Philippine Journal of Science
149 (2): 431-449, June 2020
ISSN 0031 - 7683
Date Received: 30 Jan 2020

Variational Approach to Data Graduation

James Lee Manejero and **Renier Mendoza***

Institute of Mathematics, University of the Philippines Diliman
Quezon City 1101 Philippines

The Whittaker-Henderson graduation is a data smoothing method commonly used in actuarial science for mortality table computation. This method is a finite-dimensional minimization problem, which aims to give a smooth approximation of the crude data. Our goal is to treat the Whittaker-Henderson method as a minimization problem in a Sobolev space. We take advantage of some results in the study of Sobolev spaces to analyze the solution to the graduation problem. Solving the arising minimization problem is equivalent to solving a partial differential equation (PDE). To solve the PDE numerically, we use the finite element method (FEM). We apply this variational approach to the data sets found in the literature. In particular, we test our method to check for patterns in (1) mortality rates for Filipino males, (2) global average temperature anomaly, and (3) monthly eBay share price. We also compare the results with that of the Whittaker-Henderson method.

Keywords: finite element method, optimization, variational formulation, Whittaker-Henderson graduation method

INTRODUCTION

In the actuarial industry, if deaths are due to accidents then mortality rates are expected to peak around the stages of adolescence to early adulthood. Past this phase, mortality rates are presumed to be somehow level. To check if such a pattern can be seen, one can determine how far one can depart from the crude mortality data while obtaining smoothness and fidelity. To do this, a process called graduation is employed. Graduation is simply the process of finding estimates for the crude data in such a way that these estimates exhibit smoothness while being consistent with the crude data (Miller 1949). The estimates produced after the graduation process are called graduated values. From the definition of graduation, we must take two things into account: we want the new data to be close to the crude data and, at the same time, exhibit some smoothness.

A popular graduation technique which considers the problem on fit and smoothness is the so-called Whittaker-Henderson graduation. A lot of actuaries prefer to use this technique because it gives them the freedom to choose what the graduated values should favor: smoothness, fidelity to the crude/actual data, or a balance between the two (Henderson 1924). A chapter dedicated to the discussion of the Whittaker-Henderson method of graduation can be

*Corresponding Author: rmendoza@math.upd.edu.ph

found in the book by MacAulay (1931). For a historical background on this graduation method, one may read on Nocon and Scott (2012).

The one-dimensional Whittaker-Henderson graduation method aims to find a set of vectors $\{u_i\}_{i=1}^{n}$ given the crude data, $\{v_i\}_{i=1}^{n}$, and some $\lambda \geq 0$ that will minimize the composite measure given by:

$$F(u) + \lambda S(u) = \sum_{i=1}^{n} w_i(u_i - v_i)^2 + \lambda \sum_{i=1}^{n-1} (\Delta u_i)^2, \quad (1)$$

where the w_i's are weights designated to each data point. The weights w_i's are positive numbers, which are user-defined. The symbol Δ denotes the Newton's advancing operator or differencing operator defined as:

$$\Delta u_i = u_{i+1} - u_i.$$

In the context of mortality rate graduation, the indexing i usually corresponds to time or to an individual's age. The vector that provides the minimum of expression (Equation 1) will serve as our graduated values.

The F component in (Equation 1) serves as a measurement of fitness, *i.e.* the fidelity of the graduated values to the crude data set. Indeed, the F component penalizes the weighted sum of the square of the deviation of the graduated values from the crude data. The weight parameter, w_i, dictates how close the graduated values must be to the crude data. For instance, if w_i is large for some i then in order to minimize the first term in (Equation 1), we expect that the graduated value is close to the crude data. The values of w_i's are set by the user. In some cases, the weights are uniformly set to one (Nocon and Scott 2012). There are instances when the user would like to put more emphasis on some data points and less on others. For example, in Miller (1949), the weights vary from 1–23.

The latter component in Equation 1, S, serves as a measure of the smoothness of the graduated values. It is somehow intuitive that F and S have a push-pull reaction. Wanting the graduated data to be close to the crude data might compromise smoothness. Similarly, wanting the graduated data to be as smooth as possible might compromise the fidelity of the result to the actual data. Thus, the entire graduation process must assure that the graduated values will be as close to the crude data as possible while attempting to maintain smoothness (Knorr 1984).

As stated, the expression being minimized contains a parameter $\lambda \geq 0$. This constant plays an important role in determining the smoothness that will be exhibited by the graduated values. When $\lambda \ll 1$, the emphasis will be more on the fit. If $\lambda = 0$, then the minimizer of the function (Equation 1) will be the crude data points themselves. On the other hand, when $\lambda \to \infty$, the graduation process is inclined to smoothing. In fact, as $\lambda \to \infty$, the graduated values tend to converge to a constant polynomial (Knorr 1984). We present in the next section an approach to determine this constant polynomial (see Theorem 6). So, the question now is how can one find the appropriate value for λ? An approach done by Brooks *et al.* (1988) to determine this λ is to minimize what is called the generalized cross-validation (GCV) score (Nocon and Scott 2012). An efficient algorithm for the computation of the GCV score has been proposed by Weinert (2007).

It is well-known that the graduated values from the Whittaker-Henderson graduation can be obtained by converting the problem into a linear system (Nocon and Scott 2012; MacLeod 1989). A proof for this was shown by Greville in his study notes (Greville 1974). He proceeded by first setting:

W : an $n \times n$ matrix containing the weights w_i on its diagonals,

u : the vector $(u_1, u_2, \ldots, u_n)^T$: the vector of graduated values,

v : the vector $(v_1, v_2, \ldots, v_n)^T$: the vector of crude data,

K : an $(n-1) \times n$ matrix such that $Ku = (\Delta u_1, \Delta u_2, \ldots, \Delta u_{n-1})^T$.

Then:

$$K_{i,j} = (-1)^{1+j-i} \binom{1}{j-i}, i = 1, \ldots, n-1, \qquad j = i, \ldots, i+1.$$

The graduated values are computed by:

$$u = (W + \lambda K^T K)^{-1} W v. \tag{2}$$

If we set $W = I_n$ in Equation 2, then the quantity $(W + \lambda K^T K)^{-1}$ is called the smoother matrix and the elements of this matrix are called the smoother weights for the smoothing. Explicit formulas have been given for the smoother weights of the Whittaker-Henderson graduation of degree 1 (Yamada and Jahra 2019). The smoother matrix has been shown to be bisymmetric (Yamada 2020). Some extensions to this graduation method have been introduced. For example, in Nocon and Scott (2012), a graduation process by minimizing the criterion:

$$(u-v)^T \ (u-v) + \lambda (Ku)^T \ (Ku),$$

and subjecting it to some constraints is proposed. A multi-dimensional approach to this graduation method is proposed in Knorr (1984). The rationale for this approach is similar to the one-dimensional case in the sense that it also aims to minimize the composite measure $F(u) + \lambda S(u)$. Moreover, the multi-dimensional setup maintains the relationship between the two addends. The only difference is in the form of F and S. For the two-dimensional setting, the fitness component is given by:

$$F(u) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_{ij} (u_{ij} - v_{ij})^2,$$

where $\{v_{ij}\}$ represents the crude data and $\{u_{ij}\}$ represents the graduated data. This means that if we have a grid of data, we simply take the difference of the crude and the graduated values across each cell, square it, and then multiply it by the weight function associated with the cell. Smoothness in the two-dimensional setting is obtained both vertically and horizontally so that we can balance the smoothness in all directions possible. Thus, the constant λ is a vector with two entries: an entry for smoothing down the vertical and another for smoothing across the horizontal. In general, these two entries need not be equal.

