Kapitel 8

Numerische Integration und Differentiation

8.1 Die Formeln von Newton-Cotes

Sei $f \in C[a,b]$. Wir wollen das Integral $I = \int\limits_a^b f(x)dx$ numerisch berechnen und dabei nur Auswertungen von f benutzen. Solche linearen Integrationsformeln haben die allgemeine Form:

$$I \simeq \sum_{k=0}^{n} A_k f(x_k)$$

mit $x_k \in [a.b]$, $A_k \in \mathbb{R}^1$. Die x_k heißen Stützstellen, die von f unabhängigen A_k heißen Gewichte.

Wir betrachten die geschlossenen Newton-Cotes Formeln, d.h. die Randpunkte des Intervalls sind Stützstellen:

$$\begin{aligned} x_k &= a + k \cdot h \qquad , \qquad h = \tfrac{b-a}{n} \qquad k = 0, \ldots, n \\ I_n &= \sum_{k=0}^{n} A_k f(x_k) \end{aligned}$$

Eine Möglichkeit, die Gewichte A_k zu bestimmen, ist, die Funktion f durch ihr Interpolationspolynom p zu ersetzen, das an ihrer Stelle integriert wird:

In der Form von Lagrange lautet p

$$p(x) = \sum_{k=0}^{n} f(x_k)\omega_k(x) \quad , \quad \omega_k(x) = \prod_{\substack{\ell=0 \\ \ell \neq k}}^{n} \frac{x - x_\ell}{x_k - x_\ell} \; .$$

Es folgt

$$\begin{aligned} I_n &= \int_a^b p(x)dx = \sum_{k=0}^{n} \int_a^b \omega_k(x)dx f(x_k) \; , \\ A_k &= \int_a^b \omega_k(x)dx = \int_a^b \prod_{\substack{\ell=0 \\ \ell \neq k}}^{n} \frac{x - x_\ell}{x_k - x_\ell} dx \; . \end{aligned}$$

Wir substituieren $x = ht + a$ und erhalten

$$A_k = h \int_0^n \prod_{\substack{\ell=0 \\ \ell \neq k}}^{n} \frac{t - \ell}{k - \ell} dt = h a_k \; .$$

Für $n = 1$ erhalten wir damit die "Trapezregel"

$$\begin{aligned} a_0 &= \int_0^1 \frac{t-1}{-1} dt = \frac{1}{2} \, , \quad a_1 = \int_0^1 t dt = \frac{1}{2} \\ I_1 &= \frac{h}{2}(f(a) + f(b)) = \frac{b-a}{2}(f(a) + f(b)) \; . \end{aligned}$$

Das Integral wird also durch die Fläche eines Trapezes angenähert. Für $n = 2$ ergibt sich die trotz ihrer Einfachheit relativ genaue Simpson'sche Regel:

$$\begin{aligned} a_0 = \frac{1}{3} \quad &, \quad a_1 = \frac{4}{3} \quad , \quad a_2 = \frac{1}{3} \\ I_2 = \frac{h}{3} &\left(f(a) + 4f\left(\frac{b+a}{2}\right) + f(b) \right) \; . \end{aligned}$$

Die folgende Tabelle enthält die Koeffizienten a_j für $n \leq 4$:

n	a_0	a_1	a_2	a_3	a_4		Bezeichnung
1	1	1				$\cdot\ \frac{1}{2}$	Trapezregel
2	1	4	1			$\cdot\ \frac{1}{3}$	Simpson-Regel
3	1	3	3	1		$\cdot\ \frac{3}{8}$	Newton'sche $\frac{3}{8}$ - Regel
4	7	32	12	32	7	$\cdot\ \frac{2}{45}$	Milne - Regel

Für $n \geq 8$ können negative Gewichte auftreten, was aus Rundungsfehlergründen nicht gut ist. Wie wir später sehen werden, kann man Formeln höherer Genauigkeit konstruieren, indem man die oben angegebenen Regeln auf Teilintervalle anwendet.

Beispiel:

$$\begin{aligned} I &= \int_0^1 e^x dx = e - 1 &&= 1.7183 \\ I_1 &= \tfrac{1}{2}(1 + e) &&= 1.8591 \\ I_2 &= \tfrac{1}{6}(1 + 4e^{1/2} + e) &&= 1.7189 \\ I_3 &= \tfrac{1}{8}(1 + 3e^{1/3} + 3e^{2/3} + e) &&= 1.7185 \end{aligned}$$

Man sieht, daß der Übergang von I_1 nach I_2 einen großen Gewinn an Genauigkeit ergibt.

