leq 0.1 \left(\frac{1}{2}e^2 - 2 \right) (e^{t_i} - 1)$$

On obtient le tableau suivant pour $i=0, \dots, 5$:

t_i	0.2	0.4	0.6	0.8	1
Erreurs locales $ y(t_i) - w_i $	0.02930	0.06209	0.09854	0.13875	0.18268
Borne sur les erreurs locales	0.03752	0.08334	0.13931	0.20767	0.29117

On remarque que les erreurs locales e_i commises par la Méthode d'Euler restent largement inférieures à l'estimation d'erreur donnée par le Théorème 1.

L'information importante donnée par le Théorème 1 est que la borne sur les erreurs locales dépend linéairement du pas h . Par conséquent, si on diminue le pas de discrétisation, on aboutira à une meilleure précision des résultats numériques, pour la Méthode d'Euler.

2.3.4 Erreur d'approximation globale

Puisqu'on approche la solution du problème de Cauchy $y(t)$ en s'appuyant sur plusieurs points de discrétisation t_i , $\forall i = 0, 1, \dots, n$, l'erreur d'approximation n'est pas une simple valeur, mais elle repose sur une famille de valeurs $\{y(t_i) - w_i\}_{0 \leq i \leq n}$.

Définition 2 - Erreur d'approximation pour une Méthode numérique

On définit l'erreur d'approximation globale pour une Méthode numérique par :

$$e = \max_{0 \leq i \leq n} e_i = \max_{0 \leq i \leq n} |y(t_i) - w_i|$$

Dans le cas de la Méthode d'Euler, on a une estimation de l'erreur globale produite par la Méthode d'Euler.

Théorème 2 - Estimation de l'erreur globale commise par la Méthode d'Euler

On suppose que toutes les hypothèses du Théorème 1 sont vérifiées.

Alors, l'erreur d'approximation de la Méthode d'Euler vérifie :

$$e = \max_{0 \leq i \leq n} e_i = \max_{0 \leq i \leq n} |y(t_i) - w_i| \leq \frac{hM}{2L} (e^{L(b-a)} - 1)$$

Ainsi, l'erreur d'approximation produite par la Méthode d'Euler vérifie :

$$e \leq C.h$$

où $C = \frac{M}{2L} (e^{L(b-a)} - 1) > 0$ est une constante indépendante du pas h et de t .

On déduit du Théorème 2 que l'ordre de précision de la Méthode d'Euler est égal à 1.

En effet, l'erreur d'approximation vérifie :

$$e = O(h)$$

On donnera des définitions plus précises de l'erreur d'approximation d'une Méthode numérique dans le Chapitre 3.

2.4 Méthode de Taylor

2.4.1 Introduction

La Méthode d'Euler n'est que d'ordre 1 en temps, et nous avons vu que les erreurs locales e_i sont de l'ordre du pas de discrétisation h , et qu'elles augmentent rapidement lorsque t_i augmente. Pour construire la Méthode d'Euler, nous avons effectué un développement limité d'ordre 1 de $y(t_{i+1})$ au point t_i .

La question que nous nous posons est la suivante : si on augmente l'ordre du développement limité, est-ce que cela va améliorer la précision de la Méthode numérique ?

2.4.2 Construction de la Méthode de Taylor

La Méthode de Taylor repose sur un développement limité d'ordre 2 de la solution $y(t)$ au point t_i .

On a :

$$y(t) = y(t_i) + y'(t_i)(t - t_i) + y''(t_i)\frac{(t - t_i)^2}{2} + o((t - t_i)^2)$$

On pose $t = t_{i+1}$:

$$y(t_{i+1}) = y(t_i) + y'(t_i)(t_{i+1} - t_i) + y''(t_i)\frac{(t_{i+1} - t_i)^2}{2} + o((t_{i+1} - t_i)^2)$$

Comme $t_{i+1} - t_i = h$, il vient :

$$y(t_{i+1}) = y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(t_i) + o(h^2)$$

Or dans la Méthode d'Euler, on a posé :

$$y'(t_i) = f(t_i, y(t_i))$$

Mais comment doit-on calculer $y''(t_i)$?

On a :

$$y''(t_i) = \frac{df}{dt}(t_i, y(t_i))$$

Or la fonction $f(t, y(t))$ est une fonction de 2 variables t et y et la deuxième variable y dépend elle-même de t .

Pour évaluer $\frac{df}{dt}(t, y(t))$, on a recours à la Dérivée Particulaire de $f(t, y(t))$.

Définition 3 - Dérivée Particulaire

La dérivée particulaire de $f(t, y(t))$ s'écrit :

$$\frac{df}{dt}(t, y(t)) = \frac{\partial f}{\partial t}(t, y(t)) + \frac{\partial f}{\partial y}(t, y(t)) \cdot \frac{dy(t)}{dt}$$

Or d'après le Problème de Cauchy initial, nous avons :

$$y'(t) = f(t, y(t))$$

donc :

$$\frac{dy(t)}{dt} = f(t, y(t))$$

En remplaçant dans l'expression de la Dérivée Particulaire, on obtient :

$$\frac{df}{dt}(t, y(t)) = \frac{\partial f}{\partial t}(t, y(t)) + \frac{\partial f}{\partial y}(t, y(t)) \cdot f(t, y(t))$$

Donc, la Dérivée particulaire de $f(t, y(t))$ au point t_i s'écrit :

$$\frac{df}{dt}(t_i, y(t_i)) = \frac{\partial f}{\partial t}(t_i, y(t_i)) + \frac{\partial f}{\partial y}(t_i, y(t_i)) \cdot f(t_i, y(t_i))$$

Théorème 3 - Méthode de Taylor

La Méthode de Taylor construit les approximations w_i avec $0 \leq i \leq n$, qui approchent les valeurs de la solution exacte $y(t_i)$ telle que :

$$\begin{cases} w_{i+1} = w_i + hf(t_i, w_i) + \frac{h^2}{2} \frac{df}{dt}(t_i, w_i) \\ w_0 = y(a) = \alpha \end{cases}$$

où $\frac{df}{dt}(t_i, w_i)$ désigne la Dérivée Particulaire de $f(t, y(t))$ évaluée aux points (t_i, w_i) .

On a :

$$\frac{df}{dt}(t_i, w_i) = \frac{\partial f}{\partial t}(t_i, w_i) + \frac{\partial f}{\partial y}(t_i, w_i) \cdot f(t_i, w_i)$$

2.4.3 Application

Reprenons l'Exemple précédent.

On considère le problème de Cauchy :

$$\begin{cases} y'(t) = y(t) - t^2 + 1; & 0 \leq t \leq 2 \\ y(0) = \frac{1}{2} \end{cases}$$

Ce problème admet une solution $y(t)$ que l'on a déjà calculé :

$$y(t) = (t+1)^2 - \frac{1}{2}e^t$$

Nous allons approcher cette solution $y(t)$ par la Méthode de Taylor.

On a :

$$f(t, y) = y - t^2 + 1$$

1. Calculons la Dérivée Particulaire de $f(t, y)$:

On a :

$$\frac{df}{dt}(t, y) = \frac{\partial f}{\partial t}(t, y) + \frac{\partial f}{\partial y}(t, y) \cdot f(t, y)$$

$$\text{Ici } \frac{\partial f}{\partial t} = -2t \text{ et } \frac{\partial f}{\partial y} = 1.$$

Donc la Dérivée Particulaire de $f(t, y)$ s'écrit :

$$\frac{df}{dt}(t, y) = -2t + f(t, y) = -2t + (y - t^2 + 1) = y - t^2 - 2t + 1$$

2. Comment s'écrit la Méthode de Taylor ?

D'après le Théorème 3, la Méthode de Taylor s'écrit :

$$\begin{cases} w_{i+1} = w_i + hf(t_i, w_i) + \frac{h^2}{2} \frac{df}{dt}(t_i, w_i) \\ w_0 = \frac{1}{2} \end{cases}$$

On a :

$$f(t_i, w_i) = w_i - t_i^2 + 1$$

et

$$\frac{df}{dt}(t_i, w_i) = w_i - t_i^2 - 2t_i + 1$$

Ici, on obtient :

$$\begin{cases} w_{i+1} = w_i + h(w_i - t_i^2 + 1) + \frac{h^2}{2}(w_i - t_i^2 - 2t_i + 1) \\ w_0 = \frac{1}{2} \end{cases}$$

En factorisant $(w_i - t_i^2 + 1)$, on obtient :

$$\begin{cases} w_{i+1} = w_i + \left(h + \frac{h^2}{2}\right)(w_i - t_i^2 + 1) - h^2 t_i \\ w_0 = \frac{1}{2} \end{cases}$$

Or, nous avons choisi comme pas de discrétisation $h = 0.2$.

Donc

$$t_i = a + ih = 0 + 0.2i = 0.2i \quad \forall i = 0, 1, \dots, 10$$

$$t_i^2 = 0.04 i^2 \quad \forall i = 0, 1, \dots, 10$$

$$h^2 t_i = 0.04.0.2i = 0.008i \quad \forall i = 0, 1, \dots, 10$$

En remplaçant dans le schéma numérique, on obtient :

$$\begin{cases} w_{i+1} = w_i + 0.22 (w_i - 0.04 i^2 + 1) - 0.008 i & \forall i = 0, 1, \dots, 9 \\ w_0 = \frac{1}{2} \end{cases}$$

Ainsi, cela devient très facile de calculer toutes les solutions numériques, pour $t_i \in [0, 2]$.

On résume les résultats dans le tableau suivant pour $t_i \in [0, 1]$.

t_i	Solution exacte $y(t_i)$	Solution numérique w_i	Erreurs locales $ y(t_i) - w_i $
0	$\frac{1}{2}$	$\frac{1}{2}$	0
0.2	0.829298	0.83	0.0007014
0.4	1.214087	1.2158	0.0017123
0.6	1.648940	1.652076	0.0031354
0.8	2.127229	2.1323327	0.0051032
1	2.640859	2.6486459	0.0077868

Remarques :

Nous voyons dans le tableau que les erreurs locales $|y(t_i) - w_i|$ sont beaucoup plus faibles avec la Méthode de Taylor qu'avec la Méthode d'Euler. Elles sont de l'ordre de 10^{-3} avec la Méthode de Taylor, alors qu'elles étaient de l'ordre de 10^{-1} avec la Méthode d'Euler.

De plus, les erreurs augmentent beaucoup plus lentement avec la Méthode de Taylor qu'avec la Méthode d'Euler, ce qui garantit une meilleure précision numérique.

2.4.4 Précision de la Méthode de Taylor

Théorème 4

La Méthode de Taylor construite à partir d'un développement limité d'ordre 2 est d'ordre 2 en temps.

L'erreur d'approximation globale vérifie :

$$e = \max_{0 \leq i \leq n} |y(t_i) - w_i| \leq C.h^2$$

où C est une constante indépendante du pas de discrétisation h .

Dans l'Exemple précédent pour $t_i = 2$, on obtient comme erreur locale :

$$e_{10} = |y(t_{10}) - w_{10}| = 0.0422123$$

et nous voyons que lorsque t_i augmente, l'erreur d'approximation e est bien de l'ordre de $h^2 = 0.04$ pour $h = 0.2$.

En augmentant l'ordre du développement limité de $y(t_{i+1})$ au point t_i , on peut construire des Méthodes de Taylor encore plus précises.

