

Variational and Topological Methods: Theory, Applications, Numerical Simulations, and Open Problems (2012). *Electronic Journal of Differential Equations*, Conference 21 (2014), pp. 235–246. ISSN: 1072-6691. <http://ejde.math.txstate.edu>, <http://ejde.math.unt.edu> <ftp://ejde.math.txstate.edu>

METHODS FOR NUMERICAL DIFFERENTIATION OF NOISY DATA

IAN KNOWLES, ROBERT J. RENKA

ABSTRACT. We describe several methods for the numerical approximation of a first derivative of a smooth real-valued univariate function for which only discrete noise-contaminated data values are given. The methods allow for both uniformly distributed and non-uniformly distributed abscissae. They are compared for accuracy on artificial data sets constructed by adding Gaussian noise to simple test functions. We also display results associated with an experimental data set.

1. INTRODUCTION

The problem of approximating a derivative of a function defined by error-contaminated data points arises in several scientific computing and engineering disciplines. Among the applications are image reconstruction, probability density estimation, satellite orbit determination, and seismic profiling. The problem is notable for the large variety of methods that have been developed for its treatment. These include polynomial regression, spline smoothing, filtering with Fourier and wavelet transforms, total variation denoising, and convolution with a mollifier, as well as local methods and other global methods. We describe the problem in Section 2, discuss methods in Section 3, and present test results in Section 4.

2. THE PROBLEM

Given a set of m discrete data points (observations) $\{(x_i, y_i)\}$ taken from a smooth function $g : [a, b] \rightarrow \mathbb{R}$ with

$$a \leq x_1 \leq x_2 \leq \cdots \leq x_m \leq b, \quad (2.1)$$

we wish to construct an approximation u to the first derivative g' . The differentiation problem is the inverse of the problem of computing an integral. As with most inverse problems, it is ill-posed because the solution $u \approx g'$ does not depend continuously on g . If we use divided difference approximations to derivative values, then arbitrarily small relative perturbations in the data can lead to arbitrarily large

2000 *Mathematics Subject Classification*. 65D10, 65D25, 65R32.

Key words and phrases. Ill-posed problem; numerical differentiation; smoothing spline;

Tikhonov regularization; total variation.

©2014 Texas State University - San Marcos.

Published February 10, 2014.

relative changes in the solution, and the discretized problem is ill-conditioned. We therefore require some form of regularization in order to avoid overfitting.

In addition to constructing the derivative u , we compute a smooth function f which approximates g . If u is computed directly from the data, it may be integrated to produce f . Alternatively, we may view the problem as that of constructing a smooth approximation f which may then be differentiated to obtain $u = f'$. We assume that the data values are contaminated by independent identically distributed zero-mean noise values η_i :

$$y_i = g(x_i) + \eta_i, \quad E[\eta_i] = 0, \quad (i = 1, \dots, m).$$

Treating values of the fitting function $f(x)$ as random variables, the mean squared error at each point $x \in [a, b]$ is

$$\text{MSE}(x) = E[(g(x) - f(x))^2] = (g(x) - E[f(x)])^2 + E[(f(x) - E[f(x)])^2],$$

where the first term is the squared bias and the second term is the variance. A small error requires that both terms be small, but generally there is a tradeoff between the two terms. We can minimize the bias by choosing a very flexible model with a large number of free parameters — enough freedom to interpolate the data points as the extreme case. However, this would result in large variance due to too much sensitivity to the data. The fit would be poor on another sample data set. The inflexibility of too few parameters, on the other hand, would result in large error due to large bias. Most smoothing methods incorporate the bias-variance tradeoff into a single smoothing parameter which must be chosen to achieve the proper balance.

3. METHODS

We implemented and tested one or more methods in each of the following categories:

- (i) Least squares polynomial approximation.
- (ii) Tikhonov regularization.
- (iii) Smoothing spline.
- (iv) Convolution smoothing with a Friedrichs mollifier.
- (v) Knowles and Wallace variational method.
- (vi) Total variation regularization.

Each category is described in a separate subsection below.

3.1. Least squares polynomial approximation. A simple smoothing method consists of fitting the data points in a least squares sense with a sequence of polynomials f of degree d for $d = 0, 1, \dots, k$, where k is the smallest value for which the residual norm is bounded by a tolerance. We then define $u = f'$. The reciprocal of the polynomial degree serves as a regularization parameter with $k \geq m - 1$ corresponding to interpolation, and $k = 0$ corresponding to a constant. This method requires that the data abscissae be distinct, and produces $f, u \in C^\infty[a, b]$ for $a = x_1$ and $b = x_m$.

The polynomial is taken to be a linear combination of the monomial basis functions $\phi_j(x) = x^j$ for $j = 0, \dots, k$, and the coefficients are computed by solving the normal equations with a QR factorization of the Vandermonde matrix.