Der folgende Satz gibt eine Abschätzung für den Fehler $|I - I_n|$:

Satz 8.1.1 *i) Sei $f \in C^{n+1}[a,b]$. Dann gilt*

$$|I - I_n| \leq h^{n+2} c_n \max_{[a,b]} |f^{(n+1)}(x)| \, ,$$

$$c_n = \frac{1}{(n+1)!} \int_0^n \prod_{k=0}^n |t - k| dt \, .$$

ii) Sei n gerade und $f \in C^{n+2}[a,b]$. Dann gilt

$$|I - I_n| \leq h^{n+3} c_n^* \max_{[a,b]} |f^{(n+2)}(x)| \, , \quad c_n^* = \frac{n}{2} c_n \, .$$

Bemerkung: *Bei geradem n gewinnt man durch den Übergang zu $n+1$ keine Potenz von h. Die Potenzen von h sind optimal, die Konstanten c_n, c_n^* könnten verbessert werden.*

Beweis:

i) Es gilt

$$I - I_n = \int_a^b (f-p)(x)dx$$

und nach Satz V.2.1

$$(f-p)(x) = \frac{w(x)}{(n+1)!}\, f^{(n+1)}(\xi)$$

mit $w(x) = \prod_{k=0}^n (x - x_k)$. Also folgt

$$|I - I_n| \leq \frac{1}{(n+1)!} \int_a^b |w(x)|dx \; \max_{[a,b]} |f^{(n+1)}(x)| \, .$$

c_n ergibt sich aus der Berechnung von $\int_a^b |w(x)|dx$:

$$\begin{aligned} \int_a^b w(x)dx &= \int_a^b \prod_{k=0}^n |x - x_k|dx \\ &= h^{n+2} \int_0^n \prod_{k=0}^n |t-k|dt \, . \end{aligned}$$

ii) Für gerades n ist w ungerade bezüglich der Intervallmitte $c = \frac{a+b}{2}$, es gilt also

$$\int_a^b w(x)dx = 0 \, .$$

Damit hat man

$$\begin{aligned}
\int_a^b (f-p)(x)dx &= \frac{1}{(n+1)!}\int_a^b w(x)f^{(n+1)}(\xi)dx \\
&= \frac{1}{(n+1)!}\int_a^b w(x)\left\{f^{(n+1)}(c)+(\xi-c)f^{(n+2)}(\eta)\right\}dx \\
&= \frac{1}{(n+1)!}\int_a^b w(x)(\xi-c)f^{(n+2)}(\eta)dx
\end{aligned}$$

Wegen $|\xi - c| \le \frac{b-a}{2} = \frac{nh}{2}$ gilt

$$\begin{aligned}
\left|\int_a^b (f-p)(x)dx\right| &\le \frac{1}{(n+1)!}\int_a^b |w(x)|dx \cdot \frac{nh}{2}\max_{[a,b]}|f^{(n+2)}(x)| \\
&= h^{n+2}c_n\frac{nh}{2}\max_{[a,b]}|f^{(n+2)}(x)| \\
&= h^{n+2}c_n^*\max_{[a,b]}|f^{(n+2)}(x)|\ .
\end{aligned}$$

Da das Maximum hoher Ableitungen von f sehr schwer zu bestimmen ist, sind diese Formeln zur praktischen Abschätzung des Fehlers unbrauchbar. Ihr Nutzen liegt in der Information, mit welcher Potenz von h der Fehler abfällt und daß er vom Maximum einer höheren Ableitung abhängt.

Wir konstruieren nun Formeln höherer Genauigkeit. Gegeben sei wieder eine äquidistante Unterteilung $x_k = a + kh$, $h = \frac{b-a}{n}$, $k = 0, \ldots, n$.

Wir integrieren stückweise mit der Trapezregel:

$$\int_{x_k}^{x_{k+1}} f(x)dx \simeq \frac{h}{2}(fx_k) + f(x_{k+1}))\ ,$$

$$I = \int_{x_0}^{x_n} f(x)dx = \sum_{k=0}^{n-1}\int_{x_k}^{x_{k+1}} f(x)dx$$

$$
\begin{aligned}
&\simeq \frac{h}{2}\sum_{k=0}^{n-1}(f(x_k)+f(x_{k+1})) \\
&= \frac{h}{2}(f_0+2f_1+2f_2+\ldots+f_{n-1}+f_n)= \\
&= T_h
\end{aligned}
$$

mit der Abkürzung $f_k := f(x_k)$.

