

Numerical differentiation by means of Legendre polynomials in the presence of square summable noise

S. Lu, V. Naumova, S. Pereverzyev

RICAM-Report 2012-15

Numerical differentiation by means of Legendre polynomials in the presence of square summable noise

Shuai Lu, Valeriya Naumova and Sergei V. Pereverzyev

Abstract

We consider one of the classical ill-posed problems: estimating the derivative f' of a function $f \in C^1[-1, 1]$ from its noisy version f^δ , which is not assumed to be differentiable. In this paper we present and analyze an efficient method for the reconstruction of f' by the derivatives of the partial sums of Fourier-Legendre series of $f^\delta \in L_2$. We argue that in certain relevant cases this method has advantage over the standard approach, when a derivative is reconstructed as the solution of the corresponding ill-posed Volterra equation. One of the novelties of the paper is that the reconstruction of the derivative is performed not only in the observation space L_2 , but also in the space of continuous functions C . We discuss the accuracy of the proposed method in L_2 and C and provide guideline for the adaptive choice of the regularization parameter by means of the balancing principle. The numerical experiments confirm the robustness of the method for the numerical differentiation problem, and demonstrate its competitiveness with respect to other techniques proposed recently in the literature.

1 Introduction

One of the classic ill-posed problems consists in estimating the derivative of a function $f \in C^1[-1, 1]$ from its noisy version f^δ , which is not assumed to be differentiable. If f^δ admits the evaluation at any point $x \in [-1, 1]$, then one may think of it as $f^\delta \in C[-1, 1]$ and apply a variety of properly regularized numerical differentiation techniques (see, e.g. [1, 2] and references therein). But suppose that even f^δ can only be observed in discretized or binned form.

To be more precise, we are given only a vector of noisy Fourier coefficients

$$f_k^\delta = \langle f^\delta, \varphi_k \rangle := \int_{-1}^1 f^\delta(u) \varphi_k(u) du, \quad k = 0, 1, \dots, N, \quad (1)$$

with respect to some L_2 -orthonormal system $\{\varphi_k\}_{k=0}^\infty$, which is sometimes called design. Then a natural assumption is that $f^\delta \in L_2(-1, 1)$, and

$$\|f - f^\delta\|_{L_2(-1,1)} \leq \delta, \quad (2)$$

where $\delta \in (0, 1)$ is a small number used for measuring the noise level.

Note that the assumption (2) is standard for the theory of the regularization [3], where one is sometimes advised to find $f'(t)$ from the ill-posed operator equation

$$Ax(t) := \int_{-1}^t x(u) du = f^\delta(t) - f^\delta(-1), \quad (3)$$

or from its discretization by means of (1). At the same time, as it follows from Example 5 of [4], even if one uses the native design $\{\varphi_k\}$ of the integration operator A consisting of its singular functions, the derivatives of some simple and analytic functions, such as $f(t) = at + b$, for example, can not be reconstructed from (3) by means of the standard regularization methods with L_2 -accuracy better than $O(\delta^{\frac{1}{3}})$, that is far from to be the best order under the assumption (2).

This observation can be explained by the fact that the native design of the integration operator A is formed by the trigonometric functions that vanish at the point $t = 1$ and can not approximate well the functions not having such feature.

Thus, it is not always reasonable to transform the problem of numerical differentiation into the equation (3), especially when f is expected to be infinitely differentiable function. It hints at the use of a design $\{\varphi_k\}$ that is more universal than the singular system of the integration operator. One of the first candidates for such a design is the classical system of Legendre polynomials studied in the present paper. Note that the use of Legendre polynomials in numerical differentiation has its origin in the paper [5]. Here we consider the approximation of $f'(x)$ by the derivatives of the partial sums of Fourier-Legendre series of $f^\delta \in L_2(-1, 1)$

$$S_n f^\delta(x) := \sum_{k=1}^n f_k^\delta P_k(x),$$

where $f_k^\delta = \langle f^\delta, P_k \rangle$,

$$P_k(x) = \sqrt{k+1/2} (2^k k!)^{-1} \frac{d^k}{dx^k} [(x^2 - 1)^k],$$

and $n \in \{1, 2, \dots, N\}$ plays the role of the regularization parameter.

One of the novelties of the paper is that the reconstruction of f' is considered not only in the observation space $L_2 = L_2(-1, 1)$, but also in the space of continuous functions $C = C[-1, 1]$. The latter reconstruction problem can be seen as an extension of another classical ill-posed problem, namely the recovery of a continuous function from its noisy Fourier coefficients [6]. This time, from the same data we recover a continuous derivative of the considered function.

In this paper the smoothness of the function f to be differentiated is measured by decay of its Fourier-Legendre coefficients $f_k = \langle f, P_k \rangle$, $k = 0, 1, \dots$. More precisely, in the spirit of [7] we assume that

$$f \in W_2^\psi := \{g : g \in L_2(-1, 1), \|g\|_\psi^2 := \sum_{k=0}^{\infty} \psi^2(k) |\langle g, P_k \rangle|^2 < \infty\},$$

where $\psi : [0, \infty) \rightarrow (0, \infty)$ is called the smoothness index function and supposed to be non-decreasing continuous, with $\lim_{t \rightarrow \infty} \psi(t) = \infty$. This assumption allows us to treat simultaneously the cases of finite and infinite smoothness, which correspond to ψ increasing with polynomial and exponential rate.