We implemented only the simple global method, but a method that could be more effective for dense data sets is locally weighted polynomial regression, referred to

as LOESS or LOWESS [3]. The idea is to fit a low-degree polynomial (degree 1 or 2) to the nearest neighbors of each data point with weight proportional to inverse distance from the abscissa of the neighbor to that of the point. The number of points used in each fit increases with the value of a smoothing parameter. Another method that uses local polynomial regression (for uniformly distributed data points) is the Savitsky-Golay smoothing filter [17].

3.2. Tikhonov regularization. Tikhonov regularization was introduced in [20] and first applied to the numerical differentiation problem in [5]. Refer also to [21]. We include three methods, indexed by $k = 0, 1$, and 2 , corresponding to the smoothness of the solution u . We assume that $g(a)$ is given and that g is in the Sobolev space $H^{k+1}(a, b)$. We formulate the problem as that of computing an approximate solution $u \in H^k(a, b)$ to the operator equation

$$Au(x) = \int_a^x u(t) dt = \hat{g}(x), \quad x \in [a, b], \quad (3.1)$$

where $\hat{g}(x) = g(x) - g(a)$. Once u is computed, f can be obtained from $f(x) = \int_a^x u(t) dt + g(a)$.

We discretize the problem by representing f as a vector \mathbf{f} of function values on a uniform grid that partitions $[a, b]$ into n subintervals of length $\Delta t = (b - a)/n$:

$$f_j = f(t_j), \quad t_j = a + (j - 1)\Delta t, \quad (j = 1, \dots, n + 1). \quad (3.2)$$

The function u is represented by derivative values at the midpoints:

$$u_j = f'(t_j + \Delta t/2), \quad (j = 1, \dots, n).$$

The discretized system is $A\mathbf{u} = \hat{\mathbf{y}}$, where $\hat{y}_i = y_i - g(a)$ and A_{ij} is the length of $[a, x_i] \cap [t_j, t_{j+1}]$:

$$A_{ij} = \begin{cases} 0 & \text{if } x_i \leq t_j \\ x_i - t_j & \text{if } t_j < x_i < t_{j+1} \\ \Delta t & \text{if } t_{j+1} \leq x_i. \end{cases}$$

This linear system may be underdetermined or overdetermined and is likely to be ill-conditioned. We therefore use a least squares formulation with Tikhonov regularization. We minimize the convex functional

$$E(\mathbf{u}) = \|A\mathbf{u} - \hat{\mathbf{y}}\|^2 + \alpha \|D\mathbf{u}\|^2, \quad (3.3)$$

where α is a nonnegative regularization parameter, $\|\cdot\|$ denotes the Euclidean norm, and D is a differential operator of order k defining a discretization of the H^k Sobolev norm:

$$D^t = \begin{cases} I & \text{if } k = 0 \\ (I D_1^t) & \text{if } k = 1 \\ (I D_1^t D_2^t) & \text{if } k = 2 \end{cases}$$

where I denotes the identity matrix, and D_1 and D_2 are first and second difference operators. We use second-order central differencing so that D_1 maps midpoint values to interior grid points, and D_2 maps midpoint values to interior midpoint values. The regularization (smoothing) parameter α defines a balance between fidelity to the data (with high fidelity corresponding to low bias) on the one hand, and the size of the solution norm, which is related to variance, on the other hand. The optimal value depends on the choice of norm. Larger values of k enforce more

smoothness on the solution. Setting the gradient of E to zero, we obtain a linear system with an order- n symmetric positive definite matrix:

$$(A^t A + \alpha D^t D) \mathbf{u} = A^t \hat{\mathbf{y}}.$$

In the case that the error norm $\|\eta\|$ is known, a good value of α is obtained by choosing it so that the residual norm $\|A\mathbf{u} - \hat{\mathbf{y}}\|$ agrees with $\|\eta\|$. This is Morozov's discrepancy principle [9].

3.3. Smoothing spline. The cubic smoothing spline was introduced by Schoenberg [18] and Reinsch [12, 13]. The method of generalized cross validation for automatic selection of a smoothing parameter first appeared in [4]. The cubic spline was generalized to a spline under tension by Schweikert [19], and methods for automatic selection of tension factors were developed in [14]. This method is effective for relatively sparse data sets.

