Journal of Mathematical Analysis
ISSN: 2217-3412, URL: http://www.ilirias.com
Volume 5 Issue 4(2014), Pages 1-11.

THE METHOD OF UNDETERMINED COEFFICIENTS: GENERAL APPROCH AND OPTIMAL ERROR BOUNDS

FRANÇOIS DUBEAU

Abstract. We revisit and present a unified framework for the method of undetermined coefficients for numerical integration, numerical differentiation, and pointwise estimation of a function. Optimal error bounds are given which depends on the regularity of the function. Corrected integration and differentiation rules are also discussed.

1. Introduction

Many numerical analysis textbooks present the method of undetermined coefficients to obtain numerical approximation processes for the integral or the derivative of a function (see for example [3, 4, 6, 7, 10]).

The numerical integration processes obtained include not only midpoint, trapezoidal, Simpson, Newton-Cotes rules but also Gaussian quadrature formulas. Numerical differentiation formulas that can be analyzed by this method also include differentiation methods obtained by the Richardson's extrapolation process. Moreover it can be used to solve the pointwise estimation problem, approximating a function at a given point where the function cannot be evaluated directly.

In this paper we present a general framework for the method of undetermined coefficients. In particular it is applied to the three problems: numerical integration, pointwise estimation, and numerical differentiation but could be applied to any process which satisfies the assumptions given along the paper. The analysis is based on Taylor's expansions for absolutely continuous functions to obtain Peano's kernels, and leads to optimal error bounds which depend on the regularity of the function.

For the numerical integration we will consider the following expression

$$QI_k(f;h) = h^{k-1}\int_{-h}^{h} f^{(k)}(x)dx \tag{1.1}$$

2000 *Mathematics Subject Classification.* 65D25, 65D30, 65D32, 41A55, 26A42, 26A46.

Key words and phrases. Undetermined coefficients; numerical integration; pointwise estimation; numerical differentiation; Taylor's expansion; Peano Kernel, corrected integration rules, corrected differentiation rules.

Submitted August 27, 2014. Published October 21, 2014.

F. Dubeau was supported by the Natural Sciences and Engineering Research Council of Canada.

even if for $k \geq 1$ we have

$$QI_k(f;h) = h^{k-1}\int_{-h}^{h} f^{(k)}(x)dx = h^{k-1}\left[f^{(k-1)}(h) - f^{(k-1)}(-h)\right].$$

For example this expression with $k = 2$ will be useful for endpoint corrected integration rules. For pointwise estimation ($k = 0$), and numerical differentiation ($k \geq 1$), we will consider the expression

$$QD_k(f;h) = h^k f^{(k)}(0). \tag{1.2}$$

In the sequel, to simplify we will consider the expression

$$Q_k(f;h) = \begin{cases} QI_k(f;h), \\ \\ QD_k(f;h). \end{cases} \tag{1.3}$$

The method of undetermined coefficients consists in finding a $(n+1)$-dimensional weight vector $\vec{a}(k) = (a_0(k), \ldots, a_n(k))$ associated to a given $(n+1)$-dimensional vector of distinct coordinates (or nodes) $\vec{x} = (x_0, \ldots, x_n)$ such that $Q_k(f;h)$ is approximated by

$$Q^d_{n,k}(f;h) = \sum_{i=0}^{n} a_i(k) f(hx_i) = \begin{cases} QI^d_{n,k}(f;h), \\ \\ QD^d_{n,k}(f;h). \end{cases} \tag{1.4}$$

(where $x_i \neq 0$ in the case of the pointwise estimation problem because $f(0)$ is not supposed to be accessible).

The truncation error of the process is

$$R_{n,k}(f;h) = Q_k(f;h) - Q^d_{n,k}(f;h), \tag{1.5}$$

more specifically

$$R_{n,k}(f;h) = \begin{cases} RI_{n,k}(f;h) & = & QI_k(f;h) - QI^d_{n,k}(f;h), \\ \\ RD_{n,k}(f;h) & = & QD_k(f;h) - QD^d_{n,k}(f;h). \end{cases} \tag{1.6}$$

and the method is based on the requirement that

$$R_{n,k}(f;h) = o(h^{r(n)}), \tag{1.7}$$

where $r(n) \geq n$, and which also depends on the regularity of $f(x)$.