Diese Formel heißt zusammengesetzte Trapezregel. Für den Fehler gilt:

$$|I-T_h| = \sum_{k=0}^{n-1}\left|\int_{x_k}^{x_{k+1}} f(x)dx - \frac{h}{2}(f(x_k)+f(x_{k+1}))\right|$$

Nach Satz 8.1 gilt mit $c_1 = \frac{1}{2}\int_0^1 t(1-t)dt = \frac{1}{12}$:

$$
\begin{aligned}
|I-T_h| &\leq \frac{1}{12}\sum_{k=0}^{n-1} h^3 \max_{[x_k,x_{k+1}]}|f''(x)| \\
&\leq \frac{n}{12}\, h^3 \max_{[a,b]}|f''(x)| \\
&= \frac{b-a}{12}h^2 \max_{[a,b]}|f''(x)|\,.
\end{aligned}
$$

Für gerades n bilden wir nun analog die zusammengesetzte Simpson-Regel, indem wir die Simpson-Regel auf jeweils zwei aufeinanderfolgende Teilintervalle anwenden und anschließend summieren:

$$
\begin{aligned}
I = \int_a^b f(x)dx &= \sum_{k=0}^{\frac{n}{2}-1}\int_{x_{2k}}^{x_{2k+2}} f(x)dx \\
&\simeq \frac{h}{3}(f_0+4f_1+f_2+f_2+4f_3+f_4+\ldots) \\
&= \frac{h}{3}(f_0+4f_1+2f_2+4f_3+2f_4+\ldots+2f_{n-2}+4f_n) \\
&= S_h
\end{aligned}
$$

8.2 Das Romberg-Verfahren

Das Romberg-Verfahren beruht auf der Trapezregel. Durch Berechnen von Integrationsformeln mit verschiedener Schrittweite h lassen sich Formeln konstruieren, deren Fehler mit einer hohen Potenz von h abfällt.

Für das Integral

$$I = \int_a^b f(x)dx$$

ergibt die Trapezregel:

$$T_1(h) = \frac{h}{2}(f_0 + 2f_1 + \ldots + 2f_{n-1} + f_n)$$

mit

$$f_i := f(x_i)\ , \quad x_i = a + ih\ , \quad i = 0, \ldots, n\ , \quad h = \frac{b-ab}{n}\ .$$

Wir entwickeln nun $T_1(h)$ nach Potenzen von h.

Satz 8.2.1 *Sei $f \in C^{2m+2}[a,b]$. Dann gilt*

$$T_1(h) = I + c_1h^2 + c_2h^4 + \ldots + c_mh^{2m} + O(h^{2m+2})$$

mit

$$c_k(-1)^{k+1}\frac{B_{2k}}{(2k)!}(f^{(2k-1)}(b) - f^{(2k-1)}(a))$$

und den Bernoulli'schen Zahlen B_k:

$$B_1 = \frac{1}{6} \quad , \quad B_2 = \frac{1}{30} \quad , \quad B_3 = \frac{1}{42}, \ldots$$

Folgerung: $f \in C^{2m+2}(\mathbb{R}^1)$ 2π-periodisch. Dann gilt

$$\int_a^{a+2\pi} f(x)dx = T_1(h) + \mathrm{O}(h^{2m+2})\ .$$

Wenn f sogar $\in C^\infty(\mathbb{R}^1)$, so gilt

$$\int_a^{a+2\pi} f(x)dx - T_1(h) \to 0 \quad \text{schneller als jede Potenz von } h\ .$$

Periodische Funktionen lassen sich also über die volle Periode außerordentlich gut mit der Trapezregel integrieren. Diese vereinfacht sich in diesem Fall zu

$$T_1(h) = h \sum_{j=0}^{n-1} f_j \ .$$

Wir konstruieren nun Formeln hoher Genauigkeit für beliebige Funktionen. Wir bilden dazu

$$\begin{aligned} T_1(h) &= I + c_1 h^2 + c_2 h^4 + \ldots + c_m h^{2m} + \ \mathrm{O}(h^{2m+2}) \ , \\ T_1(\frac{h}{2}) &= I + c_1 2^{-2} h^2 + c_2 2^{-4} h^4 + \ldots + c_m 2^{-2m} h^{2m} + \ \mathrm{O}(h^{2m+2}) \ . \end{aligned}$$

Damit ist

$$2^2 T_1(\frac{h}{2}) - T_1(h) = (2^2-1)I + c_2(2^{-2}-1)h^4 + \ldots + c_m(2^{2-2m}-1)h^{2m} + \ \mathrm{O}(h^{2m+2}) \ ,$$

$$I = \frac{1}{2^2-1}(2^2 T_1(\frac{h}{2}) - T_1(h)) - c_2 \frac{2^{-2}-1}{2^2-1} h^4 - \ldots - c_m \frac{2^{2-2m}-1}{2^2-1} h^{2m} + \ \mathrm{O}(h^{2m+2}) \ .$$

Mit

$$T_2(h) := \frac{1}{2^2-1}(2^2 T_1(\frac{h}{2}) - T_1(h))$$

erhalten wir also

$$I = T_2(h) + c_2' h^4 + \ldots + c_m' h^{2m} + \ \mathrm{O}(h^{2m+2}) \ .$$

Durch Wiederholen dieses Verfahrens erhalten wir Formeln höherer Ordnung: Sei $T_k(h)$ eine Formel der Ordnung h^{2k}, d.h.

