

Mathématiques assistées par ordinateur

Chapitre 8 : Intégration numérique

Michael Eisermann

Mat249, DLST L2S4, Année 2008-2009

www-fourier.ujf-grenoble.fr/~eiserm/cours#mao

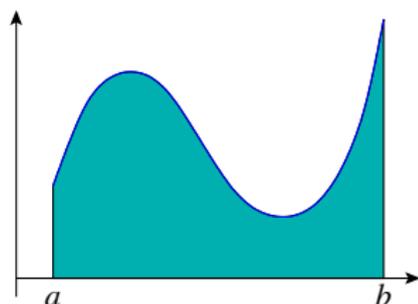
Document mis à jour le 6 juillet 2009



Motivation et objectifs

Géométriquement, l'intégrale $\int_a^b f(x)dx$ d'une fonction continue $f: [a, b] \rightarrow \mathbb{R}$ mesure l'aire entre l'axe des abscisses et f .

L'intégration est aussi l'opération « inverse » à la dérivation. Si $F: [a, b] \rightarrow \mathbb{R}$ vérifie $F' = f$, alors $\int_a^b f(x)dx = F(b) - F(a)$.



Dans des rares cas favorables on arrive à **expliciter** une telle fonction F , dite primitive de f . Ceci permet de calculer certaines intégrales.

En général c'est trop dur, voire impossible : la plupart des fonctions n'admet pas de primitives s'exprimant à l'aide des fonctions usuelles.

Motivé par cette problématique, ce cours présente quelques méthodes pour calculer **numériquement** de telles intégrales.

Comme toujours, avant de calculer quoi que ce soit, il faut d'abord **définir** l'objet en question, puis établir ses principales propriétés. On commencera donc avec une révision de l'intégrale de Riemann.

Sommaire

- 1 L'intégrale de Riemann : construction et propriétés
 - Construction de l'intégrale de Riemann
 - Propriétés principales de l'intégrale
 - Différentiation et intégration
- 2 Méthodes numériques basiques
 - Méthodes des rectangles, majoration d'erreur
 - Méthode des trapèzes, majoration d'erreur
 - Comparaison des méthodes basiques
- 3 Méthodes de Newton-Cotes
 - Méthodes élémentaires, majoration d'erreur
 - Méthodes composées, majoration d'erreur
 - Méthodes de Simpson, Boole, Weddle
- 4 La méthode de Romberg
 - La formule d'Euler-Maclaurin
 - Extrapolation de Richardson
 - La méthode de Romberg

Qu'est-ce que l'intégrale ?

On veut associer à $f: \mathbb{R} \rightarrow \mathbb{R}$ et $a < b$ un nombre réel $\int_a^b f$, appelé l'intégrale de f sur $[a, b]$, vérifiant quelques exigences naturelles :

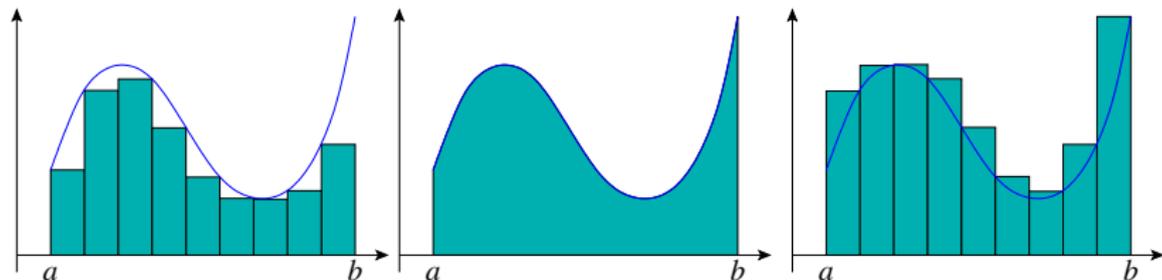
Constantes : $\int_a^b \lambda = (b - a)\lambda$ pour tout $\lambda \in \mathbb{R}$

Subdivision : $\int_a^c f = \int_a^b f + \int_b^c f$ pour tout $a < b < c$

Monotonie : $\int_a^b f \leq \int_a^b g$ si $f(x) \leq g(x)$ pour $a < x < b$

De ces axiomes découlent toutes les propriétés de l'intégrale !

Tout d'abord, ils définissent l'intégrale pour toute fonction en escalier :



Puis tout s'étend aux autres fonctions intégrables par encadrement.

Construction de l'intégrale : fonctions en escaliers

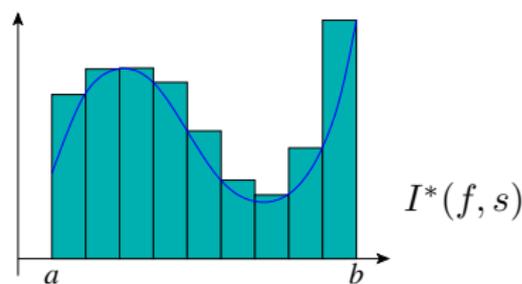
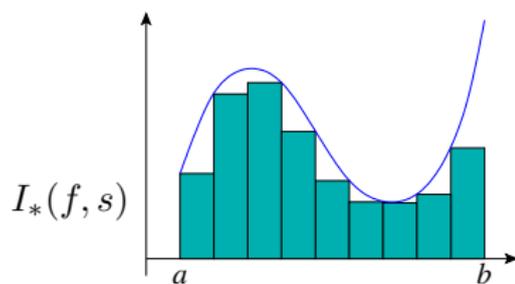
Une **subdivision** de $[a, b]$ en n sous-intervalles est une famille $s = \{s_0 < s_1 < \dots < s_{n-1} < s_n\}$ telle que $s_0 = a$ et $s_n = b$.

Pour toute fonction bornée $f: [a, b] \rightarrow \mathbb{R}$ on définit

$$I_*(f, s) := \sum_{k=1}^n (s_k - s_{k-1}) \inf_{[s_{k-1}, s_k]} f,$$

$$I^*(f, s) := \sum_{k=1}^n (s_k - s_{k-1}) \sup_{[s_{k-1}, s_k]} f.$$

Ces deux formules explicitent notre heuristique :



Observation : Si $s' \supset s$ est une subdivision plus fine de $[a, b]$, alors

$$I_*(f, s) \leq I_*(f, s') \leq I^*(f, s') \leq I^*(f, s).$$

Construction de l'intégrale : passage à la limite

On passe alors à des subdivisions de plus en plus fines :

$$I_*(f) := \sup \{ I_*(f, s) \mid s \text{ une subdivision de } [a, b] \}$$

$$I^*(f) := \inf \{ I^*(f, s) \mid s \text{ une subdivision de } [a, b] \}$$

Si jamais l'intégrale de f existe, elle est encadrée par $I_*(f)$ et $I^*(f)$.

On a toujours $I_*(f) \leq I^*(f)$, mais cette inégalité peut être stricte !
Dans ce cas on n'arrive pas à définir l'intégrale cherchée.

Si $I_*(f) = I^*(f)$, alors on n'a qu'une seule possibilité pour l'intégrale !
Ce cas favorable mérite une définition digne de son importance :

Définition (intégrale de Riemann)

On dit que $f : [a, b] \rightarrow \mathbb{R}$ est **intégrable** si $I_*(f) = I^*(f)$.

Dans ce cas son **intégrale** est $\int_a^b f(x)dx := I_*(f) = I^*(f)$.

Caractérisation des fonctions intégrables

Proposition

Une fonction $f: [a, b] \rightarrow \mathbb{R}$ est intégrable si et seulement si pour tout $\varepsilon > 0$ il existe une subdivision s telle que $I^*(f, s) - I_*(f, s) \leq \varepsilon$.

