Comput. Math. Applic. Vol. 13, No. 4, pp. 381–386, 1987
Printed in Great Britain.

0097-4943/87 $3.00 + 0.00

AUTOMATIC NUMERICAL DIFFERENTIATION BY DISCRETE MOLLIFICATION

D. A. Murio
Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221, U.S.A.

(*Received* 26 *August* 1986)

Communicated by E. Y. Rodin

Abstract—A new, very simple, totally automated and powerful technique for numerical differentiation based on the computation of the derivative of a suitable filtered version of the noisy data by discrete mollification is presented. Several numerical examples of interest are also analyzed.

1. INTRODUCTION

In several practical contexts, it is sometimes necessary to estimate the derivative of a function whose values are given approximately at discrete points. It is well known that the process of differentiation is an ill-posed problem because small errors in the data might cause large errors in the computed derivative. Numerical differentiation has been discussed by many authors and a number of solution methods have been proposed. Finite differences approaches have been used, for example, in Refs [1]–[4]. Regularization procedures have been analyzed in Refs [5]–[7] and more recently in [8]. For methods of statistical nature see Refs [9]–[12]. In this paper we present yet another method of solution based on attempting to reconstruct a mollified version of the derivative. This approach was first introduced, for a different algorithm, in Ref. [13] and leads naturally to a very simple and powerful computational technique. Our approximation is generated initially by filtering the noisy data by discrete convolution with a suitable averaging kernel and then using centered differences to numerically solve the associated well-posed problem. The efficiency of this method is demonstrated in Section 5 where we describe several examples, presented in Ref. [8].

In Section 2 we describe the stabilized problem and in Section 3 we introduce a new automatic algorithm to uniquely determine the radius of mollification, depending on the amount of noise in the data. Section 4 describes in detail the computational procedures related with our method.

2. STABILIZED PROBLEMS

Let $I = [0, 1]$ and let $C^{\circ}(I)$ denote the set of continuous functions over I with

$$\|f(x)\|_{\infty,I} = \max_{x\in I}|f(x)| < \infty.$$

Given a compact set K, $K \subset I$, we consider the problem of estimating in K, the derivative $g'(x)$ of a function $g(x)$ defined on I and observed with error. We assume $g(x) \in C^{\circ}(I)$ and $g''(x) \in C^{\circ}(K)$. Instead of $g(x)$, we know some data function $g^{\epsilon}(x) \in C^{\circ}(I)$ such that $\|g^{\epsilon}(x) - g(x)\|_{\infty,I} \leqslant \epsilon$.

Following Ref. [13], in order to stabilize the differentiation problem, we introduce the function

$$\rho_\delta(x) = \begin{cases} \left\{\int_{-\delta}^{\delta} \exp[s^2/(s^2-\delta^2)]\,\mathrm{d}s\right\}^{-1} \exp[x^2/(x^2-\delta^2)] & \text{if} \quad |x| < \delta \\ 0 & \text{if} \quad |x| \geqslant \delta. \end{cases} \tag{1}$$

$\rho_\delta(x)$ is a C^∞ (infinitely differentiable) function in R with support in $|x| \leqslant \delta$, such that $\rho_\delta(x) \geqslant 0$ and

$$\int_R \rho_\delta(x)\,\mathrm{d}x = 1.$$

If $\delta > 0$, δ smaller than the distance from K to ∂I, the function

$$J_\delta f(x) = (\rho_\delta * f)(x) = \int_{x-\delta}^{x+\delta} \rho_\delta(x-s) f(s)\,\mathrm{d}s \tag{2}$$

is a C^∞ function in R and for fixed $x \in K$, $\rho_\sigma(x-s)$ has compact support in I. $J_\delta f$ is the mollifier of f and δ is the radius of mollification. Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}x} J_\delta f(x) = (\rho_\delta * f)'(x) \tag{3a}$$

$$= (\rho_\delta' * f)(x). \tag{3b}$$

The following two Lemmas and Theorem 1 are proved, with minor modifications, in Ref. [13] and are included here for completeness.