$$\begin{aligned} T_k(h) &= I + c_k h^{2k} + c_{k+1} h^{2k+2} + \ldots + c_m h^{2m} + \ \mathrm{O}(h^{2m+2}) \ , \\ T_k(\frac{h}{2}) &= I + c_k 2^{-2k} h^{2k} + \ldots + c_m 2^{-2m} h^2 2m \quad + \ \mathrm{O}(h^{2m+2}) \ . \end{aligned}$$

Damit wird

$$\begin{aligned} & 2^{2k} T_k(\frac{h}{2}) - T_k(h) = \\ & \quad = (2^{2k}-1)I + c_{k+1}(2^{-2}-1)h^{2k+2} + \ldots + c_m(2^{2k-2m})h^{2m} + \ \mathrm{O}(h^{2m+2}) \ . \end{aligned}$$

Mit
$$T_{k+1}(h) := \frac{1}{2^{2k}-1}(2^{2k}T_k(\frac{h}{2}) - T_k(h))$$
gilt also
$$I = T_{k+1}(h) + \mathrm{O}(h^{2k+2}) .$$
Diese Konstruktion von Formeln höherer Ordnung läßt sich im sogenannten Romberg-Schema darstellen:

$$\begin{array}{cccccccc}
h & T_1(h) & & & & & & \\
 & & \diagdown & & & & & \\
\frac{h}{2} & T_1(\frac{h}{2}) & \text{—} & T_2(h) & & & & \\
 & & \diagdown & & \diagdown & & & \\
\frac{h}{4} & T_1(\frac{h}{4}) & \text{—} & T_2(\frac{h}{2}) & \text{—} & T_3(h) & & \\
 & & \diagdown & & \diagdown & & \diagdown & \\
\frac{h}{8} & T_1(\frac{h}{8}) & \text{—} & T_2(\frac{h}{4}) & \text{—} & T_3(\frac{h}{2}) & \text{—} & T_4(h)
\end{array}$$

Die Rekursion schreibt sich am besten in der Form
$$T_{k+1}(h) = T_k(\frac{h}{2}) + \frac{1}{2^{2k}-1}(T_k(\frac{h}{2}) - T_k(h)) .$$
Da $T_k(\frac{h}{2})$ und $T_k(h)$ für große k und kleines h jeweils gute Näherungen für I sind, tritt beim Bilden der Differenz Auslöschung auf. Es lohnt im allgemeinen nicht, über $k = 6$ hinauszugehen.

Bei der Berechnung von $T_1(\frac{h}{2})$ wird man die schon in $T_1(h)$ benötigten Funktionswerte mitbenutzen: Mit $f_{i+1/2} = f(x_i + \frac{h}{2})$ ist
$$\begin{aligned}
T_1(\frac{h}{2}) &= \frac{h}{4}(f_0 + 2f_{1/2} + 2f_1 + \ldots + 2f_{n-1/2} + f_n) \\
&= \frac{h}{4}(f_0 + 2f_1 + 2f_2 + \ldots + 2f_{n-1} + f_n) \\
&\quad + \frac{h}{2}(f_{1/2} + f_{3/2} + \ldots + f_{n-1/2}) \\
&= \frac{1}{2}T_1(h) + \frac{h}{2}(f_{1/2} + f_{3/2} + \ldots + f_{n-1/2}) .
\end{aligned}$$

Der erste Teil ist bereits bekannt. Nur der zweite Teil muß noch berechnet werden.

Beispiel: $I = \int\limits_0^1 e^x dx = 1.718281828$

h	T_1	T_2	T_3	T_4
1	1.859140914			
$\frac{1}{2}$	1.753931092	1.718861151		
$\frac{1}{4}$	1.727221904	1.718318841	1.718282687	
$\frac{1}{8}$	1.720518592	1.718284155	1.718281842	1.718281829

Die Romberg-Integration wird vor allem in Programmen zur automatischen Integration benutzt. Sie sind etwa folgendermaßen aufgebaut:

Eingabe: Gewünschte relative Genauigkeit ε, Intervallgrenzen a, b, ein Unterprogramm zur Auswertung von f, maximale Anzahl n der Auswertung von f.

Ausgabe: Eine Näherung $\tilde{I}$ für das Integral I mit $|I - \tilde{I}| \leq \varepsilon |I|$, Zuverlässigkeitsindex.