It is clear that the choice of the regularization parameter n depends crucially on the smoothness of the underlying function f , which is coded in the index function ψ and rarely known in advance. Therefore, the numerical differentiation scheme

$$D_n f^\delta = \frac{d}{dx} (S_n f^\delta) \tag{4}$$

should be equipped with a parameter choice rule $n = n(\delta; f^\delta)$ for automatically adapting to unknown ψ .

In the paper we construct such a rule on the base of the balancing principle (deterministic oracle inequality) developed during the last decade in the series of papers (see, e.g. [8, 9, 10] and references therein). This principle allows a parameter choice $n = n_+ = n_+(\delta; f^\delta)$ guaranteeing an accuracy, which only by a constant factor worse than the best one that could be obtained provided ψ was known.

The outline of the paper is as follows. In the next section we introduce our approach to numerical differentiation of a function in the presence of square summable noise and we estimate its accuracy in the L_2 -sense. In

section 3, we reconstruct the derivative of a function in the space of continuous functions and prove the corresponding error estimates. Section 4 deals with an adaptive rule for choosing the regularization parameter n , which is based on the balancing principle. In last section, the results of the numerical experiments supporting the theoretical findings of previous sections are reported.

2 Numerical differentiation in L_2

From the above discussion, one can easily see that the error between $f'(x)$ and its approximation by means of the Legendre polynomials can be bounded as follows

$$\|f' - D_n f^\delta\|_{L_2} \leq \|f' - D_n f\|_{L_2} + \|D_n f - D_n f^\delta\|_{L_2}, \quad (5)$$

where the first term in the right-hand side is the approximation error, whereas the second term is the noise propagation error. In the next two lemmas we estimate these terms for $f \in W_2^\psi$ and specify our estimates for ψ increasing with polynomial or exponential rate.

We will use the following representation of the derivatives of the Legendre polynomials [11]

$$P'_k(x) = 2\sqrt{k+1/2} \sum_{i=0}^{(k-q_k-1)/2} \sqrt{2i+q_k+1/2} P_{2i+q_k}(x), \quad (6)$$

where

$$q_k = \begin{cases} 0 & \text{if } k = 2\nu + 1 \\ 1 & \text{if } k = 2\nu \end{cases} \quad \nu = 0, 1, 2, \dots$$

Moreover, we follow the convention that the symbol C denotes an absolute constant, which may not be the same at different occurrences.

Lemma 1. *For $f \in W_2^\psi$ the approximation error has the following bound, provided the integral below exists:*

$$\|f' - D_n f\|_{L_2} \leq C \left(\iint_{\Omega_n} \frac{t\tau}{\psi^2(\tau)} dt d\tau \right)^{1/2} \|f\|_\psi, \quad (7)$$

where $\Omega_n = [0, n] \times [n, \infty] \cup \{(t, \tau) : n \leq t \leq \tau < \infty\}$.

In cases $\psi(t) = t^\mu$ and $\psi(t) = e^{th}$, $h > 0$, the bound (7) reduces to the following ones respectively:

$$\|f' - D_n f\|_{L_2} \leq C n^{2-\mu} (\mu - 2)^{-1/2} \|f\|_\psi, \quad \mu > 2, \quad (8)$$

and

$$\|f' - D_n f\|_{L_2} \leq C \left(\frac{n^3}{h} + \frac{1}{h^4} \right)^{1/2} e^{-nh} \|f\|_\psi. \quad (9)$$

Proof. From (6) we obtain the representation

$$\begin{aligned} f'(x) - D_n f(x) &= f'(x) - \sum_{k=1}^n \langle f, P_k \rangle P'_k(x) = \sum_{k=n+1}^{\infty} \langle f, P_k \rangle P'_k(x) \\ &= \sum_{k=n+1}^{\infty} \langle f, P_k \rangle \sum_{i=0}^{(k-q_k-1)/2} 2\sqrt{k+1/2} \sqrt{2i+q_k+1/2} P_{2i+q_k}(x). \end{aligned} \quad (10)$$

Therefore, using the orthonormality of $\{P_{2i+q_k}(x)\}$ we can estimate the approximation error as follows

$$\begin{aligned} \|f' - D_n f\|_{L_2}^2 &\leq \sum_{j=0}^{\infty} \left(\sum_{l=\max\{n+1, j+1\}}^{\infty} 2\sqrt{j+1/2} \sqrt{l+1/2} \langle f, P_l \rangle \right)^2 \\ &\leq \sum_{j=0}^{\infty} \sum_{l=\max\{n+1, j+1\}}^{\infty} \frac{4(j+1/2)(l+1/2)}{\psi^2(l)} \sum_{m=\max\{n+1, j+1\}}^{\infty} \langle f, P_m \rangle^2 \psi^2(m) \\ &\leq 4 \left(\sum_{j=0}^{\infty} \sum_{l=\max\{n+1, j+1\}}^{\infty} \frac{(j+1/2)(l+1/2)}{\psi^2(l)} \right) \|f\|_\psi^2 \leq 4S \|f\|_\psi^2, \end{aligned}$$

where the value of S is bounded up to an absolute factor C by the integral such that

$$\begin{aligned} S &= \sum_{j=0}^n (j+1/2) \sum_{l=n+1}^{\infty} \frac{(l+1/2)}{\psi^2(l)} + \sum_{j=n+1}^{\infty} (j+1/2) \sum_{l=j+1}^{\infty} \frac{(l+1/2)}{\psi^2(l)} \\ &\leq C \left(\int_0^n t \left(\int_n^{\infty} \frac{\tau}{\psi^2(\tau)} d\tau \right) dt + \int_n^{\infty} t \left(\int_t^{\infty} \frac{\tau}{\psi^2(\tau)} d\tau \right) dt \right) \\ &= C \iint_{\Omega_n} \frac{t\tau}{\psi^2(\tau)} dt d\tau, \end{aligned}$$

that proves the bound (7).