For this method we take $m = n$, and require that the abscissae are distinct and include the endpoints:

$$a = x_1 < x_2 < \cdots < x_n = b.$$

The smoothing function f is a spline with knots at the abscissae. Denote the knot values by $v_i = f(x_i)$, and for each subinterval $[x_i, x_{i+1}]$, denote the subinterval length by Δx_i , denote the slope of the piecewise linear interpolant by $s_i = (v_{i+1} - v_i)/\Delta x_i$, and let τ_i be a nonnegative tension factor associated with the subinterval. Then our measure of smoothness for $f \in H^2[a, b]$ is

$$q_1(f) = \int_a^b f''^2 + \sum_{i=1}^{n-1} \left(\frac{\tau_i}{\Delta x_i} \right)^2 \int_{x_i}^{x_{i+1}} (f' - s_i)^2.$$

For a given set of knot values $\{v_i\}$ the minimizer of q_1 is in $C^2[a, b]$ and satisfies

$$f^{(4)} - (\tau_i/\Delta x_i)^2 f'' = 0 \tag{3.4}$$

on each subinterval so that f is cubic in subintervals for which $\tau_i = 0$, and f approaches the linear interpolant of (x_i, v_i) and (x_{i+1}, v_{i+1}) as $\tau_i \rightarrow \infty$. Solutions to (3.4) lie in the span of $\{1, x, e^{(\tau_i/\Delta x_i)x}, e^{-(\tau_i/\Delta x_i)x}\}$ so that f is piecewise exponential. Using a Hermite interpolatory representation of f , the four degrees of freedom in each subinterval are taken to be the endpoint function values v_i, v_{i+1} and first derivative values $d_i = f'(x_i), d_{i+1} = f'(x_{i+1})$. Let q_1 now denote the quadratic functional defined on the pair of n -vectors of knot function values \mathbf{v} and derivatives \mathbf{d} .

We obtain a smoothing spline by minimizing q_1 subject to the constraint

$$q_2(\mathbf{v}) = \sum_{i=1}^n \left(\frac{y_i - v_i}{\sigma_i} \right)^2 \leq S,$$

where σ_i is the standard deviation in y_i and S is a nonnegative number with nominal value in the confidence interval $[n - \sqrt{2n}, n + \sqrt{2n}]$ for the weighted sum of squared deviations from the data values $q_2(\mathbf{v})$. For S sufficiently large, f is linear and $q_1 = 0$. For an active constraint the problem is equivalent to minimizing

$$E(\mathbf{v}, \mathbf{d}, \lambda) = q_1(\mathbf{v}, \mathbf{d}) + \lambda(q_2(\mathbf{v}) - S)$$

for Lagrange multiplier λ (whose reciprocal serves as a regularization parameter). This problem is treated by finding a zero of

$$\phi(\lambda) = \frac{1}{\sqrt{q_2(\mathbf{v})}} - \frac{1}{\sqrt{S}},$$

where \mathbf{v} and \mathbf{d} satisfy the order- $2n$ symmetric positive definite linear system obtained by setting the gradient of E to zero with λ fixed.

Our test suite includes two methods in this category: a cubic spline in which all tension factors are zero, and a tension spline with just enough tension in each subinterval to preserve local monotonicity and convexity of the data [14, 15].

3.4. Convolution smoothing with a Friedrichs mollifier. We compute a smooth approximation $f \in C^\infty[a, b]$ by convolution of the piecewise linear interpolant p of the data values with the positive symmetric Friedrichs mollifier function [6]

$$\rho(x) = \begin{cases} ce^{1/(x^2-1)} & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

where c is chosen so that ρ integrates to 1. Then for $x \in [a, b]$,

$$f(x) = \frac{1}{h} \int_{a-h}^{b+h} \rho\left(\frac{x-s}{h}\right) p(s) ds$$

for small positive h . The derivative g' is then approximated by $u = f'$. Note that h is a smoothing parameter: $f \rightarrow p$ in $L^r(a, b)$, $1 \leq r < \infty$, as $h \rightarrow 0$ (interpolation), and $f \rightarrow 0$ as $h \rightarrow \infty$.

We represent the smoothed function f by a discrete set of values on a uniform grid with mesh width $\Delta t = (b-a)/(n-1)$:

$$f_i = f(t_i), \quad t_i = a + (i-1)\Delta t \quad \text{for } i = 1, \dots, n.$$

We use second-order central difference approximations to midpoint first derivative values

$$f'(t_i + \Delta t/2) = (f_{i+1} - f_i)/\Delta t \quad \text{for } i = 1, \dots, n-1.$$

This method has the limitation that the abscissae must be distinct, and the approximation is valid only on the subdomain $[a+h, b-h]$. To our knowledge, this novel numerical application of the well-known Friedrichs mollifier was first suggested by Aimin Yan as a way to handle the numerical differentiation of a measurement dataset arising from a function of two variables. This was needed for the solution of the inverse groundwater problem in [8], and elsewhere, and found to be quite effective in those contexts.