In the following discussion, we will adapt the notations used in Mamon (1996). Suppose that Δ_h is the differencing operator across the horizontal and that Δ_v is the differencing operator down the vertical. Then, a measure of smoothness on the ith row (across the horizontal) will be given by:

$$\sum_{j=1}^{n_2-1} (\Delta_h u_{ij})^2.$$

Consequently, the overall measure of smoothness across the horizontal is given by:

$$^{h}S(u) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2-1} (\Delta_h u_{ij})^2.$$

In a similar fashion, a measure of smoothness on the j^{th} column (down the vertical) is given by:

$$\sum_{i=1}^{n_1-1} (\Delta_v u_{ij})^2.$$

This means that the overall measure of vertical smoothness is given by:

$$^{v}S(u) = \sum_{j=1}^{n_2} \sum_{i=1}^{n_1-1} (\Delta_v u_{ij})^2.$$

The smoothness term, S, will be the sum of these two smoothing terms, *i.e.*:

$$\lambda S(u) = \lambda_1 {}^{v}S(u) + \lambda_2 {}^{h}S(u), \qquad \lambda_1, \lambda_2 > 0.$$

Thus, the Whittaker-Henderson Graduation for 2D minimizes the expression:

$$\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_{ij}(u_{ij} - v_{ij})^2 + \lambda_1 \sum_{j=1}^{n_2} \sum_{i=1}^{n_1-1} (\Delta_v u_{ij})^2 + \lambda_2 \sum_{i=1}^{n_1} \sum_{j=1}^{n_2-1} (\Delta_h u_{ij})^2. \tag{3}$$

Note that the Whittaker-Henderson graduation method is a finite-dimensional minimization problem. For this work, our main goal is to propose an alternative way to graduate the crude data. What we do is treat the crude data as a function using spline interpolation. Then, we convert the graduation problem as a minimization in a suitable function space.

We define the following functions:

$w(x) :=$ the linear spline interpolation of the weights w_i,

$f(x) :=$ the linear spline interpolation of the crude data v_i,

$u(x) :=$ the linear spline interpolation of the graduated data u_i.

Suppose that there are n observations, *i.e.* our crude data set is given by $\{v_i\}_{i=1}^{n}$. What we do is treat the points on each line segment connecting v_{i-1} and v_i, $i = 2, \dots, n$, as an observation or a part of the crude data. Because we are now treating the crude data as a piecewise linear function, we transform the F component of Equations 1 and 3 as follows:

$$F \longrightarrow \int_{\Omega} w(x)[u(x) - f(x)]^2 \; dx, \tag{4}$$

where Ω is a region in $\mathbb{R}^n$, $n \in \mathbb{N}$. Note that for $n \geq 2$, the integral is viewed as a multiple integral. The case $n = 2$ is consistent with the double summation in Equation 3.

Next, we recall the concept of the forward difference formula., *i.e.* for a given function $f: \mathbb{R} \to \mathbb{R}$ and $\Delta x > 0$:

$$f'(x) \approx \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$

Setting $\Delta x = 1$, we can see the resemblance of the resulting expression with Newton's advancing operator Δ. We use this to motivate our identification of the smoothness term in Equation 1 with the derivative. Hence, we identify the smoothness term with the gradient, *i.e.*:

$$\lambda S \longrightarrow \lambda \int_{\Omega} [\nabla u(x)]^2 dx. \tag{5}$$

Recall that for the two-dimensional Whittaker-Henderson graduation, there are two smoothness constants: the one along the horizontal direction and the other along the vertical. Again, these two constants need not be equal. In this work, the integral expression will assume that these two constants are equal. Adding the two integrals from Equations 4 and 5, we will get the following quantity:

$$J(u) = \int_{\Omega} w(x)[u(x) - f(x)]^2 \; dx + \lambda \int_{\Omega} [\nabla u(x)]^2 dx. \tag{6}$$

The expression above can be viewed as an infinite-dimensional counterpart of Equations 1 and 3. In a way, the functional J in (Equation 6) generalizes the approach by Henderson (1924) (n = 1) and Knorr (1984) (n = 2). Like the rationale of the Whittaker-Henderson graduation, we aim to minimize $J(u)$. In other words, we ought to find the value u^*, which minimizes $J(u)$. Since the argument of Equation 6 is a function, then we can reformulate our problem into a problem of finding a function u^* in some function space V such that:

$$u^* = argmin_{u \in V} \; J(u). \tag{7}$$

In what function space V will the minimizer of $J(u)$ belong? We will see in the next sections that the answer to this question is the function space called the Sobolev space. In Section 2, we will show that the minimization of $J(u)$ is equivalent to solving a PDE – in particular, a boundary value problem (BVP) – under certain assumptions. We use mathematical results from functional analysis to show that the proposed minimization problem has a unique solution in the chosen function space. Furthermore, we study the behavior of the solution as we perturb the smoothing parameter λ. In Section 3, we present a numerical tool called the FEM to approximate the solution of the PDE. In Section 4, we present our numerical results. Lastly in Section 5, we give a summary and possible future work.

THEORETICAL RESULTS

In this section, we discuss the mathematical concepts needed to solve u^* in Equation 7. We do not include the discussion of linear operators and function spaces, as they are standard in real analysis. We refer the readers to Adams (1975), Cioranescu *et al*. (2018), and Evans (1998) for the proofs and examples of these concepts. In Subsection 2.1, we analyze J in Equation 6. We also show that the minimizer of J exists in a suitable space and is a solution to a variational formulation. In Subsection 2.2, we analyze the properties of u^* as we vary the smoothing parameter λ.

Analysis of the Functional *J*

Recall from Equation 6 that:

$$J(u) = \int_{\Omega} w(x)[u(x) - f(x)]^2 \; dx + \lambda \int_{\Omega} [\nabla u(x)]^2 dx, \tag{8}$$

where $\Omega \subset \mathbb{R}^n, n \in \{1,2\}$. The second term of J requires the argument u to be differentiable at least once, then it makes sense to take its minimizer from the Sobolev space $H^1(\Omega)$. For the discussion of Sobolev spaces and their properties, one may refer to Adams (1975). Therefore, we wish to solve the optimization problem:

$$u^* = arg\ min_{u \in H^1(\Omega)} J(u). \tag{9}$$

In view of Equation 9, we can call J a functional on $H^1(\Omega)$. Getting the Gâteaux derivative (Evans 1998) of J along an arbitrary direction $v \in H^1(\Omega)$, we get:

$$\begin{aligned} \frac{\partial J}{\partial u}(u; v) &= \lim_{s \to 0} \frac{J(u + sv) - J(u)}{s} \\ &= \frac{d}{ds}\left(\int_\Omega w(x)[u(x) + sv(x) - f(x)]^2 + \lambda[\nabla(u(x) + sv(x))]^2 dx\right)\Big|_{s=0} \\ &= \int_\Omega 2w(x)[u(x) - f(x)]v(x)\ dx + 2\lambda \int_\Omega \nabla u(x) \cdot \nabla v(x)\ dx. \end{aligned} \tag{10}$$