The bounds (8), (9) are obtained by direct calculation from (7). \square

Lemma 2. *Under the assumption (2) the following bound holds true*

$$\|D_n f - D_n f^\delta\|_{L_2} \leq \frac{\delta}{2} n(n^2 + 6n + 5)^{1/2}. \quad (11)$$

Proof. Let $\xi_k = \langle f - f^\delta, P_k \rangle$, $k = 1, 2, \dots$. In view of (2) it is clear that

$$\sum_{k=1}^{\infty} \xi_k^2 \leq \delta^2.$$

Moreover, from (6) it follows that

$$\begin{aligned} D_n f(x) - D_n f^\delta(x) &= \sum_{k=1}^n \xi_k \sum_{i=0}^{(k-q_k-1)/2} 2\sqrt{k+1/2}\sqrt{2i+q_k+1/2} P_{2i+q_k}(x) \\ &= \sum_{j=0}^{n-1} P_j(x) \sum_{i=0}^{\lfloor (n-j-1)/2 \rfloor} 2\sqrt{j+1/2}\sqrt{j+2i+3/2} \xi_{j+2i+1}, \end{aligned}$$

where we use the notation $\lfloor a \rfloor = \max\{n \in \mathbb{Z} : n \leq a\}$.

$$\begin{aligned} \|D_n f - D_n f^\delta\|_{L_2}^2 &\leq 4\delta^2 \sum_{j=0}^{n-1} \sum_{i=0}^{\lfloor (n-j-1)/2 \rfloor} (j+1/2)(j+2i+3/2) \\ &\leq 2\delta^2 \sum_{j=0}^{n-1} (j+1/2)(n+j+2)(\lfloor (n-j-1)/2 \rfloor + 1) \\ &\leq \delta^2 \sum_{j=0}^{n-1} (j+1/2)(n+j+2)(n-j+1) \\ &\leq \frac{1}{4} \delta^2 n^2 (n^2 + 6n + 5) \end{aligned}$$

□

Now we will formulate the main result of this section that follows directly from Lemmas 1, 2 and (5).

Theorem 1. *Let the assumption (2) be satisfied. Assume that $f \in W_2^\psi$ with $\psi(t) = t^\mu$. Then for $\mu > 2$ and $n = C\delta^{-1/\mu}$ we have*

$$\|f' - D_n f^\delta\|_{L_2} = O(\delta^{\frac{\mu-2}{\mu}}). \quad (12)$$

If $f \in W_2^\psi$ with $\psi(t) = e^{th}$, $h > 0$, then for $n = \frac{C}{h} \log(\frac{1}{\delta})$ we obtain

$$\|f' - D_n f^\delta\|_{L_2} = O(\delta \log^2 \delta). \quad (13)$$

Remark 1. *In the presence of square summable noise (2) the accuracy of the numerical differentiation is usually quantified under the assumption that the function to be differentiated belongs to the Sobolev space H_2^r of L_2 -functions whose weak derivatives of order up to r are also in $L_2 = L_2(-1, 1)$. For example, in [12] the error bound*

$$\|f' - D_n f^\delta\|_{L_2} = O(\delta^{\frac{r-2}{r}}) \quad (14)$$

has been proven recently for $f \in H_2^r$ and $n = C\delta^{-1/r}$, $r > 2$, under the assumption (2).

Let us denote by W_2^μ the space W_2^ψ with $\psi(t) = t^\mu$. From [13, 14] (see also [8]) it follows that for $r \geq \mu$ $H_2^r \subset W_2^\mu$. Moreover, for $f \in W_2^\mu$ the derivative $f^{(\mu)}$ may not belong to $L_2(-1, 1)$, but

$$\int_{-1}^1 |f^{(\mu)}(x)|^2 (1-x)^\mu x^\mu dx < \infty.$$

Therefore, the estimation (12) may be seen as an improvement of the result (14) of [12], since the same order of the accuracy is justified by (12) for a wider space of functions.

On the other hand, it is a common belief that under the assumption (2) noisy data f^δ allow a reconstruction of the derivative f' of a function $f \in H_2^r$ with L_2 -accuracy of order $O(\delta^{\frac{r-1}{r}})$ (see, e. g. [15]). This order of accuracy can be achieved at least for functions $f \in H_2^r$ that admit a periodic continuation preserving differentiable properties.

At the same time, from (8) it follows that for $\mu > 2$ any function $f \in W_2^\mu$ belongs to H_2^1 , i. e. $W_2^\mu \subset H_2^1$ for $\mu > 2$. Moreover, differentiating the relation (10) and using (6) in the same way as in the proof of Lemma 1 one may show by induction that $W_2^\mu \subset H_2^r$ for $\mu > 2r$. Thus, for $f \in W_2^\mu$, $\mu > 2r$, one may expect that the order of the accuracy of the reconstruction of the derivative f' from noisy data f^δ is better than $O(\delta^{\frac{r-1}{r}})$. The bound (12) shows that it is really the case since for $r < \mu/2$ we have $\frac{r-1}{r} < \frac{\mu/2-1}{\mu/2} = \frac{\mu-2}{\mu}$ and $\delta^{\frac{r-1}{r}} > \delta^{\frac{\mu-2}{\mu}}$.