The plan of the paper is the following. In the next section we present preliminaries about polynomials, Vandermonde matrix, and Taylor's expansions. The method of undetermined coefficients is the object of Section 3. In Section 4 we obtain optimal error bounds using Taylor's expansions and Peano's kernels. Total error including not only the truncation error but also the roundoff error is discussed in Section 5. Corrected integration rules are discussed in Section 6. Finally a way to increase the order of differentiation rules is indicated in Section 7.

We will use $f^{(l)}(x)$ for the l-th derivative of $f(x)$ for $l = 0, 1, 2, \ldots$, where $f^{(0)}(x) = f(x)$. Let $1 \leq p \leq \infty$, if $f(x)$ is defined on a set E, $\|f\|_{p,E}$ will be its p-norm on E, and if $\vec{v}$ is a vector in $\mathbb{R}^n$, its p-norm will be $\|\vec{v}\|_p$.

2. Preliminaries

2.1. **Small o and big O notations.** Let $f(x)$ be a function such that $\lim_{x\to\alpha} f(x) = 0$. We say that $g(x)$ is a small o of $f(x)$ around α, and write $g(x) = o(f(x))$, if for any $\epsilon > 0$ there exists a $\delta_\epsilon > 0$ such that

$$|g(x)| \leq \epsilon \, |f(x)| \,. \tag{2.1}$$

holds for $0 < |x-\alpha| < \delta_\epsilon$. We say that $g(x)$ is a big O of $f(x)$ around α, and write $g(x) = O(f(x))$, if there exist a constant C and a $\delta > 0$ such that

$$|g(x)| \leq C \, |f(x)| \,. \tag{2.2}$$

holds for $0 < |x-\alpha| < \delta$.

2.2. **Polynomials and small o notation.** The next lemma is a direct consequence of the small o notation for polynomials and will be useful to obtain necessary conditions in section 4.

Lemma 2.1. *Let r be a positive real number, and $n = \lfloor r \rfloor$. Let $\pi_m(x)$ be a polynomial of degree m such that*

$$\pi_m(x) = o(|x-\alpha|^r).$$

Then,

$$\pi_m(x) = \begin{cases} (x-\alpha)^n \pi_{m-n}(x) & \text{if} \quad m > r, \\ \\ 0 & \text{if} \quad m \leq r, \end{cases}$$

where $\pi_{m-n}(x)$ is a polynomial of degree $m-n$.

2.3. **Vandermonde matrix.** The Vandermonde matrix associated to $\vec{x}$ is given by

$$V(\vec{x}) = \begin{bmatrix} 1 & \dots & 1 \\ x_0 & & x_n \\ \vdots & & \vdots \\ x_0^n & \dots & x_n^n \end{bmatrix}.$$

Let $\vec{e}_l$ be the $(n+1)$-column vector, the transpose of $(\delta_{l,0}, \cdots, \delta_{l,j} \cdots, \delta_{l,n})$, where

$$\delta_{l,j} = \begin{cases} 1 & \text{if} \quad j = l, \\ \\ 0 & \text{if} \quad j \neq l, \end{cases}$$

for $0 \leq l, j \leq n$.

Lemma 2.2. [9] *If the $n+1$ coordinates of $\vec{x}$ are distinct, the Vandermonde matrix $V(\vec{x})$ is invertible and the l-th column of $V^{-1}(\vec{x})$ is*

$$V^{-1}(\vec{x})\vec{e}_l = \frac{1}{l!} \begin{bmatrix} w_{n,0}^{(l)}(0) \\ w_{n,1}^{(l)}(0) \\ \vdots \\ w_{n,n}^{(l)}(0) \end{bmatrix}, \tag{2.3}$$

for $l = 0, \dots, n$, where $\{w_{n,j}(x)\}_{j=0}^{n}$ is the Lagrange's basis of the space of polynomial of degree at most n given by

$$w_{n,j}(x) = \prod_{\substack{i=0 \\ i \neq j}}^{n} \frac{(x - x_i)}{(x_j - x_i)} \quad \textit{for} \quad j = 0, \dots, n.$$

2.4. **Taylor's expansion.** Let $I_H = [-H, H]$, for $H = 1$ we will simply use $I = [-1, 1]$, and $I_H^+ = [0, H]$ and $I_H^- = [-H, 0]$. Let $C^l(I_H)$ be the set of continuously differentiable functions up to order l on I_H, let $p \in [1, +\infty]$, and let the set of absolutely continuous function on I_h be defined by

$$AC^{l+1,p}(I_H) = \left\{ f \in C^l(I_H) \;\middle|\; \begin{array}{l} (a)\ f^{(l+1)} \in L^p(I_H), \quad and \\ (b)\ f^{(l)}(s) = f^{(l)}(r) + \int_r^s f^{(l+1)}(\xi) d\xi,\ \forall r, s \in I_H \end{array} \right\}.$$