We use Green's theorem (Evans 1998) and obtain:

$$\int_{\partial\Omega} v \frac{\partial u}{\partial n}\ ds = \int_\Omega v\Delta u\ dx + \int_\Omega \nabla u \cdot \nabla v\ dx. \tag{11}$$

Suppose that u is sufficiently smooth. Then, equating Equation 10 to 0 and incorporating Equation 11, we will obtain the weak necessary optimality condition in minimizing functional (Equation 8). In other words, a sufficiently smooth minimizer of functional (Equation 8) satisfies:

$$\int_\Omega w(x)[u(x) - f(x)]v(x) - \lambda v(x)\Delta u(x) dx + \lambda \int_{\partial\Omega} v(x) \frac{\partial u(x)}{\partial n}\ ds\ = 0. \tag{12}$$

Since we have assumed that u is sufficiently smooth, we may apply the fundamental lemma of the calculus of variations (Adams 1975) to arrive at:

$$\begin{cases} -\lambda\Delta u(x) + w(x)u(x) = w(x)f(x), & x \in \Omega, \\ \dfrac{\partial u}{\partial n} = 0, & x \in \partial\Omega. \end{cases} \tag{13}$$

This means that a sufficiently smooth minimizer of functional (Equation 8) will satisfy the problem given by the PDE (Equation 13). In Henderson (1924), a difference equation was used to compute the graduated values. Using linear spline interpolation instead of discrete data points means that we are working with functions instead of vectors. Thus, a differential equation (Equation 13), instead of a difference equation, is expected. Solving Equation 13 requires u to be at least twice-differentiable because of the Laplacian operator. Furthermore, the functions $f(x)$ and $w(x)$ are piecewise linear functions and are not smooth. Traditional techniques won't work. We resolve this by solving an equivalent problem called the variational or weak formulation. Our next task is to find the variational/weak formulation of Equation 13. We present this formulation in the following theorem. This main result provides a way of solving u^* in Equation 9. This theorem is similar to what was shown in Recio and Mendoza (2019). The difference is the inclusion of the weight function $w(x)$.

Theorem 1. Let Ω be a closed and bounded region in $\mathbb{R}^n$, for $n = 1,2.$ Let:

$$K = \left\{ v \in H^1(\Omega) \mid \frac{\partial v}{\partial n} = 0 \ \ in \ \ \partial\Omega \right\}.$$

Then, the weak formulation of the necessary optimality condition of Equation 9 is given by:

$$\int_{\Omega} w(x)u(x)v(x) + \lambda \nabla u(x) \cdot \nabla v(x) \, dx = \int_{\Omega} w(x)f(x)v(x) \ dx, \forall v \in K. \tag{14}$$

This formulation admits a unique solution in K. Furthermore, u is the solution of Equation 14 if and only if u is the unique minimizer of the functional (Equation 8) in K.

Proof. Applying the boundary condition of Equations 13, we get:

$$-\int_{\Omega} v\Delta u \ \ dx = \int_{\Omega} \nabla u \cdot \nabla v \ \ dx. \tag{15}$$

Moreover, Equation 12 will reduce to:

$$\int_{\Omega} w(x)[u(x) - f(x)]v(x) \, dx - \lambda \int_{\Omega} v(x)\Delta u(x) \, dx = 0. \tag{16}$$

Applying Equation 15 into Equation 16 will yield equation Equation 14. Define:

$$a(u, v) := \int_{\Omega} w(x)u(x)v(x) \ \ dx + \lambda \int_{\Omega} \nabla u(x) \cdot \nabla v(x) \ \ dx, \tag{17}$$

$$b(v) := \int_{\Omega} w(x)f(x)v(x) \ \ dx. \tag{18}$$

Observe that:

$$a(u, v) = b(v) \quad \forall v \in K,$$

is equivalent to Equation 14. Because the integral operator is linear, it can easily be shown that a in Equation 17 is bilinear. Similarly, b in Equation 18 is linear. To show that Equation 14 has a unique solution, we use the Lax-Milgram theorem (Cioranescu *et al*. 2018), *i.e.* we need to prove that a and b in Equations 17 and 18 are both continuous. We also need to show that a is coercive. First, since $w(x)$ is piecewise linear, it is bounded, and we have:

$$\| w \|_{\infty} < \infty \tag{19}$$

We now show the following:

The continuity of $a(u, v)$, *i.e.* $\exists c_1 > 0$ such that $|a(u, v)| \leq c_1 \| u \|_{H^1(\Omega)} \| v \|_{H^1(\Omega)}$.

We use the triangle inequality, Hölder's inequality (Cioranescu *et al*. 2018), Equation 19, and the definition of the H^1 norm to obtain:

$$|a(u,v)| \leq \| w \|_{\infty} \| u \|_{H^1(\Omega)} \| v \|_{H^1(\Omega)} + \lambda \| u \|_{H^1(\Omega)} \| v \|_{H^1(\Omega)}$$

$$= \underbrace{(\| w \|_{\infty} + \lambda)}_{c_1} \| u \|_{H^1(\Omega)} \| v \|_{H^1(\Omega)}.$$

The continuity of $b(v)$, *i.e.* $\exists c_2 > 0$ such that $|b(v)| \leq c_2 \| v \|_{H^1(\Omega)}$.

Applying the techniques from (a) and setting the constant $c_2 = \| w \|_{\infty} \| f \|_2$, we have a similar result:

$$b(v) \leq \underbrace{(\| w \|_{\infty} \| f \|_2)}_{c_2} \| v \|_{H^1(\Omega)}.$$

The coercivity of $a(u,v)$, *i.e.* $\exists c_3 > 0$ with $|a(u,u)| \geq c_3 \| u \|^2_{H^1(\Omega)}\ \forall u \in H^1(\Omega)$.

Since $w(x)$ is piecewise linear and positive, then $\exists\ \underline{w} > 0$ such that:

$$\underline{w} \leq w(x). \tag{20}$$

Then:

$$|a(u,u)| \geq \int_\Omega \underline{w} |u^2(x)|\ dx + \lambda \int_\Omega |(\nabla u(x))^2|\ dx$$

$$\geq \underbrace{\min\{\underline{w}, \lambda\}}_{c_3} \| u \|^2_{H^1(\Omega)}.$$

and so, the inequality is satisfied.