3 Numerical differentiation in the space of continuous functions

In this section we consider the problem of numerical differentiation when a derivative needs to be reconstructed in the space $C = C[-1, 1]$ of continuous

functions, while a function to be differentiated is observed in the L_2 -space. As we already mentioned in the Introduction, this problem can be seen as an extension of the classical ill-posed problem of reconstruction of a continuous function from its observations in L_2 [6]. In this section we, in particular, show that under the assumption (2) for a wide variety of analytic functions f their noisy data f^δ allow the recovery of f' with the accuracy of order $O(\delta \log^3 \frac{1}{\delta})$. This is in contrast to the recovery of f' from the ill-posed equation (3) by means of the standard regularization methods, where even such a simple analytic function as $f(x) = ax + b$ can not be in general reconstructed from f^δ with the accuracy better than $O(\delta^{\frac{1}{3}})$ [4].

We first note the following property of the derivatives of the Legendre polynomials [11], which will be useful below:

$$|P'_k(x)| \leq |P'_k(1)| = \frac{k(k+1)}{2} \sqrt{k+1/2}. \quad (15)$$

Lemma 3. For $f \in W_2^\psi$ the approximation error has the following bound, provided the integral below exists:

$$\|f' - D_n f\|_C \leq C \left(\int_n^\infty \frac{t^5}{\psi^2(t)} dt \right)^{1/2} \|f\|_\psi. \quad (16)$$

In cases $\psi(t) = t^\mu$ and $\psi(t) = e^{th}$, $h > 0$, from (16) we can derive the following bounds respectively:

$$\|f' - D_n f\|_C \leq C n^{3-\mu} (2\mu - 6)^{-1/2} \|f\|_\psi, \quad \mu > 3, \quad (17)$$

and

$$\|f' - D_n f\|_C \leq C \left(\frac{n^5}{h} + \frac{n^5}{h^6} \right)^{1/2} e^{-nh} \|f\|_\psi. \quad (18)$$

Proof. Using (15) we can show that

$$\begin{aligned} \|f' - D_n f\|_C &\leq \sup_x \sum_{k=n+1}^\infty |\langle f, P_k \rangle| |\psi(k)| |P'_k(x)| \frac{1}{|\psi(k)|} \\ &\leq \sup_x \left(\sum_{k=n+1}^\infty \left(\frac{|P'_k(x)|}{|\psi(k)|} \right)^2 \right)^{1/2} \|f\|_\psi \\ &\leq \left(\sum_{k=n+1}^\infty \frac{k^2(k+1)^2(k+1/2)}{4\psi^2(k)} \right)^{1/2} \|f\|_\psi \\ &\leq C \left(\int_n^\infty \frac{t^5}{\psi^2(t)} dt \right)^{1/2} \|f\|_\psi. \end{aligned}$$

The bounds (17), (18) are obtained by direct calculation from (16). \square

Lemma 4. *Under the assumption (2) the following bound holds true*

$$\|D_n f - D_n f^\delta\|_C \leq \frac{\delta}{2\sqrt{6}} n(n+1)(n+2). \quad (19)$$

Proof. As in the previous section, let $\xi_k = \langle f - f^\delta, P_k \rangle$, $k = 1, 2, \dots$. In view of (2) it is clear that

$$\sum_{k=1}^{\infty} \xi_k^2 \leq \delta^2.$$

Moreover, using (15) we obtain the following bound

$$\begin{aligned} \|D_n f - D_n f^\delta\|_C^2 &\leq \sup_x \sum_{k=1}^n |\xi_k| |P'_k(x)| \\ &\leq \delta \sup_x \left(\sum_{k=1}^n |P'_k(x)|^2 \right)^{1/2} \\ &\leq \delta \left(\sum_{k=1}^n \frac{k^2 (k+1)^2 (k+1/2)}{4} \right)^{1/2} \\ &= \frac{\delta}{2} \left(\frac{1}{6} n^2 (n+1)^2 (n+2)^2 \right)^{1/2} \end{aligned}$$

\square

Now we will formulate the main result of this section that follows directly from Lemmas 3 and 4.

Theorem 2. *Let the assumption (2) be satisfied. Assume that $f \in W_2^\psi$ with $\psi(t) = t^\mu$. Then for $\mu > 3$ and $n = C\delta^{-1/\mu}$ we have*

$$\|f' - D_n f^\delta\|_C = O(\delta^{\frac{\mu-3}{\mu}}). \quad (20)$$

If $f \in W_2^\psi$ with $\psi(t) = e^{th}$, $h > 0$, then for $n = \frac{C}{h} \log \frac{1}{\delta}$ we obtain

$$\|f' - D_n f^\delta\|_C = O\left(\delta \log^3 \frac{1}{\delta}\right). \quad (21)$$

Remark 2. *To the best of our knowledge the recovery of the derivative f' in the space of continuous functions has been mainly studied under the assumptions that f belongs to the space $C^s = C^s[-1, 1]$ of s -times continuously differentiable functions, and noisy data $f_\varepsilon \in C$ are available such that*

$$\|f - f_\varepsilon\| \leq \varepsilon. \quad (22)$$

It is well known (see, e. g. [2] and references therein) that under these assumptions noisy data f_ε allow the reconstruction f' with the accuracy of order $O(\varepsilon^{\frac{s-1}{s}})$ in C -norm.

