Trigonometric mollification

D A Murio\(^1\) and C E Mejía\(^2\)

\(^1\) Dept. of Mathematical Sciences, University of Cincinnati, Cincinnati, Ohio, USA
\(^2\) Dept. of Mathematics, Universidad Nacional de Colombia, Medellín, Colombia

E-mail: \(^1\) diego@dmurio.csm.uc.edu \(^2\) cemejia@unal.edu.co

Abstract. An automatic method for data smoothing and numerical differentiation, based on mollification techniques, trigonometric interpolation and the principle of generalized cross validation is presented. With data measured at a discrete set of equidistant points in a given interval, the method allows for the approximate recovery of the function and its first derivative on the entire domain.

The trigonometric mollification algorithm provides a functional representation for the mollified computed approximations that guarantees good fitting near the endpoints of the interval and the usage of small values of the regularization parameters –radii of mollification– if necessary. These parameters are chosen automatically although no information about the noise distribution is assumed. Error estimates are included together with several numerical examples of interest.

1. Introduction

Numerical differentiation is an ill-posed problem in the sense that small errors in the data might induce large errors in the computed derivative. The method that we present in this paper allows for the stable reconstruction of a function and its first derivative if the function is known approximately at a discrete set of data points. The implementation of the numerical method is based on the mollification method (\([1, 2]\)) to filter the noisy data, generalized cross validation (GCV) (\([3, 4]\)) to automatically select the radii of mollification as functions of the perturbation level in the data, which is generally not known, and trigonometric interpolation (\([5, 6]\)) to provide a functional representation of the mollified function and its first derivative on the entire domain. This last characteristic guarantees good fitting near the endpoints of the interval and the usage of small values of the regularization parameters if necessary. The need of functional representations was motivated for applications requiring successive (adaptive) highly accurate Gaussian quadrature formulae involving the reconstructed derivatives from inexact data. The new algorithm was successfully applied to the crucial task of estimating Caputo fractional derivatives near initial times from noisy data, in the numerical solution of the generalized time fractional inverse heat conduction problem (GTFIHCP) \([7]\).

Previous results, relating mollification and differentiation of noisy data, can be found in \([8, 9]\). Many inverse problems of importance in Science and Engineering are ultimately based on numerical differentiation procedures. Due to the ill-posedness of the problem, regularization of some kind is required in order to compute meaningful approximations of derivatives. With respect to regularization, it is important to remember Hansen (\([10]\), p. 2), who wrote: “No regularization method is superior to the other methods. Rather, each method has its advantages, depending on the application in which it is used.” Accordingly, numerical differentiation is an
inverse problem that has been discussed by many authors and a great number of different solution methods has been proposed. A general introduction to the subject with a generous set of references is available in [1].

This paper is organized as follows: estimates for consistency, stability and convergence of trigonometric mollification and the numerical differentiation procedure are discussed in section 2. Section 3 includes considerations on the implementation of the numerical method and computational results.

2. Background

It is well known [6], that if a real function \( f(t) \) defined on the closed interval \( I = [0, 1] \) and belonging to the class \( C^1(0, 1) \) of continuously differentiable functions for \( 0 < t < 1 \), is extended as a periodic function of period 2, then its Fourier series representation

\[
f(t) = a_0 + \sum_{j=1}^{\infty} (a_j \cos j\pi t + b_j \sin j\pi t)
\]

with Fourier coefficients given by

\[
a_0 = \frac{1}{2} \int_{-1}^{1} f(t) dt, \quad a_j = \int_{-1}^{1} f(t) \cos j\pi t dt, \quad b_j = \int_{-1}^{1} f(t) \sin j\pi t dt, \quad j \geq 1,
\]

converges uniformly to \( f(t) \) in \( (0, 1) \). At the end points \( t = 0 \) and \( t = 1 \), the Fourier series converges to \( \frac{1}{2} (f(0^+) + f(0^-)) \) and \( \frac{1}{2} (f(1^+) + f(1^-)) \) and the continuity of the extended function and/or its first derivative are not guaranteed.

The convergence of the Fourier series is determined by the properties of the function at the boundary points and the number of derivatives that the extended function possesses.