We now show that the minimizer of functional (Equation 8) in K is the unique solution of the weak formulation in Equation 14. Suppose that u is a minimizer of J in K. Let $v \in K$. Then, for any $t \in \mathbb{R}$, $u + tv \in K$. Define $z: \mathbb{R} \to \mathbb{R}$ by:

$$z(t) = J(u + tv) - J(u).$$

Since u is the minimizer of J, then $z(t) \geq 0$. Moreover, z attains its minimum when $z = 0$. Using the definition of J, we can simplify $z(t)$:

$$z(t) = 2t\left[\int_\Omega [w(x)[u(x)v(x) - f(x)v(x)]\ dx + \lambda \int_\Omega \nabla u(x) \cdot \nabla v(x)\ dx\right]$$

$$+ t^2 \int_\Omega w(x)v^2(x) + [\nabla v(x)]^2\ dx.$$

This implies that:

$$z'(t) = 2\int_\Omega [w(x)u(x)v(x) - w(x)f(x)v(x)]\ dx + 2\lambda \int_\Omega \nabla u(x) \cdot \nabla v(x)\ dx$$

$$+ 2t \int_\Omega w(x)v^2(x) + [\nabla v(x)]^2\ dx.$$

Since 0 is the minimum of z then $z'(0) = 0$. So:

$$0 = z'(0) = \int_\Omega [2w(x)u(x)v(x) - 2w(x)f(x)v(x)]\,dx + 2\lambda \int_\Omega \nabla u(x) \cdot \nabla v(x)\;dx.$$

If we divide both sides by two, then the resulting expression is precisely Equation 14. Since u solves the variational formulation, then it must be unique. Conversely, let u be the unique solution of Equation 14 in K. Let $q(x) \in K$ and define $v(x) = q(x) - u(x) \in H^1(\Omega)$. To simplify typesetting, we omit the independent variable x. Then, by using the definition of J:

$$J(q) - J(u) = 2\left[\int_\Omega (wuv - fv)\,dx + \lambda \int_\Omega \nabla u \cdot \nabla v\,dx\right] + \int_\Omega wv^2\,dx + \lambda \int_\Omega [\nabla v]^2\;dx.$$

Because of Equation 14, the first two integral terms of the above equation is 0. Thus, we have:

$$J(q) - J(u) = \int_\Omega wv^2\;dx + \lambda \int_\Omega [\nabla v]^2 dx \geq 0.$$

Thus, $J(q) \geq J(u)\ \forall q \in K$ and u is the unique minimizer of functional (Equation 8) in K. ■

Properties of the Graduated Values

Theorem 1 tells us that to calculate the graduated values, it is sufficient to solve Equation 14. In this subsection, we present some results on how the smoothing parameter λ affects the calculated graduated values.

Definition 2. Let $\lambda \in \mathbb{R}^+$. We define the operator $T_\lambda : L^2(\Omega) \to L^2(\Omega)$ by:

$$T_\lambda(f) = u_\lambda,$$

where u_λ is the solution of the variational formulation (Equation 14) given the piecewise linear polynomial $f \in L^2(\Omega)$ interpolating the crude data.

From Section 1, we have seen that as $\lambda \to 0$ in Equations 1 and 3, the graduated values will approach the crude data. Interestingly, we will show in the next theorem that this property carries over to the solution of Equation 14. But first, we prove the following lemma.

Lemma 3. Let f be the crude data interpolation and $u_\lambda(x) = T_\lambda(f)$ for $\lambda \in \mathbb{R}^+$, then:

$$\lim_{\lambda \to +\infty} \| \nabla u_\lambda(x) \|_2 = 0.$$

Proof. Because u_λ is the minimizer of the functional J on K and since $0 \in K$, then $J(u_\lambda) \leq J(0)$. Using this, we calculate the inequality:

$$\begin{aligned}\lambda \int_\Omega |\nabla u_\lambda(x)|^2 dx &\leq \int_\Omega w(x)|u_\lambda(x) - f(x)|^2 dx + \lambda \int_\Omega |\nabla u_\lambda(x)|^2 dx \\ &= J(u_\lambda) \\ &\leq J(0)\end{aligned}$$

$$= \int_{\Omega} w(x) f^2(x) dx$$

$$\leq \|w\|_{\infty} \|f\|_2^2.$$

Therefore:

$$\|\nabla u_\lambda\|_2 \leq \sqrt{\frac{\|w\|_{\infty}}{\lambda}} \|f\|_2. \tag{21}$$

We achieve our desired result if we let $\lambda \to \infty$. ■

Theorem 4. Let f be the crude data interpolation and $u_\lambda(x) = T_\lambda(f)$ for $\lambda \in \mathbb{R}^+$, then:

$$\lim_{\lambda \to 0} \| u_\lambda(x) \|_2 = f.$$

Proof. Because the variational formulation holds for all $v \in K$, we can set $vf - u_\lambda$. Using Equation 14, Equation 20, the linearity of the gradient operator, the Cauchy-Schwarz inequality (Cioranescu *et al.* 2018), and Equation 21, we have:

$$\underline{w} \| u_\lambda(x) - f(x) \|_{L^2(\Omega)}^2 \leq \int_{\Omega} w(x)(u_\lambda(x) - f(x))^2 \; dx$$

$$= \lambda \int_{\Omega} \nabla u_\lambda(x) \cdot \nabla(f(x) - u_\lambda(x)) \; dx$$

$$\leq \lambda \int_{\Omega} \nabla u_\lambda \cdot \nabla f dV$$

$$\leq \sqrt{\lambda \|w\|_{L^\infty(\Omega)}} \|f\|_{L^2(\Omega)} \|\nabla f\|_{L^2(\Omega)}.$$

Hence, if we let $\lambda \to 0$, we arrive at the result. ■

Intuitively, this result makes sense since letting the smoothing parameter approach zero will favor fidelity over smoothness. Next, we show what happens when we make $\lambda \to \infty$ if the weights are uniform. To do so, we need to establish the following lemma.

Lemma 5 (Invariance Property). *For any* $\lambda \in \mathbb{R}^+$*:*

$$\int_{\Omega} w(x) u_\lambda(x) \; dx = \int_{\Omega} w(x) f(x) \; dx.$$

Proof. Once again we recall that u_λ satisfies the following BVP:

$$\begin{cases} -\lambda \Delta u_\lambda(x) + w(x) u_\lambda(x) &= w(x) f(x), \quad x \in \Omega, \\ \dfrac{\partial u_\lambda}{\partial n} &= 0, \quad x \in \partial\Omega. \end{cases}$$

Integrating the given differential equation over Ω and using the definition of the Laplacian operator Δ, we obtain:

$$-\lambda \int_{\Omega} \text{Div}(\nabla u_\lambda(x)) \ dx + \int_{\Omega} w(x) u_\lambda(x) \ dx = \int_{\Omega} w(x) f(x) \ dx.$$

By Gauss' divergence theorem (Evans 1998), we get:

$$-\lambda \int_{\partial\Omega} \frac{\partial u_\lambda}{\partial n} \ ds + \int_{\Omega} w(x) u_\lambda(x) \, dx = \int_{\Omega} w(x) f(x) \, dx. \tag{22}$$

Applying the boundary condition to Equation 22 will yield the desired result. ■

We now show if we set λ to be too large, the graduated value becomes a constant function.

Theorem 6. Suppose that for all $x \in \Omega$, $w(x) = k$, for some constant k. Then u_λ converges in $L^1(\Omega)$ to the average of the crude data interpolation f, *i.e.*:

$$\| u_\lambda - \bar{f} \|_1 \rightarrow 0 \ \text{ as } \ \lambda \rightarrow \infty,$$

where:

$$\bar{f} = \frac{1}{|\Omega|} \int_{\Omega} f(x) \ dx$$

and $|\Omega|$ is the size of Ω.