Lemma 1. If $\| g'' \|_{\infty, K} \leqslant M$, then $\| (\rho_\delta * g)' - g' \|_{\infty, K} \leqslant \delta M$.

Proof. For $x \in K$,

$$(\rho_\delta * g)'(x) = \int_{x-\delta}^{x+\delta} \rho_\delta(x-s) g'(s)\,\mathrm{d}s$$

and

$$g'(x) = \int_{x-\delta}^{x+\delta} \rho_\delta(x-s) g'(x)\,\mathrm{d}s.$$

Subtracting and using the mean value theorem,

$$[(\rho_\delta * g)'(x) - g'(x)] \leqslant \delta M \int_{x-\delta}^{x+\delta} \rho_\delta(x-s)\,\mathrm{d}s = \delta M.$$

Thus,

$$\| (\rho_\delta * g)' - g' \|_{\infty, K} \leqslant \delta M.$$

Lemma 2. If $g^\epsilon(x) \in C^\circ(I)$ and $\| g^\epsilon - g \|_{\infty, I} \leqslant \epsilon$, then

$$\| (\rho_\delta * g^\epsilon)' - (\rho_\delta * g)' \|_{\infty, K} \leqslant \frac{\epsilon \alpha}{\delta},$$

where

$$\alpha = \left\{ \int_0^1 \exp[s^2/(s^2 - \delta^2)]\,\mathrm{d}s \right\}^{-1} \cong 1.65.$$

Proof. For $x \in K$,

$$|(\rho_\delta * g^\epsilon)'(x) - (\rho_\delta * g)'(x)| \leqslant \int_{x-\delta}^{x+\delta} \left| \frac{\mathrm{d}}{\mathrm{d}x} \rho_\delta(x-s) \right| |g^\epsilon(x) - g(s)|\,\mathrm{d}s$$

$$\leqslant \epsilon \int_{x-\delta}^{x+\delta} \left| \frac{\mathrm{d}}{\mathrm{d}x} \rho_\delta(x-s) \right| \mathrm{d}s = \frac{\epsilon \alpha}{\delta}.$$

It follows that

$$\| (\rho_\delta * g^\epsilon)' - (\rho_\delta * g)' \|_{\infty, K} \leqslant \frac{\epsilon \alpha}{\delta}.$$

Lemma 2 shows that attempting to reconstruct the derivative of the mollified data function is a stable problem with respect to perturbations in the data, in the maximum norm, for δ fixed.

Theorem 1. Under the conditions of Lemmas 1 and 2,

$$\| (\rho_\delta * g^\epsilon)' - g' \|_{\infty, K} \leqslant \delta M + \epsilon \alpha / \delta. \tag{4}$$

Proof. The estimate follows from Lemmas 1 and 2 and the triangle inequality.

We observe that the r.h.s. of inequality (4) is minimized by choosing $\delta = (\epsilon \alpha / M)^{1/2}$, but this optimal selection of the radius of mollification is, in actual computations, impossible because M is not known in general.

3. PARAMETER SELECTION

In this section we indicate a procedure to determine the radius of mollification, δ, based on properties of the filtered data $\rho_\delta * g^\epsilon$.

We need to extend the data function $g^\epsilon(x)$ defined on $I = [0, 1]$ to all R in such a way that $g^\epsilon(x)$ decays smoothly to zero in $I_a - I$, where $I_a = [-a, 1 + a]$ for some $a > 0$, and it is zero in $R - I_a$. For instance, we can define

$$g^\epsilon(x) = g^\epsilon(0)\exp[x^2/(x^2 - a^2)], \quad -a \leqslant x \leqslant 0$$
$$g^\epsilon(x) = g^\epsilon(0)\exp\{(x-1)^2/[(x-1)^2 - a^2]\}, \quad 1 \leqslant x \leqslant 1 + a. \tag{5}$$

The mollification of the extended data function $g^\epsilon(x)$ by convolution with the kernel $\rho_\delta(x)$ is actually an averaging process that satisfies

Lemma 3. If $\delta_1 > \delta_2 \geqslant 0$, then $\|J_{\delta_1}g^\epsilon\|_{\infty,I} \leqslant \|J_{\delta_2}g^\epsilon\|_{\infty,I}$ and $\|J_{\delta_1}g^\epsilon - g^\epsilon\|_{\infty,I} \geqslant \|J_{\delta_1}g^\epsilon\|_{\infty,I}$.