Following [5], it is possible to obtain better convergence by subtracting a linear trend from the given function \( f \) and operating instead with

\[
g(t) = f(t) - [f(0) + (f(1) - f(0))t]. \quad (1)
\]

Expanding \( g \) as an odd function of period 2, the function \( g \) and its derivative \( g' \) are both continuous at the two endpoints of the period. Observe that \( g(-1) = g(1) = 0 \), \( g'(-1) = g'(1) \) and the new function satisfies the boundary conditions \( g(0) = g(1) = 0 \). The discontinuities appear now in the second derivative at the endpoints \( t = 0 \) and \( t = 1 \).

Thus, the Fourier series of \( g(t) \) and \( g'(t) \) converge uniformly to the functions \( g(t) \) and \( g'(t) \) respectively, in \( I \).

The coefficients \( b_j \) of the sine series of \( g \),

\[
g(t) = \sum_{j=1}^{\infty} b_j \sin j\pi t, \quad (2)
\]

behave asymptotically as \( O(j^{-3}) \) for \( j \to \infty \).

Integrating by parts three times, for large \( j \),

\[
\int_{0}^{1} g(t) \sin j\pi t \ dt \simeq \frac{1}{j^3}((-1)^j f''(1) - f''(0)).
\]

More precisely [6], if \( g(t) \) is a 2-periodic function with its \( p \)th derivative a piecewise \( C^1 \)-function, then there exists a constant \( C \), independent of \( j \), such that

\[
|b_j| \leq \frac{C}{j^{p+1}}. \quad (3)
\]
2.1. Trigonometric Interpolation

Let \( h = \frac{1}{N-1} \), where \( N \) is an integer greater than 1, denote a grid interval and define a grid on \( I \) to be the set of gridpoints

\[
K = \{ t_i = ih, \ i = 0, 1, ..., N - 1 \}.
\]

(4)

The restriction of the sine series representation of \( g \) to \( K \) defines a 2-periodic gridfunction

\[
G = \{ G_i \}_{i=0}^{N-1} \equiv \{ g(t_i) \}_{i=0}^{N-1},
\]

(5)

and we consider a trigonometric polynomial which interpolates \( g \), that is,

\[
\text{Int}_{N-2} g(t) = \sum_{j=1}^{N-2} \hat{b}_j \sin j\pi t,
\]

(6)

with

\[
G_i = \text{Int}_{N-2} g(t_i), \ i = 1, 2, ..., N - 2,
\]

(7)

after dropping the zero values associated with \( G_0 = g(0) = G_{N-1} = g(1) \).

Equation (7) represents a system of \( N - 2 \) equations for the \( N - 2 \) unknowns \( \hat{b}_j \) and we need to show that the system has a unique solution. Since the gridfunctions

\[
\{ \varphi_j \}_{j=1}^{N-2} = \{ \sin j\pi t_1, ..., \sin j\pi t_{N-2} \}_{j=1}^{N-2}
\]

(8)

form an orthogonal system with respect to the usual \( l_2 \) inner product,

\[
(\varphi_j, \varphi_i) = \sum_{n=1}^{N-2} (\sin j\pi t_n \sin i\pi t_n) = \left\{ \begin{array}{cl} 0, & i \neq j, \\ \frac{1}{2}(N - 1), & i = j, \end{array} \right.
\]

the unique solution is obtained by the method of least squares via the Fourier coefficients formula

\[
\hat{b}_j = \frac{(G, \varphi_j)}{(\varphi_j, \varphi_j)}, \ j = 1, 2, ..., N - 2,
\]

(9)

which provides an inexpensive and efficient method of calculation.

The relationship between the coefficients \( b_j \) in (2) and the Fourier coefficients \( \hat{b}_j \) given by (9) is stated in the following theorem. For a proof, see [6].

**Theorem 1** Let \( g \) be a 2-periodic function that is \( C^1 \) with piecewise smooth second derivative and assume that the coefficients in its series representation satisfy estimate (3). Then there are constants \( C_0 \) and \( C_1 \) such that

\[
\| g - \text{Int}_{N-2} g \|_{\infty, I} \leq C_0 h^2 \quad \text{and} \quad \left\| \frac{d}{dt} g - \frac{d}{dt} \text{Int}_{N-2} g \right\|_{\infty, I} \leq C_1 h.
\]

Theorem 1 establishes, in the absence of noise in the data, the uniform convergence of the trigonometric interpolating polynomial and its derivative to the function \( g \) and its derivative \( g' \) on a uniform (equidistant points) grid of the unit interval, including the boundary points \( t = 0 \) and \( t = 1 \).
2.2. Mollification