At the same time, from [8] it follows that under the assumption (2) a function $f \in C^s$ can be recovered in C -space from f^δ with the accuracy of order $O(\delta^{\frac{s-1}{s+1/2}})$. Therefore, under the assumption (2) the derivative f' of a function $f \in C^s$ can be potentially reconstructed in C -space from f^δ in two steps. At first, using $f^\delta \in L_2$ one may find a function $f_\varepsilon \in C$ such that (22) is satisfied with $\varepsilon = O(\delta^{\frac{s-1}{s+1/2}})$. Then, using the results [2], for example, one may reconstruct f' from f_ε with the accuracy of order $\varepsilon^{\frac{s-1}{s}} = O(\delta^{\frac{s-1}{s+1/2}})$ in C -norm. On the other hand, the bound (20) gives an estimation of the accuracy of a one-step procedure $D_n f^\delta$ that reconstructs f' in C -space directly from $f^\delta \in L_2$. To compare this direct reconstruction with the above mentioned two-steps procedure one may use the embedding $H_2^r \subset C^s$, $r > s + \frac{1}{2}$. Moreover, as it has been mentioned in Remark 1, $W_2^\mu \subset H_2^r$, $\mu > 2r$. Therefore, for $\mu > 2s + 1$ we have $W_2^\mu \subset C^s$. This means that for $f \in W_2^\mu$, $\mu > 2s + 1$, one should expect that the order of the accuracy of the reconstruction of the derivative f' from noisy data f^δ is better than $O(\delta^{\frac{s-1}{s+1/2}})$. The bound (20) shows that the expectation is true, since for $s < (\mu - 1)/2$ we have $\frac{s-1}{s+1/2} < \frac{(\mu-1)/2-1}{(\mu-1)/2+1/2} = \frac{\mu-3}{\mu}$, and $\delta^{\frac{s-1}{s+1/2}} > \delta^{\frac{\mu-3}{\mu}}$. This reasoning can be seen as a support for the adequateness of the bound (20).

4 Adaptation to the unknown form of the approximation error

It is clear that a regularization effect of the method $D_n f^\delta$ can only be achieved if the truncation level n , playing the role of the regularization parameter, is properly chosen depending on noise level δ and smoothness of the function f to be differentiated. The latter one influences the approximation error that should be balanced with the noise propagation error as it has been made, for example, in Theorems 1 and 2. But the choice of n presented in these theorems is essentially based on the knowledge of the form of the smoothness index function ψ , which may be rather complex and a priori unknown.

Therefore, to be of practical interest, the method $D_n f^\delta$ should be equipped with a posteriori parameter choice rule that automatically adjusts the value n of the truncation level to unknown smoothness index function ψ or, that is the same, to unknown form of the approximation error $\|f' - D_n f\|$.

In this section we present such a posteriori choice rule, which is based on the so-called balancing principle that has been extensively studied recently

(see, e.g. [8, 10] and references therein). In contrast to the well-known discrepancy principle, which has been used in [12] for choosing n in the method $D_n f^\delta$, the rule presented in this section, as well as the balancing principle itself, can be used not only in L_2 -space, but also in any space, where reliable bounds for the noise propagation error are available. Note that for the method $D_n f^\delta$ such bounds have been obtained in Lemmas 2 and 4.

In the following we assume that the noise propagation error is controlled by some known increasing continuous function λ such that

$$\|D_n f - D_n f^\delta\| \leq \lambda(n)\delta, \quad (23)$$

and the largest value $n = N$ of the truncation level is given as $N = \lfloor (\lambda^{-1}(1/\delta)) \rfloor$. This bound for the truncation level is natural, since for $n > \lfloor (\lambda^{-1}(1/\delta)) \rfloor$ only a very rough estimation $\|D_n f - D_n f^\delta\| = O(1)$ can be guaranteed.

Note that the bound (23) is typically satisfied. For example, in case of L_2 -norm, i.e. $\|\cdot\| = \|\cdot\|_{L_2}$, from Lemma 2 it follows that $\lambda(n) = \frac{1}{2}n(n^2 + 6n + 5)^{1/2}$, and in case of $\|\cdot\| = \|\cdot\|_C$ Lemma 4 tells that $\lambda(n) = \frac{n}{2\sqrt{6}}(n+1)(n+2)$.

Definition 1. *Following [8], we say that a function $\varphi(n) = \varphi(n; \lambda, f, \delta)$ is admissible for given λ , f and δ if the following holds*

1. $\varphi(n)$ is a non-increasing function on $[1, N]$,
2. $\varphi(N) < \lambda(N)\delta$,
3. $\forall n \in \{1, \dots, N\}$

$$\|f' - D_n f\| \leq \varphi(n). \quad (24)$$

For given λ , f , δ the set of admissible functions is denoted by $\Phi(\lambda, f, \delta)$.

Now we are ready to present the adaptive procedure for choosing the value n of the truncation level.

Theorem 3. *Let the truncation level n_+ be chosen as*

$$n_+ = \min\{n : \|D_n f^\delta - D_m f^\delta\| \leq 4\lambda(m)\delta, m = N, N-1, \dots, n+1\}. \quad (25)$$

Then the following error bound holds true

$$\|f' - D_{n_+} f^\delta\| \leq 6\rho \inf_{\varphi \in \Phi(\lambda, f, \delta)} \min_{n=1, \dots, N} \{\varphi(n) + \lambda(n)\delta\}, \quad (26)$$

where $\rho = \max \left\{ \frac{\lambda(n+1)}{\lambda(n)}, n = 1, 2, \dots, N \right\}$.