Taylor's expansion of $f(x) \in AC^{l+1,p}(I_H)$ around $x = 0$ of order $l + 1$ is

$$f(x) = \sum_{j=0}^{l} \frac{f^{(j)}(0)}{j!} x^j + \int_{-H}^{H} f^{(l+1)}(y) K_{T,l}(x, y; H) dy, \tag{2.4}$$

where $K_{T,l}(x, y; H)$ is the kernel

$$K_{T,l}(x, y; H) = \frac{1}{l!} \left[(x - y)_+^l \mathbf{1}_{[0,H]}(y) + (-1)^{l+1} (y - x)_+^l \mathbf{1}_{[-H,0]}(y) \right], \tag{2.5}$$

for any x, y in I_H (see [2, 11]). This kernel is a piecewise polynomial function of degree l. In this expression, if E is a set, then

$$\mathbf{1}_E(y) = \begin{cases} 1 & \text{if} \quad y \in E, \\ 0 & \text{if} \quad y \notin E, \end{cases}$$

and for any non negative integer l,

$$(z)_+^l = \begin{cases} z^l & \text{if} \quad z > 0, \\ 0 & \text{if} \quad z \leq 0. \end{cases}$$

If we set $H = h\tau$, $x = h\xi$, and $y = h\eta$, then the kernel becomes

$$K_{T,l}(x, y; H) = K_{T,l}(h\xi, h\eta; h\tau) = h^l K_{T,l}(\xi, \eta; \tau), \tag{2.6}$$

for any ξ, η in I_τ.

3. Method of undetermined coefficients

Let τ be any fixed real number such that $\tau \geq \max\{1, \max\{|x_i| \mid i = 0, \dots, n\}\} \geq 1$, and set $H = \tau h$. Using h and H with $H > h$ allows the possibility to have $|x_i| > 1$, which means that hx_i is outside of $[-h, h]$. In particular, for numerical integration, it includes integration rules with nodes outside the interval of integration.

The method of undetermined coefficients is based on the following properties of $Q_k(f; h)$ (or both $QI_k(f; h)$ and $QD_k(f; h)$):

(P.1) For any k there exists $\underline{n}_k$ such that $l \geq \underline{n}_k$, the expression $Q_k(f;h)$ is well defined for $f(x) \in C^l(I_H)$;

(P.2) $Q_k(f;h)$ is linear with respect to $f(x)$;

(P.3) If $f(x)$ is a polynomial of degree $\leq m$ with respect to x, then $Q_k(f;h)$ is a polynomial of degree $\leq m$ with respect to h.

(P.4) If $f(x) = x^l$ then $Q_k(x^l;h) = h^l Q_k(x^l;1)$.

From (P.2) and (P.3), we observe that $R_{n,k}(f;h)$ is linear with respect to $f(x)$, and if $f(x)$ is a polynomial of degree $\leq m$ with respect to x, then $R_{n,k}(f;h)$ is a polynomial of degree $\leq m$ with respect to h. From Lemma 2.1, the condition (1.7) implies that $R_{n,k}(f;h) = 0$ for any polynomial $f(x)$ of degree $\leq n$. From (P.4), using the standard basis $\left\{x^l\right\}_{l=0}^{n}$, we have to solve the linear system

$$\sum_{i=0}^{n} a_i \left(x_i h\right)^l = Q_k(x^l;h) \quad \text{for} \quad l = 0, \ldots, n, \tag{3.1}$$

which has the solution

$$\vec{a}(k) = V^{-1}(\vec{x})\vec{q}(k),$$

where the $n+1$ components of $\vec{q}(k)$ are

$$q_l(k) = \frac{Q_k(x^l;h)}{h^l} = Q_k(x^l;1) \quad \text{for} \quad l = 0, \ldots, n.$$

It might happen that $R_{n,k}(f;h) = 0$ for some polynomials $f(x)$ of degree $> n$, hence let us define the degree of accuracy (or precision) $\overline{n}_k$ of the approximation process (1.4) as the largest integer $\overline{n}_k \geq n$ such that $R_{n,k}(f;h) = 0$ holds for any polynomial $f(x)$ of degree $l \leq \overline{n}_k$.