In practice, we work with noisy sample data. Instead of \( f(t_i) \) we know \( f^\epsilon(t_i) = f(t_i) \pm \epsilon_i \), a perturbed version of \( f \), where \( \{\epsilon_i\}_{i=0}^{N-1} \) is a sequence of independent Gaussian random variables satisfying
\[
\max_k |\epsilon_i| \leq \epsilon.
\]

Likewise, we should consider the noisy version of \( g \), denoted \( g^\epsilon \), and the gridfunction \( G^\epsilon \) with \( \|G-G^\epsilon\|_{K,\infty} \leq \epsilon \). In this situation the computation of \( \frac{d}{dt} \text{Int}_{N-2} g^\epsilon \) becomes an ill-posed problem and we need to introduce some type of regularization to restore stability with respect to perturbations in the data.

In this section we use the mollification method to stabilize the differentiation problem. Introduce the function
\[
\rho_\delta(t) = \frac{1}{\sqrt{2\pi}} \exp(-t^2/\delta^2),
\]
the one dimensional Gaussian kernel of “blurring” radius \( \delta > 0 \). The kernel \( \rho_\delta \) is always positive, nearly vanishes outside the interval \((-3\delta, 3\delta)\) and \( \int_{-\infty}^{\infty} \rho_\delta(s) \, ds = 1 \).

The convolution of a real function \( w \) with \( \rho_\delta \) is defined by
\[
J_\delta w(t) = (\rho_\delta \ast w)(t) = \int_{-\infty}^{\infty} \rho_\delta(t-s) \, w(s) \, ds \approx \int_{t-3\delta}^{t+3\delta} \rho_\delta(t-s) \, w(s) \, ds.
\]

For any locally integrable function \( w \), \( J_\delta w \) is infinitely differentiable and \( \|J_\delta w - w\|_2 \to 0 \) as \( \delta \to 0 \), the convergence being uniform on any compact set where \( w \) is continuous [2]. The basic idea of the mollification method is that instead of attempting to find the point values of the function \( w \), we attempt to reconstruct the \( \delta \)-mollification (weighted average) of the functions with the high frequency components suitably filtered out and leaving the low frequency components essentially unmodified.

Below we indicate the main properties of the mollification method. Their proofs can be found in [9].

For \( t \in I_\delta = [3\delta, 1 - 3\delta] \), if \( w \) is integrable on \( I \), we define its \( \delta \)-mollification on \( I_\delta \) by the convolution
\[
J_\delta w(t) = \int_0^1 \rho_\delta(t-s) \, w(s) \, ds.
\]

In what follows \( C \) represents a generic constant independent of \( \delta \).

**Lemma 2** (Maximum norm consistency) If \( w \in C^1(I) \), then there exists a constant \( C \) such that
\[
\|J_\delta w - w\|_{\infty, I_\delta} \leq C\delta \quad \text{and} \quad \left\| \frac{d}{dt} J_\delta w - \frac{d}{dt} w \right\|_{\infty, I_\delta} \leq C\delta.
\]

**Lemma 3** (Maximum norm stability) If \( w, w^\epsilon \in C^0(I) \) and \( \|w - w^\epsilon\|_{\infty, I} \leq \epsilon \), then there exists a constant \( C \) such that
\[
\|J_\delta w - J_\delta w^\epsilon\|_{\infty, I_\delta} \leq \epsilon \quad \text{and} \quad \left\| \frac{d}{dt} J_\delta w - \frac{d}{dt} J_\delta w^\epsilon \right\|_{\infty, I_\delta} \leq C\epsilon.
\]

**Theorem 4** (Maximum norm convergence) Under the conditions of Lemmas 2 and 3,
\[
\|J_\delta w^\epsilon - w\|_{\infty, I_\delta} \leq C\delta + \epsilon \quad \text{and} \quad \left\| \frac{d}{dt} J_\delta w^\epsilon - \frac{d}{dt} w \right\|_{\infty, I_\delta} \leq C(\delta + \frac{\epsilon}{\delta}).
\]

We observe that in order to obtain convergence as \( \epsilon \to 0 \), in the first case it suffices to consider \( \delta \to 0 \) but in the second case we need to relate both parameters. For example, we can choose \( \delta = O(\sqrt{\epsilon}) \). This is a consequence of the ill-posedness of differentiation of noisy data.
2.3. Trigonometric Mollification

We assume that the underlying function \( g \), that we are trying to estimate, and the interpolating function \( \text{Int}_{N-2}g \) satisfy the conditions of Theorem 1. Setting \( w = \text{Int}_{N-2}g \) and \( w^{\epsilon} = \text{Int}_{N-2}g^{\epsilon} \) the conditions of Lemmas 2 and 3 are satisfied and the conclusions of Theorem 4 hold with the norm estimates extended to the whole interval \( I = [0,1] \) instead of \( I_3 = [3\delta, 1-3\delta] \).