3.5. Knowles and Wallace variational method. The method of Knowles and Wallace [7] was designed with the goal of eliminating the difficult problem of determining an effective regularization parameter when nothing is known about the errors in the data. Let $p(x) = g(x) + k$ for $x \in [a, b]$ and a constant k large enough that p is bounded below by a positive constant and $Q(x) = p''(x)/p(x)$ is bounded above by a curvature-limiting constant $M < (\pi/(b-a))^2$ for $x \in [a, b]$. Then $u = g' = p'$ can be obtained by solving $-p'' + Qp = 0$ with specified endpoint

values of p , where Q is the unique global minimizer of the strictly convex nonlinear energy functional

$$E(q) = \int_a^b (p' - f')^2 + q(p - f)^2,$$

with f as the solution to the two-point boundary value problem

$$\begin{aligned} A_q f &= -f'' + qf = 0 \quad \text{in } (a, b), \\ f(a) &= p(a), \quad f(b) = p(b). \end{aligned}$$

Note that $E(Q) = 0$ and $f = p$ so that $f - k = g$ at the solution. Since g is specified only by discrete data points, possibly contaminated by noise, using the data to discretize E will not result in a zero at the minimizer, but the computed function f will be a smooth approximation to p and can be used to compute u . The idea is thus to compute the best approximation to p in the space of solutions to the two-point boundary value problem.

An equivalent expression for the functional is

$$E(q) = \int_a^b p'^2 + qp^2 - (f'^2 + qf^2).$$

The gradient of E is defined by the first Gateaux derivative

$$E'(q)[h] = \int_a^b (p^2 - f^2)h,$$

and the Hessian is defined by the second derivative

$$E''(q)[h][k] = 2 \int_a^b (fk)A_q^{-1}(fh),$$

where A_q is restricted to functions satisfying homogeneous Dirichlet boundary conditions.

In order to discretize the problem, we take p to be the piecewise linear interpolant of $\{(x_i, y_i + k)\}$, and we represent functions p, f , and q as n -vectors of grid-point values on the uniform grid with mesh width $\Delta t = (b - a)/(n - 1)$:

$$f_i = f(t_i), \quad t_i = a + (i - 1)\Delta t \quad (i = 1, \dots, n).$$

We use second-order central difference approximations to midpoint first derivative values

$$(D_1 \mathbf{f})_i = f'(t_i + \Delta t/2) = (f_{i+1} - f_i)/\Delta t \quad (i = 1, \dots, n - 1).$$

We then have second-order approximations to $A_q \mathbf{f}$ at the interior grid points

$$(A_q \mathbf{f})_i = \frac{-f_{i-1} + 2f_i - f_{i+1}}{\Delta t^2} + q_i f_i$$

for $i = 2, \dots, n - 1$ with $f_1 = p(a)$ and $f_n = p(b)$. The solution to the two-point boundary value problem is $\mathbf{f} = A_q^{-1} \mathbf{c}$ for order- $(n-2)$ matrix $A_q = \text{diag}(\mathbf{q}) - D_2$ and $\mathbf{c} = (p(a)/\Delta t^2)\mathbf{e}_1 + (p(b)/\Delta t^2)\mathbf{e}_{n-2}$, where D_2 denotes the second difference operator, and \mathbf{e}_1 and \mathbf{e}_{n-2} are standard basis vectors. We approximate $E(\mathbf{q})$ with a rectangle rule for the midpoint derivative values (first term) and a trapezoidal rule for the interior grid-point values (second term):

$$E(\mathbf{q}) = \frac{1}{\Delta t} \sum_{i=1}^{n-1} (w_{i+1} - w_i)^2 + \Delta t \sum_{i=2}^{n-1} q_i w_i^2$$

for $\mathbf{w} = \mathbf{p} - \mathbf{f}$. A Newton-like method for minimizing E is not feasible because every component of the Hessian would require solution of a linear system. We therefore use a gradient descent iteration, but not with the ordinary gradient. The standard method of steepest descent is notoriously slow, but we obtain an effective method by using the Sobolev gradient [10]:

$$\mathbf{q}_{k+1} = \mathbf{q}_k - \beta \nabla_S E(\mathbf{q}_k),$$

where β denotes a positive step size, and $\nabla_S E(\mathbf{q})$ is the H^1 Sobolev gradient of E at \mathbf{q} defined by

$$\nabla_S E(\mathbf{q}) = (I + D_1^t D_1)^{-1} \nabla E(\mathbf{q})$$

for Euclidean gradient $\nabla E(\mathbf{q})$ with components $p_i^2 - f_i^2$. Note that D_1 is restricted to the interior grid points, and $I + D_1^t D_1 = I - D_2$. The inverse of this matrix is a smoothing operator which serves as a preconditioner for the descent iteration.