Théorème 5

La Méthode de Taylor construite à partir d'un développement limité d'ordre p est d'ordre p en temps.

L'erreur d'approximation globale vérifie :

$$e = \max_{0 \leq i \leq n} |y(t_i) - w_i| \leq C.h^p$$

où C est une constante indépendante du pas de discrétisation h .

Ainsi, si on construit une Méthode de Taylor d'ordre 4 à partir d'un développement limité d'ordre 4 de $y(t_{i+1})$ au point t_i , on obtient les résultats numériques suivants :

t_i	Solution exacte $y(t_i)$	Solution numérique w_i	Erreurs locales $ y(t_i) - w_i $
0	$\frac{1}{2}$	$\frac{1}{2}$	0
0.2	0.829298	0.8293	0.0000014
0.4	1.214087	1.214091	0.0000034
0.6	1.648940	1.6489468	0.0000062
0.8	2.127229	2.1272396	0.0000101
1	2.640859	2.6408744	0.0000153

Nous voyons dans le tableau que les erreurs locales $|y(t_i) - w_i|$ sont de l'ordre de 10^{-6} puis 10^{-5} , et qu'elles augmentent très lentement.

Pour $t_i = 2$, on obtient comme erreur locale :

$$e_{10} = |y(t_{10}) - w_{10}| = 0.0000834$$

Ainsi, nous voyons que lorsque t_i augmente, l'erreur d'approximation e est largement inférieure à $h^4 = 0.0016$ pour $h = 0.2$, ce qui prouve que la Méthode de Taylor d'ordre 4 est largement supérieure à la Méthode d'Euler et à la Méthode de Taylor d'ordre 2, en terme de précision numérique.

2.5 Méthodes de Runge-Kutta

2.5.1 Introduction

La Méthode de Taylor d'ordre 2 repose sur un développement limité d'ordre 2 de $y(t_{i+1})$ au point t_i . On a :

$$y(t_{i+1}) = y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(t_i) + o(h^2)$$

Dans la Méthode de Taylor, on a posé :

$$y'(t_i) = f(t_i, y(t_i)),$$

$$y''(t_i) = \frac{df}{dt}(t_i, y(t_i))$$

où $\frac{df}{dt}(t_i, y(t_i))$ désigne la Dérivée Particulaire de la fonction $f(t, y)$ au point $(t_i, y(t_i))$.

Par définition de la Dérivée Particulaire, on a vu :

$$\frac{df}{dt}(t_i, y(t_i)) = \frac{\partial f}{\partial t}(t_i, y(t_i)) + \frac{\partial f}{\partial y}(t_i, y(t_i)) \cdot f(t_i, y(t_i))$$

La Méthode de Taylor a un inconvénient : il est nécessaire de calculer les dérivées partielles de $f(t, y) : \frac{\partial f}{\partial t}$ et $\frac{\partial f}{\partial y}$, et cela peut s'avérer fastidieux, surtout pour des méthodes de Taylor d'ordre plus élevé.

On souhaite s'affranchir de cette contrainte. Pour cela, nous allons proposer des approximations pour les dérivées partielles $\frac{\partial f}{\partial t}(t_i, y(t_i))$ et $\frac{\partial f}{\partial y}(t_i, y(t_i))$.

2.5.2 Construction de la Méthode de Runge-Kutta 2

Pour approcher le terme $y''(t_i)$, nous allons utiliser un développement limité d'ordre 1 de la fonction à 2 variables :

$$f(t_i + \beta, y(t_i) + \delta)$$

où β et δ représentent des constantes que nous devons identifier, de manière à ce que le terme $f(t_i + \beta, y(t_i) + \delta)$ approche bien le terme $y''(t_i)$.

Le développement limité d'ordre 1 de $f(t_i + \beta, y(t_i) + \delta)$ au point $(t_i, y(t_i))$ s'écrit :

$$f(t_i + \beta, y(t_i) + \delta) = f(t_i, y(t_i)) + \beta \frac{\partial f}{\partial t}(t_i, y(t_i)) + \delta \frac{\partial f}{\partial y}(t_i, y(t_i)) + o\left(\sqrt{\beta^2 + \delta^2}\right) \quad (2.6)$$

On identifie les constantes β et δ comme suit :

$$\begin{cases} \beta = h, \\ \delta = hf(t_i, y(t_i)) \end{cases}$$

Le développement limité (2.6) se réécrit donc :

$$f(t_i + h, y(t_i) + hf(t_i, y(t_i))) = f(t_i, y(t_i)) + h \frac{\partial f}{\partial t}(t_i, y(t_i)) + hf(t_i, y(t_i)) \frac{\partial f}{\partial y}(t_i, y(t_i)) + o\left(\sqrt{h^2 + h^2(f(t_i, y(t_i)))^2}\right)$$

ce qui s'écrit encore :

$$f(t_i + h, y(t_i) + hf(t_i, y(t_i))) = f(t_i, y(t_i)) + h \frac{\partial f}{\partial t}(t_i, y(t_i)) + hf(t_i, y(t_i)) \frac{\partial f}{\partial y}(t_i, y(t_i)) + o(h)$$

Dans cette équation, nous voyons apparaître la Dérivée Particulière de $f(t, y(t))$ au point $(t_i, y(t_i))$. En effet, nous avons :

$$\frac{\partial f}{\partial t}(t_i, y(t_i)) + f(t_i, y(t_i)) \frac{\partial f}{\partial y}(t_i, y(t_i)) = \frac{1}{h} [f(t_i + h, y(t_i) + hf(t_i, y(t_i))) - f(t_i, y(t_i)) - o(h)] \quad (2.7)$$

Or comment s'écrit la Méthode de Taylor d'ordre 2 ?

La Méthode de Taylor construit les approximations w_i telle que :

$$\begin{cases} w_{i+1} = w_i + hf(t_i, w_i) + \frac{h^2}{2} \frac{df}{dt}(t_i, w_i) \\ w_0 = y(a) = \alpha \end{cases}$$

Elle s'écrit encore :

$$\begin{cases} w_{i+1} = w_i + hf(t_i, w_i) + \frac{h^2}{2} \left[\frac{\partial f}{\partial t}(t_i, w_i) + \frac{\partial f}{\partial y}(t_i, w_i) \cdot f(t_i, w_i) \right] \\ w_0 = y(a) = \alpha \end{cases} \quad (2.8)$$

En remplaçant l'équation (2.7) dans le schéma numérique (2.8), on obtient la Méthode de Runge-Kutta d'ordre 2 :

$$\begin{cases} w_{i+1} = w_i + hf(t_i, w_i) + \frac{h}{2} [f(t_i + h, w_i + hf(t_i, w_i)) - f(t_i, w_i)] \\ w_0 = y(a) = \alpha \end{cases}$$

En simplifiant, on obtient :

$$\begin{cases} w_{i+1} = w_i + \frac{h}{2} [f(t_i, w_i) + f(t_i + h, w_i + hf(t_i, w_i))] \\ w_0 = y(a) = \alpha \end{cases}$$

Théorème 6 - Méthode de Runge-Kutta 2

La Méthode de Runge-Kutta 2 construit les approximations w_i telles que :

$$\begin{cases} w_{i+1} = w_i + \frac{h}{2} [f(t_i, w_i) + f(t_{i+1}, w_i + hf(t_i, w_i))] & \forall i = 0, 1, \dots, n-1 \\ w_0 = y(a) = \alpha \end{cases}$$

On appelle aussi la Méthode de Runge-Kutta 2, Méthode d'Heun.

2.5.3 Précision de la Méthode de Runge-Kutta 2

Théorème 7

La Méthode de Runge-Kutta 2 est d'ordre 2 en temps.

L'erreur d'approximation globale vérifie :

$$e = \max_{0 \leq i \leq n} |y(t_i) - w_i| \leq C.h^2$$

où C est une constante indépendante du pas de discrétisation h .

Reprenons l'Exemple précédent avec $h = 0.2$.

On obtient les résultats numériques suivants pour la Méthode de Runge-Kutta 2 :

t_i	Solution exacte $y(t_i)$	Solution numérique w_i	Erreurs locales $ y(t_i) - w_i $
0	$\frac{1}{2}$	$\frac{1}{2}$	0
0.2	0.829298	0.827333	0.001965
0.4	1.214087	1.209880	0.004207
0.6	1.648940	1.642187	0.006754
0.8	2.127229	2.117601	0.009628
1	2.640859	2.628007	0.012852

Remarques :

Si on compare les erreurs obtenues dans le tableau avec celles obtenues pour la Méthode de Taylor d'ordre 2, on voit que la Méthode de Taylor d'ordre 2 donne une meilleure approximation de la première valeur w_1 , et aussi une meilleure précision que la Méthode de Runge-Kutta 2.

En effet, les erreurs numériques augmentent plus lentement avec la Méthode de Taylor, bien que ces 2 méthodes soient d'ordre 2.

Cependant, la Méthode de Runge-Kutta 2 est la plus utilisée en pratique, car elle est beaucoup plus facile à programmer que la Méthode de Taylor.

2.5.4 Algorithme de la Méthode de Runge-Kutta 2

1. Initialisation : $h = \frac{b-a}{n}$; $t_0 = a$, $w_0 = \alpha$ et $i = 0$
2. Tant que $i < n$, répéter les étapes 3 à 5 :
3. $t_{i+1} = t_i + h$
4. $w_{i+1} = w_i + \frac{h}{2} [f(t_i, w_i) + f(t_{i+1}, w_i + hf(t_i, w_i))]$
5. $i = i + 1$

2.5.5 Méthodes de Runge-Kutta 3 et 4

En procédant de la même manière, on peut obtenir les Méthodes de Runge-Kutta 3 et 4 (RK3 et RK4) à partir des Méthodes de Taylor d'ordre 3 et d'ordre 4.

On cherche dans ces Méthodes des approximations pour les dérivées partielles d'ordre 2 et 3 de $f(t, y)$: $\frac{\partial^2 f}{\partial t^2}$, $\frac{\partial^2 f}{\partial y^2}$ et $\frac{\partial^3 f}{\partial t^3}$, $\frac{\partial^3 f}{\partial y^3}$.

Les dérivations des Méthodes de Runge-Kutta 3 et Runge-Kutta 4 sont longues et fastidieuses.

Heureusement, leurs formules sont faciles à programmer, et nous nous contenterons d'utiliser leurs résultats.

Théorème 8 - Méthode de Runge-Kutta 3

$$\left\{ \begin{array}{l} w_0 = y(a) = \alpha \\ K_1 = h f(t_i, w_i) \\ K_2 = f(t_i + \frac{h}{2}, w_i + \frac{K_1}{2}) \\ K_3 = f(t_i + h, w_i + 2K_2 - K_1) \\ w_{i+1} = w_i + \frac{h}{6} (K_1 + 4K_2 + K_3) \quad \forall i = 0, 1, \dots, n-1 \end{array} \right.$$

On voit que cette Méthode nécessite 3 évaluations de la fonction $f(t, y)$ par pas de temps, ce qui a un coût non négligeable par-rapport à la Méthode d'Euler ou la Méthode de Runge-Kutta 2. Mais comme on le verra, ce coût est largement compensé par le gain de précision numérique obtenu.