4. Truncation error

For the error analysis, we add the following property on $Q_k(f;h)$.

(P.5) If $K(x,y) = g(y)P(x,y)$ where $g(y) \in L^p(I_H)$, and $P(x,y)$ is piecewise polynomial function in terms of x and y, we have

$$Q_k\left(\int_{-H}^{H} K(\cdot,y)dy;h\right) = \int_{-H}^{H} Q_k(K(\cdot,y);h)dy. \tag{4.1}$$

For any $\underline{n}_k \leq l \leq \overline{n}_k$, let $f(x) \in AC^{l+1,p}(I_H)$. From (P.1)-(P.5), taking the Taylor's expansion (2.4) of $f(x)$ of order $l+1$, and using the fact that the process is exact for polynomials of degree $\leq l$, we obtain

$$R_{n,k}(f;h) = \int_{-H}^{H} f^{(l+1)}(y) K_{Q_k,l}(y;H,h)dy, \tag{4.2}$$

where $K_{Q_k,l}(y;H,h)$ is the Peano kernel associated to the process $Q_k(f;h)$, given by

$$K_{Q_k,l}(y;H,h) = R_{n,k}(K_{T,l}(\cdot,y;H);h). \tag{4.3}$$

Let the number $q \in [1, +\infty]$ be the conjugate of p, $\frac{1}{p} + \frac{1}{q} = 1$, it follows from the Holder's inequality that

$$|R_{n,k}(f;h)| \leq \left\| f^{(l+1)} \right\|_{p,I_H} \left\| K_{Q_k,l}(\cdot;H,h) \right\|_{q,I_H}. \tag{4.4}$$

because $K_{Q_k,l}(y;H,h) \in L^{\infty}(I_H) \subseteq L^q(I_H)$ for any $1 \leq q \leq \infty$.

We also need the following property of the kernel.

(P.6) If $H = h\tau$ and $y = h\eta$, then $K_{Q_k,l}(y;H,h) = h^l K_{Q_k,l}(\eta;\tau,1)$.

It follows that

$$\left\| K_{Q_k,l}(\cdot;H,h) \right\|_{q,I_H} = h^{l+1-\frac{1}{p}} \left\| K_{Q_k,l}(\cdot;\tau,1) \right\|_{q,I_\tau}, \tag{4.5}$$

and so

$$|R_{n,k}(f;h)| \leq h^{l+1-\frac{1}{p}} C_{l,p}(k) \left\| f^{(l+1)} \right\|_{p,I_H}, \tag{4.6}$$

where

$$C_{l,p}(k) = \left\| K_{Q_k,l}(\cdot;\tau,1) \right\|_{q,I_\tau}, \tag{4.7}$$

does not depend on h. Since for $H = h\tau$ we have

$$\lim_{h \to 0} \left\| f^{(l+1)} \right\|_{p,I_H} = \begin{cases} 0 & \text{for} \quad 1 \leq p < \infty, \\ C & \text{for} \quad p = \infty, \end{cases}$$

then this result says that

$$R_{n,k}(f;h) = \begin{cases} o\left(h^{l+1-\frac{1}{p}}\right) & \text{for} \quad 1 \leq p < \infty, \\ O\left(h^{l+1}\right) & \text{for} \quad p = \infty, \end{cases} \tag{4.8}$$

Since an $o\left(h^{l+1-\frac{1}{p}}\right)$ and an $O\left(h^{l+1}\right)$ are $o\left(h^l\right)$, it means that $R_{n,k}(f;h) = o\left(h^l\right)$. In summary we have proved the following theorem which presents a necessary and sufficient condition to obtain the desired error order.

Theorem 4.1. *Let the real number $\tau \geq 1$ be fixed and $H = h\tau$. A necessary and sufficient condition to have $R_{n,k}(f;h) = o(h^n)$ for any $f \in AC^{n+1,p}(I_H)$ is that $R_{n,k}(f;h) = 0$ for any polynomial $f(x)$ of degree $\leq n$.*

As an extension of this first result, we have the next one which shows the dependance of the error in terms of the regularity of the function.