Démonstration. Supposons f intégrable et notons $I := \int_a^b f(x)dx$. Alors pour $\varepsilon > 0$ il existe des subdivisions s_* et s^* de $[a, b]$ telles que

$$I^*(f, s^*) \leq I + \varepsilon/2 \quad \text{et} \quad I_*(f, s_*) \geq I - \varepsilon/2.$$

Pour la subdivision $s = s_* \cup s^*$ on obtient ainsi

$$I_*(f, s_*) \leq I_*(f, s) \leq I \leq I^*(f, s) \leq I^*(f, s^*).$$

Réciproquement, $I^*(f, s) - I_*(f, s) \leq \varepsilon$ et

$$I_*(f, s) \leq I_*(f) \leq I^*(f) \leq I^*(f, s),$$

entraînent $0 \leq I^*(f) - I_*(f) \leq \varepsilon$. Si cette inégalité est vraie pour tout $\varepsilon > 0$, alors $I^*(f) - I_*(f) = 0$, et on conclut que f est intégrable. \square

Fonctions monotones, fonctions continues

Théorème

Toute fonction monotone $f : [a, b] \rightarrow \mathbb{R}$ est intégrable.

Démonstration. Supposons que f est croissante. Soit $s_k = a + kh$ la subdivision équidistante où $h = \frac{b-a}{n}$ et $k = 0, \dots, n$.

$$\begin{aligned} I^*(f, s) - I_*(f, s) &= \sum_{k=1}^n (s_k - s_{k-1}) [f(s_k) - f(s_{k-1})] \\ &= h[f(b) - f(a)] \rightarrow 0 \quad \text{pour } n \rightarrow \infty. \end{aligned}$$

On conclut en faisant appel à la caractérisation précédente. □

Théorème

Toute fonction continue $f : [a, b] \rightarrow \mathbb{R}$ est intégrable.

Démonstration. Puisque $[a, b]$ est compact, la fonction f est **uniformément continue**, càd pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que $|x - x'| \leq \delta$ implique $|f(x) - f(x')| \leq \varepsilon$. Pour un pas $h \leq \delta$ on obtient

$$I^*(f, s) - I_*(f, s) = \sum_{k=1}^n (s_k - s_{k-1}) \varepsilon \leq (b - a) \varepsilon.$$

On conclut en faisant appel à la caractérisation précédente. □

C Exemples de fonctions discontinues

Exemple (trop sauvage pour être Riemann-intégrable)

La fonction de Dirichlet $f: [0, 1] \rightarrow \mathbb{R}$ est définie par

$$f(x) = \begin{cases} 1 & \text{si } x \in \mathbb{Q}, \\ 0 & \text{si } x \notin \mathbb{Q}. \end{cases}$$

On trouve $I_*(f) = 0$ et $I^*(f) = 1$, donc f n'est pas intégrable.

Exemple (encore sauvage mais Riemann-intégrable)

Considérons $g: [0, 1] \rightarrow \mathbb{R}$ définie par

$$g(x) = \begin{cases} \frac{1}{b} & \text{si } x = \frac{a}{b} \text{ où } a, b \in \mathbb{Z}, b > 0, \text{pgcd}(a, b) = 1, \\ 0 & \text{si } x \notin \mathbb{Q}. \end{cases}$$

On trouve $I_*(g) = 0$ et $I^*(g) = 0$, donc g est intégrable et $\int_0^1 g dx = 0$.

À noter que g n'est continue sur aucun intervalle $[a, b]$ où $a < b$.
Surprenant mais vrai : g est continue en tout $x \in [0, 1] \setminus \mathbb{Q}$.

Propriétés de l'intégrale

Théorème (propriétés caractérisantes)

Soit \mathcal{R}_a^b l'ensemble des fonctions intégrables $f: [a, b] \rightarrow \mathbb{R}$.

- Si $f(x) = \lambda$, alors $f \in \mathcal{R}_a^b$ et $\int_a^b f dx = \lambda \cdot (b - a)$.
- Pour $a < b < c$ on a $f \in \mathcal{R}_a^c \iff f|_{[a,b]} \in \mathcal{R}_a^b \wedge f|_{[b,c]} \in \mathcal{R}_b^c$.
Dans ce cas $\int_a^c f dx = \int_a^b f dx + \int_b^c f dx$.
- Si $f, g \in \mathcal{R}_a^b$ et $f \leq g$ sur $]a, b[$, alors $\int_a^b f dx \leq \int_a^b g dx$. □

Autrement dit, l'intégrale de Riemann satisfait aux exigences naturelles formulées au début. On récolte bien plus encore :

Théorème (linéarité et monotonie)

- Si $f \in \mathcal{R}$ et $\lambda \in \mathbb{R}$, alors $\lambda f \in \mathcal{R}$ et $\int_a^b \lambda f dx = \lambda \int_a^b f dx$.
- Si $f, g \in \mathcal{R}$, alors $f + g \in \mathcal{R}$ et $\int_a^b (f + g) dx = \int_a^b f dx + \int_a^b g dx$.
- Si $f, g \in \mathcal{R}$, alors $fg \in \mathcal{R}$. En général $\int_a^b fg dx \neq \int_a^b f dx \cdot \int_a^b g dx$!
- Si $f \in \mathcal{R}$, alors $|f| \in \mathcal{R}$ et $|\int_a^b f dx| \leq \int_a^b |f| dx$. □

Différentiation et intégration

Théorème

Supposons que $f: [a, b] \rightarrow \mathbb{R}$ est intégrable. Si $f = F'$ pour une fonction dérivable $F: [a, b] \rightarrow \mathbb{R}$, alors $\int_a^b f(x)dx = F(b) - F(a)$.

Démonstration. Pour toute partition $s = \{s_0 < s_1 < \dots < s_{n-1} < s_n\}$ de $[a, b]$ il existe $s_0 < t_1 < s_1 < \dots < s_{n-1} < t_n < s_n$ tels que

$$F(s_k) - F(s_{k-1}) = (s_k - s_{k-1})f(t_k),$$

par le théorème des accroissements finis. Ainsi

$$F(b) - F(a) = \sum_{k=1}^n F(s_k) - F(s_{k-1}) = \sum_{k=1}^n (s_k - s_{k-1})f(t_k).$$

On en déduit que $I_*(f, s) \leq F(b) - F(a) \leq I^*(f, s)$.

Puis $I_*(f) \leq F(b) - F(a) \leq I^*(f)$ par passage aux limites.

Si f est intégrable, alors on a l'égalité $F(b) - F(a) = \int_a^b f(x)dx$. □

La propriété de la moyenne

Théorème

Soit $f: [a, b] \rightarrow \mathbb{R}$ une fonction continue. Pour toute fonction intégrable $\phi: [a, b] \rightarrow \mathbb{R}$ vérifiant $\phi \geq 0$ il existe $\xi \in [a, b]$ tel que

$$(*) \quad \int_a^b f(x)\phi(x)dx = f(\xi) \int_a^b \phi(x)dx.$$

En particulier, pour $\phi = 1$, on obtient $\int_a^b f(x)dx = f(\xi)(b - a)$.
Si $f = F'$ c'est le théorème des accroissements finis appliqué à F .

Démonstration.

Pour $m = \min f$ et $M = \max f$ on a $m\phi \leq f\phi \leq M\phi$. Ainsi

$$m \int_a^b \phi(x)dx \leq \int_a^b f(x)\phi(x)dx \leq M \int_a^b \phi(x)dx.$$

Si $\int_a^b \phi(x)dx = 0$, alors l'équation (*) est vraie pour tout $\xi \in [a, b]$.

Sinon, il existe $\xi \in [a, b]$ tel que $f(\xi) = \int_a^b f(x)\phi(x)dx / \int_a^b \phi(x)dx$. \square

Intégration et différentiation

Théorème

Pour $f: \mathbb{R} \rightarrow \mathbb{R}$ continue on peut définir $F: \mathbb{R} \rightarrow \mathbb{R}$ par $F(x) := \int_a^x f(t)dt$. Alors F est dérivable et $F' = f$.

Démonstration. Pour tout $x \in \mathbb{R}$ et $h > 0$ on trouve

$$\begin{aligned} F(x+h) - F(x) &= \int_a^{x+h} f(t)dt - \int_a^x f(t)dt \\ &= \int_x^{x+h} f(t)dt = hf(\xi_h) \end{aligned}$$

pour un $\xi_h \in [x, x+h]$ d'après la propriété de la moyenne.