Verfahren:

1) Man berechnet etwa 4-6 Spalten des Romberg-Schemas für $h = b - a$, $(b-a)/2, \ldots$

2) Prüfung der Genauigkeit in Spalte k. Es gilt

$$\begin{aligned} T_k(h) &= I + c_k h^{2k} + \mathrm{O}(h^{2k+2}) \,, \\ T_k(\frac{h}{2}) &= I + c_k 2^{-2k} h^{2k} + \mathrm{O}(h^{2k+2}) \,. \end{aligned}$$

Also folgt

$$T_k(\frac{h}{2}) - T_k(h) = c_k(2^{-2k} - 1)h^{2k} + \mathrm{O}(h^{2k+2}) \,,$$

$$c_k h^{2k} = \frac{1}{2^{-2k} - 1}(T_k(\frac{h}{2}) - T_k(h)) + \mathrm{O}(h^{2k+2})$$

$$= \qquad \varepsilon_k(\frac{h}{2}) + \ O(h^{2k+2}) \ .$$

Man prüft, ob $\varepsilon_k(\frac{h}{4} \sim 2^{-2k}\varepsilon_k(\frac{h}{2})$. Falls die Abweichung unter 10% liegt, wird ε_k als Fehlerschätzung akzeptiert.

3) Man führt das Romberg-Schema so lange fort, bis in der Spalte ganz rechts $|\varepsilon_k(\frac{h}{2})| \leq \varepsilon |T_k(\frac{h}{2}|$ gilt.

Wir müssen noch Satz 1 beweisen. Dazu einige Vorbereitungen. Seien B_k die Bernoulli-Polynome, d.h.

$$\begin{aligned} B_0 &= 1 \ , \\ B_k' &= B_{k-1} \ , \quad \int_0^1 B_k dx = 0 \ , \quad k = 1, 2, \dots \ . \end{aligned}$$

Z.B. ist $B_1 = x - 1/2$, $B_2 = \frac{1}{2}x^2 - \frac{1}{2}x + 1/12$. Die $B_k = k! B_k(0)$ heißen Bernoulli-Zahlen.

Wir leiten einige Eigenschaften B_k her.

1) Für $k \geq 2$ ist $B_k(0) = B_k(1)$. Dies ist klar wegen

$$B_k(1) - B_k(0) = \int_0^1 B_k' dx = \int_0^1 B_{k-1} dx = 0 \quad \text{für} k \geq 2 \ .$$

2) Für $k \geq 1$ ist $B_{2k+1}(1) = B_{2k+1} = 0$. Dazu setzen wir $P_k(x) = (-1)^k B_k(1-x)$. Diese erfüllen die gleiche Rekursion wir die B_k. Also ist $P_k = B_k$, woraus die Behauptung folgt.

Satz 8.2.2 *(Euler'sche Summenformel) Sei $g \in C^{2m+2}[0,n]$, n ganz. Dann gilt*

$$\begin{aligned} & g(0) + g(1) + \cdots + g(n-1) + \frac{1}{2}g(n) - \int_0^n g dx \\ & \quad = -\int_0^n \overline{B}_{2m+2} g^{(2m+2)} dx + \sum_{i=0}^{m} B_{2k+2}(0) \left(g^{(2n+1)}(n) - g^{(2k+1)}(0) \right) \ . \end{aligned}$$

Dabei sind $\overline{B}_k$ die Funktionen in $\mathbb{R}^1$, welche aus B_k durch periodische Fortsetzung aus $[0,1]$ mit der Periode 1 entstehen.

Beweis: Die Eigenschaften der B_k führen zu folgenden Eigenschaften der $\overline{B}_k$: $\overline{B}_k$ ist stetig für $k \geq 2$, $\overline{B}_{2k+1}(0) = \overline{B}_{2k+1}(n) = 0$, $\overline{B}_k = \overline{B}_{k+1}$ für $k \geq 1$. Der Beweis beruht auf zwei verschiedenen Auswertungen von

$$\int_0^n \overline{B}_1 g' dx \; .$$

Einerseits haben wir

$$\begin{aligned} \int_0^n \overline{B}_1 g' dx &= \sum_{i=0}^{n-1} \int_i^{i+1} \overline{B}_1 g' dx = \sum_{i=0}^{n-1} \left\{ \overline{B}_1 g|_i^{i+1} - \int_i^{i+1} \overline{B}_1' g dx \right\} \\ &= \sum_{i=0}^{n-1} \left\{ \frac{g(i+1)+g(i)}{2} - \int_i^{i+1} g dx \right\} \\ &= \frac{1}{2} g(0) + g(1) + \cdots + g(n-1) + \frac{1}{2} g(n) - \int_0^n g dx \; . \end{aligned}$$

Andererseits gilt

$$\begin{aligned} \int_0^n \overline{B}_1 g' dx &= \int_0^n \overline{B}_2' g dx = -\int_0^n \overline{B}_2 g'' dx + \overline{B}_2 g' \Big|_0^n \\ &= -\int_0^n \overline{B}_3' g'' dx + \overline{B}_2 g' \Big|_0^n \\ &= -\int_0^n \overline{B}_3 g''' dx - \overline{B}_3 g'' \Big|_0^n + \overline{B}_2 g'|_0^n \; . \end{aligned}$$