Remark 3. Note that the triangle inequality combined with (23), (24) gives the estimation

$$\|f' - D_n f^\delta\| \leq \varphi(n) + \lambda(n)\delta. \quad (27)$$

In view of (27) the quantity

$$e(\lambda, f, \delta) = \inf_{\varphi \in \Phi(\lambda, f, \delta)} \min_{n=1, \dots, N} \{\varphi(n) + \lambda(n)\delta\}$$

is the best possible accuracy that can be guaranteed for approximation of $f'(x)$ by $D_n f^\delta$ under the assumptions (2), (23).

Thus, the choice $n = n_+$ gives an error bound that only by a constant factor worse than the best possible one.

Remark 4. The factor ρ in the error bound (26) can be easily found using the bounds of Lemmas 2 and 4. For example, in case of $\|\cdot\| = \|\cdot\|_{L_2}$ we have $\rho = \sqrt{7}$, while in case $\|\cdot\| = \|\cdot\|_C$ we can put $\rho = 4$.

Note that the Theorem 3 can be proven similar to the deterministic oracle inequality in [8], but since our definition of the admissible function is different from that of [8], we present the proof of the theorem for the sake of self-containedness.

Proof. Let $\varphi \in \Phi(\lambda, f, \delta)$ be any admissible function and let us temporarily introduce the values

$$\begin{aligned} n_0 &= \min\{n : \varphi(n) \leq \lambda(n)\delta\}, \\ n_1 &= \arg \min\{\varphi(n) + \lambda(n)\delta, n \in \{1, \dots, N\}\}. \end{aligned}$$

Observe that

$$\lambda(n_0)\delta \leq \rho(\varphi(n_1) + \lambda(n_1)\delta), \quad (28)$$

because either $n_1 \leq n_0 - 1 < n_0$, so that by definition we have $\lambda(n_0 - 1)\delta \leq \varphi(n_0 - 1)$ and

$$\lambda(n_0)\delta = \frac{\lambda(n_0)}{\lambda(n_0 - 1)}\lambda(n_0 - 1)\delta \leq \rho\varphi(n_0 - 1) \leq \rho\varphi(n_1) \leq \rho(\varphi(n_1) + \lambda(n_1)\delta),$$

or $n_0 \leq n_1$ in which case

$$\lambda(n_0)\delta \leq \lambda(n_1)\delta < \rho(\varphi(n_1) + \lambda(n_1)\delta).$$

Now we also show that $n_0 \geq n_+$. Indeed, for any $m \in \{n_0 + 1, \dots, N\}$,

$$\begin{aligned} \|D_m f^\delta - D_{n_0} f^\delta\| &\leq \|f' - D_m f^\delta\| + \|f' - D_{n_0} f^\delta\| \\ &\leq \varphi(m) + \lambda(m)\delta + \varphi(n_0) + \lambda(n_0)\delta \\ &\leq 2\varphi(n_0) + \lambda(m)\delta + \lambda(n_0)\delta \\ &\leq 3\lambda(n_0)\delta + \lambda(m)\delta \leq 4\lambda(m)\delta. \end{aligned}$$

It means that

$$n_0 \geq n_+ = \min\{n : \|D_n f^\delta - D_m f^\delta\| \leq 4\lambda(m)\delta, m = N, N-1, \dots, n+1\}.$$

Using this and (28) one can finally obtain

$$\begin{aligned} \|f' - D_{n_+} f^\delta\| &\leq \|f' - D_{n_0} f^\delta\| + \|D_{n_0} f^\delta - D_{n_+} f^\delta\| \\ &\leq \varphi(n_0) + \lambda(n_0)\delta + 4\lambda(n_0)\delta \\ &\leq 6\lambda(n_0)\delta \leq 6\rho(\varphi(n_1) + \lambda(n_1)\delta) \\ &\leq 6\rho \min_{n=1, \dots, N} \{\varphi(n) + \lambda(n)\delta\}. \end{aligned}$$

This estimation holds true for an arbitrary admissible function $\varphi \in \Phi(\lambda, f, \delta)$. Therefore, we conclude that

$$\|f' - D_{n_+} f^\delta\| \leq 6\rho \inf_{\varphi \in \Phi(\lambda, f, \delta)} \min_{n=1, \dots, N} \{\varphi(n) + \lambda(n)\delta\}.$$

The proof is complete. \square

Remark 5. From the proof of Theorem 3 it can be seen that the error bound (26) can also be obtained for $n = n_+$ chosen as follows

$$n_+ = \min\{n : \|D_n f^\delta - D_m f^\delta\| \leq 3\lambda(n)\delta + \lambda(m)\delta, m = N, N-1, \dots, n+1\}. \quad (29)$$

This choice rule is used in our numerical examples.

5 Numerical examples

In this section, we demonstrate how the considered approach (4) together with the adaptive parameter choice rule (29) can be easily and effectively used for reconstructing the derivative of functions both with finite and infinite smoothness. Moreover, it will be obvious from the experiments that the bounds derived in the theoretical analysis do manifest themselves in practice.

In our numerical experiments, we consider the approximations (4) with truncation levels $n \in \{1, 2, \dots, 50\}$. Noisy coefficients $\{f_i^\delta\}_{i=1}^n$ are simulated as follows. At first, the values of a function f are calculated at 400 points x_i randomly distributed in $[-1, 1]$. Then the data $(x_i, f(x_i))$, $i = 1, 2, \dots, 400$, are used to find the coefficients of the linear combination $\sum_{i=1}^{50} c_i P_i(x)$ by the least squares method. Finally, we take $f_i^\delta = c_i + \xi_i^\delta$, $i = 1, 2, \dots, 50$, where the vector $(\xi_i^\delta)_{i=1}^{50}$ is drawn from the normal distribution with zero mean and the variance δ^2 .