Proof. From Lemma 5, if $w(x) = k$, with k being a constant, we will get:

$$\int_{\Omega} k u_\lambda(x) \, dx = \int_{\Omega} k f(x) \, dx, \quad \Rightarrow \quad \int_{\Omega} u_\lambda(x) \, dx = \int_{\Omega} f(x) \, dx. \tag{23}$$

Next, we see that:

$$\| u_\lambda - \bar{f} \|_1 = \int_{\Omega} |u_\lambda(x) - \bar{f}(x)| \ dx = \int_{\Omega} \left| u_\lambda(x) - \frac{1}{|\Omega|} \int_{\Omega} f(x) \ dx \right| \ dx.$$

By using Equation 23, the integral expression becomes:

$$\| u_\lambda - \bar{f} \|_1 = \int_{\Omega} \left| u_\lambda(x) - \frac{1}{|\Omega|} \int_{\Omega} u_\lambda(x) \, dx \right| \ dx = \int_{\Omega} |u_\lambda(x) - \bar{u}_\lambda(x)| \, dx.$$

By the Poincaré-Wirtinger inequality (Cioranescu *et al.* 2018), $\exists C_\Omega$ such that:

$$\| u_\lambda - f \|_1 \leq C_\Omega \| \nabla u_\lambda \|_1.$$

Moreover, because $L^2(\Omega)$ is embedded in $L^1(\Omega)$, $\exists C_2 > 0$ dependent on Ω such that:

$$\| u_\lambda - f \|_1 \leq C_\Omega C_2 \| \nabla u_\lambda \|_2.$$

The right-hand side of the inequality above tends to 0 as $\lambda \rightarrow \infty$ because of Lemma 3. ■

The theorem above suggests that as $\lambda \to \infty$, the graduated values becomes a constant function equal to the average of the crude data interpolation.

FINITE ELEMENT FORMULATION

The FEM is a numerical technique used in solving differential equations. FEM is a popular method to solve PDEs because of its capability to handle complex geometries. The accuracy of the solution using FEM is often higher compared to other techniques like the finite difference method. Furthermore, the solution using FEM is obtained from the variational formulation of the PDE. This makes the theoretical and convergence analysis of the solution easier because one can take advantage of the tools from the study of Sobolev spaces. One may read on Chen (2007), Pepper and Heinrich (2005), and Reddy (1993) for the discussion of this technique. For the use of FEM in applications, one may refer to Cockburn *et al.* (2000), Hecht (2019), and Madenci and Guven (2015).

The FEM can be summarized by the following steps:

- State the governing equations for the problem.
- Formulate the variational or weak form of the problem.
- Discretize the domain where the equation is defined. For our purpose, the discretization technique that we employed in the two-dimensional setting is called triangulation.
- Determine the (finite) function space where the approximate solution must lie.
- Assemble the necessary matrices. Impose the boundary conditions whenever necessary.
- Perform the calculations.

Recall that the variational formulation for the optimization problem (Equation 9) is given in Equation 14:

$$\int_\Omega w(x)u(x)v(x)\ dx + \lambda \int_\Omega \nabla u(x) \cdot \nabla v(x)\ dx = \int_\Omega w(x)f(x)v(x)\ dx, \quad \forall v \in K.$$

Since K is an infinite-dimensional space, solving for u satisfying the above equation $\forall v \in K$ is equivalent to solving an infinite-dimensional system. To overcome this, we can project the problem into a finite-dimensional subspace of K, say V_h. Hence, we restate the variational problem as follows:

$$\int_\Omega w(x)u_h(x)v_h(x)\ dx + \lambda \int_\Omega \nabla u_h(x) \cdot \nabla v_h(x)\ dx = \int_\Omega w(x)f(x)v_h(x)\ dx, \tag{24}$$

$\forall v_h \in V_h$. Since V_h is a finite-dimensional subspace of K, it is enough if the equation above holds for each of the basis functions of V_h. These basis functions are constructed based on the discretization of the domain Ω. In this work, we consider the Lagrange basis functions, which are piecewise linear functions. This is one of the standard basis functions used in FEM. For the discussion about this, one may read Pepper and Heinrich (2005). We denote these basis functions by $\{\phi_i(x)\}_{i=1}^n$, where $n = dim(V_h)$. Restricting the formulation (Equation 24) to just the basis functions yields a system of n equations. Note that $u_h \in V_h$ can be expressed as a linear combination of the basis functions of V_h, *i.e.*:

$$u_h = \sum_{i=1}^{n} u_i \phi_i(x), \tag{25}$$

where u_i's are the unknown coefficients. Similarly, we express the crude data interpolation f using the basis functions of V_h, *i.e.*:

$$f = \sum_{i=1}^{n} f_i \phi_i(x),$$

where the values of f_i's depend on the discretization of Ω. Limiting Equation 24 to the basis functions and taking into account the alternate expressions for u_h and f, formulation (Equation 24) becomes:

$$\sum_{i=1}^{n} u_i \int_{\Omega} w(x)\phi_i(x)\phi_j(x) + \lambda \nabla\phi_i(x) \cdot \nabla\phi_j(x) dx$$
$$= \sum_{i=1}^{n} f_i \int_{\Omega} w(x)\phi_i(x)\phi_j(x) dx, \qquad \forall j = 1,2,\dots,n. \tag{26}$$

Formulation (Equation 26) is a linear system with n equations and n unknowns. Here, the unknowns are the coefficients u_i's of the approximate solution u_h. In matrix form, Equation 26 is equivalent to:

$$KU = MF, \tag{27}$$

where:

K : a symmetric $n \times n$ matrix defined by $K_{ij} = \int_{\Omega} \phi_{i'}(x)\phi_{j'}(x) \; dx$,

U : the matrix of unknown coefficients $(u_1, \dots, u_n)^T$,

M : a symmetric $n \times n$ matrix defined by $M_{ij} = \int_{\Omega} \phi_i(x)\phi_j(x) \; dx$, and

F : the matrix defined by $(f(x_1), \dots, f(x_n))^T$.

Solving Equation 27 gives the coefficients u_i's of u_h. Plugging the values of u_i's to Equation 25 gives the numerical approximation of the data graduation.

RESULTS AND DISCUSSION

We now test our approach to some of the datasets we found in the literature. In Subsection 4.1, we test our approach for varying parameter λ. In Subsection 4.2, we compare our approach to the Whittaker-Henderson method. In Subsection 4.3, we present an alternative approach in solving the one-dimensional graduation problem using a MATLAB BVP solver. Finally, we test our approach to a two-dimensional dataset.

Varying Parameter λ

We now apply our proposed method to graduate the data of mortality rates for Filipino males aged 0–20 presented in ASP (2017). In this simulation, we used unit weights.

In Figure 1, we plot the computed graduated values using a variational approach for different values of λ. We can observe that when λ is smaller, the graduated values that we obtain resemble that of the crude data. This is consistent with Theorem 4. On the other hand, if we make the value of λ larger, it can be observed that the resulting graph of the graduated values becomes smoother and flatter. In particular, the graph resembles a constant function. For $\lambda = 500$, the graduated rates are approximately 0.32. The average value of the crude data interpolation is 0.316. This validates Theorem 6. We also implement our method using a smoothing parameter obtained through the minimized GCV score. For the discussion of the numerical scheme in computing the GCV score, we refer the reader to Weinert (2007) and Garcia (2010). For this dataset, the computed smoothing parameter using the minimized GCV score is $\lambda = 5.56$. The graduated values are also illustrated in Figure 1. Observe that this value of λ gives graduated rates, which are not too far from the crude data while exhibiting smoothness.