Proof. It follows from expressions (2) and (5) and the fact [14] that $J_\delta g^\epsilon(x) \to g^\epsilon(x)$ uniformly on I, if $\delta \to 0$.

The monotonicity properties in Lemma 3 show that there is a unique $\tilde{\delta}$ such that

$$\|J_{\tilde{\delta}}g^\epsilon - g^\epsilon\|_{\infty,I} = \epsilon. \tag{6}$$

This particular parameter choice criterion is characterized by selecting, among all possible mollifiers of g^ϵ satisfying $\|J_\delta g^\epsilon - g^\epsilon\|_{\infty,I} \leqslant \epsilon$, the mollifier with minimum maximum norm. More precisely,

Theorem 2. $\min_{\|J_\delta g^\epsilon - g^\epsilon\|_{\infty,I} \leqslant \epsilon} \|J_\delta g^\epsilon\|_{\infty,I} = \|J_{\tilde{\delta}}g^\epsilon\|_{\infty,I}$, where $\tilde{\delta}$ is such that $\|J_{\tilde{\delta}}g^\epsilon - g^\epsilon\|_{\infty,I} = \epsilon$,

Proof. Suppose there exists δ such that $\|J_\delta g^\epsilon - g^\epsilon\| < \epsilon$. It follows then that $\delta < \tilde{\delta}$ and by Lemma 3, $\|J_{\tilde{\delta}}g^\epsilon\|_{\infty,I} \leqslant \|J_\delta g^\epsilon\|_{\infty,I}$.

The parameter selection (6) determines $\tilde{\delta}$ in a manner consistent with the amount of noise in the data function g^ϵ. Note that if $\|g^\epsilon - g\| \leqslant \epsilon$, then $\|J_{\tilde{\delta}}g^\epsilon - g\| \leqslant 2\epsilon$. Furthermore, the bisection method can easily be implemented to numerically determine $\tilde{\delta}$. The computational details are presented in the next section.

4. NUMERICAL PROCEDURE

To numerically approximate $d[J_\delta g^\epsilon(x)]/dx$, a quadrature formula for the convolution equation (6) is used in Ref. [13]. The actual implementation of this well-posed problem assumes that δ is known. Instead, our procedure for numerical differentiation is based on formula (3a), a particular choice of δ and centered differences.

Since in practice only a discrete set of data points is available, we assume in what follows that the data function g^ϵ is a discrete function in $I = [0, 1]$, measured at the $N + 1$ sample points $x_i = i\Delta x$, $i = 0, 1, \ldots, N$, $N\Delta x = 1$. Given $a > \Delta x$, we use equation (5) to extend the data to $I_a = [-a, 1 + a]$ and since the data is defined to be zero in $R - I_a$, we consider the extended discrete data function g^ϵ defined at equally spaced sample points on any interval of interest.

The parameter selection is implemented by solving the discrete version of equation (6) using the bisection method. The following steps summarize the method.

Step 1: Set $\delta_{min} = \Delta x$, $\delta_{max} = 0.5$ and choose an initial value of δ between δ_{min} and δ_{max}.

Step 2: Compute $J_\delta g^\epsilon = \rho_\delta * g^\epsilon$ by discrete convolution on a sufficiently large interval.

Step 3: If $F(\delta) = \max_{0 \leqslant i \leqslant n} |J_\delta g^\epsilon(x_i) - g^\epsilon(x_i)| = \epsilon \pm \eta$, where η is a given tolerance, exit.

Step 4: If $F(\delta) - \epsilon < -\eta$, set $\delta_{\min} = \delta$. If $F(\delta) - \epsilon > \eta$, set $\delta_{\max} = \delta$. The updated value of δ is always given by $\frac{1}{2}(\delta_{\max} + \delta_{\min})$.

Step 5: Return to Step 2.