Pour $h \rightarrow 0$ on a donc $\xi_h \rightarrow x$, puis

$$\frac{F(x+h) - F(x)}{h} = f(\xi_h) \rightarrow f(x)$$

par continuité de f . Ceci prouve le théorème. □

Résumé : construction de l'intégrale de Riemann

Si $f: [a, b] \rightarrow \mathbb{R}$ est continue (ou modérément discontinue)
nous savons définir l'intégrale (dite de Riemann)

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

Théorème fondamental : si nous disposons d'une fonction primitive
 $F: [a, b] \rightarrow \mathbb{R}$ vérifiant $F' = f$, alors $\int_a^b f(x)dx = F(b) - F(a)$.

Toute fonction continue f admet une telle primitive F .

Cette relation entre différentiation et intégration est très utile pour le calcul, notamment elle mène directement à

- l'intégration par partie,
- la substitution des variables,
- le calcul explicite de certaines intégrales.

Révisez votre cours d'intégration pour un développement complet.
Dans ce cours-ci on utilise ces outils sans retenue si besoin en est.

À quoi sert l'intégration numérique ?

Pour $f: [a, b] \rightarrow \mathbb{R}$ nous souhaitons calculer l'intégrale $\int_a^b f(x)dx$.

Si nous disposons d'une primitive F , alors $\int_a^b f(x)dx = F(b) - F(a)$, ce qui résout notre problème d'une manière très satisfaisante.

① Peut-on toujours expliciter une primitive sous « forme close » ?

Non ! Pour $f(x) = \exp(-x^2/2)$, par exemple, il n'existe pas de fonction primitive exprimable par des fonctions usuelles \exp, \ln, \dots

② Dans ce cas, comment être sûr que l'intégrale existe ?

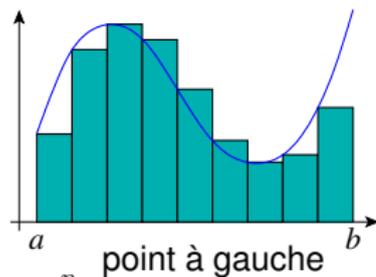
Heureusement, la construction précédente prouve que l'intégrale existe, même si nous n'arrivons pas à expliciter une formule close.

③ Peut-on au moins calculer des valeurs approchées de l'intégrale ?

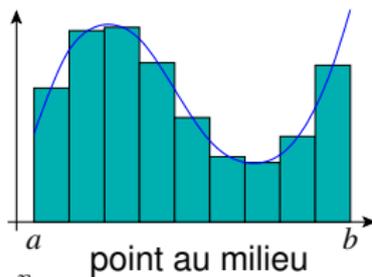
Oui ! Dans la suite nous développons des méthodes d'intégration numérique qui sont plus profitables et plus efficaces dans la pratique que les formules utilisées pour la construction ci-dessus.

Méthodes des rectangles

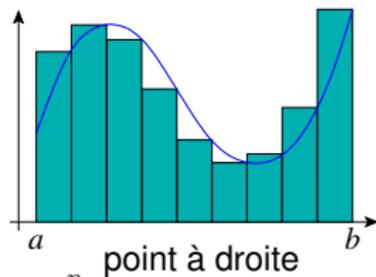
On approche l'intégrale par $\sum_{k=1}^n (s_k - s_{k-1}) f(\xi_k)$ où $\xi_k \in [s_{k-1}, s_k]$.



$$\sum_{k=1}^n (s_k - s_{k-1}) f(s_{k-1})$$



$$\sum_{k=1}^n (s_k - s_{k-1}) f\left(\frac{s_{k-1} + s_k}{2}\right)$$



$$\sum_{k=1}^n (s_k - s_{k-1}) f(s_k)$$

Cas particulier d'une subdivision équidistante $s_k = a + kh$ où $h = \frac{b-a}{n}$:

$$h \sum_{k=0}^{n-1} f(a + kh)$$

$$h \sum_{k=0}^{n-1} f\left(a + \frac{h}{2} + kh\right)$$

$$h \sum_{k=1}^n f(a + kh)$$

Exemple : Pour $f(x) = x^3$ on connaît l'intégrale $\int_0^1 f(x) dx = 0.25$.

Une subdivision équidistante à $n = 10$ pas donne respectivement les approximations $\frac{81}{400} = 0.2025$ et $\frac{199}{800} = 0.24875$ et $\frac{121}{400} = 0.3025$.

On constate que l'approximation par le point au milieu est nettement plus précise. Ceci n'est pas un hasard, et on verra pourquoi. . .

Point à gauche ou à droite : majoration de l'erreur

Proposition

Soit $f: [a, b] \rightarrow \mathbb{R}$ de classe C^1 . Pour $n \geq 1$ soit $h = \frac{b-a}{n}$ et soit

$$R_n^{\text{gauche}}(f) := h \sum_{k=0}^{n-1} f(a + kh)$$

l'approximation obtenue par le point à gauche.

Alors on a la majoration d'erreur

$$\left| \int_a^b f(x) dx - R_n^{\text{gauche}}(f) \right| \leq \frac{(b-a)^2}{2} \cdot \frac{1}{n} \cdot \max_{[a,b]} |f'|.$$

La même majoration est valable pour le point à droite.

Pour augmenter la précision obtenue on augmentera n , ce qui revient à passer à des subdivisions de plus en plus fines.

Or, la précision n'augmente que très lentement : $\sim \frac{1}{n}$.

Pour gagner un bit de précision il faut deux fois plus de travail.

Point à gauche ou à droite : majoration de l'erreur

Démonstration.

On considère l'intervalle $[\alpha, \beta]$ où $\alpha = a + (k - 1)h$ et $\beta = a + kh$.

Le théorème des accroissements finis (Taylor d'ordre 0) assure que

$$f(x) - f(\alpha) = f'(\xi)(x - \alpha)$$

pour un $\xi \in [\alpha, x]$. En particulier on a

$$|f(x) - f(\alpha)| \leq \max_{[\alpha, \beta]} |f'| \cdot (x - \alpha).$$

On en déduit que

$$\begin{aligned} \left| \int_{\alpha}^{\beta} f(x) dx - (\beta - \alpha) f(\alpha) \right| &= \left| \int_{\alpha}^{\beta} f(x) dx - \int_{\alpha}^{\beta} f(\alpha) dx \right| \\ &= \left| \int_{\alpha}^{\beta} f(x) - f(\alpha) dx \right| \leq \int_{\alpha}^{\beta} |f(x) - f(\alpha)| dx \\ &\leq \max_{[\alpha, \beta]} |f'| \cdot \int_{\alpha}^{\beta} (x - \alpha) dx = \max_{[\alpha, \beta]} |f'| \cdot \frac{(\beta - \alpha)^2}{2}. \end{aligned}$$

La proposition s'en déduit en sommant sur $k = 1, \dots, n$. □

Point au milieu : majoration de l'erreur

Proposition

Soit $f : [a, b] \rightarrow \mathbb{R}$ de classe C^2 . Pour $n \geq 1$ soit $h = \frac{b-a}{n}$ et soit

$$R_n(f) := h \sum_{k=0}^{n-1} f\left(a + \frac{h}{2} + kh\right)$$

l'approximation obtenue par le point au milieu.

Alors on a la majoration d'erreur

$$\left| \int_a^b f(x) dx - R_n(f) \right| \leq \frac{(b-a)^3}{24} \cdot \frac{1}{n^2} \cdot \max_{[a,b]} |f''|.$$

Pour augmenter la précision obtenue on augmentera n , ce qui revient à passer à des subdivisions de plus en plus fines.

Ici la précision augmente un peu plus rapidement : $\sim \frac{1}{n^2}$.

Pour gagner deux bits de précision il faut deux fois plus de travail.

Point au milieu : majoration de l'erreur

Démonstration.

On considère l'intervalle $[\alpha, \beta]$ où $\alpha = a + (k-1)h$ et $\beta = a + kh$.

On fait un développement d'ordre 1 autour du milieu $x_0 = \frac{\alpha + \beta}{2}$:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(\xi)(x - x_0)^2.$$

Ainsi $|f(x) - f(x_0) - f'(x_0)(x - x_0)| \leq \frac{1}{2} \max |f''|(x - x_0)^2$.