Méthode de Runge-Kutta 4

$$\left\{ \begin{array}{l} w_0 = y(a) = \alpha \\ K_1 = h f(t_i, w_i) \\ K_2 = f(t_i + \frac{h}{2}, w_i + \frac{K_1}{2}) \\ K_3 = f(t_i + h, w_i + \frac{K_2}{2}) \\ K_4 = f(t_i + h, w_i + K_3) \\ w_{i+1} = w_i + \frac{h}{6} (K_1 + 2K_2 + 2K_3 + K_4) \quad \forall i = 0, 1, \dots, n-1 \end{array} \right.$$

La Méthode de Runge-Kutta 4 a une forme assez symétrique.
Elle nécessite 4 évaluations de la fonction $f(t, y)$ par pas de temps.
Elle est de loin, la plus utilisée dans les codes de calcul numériques.

2.5.6 Précision des Méthodes de Runge-Kutta

Théorème 9

Les Méthodes Runge-Kutta p (RK p) sont d'ordre p en temps.

L'erreur d'approximation globale vérifie :

$$e = \max_{0 \leq i \leq n} |y(t_i) - w_i| \leq C.h^p$$

où C est une constante indépendante du pas de discrétisation h .

En particulier, si on divise le pas h par 2, l'erreur d'approximation globale sera grosso-modo divisée par 2^p .

Ainsi la Méthode de Runge-Kutta 3 est d'ordre 3 en temps et la Méthode de Runge-Kutta 4 est d'ordre 4 en temps.

Si on reprend l'Exemple précédent avec la Méthode de Runge-Kutta 4, on obtient les résultats numériques suivants :

t_i	Solution exacte $y(t_i)$	Solution numérique w_i	Erreurs locales $ y(t_i) - w_i $
0	$\frac{1}{2}$	$\frac{1}{2}$	0
0.2	0.829298	0.829293	0.0000053
0.4	1.214087	1.214076	0.000011
0.6	1.648940	1.648922	0.0000186
0.8	2.127229	2.127203	0.0000269
1	2.640859	2.640823	0.0000364

Nous voyons dans le tableau que les erreurs numériques obtenues avec la Méthode de Runge-Kutta 4 sont du même ordre que celles obtenues avec la Méthode de Taylor d'ordre 4.

Pour $t_i = 2$, on obtient comme erreur locale :

$$e_{10} = |y(t_{10}) - w_{10}| = 0.0001089$$

Ainsi, l'erreur d'approximation e reste bien inférieure à $h^4 = 0.0016$ quand t_i augmente, ce qui montre que la Méthode de Runge-Kutta 4 est une bonne alternative aux Méthodes de Taylor, en terme de précision numérique, même si elle nécessite 4 évaluations de la fonction $f(t, y)$ par pas de temps.

Chapitre 3

Propriétés des Méthodes numériques à un pas

3.1 Définitions générales

Les Méthodes numériques que nous avons présentées sont des méthodes à un pas : cela signifie qu'à l'instant t_{i+1} , l'inconnue w_{i+1} ne dépend que de la valeur à l'instant précédent w_i . Il n'y a donc qu'un seul pas qui entre en jeu.

3.1.1 Définition 1 - Forme des Méthodes numériques à un pas

La forme générale des Méthodes à un pas s'écrit de la façon suivante.

On construit une suite $(w_i)_{0 \leq i \leq n}$ qui approche la solution exacte $y(t_i)_{0 \leq i \leq n}$ telle que :

$$\begin{cases} w_0 = \alpha \\ w_{i+1} = w_i + h \Phi(t_i, w_i, h) \quad \forall i \in \{0, 1, \dots, n-1\} \end{cases}$$

3.1.2 Remarques

1. Méthode d'Euler :

Elle s'écrit :

$$\begin{cases} w_0 = \alpha \\ w_{i+1} = w_i + h f(t_i, w_i) \quad \forall i \in \{0, 1, \dots, n-1\} \end{cases}$$

Ici :

$$\Phi(t_i, w_i, h) = f(t_i, w_i)$$

et la fonction Φ ne dépend pas du pas h .

2. Méthode de Runge-Kutta 2 :

Elle s'écrit :

$$\begin{cases} w_0 = \alpha \\ w_{i+1} = w_i + \frac{h}{2} [f(t_i, w_i) + f(t_{i+1}, w_i + hf(t_i, w_i))] \quad \forall i = 0, 1, \dots, n-1 \end{cases}$$

Ici :

$$\Phi(t_i, w_i, h) = \frac{1}{2} [f(t_i, w_i) + f(t_i + h, w_i + hf(t_i, w_i))]$$

et la fonction Φ dépend de t_i , w_i et du pas h .

Dans le chapitre 2, nous avons construit différentes méthodes numériques : Méthode d'Euler, Méthodes de Taylor d'ordre 2, Méthodes de Runge-Kutta d'ordre 2, 3 ou 4, qui approchent le problème de Cauchy.

Il faut ensuite s'assurer que la solution numérique w_i converge vers la solution exacte $y(t_i)$ lorsque le pas de discrétisation h tend vers 0.

3.2 Convergence d'une Méthode numérique à un pas

3.2.1 Définition 2 - Convergence

On dira qu'une Méthode numérique à un pas est convergente si et seulement si :

$$\max_{0 \leq i \leq n} |w_i - y(t_i)| \longrightarrow 0 \text{ quand } h \longrightarrow 0$$

Afin d'étudier la convergence d'une méthode numérique, il faut analyser les sources des erreurs qui interviennent dans l'approximation du problème de Cauchy.

1. L'erreur de troncature : il s'agit de l'erreur de troncature du développement limité que l'on fait pour construire la méthode numérique.

Par exemple, pour construire la Méthode d'Euler, on effectue un développement limité d'ordre 1 de $y(t_{i+1})$ au point t_i . Le reste du développement limité, c'est-à-dire les termes d'ordre 2, 3, 4, ..., n, sont négligés.

Cette approximation introduit naturellement des erreurs d'approximation, que l'on appelle erreur de troncature.

2. Le cumul des erreurs : il s'agit des erreurs numériques accumulées au cours du calcul, quand on avance dans les itérations de la méthode numérique.

Comme on l'a vu par exemple dans la Méthode d'Euler, les erreurs augmentent linéairement quand t_i augmente, et deviennent de plus en plus importantes au fur et à mesure des itérations.

Deux notions sont introduites afin de caractériser ces deux sources d'erreurs : la notion de consistance pour l'erreur de troncature, et la notion de stabilité théorique pour le cumul des erreurs.

3.2.2 Théorème 1 - Convergence d'une Méthode numérique

Pour garantir la convergence d'une méthode numérique à un pas, il faut vérifier ces deux conditions :

- Consistance de la méthode numérique
 - Stabilité théorique de la méthode numérique
- puisque l'on :

$$\text{Consistance} + \text{Stabilité théorique} \implies \text{Convergence de la méthode numérique}$$

3.3 Consistance d'une Méthode numérique

3.3.1 Introduction

Soit une Méthode numérique définie par :

$$\begin{cases} w_{i+1} = w_i + h \Phi(t_i, w_i, h) & \forall i \in \{0, 1, \dots, n-1\} \\ w_0 = y(a) = \alpha \end{cases} \quad (3.1)$$

La suite $(w_i)_{0 \leq i \leq n}$ est donc construite de manière à vérifier l'égalité :

$$\frac{w_{i+1} - w_i}{h} - \Phi(t_i, w_i, h) = 0 \quad (3.2)$$

La solution exacte $y(t_i)$ n'a pas de raison de vérifier l'équation (3.2) à l'instant t_i , mais on espère que l'erreur commise en remplaçant w_i par $y(t_i)$ dans l'équation (3.2) reste petite.

3.3.2 Définition 3 - Erreur de troncature locale

On appelle erreur de troncature locale, la quantité $\epsilon_i(h)$ définie par :

$$\epsilon_i(h) = \frac{y(t_{i+1}) - y(t_i)}{h} - \Phi(t_i, y(t_i), h)$$

où $y(t_i)$ représente la solution du problème de Cauchy $y'(t) = f(t, y(t))$ à l'instant t_i .

3.3.3 Définition 4 - Consistance d'une Méthode numérique

La Méthode numérique (3.1) est consistante avec l'équation différentielle $y'(t) = f(t, y(t))$ si et seulement si :

$$\max_{0 \leq i \leq n} |\epsilon_i(h)| \longrightarrow 0 \quad \text{quand} \quad h \longrightarrow 0$$

3.3.4 Remarques :

1. Pour vérifier la consistance d'une Méthode numérique, on injecte la solution exacte du problème de Cauchy $y(t)$, évaluée aux points t_i et t_{i+1} , dans le schéma numérique (3.1). Le résultat obtenu, qui dépend de t_i , et de h , doit tendre vers 0 quand h tend vers 0.
2. Cela signifie que la méthode numérique doit être une approximation du problème de Cauchy, bien construite et vraisemblable.

3.3.5 Exemple

On considère le problème de Cauchy suivant :

$$\begin{cases} y' = y - t + 1; & 0 \leq t \leq 1 \\ y(0) = 1 \end{cases} \quad (3.3)$$

La solution exacte vérifie :

$$y(t) = e^t + t$$

Ecrivons la Méthode d'Euler pour ce problème :

$$\begin{cases} w_{i+1} = w_i + h f(t_i, w_i) & \forall i \in \{0, 1, \dots, n-1\} \\ w_0 = 1 \end{cases} \quad (3.4)$$

Ici :

$$f(t, y) = y - t + 1$$

Le schéma numérique (3.4) se réécrit :

$$\frac{w_{i+1} - w_i}{h} - f(t_i, w_i) = 0$$

Remplaçons les valeurs numérique (w_i) par les solutions exactes calculées au point t_i et t_{i+1} : $y(t_i)$ et $y(t_{i+1})$.

On obtient alors :

$$\frac{y(t_{i+1}) - y(t_i)}{h} - f(t_i, y(t_i)) = \epsilon_i(h) \quad (3.5)$$

Dans notre cas :

$$y(t_i) = e^{t_i} + t_i$$

$$y(t_{i+1}) = e^{t_{i+1}} + t_{i+1}$$

et

$$f(t_i, y(t_i)) = y(t_i) - t_i + 1$$

On remplace dans l'équation (3.5) et on obtient :

$$\frac{e^{t_{i+1}} + t_{i+1} - e^{t_i} - t_i}{h} - (e^{t_i} + t_i - t_i + 1) = \epsilon_i(h)$$

En simplifiant, on obtient :

$$\frac{e^{t_{i+1}} - e^{t_i} + h}{h} - (e^{t_i} + 1) = \epsilon_i(h) \quad (3.6)$$

Or :

$$e^{t_{i+1}} = e^{t_i+h} = e^{t_i} \cdot e^h$$

Or par définition du pas de discrétisation h , on a : $h \ll 1$ et h est petit, proche de 0.