So fortfahrend erhält man schließlich

$$\begin{aligned} \int_0^n \overline{B}_1 g' dx &= -\int_0^n \overline{B}_{2m+2} g^{(2m+2)} dx \\ &+ \left[-\overline{B}_{2m+2} g^{(2m+1)} - \overline{B}_{2m+1} g^{(2m)} + \cdots \overline{B}_2 g' \right]_0^n \; . \end{aligned}$$

Aus dem Vergleich der beiden Darstellungen folgt die Behauptung.

□

Satz 1 folgt nun durch lineare Transformation des Intervalls $[0, b]$ auf $[0, n]$.

8.3 Integration nach Gauß

Wir suchen eine Integrationsformel für das Integral

$$If = \int_a^b w(x)f(x)dx$$

mit einer in (a, b) streng positiven Gewichtsfunktion w. Eine Integrationsformel der Form

$$G_n f = \sum_{j=1}^{n} A_j f(x_j)$$

hat die $2n$ freien Parameter A_j und x_j. Die Formeln von Newton-Cotes integrieren ein Polynom n-ten Grades exakt. Wir wollen nun fordern, daß

$$G_n f = If \quad \text{für} \quad f \in \mathcal{P}_{2n-1} \; .$$

Dies ergibt gerade $2n$ Bedingungen für die $2n$ Parameter. Der folgende Satz zeigt, daß diese Forderung maximal ist:

Satz 8.3.1 *Es gibt keine Formel G_n, die in $\mathcal{P}_{2n}$ exakt ist.*

Beweis: Wir nehmen an, G_n sei eine solche Formel, also $G_n f = If$ für alle $f \in \mathcal{P}_{2n}$. Dies wäre dann auch richtig für das Polynom

$$f(x) = \prod_{j=1}^{n} (x - x_j)^2 \; .$$

Offenbar ist aber $G_n f = 0$, $If \neq 0$.

□

Zur Konstruktion einer in $\mathcal{P}_{2n-1}$ exakten Formel G_n benutzen wir die nach dem Schmidt'schen Verfahren konstruierten orthonormalen Polynome p_n:

$$\int_a^b w p_n p_m dx = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

Wir wissen: p_n hat in (a, b) genau n einfache Nullstellen.

Satz 8.3.2 *Es gibt Formeln G_n, welche auf $\mathcal{P}_{2n-1}$, exakt sind. Die x_j sind die Nullstellen von p_n und es gilt*

$$A_j = \int_a^b w(x) \prod_{\substack{i=1 \\ i \neq j}} \left(\frac{x - x_i}{x_j - x_i} \right)^2 dx \,.$$

Beweis: Sei G_n die Newton-Cotes Formel zu den Nullstellen $x_1, \ldots, x_n$ von p_n, also

$$G_n f = \sum_{j=1}^{n} \int_a^b w \prod_{\substack{i=1 \\ i \neq j}} \frac{x - x_i}{x_j - x_i} dx f(x_j) \,.$$

G_n ist offenbar exakt in $\mathcal{P}_{n-1}$, da ja gerade das Interpolationspolynom vom Grad $n - 1$ integriert wird. Ist $f \in \mathcal{P}_{2n-1}$, so schreiben wir

$$f = q p_n + r \quad \text{mit} \quad q, r \in \mathcal{P}_{n-1} \qquad \text{(Division mit Rest)} \,.$$

Dann ist

$$\int_a^b w f dx = \int_a^b w q p_n dx + \int_a^b w r dx \,.$$

Das erste Integral verschwindet wegen der Orthogonalitätseigenschaften der p_n. Das zweite ist $G_n r$, weil G_n ja auf $\mathcal{P}_{n-1}$ exakt ist. Also folgt

$$\int_a^b w f dx = G_n r = G_n(r + q p_n) = G_n f \,,$$

weil p_n an den Stützstellen von G_n verschwindet.