Theorem 4.2. *Let the real number $\tau \geq 1$ be fixed and $H = h\tau$. If $R_{n,k}(f;h) = 0$ for any polynomial of degree $\leq n_k$, then (4.6), (4.7), and (4.8) hold for any $f \in AC^{l+1,p}(I_H)$, for $\underline{n}_k \leq l \leq \overline{n}_k$.*

Remark. *Let us specify the kernels $K_{Q_k,l}(y;H,h)$ and $K_{Q_k,l}(\eta;\tau,1)$ for the $Q_k(f;h)$ given by (1.3) for different values of k .*

(a) Numerical integration:

$$\begin{aligned} K_{QI_k,l}(y;H,h) &= h^{k-1}\int_{-h}^{h} K_{T,l-k}(x,y;H)dx - \sum_{i=0}^{n} a_i(k)K_{T,l}(hx_i,y;H) \\ &= \frac{h^k}{(l+1-k)!}\left[(h-y)_+^{l+1-k}\boldsymbol{1}_{I_H^+}(y) + (-1)^{l+1-k}(y-h)_+^{l+1-k}\boldsymbol{1}_{I_H^-}(y)\right] \\ &\quad -\frac{1}{l!}\sum_{i=0}^{n} a_i(k)\left[(hx_i-y)_+^l\boldsymbol{1}_{I_H^+}(y) + (-1)^{l+1}(y-hx_i)_+^l\boldsymbol{1}_{I_H^-}(y)\right], \end{aligned}$$

and

$$\begin{aligned} K_{QI_k,l}(\eta;\tau,1) &= \frac{1}{(l+1-k)!}\left[(1-\eta)^{l+1-k}\boldsymbol{1}_{I_\tau^+}(\eta) + (-1)^{l+1-k}(\eta+1)^{l+1}\boldsymbol{1}_{I_\tau^-}(\eta)\right] \\ &\quad -\frac{1}{l!}\sum_{i=0}^{n} a_i(k)\left[(x_i-\eta)_+^l\boldsymbol{1}_{I_\tau^+}(\eta) + (-1)^{l+1}(\eta-x_i)_+^l\boldsymbol{1}_{I_\tau^-}(\eta)\right]. \end{aligned}$$

(b) Numerical differentiation:

$$\begin{aligned} K_{QD_k,l}(y;H,h) &= h^k K_{T,l-k}(x,y;H) - \sum_{i=0}^{n} a_i(k)K_{T,l}(hx_i,y;H) \\ &= \frac{h^k}{(l-k)!}\left[(0-y)_+^{l-k}\boldsymbol{1}_{I_H^+}(y) + (-1)^{l-k+1}(y-0)_+^{l-k}\boldsymbol{1}_{I_H^-}(y)\right] \\ &\quad -\frac{1}{l!}\sum_{i=0}^{n} a_i(k)\left[(hx_i-y)_+^l\boldsymbol{1}_{I_H^+}(y) + (-1)^{l+1}(y-hx_i)_+^l\boldsymbol{1}_{I_H^-}(y)\right], \end{aligned}$$

and

$$K_{QD_k,l}(\eta;\tau,1) = -\frac{1}{l!}\sum_{i=0}^{n} a_i(k)\left[(x_i-\eta)_+^l\boldsymbol{1}_{I_\tau^+}(\eta) + (-1)^{l+1}(\eta-x_i)_+^l\boldsymbol{1}_{I_\tau^-}(\eta)\right].$$

Remark. *The bounds given by (4.5), (4.6) and (4.7) are the best one. To show this fact we can use a standard construction for $f^{(l+1)}(x)$ given in* [1]. *Indeed, for $1 < p \le \infty$, $1 \le q < \infty$, let $f(x) \in AC^{l+1,p}(I_H)$ be such that*

$$f^{(l+1)}(y) = \begin{cases} |K_{Q_k,l}(y;H,h)|^{q-2}K_{Q_k,l}(y;H,h) & \text{for} \quad K_{Q_k,l}(y;H,h) \neq 0, \\ 0 & \text{for} \quad K_{Q_k,l}(y;H,h) = 0, \end{cases}$$

we get an equality

$$\int_{-H}^{H} f^{(l+1)}(y)K_{Q_k,l}(y;H,h)dy = \left\|f^{(l+1)}\right\|_{p,I_H}\|K_{Q_k,l}(\cdot;H,h)\|_{q,I_H},$$

because

$$\left\|f^{(l+1)}\right\|_{p,I_H} = \|K_{Q_k,l}(\cdot;H,h)\|_{q,I_H}^{q-1},$$

and

$$\int_{-H}^{H} f^{(l+1)}(y)K_{Q_k,l}(y;H,h)dy = \|K_{Q_k,l}(\cdot;H,h)\|_{q,I_H}^{q}.$$