We start with the numerical experiments confirming robustness of the adaptive choice $n = n_+$ and our error bounds for infinitely smooth functions. Consider at first the function

$$f_1(x) = (1 - 2x\eta + \eta^2)^{-1/2} = \sum_{k=0}^{\infty} \eta^k \sqrt{k + 1/2} P_k(x);$$

$$f_1'(x) = \eta(1 - 2x\eta + \eta^2)^{-3/2}$$

with $\eta = 1/3$. In the left panel of Figure 1 we depict the C -norm error of the derivative approximation for different truncation levels n and $\delta = 10^{-3}$. As it can be seen from the figure, the proposed adaptive choice rule $n = n_+$ allows us to pick one of the truncation levels giving the best accuracy (see the middle panel of the figure). For comparison, in the right panel we show the derivative approximation by the method (4) with a priori parameter choice $n = 19$. As one can see, near $x = -1$ the approximation based on the adaptive method (29) performs much better compared to the one in the right panel.

In order to verify the convergence rate in Theorem 2, we estimate the constants from the error bound (21) as follows $C = C_1 = \frac{\|f_1' - D_{n_+} f_1^\delta\|_C}{|\delta \log^3 \delta|}$. As it can be seen from Table 1, for different noise levels the values of these constants exhibit a rather stable behavior supporting the convergence rate $O(\delta \log^3 \frac{1}{\delta})$ indicated in the Theorem 2.

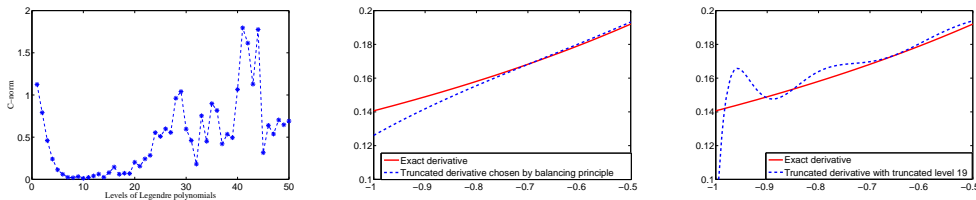


Figure 1: Numerical example for f_1 . **Left**: the C -norm error for truncation levels from 1 to 50. **Middle**: the approximated derivative with the truncation level $n_+ = 9$ chosen in accordance with (29). **Right**: the approximated derivative with the truncation level $n = 19$ chosen a priori.

These observations are also true for the second example with another infinitely smooth function

$$f_2(x) = x^2; \quad f_2'(x) = 2x.$$

From Figure 2 it can be seen that the proposed adaptive choice rule $n = n_+$ automatically picks up one of the truncation levels giving the best accuracy, whereas the derivative approximation with a priori parameter choice $n = 23$ exhibits poor performance near $x = -1$.

Table 1: Numerical example for f_1 . Estimation of the constants from the error bound (21) for different noise levels.

Noise level	$10^{-2.5}$	$10^{-2.9}$	$10^{-3.3}$	$10^{-3.7}$	$10^{-4.1}$	$10^{-4.5}$	$10^{-4.9}$	$10^{-5.3}$
Constant	0.2497	0.3003	0.1363	0.2232	0.7072	1.8032	2.0303	8.6845
Truncation level n_+	6	9	13	17	23	32	44	61

Note that in one of the recent papers [16] on numerical differentiation, where the authors consider a method based on the Discrete Cosine Transform, the Gibbs phenomenon was observed in the approximated derivative near the endpoint of the interval of the definition. From Figure 2 one may conclude that the method (4) with $n = n_+$ is superior to [16] in weakening Gibbs phenomenon.

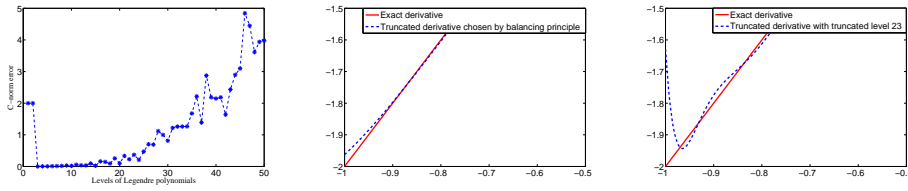


Figure 2: Numerical example for f_2 . **Left:** the C -norm error for truncation levels from 1 to 50. **Middle:** the approximated derivative with the truncation level $n_+ = 13$ chosen in accordance with (29). **Right:** the approximated derivative with the truncation level $n = 23$ chosen a priori.

To show the tightness of the error bounds in Theorem 2, we consider another two functions of a final smoothness

$$f_3(x) = \begin{cases} -x, & x \in [-1, 0), \\ x, & x \in [0, 1], \end{cases} \quad f_4(x) = \begin{cases} -x^2, & x \in [-1, 0), \\ x^2, & x \in [0, 1]. \end{cases}$$

After a first glance at Figure 3 (upper left panel), one may think that the choice of the truncation level $n = n_+ = 15$ made in accordance with (29) for $\delta = 10^{-3}$ is not optimal in the considered case. However, once the value of the truncation level n becomes smaller, the method (4) shows low accuracy of the derivative approximation near the point of discontinuity, while the truncation level $n = n_+$ enhances the accuracy near this point. This effect is even more pronounced for smaller noise level $\delta = 10^{-4}$, as it can be seen from

the lower right panel of Figure 3, where the approximated derivative with the truncation level $n_+ = 40$ chosen in accordance with (29) is displayed, and this is the best choice in the considered case.