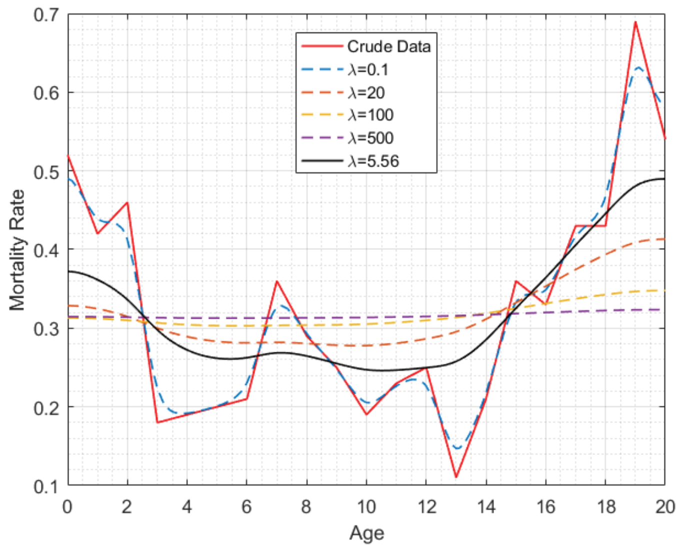

Figure 1. Graduated values of the mortality rates data for different values of λ**.**

Comparison with the Whittaker-Henderson Graduation

In Section 1, we formulated $J(u)$ by using the composite measure minimized by the Whittaker-Henderson Method. Thus, it is interesting to see how our proposed method will compare to Whittaker-Henderson graduation. We apply our technique to graduate the crude data regarding the global average temperature anomaly (in hundredths of $1^{\circ} C$) from 1989–2009, inclusive (Nocon and Scott 2012). To solve the graduated values using the Whittaker-Henderson method, we use the formula Equation 2 from the first section.

The smoothing parameter calculated by minimizing the GCV score for this dataset is $\lambda = 4.16$. We apply our method using this parameter value and compare the results computed using the Whittaker-Henderson graduation method. The results are illustrated in Figure 2. It can be observed that the two results are almost equal. We also consider the value of $\lambda = 97$. This value of λ was chosen since it is the one used in Nocon and Scott (2012). Now, we again make use of Equation 2 to find the Whittaker graduated values. A comparison of these graduated values with our proposed method is also shown in Figure 2. We can also observe how the obtained results are close to each other.

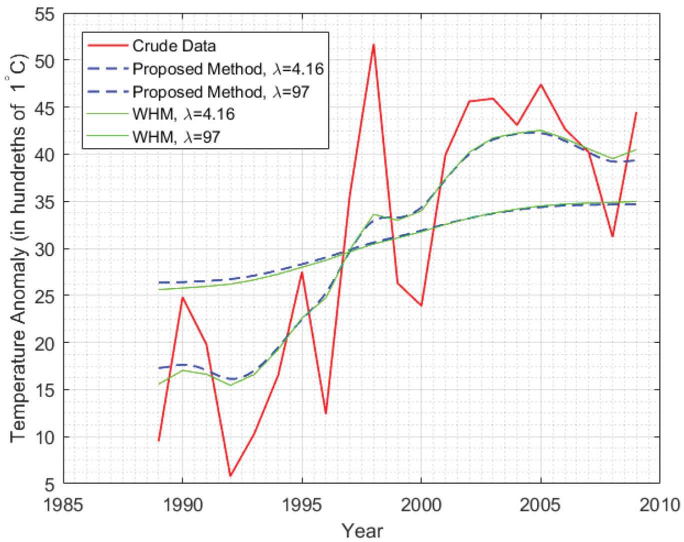

Figure 2. Comparison of the Whittaker-Henderson method with the proposed method using temperature anomaly data.

Using a MATLAB BVP Solver

In this part, we look at the behavior of the solution of the one-dimensional version of the problem (Equation 13). We have shown that the problem (Equation 13) is an equivalent way of solving the minimizer of $J(u)$.

Recall that the problem (Equation 13) in one dimension is given by:

$$\begin{cases} -\lambda u''(x) + w(x)u(x) = w(x)f(x), & x \in (a,b), \\ u'(x) = 0, & x = a, b, \end{cases} \tag{28}$$

where a and b are the endpoints of the data set. The differential equation above is a BVP. Hence, we can use the built-in BVP solver of MATLAB called bvp5c. For the discussion on how to use this solver, we refer the reader to www.mathworks.com/help/matlab/ref/bvp5c.html. For the description of the algorithms and implementations of bvp5c, one may refer to Kierzenka and Shampine (2008).

We apply this approach to the monthly eBay share price from January 2009 to August 2010 (Nocon and Scott 2012). We used the smoothing parameter $\lambda = 4.38$, obtained using the minimized GCV score. We see in Figure 3 that the solution of the BVP (Equation 28) matches the graduated values obtained by the Whittaker-Henderson method. Thus, for the one-dimensional case, one may choose to use the built-in MATLAB BVP solver instead of using FEM. This approach is simpler because setting up the linear system arising from FEM is not trivial.

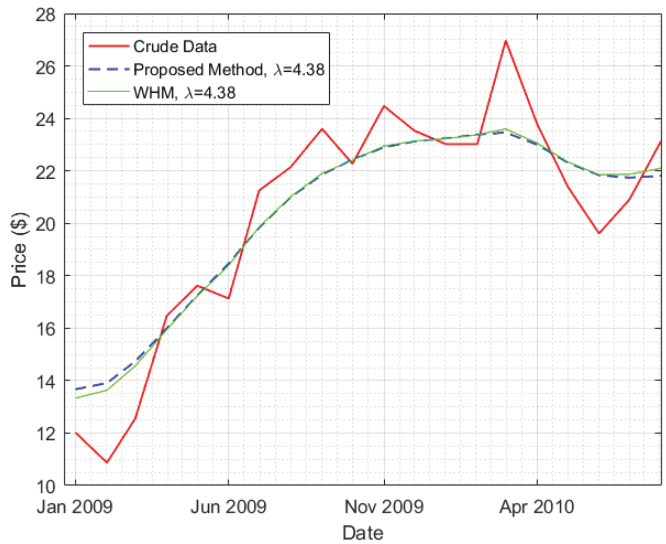

Figure 3. The graph of the solution to the BVP in Equation 30 with a value of $\lambda = 4.38$ compared to that of the Whittaker-Henderson method using the monthly eBay share price.

Two-dimensional Case

For this case, we consider a dataset from Knorr (1984) shown in Figure 4a. Applying the Whittaker-Henderson graduation on the given crude data, the graduated values are shown in Figure 4b. It is observed that the graph is not that "smooth" as there are jump discontinuities across ages. Moreover, it has somehow neglected the trend across the first five years of the policy corresponding to ages 45–50 (see Figure 4b).

Applying our method to the crude data, we again observe that for a small λ (shown in Figure 4d), the graduated values resemble that of the crude data. Setting $\lambda = 19.1$, we get a result that is smooth and is still close to the data. This value of λ is computed using the minimized GCV score. The plot is shown in Figure 4c. A difference between results obtained using the Whittaker-Henderson method and using FEM is that the graduated data using FEM appear to be smoother as there are no jump discontinuities.