Ecrivons le développement limité d'ordre 2 de e^h au voisinage de 0 :

$$e^h = 1 + h + \frac{h^2}{2} + o(h^2)$$

et :

$$e^{t_i+h} = e^{t_i} \cdot \left(1 + h + \frac{h^2}{2} + o(h^2)\right) = e^{t_i} + e^{t_i}h + e^{t_i}\frac{h^2}{2} + o(h^2)$$

En remplaçant dans l'équation (3.6), on obtient :

$$\frac{e^{t_i} + e^{t_i}h + e^{t_i}\frac{h^2}{2} - e^{t_i} + h}{h} - (e^{t_i} + 1) + o(h) = \epsilon_i(h)$$

En simplifiant, on obtient :

$$e^{t_i} + e^{t_i}\frac{h}{2} + 1 - (e^{t_i} + 1) + o(h) = \epsilon_i(h)$$

ou encore :

$$e^{t_i}\frac{h}{2} + o(h) = \epsilon_i(h)$$

L'erreur de troncature pour la Méthode d'Euler appliquée au problème de Cauchy (3.3) vaut donc :

$$\epsilon_i(h) = e^{t_i}\frac{h}{2} + o(h)$$

et on a :

$$\forall i \in \{0, 1, \dots, n\}, |\epsilon_i(h)| \longrightarrow 0 \quad \text{quand } h \longrightarrow 0$$

et donc, on a bien :

$$\max_{0 \leq i \leq n} |\epsilon_i(h)| \longrightarrow 0 \quad \text{quand } h \longrightarrow 0$$

Conclusion : nous avons démontré que la Méthode d'Euler est bien consistante avec le problème de Cauchy (3.3).

3.3.6 Théorème 2 - Outil pour vérifier la consistance :

On dispose d'un outil très simple pour montrer qu'une Méthode numérique est consistante avec le problème de Cauchy.

Cette méthode nous évite tous les calculs effectués dans l'Exemple précédent.

Soit une Méthode numérique définie par :

$$\begin{cases} w_0 = \alpha \\ w_{i+1} = w_i + h \Phi(t_i, w_i, h) \quad \forall i \in \{0, 1, \dots, n-1\} \end{cases}$$

Cette méthode numérique est consistante si et seulement si :

$$\forall t \in [a, b], \quad \forall y \in \mathbb{R}, \quad \Phi(t, y, 0) = f(t, y)$$

3.3.7 Exemples

1. La Méthode d'Euler est bien consistante avec le problème de Cauchy car :

$$\Phi(t_i, w_i, h) = f(t_i, w_i)$$

et Φ ne dépend pas du pas h .

$$\text{Donc } \Phi(t, y, 0) = f(t, y)$$

2. La Méthode de Runge-Kutta 2 est bien consistante avec le problème de Cauchy, car :

$$\Phi(t_i, w_i, h) = \frac{1}{2} [f(t_i, w_i) + f(t_i + h, w_i + hf(t_i, w_i))]$$

$$\Phi(t, y, h) \Big|_{h=0} = \frac{1}{2} [f(t, y) + f(t + h, y + hf(t, y))] \Big|_{h=0} = \frac{1}{2} [f(t, y) + f(t + 0, y + 0 \cdot f(t, y))] = f(t, y)$$

3.4 Stabilité théorique d'une Méthode numérique

3.4.1 Définition 5 - Stabilité théorique

Soit une Méthode numérique à un pas, approchant le problème de Cauchy $y'(t) = f(t, y(t))$:

$$\begin{cases} w_0 = \alpha \\ w_{i+1} = w_i + h \Phi(t_i, w_i, h) \quad \forall i \in \{0, 1, \dots, n-1\} \end{cases}$$

On considère son problème perturbé associé :

$$\begin{cases} z_0 = \beta \\ z_{i+1} = z_i + h (\Phi(t_i, z_i, h) + \delta_i) \quad \forall i \in \{0, 1, \dots, n-1\} \end{cases}$$

Pour créer le problème perturbé, on rajoute au schéma numérique :

- une faible perturbation δ_i sur la fonction $\Phi(t_i, w_i, h)$ à l'instant t_i ,
- une faible perturbation $\beta - \alpha$ sur la condition initiale w_0 .

On dira que la Méthode numérique (3.1) est théoriquement stable si et seulement si il existe deux constantes M_1 et M_2 indépendantes du pas h , telles que :

$$\max_{0 \leq i \leq n} |z_i - w_i| \leq M_1 |\beta - \alpha| + M_2 \max_{0 \leq i \leq n-1} |\delta_i|$$

La Méthode numérique sera théoriquement stable si et seulement si une petite perturbation sur la fonction Φ ou sur la condition initiale w_0 n'entraîne qu'une légère variation de la solution numérique w_i , et ceci, quel que soit $0 \leq i \leq n$, et quel que soit le pas de discrétisation h .

Cela signifie que les erreurs numériques réalisées au cours d'un calcul ne seront pas amplifiées au fur et à mesure des itérations, et que le cumul des erreurs restera limité.

3.4.2 Théorème 3 - Outil pour vérifier la stabilité théorique

La Définition 5 n'est pas facile à utiliser en pratique.

Pour montrer qu'une Méthode numérique est théoriquement stable, on dispose d'un outil très simple.

Soit une Méthode numérique définie par :

$$\begin{cases} w_0 = \alpha \\ w_{i+1} = w_i + h \Phi(t_i, w_i, h) \quad \forall i \in \{0, 1, \dots, n-1\} \end{cases}$$

Si la fonction Φ est lipschitzienne par-rapport à y , c'est-à-dire s'il existe une constante $M > 0$ telle que :

$$\forall t \in [a, b], \forall (y_1, y_2) \in \mathbb{R}^2, \forall h > 0, \text{ avec } h \text{ très petit devant } 1, \text{ on ait : } \left| \Phi(t, y_1, h) - \Phi(t, y_2, h) \right| \leq M |y_1 - y_2|$$

alors la Méthode numérique est théoriquement stable.

3.4.3 Exemples

1. Méthode d'Euler : dans la Méthode d'Euler, on a :

$$\Phi(t, y, h) = f(t, y)$$

et donc : si la fonction $f(t, y)$ est lipschitzienne par-rapport à y , alors la Méthode d'Euler est stable.

2. Méthode de Runge-Kutta 2 : dans la Méthode de Runge-Kutta 2, on a :

$$\Phi(t, y, h) = \frac{1}{2} [f(t, y) + f(t + h, y + hf(t, y))]$$

Si la fonction $f(t, y)$ est lipschitzienne par-rapport à y , alors la Méthode de Runge-Kutta 2 est stable.

Démonstration : Supposons que $f(t, y)$ est lipschitzienne par-rapport à y de constante $L > 0$ sur D :

$$D = \{(t, y) \text{ tels que } a \leq t \leq b, y \in \mathbb{R}\}$$

$$\exists L > 0, \forall (t, y_1) \in D, \forall (t, y_2) \in D, \left| f(t, y_1) - f(t, y_2) \right| \leq L |y_1 - y_2|$$

Montrons qu'il existe une constante $M > 0$ telle que :

$\forall (y_1, y_2) \in \mathbb{R}^2, \forall h > 0$, avec h très petit devant 1, on ait : $\left| \Phi(t, y_1, h) - \Phi(t, y_2, h) \right| \leq M |y_1 - y_2|$

On a :

$$\left| \Phi(t, y_1, h) - \Phi(t, y_2, h) \right| = \left| \frac{1}{2} (f(t, y_1) + f(t+h, y_1 + hf(t, y_1))) - \frac{1}{2} (f(t, y_2) + f(t+h, y_2 + hf(t, y_2))) \right|$$

D'après l'inégalité triangulaire, on obtient :

$$\left| \Phi(t, y_1, h) - \Phi(t, y_2, h) \right| \leq \frac{1}{2} \left(\left| f(t, y_1) - f(t, y_2) \right| + \left| f(t+h, y_1 + hf(t, y_1)) - f(t+h, y_2 + hf(t, y_2)) \right| \right)$$

Comme $f(t, y)$ est lipschitzienne par-rapport à y de constante $L > 0$ sur D , on a :

$$\left| f(t, y_1) - f(t, y_2) \right| \leq L |y_1 - y_2|$$

et

$$\left| f(t+h, y_1 + hf(t, y_1)) - f(t+h, y_2 + hf(t, y_2)) \right| \leq L \left| y_1 + hf(t, y_1) - (y_2 + hf(t, y_2)) \right|$$

Donc, on peut écrire :

$$\left| \Phi(t, y_1, h) - \Phi(t, y_2, h) \right| \leq \frac{1}{2} \left(L|y_1 - y_2| + L \left| y_1 + hf(t, y_1) - (y_2 + hf(t, y_2)) \right| \right)$$

Appliquons une deuxième fois l'inégalité triangulaire au dernier terme de l'inégalité.

On obtient :

$$\left| \Phi(t, y_1, h) - \Phi(t, y_2, h) \right| \leq \frac{1}{2} (L|y_1 - y_2| + L|y_1 - y_2| + hL |f(t, y_1) - f(t, y_2)|)$$

Comme $f(t, y)$ est lipschitzienne par-rapport à y de constante $L > 0$ sur D , on obtient finalement :

$$\left| \Phi(t, y_1, h) - \Phi(t, y_2, h) \right| \leq \frac{1}{2} (2L|y_1 - y_2| + hL^2 |y_1 - y_2|)$$

et on a :

$$\left| \Phi(t, y_1, h) - \Phi(t, y_2, h) \right| \leq \frac{1}{2} (2L + hL^2) |y_1 - y_2| \iff \left| \Phi(t, y_1, h) - \Phi(t, y_2, h) \right| \leq \left(L + h \frac{L^2}{2} \right) |y_1 - y_2|$$

Or le pas de discrétisation h est par hypothèse, très petit devant 1 : $h \ll 1$ et on a la majoration suivante :

$$L + h \frac{L^2}{2} \leq L + \frac{L^2}{2}$$

Donc il existe bien une constante $M = L + \frac{L^2}{2} > 0$ indépendante du pas h , telle que la fonction $\Phi(t, y, h)$ soit lipschitzienne par-rapport à y de constante $M > 0$ sur D .

Pour conclure, nous avons démontré par ce calcul, que la Méthode de Runge-Kutta 2 est stable.

3.5 Ordre de précision

3.5.1 Définition 6 - Précision d'une Méthode numérique

Soit une Méthode numérique définie par :

$$\begin{cases} w_{i+1} = w_i + h \Phi(t_i, w_i, h) \quad \forall i \in \{0, 1, \dots, n-1\} \\ w_0 = y(a) = \alpha \end{cases} \quad (3.7)$$

On rappelle la définition de l'erreur de troncature locale d'une Méthode numérique :

$$\epsilon_i(h) = \frac{y(t_{i+1}) - y(t_i)}{h} - \Phi(t_i, y(t_i), h)$$

où $y(t_i)$ est la solution du problème de Cauchy $y'(t) = f(t, y(t))$ à l'instant t_i .

On dira que la Méthode numérique (3.7) est d'ordre p en temps si et seulement si :

$$\max_{0 \leq i \leq n} |\epsilon_i(h)| \leq C h^p$$

où C est une constante indépendante du pas h .