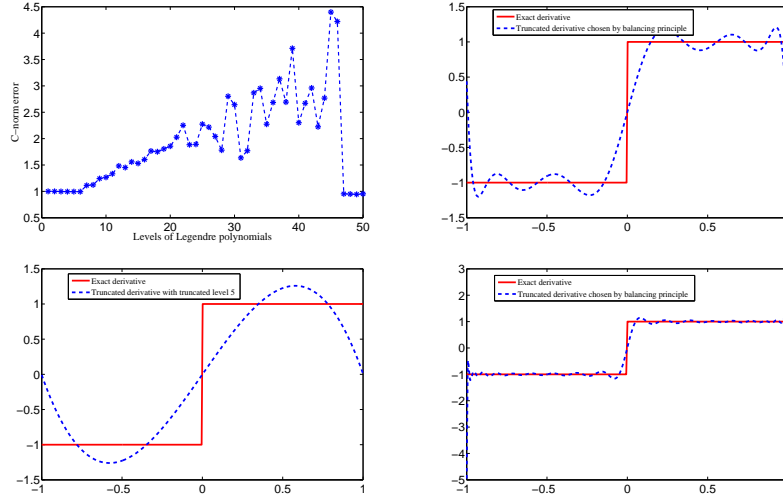


Figure 3: Numerical example for f_3 . **Upper left:** the C -norm error for truncation levels from 1 to 50. **Upper right:** the approximated derivative with the truncation level $n_+ = 15$ chosen in accordance with (29). **Lower left:** the approximated derivative with the truncation level $n = 5$ chosen a priori. **Lower right:** the approximated derivative with the truncation level $n_+ = 40$ chosen in accordance with (29) for $\delta = 10^{-4}$.

The last example has been inspired by the paper [12], where the well-known discrepancy principle is used to choose the truncation level n . As we already mentioned above, in the case of square-summable noise the discrepancy principle can be used only in L_2 -space. In contrast to this, the rule (29), as well as the balancing principle itself, can be used in any space, where reliable bounds for the noise propagation error are available, such as those given, for example, in Lemmas 2 and 4. In order to perform the comparison of both rules, in Figure 4 we display the approximation of the derivative f_4' for $x \in [-1, -0.9]$. From the figure it is clearly seen that the balancing principle outperforms the discrepancy principle in the sense of C -norm error.

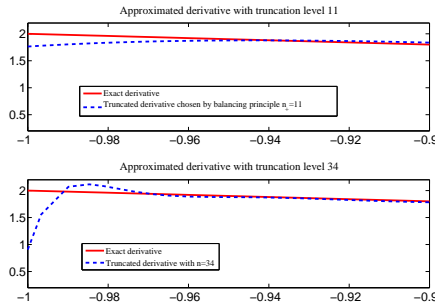


Figure 4: Comparison of the approximated derivative f'_4 with the truncation levels $n_+ = 11$ and $n = 34$ chosen, respectively, in accordance with (29) and the discrepancy principle.

References

- [1] Anderssen R S and Hegland M 2010 Derivative spectroscopy – an enhanced role for numerical differentiation *J. Integral Equations Appl.* **22** 355–67
- [2] Lu S and Pereverzev S V 2006 Numerical differentiation from a viewpoint of regularization theory *Math. Comp.* **75** 1853–70
- [3] Engl H, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)
- [4] Flemming J, Hofmann B and Mathé P 2011 Sharp converse results for the regularization error using distance functions *Inverse Problems* **27** 025006
- [5] Dolgoplova T F and Ivanov V K 1966 Numerical differentiation *Comp. Math. and Math. Physics* **6** 570–76
- [6] Tikhonov A N and Arsenin V Y 1977 *Solutions of ill posed problems* (V. H. Winston and Sons, distributed by Wiley, New York)
- [7] Ivanov V K and Koroljuk T I 1969 The estimation of errors in the solution of linear ill-posed problems *Z. Vycisl. Mat. i Mat. Fiz.* **9** 30–41
- [8] Mathé P and Pereverzev S V 2006 Regularization of some linear ill-posed problems with discretized random noisy data *Math. Comp.* **75** 1913–29

- [9] Lazarov R D, Lu S and Pereverzev S V 2007 On the balancing principle for some problems of numerical analysis *Numer. Math.* **106** 659–89
- [10] Hämarik U and Raus T 2009 About the balancing principle for choice of the regularization parameter *Numer. Functional Anal. and Optimiz.* **30** 951–70
- [11] Müller C 1969 *Foundations of the Mathematical Theory of Electromagnetic Waves* (Springer-Verlag, Berlin, Heidelberg, New York)
- [12] Zhao Z 2010 A truncated Legendre spectral method for solving numerical differentiation *International Journal of Computer Mathematics* **87** 3209–17
- [13] Rafal'son S Z 1969 The approximation of functions by Fourier-Jacobi sums *Izv. Vysš. Učebn. Zaved. Matematika* **4** 54–62
- [14] Tomin N G 1973 An application of the interpolation of linear operators to questions of the convergence of series of Fourier coefficients with respect to classical orthogonal polynomials *Dokl. Akad. Nauk SSSR* **212** 1074–77
- [15] Nussbaum M and Pereverzev S V 1999 The degree of ill-posedness in stochastic and deterministic noise models (*Preprint No 509, WIAS-Berlin*)
- [16] Hu B and Lu S 2012 Numerical differentiation by a Tikhonov regularization method based on the discrete cosine transform *To appear in Applicable Analysis*