3.6. Total variation regularization. In some applications it may be necessary to allow for discontinuities in the derivative u corresponding to corners in g . To this end, we assume only that $g \in L^2[a, b]$ and we take the measure of regularity to be the total variation in u : $F(u) = \int_a^b |u'|$ for u in the space $BV[a, b]$ of functions of bounded variation. The method of total variation regularization was introduced by Rudin, Osher, and Fatemi [16] for the purpose of removing noise from images without smearing edges. It was applied to the problem of differentiating noisy data by Chartrand [2] using an iterative method of Vogel and Oman [22]. Here we use a generalized Sobolev gradient method to minimize the energy functional.

As in the case of Tikhonov regularization we assume that $g(a)$ is given and we seek an approximate solution u to the operator equation

$$Au(x) = \int_a^x u(t) dt = \hat{g}(x), \quad x \in [a, b],$$

where $\hat{g}(x) = g(x) - g(a)$. We again represent u by a vector \mathbf{u} of discrete values at the midpoints of a uniform grid:

$$u_j = f'(a + (j-1)\Delta t + \Delta t/2), \quad (j = 1, \dots, n)$$

for $\Delta t = (b-a)/n$. The discretized system consists of m equations in n unknowns $A\mathbf{u} = \hat{\mathbf{y}}$, where A_{ij} is the length of $[a, x_i] \cap [t_j, t_{j+1}]$ and $\hat{y}_i = y_i - g(a)$. The energy functional is

$$E(\mathbf{u}) = \frac{1}{2} \|A\mathbf{u} - \hat{\mathbf{y}}\|^2 + \alpha \sum_{j=1}^{n-1} \sqrt{(u_{j+1} - u_j)^2 + \epsilon}, \quad (3.5)$$

where $\|\cdot\|$ denotes the Euclidean norm, and α is a nonnegative regularization parameter. The regularization term is the discretization of the BV seminorm $F(u) = \int_a^b |u'|$ using the trapezoidal rule and natural end conditions: $u'(a) = u'(b) = 0$. Since the gradient of F is $\nabla F(u) = -(u'/|u'|)'$, the perturbation by a small positive number ϵ is necessary for differentiability of E where $u' = 0$.

The gradient of E is

$$\nabla E(\mathbf{u}) = A^t(A\mathbf{u} - \hat{\mathbf{y}}) - \alpha \mathbf{s}',$$

where the sign vector \mathbf{s} is the approximation to $u'/|u'|$:

$$s_i = (u_{i+1} - u_i) / \sqrt{(u_{i+1} - u_i)^2 + \epsilon} \quad (i = 1, \dots, n-1),$$

and

$$(s')_1 = s_1, \quad (s')_i = s_i - s_{i-1} \quad (i = 2, \dots, n-1), \quad (s')_n = -s_{n-1}.$$

The Hessian of E at \mathbf{u} is

$$H(\mathbf{u}) = A^t A + \alpha D_u^t (I - \text{diag}(\mathbf{s})^2) D_u,$$

where D_u is the discretization of the operator that maps v to $v'/\sqrt{|u'|}$, and $1 - s_i^2 = \epsilon/((u_{i+1} - u_i)^2 + \epsilon)$. We approximate $H(\mathbf{u})$ by the preconditioner

$$A_u = A^t A + \alpha \Delta t D_u^t D_u.$$

The Sobolev gradient $A_u^{-1} \nabla E(\mathbf{u})$ is the gradient associated with the weighted Sobolev inner product

$$\langle \mathbf{v}, \mathbf{w} \rangle_u = (A\mathbf{v})^t (A\mathbf{w}) + \alpha \Delta t (D_u \mathbf{v})^t (D_u \mathbf{w}).$$

The steepest descent iteration is

$$\mathbf{u}_{k+1} = \mathbf{u}_k - \beta A_u^{-1} \nabla E(\mathbf{u}_k)$$

with constant step size β , and initial value \mathbf{u}_0 obtained by differencing y values. Since the gradient and inner product at step k depend on \mathbf{u}_k , this is a variable-metric method. The gridpoint values of \mathbf{f} are computed as

$$f_j = \Delta t \sum_{i=1}^{j-1} u_i + g(a) \quad (j = 1, \dots, n+1).$$

4. TEST RESULTS

Our first test function is $g(x) = \cos(x)$ on $[-.5, .5]$. We generated two data sets, one with $m = 100$ points, and one with $m = 10$ points, and both with uniformly distributed abscissae x_i and data values $y_i = g(x_i) + \eta_i$, where η_i is taken from a normal distribution with mean 0 and standard deviation σ . We tested with $\sigma = 0.01$ on both data sets, and $\sigma = 0.1$ on the dense data set only. For the Tikhonov regularization method we used the known value of $\|\eta\|$ to compute the optimal parameter value α . In the case of smoothing splines we used the known value of σ to define the smoothing parameter. Table 1 displays the maximum relative error in the computed derivative for most of the methods. We omitted the tension spline because the results are almost identical to those of the cubic spline, and we omitted the total variation regularization method because it is not designed to fit data from a smooth function.