Once the radius of mollification $\bar{\delta}$ and the discrete data function $J_{\bar{\delta}} g^\epsilon$ are determined, we use centered differences to approximate the derivative of $J_{\bar{\delta}} g^\epsilon$ at the sample points of the interval $[\bar{\delta}, 1 - \bar{\delta}]$.

5. NUMERICAL RESULTS

In this section we discuss the implementation of our numerical method and the tests which we have performed in order to investigate the accuracy and stability of the numerical differentiation procedure.

In all the examples, $\Delta x = 0.01$, $a = 0.1$ and $I = [0, 1]$. The exact data function is denoted by $g(x)$ and the noisy data function $g^\epsilon(x)$ is obtained by adding an ϵ random error to $g(x)$, i.e.

$$g^\epsilon(x_i) = g(x_i) + \epsilon\theta_i, \tag{7}$$

where $x_i = i\Delta x$, $i = 0, 1, \ldots, N$, $N\Delta x = 1$ and θ_i is a uniform random variable with values in $[-1, 1]$ such that

$$\max_{0 \leqslant i \leqslant N} |g^\epsilon(x_i) - g(x_i)| \leqslant \epsilon.$$

After extending the discrete data function as explained in Section 4, the parameter selection criterion was implemented with the tolerance η, used in Step 3 of the algorithm, set to reflect a 0.05 error in the satisfaction of the constraint. In all cases, independently of the initial choice of δ, convergence to the value $\bar{\delta}$ determined by the selection criterion was reached in no more than eight iterations. The discrete numerical approximation to the derivative $g'(x)$, denoted $\tilde{g}'(x_i)$, is then reconstructed by means of centered differences in $I_{\bar{\delta}} = [\bar{\delta}, 1 - \bar{\delta}]$. In what follows we use $\| f \|_{\infty, I_{\bar{\delta}}}$ and $\| f \|_{2, I_{\bar{\delta}}}$ to represent

$$\max_{x_i \in I_{\bar{\delta}}} |f(x_i)| \quad \text{and} \quad \left[\frac{1}{n} \sum_{i=1}^{n} |f(x_i)|^2\right]^{1/2}, \quad x_i \in I_{\bar{\delta}},$$

respectively.

Example 1

As a first example we consider $g(x) = x(1 - \frac{1}{2}x)$. Figure 1 shows the solution obtained by our method (+) and the exact derivative $g'(x) = 1 - x$. With $\epsilon = 0.05$ and $\Delta x = 0.01$, the correspond-

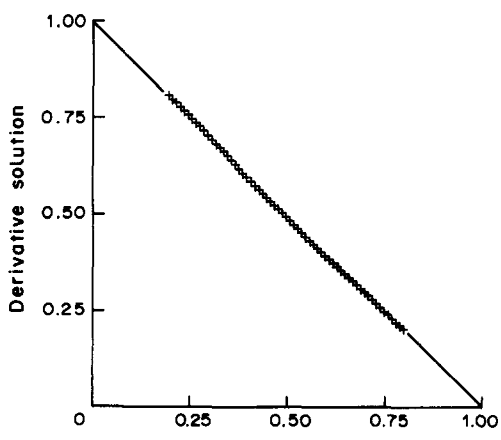

Fig. 1. $g(x) = x(1 - \frac{1}{2}x)$, $\epsilon = 0.05$, $\Delta x = 0.01$, $\bar{\delta} = 0.2$. Computed solution, + + +; true solution, ——.

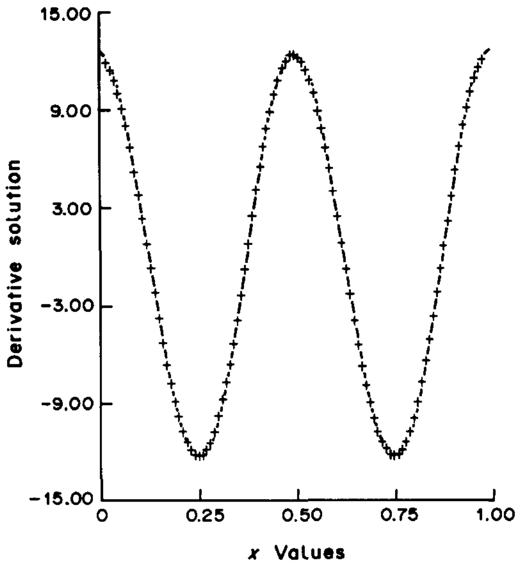

Fig. 2. $g(x) = \sin 4\pi x$, $\epsilon = 0.01$, $\Delta x = 0.01$, $\bar{\delta} = 0.02$. Computed solutions, + + +; true solution, ——.