If we set $\lambda = 1{,}000{,}000$, the result produces graduated values that are somewhat flat, and this is evident from the z-coordinates on the plot in Figure 4e. The average value of the crude data is computed to be 2.1475. Observe that the z-values of Figure 4e is between 2.1474–2.1476. This again validates Theorem 6.

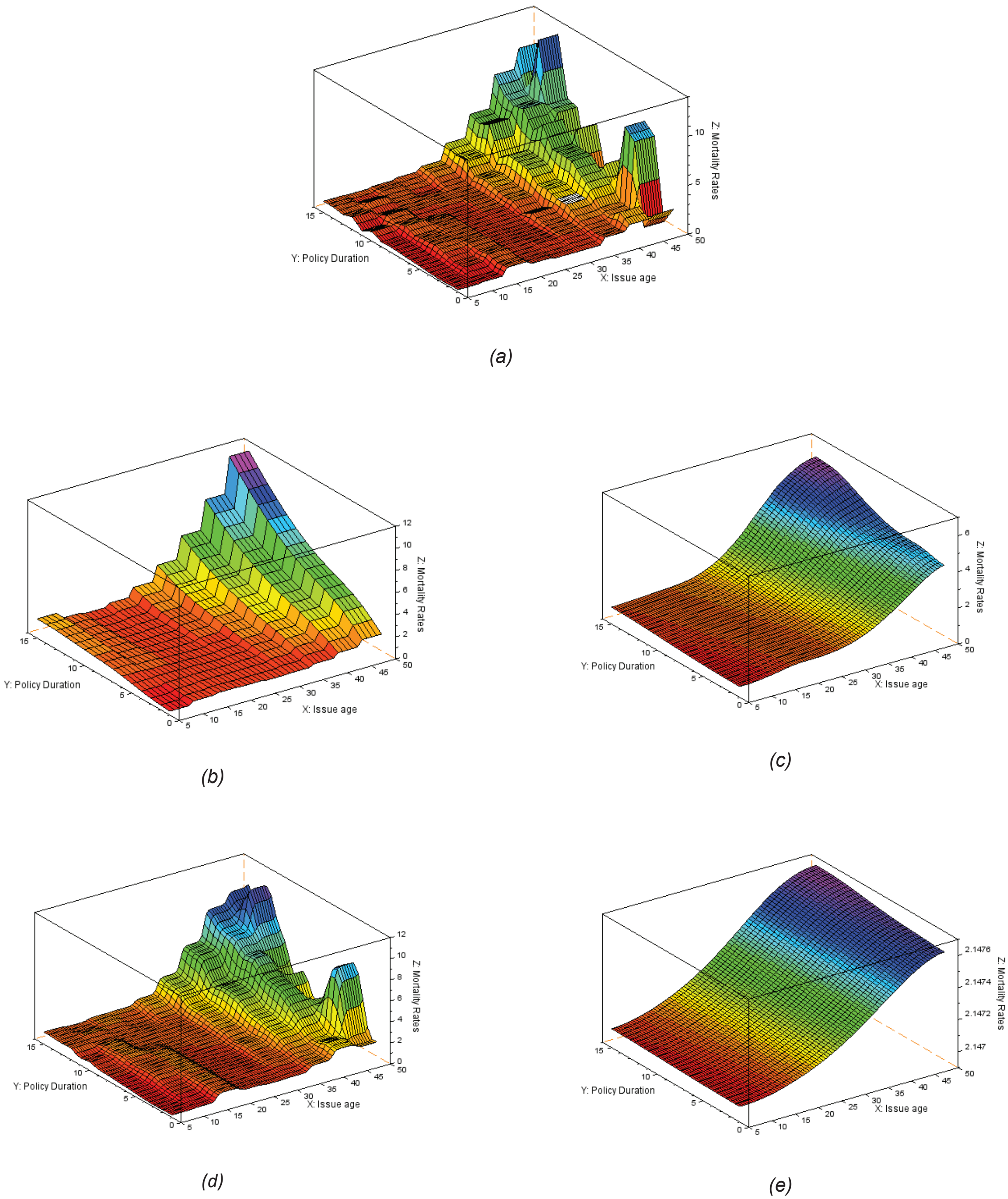

Figure 4. Plots of (a) two-dimensional crude mortality data, (b) graduated data using WHM, and (c) graduated data using a variational approach with $\boldsymbol{\lambda} = 19.1$, (d) $\lambda = 0.1$, and (e) $\lambda = 1000000$. The policy duration and issue age are both in years.

CONCLUSIONS AND RECOMMENDATIONS

In this work, we were able to express the Whittaker-Henderson graduation method as a minimization problem in the function space $H^1(\Omega)$. We have shown that this problem is equivalent to solving a PDE. To solve the PDE, we obtained the variational formulation and proved that this yields a unique solution in $H^1(\Omega)$. Furthermore, tools from the functional analysis were used to understand the behavior of our solution for varying parameters λ. For small λ, the graduated values resemble the crude data. For large λ, the graduated values tend to the average of the crude data interpolation. To approximate the solution of the variational formulation, we used the FEM. We test our methods to data sets found in the literature and the obtained results are similar to that of the Whittaker-Henderson method. In the one-dimensional case, an alternative numerical approach, the MATLAB built-in BVP solver, was also used to compute the graduated values. In the two-dimensional case, we treated the parameter λ as a vector whose entries are identical. The obtained graduated values in 2D using the variational approach gives smoother graduation. For higher-dimensional cases, $n \geq 3$, setting up the linear system can be tricky. One may use FEM solvers like FreeFEM++ (Hecht 2019) to solve this type of problem.

If the only goal of the researcher is to perform data graduation using a fixed smoothing parameter, he/she may use either the Whittaker-Henderson graduation or our proposed approach. Our numerical tests suggest that our obtained results are comparable to that of the Whittaker-Henderson method. However, if one wants to analyze the behavior of the solution of the graduation problem if the smoothing parameter λ is varied or if the interpolated data f is perturbed, our proposed approach can be more useful. Theoretical results from functional analysis and study of PDEs can be used for these types of analyses.

For future work, one can study what happens if the horizontal and vertical smoothing components are not equal. One can also consider a higher degree of the regularization, *i.e.* minimize the functional:

$$J(u) = \int_\Omega w(x)[u(x) - f(x)]^2 \ dx + \lambda \int_\Omega [\nabla^m u(x)]^2 dx,$$

where $m \geq 2$. Will the obtained result be similar to the extension of the Whittaker-Henderson method done by Nocon and Scott (2012)?

REFERENCES

ADAMS R. 1975. Sobolev Spaces. New York: Academic Press.

[ASP] Actuarial Society of the Philippines. 2017. Philippine Intercompany Mortality Study by the Actuarial Society of the Philippines (ASP) Life Insurance Committee.

BROOKS RJ, STONE M, CHAN FY, CHAN LK. 1988. Cross-validatory graduation. Insurance: Mathematics and Economics 7: 59–66.

CHEN L. 2007. Lecture Notes in Introduction to Finite Element Method.

CIORANESCU D, DONATO P, ROQUE M. 2018. Introduction to Classical and Variational Partial Differential Equations. World Scientific.

COCKBURN B, KARNIADAKIS G, SHU C eds. 2000. Discontinuous Galerkin Methods: Theory, Computation and Applications.

EVANS LC. 1998. Partial Differential Equations. American Mathematical Society.