Method	$m = 100, \sigma = .01$	$m = 100, \sigma = .1$	$m = 10, \sigma = .01$
Degree-2 polynomial	.0287	.3190	.2786
Tikhonov, $k = 0$.7393	.8297	.7062
Tikhonov, $k = 1$.1803	.3038	.6420
Tikhonov, $k = 2$.0186	.0301	.4432
Cubic spline	.1060	1.15	.3004
Convolution smoothing	.1059	.8603	.2098
Variational method	.1669	.7149	.3419

TABLE 1. Relative errors in derivative approximations

The best method on the dense data set is clearly Tikhonov regularization with $k = 2$. The solution is graphed in Figure 1. Note that the curvature is exaggerated by the mismatched axis scaling. The large error in the cubic spline with $\sigma = .1$ is the result of fitting with a polynomial of degree 1 because the constraint was not active. The last three methods outperform Tikhonov regularization on the sparser data set, with convolution smoothing producing the smallest error. The fitting function is graphed in Figure 2. The polynomial of degree 2 does reasonably well on both data sets for this particular test function.

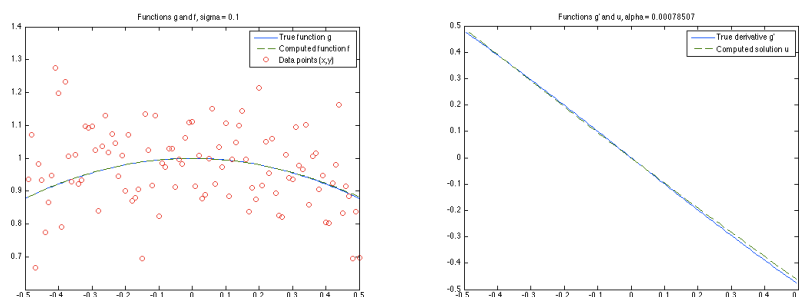


FIGURE 1. Tikhonov regularization, $k = 2, g(x) = \cos(x)$

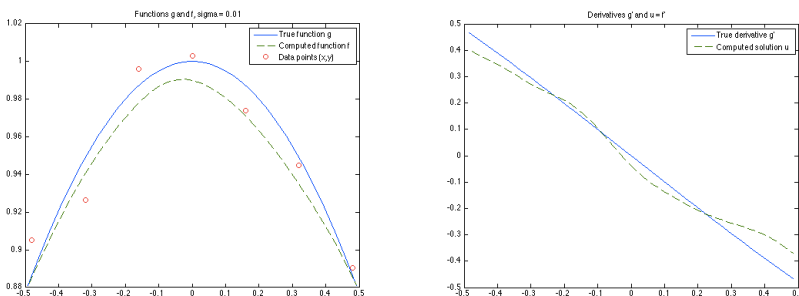


FIGURE 2. Convolution smoothing, $h = .3, g(x) = \cos(x)$

To demonstrate the benefits of tension factors we chose test function $g(x) = x^7$ on $[-1, 1]$ and created a data set consisting of $m = 10$ points with a uniform random distribution of the abscissae and data values with Gaussian noise, $\sigma = .01$. Figure 3 depicts derivative curves associated with zero tension on the left, and shape-preserving tension factors on the right.

Figure 4 demonstrates the effectiveness of the total variation regularization method for test function $g(x) = |x - .5|$ on $[0, 1]$. The data set is again 100 uniformly distributed points with $\sigma = .01$.

The variational method of Knowles and Wallace was compared with the spectral method of Anderssen and Bloomfield [1] on a data set provided by Pallaghy and Lüttge [11] in [7]. The data set consists of 101 points with uniformly distributed

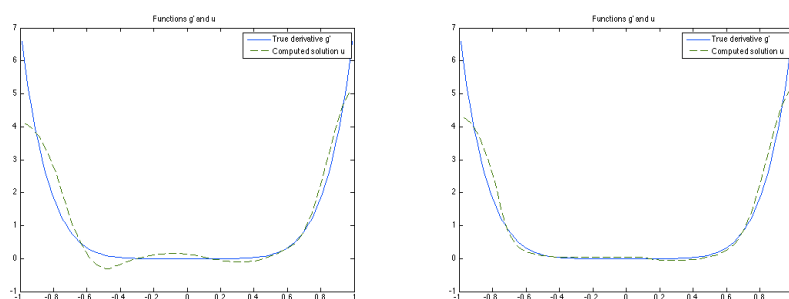


FIGURE 3. Smoothing splines with no tension (left), nonzero tension (right)