On constate que $\int_{\alpha}^{\beta} f'(x_0)(x - x_0)dx = 0$. Par conséquent

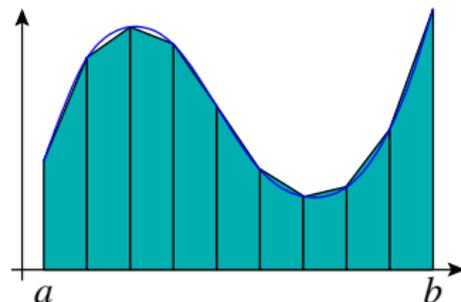
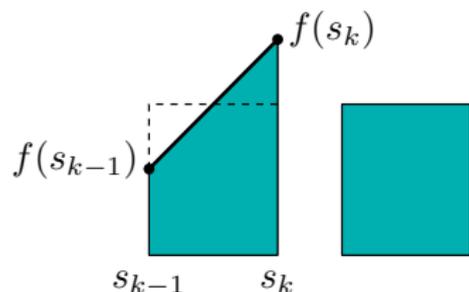
$$\begin{aligned} \left| \int_{\alpha}^{\beta} f(x)dx - (\beta - \alpha)f(x_0) \right| &= \left| \int_{\alpha}^{\beta} f(x) - f(x_0) - f'(x_0)(x - x_0)dx \right| \\ &\leq \int_{\alpha}^{\beta} |f(x) - f(x_0) - f'(x_0)(x - x_0)|dx \leq \frac{1}{2} \max_{[\alpha, \beta]} |f''| \int_{\alpha}^{\beta} (x - x_0)^2 dx \\ &= \frac{1}{2} \max_{[\alpha, \beta]} |f''| \cdot \left[\frac{1}{3}(x - x_0)^3 \right]_{\alpha}^{\beta} = \frac{1}{24} \max_{[\alpha, \beta]} |f''| \cdot (\beta - \alpha)^3. \end{aligned}$$

La proposition s'en déduit en sommant sur $k = 1, \dots, n$. □

Méthode des trapèzes

Idée : on interpole entre les points gauche et droite par un segment.

La figure obtenue est un trapèze d'aire $(s_k - s_{k-1}) \frac{f(s_{k-1}) + f(s_k)}{2}$.



Pour une subdivision à pas constant h on obtient $\frac{h}{2} f(s_{k-1}) + \frac{h}{2} f(s_k)$.

Dans la somme sur $k = 1, \dots, n$ tout terme apparaît deux fois, à l'exception des termes $\frac{h}{2} f(a)$ et $\frac{h}{2} f(b)$ aux bords. Ainsi

$$T_n(f) = h \left[\frac{1}{2} f(a) + f(a+h) + \dots + f(b-h) + \frac{1}{2} f(b) \right]$$

Dans notre exemple $\int_0^1 x^3 dx$ on trouve l'approximation $\frac{101}{400} = 0.2525$.

Méthode des trapèzes : majoration de l'erreur

Lemme

Soit $f: [\alpha, \beta] \rightarrow \mathbb{R}$ de classe C^2 . Alors il existe $\xi \in [\alpha, \beta]$ tel que

$$\int_{\alpha}^{\beta} f(x)dx = \frac{\beta - \alpha}{2} (f(\alpha) + f(\beta)) - \frac{(\beta - \alpha)^3}{12} f''(\xi).$$

Démonstration. La fonction $\phi(x) = -\frac{1}{2}(x - \alpha)(x - \beta)$ vérifie $\phi(x) \geq 0$ pour tout $x \in [\alpha, \beta]$. On a $\phi'(x) = \frac{1}{2}(\alpha + \beta) - x$ puis $\phi''(x) = -1$. Ainsi

$$\begin{aligned} \int_{\alpha}^{\beta} f(x)dx &= \int_{\alpha}^{\beta} -\phi''(x)f(x)dx = \left[-\phi'(x)f(x)\right]_{\alpha}^{\beta} + \int_{\alpha}^{\beta} \phi'(x)f'(x)dx \\ &= \frac{\beta - \alpha}{2} (f(\alpha) + f(\beta)) + \left[\phi(x)f'(x)\right]_{\alpha}^{\beta} - \int_{\alpha}^{\beta} \phi(x)f''(x)dx \\ &= \frac{\beta - \alpha}{2} (f(\alpha) + f(\beta)) - f''(\xi) \int_{\alpha}^{\beta} \phi(x)dx \\ &= \frac{\beta - \alpha}{2} (f(\alpha) + f(\beta)) - f''(\xi) \frac{(\beta - \alpha)^3}{12}. \end{aligned}$$

Méthode des trapèzes : majoration de l'erreur

Proposition

Soit $f : [a, b] \rightarrow \mathbb{R}$ de classe C^2 . Pour $n \geq 1$ soit $h = \frac{b-a}{n}$ et soit

$$T_n(f) := h \left[\frac{1}{2} f(a) + f(a+h) + \cdots + f(b-h) + \frac{1}{2} f(b) \right]$$

l'approximation obtenue par la méthode des trapèzes. Alors on a

$$\left| \int_a^b f(x) dx - T_n(f) \right| \leq \frac{(b-a)^3}{12} \cdot \frac{1}{n^2} \cdot \max_{[a,b]} |f''|.$$

Démonstration. Le lemme s'applique à chaque sous-intervalle :

$$\left| \int_{a+(k-1)h}^{a+kh} f(x) dx - \frac{1}{2} \left(f(a+(k-1)h) + f(a+kh) \right) \right| \leq \frac{h^3}{12} \max |f''|$$

Le théorème s'en déduit en sommant sur $k = 1, \dots, n$. □

À nouveau la précision augmente proportionnellement à $\frac{1}{n^2}$.
Pour gagner deux bits de précision il faut deux fois plus de travail.

Comparaison des méthodes basiques

Approximation de $\int_a^b f(x)dx$ avec n pas de longueur $h = \frac{b-a}{n}$.

Méthode	Ordre	Majoration d'erreur
Point à gauche	$d = 0$	$\frac{(b-a)^2}{2} \cdot \frac{1}{n} \cdot \max f' $
Point à droite	$d = 0$	$\frac{(b-a)^2}{2} \cdot \frac{1}{n} \cdot \max f' $
Point au milieu	$d = 1$	$\frac{(b-a)^3}{24} \cdot \frac{1}{n^2} \cdot \max f'' $
Trapèzes	$d = 1$	$\frac{(b-a)^3}{12} \cdot \frac{1}{n^2} \cdot \max f'' $

On suppose tacitement que f est suffisamment dérivable.
Les bornes sont atteintes pour un polynôme de degré $d + 1$.

Exemple : Pour approcher $\int_0^1 \exp(-x^2/2)dx$ avec ces méthodes à 10^{-6} près il faudra environ $n = 10^6$ respectivement $n = 10^3$, à 10^{-12} près il faudra environ $n = 10^{12}$ respectivement $n = 10^6$, à 10^{-18} près il faudra environ $n = 10^{18}$ respectivement $n = 10^9$. (C'est souvent trop lent, même sur des ordinateurs puissants.)

Méthodes élémentaires d'intégration

On cherche des formules pour approcher l'intégrale $\int_{\alpha}^{\beta} f(x)dx$.
Or, l'intégrale est un « procédé infini » qui ne s'implémente pas tel quel sur ordinateur : on ne peut effectuer qu'un nombre fini d'opérations.

Nos formules approchées d'intégration seront donc de la forme

$$S(f) = (\beta - \alpha) \sum_{j=0}^m w_j f(\xi_j).$$

Ici $\xi_0, \dots, \xi_m \in [\alpha, \beta]$ sont les points où l'on évalue la fonction f ,
et $w_0, \dots, w_m \in \mathbb{R}$ sont certains poids tels que $w_j > 0$ et $\sum_{j=0}^m w_j = 1$.

Une telle somme définit une application $S: C([\alpha, \beta]) \rightarrow \mathbb{R}$.