For $p=1$, $q=\infty$, *we suppose that* $\|K_{Q_k,l}(\cdot;H,h)\|_{\infty,I_H}>0$. *Let us consider* $0<\epsilon<\|K_{Q_k,l}(\cdot;H,h)\|_{\infty,I_H}$, *define the set* A_ϵ *by*

$$A_\epsilon=\left\{y\in I_H\,|\,|K_{Q_k,l}(y;H,h)|\geq\|K_{Q_k,l}(\cdot;H,h)\|_{\infty,I_H}-\epsilon\right\},$$

which is such that its Lebesque's measure is $0<\mu(A_\epsilon)\leq 2H$. *If we set*

$$f^{(l+1)}(y)=\begin{cases}|K_{Q_k,l}(y;H,h)|^{-1}K_{Q_k,l}(y;H,h) & \text{for} \quad y\in A_\epsilon,\\ 0 & \text{elsewhere},\end{cases}$$

then

$$\begin{aligned}\left\|f^{(l+1)}\right\|_{1,I_H}\left(\|K_{Q_k,l}(\cdot;H,h)\|_{\infty,I_H}-\epsilon\right) &\leq \int_{-H}^{H}f^{(l+1)}(y)K_{Q_k,l}(y;H,h)dy\\ &\leq \left\|f^{(l+1)}\right\|_{1,I_H}\|K_{Q_k,l}(\cdot;H,h)\|_{\infty,I_H},\end{aligned}$$

because

$$\left\|f^{(l+1)}\right\|_{1,I_H}=\mu(A_\epsilon).$$

5. Total error

We are interested by the quantity $A_k(f;h)=\frac{1}{h^k}Q_k(f;h)$ for which we have the approximation $A_n^d(\tilde{f};h)=\frac{1}{h^k}Q_n^d(\tilde{f};h)$ which uses $\tilde{f}(hx_i)$ instead of $f(hx_i)$, and hence introducing roudoff error $\epsilon_i=f(hx_i)-\tilde{f}(hx_i)$. The total error $E_{n,k}(f;h)$ is decomposed in two types of error, the truncation error $R_{n,k}(f;h)$ and the roundoff error $S_{n,k}(f;h)$, hence

$$\begin{aligned}E_{n,k}(f;h) &= A_k(f;h)-A_n^d(\tilde{f};h)\\ &= \left[\left(A_k(f;h)-A_n^d(f;h)\right)+\left(A_n^d(f;h)-A_n^d(\tilde{f};h)\right)\right]\\ &= \frac{1}{h^k}\left[\left(Q_k(f;h)-Q_n^d(f;h)\right)+\left(Q_n^d(f;h)-Q_n^d(\tilde{f};h)\right)\right]\\ &= \frac{1}{h^k}\Big[R_{n,k}(f;h)+S_{n,k}(f;h)\Big].\end{aligned}$$

For the truncation error we already have

$$|R_{n,k}(f;h)|\leq h^{l+1-\frac{1}{p}}C_{l,p}\left\|f^{(l+1)}\right\|_{p,I_H}.$$

For the roundoff error

$$S_{n,k}(f;h)=Q_n^d(f;h)-Q_n^d(\tilde{f};h)=\sum_{i=0}^{n}a_i(k)(f(hx_i)-\tilde{f}(hx_i))=\sum_{i=0}^{n}a_i(k)\epsilon_i,$$

then

$$|S_{n,k}(f;h)|=\left|\sum_{i=0}^{n}a_i(k)\epsilon_i\right|\leq\sum_{i=0}^{n}|a_i(k)|\,|\epsilon_i|\leq\|\vec{a}(k)\|_q\,\|\vec{\epsilon}\|_p,$$

where $\|\vec{a}(k)\|_q$ is independant of h.

Consequently, the total error is bounded by

$$|E_{n,k}(f;h)|\leq h^{l+1-\frac{1}{p}-k}C_{l,p}\left\|f^{(l+1)}\right\|_{p,I_H}+\frac{\|\vec{a}(k)\|_q\,\|\vec{\epsilon}\|_p}{h^k},$$

for $l = \underline{n}_k, \ldots, \overline{n}_k$ and $f \in AC^{l+1,p}(I_H)$. This expression shows that the process is numerically stable for $k = 0$, and unstable for $k \geq 1$.