Also ist G_n exakt in $\mathcal{P}_{2n-1}$ und wir müssen nur noch die Formel für die Gewichte bestätigen. Mit

$$w_j = \prod_{\substack{i=1 \\ i \neq j}} \frac{x - x_i}{x_j - x_i}$$

ist $w_j^2 \in \mathcal{P}_{2n-2}$, also

$$\int_a^b w w_j^2 dx = G_n(w_j^2) = \sum_{k=1}^{n} A_k w_j^2(x_k) = A_j \ .$$

$\square$

Beispiel: $[a, b] = [-1, +1], \ w = 1.$

Die x_j sind die Nullstellen der Legendre-Polynome p_n.

n	x_1	x_2	x_3	A_1	A_2	A_3
1	0			2		
2	$-\sqrt{\frac{1}{3}}$	$+\sqrt{13}$		1	1	
3	$-\sqrt{35}$	0	$\sqrt{35}$	$\frac{5}{9}$	$\frac{8}{9}$	$\frac{5}{9}$

$$I = \int_{-1}^{1} e^x dx = 2.350402$$

Die Simpson-Regel liefert:

$$I_2 = 2.362054$$

Dagegen ist mit gleich vielen Funktionswertungen

$$G_3 = 2.350337$$

8.4 Numerische Differentiation

Sei $f \in C^{p+1}(\mathbb{R}^1)$. Wir interessieren uns für eine Näherung für $f^{(k)}(0)$, die mit Hilfe der Werte $f(x_j)$ an den Stützstellen $x_j = jh$ berechnet wird. Der Satz von Taylor liefert

$$f(ih) = \sum_{\ell=0}^{p} f^{(\ell)}(0)\frac{(ih)^\ell}{\ell!} + \mathrm{O}(h^{p+1}) \,.$$

Wir bilden die Linearkombination

$$\sum_{i=-q}^{q} \alpha_i f(ih) = \sum_{\ell=0}^{p} f^{(\ell)}(0)\frac{1}{\ell!}h^\ell \sum_{i=-q}^{q} i^\ell \alpha_i + \mathrm{O}(h^{p+1})$$

und wählen die α_i so, daß

$$\sum_{i=-q}^{q} i^\ell \alpha_i = \begin{cases} 1 & , \quad \ell = k < p \\ 0 & , \quad \text{sonst} \end{cases} \quad , \qquad \ell = 0, \ldots, p \,.$$

Wir erhalten

$$\frac{1}{k!}h^k f^{(k)}(0) = \sum_{i=-q}^{q} \alpha_i f(ih) + \mathrm{O}(h^{p+1}) \,.$$

Für $2q+1 = p+1$ führt das Gleichungssystem für die α_i auf eine Vandermonde-Matrix, ist also eindeutig lösbar. Für $2q+1 > p+1$ setzt man einige $\alpha_i = 0$, bis man wieder eine Vandermonde-Matrix erhält.

Wir haben also mit

$$D_h^{(k)} f = \frac{k!}{h^k} \sum_{i=-q}^{q} \alpha_i f(ih) = f^{(k)}(0) + \mathrm{O}(h^{p+1-k})$$

eine Differentiationsformel der Ordnung h^{p+1-k}.

Beispiele: $k = 1$:

$$p = 1 \,,\; q = 1 \,,\; \alpha_{-1} = 0 \,,\; \alpha_0 = -1 \,,\; \alpha_1 = +1 \,,$$

$$D_h^{(1)} f = \frac{f(h) - f(0)}{h} \,, \quad f'(0) = D_h^{(1)} f + \mathrm{O}(h) \quad \text{für} \quad f \in C^2 \,.$$

$$p = 2\ ,\ q = 1\ ,\ \alpha_{-1} = -\frac{1}{2}\ ,\ \alpha_0 = 0\ ,\ \alpha_1 = \frac{1}{2}\ ,$$

$$D_h^{(1)} f = \frac{f(h) - f(-h)}{2h}\ , \quad f'(0) = D_h^{(1)} f + \mathrm{O}(h^2) \quad \text{für} \quad f \in C^3\ .$$

$k = 2$:

$$p = 3\ ,\ q = 1\ ,\ \alpha_{-1} = \frac{1}{2}\ ,\ \alpha_0 = -1\ ,\ \alpha_1 = \frac{1}{2}\ ,$$

$$D_h^{(2)} f = \frac{f(h) - 2f(0) + f(-h)}{h^2}\ ,\ f''(0) = D_h^{(2)} f + \mathrm{O}(h^2) \quad \text{falls} \ \ f \in C^4\ .$$

8.5 Der Fehler bei Integration und Differentiation

Sei $If = \int\limits_a^b f(x)dx$ zu berechnen. Steht anstelle von f nur eine Näherung $\tilde{f}$ mit relativem Fehler ε (also $|(f - \tilde{f})| \leq \varepsilon|f(x)|$) zur Verfügung, so kann nur die Näherung $I\tilde{f}$ für If berechnet werden. Es gilt

$$|If - \tilde{I}f| \leq \int_a^b |(f - \tilde{f})(x)|dx \leq \varepsilon I|f|\ . \tag{5.1}$$

Falls $I|f|$, $|If|$ die gleiche Größenordnung haben (z.B. falls $f \geq 0$), so haben $|If - \tilde{I}f|$, $\varepsilon|If|$ die gleiche Größenordnung, also $\tilde{I}f$ einen relativen Fehler der Größenordnung ε. Die Integration ist in diesem Fall also eine gut konditionierte Aufgabe.