Linéarité : $S(\lambda f + \mu g) = \lambda S(f) + \mu S(g)$, par définition.

Constantes : $S(\lambda) = (\beta - \alpha)\lambda$, parce que $\sum_{j=0}^m w_j = 1$.

Monotonie : Si $f \leq g$ alors $S(f) \leq S(g)$, parce que $w_j > 0$.

En particulier on a l'inégalité $|S(f)| \leq S(|f|)$.

Ordre d'une méthode d'intégration

Dans la suite notre objectif est de choisir les points ξ_j et les poids w_j dans le souci de minimiser le travail et l'erreur de l'approximation.

On a déjà vu dans les exemples basiques que ce choix détermine la qualité de l'approximation et la vitesse de convergence pour $n \rightarrow \infty$.

La classification suivante nous guidera dans notre démarche :

Définition

Une méthode d'intégration $S: C([\alpha, \beta]) \rightarrow \mathbb{R}$ est **d'ordre** d si elle est exacte pour tout polynôme f de degré $\leq d$, c'à d $S(f) = \int_{\alpha}^{\beta} f(x)dx$.

Grâce à notre convention $\sum w_j = 1$ nos formules sont exactes pour les constantes (degré ≤ 0). Elles sont donc au moins d'ordre 0.

Les méthodes du point au milieu et des trapèzes sont d'ordre 1.

On construira judicieusement des méthodes d'ordre supérieur.
Tout d'abord on montre comment l'ordre se traduit en précision.

Majoration d'erreur d'une méthode élémentaire (1/2)

Proposition

Soit $S: C([\alpha, \beta]) \rightarrow \mathbb{R}$ une méthode d'intégration d'ordre d . Alors

$$\left| \int_{\alpha}^{\beta} f(x) dx - S(f) \right| \leq \frac{(\beta - \alpha)^{d+2}}{2^d (d+1)!} \max_{[\alpha, \beta]} |f^{(d+1)}|$$

pour toute fonction $f: [\alpha, \beta] \rightarrow \mathbb{R}$ de classe C^{d+1} .

Démonstration. Soit $x_0 = \frac{\alpha + \beta}{2}$ le centre de l'intervalle $[\alpha, \beta]$ et soit

$$g(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots + \frac{f^{(d)}(x_0)}{d!} (x - x_0)^d$$

le polynôme de Taylor en x_0 . Nous avons la majoration d'erreur

$$\begin{aligned} |f(x) - g(x)| &\leq \frac{|x - x_0|^{d+1}}{(d+1)!} \max |f^{(d+1)}| \\ &\leq \frac{(\beta - \alpha)^{d+1}}{2^{d+1} (d+1)!} \max |f^{(d+1)}|. \end{aligned}$$

Majoration d'erreur d'une méthode élémentaire (2/2)

Puisque S est d'ordre d on a $\int_{\alpha}^{\beta} g(x)dx - S(g) = 0$. Ainsi

$$\begin{aligned} \left| \int_{\alpha}^{\beta} f(x)dx - S(f) \right| &= \left| \int_{\alpha}^{\beta} f(x) - g(x)dx - S(f - g) \right| \\ &\leq \int_{\alpha}^{\beta} |f(x) - g(x)|dx + S(|f - g|). \end{aligned}$$

En majorant $|f - g|$ deux fois par $\frac{(\beta - \alpha)^{d+1}}{2^{d+1}(d+1)!} \max |f^{(d+1)}|$ on obtient

$$\left| \int_{\alpha}^{\beta} f(x)dx - S(f) \right| \leq \frac{(\beta - \alpha)^{d+2}}{2^d(d+1)!} \max |f^{(d+1)}|.$$

On obtient ainsi la majoration souhaitée. □

 On ne prétend pas que la constante $\frac{1}{2^d(d+1)!}$ soit optimale.

- Le polynôme de Taylor peut être remplacé par une meilleure approximation : c'est le problème de l'approximation polynomiale.
- Pour optimiser la constante il faudra étudier de plus près la méthode d'intégration, comme le point milieu, les trapèzes, etc.

Méthodes composées d'intégration

L'erreur d'approximation d'une méthode élémentaire d'intégration d'ordre d dépend de deux quantités : $\max |f^{(d+1)}|$ et $(\beta - \alpha)^{d+2}$.

On a intérêt à l'appliquer à des **petits** intervalles $[\alpha, \beta]$!

Pour un **grand** intervalle $[a, b]$ on choisit une subdivision assez fine

$$a = s_0 < s_1 < \dots < s_{n-1} < s_n = b.$$

Souvent on prend des points équirépartis $s_k = a + kh$ où $h = \frac{b-a}{n}$.

Ceci permet de **décomposer** l'intégrale en une somme

$$\int_a^b f(x)dx = \sum_{k=1}^n \int_{s_{k-1}}^{s_k} f(x)dx.$$

On est ainsi ramené à intégrer f sur un petit intervalle $[s_{k-1}, s_k]$.

Sur chacun des petits intervalles, l'intégrale sera approchée par une méthode élémentaire d'intégration :

$$\int_{s_{k-1}}^{s_k} f(x)dx \approx (s_k - s_{k-1}) \sum_{j=0}^{m_k} w_{k,j} f(\xi_{k,j})$$

pour certains points $\xi_{k,j} \in [s_{k-1}, s_k]$ et des poids $w_{k,j} \in \mathbb{R}$.

Méthodes composées d'intégration

On obtient ainsi une **méthode composée d'intégration** :

$$\int_a^b f(x)dx = \sum_{k=1}^n \int_{s_{k-1}}^{s_k} f(x)dx \approx \sum_{k=1}^n (s_k - s_{k-1}) \sum_{j=0}^{m_k} w_{k,j} f(\xi_{k,j})$$

A priori on pourrait prendre une subdivision s_0, \dots, s_n quelconque et une méthode élémentaire différente sur chacun des intervalles.

Pour simplifier nous prenons la subdivision à pas constant $h = \frac{b-a}{n}$ et la « même » méthode élémentaire sur chaque sous-intervalle :

$$S: C([\alpha, \beta]) \rightarrow \mathbb{R}, \quad S(f) = (\beta - \alpha) \sum_{j=0}^m w_j f(\xi_j),$$

se transporte sur n'importe quel intervalle $[\alpha', \beta']$ par la bijection affine $\phi: [\alpha, \beta] \rightarrow [\alpha', \beta']$, $\phi(x) = (x - \alpha) \frac{\beta' - \alpha'}{\beta - \alpha} + \alpha'$. On obtient ainsi

$$S': C([\alpha', \beta']) \rightarrow \mathbb{R}, \quad S'(g) = (\beta' - \alpha') \sum_{j=0}^m w_j g(\xi'_j),$$

avec les mêmes poids w_j mais g étant évaluée en $\xi'_j = \phi(\xi_j)$.

Majoration d'erreur d'une méthode composée

Proposition

Soit S une méthode élémentaire d'intégration d'ordre d .

Soit $S_n: C([a, b]) \rightarrow \mathbb{R}$ la méthode composée associée à la subdivision équidistante $s_k = a + kh$ où $h = \frac{b-a}{n}$. Alors

$$\begin{aligned} \left| \int_a^b f(x) dx - S_n(f) \right| &\leq \frac{(b-a)^{d+2}}{2^d(d+1)!} \cdot \frac{1}{n^{d+1}} \cdot \max_{[a,b]} |f^{(d+1)}| \\ &= 2(b-a) \cdot \left(\frac{h}{2}\right)^{d+1} \cdot \max_{[a,b]} \left| \frac{f^{(d+1)}}{(d+1)!} \right| \end{aligned}$$

pour toute fonction $f: [a, b] \rightarrow \mathbb{R}$ de classe C^{d+1} .

Démonstration. Cette majoration découle de celle obtenue pour la méthode élémentaire S en sommant sur $k = 1, \dots, n$. □

◆ Pour approcher $\int_a^b f(x) dx$ on procède donc comme suit :

- On détermine ou majore $\max |f^{(d+1)}|$, qui ne dépend que de f .
- On choisit n assez grand pour assurer que l'erreur soit petite.