GARCIA D. 2010. Robust Smoothing of Gridded Data in One and Higher Dimensions with Missing Values. Computational Statistics and Data Analysis 54: 1167–1178.

GREVILLE TNE. 1974. Part 5 Study Notes: Graduation. Society of Actuaries.

HECHT F. 2019. FreeFEM Documentation Release, version 4.2.1.

HENDERSON R. 1924. A new method of graduation. Transactions of the Actuarial Society of America 25: 29–40.

KIERZENKA J, SHAMPINE L. 2008. A bvp solver that controls residual and error. Journal of Numerical Analysis, Industrial and Applied Mathematics 3: 27–41.

KNORR FE. 1984. Multidimensional Whittaker-Henderson Graduation. Transactions of Society of Actuaries 36: 213–255.

MACAULAY F. 1931. The Smoothing of Time Series. National Bureau of Economic Research.

MACLEOD AJ. 1989. A Note on the Computation in Whittaker-Henderson Graduation. Scandinavian Actuarial Journal 2: 115–117.

MADENCI E, GUVEN I. 2015. The Finite Element Method and Applications in Engineering using ANSYS. Springer.

MAMON RS. 1996. Estimating Select and Ultimate Mortality Rates using a Whittaker-Henderson Graduation [M.S. Thesis]. University of the Philippines Diliman, Quezon City.

MILLER M. 1949. Elements of Graduation. The Actuarial Society of America and American Institute of Actuaries.

NOCON AS, SCOTT WF. 2012. An Extension of the Whittaker-Henderson Method of Graduation. Scandinavian Actuarial Journal 1: 70–79.

PEPPER D, HEINRICH J. 2005. The Finite Element Method: Basic Concepts and Applications. CRC Press.

RECIO KR, MENDOZA R. 2019. Three-step Approach to Edge Detection of Texts. Philippine Journal of Science 148(1): 193–211.

REDDY JN. 1993. An Introduction to the Finite Element Method, 2nd edition. McGraw-Hill Inc.

WEINERT HL. 2007. Efficient Computation for Whittaker-Henderson Smoothing. Computational Statistics and Data Analysis 52(2): 959–974.

YAMADA H. 2020. A note on Whittaker-Henderson graduation: Bisymmetry of the smoother matrix. Communications in Statistics – Theory and Methods 49(7): 1629–1634.

YAMADA H, JAHRA FT. 2019. Explicit formulas for the smoother weights of the Whittaker-Henderson graduation of order 1. Communications in Statistics – Theory and Methods 48(12): 3153–3161.

APPENDICES

The following are the MATLAB codes used in the paper.

One-dimensional Case

```
function [x_graduated,y_graduated]=onedim_
grad(x,y,lambda)
%x is 1xn
%y is 1xn
%y is the data vector
%lambda is the smoothing parameter

fcn=@(t) interp1q(x,y,t);
bvpfcn =@(t,y) [y(2);(y(1)-fcn(t))/lambda];
bcfcn =@(ya,yb) [ya(2);yb(2)];
guess = @(t) [1 1];
xmesh=linspace(x(1),x(end),500);
solinit = bvpinit(xmesh, guess);
sol = bvp5c(bvpfcn, bcfcn, solinit);
x_graduated=sol.x;
y_graduated=sol.y(1,:);

%You may omit the rest of the code if you want
your plots to look finer.
fcn2=@(t) interp1q(x_graduated,y_graduated',t);
x_graduated=x;
y_graduated=zeros(size(x));
for i=1:length(x);
    y_graduated(i)=fcn2(x(i));
end
```

Two-dimensional Case

```
function [x_graduated,y_graduated,z_
graduated]=twodim_grad(x,y,z,lambda)
%x is 1xn
%y is 1xm
%z is nxm
%z is the data matrix
%lambda is the smoothing parameter

h=x(2)-x(1);
f=@(t) interp2(x,y,z',t(1),t(2));

N=size(x,2)-1;
M=size(y,2)-1;

K=zeros((N+1)*(M+1), (N+1)*(M+1));
P=zeros((N+1)*(M+1), (N+1)*(M+1));
z_graduated=zeros(M+1,N+1);
F=zeros((N+1)*(M+1), 1);

xnodes=zeros((N+1)*(M+1), 1);
ynodes=zeros((N+1)*(M+1), 1);

for i=1:(N+1)*(M+1)
    j=mod(i-1,N+1)+1;
    k=floor((i-1)/(N+1))+1;

    xnodes(i,1)= x(1,j);
    ynodes(i,1)= y(1,k);
    F(i,1)=f([xnodes(i,1) ynodes(i,1)]);
end


E=zeros(2*N*M, 3);

for i=1:2:2*N*M-1
    E(i,1)= (i-1)/2 + ceil(i/(2*N));
    E(i,2)= E(i,1)+1;
    E(i,3)= E(i,1)+N+1;
    E(i+1,1)= E(i,2);
    E(i+1,2)= E(i,3);
    E(i+1,3)= E(i+1,2)+1;
end


for i=1:2*M*N
    for k=1:3
        for l=1:3
            if k==l
                P(E(i,k), E(i,l))= P(E(i,k),
E(i,l))+ (h^2)/12;
            elseif k~=l
                P(E(i,k), E(i,l))= P(E(i,k),
E(i,l))+ (h^2)/24;
            end
        end
    end
end

for i=1:2*M*N-1
    K(E(i,1),E(i,1))= K(E(i,1),E(i,1))+ 1;
    K(E(i,2),E(i,2))= K(E(i,2),E(i,2))+ 0.5;
    K(E(i,3),E(i,3))= K(E(i,3),E(i,3))+ 0.5;
    K(E(i,1),E(i,2))= K(E(i,1),E(i,2))- 0.5;
    K(E(i,1),E(i,3))= K(E(i,1),E(i,3))- 0.5;
    K(E(i,2),E(i,3))= K(E(i,2),E(i,3))+ 0;
    K(E(i,2),E(i,1))= K(E(i,2),E(i,1))- 0.5;
    K(E(i,3),E(i,1))= K(E(i,3),E(i,1))- 0.5;
    K(E(i,3),E(i,2))= K(E(i,3),E(i,2))+ 0;
end

for i=2:2:2*M*N
    K(E(i,1),E(i,1))= K(E(i,1),E(i,1))+ 0.5;
    K(E(i,2),E(i,2))= K(E(i,2),E(i,2))+ 0.5;
    K(E(i,3),E(i,3))= K(E(i,3),E(i,3))+ 1;
    K(E(i,1),E(i,2))= K(E(i,1),E(i,2))+ 0;
    K(E(i,1),E(i,3))= K(E(i,1),E(i,3))- 0.5;
    K(E(i,2),E(i,3))= K(E(i,2),E(i,3))- 0.5;
    K(E(i,2),E(i,1))= K(E(i,2),E(i,1))+ 0;
    K(E(i,3),E(i,1))= K(E(i,3),E(i,1))- 0.5;
    K(E(i,3),E(i,2))= K(E(i,3),E(i,2))- 0.5;
end

Q= P+lambda*K;
U= Q\(P*F);

for l=1:(N+1)*(M+1)
    z_graduated(floor((l-1)/(N+1))+1,mod(l-1,
N+1)+1)= U(l,1);
end
x_graduated=x;
y_graduated=y;
```