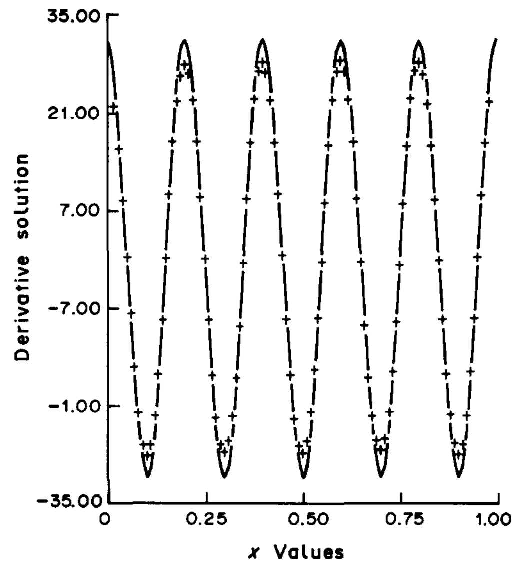

Fig. 3. $g(x) = \sin 10\pi x$, $\epsilon = 0.1$, $\Delta x = 0.01$, $\bar{\delta} = 0.02$. Computed solution, + + +; true solution, ——.

ing radius of mollification in this case is $\bar{\delta} = 0.2$. The resolution in this problem is quite good considering the relative high noise level which we used. The associated error norms are given by $\|g' - \tilde{g}'\|_{\infty, I_\delta} = 0.001108$ and $\|g' - \tilde{g}'\|_{2, I_\delta} = 0.00576$.

Example 2

Our second example is rather oscillatory on [0, 1]. We chose $g(x) = \sin 4\pi x$ and we use $\epsilon = 0.01$. With $\Delta x = 0.01$, $\bar{\delta} = 0.02$. In Fig. 2, we plot the numerical solution obtained by our method (+) and the exact derivative $g'(x) = 4\pi \cos 4\pi x$. The corresponding relative error norms are given by $\|g' - \tilde{g}'\|_{\infty, I_\delta}/\|g'\|_{\infty, I_\delta} = 0.02439$ and $\|g' - \tilde{g}'\|_{2, I_\delta}/\|g'\|_{2, I_\delta} = 0.01976$.

Example 3

In Fig. 3 we plot the exact derivative of $g(x) = \sin 10\pi x$ and the solution obtained by our method (+). The data function is highly oscillatory and the noise level $\epsilon = 0.1$ very high. With

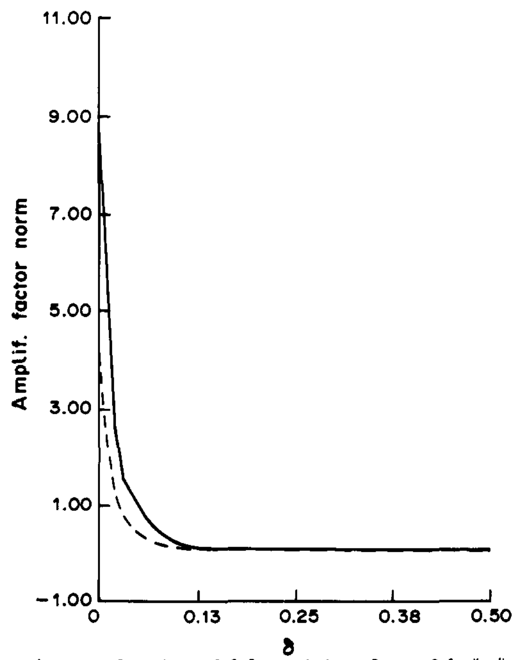

Fig. 4. Response to noise as a function of δ for $g^\epsilon(x_i) = \epsilon\theta_i$, $\epsilon = 0.1$. $\|\cdot\|_{\infty, I}$, ——; $\|\cdot\|_{2, I}$, – – –.