Il reste encore à trouver des méthodes d'ordre plus élevé.

Méthodes de Newton-Cotes

On souhaite construire des méthodes élémentaires d'intégration

$$S: C([\alpha, \beta]) \rightarrow \mathbb{R}, \quad S(f) = (\beta - \alpha) \sum_{j=0}^m w_j f(\xi_j)$$

pour des points équirépartis $\xi_j = \alpha + jh$ où $h = \frac{\beta - \alpha}{m}$.

L'idée est d'interpoler f par son polynôme de Lagrange

$$p = \sum_{j=0}^m f(\xi_j) p_j \quad \text{où} \quad p_j(x) = \prod_{i \neq j} \frac{x - \xi_i}{\xi_j - \xi_i}$$

puis d'intégrer p (de manière exacte) sur $[\alpha, \beta]$:

$$\int_{\alpha}^{\beta} p(x) dx = (\beta - \alpha) \sum_{j=0}^m f(\xi_j) w_j \quad \text{où} \quad w_j = \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} p_j(x) dx.$$

Exemple. Prenons $[-1, +1]$ et les extrémités $\xi_0 = -1$, $\xi_1 = +1$.

On a $p = f(\xi_0)p_0 + f(\xi_1)p_1$ où $p_0(x) = \frac{1-x}{2}$ et $p_1(x) = \frac{x+1}{2}$.

On trouve $w_0 = w_1 = \frac{1}{2}$ et ainsi $S(f) = (\beta - \alpha) \left[\frac{1}{2} f(\alpha) + \frac{1}{2} f(\beta) \right]$.

C'est la méthode des trapèzes !

La méthode de Simpson

Appliquons cette recette à $[-1, +1]$ et $\xi_0 = -1$, $\xi_1 = 0$, $\xi_2 = +1$.

$$p_0(x) = \frac{x-0}{-1-0} \cdot \frac{x-1}{-1-1} = \frac{1}{2}x(x-1) \quad \Rightarrow \quad w_0 = \frac{1}{2} \int_{-1}^1 p_0(x) dx = \frac{1}{6}$$

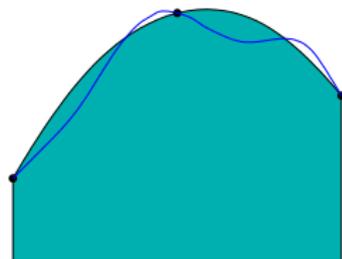
$$p_1(x) = \frac{x-(-1)}{0-(-1)} \cdot \frac{x-1}{0-1} = (x+1)(1-x) \quad \Rightarrow \quad w_1 = \frac{1}{2} \int_{-1}^1 p_1(x) dx = \frac{4}{6}$$

$$p_2(x) = \frac{x-(-1)}{1-(-1)} \cdot \frac{x-0}{1-0} = \frac{1}{2}(x+1)x \quad \Rightarrow \quad w_2 = \frac{1}{2} \int_{-1}^1 p_2(x) dx = \frac{1}{6}$$

On obtient ainsi la méthode de Simpson :

$$S(f) = (\beta - \alpha) \left[\frac{1}{6} f(\alpha) + \frac{4}{6} f\left(\frac{\alpha + \beta}{2}\right) + \frac{1}{6} f(\beta) \right]$$

Elle interpole f par une parabole.



La méthode de Simpson est d'ordre 3 :

Par construction elle est exacte si f est un polynôme de degré ≤ 2 .

Petit miracle : par symétrie elle est aussi exacte pour $f(x) = x^3$ car

$$\int_{-1}^1 f(x) dx = 0 = S(f).$$

Majoration d'erreur : cas général

On considère une méthode de Newton-Cotes

$$S : C([\alpha, \beta]) \rightarrow \mathbb{R}, \quad S(f) = \int_{\alpha}^{\beta} p(x) dx = (\beta - \alpha) \sum_{j=0}^m w_j f(\xi_j).$$

utilisant le polynôme interpolateur p en $m + 1$ points équirépartis.

Pour tout $x \in [\alpha, \beta]$ on a la majoration d'approximation de Lagrange

$$|f(x) - p(x)| \leq \frac{(\beta - \alpha)^{m+1}}{m^{m+2}} \max_{[\alpha, \beta]} |f^{(m+1)}|.$$

Pour l'intégration approchée on en déduit la majoration

$$\left| \int_{\alpha}^{\beta} f(x) dx - S(f) \right| \leq \frac{(\beta - \alpha)^{m+2}}{m^{m+2}} \max_{[\alpha, \beta]} |f^{(m+1)}|.$$

Pour la méthode composée $S_n : C([a, b]) \rightarrow \mathbb{R}$ on conclut que

$$\left| \int_a^b f(x) dx - S_n(f) \right| \leq \frac{(b - a)^{m+2}}{m^{m+2}} \cdot \frac{1}{n^{m+1}} \cdot \max_{[\alpha, \beta]} |f^{(m+1)}|.$$

Majoration d'erreur : cas où m est pair

Proposition

Si m est un nombre pair (par exemple $m = 2$ pour Simpson), alors la méthode de Newton-Cotes $S: C([\alpha, \beta]) \rightarrow \mathbb{R}$ est d'ordre $m + 1$.

Pour la méthode composée $S_n: C([a, b]) \rightarrow \mathbb{R}$ on conclut que

$$\left| \int_a^b f(x) dx - S_n(f) \right| \leq \frac{(b-a)^{m+3}}{2^{m+1}(m+2)!} \cdot \frac{1}{n^{m+2}} \cdot \max_{[\alpha, \beta]} |f^{(m+2)}|.$$

Démonstration. La méthode S est exacte si f est un polynôme de degré $\leq m$: par construction on a $f = p$ et donc $\int_a^b f(x) dx = S(f)$.

Pour simplifier considérons l'intervalle $[-1, +1]$. Si m est pair, alors $f(x) = x^{m+1}$ est impaire, donc $f(-x) = -f(x)$ et $\int_{-1}^1 f(x) dx = 0$.

D'un autre coté on a $w_{m-j} = w_j$ et $\xi_{m-j} = -\xi_j$, donc $S(f) = 0$. □

On ne prétend pas que la constante $\frac{(b-a)^{m+3}}{2^{m+1}(m+2)!}$ soit optimale. Une analyse plus fine améliorera légèrement cette majoration.

Résumé des méthodes de Newton-Cotes

Pour tout $m \geq 1$ on construit une méthode élémentaire d'intégration

$$S: C([\alpha, \beta]) \rightarrow \mathbb{R}, \quad S(f) = (\beta - \alpha) \sum_{j=0}^m w_j f(\xi_j)$$

en des points équirépartis $\xi_j = \alpha + jh$ où $h = \frac{\beta - \alpha}{m}$.

Méthode	m	Poids (w_0, \dots, w_m)	Ordre	Majoration d'erreur
Trapèzes	1	$\frac{1}{2}, \frac{1}{2}$	1	$\frac{(b-a)^3}{12} \cdot \frac{1}{n^2} \cdot \max f'' $
Simpson	2	$\frac{1}{6}, \frac{4}{6}, \frac{1}{6}$	3	$\frac{(b-a)^5}{2880} \cdot \frac{1}{n^4} \cdot \max f^{(4)} $
Boole	4	$\frac{7}{90}, \frac{32}{90}, \frac{12}{90}, \frac{32}{90}, \frac{7}{90}$	5	$\frac{(b-a)^7}{945 \cdot 2^{11}} \cdot \frac{1}{n^6} \cdot \max f^{(6)} $
Weddle	6	$\frac{41}{840}, \frac{216}{840}, \frac{27}{840}, \frac{272}{840}, \dots$	7	$\frac{(b-a)^9}{5600 \cdot 6^7} \cdot \frac{1}{n^8} \cdot \max f^{(8)} $

 Pour $m \geq 8$ les méthodes de Newton-Cotes font apparaître aussi des poids négatifs. On perd ainsi la monotonie et risque des erreurs numériques d'annulation. Ces méthodes sont donc peu utilisées.