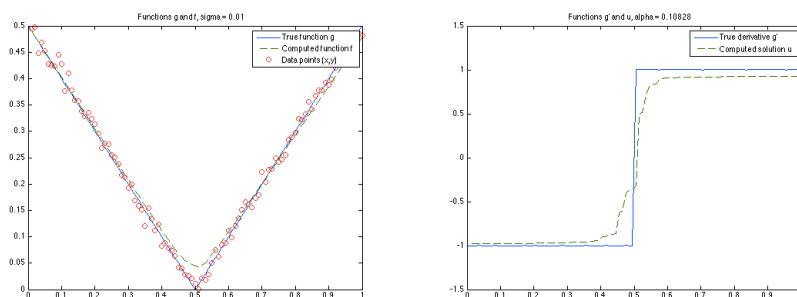


FIGURE 4. Total variation regularization: $g(x) = |x - .5|$

abscissae. Values and central difference approximations to derivative values are depicted in Figure 5.

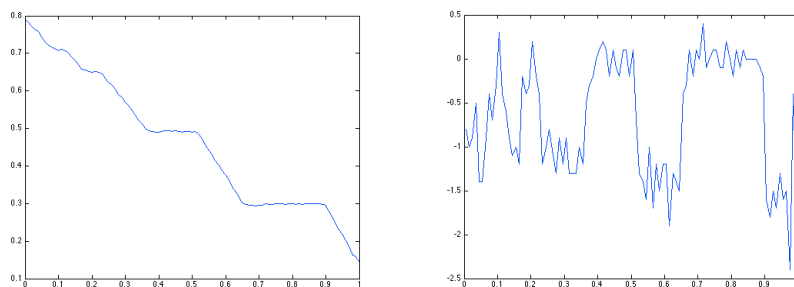


FIGURE 5. Pallaghy and Lüttge data and derivative approximations

Figures 6, 7 and 8 depict derivatives computed by some of the methods described in the previous section. Parameters were chosen experimentally to balance smoothness against fidelity to the data. Only a few tests were required to select reasonable values: tolerance .05 resulting in degree 14 for the polynomial on the left side of

Figure 6; parameters $\alpha = 10^{-3}$, $\alpha = 10^{-7}$, and $\alpha = 10^{-10}$ for Tikhonov regularization with $k = 0$, $k = 1$, and $k = 2$, respectively; standard deviation $\sigma = 3 \times 10^{-3}$ for the smoothing spline on the left side of Figure 8; and $h = .06$ for the convolution method on the right side. The exponential tension spline with optimal tension factors produced a fit nearly identical to that of the cubic spline. Refer to [7] for results with the variational method. The total variation method was not tested on this data set.

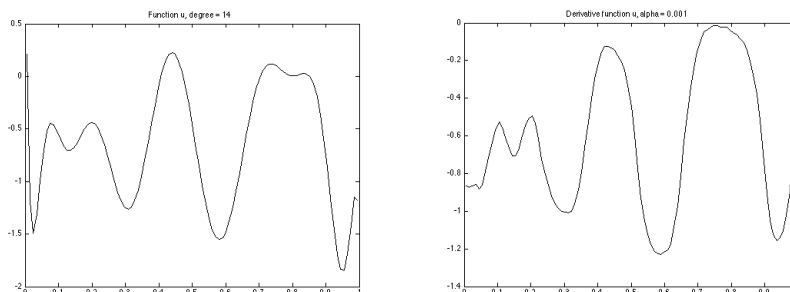


FIGURE 6. Polynomial (left), Tikhonov regularization, $k = 0$ (right)

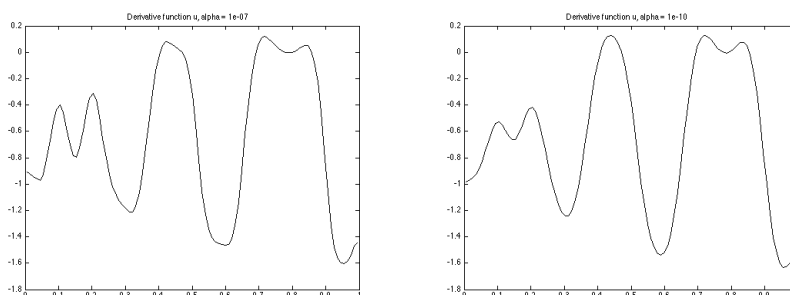


FIGURE 7. Tikhonov regularization with $k = 1$ (left), $k = 2$ (right)