3.5.2 Exemples

1. La Méthode d'Euler est d'ordre un en temps.

Démonstration :

Le développement limite à l'ordre 2 de $y(t_{i+1})$ au voisinage du point t_i s'écrit :

$$y(t_{i+1}) = y(t_i) + h y'(t_i) + \frac{h^2}{2} y''(t_i) + o(h^2)$$

Pour la Méthode d'Euler, l'erreur de troncature locale $\epsilon_i(h)$ s'écrit :

$$\epsilon_i(h) = \frac{y(t_{i+1}) - y(t_i)}{h} - f(t_i, y(t_i)) \quad (3.8)$$

En remplaçant $y(t_{i+1})$ par son développement limité dans l'équation (3.8), on obtient :

$$\epsilon_i(h) = \frac{1}{h} \left(y(t_i) + h y'(t_i) + \frac{h^2}{2} y''(t_i) + o(h^2) - y(t_i) \right) - f(t_i, y(t_i))$$

En simplifiant, on obtient :

$$\epsilon_i(h) = y'(t_i) + \frac{h}{2} y''(t_i) + o(h) - f(t_i, y(t_i))$$

Or $y(t_i)$ est solution du problème de Cauchy :

$$y'(t) = f(t, y(t))$$

et on a :

$$y'(t_i) = f(t_i, y(t_i))$$

Finalement, on obtient l'erreur de troncature locale suivante :

$$\epsilon_i(h) = \frac{h}{2} y''(t_i) + o(h)$$

On voit bien que :

$$\forall i \in \{0, 1, \dots, n\} \quad |\epsilon_i(h)| \leq C.h$$

où $C = \frac{|y''(t_i)|}{2}$ est une constante indépendante du pas h .

Nous avons bien démontré que la Méthode d'Euler appliquée à un problème de Cauchy, est d'ordre 1 en temps.

2. En procédant de la même façon par des développements limités de $y(t_{i+1})$ au point t_i , on peut montrer que la Méthode de Runge-Kutta 2 est d'ordre 2, que la Méthode de Taylor est d'ordre 2, que la Méthode Runge-Kutta 3 est d'ordre 3, et que la Méthode Runge-Kutta 4 est d'ordre 4.

On termine ce chapitre par un Théorème sur la précision globale d'une Méthode numérique.

3.5.3 Erreur d'approximation d'une Méthode numérique

Soit une Méthode numérique définie par :

$$\begin{cases} w_{i+1} = w_i + h \Phi(t_i, w_i, h) & \forall i \in \{0, 1, \dots, n-1\} \\ w_0 = y(a) = \alpha \end{cases} \quad (3.9)$$

On suppose que la fonction $\Phi(t, y, h)$ est lipschitzienne par-rapport à y .

Si cette Méthode numérique est d'ordre p en temps, alors l'erreur d'approximation e de la Méthode numérique vérifie :

$$e = \max_{0 \leq i \leq n} |y(t_i) - w_i| \leq \text{Const.} h^p$$

Cela signifie que si l'erreur de troncature locale $\epsilon_i(h)$ est d'ordre p , alors l'erreur d'approximation globale de la méthode numérique sera aussi d'ordre p .

En particulier, si on divise le pas h par 2, l'erreur d'approximation globale de la méthode numérique sera divisée grosso-modo par 2^p .

Chapitre 4

Interpolation polynômiale

4.1 Introduction

En Analyse numérique, ou en physique, une fonction f inconnue explicitement est souvent connue seulement en certains points $x_0, x_1, x_2, \dots, x_n$, par exemple, au moyen de relevés expérimentaux.

Un problème d'interpolation est défini comme suit :
étant donné un ensemble de points, et de valeurs en ces points, obtenus par exemple à la suite d'une expérience, on cherche une fonction f passant par tous ces points.

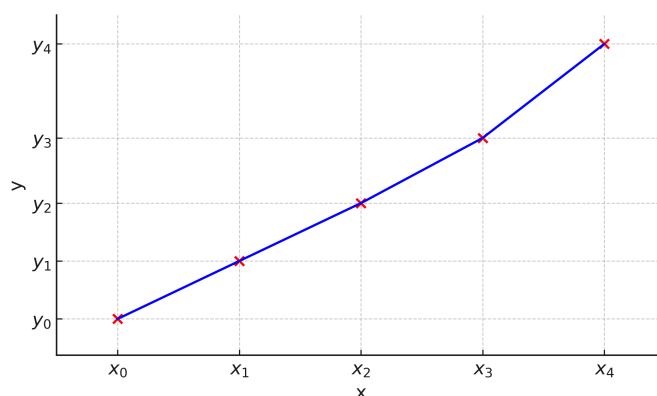


FIGURE 4.1 – Exemple de problème d'interpolation

On cherche donc à reconstruire la fonction f au moyen d'une autre fonction p plus simple, et plus facile à évaluer au moyen des données discrètes :

$$y_i = f(x_i) \quad \forall i \in \{0, 1, 2, \dots, n\}.$$

On espère que le modèle p ne sera pas trop éloigné de la fonction f aux autres points.

Comment approcher le plus simplement une fonction ?
Par des polynômes.

4.2 Interpolation polynômiale

4.2.1 Approximation polynômiale

Les polynômes sont les fonctions les plus simples que l'on puisse construire à l'aide des opérations : $+$, $-$, \times .

De plus, ils sont faciles à dériver, et faciles à intégrer : par exemple si on cherche à approcher $\int_a^b f(x) dx$.

Le Théorème de Stone-Weierstrass nous assure que l'on peut approcher toute fonction continue sur un intervalle fermé borné $[a, b]$, $a \in \mathbb{R}$, $b \in \mathbb{R}$, d'autant plus près que l'on veut par un polynôme.

Théorème 1 - Théorème de Stone-Weierstrass

Soit f une fonction définie et continue sur un intervalle fermé borné $[a, b]$, $a \in \mathbb{R}$, $b \in \mathbb{R}$. Alors :

$\forall \epsilon > 0$, il existe un polynôme P tel que : $\forall x \in [a, b]$, $|f(x) - P(x)| < \epsilon$

Plus ϵ est petit, plus le degré du polynôme P sera grand.

4.2.2 Position du problème d'interpolation

Une première condition à imposer au polynôme P est qu'il coïncide avec la fonction f en des points précis, appelés points d'interpolation.

Plus généralement le problème d'interpolation peut se formuler mathématiquement comme suit :

Définition du problème d'interpolation \mathcal{P}

Etant donné $n + 1$ points $\{x_i, f(x_i)\}_{i=0,1,\dots,n}$ où les points $\{x_i\}_{i=0,1,\dots,n}$ sont deux à deux distincts, trouver un polynôme de degré inférieur ou égal à n (noté $P_n(x)$) appartenant à $\mathbb{R}^n[X]$ tel que :

$$\forall i = 0, 1, \dots, n, P_n(x_i) = f(x_i)$$

Les points $\{x_i\}_{i=0,1,\dots,n}$ sont appelés points d'interpolation, ou noeuds d'interpolation.

Les points $\{f(x_i)\}_{i=0,1,\dots,n}$ représentent les valeurs d'interpolation.

Maintenant, il est naturel de se poser les questions suivantes :

1. Le polynôme P_n est-il unique, ou non ?
2. Quelle est l'erreur commise en remplaçant f par P_n ?
3. En particulier, que peut-on dire de $|f - P_n|$ quand le nombre de points d'interpolation $\{x_i\}$ augmente ?

4.2.3 Existence et unicité du polynôme d'interpolation

Théorème 2

Le problème d'interpolation \mathcal{P} admet une solution unique si et seulement si les points d'interpolation $\{x_i\}_{i=0,1,\dots,n}$ sont deux à deux distincts.

Autrement dit, il existe un unique polynôme de degré inférieur ou égal à n , vérifiant :

$$\forall i = 0, 1, \dots, n, P_n(x_i) = f(x_i)$$

Démonstration :

Ecrivons le polynôme $P_n(x)$ dans la base canonique :

$$P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

où les coefficients $(a_0, a_1, a_2, \dots, a_n) \in \mathbb{R}^n$.

D'après le Théorème 2, $P_n(x)$ doit vérifier $n + 1$ relations :

$$\begin{cases} P_n(x_0) = a_0 + a_1x_0 + a_2x_0^2 + \dots + a_nx_0^n = f(x_0) \\ P_n(x_1) = a_0 + a_1x_1 + a_2x_1^2 + \dots + a_nx_1^n = f(x_1) \\ \cdot \\ \cdot \\ \cdot \\ P_n(x_n) = a_0 + a_1x_n + a_2x_n^2 + \dots + a_nx_n^n = f(x_n) \end{cases}$$

Les inconnues sont les coefficients $(a_0, a_1, a_2, \dots, a_n)$, et les équations obtenues sont linéaires par-rapport à $\{a_0, a_1, a_2, \dots, a_n\}$.

Nous pouvons réécrire ce système sous forme matricielle :

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \cdot \\ \cdot \\ \cdot \\ f(x_n) \end{bmatrix}$$

A partir des données $\{(x_i, f(x_i))\}_{i=0,1,\dots,n}$, on peut résoudre ce système matriciel qui s'écrit :

$$VA = F,$$

où V désigne la matrice de Vandermonde.

A quelle condition la matrice V du système matriciel est-elle inversible ?

On peut calculer le déterminant de V , et montrer que l'on obtient :

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

Par conséquent, la matrice de Vandermonde V est inversible si et seulement si les points d'interpolation $\{x_i\}_{i=0,1,\dots,n}$ sont deux à deux distincts.

Dans ce cas, on obtient :

$$A = V^{-1}F$$

et donc, il existe un unique n -uplet $(a_0, a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ qui vérifie le système matriciel.

En conclusion, nous avons bien montré qu'il existe un unique polynôme d'interpolation P_n lié aux points $\{x_i, f(x_i)\}_{i=0,1,\dots,n}$

Remarque :

En pratique, cette méthode matricielle n'est pas utilisée, car il faut inverser la matrice de Vandermonde V , qui est une matrice pleine de taille $[n + 1, n + 1]$.

Le coût en terme de nombre d'opérations nécessaires pour inverser la matrice V (par exemple, en utilisant la méthode d'élimination de Gauss) est de l'ordre de $\frac{1}{3}(n + 1)^3$, ce qui est trop important.

Dans la partie qui suit, nous allons essayer d'écrire le polynôme d'interpolation P_n sous une autre forme.

Nous allons rechercher une expression explicite des coefficients $(a_0, a_1, a_2, \dots, a_n)$ qui nécessite moins d'opérations que la Méthode matricielle.

4.3 Interpolation de Lagrange

4.3.1 Base de Lagrange

Une première amélioration du calcul de $P_n(x)$ consiste à exprimer ce polynôme dans une autre base : la base de Lagrange.

Définition 1

Soient $\{x_i\}_{i=0,1,\dots,n}$ $n+1$ points deux à deux distincts. On appelle base de Lagrange relative aux points x_i , les polynômes :

$$l_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j} \quad \forall i = 0, 1, \dots, n$$

On peut écrire :

$$l_i(x) = \frac{(x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)}$$

4.3.2 Propriétés des polynômes de Lagrange

1. Les polynômes de Lagrange sont des polynômes de degré exactement n , qui vérifient la propriété :

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

2. Les polynômes $l_i(x)$ forment une base de l'espace des polynômes de degré inférieur ou égal à n : $\mathbb{R}^n[X]$.

Démonstration

La famille des polynômes de Lagrange $\{l_i(x)\}_{i=0,1,\dots,n}$ est composée de $n+1$ polynômes de degré n $l_i(x)$.

Pour montrer qu'ils forment une base de l'espace des polynômes $\mathbb{R}^n[X]$, il faut montrer que les polynômes $l_i(x)$ sont linéairement indépendants, c'est-à-dire :

$$\sum_{i=0}^n \lambda_i l_i(x) = 0, \quad \forall x \in \mathbb{R} \implies \lambda_i = 0 \quad \forall i = 0, 1, \dots, n$$

Il suffit de poser successivement : $x = x_0, x = x_1, x = x_2, \dots, x = x_n$.