La formule d'Euler-Maclaurin

Théorème (Euler-Maclaurin)

Soit $f: [a, b] \rightarrow \mathbb{R}$ de classe C^{2m+1} . Pour $n \geq 1$ soit $h = \frac{b-a}{n}$ et soit

$$A(h) := h \left[\frac{1}{2}f(a) + f(a+h) + \cdots + f(b-h) + \frac{1}{2}f(b) \right]$$

l'approximation obtenue par la méthode des trapèzes. Il existe des constantes $c_2, c_4, \dots, c_{2m}, c \in \mathbb{R}$ indépendantes de f et h telles que

$$A(h) = \int_a^b f(x)dx + c_2 h^2 [f']_a^b + \cdots + c_{2m} h^{2m} [f^{(2m-1)}]_a^b + r_{2m}(h)$$

avec un reste majoré par $|r_{2m}(h)| \leq c(b-a) \cdot h^{2m+1} \cdot \max |f^{(2m+1)}|$.

Dans la suite nous n'aurons pas besoin d'explicitier les constantes. Il nous suffira de connaître **la forme** du développement en h :

$$A(h) = \int_a^b f(x)dx + a_2 h^2 + \cdots + a_{2m} h^{2m} + O(h^{2m+1}).$$

 Pour un développement détaillé voir Demailly, chapitre 3, §4.

C Démonstration d'Euler-Maclaurin (1/2)

Considérons d'abord l'intervalle $[-1/2, 1/2]$ de longueur 1 centré en 0. Soit $(e_k)_{k \in \mathbb{N}}$ la famille de polynômes commençant par

$$e_0(x) = 1, \quad e_1(x) = x, \quad e_2(x) = \frac{1}{2}x^2 - \frac{1}{24}, \quad \dots$$

définis récursivement par $e'_k = e_{k-1}$ et $\int_{-1/2}^{+1/2} e_k dx = 0$ pour $k \geq 1$.

On voit que e_k est de degré k , de monôme dominant $\frac{1}{k!}x^k$.

Les polynômes e_0, e_2, e_4, \dots sont pairs, $e_{2j}(-x) = e_{2j}(x)$, alors que e_1, e_3, e_5, \dots sont impairs, $e_{2j+1}(-x) = -e_{2j+1}(x)$.

Pour $k \geq 2$ on trouve $e_k(1/2) - e_k(-1/2) = \int_{-1/2}^{+1/2} e_{k-1}(x) dx = 0$.

Par conséquent $e_{2j+1}(-1/2) = e_{2j+1}(1/2) = 0$ pour tout $j \geq 1$.

Pour les indices pairs on pose $c_{2j} := e_{2j}(-1/2) = e_{2j}(1/2)$.

Considérons ensuite un intervalle quelconque $[\alpha, \beta]$ de longueur $h = \beta - \alpha$, centré en $x_0 = \frac{\alpha + \beta}{2}$.

Pour $g_k(x) := h^k e_k\left(\frac{x-x_0}{h}\right)$ on a $g_0 = 1$ et $g'_k = g_{k-1}$ et $\int_{\alpha}^{\beta} g_k dx = 0$.

On $g_1(\alpha) = g_1(\beta) = \frac{h}{2}$. Ensuite, pour $j \geq 1$, on trouve $g_{2j}(\alpha) = g_{2j}(\beta) = c_{2j}h^{2j}$ et $g_{2j+1}(\alpha) = g_{2j+1}(\beta) = 0$.

C Démonstration d'Euler-Maclaurin (2/2)

Par intégration par partie nous obtenons successivement

$$\begin{aligned}\int_{\alpha}^{\beta} f dx &= \int_{\alpha}^{\beta} g_0 f dx = [g_1 f]_{\alpha}^{\beta} - \int_{\alpha}^{\beta} g_1 f' dx \\ &= [g_1 f]_{\alpha}^{\beta} - [g_2 f']_{\alpha}^{\beta} + \int_{\alpha}^{\beta} g_2 f'' dx = \dots \\ &= [g_1 f]_{\alpha}^{\beta} - [g_2 f']_{\alpha}^{\beta} + [g_3 f'']_{\alpha}^{\beta} - \dots - \int_{\alpha}^{\beta} g_{2m+1} f^{(2m+1)} dx \\ &= \frac{h}{2} [f(\alpha) + f(\beta)] - c_2 h^2 [f']_{\alpha}^{\beta} - c_4 h^4 [f''']_{\alpha}^{\beta} - \dots \\ &\quad - c_{2m} h^{2m} [f^{(2m-1)}]_{\alpha}^{\beta} - \int_{\alpha}^{\beta} g_{2m+1} f^{(2m+1)} dx\end{aligned}$$

Le reste est majoré par $h \cdot \max_{[\alpha, \beta]} |g_{2m+1}| \cdot \max_{[\alpha, \beta]} |f^{(2m+1)}|$.

On trouve $\max_{[\alpha, \beta]} |g_{2m+1}| = c \cdot h^{2m+1}$ où $c := \max_{[-1/2, 1/2]} |e_{2m+1}|$.

L'énoncé s'en déduit en sommant sur les sous-intervalles $[a, a+h], [a+h, a+2h], \dots, [b-2h, b-h], [b-h, b]$. □

Accélération de la convergence

Principe : Toute information sur l'erreur peut être utilisée pour augmenter la précision et accélérer la convergence.

Application : Dans notre cas nous avons le développement

$$A(h) = \int_a^b f(x)dx + a_2h^2 + a_4h^4 + \cdots + a_{2m}h^{2m} + O(h^{2m+1}).$$

Ceci veut dire que l'erreur est d'ordre h^2 , donc $\sim \frac{1}{n^2}$.

En passant du pas h au pas $h/2$, nous obtenons

$$A\left(\frac{h}{2}\right) = \int_a^b f(x)dx + a_2\frac{h^4}{4} + a_4\frac{h^4}{16} + \cdots + a_{2m}\frac{h^{2m}}{4^m} + O(h^{2m+1}).$$

Astuce. — On peut maintenant éliminer le terme d'ordre h^2 :

$$\frac{4A\left(\frac{h}{2}\right) - A(h)}{4 - 1} = \int_a^b f(x)dx + a'_4h^4 + \cdots + a'_{2m}h^{2m} + O(h^{2m+1})$$

Avantage. — Cette nouvelle formule réduit l'erreur à l'ordre h^4 .

Mieux encore, cette astuce peut être itérée pour éliminer successivement les termes d'ordre h^4, h^6, \dots, h^{2m} .

Extrapolation de Richardson : la formule magique

Soit $A(h) = a_0 + a_2h^2 + a_4h^4 + \dots + a_{2m}h^{2m} + O(h^{2m+1})$.

Ici a_0 est la valeur cherchée et a_2, a_4, \dots, a_{2m} décrivent l'erreur.

On sait calculer $A(h)$ pour $h > 0$ mais c'est coûteux si h est petit.

Posons $A^0(h) := A(h)$ puis par récurrence

$$A^j(h) := \frac{4^j A^{j-1}\left(\frac{h}{2}\right) - A^{j-1}(h)}{4^j - 1}.$$

Ceci donne des approximations d'ordre de plus en plus grand :

$$A^0(h) = a_0 + a_2^{(0)}h^2 + a_4^{(0)}h^4 + a_6^{(0)}h^6 + a_8^{(0)}h^8 + \dots + O(h^{2m+1})$$

$$A^1(h) = a_0 + a_4^{(1)}h^4 + a_6^{(1)}h^6 + a_8^{(1)}h^8 + \dots + O(h^{2m+1})$$

$$A^2(h) = a_0 + a_6^{(2)}h^6 + a_8^{(2)}h^8 + \dots + O(h^{2m+1})$$

$$A^3(h) = a_0 + a_8^{(3)}h^8 + \dots + O(h^{2m+1})$$

...