REFERENCES

- [1] R. S. Anderssen, P. Bloomfield; *Numerical differentiation procedures for non-exact data*, Numer. Math. 22 (1973), pp. 157–182.
- [2] R. Chartrand; *Numerical differentiation of noisy nonsmooth data*, ISRN Appl. Math., vol. 2011, doi:10.5402/2011/164564, 2011.
- [3] W. S. Cleveland, S. J. Devlin; *Locally-weighted regression: an approach to regression analysis by local fitting*, Journal of the American Statistical Association 83 (1988), pp. 596–610.
- [4] P. Craven, G. Wahba, *Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross validation*, Numer. Math., 31 (1979), pp. 377–403.
- [5] J. Cullum; *Numerical differentiation and regularization*, SIAM J. Numer. Anal., 8 (1971), pp. 254–265.
- [6] K. O. Friedrichs; *The identity of weak and strong extensions of differential operators*, Trans. of the Amer. Math. Soc., 55 (1944), pp. 132–151.

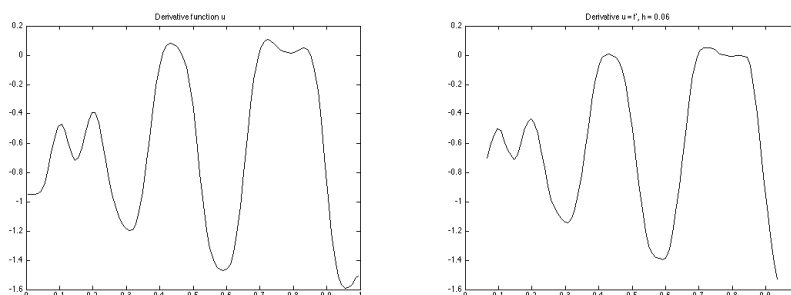


FIGURE 8. Smoothing spline (left), convolution (right)

- [7] I. Knowles, R. Wallace; *A variational method for numerical differentiation*, Numer. Math., 70 (1995), pp. 91–110.
- [8] I. Knowles, T. Le, A. Yan; *On the recovery of multiple flow parameters from transient head data*, J. Comp. Appl. Math., 169 (2004), pp. 1–15.
- [9] V. A. Morozov; *Choice of a parameter for the solution of functional equations by the regularization method*, Sov. Math. Doklady, 8 (1967), pp. 1000–1003.
- [10] J. W. Neuberger; *Sobolev Gradients and Differential Equations*, Lecture Notes in Mathematics, Springer Berlin Heidelberg, 2nd ed., 2010.
- [11] C. K. Pallaghy, U. Lüttge; *Light-induced and H⁺-ion fluxes and bioelectric phenomena in mesophyll cells of Atriplex spongiosa*, Zeit. für Pflanz., 62 (1970), pp. 417–425.
- [12] C. H. Reinsch; *Smoothing by spline functions*, Numer. Math., 10 (1967), pp. 177–183.
- [13] C. H. Reinsch; *Smoothing by spline functions. II*, Numer. Math., 16 (1971), pp. 451–454.
- [14] R. J. Renka; *Interpolatory tension splines with automatic selection of tension factors*, SIAM J. Sci. Stat. Comp., 8 (1987), pp. 393–415.
- [15] R. J. Renka; *Algorithm 716. TSPACK: Tension spline curve fitting package*, ACM Trans. Math. Softw., 19 (1993), pp. 81–94.
- [16] L. Rudin, S. Osher, E. Fatemi; *Nonlinear total variation based noise removal algorithms*, Physica D, 60 (1992), pp. 259–268.
- [17] A. Savitsky, M. J. E. Golay; *Smoothing and differentiation of data by simplified least squares procedures*, Analytical Chemistry, 36 (1964), pp. 1627–1639.
- [18] I. J. Schoenberg; *Spline functions and the problem of graduation*, Proc. Amer. Math. Soc., 52 (1964), pp. 947–950.
- [19] D. G. Schweikert; *An interpolation curve using a spline in tension*, J. Math. Phys., 45 (1966), pp. 312–317.
- [20] A. N. Tikhonov; *Regularization of incorrectly posed problems*, Sov. Math. Dokl., 4 (1963), pp. 1624–1627.
- [21] A. E. Tikhonov, V. Y. Arsenin; *Solutions of Ill-posed Problems*, John Wiley & Sons, New York, 1977.
- [22] C. R. Vogel, M. E. Oman; *Iterative methods for total variation denoising*, SIAM J. Sci. Comput., 17 (1996), pp. 227–238.

IAN KNOWLES

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ALABAMA AT BIRMINGHAM, BIRMINGHAM, AL 35294-1170, USA

E-mail address: knowles@math.uab.edu

ROBERT J. RENKA

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING, UNIVERSITY OF NORTH TEXAS, DENTON, TX 76203-1366, USA

E-mail address: robert.renka@unt.edu