Ainsi, pour $x = x_0$, on obtient :

$$\sum_{i=0}^n \lambda_i l_i(x_0) = 0 \iff \lambda_0 l_0(x_0) + \lambda_1 l_1(x_0) + \dots + \lambda_n l_n(x_0) = \lambda_0 l_0(x_0) + 0 = \lambda_0 = 0$$

D'où on obtient $\lambda_0 = 0$.

De même, pour $x = x_1$, on obtient :

$$\sum_{i=0}^n \lambda_i l_i(x_1) = 0 \iff \lambda_0 l_0(x_1) + \lambda_1 l_1(x_1) + \dots + \lambda_n l_n(x_1) = 0 + \lambda_1 l_1(x_1) + 0 = \lambda_1 = 0$$

D'où on obtient $\lambda_1 = 0$.

Et ainsi, on démontre que : $\forall i = 0, 1, \dots, n, \lambda_i = 0$.

Nous avons donc montré que les polynômes de Lagrange $l_i(x)$ sont bien linéairement indépendants, et qu'ils forment une base de l'espace des polynômes $\mathbb{R}^n[X]$.

4.3.3 Polynôme d'interpolation de Lagrange

Définition 2

Le polynôme qui interpole les valeurs $\{f(x_i)\}_{i=0,1,\dots,n}$ aux points $\{x_i\}_{i=0,1,\dots,n}$ s'écrit dans la base de Lagrange :

$$P_n(x) = \sum_{i=0}^n f(x_i) l_i(x)$$

Théorème 3

Le polynôme d'interpolation de Lagrange est l'unique polynôme de degré inférieur ou égal à n , qui vérifie :

$$P_n(x_j) = f(x_j) \quad \forall j = 0, 1, \dots, n$$

Démonstration :

$$P_n(x_j) = \sum_{i=0}^n f(x_i) l_i(x_j) = \sum_{i=0}^n f(x_i) \delta_{ij} = f(x_j) \quad \forall j = 0, 1, \dots, n$$

4.3.4 Exemples d'application

Exemple 1 :

On considère 3 points :

$$A(0, 1); \quad B(2, 5); \quad C(4, 17)$$

Déterminer le polynôme d'interpolation de Lagrange de degré inférieur ou égal à 2, passant par ces 3 points.

On pose :

$$(x_0 = 0, y_0 = 1), \quad (x_1 = 2, y_1 = 5), \quad (x_2 = 4, y_2 = 17)$$

Calculons les polynômes de Lagrange $l_i(x) \quad \forall i = 0, 1, 2$:

$$l_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 2)(x - 4)}{(0 - 2)(0 - 4)} = \frac{(x - 2)(x - 4)}{8}$$

$$l_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 0)(x - 4)}{(2 - 0)(2 - 4)} = \frac{x(x - 4)}{4}$$

$$l_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x - 0)(x - 2)}{(4 - 0)(4 - 2)} = \frac{x(x - 2)}{8}$$

Le polynôme d'interpolation de Lagrange s'écrit :

$$P_2(x) = \sum_{i=0}^2 y_i l_i(x) = y_0 l_0(x) + y_1 l_1(x) + y_2 l_2(x)$$

On obtient :

$$P_2(x) = l_0(x) + 5 l_1(x) + 17 l_2(x) = \frac{(x-2)(x-4)}{8} + \frac{5x(x-4)}{4} + \frac{17x(x-2)}{8}$$

Après simplifications, on obtient :

$$P_2(x) = 1 + x^2$$

On observe que le polynôme d'interpolation P_2 passant par les 3 points A , B , et C , est d'ordre 2.

Exemple 2 :

On souhaite interpoler la fonction $f(x) = \frac{1}{x}$ en utilisant les 3 points $x_0 = 2$, $x_1 = 2.5$, et $x_2 = 4$.

1- Ecrire le polynôme d'interpolation $P(x)$ qui interpole la fonction f pour ces 3 points.

2- Calculer l'erreur obtenue au point $x = 3$.

1- On pose :

$$\begin{cases} x_0 = 2 \\ f(x_0) = 0.5 \end{cases} \quad \begin{cases} x_1 = 2.5 \\ f(x_1) = 0.4 \end{cases} \quad \begin{cases} x_2 = 4 \\ f(x_2) = \frac{1}{4} = 0.25 \end{cases}$$

Calculons les polynômes de Lagrange $l_i(x)$ pour $i = 0, 1, 2$:

$$l_0(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} = \frac{(x-2.5)(x-4)}{(2-2.5)(2-4)} = x^2 - 6.5x + 10$$

$$l_1(x) = \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} = \frac{(x-2)(x-4)}{(2.5-2)(2.5-4)} = \frac{-4x^2 + 24x - 32}{3}$$

$$l_2(x) = \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} = \frac{(x-2)(x-2.5)}{(4-2)(4-2.5)} = \frac{x^2 - 4.5x + 5}{3}$$

Le polynôme d'interpolation de Lagrange s'écrit :

$$P(x) = \sum_{i=0}^2 f(x_i) l_i(x) = f(x_0) l_0(x) + f(x_1) l_1(x) + f(x_2) l_2(x)$$

On obtient :

$$P_2(x) = 0.5 l_0(x) + 0.4 l_1(x) + 0.25 l_2(x) = \frac{x^2 - 6.5x + 10}{2} + \frac{0.4}{3}(-4x^2 + 24x - 32) + \frac{1}{12}(x^2 - 4.5x + 5)$$

Après simplifications, on obtient :

$$P(x) = 0.05x^2 - 0.425x + 1.15$$

Là encore, on observe que le polynôme d'interpolation est d'ordre 2.

2- Calculons une valeur approchée de $f(3) = \frac{1}{3}$ à l'aide du polynôme d'interpolation :

$$P(3) = 0.325$$

L'erreur au point $x = 3$ vérifie :

$$e = |f(3) - P(3)| = |0.33333333... - 0.325| = 0.00833333...$$

4.4 Interpolation de Newton

Les polynômes de Lagrange $l_i(x)$ ne sont pas pratiques du point de vue numérique.

En effet, supposons que l'on ait calculé le polynôme d'interpolation $P_n(x)$ associé aux $n + 1$ points d'interpolation $\{x_i, f(x_i)\}_{i=0,1,\dots,n}$, et que l'on souhaite ajouter un nouveau point d'interpolation $(x_{n+1}, f(x_{n+1}))$ afin d'augmenter la précision.

Avec l'interpolation de Lagrange, on ne peut pas déduire $l_{i+1}(x)$ à partir de $l_i(x)$, et donc, on ne peut pas utiliser le polynôme d'interpolation $P_n(x)$ pour calculer le nouveau polynôme d'interpolation $P_{n+1}(x)$. Pour cette raison, on introduit le polynôme d'interpolation de Newton écrit dans la base de Newton.

4.4.1 Base de Newton

Une base de polynômes particulièrement adaptée au calcul du polynôme d'interpolation P_n est la base de Newton.

Définition 3

Soient $\{x_i\}_{i=0,1,\dots,n}$ $n + 1$ points deux à deux distincts. On appelle base de Newton relative aux points x_i , les polynômes :

$$\begin{cases} e_0(x) = 1 \\ e_i(x) = \prod_{j=0}^{i-1} (x - x_j) = (x - x_0)(x - x_1)\dots(x - x_{i-1}) \quad \forall i = 1, \dots, n \end{cases}$$

On peut écrire :

$$\begin{cases} e_0(x) = 1 \\ e_1(x) = (x - x_0) \\ e_2(x) = (x - x_0)(x - x_1) \\ \vdots \\ e_n(x) = (x - x_0)(x - x_1)(x - x_2)\dots(x - x_{n-1}) \end{cases}$$

4.4.2 Propriétés des polynômes de Newton

1. Les $n + 1$ polynômes de Newton vérifient la propriété :

$$e_i(x) = (x - x_{i-1}) e_{i-1}(x) \quad \text{pour } i \geq 1$$

2. Les polynômes $e_i(x)$ forment une base de l'espace des polynômes de degré inférieur ou égal à n : $\mathbb{R}^n[X]$.

Démonstration

Pour montrer qu'ils forment une base de l'espace des polynômes $\mathbb{R}^n[X]$, on montre comme auparavant, que les polynômes $e_i(x)$ sont linéairement indépendants, c'est-à-dire :

$$\sum_{i=0}^n \lambda_i e_i(x) = 0, \quad \forall x \in \mathbb{R} \implies \lambda_i = 0 \quad \forall i = 0, 1, \dots, n$$

Cela s'obtient en prenant successivement $x = x_j$ pour $j = 0, 1, \dots, n$.

4.4.3 Polynôme d'interpolation de Newton

Définition 4

L'unique polynôme qui interpole les valeurs $\{f(x_i)\}_{i=0,1,\dots,n}$ aux points $\{x_i\}_{i=0,1,\dots,n}$ s'écrit dans la base de Newton :

$$P_n(x) = \sum_{i=0}^n \alpha_i e_i(x) = \alpha_0 + \alpha_1(x-x_0) + \alpha_2(x-x_0)(x-x_1) + \dots + \alpha_n(x-x_0)(x-x_1)\dots(x-x_{n-1})$$

où les coefficients (α_i) doivent être calculés. Pour cela, nous allons utiliser la technique des différences divisées.

4.4.4 Différences divisées

Théorème 4

Le polynôme d'interpolation de Newton est l'unique polynôme de degré inférieur ou égal à n , qui vérifie :

$$P_n(x_j) = f(x_j) \quad \forall j = 0, 1, \dots, n$$

Ainsi, calculons successivement $P_n(x_0)$, $P_n(x_1)$, $P_n(x_2)$, ..., $P_n(x_n)$.

1. Le polynôme d'interpolation de Newton $P_n(x)$, évalué au point x_0 donne :

$$P_n(x_0) = \sum_{i=0}^n \alpha_i e_i(x_0) = \alpha_0 = f(x_0)$$

On note :

$$\alpha_0 = f[x_0]$$

$f[x_0]$ s'appelle différence divisée d'ordre 0.

D'une manière générale, on note :

$$f(x_i) = f[x_i] \quad \forall i = 0, 1, \dots, n$$

2. Le polynôme d'interpolation de Newton $P_n(x)$, évalué au point x_1 donne :

$$P_n(x_1) = \sum_{i=0}^n \alpha_i e_i(x_1) = \alpha_0 + \alpha_1(x_1 - x_0) = f(x_1)$$

On en déduit :

$$\alpha_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

On note :

$$\alpha_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{f[x_1] - f[x_0]}{x_1 - x_0} = f[x_0, x_1]$$

$f[x_0, x_1]$ s'appelle différence divisée d'ordre 1.