$$A^m(h) = a_0 + O(h^{2m+1})$$

Extrapolation de Richardson : algorithme

	étape 0		étape 1		étape 2	...	étape m
s_0	$A^0(h)$	\rightarrow	$A^1(h)$	\rightarrow	$A^2(h)$	\rightarrow	$A^m(h)$
s_1	$A^0(h/2)$	\nearrow	$A^1(h/2)$	\nearrow		\nearrow	
s_2	$A^0(h/2^2)$	\nearrow					
\vdots	\vdots		\vdots				
s_{m-2}	$A^0(h/2^{m-2})$	\rightarrow	$A^1(h/2^{m-2})$	\rightarrow	$A^2(h/2^{m-2})$		
s_{m-1}	$A^0(h/2^{m-1})$	\nearrow	$A^1(h/2^{m-1})$	\nearrow			
s_m	$A^0(h/2^m)$	\nearrow					

Algorithme 1 extrapolation de Richardson

Entrée: les valeurs $A^0(h/2^j)$ pour $j = 0, \dots, m$

Sortie: la valeur $A^m(h)$ comme spécifiée ci-dessus

$s_0 \leftarrow A^0(h)$

pour j **de** 1 **à** m **faire**

$s_j \leftarrow A^0(h/2^j)$

pour k **de** 0 **à** $j-1$ **faire** $s_k \leftarrow \frac{4^{j-k}s_{k+1}-s_k}{4^{j-k}-1}$ **fin pour**

fin pour

retourner s_0

La méthode de Romberg : algorithme

Soit $f: [a, b] \rightarrow \mathbb{R}$ de classe C^{2m+1} . Pour $n \geq 1$ soit $h = \frac{b-a}{n}$ et soit

$$A(h) := h \left[\frac{1}{2} f(a) + f(a+h) + \dots + f(b-h) + \frac{1}{2} f(b) \right].$$

Pour calculer $A(\frac{h}{2})$ il suffit de rajouter les termes manquants :

$$A(\frac{h}{2}) = \frac{1}{2} A(h) + \frac{h}{2} \left[f(a + \frac{h}{2}) + f(a + 3\frac{h}{2}) + \dots + f(b - \frac{h}{2}) \right].$$

Combinons cette astuce avec l'extrapolation de Richardson :

Algorithme 2 méthode de Romberg R_n^m

Entrée: une fonction $f: [a, b] \rightarrow \mathbb{R}$

Sortie: une valeur approchée de $\int_a^b f(x) dx$.

$$h \leftarrow \frac{b-a}{n}, \quad s_0 \leftarrow h \left[\frac{1}{2} f(a) + \sum_{k=1}^{n-1} f(a+kh) + \frac{1}{2} f(b) \right]$$

pour j **de** 1 **à** m **faire**

$$h \leftarrow h/2, \quad s_j \leftarrow \frac{1}{2} s_{j-1} + h \sum_{k=0}^{n-1} f(a + (2k+1)h), \quad n \leftarrow 2n$$

pour k **de** 0 **à** $j-1$ **faire** $s_k \leftarrow \frac{4^{j-k} s_{k+1} - s_k}{4^{j-k} - 1}$ **fin pour**

fin pour

retourner s_0

La méthode de Romberg : majoration d'erreur

Proposition

La méthode de Romberg $R_n^m : C([a, b]) \rightarrow \mathbb{R}$ est d'ordre $2m + 1$.

Pour $h = \frac{a-b}{n}$ on obtient donc la majoration d'erreur

$$\left| \int_a^b f(x) dx - R_n^m(f) \right| \leq 2(b-a) \cdot \left(\frac{h}{2}\right)^{2m+2} \cdot \max_{[a,b]} \left| \frac{f^{(2m+2)}}{(2m+2)!} \right|$$

pour toute fonction $f : [a, b] \rightarrow \mathbb{R}$ de classe C^{2m+2} .

Démonstration.

Soit f un polynôme de degré $\leq 2m$, de sorte que $f^{(2m+1)} = 0$

Le reste de la formule d'Euler-Maclaurin s'annule, donc

$$A^0(h) = \int_a^b f(x) dx + a_2 h^2 + a_4 h^4 + \dots + a_{2m} h^{2m}.$$

L'élimination successive donne finalement $A^m(h) = \int_a^b f(x) dx$.

C'est le résultat exact, la méthode de Romberg est donc d'ordre $2m$.

L'argument de symétrie montre que la méthode de Romberg est encore exacte pour $f(x) = x^{2m+1}$. Elle est donc d'ordre $2m + 1$. \square

Comparaison entre Romberg et Newton-Cotes

Sur chaque intervalle $[a + (k - 1)h, a + kh]$ la méthode de Romberg R_n^m est de la forme $h \sum_{j=0}^{2^m} w_j f(\xi_j)$ à points équirépartis.

Quant aux poids w_j , on trouve après un petit calcul :

$$\begin{array}{rcccccc} m = 0 : & \frac{1}{2} & & & & \frac{1}{2} \\ m = 1 : & \frac{1}{6} & & \frac{4}{6} & & \frac{1}{6} \\ m = 2 : & \frac{7}{90} & \frac{32}{90} & \frac{12}{90} & \frac{32}{90} & \frac{7}{90} \end{array}$$

Pour $m = 1$ on reconnaît la méthode de Simpson.

Pour $m = 2$ les poids sont ceux de la méthode de Boole.

Pour $m \geq 3$ la méthode de Romberg ne correspond plus à une méthode de Newton-Cotes.

Exercice. — Vérifier ces poids dans les cas $m = 1$ et $m = 2$.

Calculer les poids pour $m = 3$ et déterminer l'ordre de cette méthode.

Comparer avec la méthode de Newton-Cotes à 9 points équirépartis.

Rapport « qualité-prix » de l'intégration numérique

Pour les méthodes d'intégration numérique, l'évaluation de f en un point est (presque toujours) l'opération la plus coûteuse.

- La méthode du point au milieu nécessite n évaluations.
La méthode des trapèzes nécessite $n + 1$ évaluations.
Les deux méthodes sont ordre 1, donc d'erreur $\sim \frac{1}{n^2}$.
- La méthode de Simpson nécessite $2n + 1$ évaluations.
Elle est d'ordre 3, donc d'erreur $\sim \frac{1}{n^4}$.
- La méthode de Boole nécessite $4n + 1$ évaluations.
Elle est d'ordre 5, donc d'erreur $\sim \frac{1}{n^6}$.

La méthode de Romberg R_n^m prolonge cette famille.
Elle est assez performante, aisée à programmer,
et souvent préférée à toute autre dans la pratique.

À noter toutefois qu'elle nécessite $2^m n + 1$ évaluations.
Le paramètre m sera donc à choisir avec modération.

On utilisera donc plutôt un paramètre m petit, notamment si f n'a pas la régularité suffisante ou si $\max |f^{(m)}|$ croît trop rapidement avec m .

Pour conclure : synopsis des différentes approches

Comment calculer $\int_a^b f(x)dx$? Plus on sait sur f mieux c'est !

- Si f est un polynôme, $f(x) = a_0 + a_1x + \dots + a_nx^n$ dont on connaît les coefficients a_0, a_1, \dots, a_n , alors l'intégration se fait par la primitive évidente $F(x) = a_0x + \frac{a_1}{2}x^2 + \dots + \frac{a_n}{n+1}x^{n+1}$.
- Si f est une fraction rationnelle, $f(x) = \frac{p(x)}{q(x)}$, une primitive est moins évidente mais peut se calculer par une décomposition en fractions simples, comme discutée précédemment.
- Si f se développe en une série entière $f(x) = \sum_{k=0}^{\infty} a_k(x - x_0)^k$, alors on en déduit une primitive $F(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1}(x - x_0)^{k+1}$. Les séries se prêtent bien au calcul, comme déjà discuté.

L'intégration numérique sert surtout quand les traitements exacts ci-dessus sont impossibles ou trop coûteux.

- Dans notre développement nous avons seulement supposé que l'on sait évaluer $f(x)$ pour $x \in [a, b]$ à un coût raisonnable.
- Souvent les points $x = a + kh$ suffisent ($h = \frac{b-a}{n}$, $k = 0, \dots, n$). Parfois on ne dispose que de cet échantillon (« sampling »).
- Plus on sait, mieux c'est : sous l'hypothèse de régularité C^{d+1} nous avons montré comment profiter des méthodes d'ordre d .