Let us observe that a complete analysis of the roundoff error should take into account the numerical evaluation of the scalar product $\sum_{i=0}^{n} a_i(k) f(hx_i)$; see [5].

6. Corrected integration rules

In this section we consider only the integration process and $Q_k(f;h) = QI_k(f;h)$ and $\overline{n}_k$ will be noted $\overline{n}_k^I$. Corrected integration rules, see for example [4, 8, 12], use endpoint values of the first derivative to improve the quadrature formula, which is useful for the composite quadrature rules since the endpoint corrections cancelled in such a way that only two derivatives are needed for the composite rules. The corrected quadrature rules can be written on the form

$$\frac{1}{h}\int_{-h}^{h} f(x)dx \approx \sum_{i=0}^{n} \tilde{a}_i(\beta) f(hx_i) + \beta h \left[f^{(1)}(h) - f^{(1)}(-h) \right]. \qquad (6.1)$$

The truncation error of this process $R_n^c(f,h)$ can be decomposed into two truncation errors as follows. Let us rewrite the coefficients $\tilde{a}_i(\beta)$ as the difference $a_i^I(0) - \beta a_i^I(2)$, then we have

$$QI_0(f,h) = QI_{n,0}^d(f,h) + \beta \left[QI_2(f,h) - QI_{n,2}^d(f,h) \right],$$

and

$$RI_n^c(f,h) = RI_{n,0}(f,h) - \beta RI_{n,2}(f,h).$$

It follows, by linearity of the method of undetermined coefficients, that the degree of accuracy of this corrected process, noted $n(\beta)$, is such that

$$n(\beta) \geq \min \left\{ \overline{n}_0^I, \overline{n}_2^I \right\}.$$

Obviously $n(0) = \overline{n}_0^I$. The parameter β can be used to increase the degree of accuracy of the formula as follows:

- if $\overline{n}_0^I \geq \overline{n}_2^I$: let
$$\beta_* = \frac{RI_{n,0}(x^{\overline{n}_2^I+1};h)}{RI_{n,2}(x^{\overline{n}_2^I+1};h)},$$
  - for $\overline{n}_0^I > \overline{n}_2^I$: $\beta_* = 0$, $n(\beta_*) = \overline{n}_0^I$, and $n(\beta) = \overline{n}_2^I$ for any $\beta \neq \beta_* = 0$, and no improvement of the degree of accuracy is possible;
  - for $\overline{n}_0^I = \overline{n}_2^I$: $\beta_* \neq 0$, and $n(\beta_*) > \overline{n}_0^I = \overline{n}_2^I$;
- if $\overline{n}_0^I < \overline{n}_2^I$ then $n(\beta) = \overline{n}_0^I < \overline{n}_2^I$ for any $\beta \in \mathbb{R}$, and no improvement of the degree of accuracy is possible is possible.

We can also choose the parameter β to minimize the best error bound with respect to a given norm. Indeed, because the kernel of the process, noted $K_{\beta,l}^c(y;H)$, is the linear combination of two kernels

$$K_{\beta,l}^c(y;H,h) = K_{QI_0,l}(y;H,h) - \beta K_{QI_2,l}(y;H,h)$$

and the best constant for a given β is

$$C_{l,p}^c(\beta) = \left\| K_{\beta,l}^c(\cdot;\tau,1) \right\|_{q,I_\tau} = \left\| K_{QI_0,l}(\cdot;\tau,1) - \beta K_{QI_2,l}(\cdot;\tau,1) \right\|_{q,I_\tau}, \qquad (6.2)$$

this last expression can be minimized with respect to β to obtain

$$\beta_{l,p}^* = \arg\min_{\beta \in \mathbb{R}} C_{l,p}^c(\beta).$$

7. Corrected differentiation rules

In this section we consider only the differentiation process and $Q_k(f;h) = QD_k(f;h)$ and $\overline{n}_k$ will be noted $\overline{n}_k^D$. Let us assume we have a primitive of $F(x)$ of $f(x)$, that is to say $F^{(1)}(x) = f(x)$. Corrected differentiation rules use end-point values of the primitive to improve the differentiation formula. The corrected differentiation rules can be written on the form

$$h^k f^{(k)}(0) \approx \sum_{i=0}^{n} \tilde{a}_i(\beta) f(hx_i) + \frac{\beta}{h}\left[F(h) - F(-h)\right]. \tag{7.1}$$