3. Le polynôme d'interpolation de Newton $P_n(x)$, évalué au point x_2 donne :

$$P_n(x_2) = \sum_{i=0}^n \alpha_i e_i(x_2) = \alpha_0 + \alpha_1(x_2 - x_0) + \alpha_2(x_2 - x_0)(x_2 - x_1) = f(x_2)$$

En remplaçant α_0 et α_1 par leurs valeurs, on obtient :

$$P_n(x_2) = f[x_0] + f[x_0, x_1](x_2 - x_0) + \alpha_2(x_2 - x_0)(x_2 - x_1) = f[x_2]$$

On en déduit :

$$\alpha_2(x_2 - x_0)(x_2 - x_1) = f[x_2] - f[x_0] - f[x_0, x_1](x_2 - x_0)$$

ce qui s'écrit encore :

$$\alpha_2(x_2 - x_0)(x_2 - x_1) = f(x_2) - f(x_0) - \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x_2 - x_0)$$

$$\alpha_2(x_2 - x_0)(x_2 - x_1) = f(x_2) - f(x_0) + \mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_1) - \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x_2 - x_0)$$

$$\alpha_2(x_2 - x_0)(x_2 - x_1) = f(x_2) - f(x_0) + \mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_1) - \frac{f(x_1) - f(x_0)}{x_1 - x_0}(\mathbf{x}_2 - \mathbf{x}_1) - \frac{f(x_1) - f(x_0)}{x_1 - x_0}(\mathbf{x}_1 - \mathbf{x}_0)$$

En simplifiant, on obtient :

$$\alpha_2(x_2 - x_0)(x_2 - x_1) = f(x_2) - f(x_1) - \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x_2 - x_1)$$

On en déduit :

$$\alpha_2(x_2 - x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} - \frac{f(x_1) - f(x_0)}{x_1 - x_0} = f[x_1, x_2] - f[x_0, x_1]$$

Finalement, on obtient :

$$\alpha_2 = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = f[x_0, x_1, x_2]$$

On appelle $f[x_0, x_1, x_2]$ différence divisée d'ordre 2.

4. On obtient par récurrence :

$$\alpha_i = \frac{f[x_1, x_2, \dots, x_i] - f[x_0, x_1, \dots, x_{i-1}]}{x_i - x_0} = f[x_0, x_1, x_2, \dots, x_i]$$

$f[x_0, x_1, x_2, \dots, x_i]$ s'appelle la différence divisée d'ordre i .

Exemple

Supposons que l'on souhaite calculer les coefficients du polynôme d'interpolation $P_3(x)$ lié aux 4 points $(x_0, f(x_0))$, $(x_1, f(x_1))$, $(x_2, f(x_2))$ et $(x_3, f(x_3))$.

Pour calculer les coefficients α_i , $i = 0, 1, 2, 3$, on fait le tableau des différences divisées :

$$\begin{bmatrix} x_0 & f[x_0] & & & \\ x_1 & f[x_1] & f[x_0, x_1] & & \\ x_2 & f[x_2] & f[x_1, x_2] & f[x_0, x_1, x_2] & \\ x_3 & f[x_3] & f[x_2, x_3] & f[x_1, x_2, x_3] & f[x_0, x_1, x_2, x_3] \end{bmatrix}$$

Pour obtenir les coefficients α_0 , α_1 , α_2 et α_3 , nous devons calculer toutes les différences divisées présentes dans le tableau, en commençant par la gauche.

Ainsi on a par exemple :

$$f[x_1, x_2] = \frac{f[x_2] - f[x_1]}{x_2 - x_1}$$

$$f[x_2, x_3] = \frac{f[x_3] - f[x_2]}{x_3 - x_2}$$

$$f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1}$$

$$f[x_0, x_1, x_2, x_3] = \frac{f[x_1, x_2, x_3] - f[x_0, x_1, x_2]}{x_3 - x_0}$$

Les coefficients du polynôme d'interpolation $P_3(x)$ sont les différences divisées présentes dans la diagonale du tableau.

Ainsi, on obtient :

$$\begin{cases} \alpha_0 = f[x_0] \\ \alpha_1 = f[x_0, x_1] \\ \alpha_2 = f[x_0, x_1, x_2] \\ \alpha_3 = f[x_0, x_1, x_2, x_3] \end{cases}$$

4.4.5 Expression finale du polynôme d'interpolation de Newton

Théorème 5

Le polynôme d'interpolation de Newton de degré inférieur ou égal à n $P_n(x)$ lié aux points $\{x_i, f(x_i)\}_{i=0,1,\dots,n}$ s'écrit :

$$P_n(x) = f[x_0] + \sum_{i=1}^n f[x_0, x_1, x_2, \dots, x_i] e_i(x)$$

ou encore :

$$P_n(x) = f[x_0] + f[x_0, x_1](x-x_0) + f[x_0, x_1, x_2](x-x_0)(x-x_1) + f[x_0, x_1, x_2, x_3](x-x_0)(x-x_1)(x-x_2) + \dots + f[x_0, x_1, x_2, x_3, \dots, x_n](x-x_0)(x-x_1)(x-x_2)\dots(x-x_{n-1})$$

Exemple

On reprend l'Exemple 1 vu dans l'interpolation de Lagrange.

On considère 3 points :

$$A(0, 1) \quad B(2, 5) \quad C(4, 17)$$

Déterminer le polynôme d'interpolation de Newton de degré au plus 2, passant par ces 3 points.

On pose :

$$(x_0 = 0, f(x_0) = 1), \quad (x_1 = 2, f(x_1) = 5), \quad (x_2 = 4, f(x_2) = 17)$$

D'après le Théorème 5, le polynôme d'interpolation de Newton s'écrit :

$$P_2(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1)$$

$$P_2(x) = f[x_0] + f[x_0, x_1](x - 0) + f[x_0, x_1, x_2](x - 0)(x - 2)$$

Pour calculer les coefficients de $P_2(x)$, nous allons effectuer le tableau des différences divisées :

$$\left[\begin{array}{lll} x_0 = 0 & f[x_0] = 1 & \\ x_1 = 2 & f[x_1] = 5 & f[x_0, x_1] = \frac{5-1}{2-0} = 2 \\ x_2 = 4 & f[x_2] = 17 & f[x_1, x_2] = \frac{17-5}{4-2} = 6 \quad f[x_0, x_1, x_2] = \frac{6-2}{4-0} = 1 \end{array} \right]$$

Ainsi, on obtient :

$$P_2(x) = 1 + 2x + x(x-2) = 1 + x^2$$

Nous retrouvons bien le même polynôme d'interpolation qu'avec la Méthode de Lagrange. En effet, le polynôme d'interpolation est unique.

4.4.6 Réécriture du Polynôme de Newton

Cas d'une discrétisation uniforme

Dans la section précédente, nous avons donné l'expression du polynôme d'interpolation de Newton dans le cas où les points d'interpolation sont répartis d'une manière quelconque sur un intervalle $[a, b]$:

$$a = x_0 < x_1 < x_2 < \dots < x_n = b$$

Nous allons maintenant réécrire le polynôme d'interpolation de Newton dans le cas d'une répartition uniforme des points $\{x_i\}_{i=0,1,\dots,n}$ sur l'intervalle $[a, b]$.

On pose :

$$\forall i = 0, 1, \dots, n-1, \quad x_{i+1} - x_i = h = \frac{b-a}{n}$$

Soit un point $x \in \mathbb{R}$. On peut écrire :

$$\exists s \in \mathbb{R}, \quad x = x_0 + sh$$

Si $x \in [a, b] = [x_0, x_n]$ alors $s > 0$.

On peut écrire :

$$x - x_i = (x_0 + sh) - (x_0 + ih) = (s-i)h$$

Dans ce cas, le polynôme d'interpolation de Newton se réécrit :

$$P_n(x) = P_n(x_0 + sh) = f[x_0] + sh f[x_0, x_1] + s(s-1) h^2 f[x_0, x_1, x_2] + \dots + s(s-1) \dots (s-n+1) h^n f[x_0, x_1, \dots, x_n]$$

ou encore :

$$P_n(x) = f[x_0] + \sum_{k=1}^n s(s-1) \dots (s-k+1) h^k f[x_0, x_1, \dots, x_k]$$

Or, d'après l'expression du binôme de Newton, on a :

$$C_s^k = \frac{s(s-1)(s-2) \dots (s-k+1)}{k!} = \frac{s!}{k!(s-k)!}$$

Par suite, le polynôme d'interpolation de Newton s'écrit :

$$\forall x \in \mathbb{R}, \quad P_n(x) = P_n(x_0 + sh) = f[x_0] + \sum_{k=1}^n C_s^k k! h^k f[x_0, x_1, \dots, x_k]$$

Cas particulier

On traite ici le cas particulier où les points d'interpolation $\{x_i\}_{i=0,1,\dots,n}$ sont ordonnés comme suit :

$$x_n > x_{n-1} > x_{n-2} > \dots > x_1 > x_0$$

Dans ce cas, le polynôme d'interpolation de Newton s'écrit :

$$P_n(x) = f[x_n] + f[x_n, x_{n-1}](x - x_n) + f[x_n, x_{n-1}, x_{n-2}](x - x_n)(x - x_{n-1}) \\ + \dots + f[x_n, x_{n-1}, x_{n-2}, \dots, x_0](x - x_n)(x - x_{n-1}) \dots (x - x_1)$$

Supposons de plus que la répartition des points d'interpolation $\{x_i\}_{i=0,1,\dots,n}$ est uniforme :

$$\forall i = 0, 1, \dots, n-1, \quad x_{i+1} - x_i = h = \frac{b-a}{n}$$

Soit un point $x \in \mathbb{R}$. On peut écrire :

$$\exists s \in \mathbb{R}, \quad x = x_n + sh$$

Si $x \in [a, b] = [x_0, x_n]$ alors $s < 0$.

On peut écrire :

$$x - x_n = sh \\ x - x_{n-1} = (x - x_n) + (x_n - x_{n-1}) = sh + h = (s+1)h \\ x - x_{n-2} = (x - x_{n-1}) + (x_{n-1} - x_{n-2}) = (s+1)h + h = (s+2)h$$

On en déduit :

$$P_n(x) = f[x_n] + f[x_n, x_{n-1}](sh) + f[x_n, x_{n-1}, x_{n-2}]s(s+1)h^2 \\ + \dots + f[x_n, x_{n-1}, x_{n-2}, \dots, x_0]s(s+1) \dots (s+n-1)h^n$$

ce qui peut encore s'écrire :

$$P_n(x) = P_n(x_n + sh) = f[x_n] + \sum_{k=1}^n s(s+1) \dots (s+k-1) h^k f[x_n, x_{n-1}, \dots, x_{n-k}] \quad (4.1)$$

Définition de l'opérateur Gradient

On définit l'opérateur Gradient comme suit :

$$\begin{cases} \forall n \geq 1, & \nabla f(x_n) = f(x_n) - f(x_{n-1}) \\ \forall k \geq 2, & \nabla^k f(x_n) = \nabla (\nabla^{k-1} f(x_n)) \end{cases}$$

Ainsi, on a :

$$\nabla^2 f(x_n) = \nabla f(x_n) - \nabla f(x_{n-1}) \\ \nabla^2 f(x_n) = f(x_n) - f(x_{n-1}) - (f(x_{n-1}) - f(x_{n-2})) = f(x_n) - 2f(x_{n-1}) + f(x_{n-2})$$

Lien entre l'opérateur Gradient et les Différences divisées

Nous pouvons écrire :

$$\nabla f(x_n) = \frac{f(x_n) - f(x_{n-1})}{h} \cdot h = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} \cdot h$$