Ist aber $I|f|$ viel größer als $|If|$ (z.B. wenn f eine stark oszillierende Funktion ist), dann ist der relative Fehler von $I\tilde{f}$ viel größer als ε. Die Integration ist dann schlecht konditioniert.

Wir untersuchen nun, ob die Berechnung von f durch eine Integrationsformel der Ordnung p

$$I_h f = \sum_{j=1}^{n} A_j f(x_j)\ , \quad |I_h f - If| \leq ch^p$$

ein gutartiger Algorithmus ist. Es gilt

$$\begin{aligned} |I_h\tilde{f} - If| &\leq |I_h(\tilde{f} - f| + |(I_h - I)f| \\ &\leq \varepsilon \sum_{j=1}^{n} |A_j||f(x_j)| + ch^p \ . \end{aligned}$$

Sind nun alle Gewichte A_j positiv, so ist

$$\sum_{j=1}^{n} |A_j||f(x_j)| = \sum_{j=1}^{n} A_j|f(x_j)| = I_h|f| \sim I|f|$$

und damit näherungsweise

$$|I_h\tilde{f} - If| \leq \varepsilon I|f| + ch^p \ . \tag{5.2}$$

Vergleich mit (5.1) zeigt, daß hier $\varepsilon I|f|$ der unvermeidliche, ch^p der Algorithmusfehler ist. (Dabei haben wir den bei der Bildung der j-Summe entstehenden Rundungsfehler nicht berücksichtigt; er ist praktisch ohne jede Bedeutung). Ist also die Schrittweite h hinreichend klein, etwa

$$ch^p \leq \varepsilon I|f| \ , \tag{5.3}$$

so ist der Algorithmus gutartig. Dies ist nicht der Fall bei negativen Gewichten, denn dann kann

$$\varepsilon \sum |A_j||f(x_j)| \gg \varepsilon I|f|$$

sein.

Nun zur Differentiation! Im Prinzip können $\tilde{f}^{(k)}(0)$ (falls dies überhaupt Sinn hat) und $f^{(k)}(0)$ beliebig verschieden sein. Wenden wir trotzdem eine Differentiationsformel der Ordnung $p + 1 - k$

$$D_h^{((k)} f = \frac{k!}{h^k} \sum_{i=-q}^{q} \alpha_i f(ih) \ , \quad |D_h^{(k)} f - f^{(k)}(0)| \leq ch^{p+1-k}$$

an, so wird

$$\begin{aligned} |D_h^{(k)}\tilde{f} - f^{(k)}(0) &\leq |D_h^{(k)}(\tilde{f} - f)| + |D_h^{(k)} f - f^{(k)}(0)| \\ &\leq \varepsilon h^{-k} a + ch^{p+1-k} \ , \qquad (5.4) \\ a &= k! \sum_{i=-q}^{q} |\alpha_i||f(ih)| \ . \end{aligned}$$

Hier kann man nun h nicht gegen Null gehen lassen, weil dabei der Fehler über alle Grenzen wächst. Es gibt hier ein optimales $h = h_0$, bei welchem der Fehler minimal wird. Größenordnungsmäßig kann man h_0 aus der Bedingung ("balancing terms")

$$\varepsilon h^{-k} a = c h^{p+1-k}$$

bestimmen zu $h_0 = \mathrm{O}(\varepsilon^{1/(1+p)})$. Für $h = h_0$ ist dann

$$D_h^{(k)} \tilde{f} - f^{(k)}(0) = \mathrm{O}(\varepsilon^{1-k/(p+1)}) \ .$$

Ein Fehler ε in f führt also - bei einer Formel der Ordnung p und optimaler Wahl von h - zu einem Fehler ε^α, $\alpha = 1 - \frac{k}{p+1} < 1$ in $f^{(k)}$. Dies bedeutet:

1) Numerische Differentiation führt zu einem Genauigkeitsverlust.

2) Dieser Verlust ist klein (d.h. $\alpha \sim 1$) falls, $p + 1 \gg k$. Insbesondere hängt er von der Ordnung der verwendeten Differentiationsformel ab.

Beispiel: Berechnung von $f'(0)$ $(k = 1)$

Formel	p	h_0	ε
$\frac{1}{h}(f(h) - f(0))$	1	$\varepsilon^{1/2}$	$\varepsilon^{1/2}$
$\frac{1}{2h}(f(h) - (-h))$	2	$\varepsilon^{1/3}$	$\varepsilon^{2/3}$
	4	$\varepsilon^{1/5}$	$\varepsilon^{4/5}$
	6	$\varepsilon^{1/7}$	$\varepsilon^{7/8}$