The truncation error of this process $RD_n^c(f,h)$ can be decomposed into two truncation errors as follows. Let us rewrite the coefficients $\tilde{a}_i(\beta)$ as the difference $a_i^D(k) - \beta a_i^I(0)$, then we have

$$QD_k(f,h) = QD_{n,k}^d(f,h) + \beta\left[QI_0(f,h) - QI_{n,0}^d(f,h)\right],$$

and

$$RD_n^c(f,h) = RD_{n,k}(f,h) - \beta RI_{n,0}(f,h).$$

It follows, by linearity of the method of undetermined coefficients, that the degree of accuracy of this corrected process, noted $n(\beta)$, is such that

$$n(\beta) \geq \min\left\{\overline{n}_k^D, \overline{n}_0^I\right\}.$$

Obviously $n(0) = \overline{n}_k^D$. The parameter β can be used to increase the degree of accuracy of the formula as follows:

- if $\overline{n}_k^D \geq \overline{n}_0^I$: let
$$\beta_* = \frac{RD_{n,0}(x^{\overline{n}_0^I+1};h)}{RI_{n,0}(x^{\overline{n}_0^I+1};h)},$$
  - for $\overline{n}_k^D > \overline{n}_0^I$: $\beta_* = 0$, $n(\beta_*) = \overline{n}_k^D$, and $n(\beta) = \overline{n}_0^I$ for any $\beta \neq \beta_* = 0$, and no improvement of the degree of accuracy is possible;
  - for $\overline{n}_k^D = \overline{n}_0^I$: $\beta_* \neq 0$, and $n(\beta_*) > \overline{n}_k^D = \overline{n}_0^I$;
- if $\overline{n}_k^D < \overline{n}_0^I$ then $n(\beta) = \overline{n}_k^D < \overline{n}_0^I$ for any $\beta \in \mathbb{R}$, and no improvement of the degree of accuracy is possible is possible.

We can also choose the parameter β to minimize the best error bound with respect to a given norm. Indeed, because the kernel of the process, noted $K_{\beta,l}^c(y;H)$, is the linear combination of two kernels

$$K_{\beta,l}^c(y;H,h) = K_{QD_k,l}(y;H,h) - \beta K_{QI_0,l}(y;H,h)$$

and the best constant for a given β is

$$C_{l,p}^c(\beta) = \left\|K_{\beta,l}^c(\cdot;\tau,1)\right\|_{q,I_\tau} = \left\|K_{QD_k,l}(\cdot;\tau,1) - \beta K_{QI_0,l}(\cdot;\tau,1)\right\|_{q,I_\tau}, \tag{7.2}$$

this last expression can be minimized with respect to β to obtain

$$\beta_{l,p}^* = \arg\min_{\beta\in\mathbb{R}} C_{l,p}^c(\beta).$$

References

[1] R. A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.
[2] E. Asplund and L. Bungart, *A first course in integration*, Holt, Rinehart and Winston, New York, 1966.
[3] P. J. Davis, *Interpolation and Approximation*, Dover, New York, 1975.
[4] P. J. Davis and P. Rabinowitz, *Method of Numerical integration*, 2nd edition, Academic Press, Florida, 1984.
[5] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd Ed., SIAM, Philadelphia, PN, 2002.
[6] A. S. Househoulder, *Principles of Numerical Analysis*, McGraw Hill, Columbus, New York, 1953.
[7] D. Kincaid and W. Cheney, *Numerical Analysis*, Brooks/Cole Pub. Co., California, 1991.
[8] Z. Liu, Error estimate for some composite corrected quadrature rules, *Applied Mathematics Letters*, 22 (2009), 771-775.
[9] N. Macon and A. Spitzbart, Inverses of Vandermonde matrices, *The American Mathematical Monthly*, 65 (1958), 95-100.
[10] A. Ralston, *A first course in numerical analysis*, McGraw-Hill, New York, 1965.
[11] L. L. Schumaker, *Spline functions basic theory*, Wiley, New York, 1981.
[12] E. Talvila and M. Wiersma, Optimal error estimates for corrected trapezoidal rules, *Journal of Mathematical Inequalities*, 3 (2012), 431-445.

F. Dubeau
Département de mathématiques, Université de Sherbrooke, Sherbrooke (Qc), Canada, J1K2R1
E-mail address: `francois.dubeau@usherbrooke.ca`