Donc, on obtient :

$$\nabla f(x_n) = f[x_n, x_{n-1}] \cdot h$$

De même, nous pouvons écrire :

$$\nabla^2 f(x_n) = (f(x_n) - f(x_{n-1})) - (f(x_{n-1}) - f(x_{n-2})) = \left(\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} - \frac{f(x_{n-1}) - f(x_{n-2})}{x_{n-1} - x_{n-2}} \right) \cdot h$$

$$\nabla^2 f(x_n) = (f[x_n, x_{n-1}] - f[x_{n-1}, x_{n-2}]) \cdot h$$

ce qui peut encore s'écrire :

$$\nabla^2 f(x_n) = \frac{f[x_n, x_{n-1}] - f[x_{n-1}, x_{n-2}]}{x_n - x_{n-2}} \cdot h (x_n - x_{n-2})$$

$$\nabla^2 f(x_n) = f[x_n, x_{n-1}, x_{n-2}] \cdot h (2h)$$

Finalement, on obtient :

$$\nabla^2 f(x_n) = f[x_n, x_{n-1}, x_{n-2}] \cdot (2h^2)$$

On en déduit :

$$f[x_n, x_{n-1}] = \frac{\nabla f(x_n)}{h}, \text{ et}$$

$$f[x_n, x_{n-1}, x_{n-2}] = \frac{\nabla^2 f(x_n)}{2h^2}$$

Plus généralement, on a la relation :

$$\forall k \geq 1, \quad f[x_n, x_{n-1}, \dots, x_{n-k}] = \frac{\nabla^k f(x_n)}{k! h^k}$$

En remplaçant dans l'équation (4.1), on en déduit l'expression du polynôme de Newton :

$$P_n(x) = P_n(x_n + sh) = f[x_n] + \sum_{k=1}^n \frac{s(s+1)\dots(s+k-1)}{k!} \nabla^k f(x_n)$$

Or nous avons vu que si le point $x \in [a, b] = [x_0, x_n]$, alors $s < 0$.

Ecrivons maintenant la formule du binôme de Newton C_{-s}^k dans le cas où $-s > 0$:

$$C_{-s}^k = \frac{(-s)!}{k! (-s-k)!} = \frac{(-s)(-s-1)(-s-2)\dots(-s-k+1)}{k!}$$

On en déduit :

$$C_{-s}^k = (-1)^k \frac{s(s+1)(s+2)\dots(s+k-1)}{k!}$$

Théorème 6

Dans le cas particulier où les points d'interpolation $\{x_i\}_{i=0,1,\dots,n}$ sont ordonnés uniformément comme suit :

$$x_n > x_{n-1} > x_{n-2} > \dots > x_1 > x_0,$$

le polynôme d'interpolation de Newton s'écrit :

$$P_n(x) = f[x_n] + \sum_{k=1}^n (-1)^k C_{-s}^k \nabla^k f(x_n)$$

C'est cette répartition des points $\{x_i\}_{i=0,1,\dots,n}$ et cette écriture du polynôme d'interpolation de Newton que nous utiliserons dans le Chapitre suivant : Méthodes numériques à pas multiples.

4.5 Erreur d'interpolation

4.5.1 Définition

On va à présent analyser l'erreur que l'on commet lorsqu'on remplace une fonction f par son polynôme d'interpolation P_n associé aux $n + 1$ points d'interpolation x_0, x_1, \dots, x_n tous distincts. Le théorème suivant fournit une estimation de la différence.

Théorème 7

Soit f une fonction $n + 1$ fois dérivable sur \mathbb{R} , et soit P_n le polynôme d'interpolation lié aux $n + 1$ points $\{x_i, f(x_i)\}_{i=0,1,\dots,n}$ où les points $\{x_i\}_{i=0,1,\dots,n}$ sont deux-à-deux distincts. Alors, pour tout point $x \in \mathbb{R}$, on a l'estimation d'erreur suivante :

$$f(x) - P_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} (x - x_0)(x - x_1) \dots (x - x_n)$$

On voit bien dans la formule, que pour $x = x_i$, $f(x_i) = P_n(x_i)$ et l'erreur est nulle.

La question que l'on se pose à présent est de savoir si, sur un intervalle $[a, b]$ donné, l'erreur d'interpolation diminue lorsqu'on augmente le nombre de points d'interpolation, et si quand n tend vers l'infini, le polynôme d'interpolation P_n converge vers la fonction f ?

Le Théorème 7 permet de déduire la majoration de l'erreur suivante :

Corollaire

$$|f(x) - P_n(x)| \leq \frac{|(x - x_0)(x - x_1) \dots (x - x_n)|}{(n+1)!} \max_{\xi \in [a,b]} |f^{(n+1)}(\xi)|$$

où $[a, b]$ est le plus petit intervalle contenant le point x et tous les points $\{x_i\}_{i=0,1,\dots,n}$.

4.5.2 Exemple

On considère les valeurs suivantes pour la fonction $f(x) = e^x$:

$$e^{0.1} = 1.10517 \quad e^{0.2} = 1.22140 \quad e^{0.3} = 1.34986 \quad e^{0.4} = 1.49182$$

On souhaite approcher la valeur de $e^{0.22}$ à 10^{-4} près, à l'aide d'un polynôme d'interpolation. Quels points d'interpolation faut-il choisir pour approcher au mieux $e^{0.22}$?

Il faut choisir les points les plus proches de 0.22.

1. Choisissons 2 points d'interpolation : $x_0 = 0.2$ et $x_1 = 0.3$.

Le polynôme d'interpolation lié à ces points est de degré 1, et d'après le Corollaire, nous avons pour $x = 0.22$:

$$|e^{0.22} - P_1(0.22)| \leq \frac{|(x - x_0)(x - x_1)|}{2!} \max_{\xi \in [0.2, 0.3]} |e^\xi|$$

On obtient :

$$|e^{0.22} - P_1(0.22)| \leq \frac{|(0.22 - 0.2)(0.22 - 0.3)|}{2!} |e^{0.3}| \approx 0.00108 > 10^{-4}$$

Ainsi, il ne suffit pas de choisir seulement 2 points d'interpolation.

2. Rajoutons le 3ème point d'interpolation $x_2 = 0.1$: dans ce cas, le polynôme d'interpolation est de degré au plus 2.

L'estimation de l'erreur d'interpolation au point $x = 0.22$ nous donne :

$$|e^{0.22} - P_2(0.22)| \leq \frac{|(x - x_0)(x - x_1)(x - x_2)|}{3!} \max_{\xi \in [0.1, 0.3]} |e^\xi|$$

On obtient :

$$|e^{0.22} - P_2(0.22)| \leq \frac{|(0.22 - 0.1)(0.22 - 0.2)(0.22 - 0.3)|}{3!} |e^{0.3}| \approx 4.3210^{-5}$$

On a bien atteint la précision demandée avec les 3 points d'interpolation $x_0 = 0.2$, $x_1 = 0.3$ et $x_2 = 0.1$.

4.5.3 Phénomène de Runge

De par cet exemple, on peut penser qu'en augmentant le nombre de points d'interpolation x_i c'est-à-dire en augmentant le degré du polynôme d'interpolation P_n , celui-ci va approcher de mieux en mieux la fonction f sur $[a, b]$.

De plus, d'après le Corollaire, nous avons :

$$\frac{|(x - x_0)(x - x_1) \dots (x - x_n)|}{(n + 1)!} \leq \frac{(b - a)^{n+1}}{(n + 1)!} \xrightarrow{n \rightarrow +\infty} 0$$

ce qui semble confirmer cette intuition.

Cependant, il ne faut pas oublier le terme $\max_{\xi \in [a, b]} |f^{(n+1)}(\xi)|$ qui peut tendre vers l'infini quand n augmente.

Il s'agit du phénomène de Runge, découvert par l'allemand Carl Runge en 1901.

Exemple

On considère la fonction $f(x) = \frac{1}{1 + 25x^2}$ sur l'intervalle $[-1, 1]$.

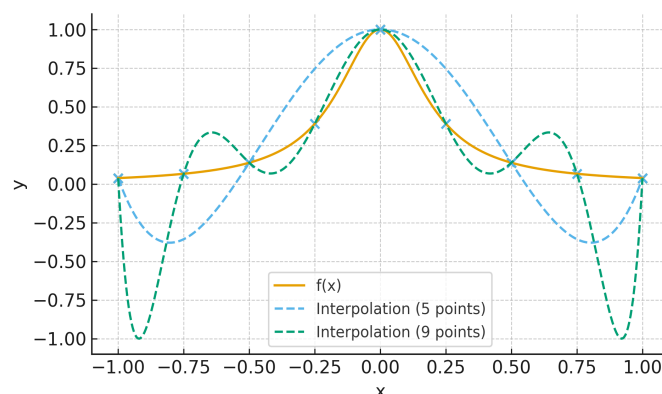


FIGURE 4.2 – Exemple de phénomène de Runge

Sur la Figure 4.2, on trace en jaune la courbe de la fonction $f(x)$.

La courbe bleue représente la courbe du polynôme d'interpolation P_5 associés aux 6 points

d'interpolation équidistants de l'intervalle $[-1, 1]$, et la courbe verte représente la courbe du polynôme d'interpolation P_9 associés aux 10 points d'interpolation équidistants de l'intervalle $[-1, 1]$.

On constate que le polynôme d'interpolation P_9 ne donne pas une meilleure approximation de $f(x)$ que P_5 , surtout dans les zones proches des extrémités.

Plus précisément, on observe des oscillations qui apparaissent aux extrémités de l'intervalle $[-1, 1]$, et on voit que les oscillations s'amplifient quand on augmente le nombre de points d'interpolation.

Pour atténuer ce phénomène, on choisit une discrétisation non uniforme pour les points d'interpolation, en utilisant les abscisses de Tchebychev :

$$x_i = \cos\left(\frac{(2i+1)\pi}{2n+2}\right) \quad \forall i = 0, 1, \dots, n$$

On peut voir que ces abscisses sont plus denses aux extrémités de l'intervalle (en 1 et en -1) et qu'elles sont réparties symétriquement autour de 0.

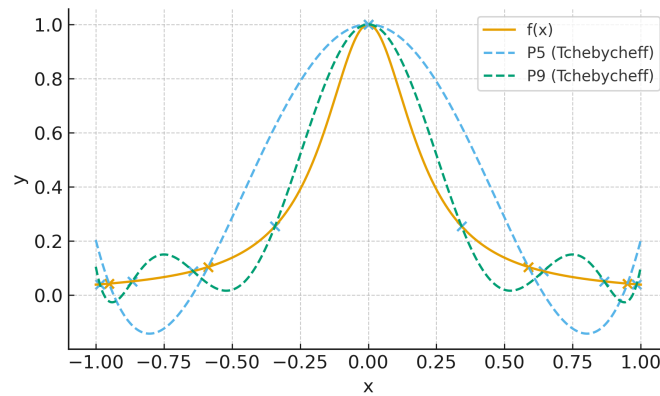


FIGURE 4.3 – Atténuation du phénomène de Runge en utilisant les abscisses de Tchebycheff

Sur la Figure 4.3, on trace la fonction $f(x)$, et les polynômes d'interpolation P_5 et P_9 , en utilisant les abscisses de Tchebycheff.

On constate que, contrairement au cas des points équidistants, P_5 et P_9 restent très proches de $f(x)$ sur tout l'intervalle $[-1, 1]$, avec beaucoup moins d'oscillations aux bords.