**Theorem 5** Under the conditions of Theorems 1 and 4, using the triangle inequality,

\[
\| J_3 \text{Int}_{N-2}g^{\epsilon} - g \|_{\infty,I} \leq C(h^{2} + \delta) + \epsilon \quad \text{and} \quad \left\| \frac{d}{dt} J_3 \text{Int}_{N-2}g^{\epsilon} - \frac{d}{dt} g \right\|_{\infty,I} \leq C(h + \delta + \frac{\epsilon}{\delta}).
\]

The evaluation of \( J_3 \text{Int}_{N-2}g^{\epsilon} \), assuming a given fixed radius of mollification \( \delta > 0 \), is obtained by convolving the Gaussian kernel (10) with the Fourier series (6) after replacing \( \hat{b}_j \) by \( (\hat{b'})_j \).

Hence, a short computation gives

\[
J_3 \text{Int}_{N-2}g^{\epsilon}(t) = (\text{Int}_{N-2} g^{\epsilon} \ast \rho_{\delta})(t) = \sum_{j=1}^{N-2} (\hat{b'})_j \int_{-\infty}^{\infty} (\sin j\pi (t-s) \ast \rho_{\delta}(s))ds = \sum_{j=1}^{N-2} (\hat{b'}_{\delta})_j \sin j\pi t,
\]

and we conclude that the Fourier series coefficients of the mollified interpolating trigonometric polynomial are determined by the single multiplication

\[
(\hat{b'}_{\delta})_j = e^{-\frac{j}{4}2^{2}\delta^{2}}(\hat{b'})_j.
\]

The derivative of the mollified trigonometric polynomial is computed directly by

\[
\frac{d}{dt} J_3 \text{Int}_{N-2}g^{\epsilon}(t) = \sum_{j=1}^{N-2} (\hat{b'}_{\delta})_j j\pi \cos j\pi t.
\]

Furthermore, the mollified approximations for the function \( f \) and its first derivative \( \frac{d}{dt}f \) are given, respectively, by

\[
J_3 \text{Int}_{N-2}g^{\epsilon}(t) + [f^{\epsilon}(0) + (f^{\epsilon}(1) - f^{\epsilon}(0))t] \quad \text{and} \quad \frac{d}{dt}(J_3 \text{Int}_{N-2}g^{\epsilon})(t) + (f^{\epsilon}(1) - f^{\epsilon}(0)).
\]

Finally, a word on regularization parameters is appropriate. The mollification radii \( \delta = \delta(\epsilon) \) are chosen automatically at each step without any information about the amount and/or character of the noise in the data. This is done by the Generalized Cross Validation Principle (GCV) presented in [3, 4], applied to the discrete linear operator associated to discrete mollification.

The idea is to balance the amount of regularization with the amount of noise in the data. Let the discrete function defined on \( K \) be \( g^{\epsilon} \) and let the finite dimensional mollification operator be \( H_{\delta} \). The right amount of mollification is achieved by using \( \delta_0 \) as radius of mollification, where

\[
\delta_0 = \arg\min_{\delta} \frac{N \| H_{\delta} g^{\epsilon} - g^{\epsilon} \|_2^2}{(\text{Trace}(I - H_{\delta}))^{\frac{1}{2}}}.\]
Table 1. Relative $l_2$ errors as functions of $\epsilon$.

| $N = 128$ | Functions | Derivatives |
|-----------|-----------|-------------|
| $\epsilon$ | Example 1 | Example 2 | Example 1 | Example 2 |
| 0.00      | 0.0053    | 0.0562      | 0.0268 | 0.0798 |
| 0.01      | 0.0092    | 0.0564      | 0.0480 | 0.0812 |
| 0.05      | 0.0309    | 0.0612      | 0.0897 | 0.0849 |
| 0.10      | 0.0667    | 0.0748      | 0.1667 | 0.0988 |

3. Numerical Results

In this section we present numerical results from two illustrative examples.