$\Delta x = 0.01$, the resulting radius of mollification is $\tilde{\delta} = 0.02$. The relative error norms are given by $\|g' - \tilde{g}'\|_{\infty, I_g}/\|g'\|_{\infty, I_g} = 0.13282$ and $\|g' - \tilde{g}'\|_{2, I_g}/\|g'\|_{2, I_g} = 0.11180$. We conclude that even in this rather difficult case, the method performs quite satisfactorily.

Finally, in order to investigate the stability of our numerical method we would like to determine the amplification factor associated with errors in the data when using the numerical differentiation procedure. If we set $g(x_i) = 0$ in equation (7), we can compute the response of the method to pure noise as a function of the radius of mollification and thereby get a measure of the amplification factor. Since all of the response norms are essentially proportional, we only show in Fig. 4 a representative curve for $\epsilon = 0.1$. The solid curve gives $\|\cdot\|_{\infty, I}$ and the dashed curve corresponds to $\|\cdot\|_{2, I}$. We notice that the "derivative of the noise" has been computed on $I = [0, 1]$ for every value of δ. To obtain these responses, we read the discrete data on the interval $[-1, 2]$ with $\Delta x = 0.01$ and performed the discrete convolution in $[-0.5, 1.5]$ for each value of δ. Then we applied centered differences to obtain the approximate derivative in $I = [0, 1]$ using the corresponding δ value.

REFERENCES

1. L. W. Johnson and R. D. Riess, An error analysis for numerical differentiation. *J. Inst. Math. Appllic.* **11,** 115–120 (1973).
2. J. N. Lyness, Has numerical differentiation a future? In *Proc. 7th Manitoba Conf. on Numerical Mathematics and Computing*, pp. 107–129. Congress. Numer., Utilitas, Math., Winnipeg, Manitoba (1978).
3. J. Oliver, An algorithm for numerical differentiation of a function of one real variable. *J. comput. appl. Math.* **6,** 145–160 (1980).
4. T. Strom and J. N. Lyness, On numerical differentiation. *BIT* **15,** 314–322 (1975).
5. J. Cullum, Numerical differentiation and regularization. *SIAM Jl numer. Analysis* **8,** 254–265, (1971).
6. N. S. Surova, An investigation of the problem of reconstructing a derivative by using an optimal regularizing integral operator. In *Numerical Methods and Programming*, Vol. 1, pp. 30–34. Saratov. Gos. Univ., Saratov, U.S.S.R. (1977).
7. E. P. Vedeneev and N. P. Zidkov, Applications of the regularization method to the differentiation of a function of one variable that is given by tables. *Computing Methods and Programming—XIII*, pp. 255–261. Izdat. Moskov. Univ., Moscow (1969).
8. J. T. King and D. A. Murio, Numerical differentiation by finite-dimensional regularization. *IMA Jl numer. Analysis* **6,** 65–85 (1986).
9. K. S. Andersen and P. Bloomfield, A time series approach to numerical differentiation. *Technometrics* **16,** 69–75 (1974).
10. M. L. Baart, Computational experience with the spectral smoothing method for differentiating noisy data. *J. comput. Phys.* **42,** 141–151 (1981).
11. D. D. Cox, Asymptotics of M-type smoothing splines. *Ann. Statist.* **11,** 530–551 (1983).
12. J. Rice and M. Rosenblatt, Smoothing splines: regression, derivatives, and disconvolution. *Ann. Statist.* **11,** 141–156 (1983).
13. V. V. Vasin, The stable evaluation of a derivative in space $C(-\infty, \infty)$. *U.S.S.R. comput. Math. math. Phys.* **13,** 16–24 (1973).
14. A. Friedman, *Partial Differential Equation.* Holt, New York (1969).