The use of average perturbation values $\epsilon$ is only necessary for the purpose of generating the noisy data for the simulations. Discretized measured approximations of the data are simulated by adding random errors to the exact data functions and the filtering procedure automatically adapts the regularization parameter to the quality of the data.

In order to test the stability and accuracy of the algorithm, we consider a selection of average noise perturbations $\epsilon$, and number of points $N$. The derivative errors are measured by the weighted $l_2$ norms defined as follows:

$$\left[ \frac{1}{N} \sum_{i=0}^{N-1} \left| \frac{d}{dt} (J_3 \text{Int}_{N-2} f^\epsilon)(t_i) - \frac{d}{dt} f(t_i) \right|^2 \right]^{1/2}.$$

The tables were prepared with $N = 64, 128$ and $256$ and maximum levels of noise, $\epsilon = 0.00, 0.01, 0.05$ and $0.1$.

**Example 1.** This prototype example represents an exponential approximation to a delayed instantaneous pulse. The equation for the exact data function is

$$f(t) = \exp \left( -40(t - 0.5)^2 \right), \; 0 \leq t \leq 1.$$

The numerical results showing the relative errors in the approximations, related with the amount of noise in the data and number of data points can be found in tables 1 and 2 respectively. A graphical comparison of the exact data function and its noisy version corresponding to $\epsilon = 0.05$ appears in figure 1. In figure 2 we show the graphs of the exact derivative function $\frac{df}{dt}$ and the computed derivative $\frac{d}{dt} J_3 \text{Int}_{N-2} f^\epsilon$ corresponding to the mentioned level of noise.

**Example 2.** Our second example consists on the approximation of the derivative of the function

$$f(t) = \sin(10\pi t), \; 0 \leq t \leq 1.$$

In this case, both $f$ and $\frac{df}{dt}$ are highly oscillatory. The global results, relating the variation of the relative errors with respect to $\epsilon$ and $N$, are shown in tables 1 and 2. The level of perturbation in the data and the mollified function can be observed in figure 3. Figure 4 exhibits the qualitative behavior of the computed solution.
Table 2. Relative $l_2$ errors as functions of $N$.

| $N$ | $\epsilon = 0.05$ Functions Example 1 | Example 2 | Derivatives Example 1 | Example 2 |
|-----|--------------------------------------|-----------|----------------------|-----------|
| 64  | 0.0254                              | 0.0587    | 0.1423               | 0.0879    |
| 128 | 0.0309                              | 0.0612    | 0.0897               | 0.0849    |
| 256 | 0.0322                              | 0.0296    | 0.1004               | 0.0678    |

Figure 1. Example 1 Data.

Figure 2. Example 1 Results.

Figure 3. Example 2 Data.

Figure 4. Example 2 Results.

Acknowledgments
D. Murio acknowledges the partial support from a C. P. Taft Fellowship.
C. Mejía acknowledges the partial support from Universidad Nacional de Colombia through Vicerrectoría de Investigación and DIME.
References

[1] Murio D A 1993 *The Mollification Method and the Numerical Solution of Ill-Posed Problems* (New York: Wiley (Interscience))

[2] Murio D A 2002 Chapter 4 in *Inverse Engineering Handbook* ed Woodbury K (Boca Raton, Florida: CRC Press) pp 220–316

[3] Craven P and Wahba G 1979 *Numer. Math.* 31 377–403

[4] Golub G, Heath M and Wahba G 1979 *Technometrics* 21 215–223

[5] Lanczos C 1938 *J. Math. Phys.* 17 123–99

[6] Gustafsson B, Kreiss O and Oliger J 1995 *Time dependent problems and difference methods* (New York: Wiley)

[7] Murio D A and Mejía C E 2008 Generalized time fractional ihcp with caputo fractional derivatives submitted

[8] Murio D A 1987 *Comput. Math. Appl.* 13 381–6

[9] Murio D A, Mejía C E and Zhan S 1998 *Comput. Math. Appl.* 35 1–16

[10] Hansen P C 1998 *Rank-Deficient and Discrete Ill-Posed Problems. Numerical Aspects of Linear Inversion* (Philadelphia: SIAM)